# The Weight of Success: The Body Mass Index and Economic Well-being in South Africa

Martin Wittenberg*
School of Economics
University of Cape Town

(Work in progress October 2005)

## 1 Introduction

Many South Africans, even poor ones, have a high body mass (Case and Deaton 2005). This has led to an increase in the prevalence of hypertension and strokes in contexts where one might not have expected to see this (Kahn and Tollman 1999). Understanding some of the correlates of high body mass would therefore be useful. Indeed the rapid increase in obesity around the world has become the focus of attention not only of health researchers. Increasing numbers of economists have also started to explore the economic correlates of the increase in body weight. In US based studies (e.g. Chou, Grossman and Saffer 2002) a negative relationship between income and obesity has been observed. On the other hand it seems clear that across countries obesity is positively correlated with income. This suggests that the relationship may be non-monotonic: increasing with income at low levels, but decreasing at higher levels.

South Africa offers an interesting setting for studying some of the these effects. Firstly it has high levels of inequality. This means that there is a section of the population (largely the White sub-population) that has incomes and standards of living comparable to those found in developed societies. Within the Black South African majority there is also a wide range of incomes which will give us some power to analyse these relationships. Secondly, obesity is becoming a demonstrable problem, even in communities in which poverty seems to be widespread.

In this paper we will be concerned with two substantive issues. Firstly we will show that among Black South Africans body mass increases more or less monotonically with income. Furthermore it seems that body mass also increases with education and employment. This is consistent with the idea that more successful individuals have greater weight. Indeed we will show that Black South Africans seem to aspire to greater body weight even in ranges where they would be classified as overweight. The picture for White South Africans, however, is quite different, with high income individuals more likely to be of lower body weight.

Secondly, we will show that there is a strong gender component to these relationships. Black women acquire body weight at a more rapid rate with increases in income than do their men. White women, on the other hand, shed it more aggressively. These "stylised facts" suggest that the images that these women have of themselves and that, perhaps, their partners have of them, are also implicated in the patterns that we observe.

In order to make these points we will also be concerned with a methodological issue. There are three publicly available data sets which have anthropometric information. Two of these are relatively small surveys that also have good socio-economic information. The third is the Demographic and Health Survey which only has asset information. A considerable part of this paper will be devoted to examining how good an estimate one can get from such asset proxies. We will show that on the whole the assets perform reasonably well – but that once one allows for the facts that assets may have independent effects on the outcome of

*I thank Susan Godlonton for brilliant research assistance. SALDRU provided funding for the research and an earlier version of this paper was presented to the Centre for Social Science Research's workshop on Family Support.

interest, the issue becomes considerably more murky. In the process, however, we will discover some points which, after the event, are fairly obvious, such as that car ownership seems to be implicated in the increase in obesity!

The structure of this paper is as follows. In the next section we will review some of the literature and, in particular, some of the literature dealing with proxy variables. In section 3 we describe our three data sets and present a number of descriptive analyses. Section 4 provides a detailed discussion of the performance of the asset proxies on these data sets. Having satisfied ourselves that these variables will not fundamentally distort the relationships, we use them in Section 5 to examine the relationship between body mass and a variety of economic and individual attributes. Section 6 concludes.

## 2 Examining the growth in obesity

A number of authors have tried to explain the increase in obesity observed internationally. Popkin (1999) has argued that urbanisation has led to changes in diets and activity patterns which are implicated in the rise of body mass. Of course this does not explain why urban people should be consuming a different diet. Chou et al. (2002) have suggested that changes in the operative prices – and in particular the opportunity costs of time spent cooking at home – are sufficient to explain increases in the rate of consumption of take-out foods and hence obesity. Cutler, Glaeser and Shapiro (2003) argue that it is technological changes in the way food is prepared, allowing it to be accessed within the home much more quickly, that are the causal factor. Philipson and Posner (1999) and Lakdawalla and Philipson (2002) argue that it is technological changes in the workplace which are to blame. In the past individuals had to expend considerable calories in order to obtain their food. Modern machinery has meant that the caloric costs of acquiring food have come down rapidly.

We will not be able to adjudicate between these explanations. In the process of looking at the correlates of high body mass we will be able to show that a number of these explanations are probably correct. For instance, access to a refrigerator seems to matter for body weight among women. Car ownership is unequivocally associated with increases in body weight, in line with the Lakdawalla and Philipson (2002) argument about the energy budgets of modern individuals. Passive forms of leisure (such as television) also seem to make a difference. Some of these effects are fairly small, however, and need a large data set for a proper analysis. The problem is that the available data set, the Demographic and Health Survey, does not have the necessary economic data.

An influential paper by Filmer and Pritchett (2001) suggested that the lack of socio-economic information in the Demographic and Health Surveys could be mitigated by constructing an asset index. They suggested that one should use the first principal component of the available asset variables. This procedure provides that linear combination of the asset variables which explains the greatest proportion of their joint variance. It is natural to think of this as extracting an index of affluence – hence the title of their paper "Estimating wealth effects without expenditure data – or tears". We will be yet more audacious and claim that we are able to estimate a lower bound on the effect of expenditure, without expenditure data!

The Filmer and Pritchett approach has been criticised by Lubotsky and Wittenberg (forthcoming), who argue that the first principal component is also likely to extract joint noise. Instead they suggest that one should add the asset variables in separately and then aggregate the coefficients up later. In particular they show that if the structural model is given by

$$\begin{aligned} y &= \beta x^* + \varepsilon && (1)\\ x_1 &= x^* + u_1 && (2)\\ x_2 &= \rho_2 x^* + u_2 && (3)\\ &\cdots \\ x_k &= \rho_k x^* + u_k && (4) \end{aligned}$$

where $x^*$ is the unobserved latent variable and $x_1, \ldots, x_k$ are asset variables that are observed, then if $b$

is the vector of OLS coefficients in the multiple regression that includes all the proxies, the estimator

$$\begin{aligned}\widehat{\beta}_{LW} &= \widehat{\rho}'b \\ &= \sum \widehat{\rho}_i b_i\end{aligned}$$

will provide an attenuated estimate of $\beta$, but one that has lower attenuation bias than any other linear combination of the proxy variables that is rescaled appropriately. They show that $\rho_i$ can be consistently estimated as

$$\widehat{\rho}_i = \frac{cov\,(y, x_i)}{cov\,(y, x_1)} \tag{5}$$

It is necessary, however, to normalise on one of the proxies and they do so by setting $\rho_1 = 1$.

In our context we will achieve this normalisation by projecting $x_1$ on $x^*$ in a national data set where we do observe both, i.e. we will estimate

$$x_1 = \delta_1 + \delta_2 x^* + v$$

and then use $\frac{1}{\delta_2} x_1$ as our rescaled version of the proxy.

A limitation of the Lubotsky-Wittenberg (LW) procedure is that it is not explicit about how to proceed if there are covariates in the model. They appeal to the Frisch-Waugh-Lovell theorem to argue that the result holds if the effect of the covariates has been stripped out, i.e. if equation 1 is the relationship between the residuals once $y$ and $x^*$ have been projected onto the covariates. Indeed the model implicitly assumes that all variables are written in deviations from their mean, i.e. have been projected on to a column of ones.

In a recent paper Wittenberg (2005b) has argued that one can do considerably better. Consider the model

$$y = \beta x^* + Z\gamma + \varepsilon \tag{6}$$

where $Z$ is a vector of covariates. He assumes that $Z'x^*$ is non-zero, but $Z'U = 0$ where $U$ is the vector of measurement errors given in equations 2 to 4. If we rewrite equation 4 as

$$x_k = \rho_k\,(x_1 - u_1) + u_k \tag{7}$$

it is clear that it can be consistently estimated by GMM techniques, using the moment conditions $Z'U = 0$ and $y'U = 0$. This amounts to estimating the relationship between $x_k$ and $x_1$ by instrumental variables, using $Z$ and $y$ as instruments for $x_1$. Indeed inspecting equation 5 it is evident that that is precisely what the original Lubotsky-Wittenberg procedure does. In the presence of covariates, however, the relationship may be overidentified and we can estimate $\rho$ more efficiently. Furthermore we can use a test of the overidentifying restrictions to test whether the assumptions of the model hold. This is a test of the validity of the assumptions for the $k-$th proxy. By stacking the $(k-1)$ equations for $\rho_2, \ldots, \rho_k$ we can estimate the vector $\rho = \begin{bmatrix} 1 & \rho_2 & \ldots & \rho_k \end{bmatrix}'$ as a system. Correspondingly there is a systems version of the overidentification test. Wittenberg notes that the test can fail for many reasons:

- A correlation between $\varepsilon$ and any of the $u_i$ terms. In this case $y$ would cease to be an appropriate instrument. This implies that the proxy variable $x_i$ is proxying not only for the latent variable $x$, but should be in the main regression. The appropriate "fix" for this is to run the regression with $x_i$ as a regressor in its own right.
- A correlation between any of the $z_j$ variables used as instruments and any of the $u_i$ terms. In this case we can "rescue" the model by removing $z_j$ from the list of instruments. However, the relationship between $z_j$ and $u_i$ means that the coefficient on $z_j$ has to be corrected, since the proxy variable will be picking up part of its effect.
- A misspecification of any of the proxy variable equations
- A misspecification of the main regression

However any of these would be reasons for being sceptical about the validity of the LW estimation procedure. Indeed, most of these would call into question **any** attempt to estimate the same regression using asset proxies. The specification test is therefore likely to be of more general interest.

In the empirical part of this paper we will devote considerable attention to validating the use of proxy variables in the context of estimating the relationships between BMI and expenditure. In particular we intend to show that:

- The FP asset index works well in proxying for expenditure, although the magnitude of the coefficient is not a reliable guide to the magnitude of the expenditure effect.
- The LW procedure does seem to provide lower bounds on the expenditure effects, although in some cases the attenuation seems large.
- The specification tests seem to work as intended and seem to reveal meaningful relationships that might otherwise not have occurred to the researcher.

# 3 Exploring BMI on three data sets

The focus of our empirical work is on the Body Mass Index. This is defined as the weight of the individual (in kilograms) divided by the height (in metres) squared. For the purposes of our analyses we only use individuals aged twenty and older, because the growth charts produced by the US National Centre for Health Statistics suggest that some individuals (particularly at the top end of the distribution) are still growing up to this age. With height basically fixed at age twenty changes in BMI are reflections mainly of changes in weight.

We will be using three data sets: The Langeberg Survey, KwaZulu-Natal Income Dynamics Study (KIDS) wave 2, and the 1998 Demographic and Health Survey (DHS). The latter two surveys were carried out in 1998 while fieldwork for the first was conducted in the second part of 1999. Our analyses should therefore be providing us a set of snapshots of South Africa in 1998/99.

Each data set has its own peculiar strengths and weaknesses. By looking across all three data sets we will be able to get a unique perspective on these strenghts and weaknesses. We will also be able to get a better feel for how robust and reliable our results are likely to be.

All of the surveys used some form of disproportional sampling. Nevertheless we report in all cases the unweighted estimates. In none of the cases that we have analysed did the weights change the results noticeably. We have chosen to report the unweighted results, because the published weights (particularly in the case of the DHS) looked somewhat odd. The results reported below should therefore be interpreted, in the first instance, as being applicable to the individuals sampled. Since the weighted results look basically the same, we are confident that the results also hold true more generally.

## 3.1 The Langeberg Survey (SAIHS)

The "South African Integrated Household Study" was conducted by a team of researchers from *inter alia* the University of Cape Town, University of the Western Cape, Princeton University and Harvard. The survey was coordinated by the South African Labour and Development Research Unit (SALDRU) at the University of Cape Town. The survey was run in three Magisterial Districts (Mossel Bay, Heidelberg and Riversdal) in the Western Cape province, which collectively make up the Langeberg health district. Altogether 294 households were interviewed with a range of instruments: detailed modules on household socio-economic information, individual information and health related information, including anthropometric data. A distinctive feature of the study design (which we will not exploit in these analyses) was that all adult members of the household were separately interviewed.

The Langeberg district is more urbanised than South Africa as a whole. Furthermore its demographic make-up is quite different. In terms of the old apartheid classifications, Black South Africans are in a minority (at 14%), with the bulk of the population classified as "Coloured" (SAIHS 2001). In order to get an adequate sample size, we pooled Coloureds and Black South Africans.

As shown in Table 1, we have 561 individuals in our overall estimation sample, i.e. adults aged 20 above with complete information on all of the variables used in the regressions. Of these 457 were Black and Coloured. The Black sub-sample was 176.

The summary statistics in Table 1 show that the average body mass index in this sample is in the overweight category. Indeed half of the sample is overweight (i.e.with a BMI in excess of 25) while around a quarter is obese (BMI of 30 or above). Black and Coloured women in particular are very heavy, with around a third being obese.

## 3.2 KIDS 1998

The KwaZulu-Natal Income Dynamics Survey (KIDS) was a re-survey of the Black and Indian sub-samples from the KwaZulu-Natal province of the 1993 Project for Statistics on Living Standards and Development (PSLSD), otherwise known as the 1993 SALDRU survey. The PSLSD was a national survey based on the model of the World Bank Living Standards Measurement Surveys. The survey instrument consisted of a variety of modules, such as Expenditures, Incomes, Labour Market status as well as anthropometrics. Since the 1993 survey was not designed as a panel, the re-survey had to confront the problem of how to track individuals and households. The decision was made to follow "core" household members only. These were the Head of Household or spouse or any adult members present in 1993. The 1998 survey can therefore not be interpreted easily as being representative of the Black and Indian population of KwaZulu-Natal province.

There are some other issues around representativeness. The original 1993 survey only measured children of six years and younger. The 1998 re-survey also measured some of the adults. As shown in Table 2, however, the coverage of the adults was highly incomplete. Less than a third of the adults in the sample have usable anthropometric information. Nevertheless the total sample size with anthropometrics is larger than in the case of the Langeberg Survey. A comparison of the summary statistics for the sample as a whole and the estimation sample (in Table 2) suggests that the indviduals in the estimation sample are notably older (a mean age of 48 when compared to 39), have less education and are less likely to be employed (38% versus 50%).

From Table 1 it is also apparent that the individuals in the KIDS sample are noticeably heavier than the ones captured in the Langeberg Survey and the DHS. Indeed fully 70% of the sample is overweight, while around half of the Black women were obese.

## 3.3 The 1998 DHS

The Demographic and Health Survey follows the template of other DHSs, i.e. it has detailed modules on child bearing, contraception and attitudes to family planning. The instrument that we will be analysing is the Adult Health Questionnaire which has information on health seeking behaviour, clinical conditions, occupational health, health-related habits as well as anthropometrics. Its socio-economic information, by contrast, is rudimentary to say the least. In particular there is no information about incomes or expenditures. There is, however, information about assets in the household questionnaire. This allows us to construct an "asset index" (labelled asset1 in Table 1) as advocated by Filmer and Pritchett (2001). We will comment on the validity of this procedure in more depth below.

The other variable that is poorly measured is labour market status. The household roster contains one question whether the individual worked for pay in the last seven days. The adult health module has a question (in the occupational health section) asking "In the last 12 months, have you worked for payment?" There is no additional information that might enable one to determine whether an individual is unemployed or not economically active, or indeed whether an individual might be employed informally or seasonally. We have chosen to work with the looser (i.e. 12 month) definition of employment, to capture any casual or seasonal workers.

The chief strength of the DHS is its sample size. As shown in Table 1 the usable sample is an order of magnitude greater than in the case of the KIDS or the Langeberg survey. Consequently this sample provides a lot more power in investigating the relationships between the "economic" variables (such as the asset index) and body mass. This is also the only sample which allows us to investigate body mass on a non-black subsample. The White sub-sample is of particular interest, since this segment of the population has been historically privileged and enjoys on the whole much higher living standards than the majority. As shown in Table 1, the usable White sub-sample of the DHS is 915 individuals which is double the size of the Black and Coloured subsample of Langeberg and only 25% smaller than the Black subsample of KIDS.

## 3.4 Comparing the data sets

The summary statistics in Table 1 provide an interesting view on the comparability of the data sets. It is striking that the Langeberg Survey and the DHS have similar mean BMI figures, obesity and overweight rates for the entire sample as well as the Black subsamples. Furthermore the age profiles and household sizes are also very similar. The Langeberg area shows lower levels of education, but higher levels of employment, which is consistent with the fact that it is a relatively urbanised population but in a non-metropolitan setting. The rural parts of the Langeberg district are commercial farms which have relatively high employment levels. The typical rural areas in the rest of South Africa are the ex-homeland areas which have quite low employment levels.

The KIDS data set is markedly different, both in the BMI values as well as in the household sizes. While the average age in the KIDS sample as a whole is comparable to the average ages recorded for the Langeberg and DHS estimation samples, the individuals with usable BMIs in the KIDS sample are around ten years older. Interestingly enough, the log expenditures values in the KIDS data set are comparable to those in the Langeberg survey. However while the Langeberg survey shows a considerable mismatch between the income and expenditure figures, this is not the case in the KIDS survey. Since we used the household aggregates calculated by the relevant research teams, instead of estimating these *de novo*, this may be due to differences in the imputation processes used in the two surveys.

The peculiar nature of the body mass information provided in the KIDS data set is confirmed by Figure 1. It is again interesting to note how similar the Langeberg and DHS BMI data is. The KIDS distributions, by contrast, are right-shifted in the case of both the Black men and women. All surveys agree, however, that there is a clear difference between Black men and women, with the latter being much heavier than the former. A quite different picture is provided by the White subsample. Figure 2 indicates that the mode of the BMI distribution for White men is higher than that for the women. The upper tail of the women's distribution extends beyond that of the men's, so that the mean BMI for White women lies just above that for the men. In summary, while White women are lighter than Black women, White men are considerably heavier than the Black ones in the DHS.

## 3.5 Some correlates of Body Mass

The series of local polynomial regressions shown in figure 3 provide a useful starting point for exploring the relationship between body mass and economic factors. Several features of these graphs invite comment. Firstly, the relationship between body mass and total (log) household expenditure or assets seems almost linear, particularly among Black South Africans. Indeed there seems little evidence in these graphs as yet to suggest that there may be any turning point in the relationship between body mass and income.

Secondly, the expenditure and asset graphs provide broadly consistent results. In the KIDS data set the asset index suggests a flattening or turning point in the relationship among wealthy individuals. This is not nearly so evident in the expenditure based graphs. This is an important point that we will return to later in the discussion.

Thirdly, the Langeberg and DHS data sets again provide remarkably congruent results. One should note that the asset indices were constructed separately for each data set, so that they provide consistent internal rankings, but there is no *a priori* reason to believe that they should be comparable across data sets. The KIDS results are again shifted upwards.

The picture of a monotonic relationship between body mass and economic resources is tempered by the results shown in figure 4. Here we have plotted the nonparametric regression functions for Black and White individuals in the DHS from the fifth to the 95th percentile of each subpopulation's asset distribution[1]. The privileged positions of Whites is underscored by the fact that the fifth percentile is far along the distribution of the Black group. Figure 4 suggests that despite many historical differences, body mass among Black and

[1] The reason for cutting off the tails of the distributions is that the nonparametric estimators become unstable in these regions.

White South Africans may very well be determined in similar kinds of ways. It is tempting to splice the two regression functions together and conclude that body mass may reach a turning point at high incomes. Of course, as the KIDS graphs show, the asset index and direct measures of economic well-being are not exactly the same thing. It may be the case that the asset index is not able to differentiate sufficiently between more and less wealthy White individuals.

Figure 5 highlights a different set of issues. The graphs show that body mass increases steeply with age, reaching a maximum around age fifty and then slowly decreasing from there. These graphs cannot tell us whether the profiles are purely a function of aging (and mortality), or whether there are cohort effects mixed in with them. If it is true that obesity is at least partially a function of childhood nutrition, it may be the case that young cohorts will have lower (or higher) BMIs over their life-cycles than older cohorts.

The Figure also suggests that the differences between KIDS and the other two data sets are not purely a function of the fact that the KIDS sample is older. Indeed it looks as though the BMI figures are about two units higher at every age than the DHS figures.

Finally the graphs confirm the weight "rankings" noted earlier. At every age Black women seem heavier than White men and women, who in turn are heavier than Black men.

These bivariate comparisons are very useful indicators of the underlying relationships, but in order to probe the relationships further, we require multivariate techniques. In particular we want to look at the relationship between BMI and economic variables, controlling for confounding effects such as age and gender. The data set that will give us maximum power in this regard is the DHS, but in order to use regression techniques on this data set we need to be certain that the asset index is a reliable proxy for household income or expenditure. We turn now to consider this question in some details.

# 4 Validating the use of asset proxies

## 4.1 The performance of asset proxies in the Langeberg survey

The Langeberg and the KIDS data sets have expenditure, income and asset information. We can therefore explore how well the asset proxies perform. Table 3 provides an initial assessment for the Langeberg survey. In column one we have regressed BMI on log total household expenditure, including some controls for household composition and personal characteristics. The coefficient on expenditure suggests that a one standard deviation increase in log expenditure (i.e. a 0.87 unit increase) will increase BMI by 1.35 units. At the average height of 1.62m (5 ft 3 inches) this amounts to an increase in weight of 3.5kg (7.7lbs). This is a significant increase, both statistically and materially.

A number of the other coefficients are also interesting. The large "female" coefficient should not come as a surprise, given the distributions shown earlier. At the average height this would suggest that women are 9.2kg (20lbs) heavier than the corresponding men. Black individuals are on average heavier than Whites and Coloureds. The quadratic in age suggests a peak at age 51, which agrees well with the nonparametric regressions shown in Figure 5.

In column 2 we have used the Filmer and Pritchett (2001) asset index, i.e. we have extracted the first principal component from a range of asset variables. Two features stand out in this column. The first is that the coefficient on the asset proxy is even larger than that on the log expenditure variable. A one standard deviation increase in the asset index would lead to a 5.6kg increase (12lbs). This is surprising since we would have expected the coefficient to be suffering from attenuation bias. One obvious explanation is that the distribution underlying the asset index is not equivalent to that of the expenditure variable, i.e. there is no reason to suppose that a standard deviation increase in the index is equivalent to a standard deviation increase in income or expenditure. Another clue is provided at the base of the column where it appears that the regression with the asset index fits better than the one with log expenditure! We will come back to this issue below.

The second point to note is that the coefficients on the other covariates in the regression are remarkably similar. Use of the asset index does not qualitatively distort the conclusions that we would draw from the regressions. This is reassuring.

In column 3 we provide the comparison with the log of household income. As Table 1 shows, there are a substantial number of individuals (22%) that live in households that reported zero income. In order not to lose these from the sample, we set these incomes at R1 and included a separate dummy variable for this category. The large coefficient on the dummy suggests that individuals that live in these households are better off than their zero income would suggest - at least when measured in terms of their girth!

Interestingly enough the coefficient on the log of income is almost 40% smaller than the coefficient on log expenditure. The most plausible explanation is that log income is a more noisy measure of the real resources available to the individual. Indeed it is often assumed that income is more poorly measured in household surveys in developing countries than expenditure (Deaton 1997). The summary statistics in Table 1 would tend to support that view.

Lubotsky and Wittenberg (forthcoming) have recently argued that it is preferable to include all the proxies separately in the regression and aggregate the coefficients afterwards. In columns 4 to 6 we implement versions of that approach. In column 4 we have included all the asset proxies separately. It is evident that the explanatory power of the regression goes up and that a number of the assets seem individually significant. This need not be evidence of the fact that these variables belong in the main regression, since they may only be capturing the impact of the omitted expenditure variable. We will show below that some of the proxies do, in fact, have independent effects.

In order to implement the Lubotsky-Wittenberg (LW) procedure we need to standardise our estimates on one of the proxy variables. We have chosen telephone ownership for a number of reasons. Firstly, the bivariate correlation between telephone ownership and BMI is fairly strong. Secondly it is harder to imagine a direct impact of telephone ownership on BMI than would be the case with some of the other variables, such as television or car ownership. In order to make the coefficients directly comparable to the expenditure coefficients in column 1, we have rescaled the telephone variable, i.e. we projected telephone ownership on expenditure as

$$telephone = b_1 + b_2 \ln \text{ expenditure} + \varepsilon$$

and then rescaled the telephone variable as $\frac{1}{b_2} telephone$. A regression in which we used only this rescaled asset proxy is given in column 5. Astonishingly this proxy performs reasonably. The coefficient shows "only" 42% attenuation.

In column 6 we aggregate up the coefficients as suggested by LW[2]. The estimated coefficient of 1.617 is remarkably close to the true coefficient of 1.554. The fact that it is larger than the "true" coefficient is troubling, however. If the assets are only proxying for income/expenditure, then the coefficient ought to have been attenuated downwards. Congruent with the LW theory, however, we note that the addition of the other asset proxies substantially increases the coefficient from that given in column 5. Indeed LW provide a procedure by which the coefficients on any other linear combination of the asset proxies can be compared with the LW estimator. The coefficient corresponding to the Filmer-Pritchett (FP) asset index is given in the last line of column 2. It suggests that adding in all the proxies separately manages to extract an appreciably stronger signal.

In summary the FP asset index does a good job of proxying for expenditure in this particular regression. The estimated coefficient, however, is biased in an unknown direction (in this case, upward) and when it is converted into a coefficient where the direction of bias can be more accurately assessed, it is more attenuated than the LW estimator. The LW estimator by contrast works better than it should!

We might be content to leave things here. However, we show at the bottom of column 6 in Table 3 that the Wittenberg (2005b) omnibus specification test soundly rejects the validity of the model. Indeed the test can be applied to individual asset proxies and these tests suggest that several variables such as television and car ownership do not function as simple proxies for the same omitted variable.

In Table 4 column 1 we provide a regression that does pass the Wittenberg (2005b) specification test. Note that these estimates are still based on the regression given in column 4 of Table 3. We have simply allowed some proxies to have independent effects, then reaggregated the coefficients of the remaining proxies. We have also allowed for some of the covariates to be correlated with the proxies, which requires us to correct their coefficients as well. The regression suggests that access to electricity has a sizable and significant impact on body mass. For an individual with the mean height (1.62m) the presence of electricity would increase weight on average by 4.6 kg (10lbs). The presence of television would add an additional 6kg (13lbs). The coefficient on car ownership, while not statistically significant, implies an increase in weight of 2.3kg (5.2lbs).

[2] To be precise we use the more efficient version of the aggregation process described in Wittenberg (2005b).

The estimated coefficient for the latent variable is now half of its previous size. Nevertheless it still has a meaningful impact. Moving from the 25th to the 75th percentile of the income distribution (a change of one unit in log expenditures) would increase average body mass by 1,9 kg (4lbs).

The second column of Table 4 confirms that the specification picked out by the test is, indeed, a sensible one. In this column we re-run the first regression, but this time with the measured log expenditure variable. The regression coefficients are very similar to those in the first column. Indeed the conventional t-tests on the coefficients suggest strongly that electricity and television do, indeed, belong in this regression. Furthermore the "true" coefficient on log expenditure is only 15% higher, suggesting that the LW procedure does an excellent job in this particular case. The direction of the bias is also in line with the *a priori* expectations.

The final column estimates the same regression, but using the FP asset proxy. As before the proxy overestimates the coefficient. In addition the impact of the assets in the main regression is more noticeably underestimated. The reason for this is, of course, that these asset variables are already embedded in the proxy.

Several lessons flow from these analyses. Firstly it is evident that asset proxies such as the FP proxy can do a good job in capturing the impact of expenditure. The main limitation of the FP proxy is that the direction of bias is uncertain and that some of the assets might belong in the main regression. Secondly, the Wittenberg (2005b) specification test performs quite well in picking out sensible regressions.

### 4.2 The asset proxies in the KIDS survey

A contrasting picture is given by the KIDS survey, as shown in Tables 5 and 6. In Table 5, column 1, we observe that the coefficient on log expenditure is similar in the KIDS survey to the Langeberg survey (1.31 as against 1.55). The FP asset proxy (column 2) in this case, however, underestimates the true coefficient, if the expenditure coefficient is taken to be the appropriate benchmark. The theoretically most efficient proxy, i.e. the LW proxy, shows massive attrition. It only manages to capture one quarter of the true effect. The estimated coefficient implies a change in weight of 0.9kg (2lbs) for an adult of average height. As shown by the specification test at the foot of column 6 in Table 5, in this case too there is strong evidence that a number of the assets belong in the main regression.

Table 6 shows that ownership of a refrigerator and ownership of a television seem to have strong independent effects on body mass, even if log expenditure is included. The coefficient on car ownership, although not statistically significant, is still sizable. Indeed the size of this coefficient is similar to that given in Table 4. Once these assets are included in the main regression, the FP asset proxy does not add much. The estimate produced by the LW proxy in this case is larger in magnitude than that of the FP proxy, but is only one eighth of the expenditure coefficient. Nevertheless the specification test in this case still rejects the hypothesis that the remaining assets (furniture, jewellery, telephone, electrical appliances, cattle and sheep ownership) could all be proxying for the same omitted variable. The problem seems to be that these proxies only contain noise once the effects of the other assets and the location variables have been accounted for.

It is clear that in this case the asset proxies are not as successful in capturing variation in affluence as they were in the Langeberg survey. Part of the reason may be that the sample is odd. We have noted that the KIDS sample does seem to be skewed when compared to the other data sets. Interestingly, however, the simple correlation coefficients between the FP asset index and log expenditure are almost identical in the two surveys: .59 in the case of Langeberg and .62 in the case of KIDS.

Before moving off the regressions in Tables 5 and 6, it is useful to note that the poorer performance of the asset proxies did not lead to major distortions on the other coefficients in the models. Some of these coefficients are interesting. We observe that body mass increases in this sample with the number of children, the opposite of the effect in the Langeberg survey. The coefficient can be explained if children do many of the chores, allowing the adults to lead a more sedentary life. Otherwise one might have assumed that increases in household size would tend to reduce the resources available and so reduce body weight. The location coefficients are large, suggesting that urban living is associated with increases in body weight. It should be noted that the inclusion of these variables is not the reason why the asset proxies performed so poorly. Regressions without these variables led to only small increases in the coefficients of the asset proxies.

### 4.3 The internal validity of the asset proxies in the DHS

The evidence from the Langeberg survey and KIDS give us reasonable confidence that the first principal component from a set of asset variables will not give a distorted picture of the correlates of BMI. One additional piece of evidence which increases our confidence in the validity of our asset index is given by the graphs in Figure 6. At high levels of the asset index hunger occurs almost never, whereas at low levels it tends to be present some times or often.

## 5 Analysing Body Mass in the DHS

### 5.1 Racial and gender differences

In our analyses thus far we have pooled different population groups, and men and women. This is not unique to us. For instance Chou et al. (2002) do the same. Table 7, however, shows that this may distort the underlying relationships. Many of the key relationships flip signs if one runs separate regressions for Blacks and Whites. In particular the asset proxy switches from being positive to negative. A test of the hypothesis that the two subsamples can be pooled is decisively rejected[3].

Figure 4 should alert us, however, to the possibility that the relationship might be non-monotonic across the entire distribution. Interestingly enough the data **accept** the hypothesis that there may be common quadratic in the FP asset proxy across the White and Black subsamples. This quadratic is estimated to reach a turning point at 1.45 of the FP asset distribution. This agrees remarkably well with the turning points suggested by the non-parametric regressions. This point would, however, be at the 97th percentile of the Black FP asset distribution!

Even allowing for a quadratic in assets, the regressions are still statistically very different. In particular the differences in the gender coefficients and the employment coefficient do not disappear. It seems clear that it is preferable to run separate regressions for each population group.

Looking at the regression for Black South Africans, the "expenditure effect" of 0.754 estimated in column 3 is substantial. It suggests that a one unit change in log expenditure would lead to an increase in average weight of 2 kg (4.4lbs) for individuals of average height. This is a lower bound. The employment coefficient is interesting also. It suggests that employed individuals are on average somewhat heavier than non-employed or unemployed ones. Since we are controlling for total household resources, this may reflect either that employed people have a higher claim on those expenditures or (equivalently) that they do not perfectly pool resources. They may, for instance, pay for meals taken outside the house. This positive coefficient certainly suggests that employed people are better off than non-employed ones.

The positive coefficients on education might be interpreted in similar ways: individuals with more education seem to have more "weight" within the household. The difference between a person with a school-leaving qualification (12 years) and an individual with no education is more than one BMI unit which translates to around 3kg (6.6lbs) at the average height.

The regressions for White South Africans seem very different, with education, income and employment all serving to reduce body weight. This is more consistent with the US pattern, where obesity is seen often seen as a problem of poverty.

Table 8 shows, however, that it is also problematic to pool men and women within each of these racial groups. Tests of the pooling hypothesis are decisively rejected for each of the two race groups. In the case of Black South Africans pooling is rejected for two reasons: Firstly, the quadratic in age is quite different. The curvature of the profile for women is much greater than it is for men. This can be seen in the nonparametric regressions in Figure 5 also. Body mass seems to increase much more rapidly with age among Black women than among Black men and then tends to fall off again at a later age. Secondly, the relationship between body mass and economic resources (as proxied by the assets) is stronger. Allowing for a quadratic in the FP asset index does not change this result.

In the case of White South Africans the intercept and the coefficient on education are significantly different at the 5% level, while the difference in the coefficients on the FP asset proxy with a t-value of 1.93 is almost significant. The age profile, however, is practically identical – a result that is consistent with the

[3] All tests of the pooling hypotheses were conducted using the first principal component asset index.

nonparametric regressions. This is interesting since it suggests that the difference in the rate at which Black women acquire body mass cannot be simply based on the difference between oestrogen and testosterone. If it is biological it must be some other factor. Several socio-economic explanations, however, come to mind. In Figure 7 we show the proportion of individuals that classified themselves as "underweight" or "overweight" as a function of BMI. There are several striking trends in graphs. Firstly, White women are much quicker to classify themselves as "overweight" than any other group in the sample. Secondly, from the point onwards where medical people classify individuals as overweight (i.e. with a BMI in excess of 25) many Black women and men still regard themselves as underweight. Indeed even at a BMI of 40 around 8% of Black women still regard themselves as underweight. Only around a half of all Black women with this BMI regard themselves as overweight.

It seems clear that body mass functions quite differently in the preference functions of White women, Black women, White men and Black men. It seems as though greater weight is preferred among Black women and relatively low weight among White women. Income among Black women seems to be invested in acquiring more of it, while White women invest to get rid of it!

## 5.2 Nonlinearities in the impact of assets

This perspective also qualifies our observation that there may be a common non-monotonic relationship between expenditure and body mass, as exemplified in Figure 4. Women at the low end of the spectrum (i.e. Black women) seem to be acquiring body mass at a rate **faster** than suggested by that picture, but then (in the case of White women) losing it more rapidly at the high end of the spectrum. Black men, by contrast put it on more slowly, while white men seem to lose it more slowly. This is shown in the nonparametric regressions in Figure 8. If we estimate the relationship between body mass and the asset proxy with a quadratic relationship (as in Table 9), the picture is broadly confirmed. In the case of Black women the relationship seems concave: body mass increasing rapidly at the lower levels and then flattening out as it nears the tail end of the distribution. The estimated turning point is at the 94th percentile of the distribution for African women, which means that the slope of the graph is still essentially flat at the point where the distribution ends. In the case of the men, the parametric estimates suggest an accelerating relationship, i.e. the relationship is flat at low levels of the asset index and then increases more rapidly towards the top end of the asset distribution.

If this is an accurate reflection, it would have very interesting implications, suggesting that in poorer Black households it is the women who initially gain weight with an increase in affluence, while the men gain relatively more once the household is somewhat better off.

The case of White South Africans is quite different. White women seem to start off with a body mass in excess of white men, but lose weight over the entire asset distribution. White men, by contrast are the only subsample where there seems to be a genuine turning point. At relatively low levels of income body mass increases. It reaches its maximum about half way along the White male asset distribution and then decreases again.

We need to add one important caveat at this stage. The case of the KIDS graphs in Figure 3 shows that nonlinearities in the relationships with the asset variables may not be reflective of nonlinearities in the relationship with expenditure. Our validation exercise above, only examined linear specifications. The nonlinearities that the nonparametric regressions in Figure 8 and the parametric regressions in Table 9 suggest may be due more to an inability of the asset index to truly capture differences at the high end of the expenditure distribution.

We take two points away from these analyses. Firstly, for most of the subsamples the relationships seem monotonic. A linear specification should therefore not distort the underlying relationships badly. Secondly, there may be trade-offs within the overall household budget constraints about weight accumulation by men and women. For instance, it may be the case that low income men receive utility from having a well-rounded wife which results in relatively more rapid body mass increases for women at low incomes. To conclusively track these tradeoffs will, however, require better data than we currently have. Our proxy variables do a competent job in highlighting the big picture. To explore these nuances would require more accurate economic measures. In particular, the LW framework does not easily lend itself to estimating a non-linear relationship.

## 5.3 Testing the validity of the model

The results of the specification tests given at the foot of the "LW" columns in Tables 7 and 8 should give pause for thought. All of those estimations are roundly rejected by the test. This is hardly surprising, because some of the covariates (such as employment status and education level) may be correlated with the asset proxies directly, not just through the latent variable household expenditure. Furthermore our validation exercises should have alerted us to the fact that a number of the asset proxies are likely to have independent impacts on body mass. Lastly there are some asset variables (such as computer ownership among Black South Africans) that have such a low incidence that it is unclear what precisely they will capture.

Table 10 provides an attempt to assess whether the assets are, in fact, just proxying for expenditure or income in the case of Black South Africans. In the first column we have kept the same specification as in the previous tables, but allowed employment and education to be correlated with the asset variables. The specification test is not impressed with this maneouvre, so we allow some of the proxies to have independent effects. We were guided in their selection by the specification tests for the individual proxy equtions. In the case of Black men the tests suggest that car, bicycle and television ownership as well as holdings of sheep and cattle are not simply proxying for total income or expenditure in the regression. The regression reported in column 2, in which these are entered as separate regressors is, however, accepted. The coefficient on car ownership is large and statistically highly significant. It suggests that car owners of average height are 2.7 kg (almost 6lbs) heavier than individuals without cars. The other coefficients are more modest and statistically not significant, but have signs that seem sensible. Cattle and sheep ownership may be a sign of more traditional values which may in turn be related to the power of the male within the household.

The estimated lower bound on the direct "expenditure effect" in this regression is 0.375 which would translate to around 1kg difference (2.2 lbs) for a one unit increase in log expenditure. The employment effect is larger than this in absolute size, as is the difference that twelve years of education make.

In the case of Black women this specification is still soundly rejected, as shown by the test in column 3. In the case of women access to electricity and to a refrigerator seem to matter in ways that do not proxy for a simple expenditure effect. It is plausible that electricity might be important – collection of fire wood has generally been women's work and access to labour saving devices may also reduce the direct physical labour that they have to perform. Refrigerator ownership may capture the greater accessibility of ready-to-eat food. Of course this begs the question why this accessibility should have a gender-specific impact.

Interestingly, television ownership seems to have a much larger impact on the weight of women than on the weight of men also. This may reflect differences in the way in which television is utilised. For instance it may be the case that lack of security outside their homes induces women to spend more of their leisure time within the confines of their homes. South Africa's time use survey, for instance, shows that girls spend less of their afternoons socialising outside of the home and more of it in front of the television when compared to boys (Wittenberg 2005a).

The specification in column 4, which is accepted by the test, has stripped out much of the substance from the asset proxy. Only four asset variables are combined in that proxy! This accounts in part also for the much reduced size of the coefficient. Of course one might argue that the television, car and electricity effects are, of course, income effects too – even if the mechanism operates not through the greater availability of calories but through the reduction in the activity levels.

Taken as a whole, the results provided in Table 10 provide some perspective on the nonlinearities explored in the previous section. It is quite plausible that the assets that are acquired first by the poorest households (such as electricity and a television) have a greater impact on the weight of the women while those that are added later (such as a car) also move those of the men.

## 5.4 Feedback effects

Throughout this discussion we have used expenditure, employment and asset holdings to explain body mass. As noted earlier, however, there is a big body of literature that uses body mass to explain earnings. This raises at least the possiblity that body mass is endogenous in these regressions. There are some reasons why we suspect that this is not likely to be a big problem in South Africa. Firstly, the US literature that has examined this issue has generally found that high body mass individuals receive lower wages or are less likely to be employed. Our findings that it is the employed that have higher body mass (at least among Black

South Africans) goes in the opposite direction. Secondly, the US literature has suggested that the penalty for high body weight is less among Black American women. This group is more likely to provide relevant comparisons for Black South Africans.

There is, however, some evidence within the South African data sets that there may be some link back from extreme body weight to employment. There are a number of body weight readings in the DHS which are truly enormous. Some of these may be data errors. However, if the BMI data is winsorized at 95% it turns out that the employment coefficient in the regression corresponding to column 4 in Table 10 increases from 0.217 to 0.29, without any of the other coefficients shifting markedly. This suggests that there is some threshold above which high BMI individuals (in this case mainly women) are restricted from doing normal work. This feedback effect serves to reduce the estimated relationship between employment and body mass.

# 6 Conclusions

The increase in body mass among poor people is being seen increasingly as an important public health issue. The purpose of this paper has been to explore some correlates of high body mass among South Africans. We have shown that among Black South Africans a higher weight is often a mark of higher societal success. This effect may partially work through the accumulation of assets that reduce the need for physical exertion, but it is also seems to be associated with a desire to be bigger. This desire seems to be particularly marked among Black women. White women, by contrast, seem to associate success with reductions in body weight.

# References

**Case, Anne and Angus Deaton**, "Health and wealth among the poor: India and South Africa compared," *American Economic Review Papers and Proceedings*, 2005.

**Chou, Shin-Yi, Michael Grossman, and Henry Saffer**, "An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System," Working Paper 9247, NBER October 2002.

**Cutler, David M., Edward L Glaeser, and Jesse M Shapiro**, "Why have Americans become more obese?," Discussion Paper 1994, Harvard Institute of Economic Research 2003.

**Deaton, Angus**, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore: Johns Hopkins University Press, 1997.

**Filmer, Deon and Lant H. Pritchett**, "Estimating Wealth Effects Without Expenditure Data – Or Tears: An Application to Educational Enrollment in States of India," *Demography*, February 2001, *38* (1), 115–132.

**Kahn, K. and S.M. Tollman**, "Stroke in rural South Africa — contributing to the little known about a big problem," *South African Journal of Medicine*, 1999, *89* (1), 63–65.

**Lakdawalla, Darius and Tomas Philipson**, "The growth of obesity and technical change: A theoretical and empirical examination," Working Paper 8946, NBER May 2002.

**Lubotsky, Darren and Martin Wittenberg**, "Interpretation of regressions with multiple proxies," *Review of Economics and Statistics*, forthcoming.

**Philipson, Tomas J. and Richard A. Posner**, "The long-run growth in obesity as a function of technological change," Working Paper 7423, NBER November 1999.

**Popkin, Barry M.**, "Urbanization, Life Style Changes and the Nutrition Transition," *World Development*, 1999, *27* (11), 1905–1916.

**SAIHS**, "Introduction to the SALDRU Langeberg 1999 Data Set," SALDRU, University of Cape Town 2001.

**Wittenberg, Martin**, "The School Day in South Africa," Working Paper 113, CSSR, University of Cape Town 2005.

——, "Testing for a common latent variable in a linear regression: Or how to 'fix' a bad variable by adding multiple proxies for it," School of Economics, University of Cape Town 2005.

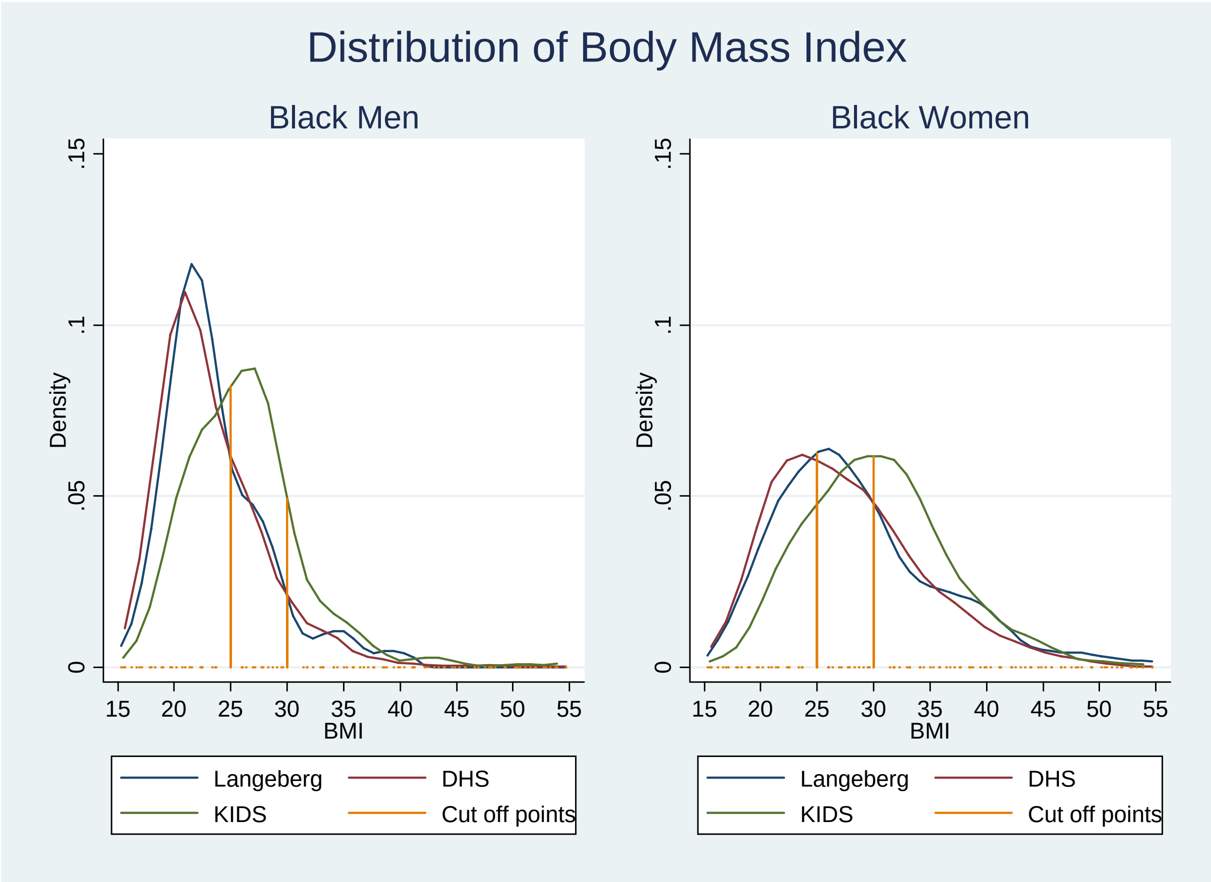


Figure 1: The Body Mass index in three surveys

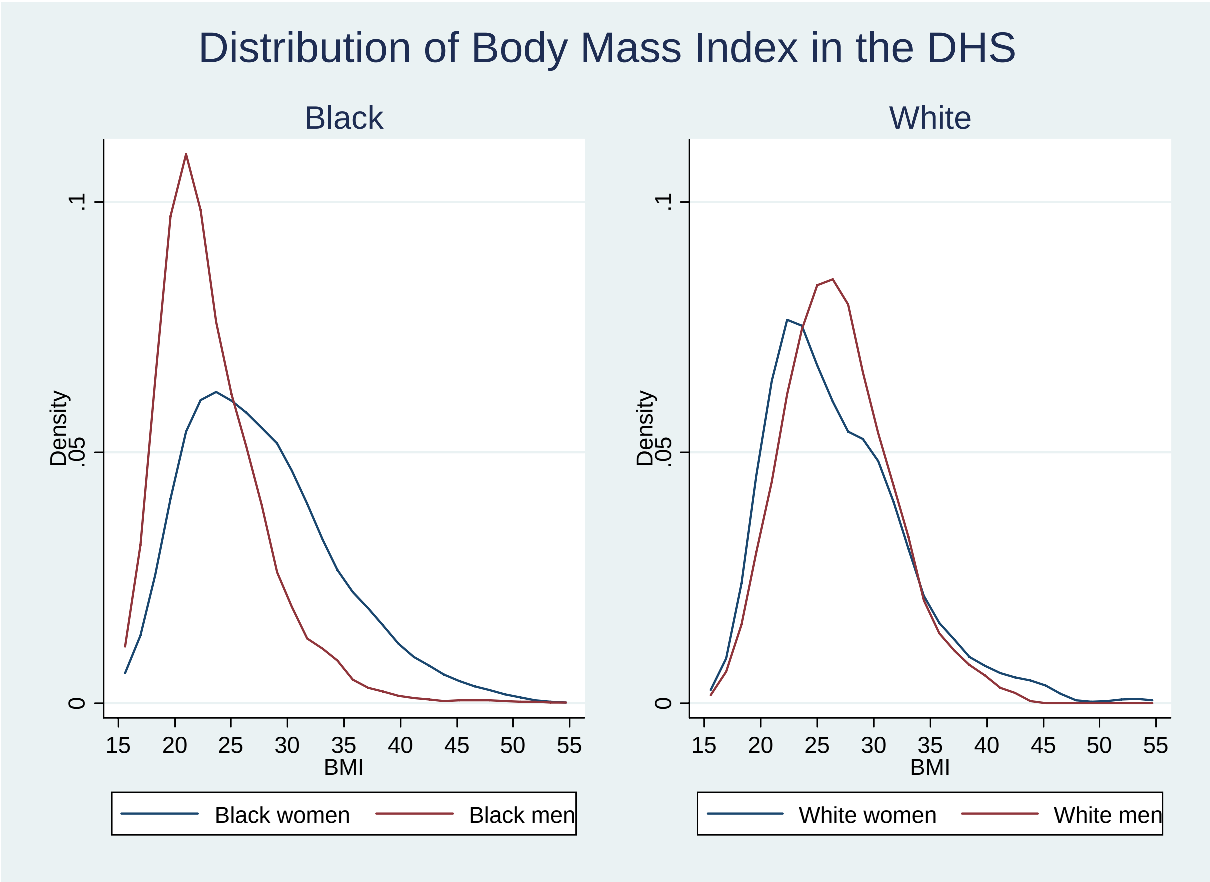


Figure 2: Distribution of BMI by Race and Gender

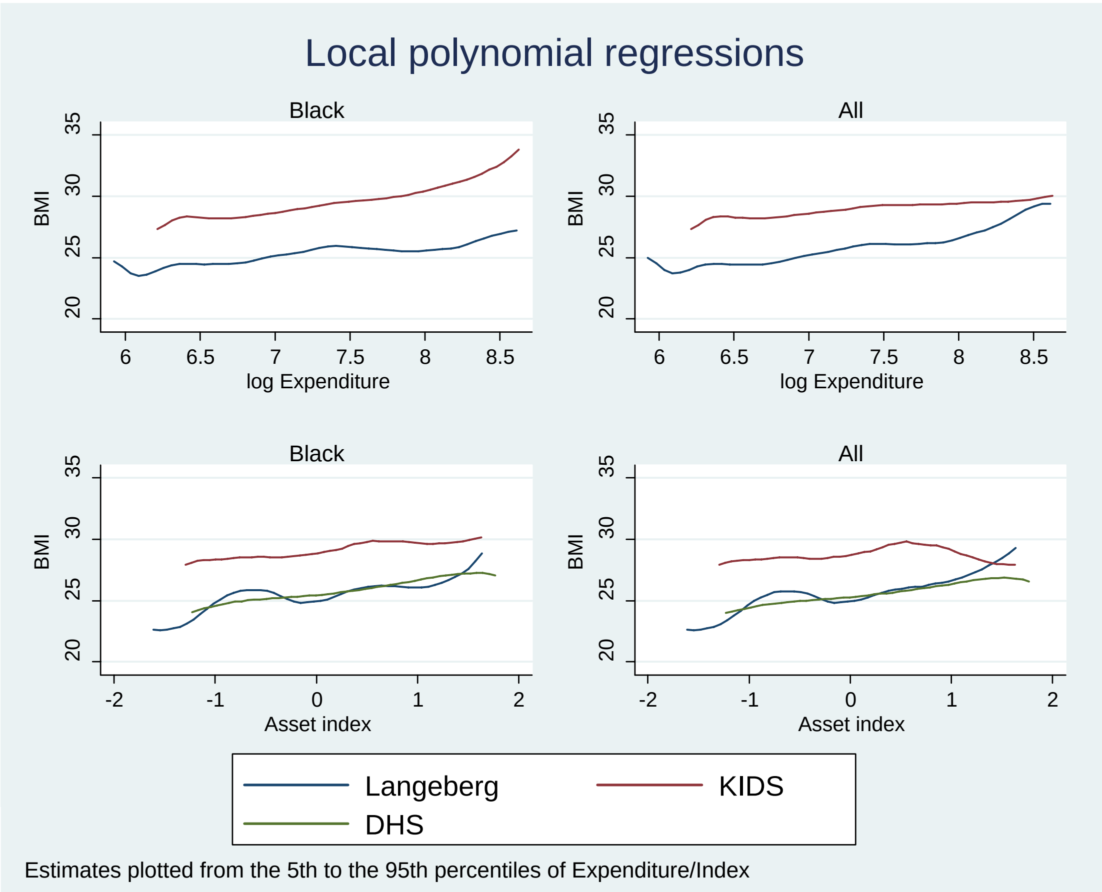


Figure 3: The relationship between body mass and affluence in three surveys

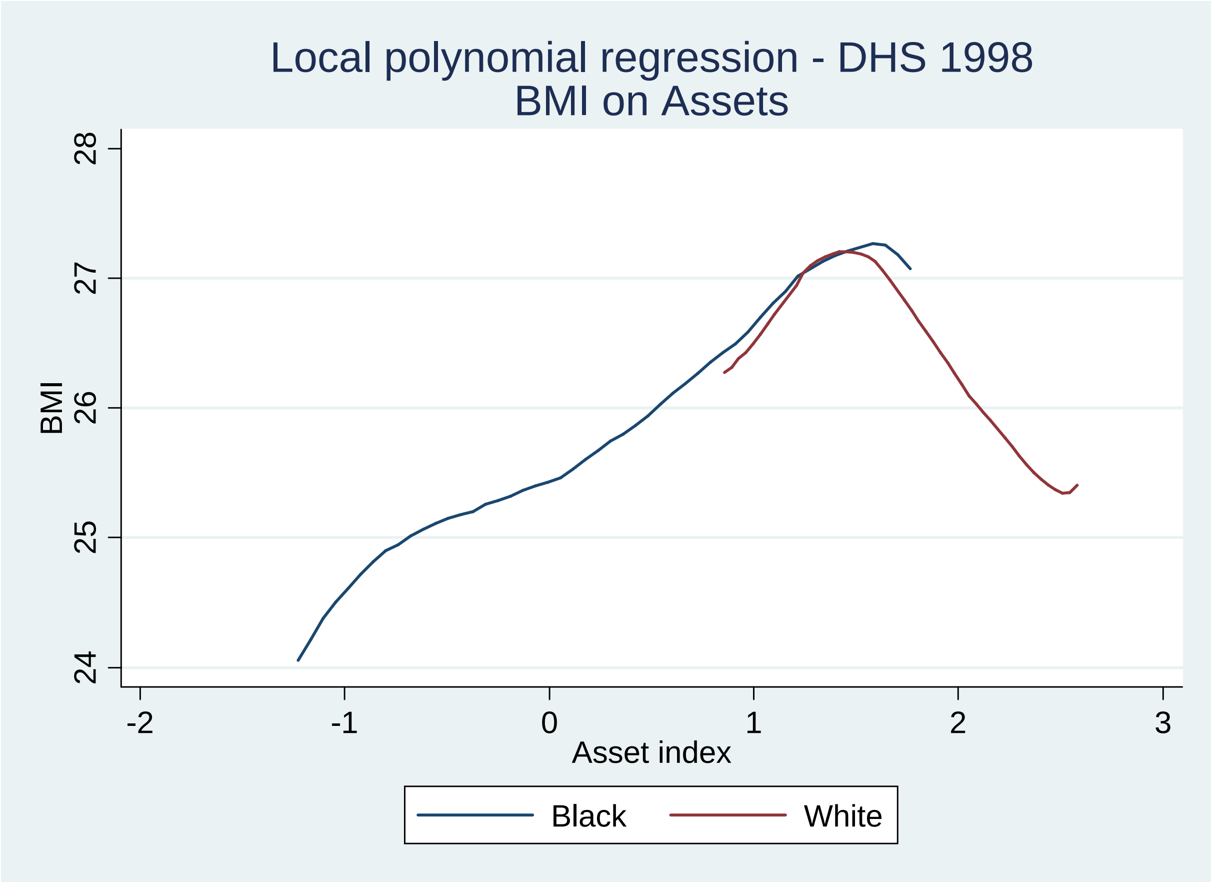


Figure 4: Relationship between BMI and the Asset index by race

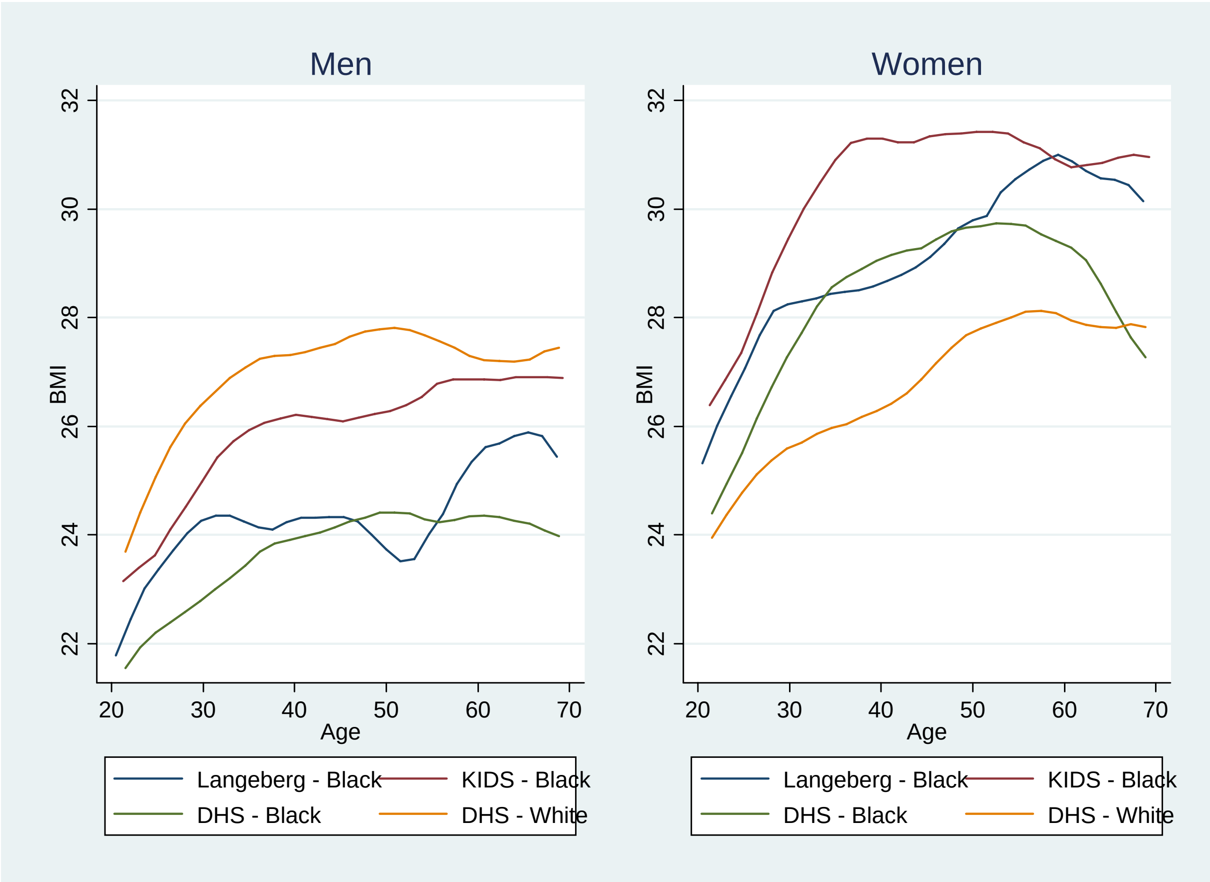


Figure 5: The Relationship between BMI and Age by Race and Gender

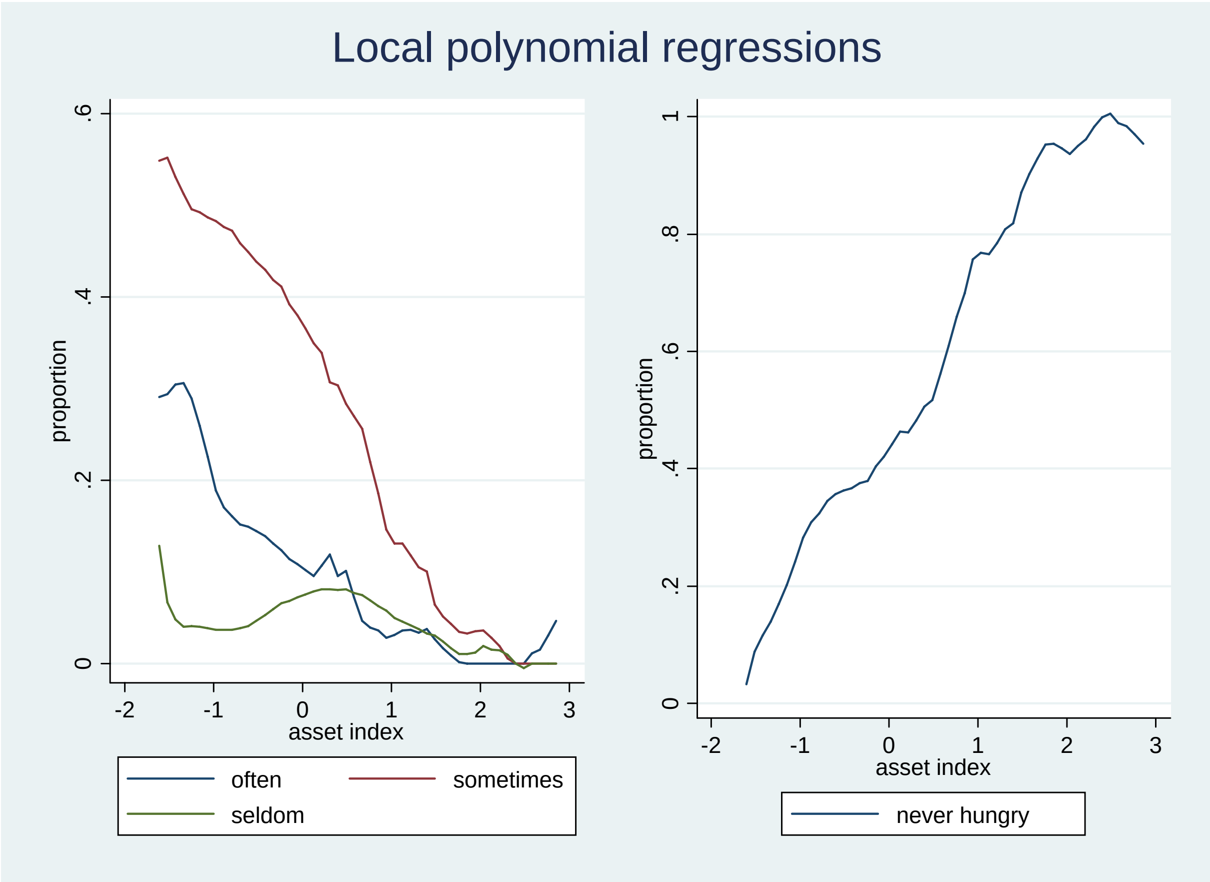


Figure 6: The relationship between reported hunger and the FP asset index

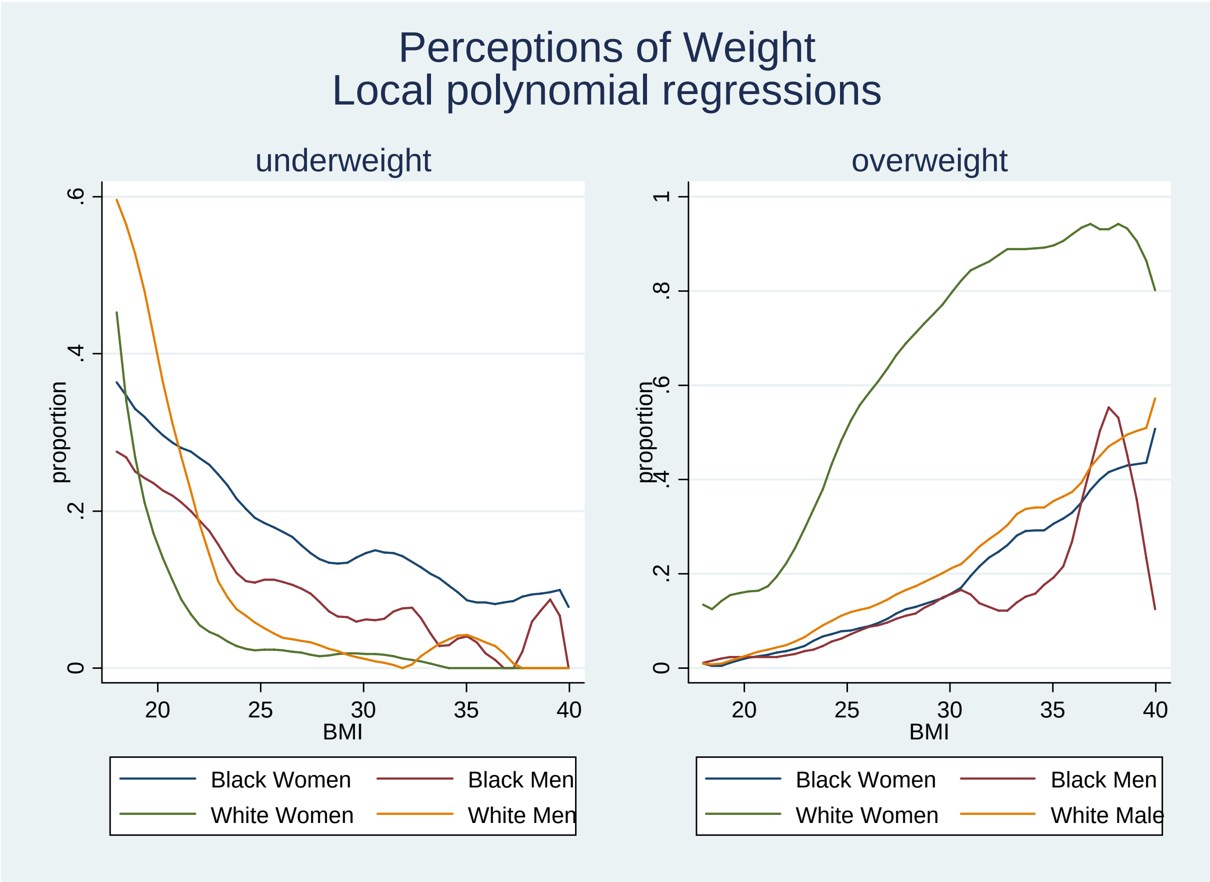


Figure 7: Differences in own perceptions of weight by race and gender

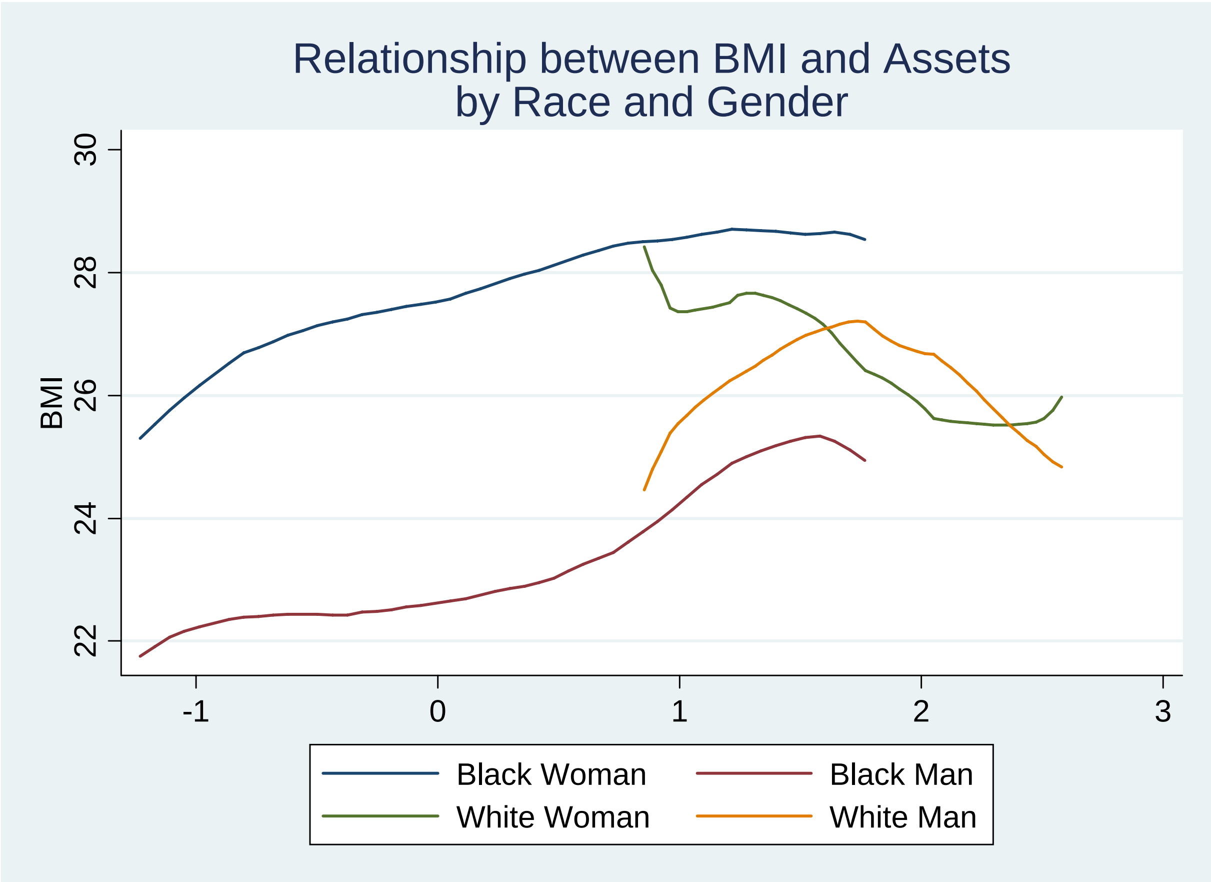


Figure 8: Local polynomial regressions of BMI on asset index by Race and Gender

## Table 1: Summary Statistics for the estimation samples

| | Langeberg | | | | DHS | | | | | | | KIDS | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **All** | **Black & Coloured** | **Women** | **Men** | **All** | **Black** | **Women** | **Men** | **White** | **Women** | **Men** | **All** | **Black** | **Women** | **Men** |
| bmi | 26.7 | 26.5 | 28.7 | 24.1 | 25.9 | 25.9 | 27.7 | 23.3 | 27.0 | 27.0 | 27.0 | 28.8 | 29.2 | 30.6 | 26.2 |
| | ( 7.57) | ( 7.83) | ( 7.11) | ( 7.89) | ( 6.59) | ( 6.54) | ( 6.85) | ( 5.10) | ( 5.44) | ( 5.97) | ( 4.74) | ( 6.52) | ( 6.58) | ( 6.67) | ( 5.30) |
| obese | 0.24 | 0.22 | 0.33 | 0.09 | 0.22 | 0.22 | 0.32 | 0.09 | 0.26 | 0.28 | 0.24 | 0.36 | 0.39 | 0.50 | 0.15 |
| | ( 0.42) | ( 0.41) | ( 0.47) | ( 0.29) | ( 0.42) | ( 0.42) | ( 0.47) | ( 0.28) | ( 0.44) | ( 0.45) | ( 0.43) | ( 0.48) | ( 0.49) | ( 0.50) | ( 0.36) |
| overweight | 0.51 | 0.49 | 0.67 | 0.30 | 0.48 | 0.47 | 0.60 | 0.29 | 0.59 | 0.55 | 0.63 | 0.70 | 0.73 | 0.79 | 0.58 |
| | ( 0.50) | ( 0.50) | ( 0.47) | ( 0.46) | ( 0.50) | ( 0.50) | ( 0.49) | ( 0.45) | ( 0.49) | ( 0.50) | ( 0.48) | ( 0.46) | ( 0.45) | ( 0.40) | ( 0.49) |
| Age | 41.2 | 39.2 | 38.6 | 39.9 | 41.7 | 40.9 | 41.1 | 40.6 | 47.4 | 48.5 | 46.1 | 48.2 | 48.7 | 48.0 | 50.3 |
| | ( 14.29) | ( 12.78) | ( 12.19) | ( 13.38) | ( 16.19) | ( 16.24) | ( 16.26) | ( 16.23) | ( 16.41) | ( 16.27) | ( 16.52) | ( 14.81) | ( 15.33) | ( 15.65) | ( 14.50) |
| hhsize | 4.9 | 5.3 | 5.5 | 5.1 | 4.9 | 5.2 | 5.4 | 4.9 | 3.2 | 3.2 | 3.2 | 7.2 | 7.7 | 7.8 | 7.3 |
| | ( 2.60) | ( 2.65) | ( 2.73) | ( 2.55) | ( 2.86) | ( 3.02) | ( 2.95) | ( 3.09) | ( 1.42) | ( 1.40) | ( 1.43) | ( 4.44) | ( 4.60) | ( 4.67) | ( 4.43) |
| yrseduc | 5.5 | 4.4 | 4.6 | 4.1 | 7.4 | 6.8 | 6.7 | 7.0 | 11.8 | 11.6 | 11.9 | 4.5 | 4.0 | 4.0 | 3.9 |
| | ( 4.21) | ( 3.40) | ( 3.42) | ( 3.37) | ( 4.32) | ( 4.31) | ( 4.34) | ( 4.27) | ( 2.15) | ( 2.10) | ( 2.20) | ( 3.71) | ( 3.50) | ( 3.45) | ( 3.61) |
| lnhhexp | 7.3 | 7.1 | 7.1 | 7.1 | | | | | | | | 7.4 | 7.2 | 7.2 | 7.2 |
| | ( 0.87) | ( 0.71) | ( 0.69) | ( 0.72) | | | | | | | | ( 0.81) | ( 0.68) | ( 0.68) | ( 0.69) |
| ln hhinc | 5.7 | 6.0 | 6.0 | 5.9 | | | | | | | | 7.4 | 7.3 | 7.2 | 7.4 |
| | ( 3.20) | ( 2.89) | ( 2.79) | ( 3.00) | | | | | | | | ( 1.06) | ( 1.02) | ( 1.03) | ( 0.98) |
| Zero income | 0.22 | 0.17 | 0.16 | 0.19 | | | | | | | | 0.0 | 0.0 | 0.0 | 0.0 |
| | ( 0.41) | ( 0.38) | ( 0.36) | ( 0.39) | | | | | | | | | | | |
| asset1 | 0.06 | -0.22 | -0.16 | -0.27 | 0.0 | -0.3 | -0.3 | -0.3 | 1.8 | 1.8 | 1.7 | 0.1 | -0.1 | -0.1 | 0.0 |
| | ( 1.03) | ( 0.91) | ( 0.89) | ( 0.92) | ( 1.01) | ( 0.78) | ( 0.79) | ( 0.77) | ( 0.55) | ( 0.53) | ( 0.58) | ( 1.03) | ( 0.94) | ( 0.94) | ( 0.95) |
| employed | 0.61 | 0.64 | 0.55 | 0.74 | 0.39 | 0.34 | 0.25 | 0.45 | 0.60 | 0.47 | 0.76 | 0.38 | 0.35 | 0.27 | 0.52 |
| | ( 0.49) | ( 0.48) | ( 0.50) | ( 0.44) | ( 0.49) | ( 0.47) | ( 0.44) | ( 0.50) | ( 0.49) | ( 0.50) | ( 0.43) | ( 0.48) | ( 0.48) | ( 0.44) | ( 0.50) |
| hhpensioner | 0.14 | 0.14 | 0.14 | 0.14 | 0.30 | 0.31 | 0.33 | 0.29 | 0.27 | 0.32 | 0.20 | 0.36 | 0.40 | 0.43 | 0.34 |
| | ( 0.40) | ( 0.38) | ( 0.38) | ( 0.39) | ( 0.46) | ( 0.46) | ( 0.47) | ( 0.45) | ( 0.44) | ( 0.47) | ( 0.40) | ( 0.48) | ( 0.49) | ( 0.50) | ( 0.47) |
| numadults | 3.4 | 3.6 | 3.7 | 3.6 | 3.1 | 3.1 | 3.0 | 3.2 | 2.6 | 2.6 | 2.6 | 3.9 | 4.1 | 4.1 | 4.1 |
| | ( 1.73) | ( 1.78) | ( 1.86) | ( 1.70) | ( 1.64) | ( 1.70) | ( 1.66) | ( 1.75) | ( 1.06) | ( 1.05) | ( 1.08) | ( 2.22) | ( 2.31) | ( 2.35) | ( 2.22) |
| children | 1.5 | 1.6 | 1.8 | 1.5 | 1.8 | 2.1 | 2.3 | 1.7 | 0.6 | 0.6 | 0.6 | 3.2 | 3.5 | 3.7 | 3.2 |
| | ( 1.45) | ( 1.47) | ( 1.48) | ( 1.46) | ( 1.90) | ( 2.00) | ( 2.02) | ( 1.92) | ( 0.92) | ( 0.93) | ( 0.91) | ( 2.82) | ( 2.90) | ( 2.94) | ( 2.78) |
| smoker | | | | | 0.30 | 0.26 | 0.08 | 0.50 | 0.31 | 0.23 | 0.40 | | | | |
| | | | | | ( 0.46) | ( 0.44) | ( 0.27) | ( 0.50) | ( 0.46) | ( 0.42) | ( 0.49) | | | | |
| female | 0.52 | 0.52 | | | 0.57 | 0.57 | | | 0.54 | | | 0.66 | 0.68 | | |
| | ( 0.50) | ( 0.50) | | | ( 0.49) | ( 0.49) | | | ( 0.50) | | | ( 0.47) | ( 0.46) | | |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| black | 0.31 | 0.39 | 0.38 | 0.39 | 0.73 | | | | | | | 0.85 | | | |
| | ( 0.46) | ( 0.49) | ( 0.49) | ( 0.49) | ( 0.44) | | | | | | | ( 0.35) | | | |
| coloured | 0.50 | | | | 0.14 | | | | | | | | | | |
| | ( 0.50) | | | | ( 0.35) | | | | | | | | | | |
| asianind | | | | | 0.04 | | | | | | | | | | |
| | | | | | ( 0.19) | | | | | | | | | | |
| n | 561 | 457 | 236 | 221 | 10299 | 7557 | 4342 | 3215 | 915 | 497 | 418 | 1444 | 1233 | 844 | 389 |

Standard deviations are given in parentheses

**Table 2: Comparison between the estimation sample and the entire sample in KIDS**

| | **All** | | **Black** | | **Women** | | **Men** | | **Estimation sample** | | **Black** | | **Women** | | **Men** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Obs | Mean | Obs | Mean | Obs | Mean | Obs | Mean | Obs | Mean | Obs | Mean | Obs | Mean | Obs | Mean |
| bmi | 1446 | 28.8 | 1235 | 29.2 | 845 | 30.6 | 390 | 26.3 | 1444 | 28.8 | 1233 | 29.2 | 844 | 30.6 | 389 | 26.2 |
| | | ( 6.52) | | ( 6.58) | | ( 6.67) | | ( 5.30) | | ( 6.52) | | ( 6.58) | | ( 6.67) | | ( 5.30) |
| age | 4266 | 38.9 | 3804 | 38.6 | 2196 | 39.5 | 1608 | 37.4 | 1444 | 48.2 | 1233 | 48.7 | 844 | 48.0 | 389 | 50.3 |
| | | ( 15.77) | | ( 15.96) | | ( 16.67) | | ( 14.86) | | ( 14.81) | | ( 15.33) | | ( 15.65) | | ( 14.50) |
| hhsizem | 4682 | 7.8 | 4193 | 8.2 | 2357 | 8.3 | 1836 | 8.1 | 1444 | 7.2 | 1233 | 7.7 | 844 | 7.8 | 389 | 7.3 |
| | | ( 4.64) | | ( 4.72) | | ( 4.71) | | ( 4.73) | | ( 4.44) | | ( 4.60) | | ( 4.67) | | ( 4.43) |
| yrseduc | 4633 | 5.5 | 4151 | 5.3 | 2355 | 5.2 | 1796 | 5.4 | 1444 | 4.5 | 1233 | 4.0 | 844 | 4.0 | 389 | 3.9 |
| | | ( 3.79) | | ( 3.73) | | ( 3.76) | | ( 3.68) | | ( 3.71) | | ( 3.50) | | ( 3.45) | | ( 3.61) |
| asset1 | 4682 | 0.1 | 4193 | -0.1 | 2357 | -0.1 | 1836 | -0.1 | 1444 | 0.1 | 1233 | -0.1 | 844 | -0.1 | 389 | 0.0 |
| | | ( 1.01) | | ( 0.95) | | ( 0.95) | | ( 0.94) | | ( 1.03) | | ( 0.94) | | ( 0.94) | | ( 0.95) |
| lnhhexp | 4682 | 7.4 | 4193 | 7.2 | 2357 | 7.2 | 1836 | 7.2 | 1444 | 7.4 | 1233 | 7.2 | 844 | 7.2 | 389 | 7.2 |
| | | ( 0.75) | | ( 0.66) | | ( 0.66) | | ( 0.65) | | ( 0.81) | | ( 0.68) | | ( 0.68) | | ( 0.69) |
| lincome | 4665 | 7.4 | 4178 | 7.3 | 2349 | 7.3 | 1829 | 7.3 | 1438 | 7.4 | 1229 | 7.3 | 841 | 7.2 | 388 | 7.4 |
| | | ( 1.05) | | ( 1.02) | | ( 1.02) | | ( 1.02) | | ( 1.06) | | ( 1.02) | | ( 1.03) | | ( 0.98) |
| empl | 4682 | 0.50 | 4193 | 0.50 | 2357 | 0.42 | 1836 | 0.61 | 1444 | 0.38 | 1233 | 0.35 | 844 | 0.27 | 389 | 0.52 |
| | | ( 0.50) | | ( 0.50) | | ( 0.49) | | ( 0.49) | | ( 0.48) | | ( 0.48) | | ( 0.44) | | ( 0.50) |
| hhpens | 4682 | 0.41 | 4193 | 0.44 | 2357 | 0.45 | 1836 | 0.43 | 1444 | 0.36 | 1233 | 0.40 | 844 | 0.43 | 389 | 0.34 |
| | | ( 0.49) | | ( 0.50) | | ( 0.50) | | ( 0.49) | | ( 0.48) | | ( 0.49) | | ( 0.50) | | ( 0.47) |
| numadults | 4682 | 4.3 | 4193 | 4.4 | 2357 | 4.4 | 1836 | 4.5 | 1444 | 3.9 | 1233 | 4.1 | 844 | 4.1 | 389 | 4.1 |
| | | ( 2.34) | | ( 2.41) | | ( 2.39) | | ( 2.44) | | ( 2.22) | | ( 2.31) | | ( 2.35) | | ( 2.22) |
| children | 4682 | 3.5 | 4193 | 3.7 | 2357 | 3.8 | 1836 | 3.6 | 1444 | 3.2 | 1233 | 3.5 | 844 | 3.7 | 389 | 3.2 |
| | | ( 3.02) | | ( 3.05) | | ( 3.03) | | ( 3.07) | | ( 2.82) | | ( 2.90) | | ( 2.94) | | ( 2.78) |
| female | 4682 | 0.56 | 4193 | 0.56 | | | | | 1444 | 0.66 | 1233 | 0.68 | | | | |
| | | ( 0.50) | | ( 0.50) | | | | | | ( 0.47) | | ( 0.46) | | | | |
| black | 4682 | 0.90 | | | | | | | 1444 | 0.85 | | | | | | |
| | | ( 0.31) | | | | | | | | ( 0.35) | | | | | | |

Standard deviations are given in parentheses

**Table 3: Comparing the performance of the asset proxies in the Langeberg survey**

| Dependent variable: BMI | (1) ln Exp | (2) FP Asset | (3) ln Inc | (4) Proxy All | (5) Proxy Tel | (6) LW |
|---|---|---|---|---|---|---|
| employed | -0.863 (0.700) | -0.522 (0.684) | -1.014 (0.738) | -0.502 (0.685) | -0.478 (0.692) | -0.502 (0.744) |
| age | 0.326 (0.124)** | 0.347 (0.121)** | 0.351 (0.124)** | 0.368 (0.122)** | 0.370 (0.122)** | 0.368 (0.107)** |
| agesq | -0.003 (0.001)* | -0.004 (0.001)** | -0.003 (0.001)* | -0.004 (0.001)** | -0.004 (0.001)** | -0.004 (0.001)** |
| yrseduc | -0.067 (0.099) | -0.153 (0.097) | -0.011 (0.096) | -0.196 (0.102)+ | -0.075 (0.096) | -0.196 (0.088)* |
| Expenditure/ proxy | 1.554 (0.458)** | 2.218 (0.386)** | 0.953 (0.346)** | 0.580 (0.228)* | 0.901 (0.201)** | 1.617 (0.209)** |
| numadults | -0.057 (0.202) | -0.035 (0.192) | 0.053 (0.197) | 0.011 (0.194) | 0.040 (0.193) | 0.011 (0.157) |
| children | -0.459 (0.231)* | -0.325 (0.228) | -0.513 (0.234)* | -0.396 (0.233)+ | -0.401 (0.229)+ | -0.396 (0.206)+ |
| female | 3.562 (0.625)** | 3.400 (0.613)** | 3.590 (0.630)** | 3.436 (0.613)** | 3.398 (0.620)** | 3.436 (0.593)** |
| black | 2.210 (1.160)+ | 2.680 (1.119)* | 1.579 (1.123) | 2.732 (1.301)* | 1.755 (1.104) | 2.732 (1.014)** |
| coloured | -0.487 (1.144) | 0.180 (1.119) | -1.576 (1.109) | 0.456 (1.201) | -0.471 (1.116) | 0.456 (0.960) |
| zeroinc | | | 5.768 (2.623)* | | | |
| electricity | | | | 2.129 (1.077)* | | |
| television | | | | 2.417 (0.788)** | | |
| car | | | | 0.777 (0.953) | | |
| bicycle | | | | 0.037 (0.718) | | |
| elecstove | | | | -1.338 (0.980) | | |
| coalstove | | | | -0.831 (0.804) | | |
| refrigerator | | | | -0.124 (0.879) | | |
| radio | | | | 1.106 (0.766) | | |
| sewingmachine | | | | 1.365 (0.921) | | |
| motorcycle | | | | -1.074 (2.294) | | |
| Constant | 7.516 (4.209)+ | 18.270 (3.086)** | 11.429 (3.877)** | 13.911 (3.229)** | 16.598 (3.114)** | 13.911 (2.648)** |
| Observations | 561 | 561 | 561 | 561 | 561 | 561 |
| R-squared | 0.12 | 0.16 | 0.12 | 0.18 | 0.14 | 0.18 |
| Turning point age: | 51.2 | 47.8 | 53.7 | 46.1 | 47.6 | 46.1 |
| System test: Chi2 | | | | | | 365.4 |
| df | | | | | | 90 |
| P value: | | | | | | 0.000 |
| adjusted coeff: | | 1.328 | | | | |

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%
Standard errors in the LW regression have not been corrected for the two-step nature of the estimation procudure

**Table 4: The performance of the asset proxies in the Langeberg Survey allowing some assets to have independent effects**

| | (1) LW2 | (2) ln Exp2 | (3) FP Asset |
|---|---|---|---|
| employed | -0.466 | -0.738 | -0.559 |
| | (0.753) | (0.687) | (0.685) |
| age | 0.368 | 0.327 | 0.351 |
| | (0.106)** | (0.122)** | (0.122)** |
| agesq | -0.004 | -0.003 | -0.004 |
| | (0.001)** | (0.001)* | (0.001)** |
| yrseduc | -0.196 | -0.181 | -0.158 |
| | (0.086)* | (0.102)+ | (0.099) |
| Expenditure/proxy | 0.724 | 0.836 | 1.150 |
| | (0.158)** | (0.480)+ | (0.690)+ |
| numadults | 0.011 | -0.105 | -0.041 |
| | (0.159) | (0.199) | (0.192) |
| children | -0.396 | -0.360 | -0.343 |
| | (0.226)+ | (0.233) | (0.233) |
| female | 3.436 | 3.469 | 3.425 |
| | (0.602)** | (0.615)** | (0.615)** |
| black | 2.174 | 2.295 | 2.391 |
| | (1.108)+ | (1.238)+ | (1.255)+ |
| coloured | -0.016 | -0.117 | -0.094 |
| | (0.975) | (1.175) | (1.180) |
| electricity | 1.756 | 2.250 | 1.108 |
| | (0.740)* | (0.926)* | (1.221) |
| television | 2.312 | 2.436 | 1.966 |
| | (0.758)** | (0.778)** | (0.854)* |
| car | 0.897 | 1.075 | 0.695 |
| | (0.764) | (0.939) | (1.024) |
| bicycle | 0.020 | 0.154 | -0.043 |
| | (0.637) | (0.703) | (0.724) |
| Constant | 13.911 | 9.554 | 16.115 |
| | (2.632)** | (4.206)* | (3.256)** |
| Observations | 561 | 561 | 561 |
| R-squared | 0.18 | 0.16 | 0.16 |
| turn age: | 46.1 | 48.6 | 47.7 |
| System test: Chi2 | 34.4 | | |
| df | 30 | | |
| P value: | 0.265 | | |

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%
Standard errors in the LW regression have not been corrected for the two-step nature of the estimation procudure

**Table 5: The performance of the asset proxies in the KIDS survey**

| Dependent variable: BMI | (1) ln exp | (2) FP Asset | (3) ln inc | (4) Proxy all | (5) proxy tel | (6) LW |
|---|---|---|---|---|---|---|
| employed | 0.279 | 0.208 | 0.089 | 0.244 | 0.320 | 0.244 |
| | (0.376) | (0.378) | (0.382) | (0.379) | (0.378) | (0.346) |
| age | 0.360 | 0.381 | 0.377 | 0.362 | 0.373 | 0.362 |
| | (0.066)** | (0.066)** | (0.066)** | (0.066)** | (0.066)** | (0.058)** |
| agesq | -0.003 | -0.003 | -0.003 | -0.003 | -0.003 | -0.003 |
| | (0.001)** | (0.001)** | (0.001)** | (0.001)** | (0.001)** | (0.001)** |
| yrs educ | 0.005 | 0.037 | 0.042 | 0.005 | 0.086 | 0.005 |
| | (0.059) | (0.059) | (0.059) | (0.060) | (0.057) | (0.054) |
| female | 4.032 | 4.022 | 4.058 | 4.040 | 4.013 | 4.040 |
| | (0.356)** | (0.357)** | (0.357)** | (0.356)** | (0.358)** | (0.328)** |
| Exp/proxy | 1.309 | 0.778 | 0.720 | 0.094 | 0.123 | 0.350 |
| | (0.280)** | (0.217)** | (0.191)** | (0.063) | (0.062)* | (0.055)** |
| children | 0.119 | 0.211 | 0.192 | 0.172 | 0.207 | 0.172 |
| | (0.072)+ | (0.070)** | (0.070)** | (0.073)* | (0.070)** | (0.075)* |
| numadults | -0.042 | 0.030 | -0.048 | 0.025 | 0.041 | 0.025 |
| | (0.088) | (0.086) | (0.090) | (0.086) | (0.086) | (0.089) |
| indian | -4.716 | -3.837 | -3.770 | -3.602 | -3.189 | -3.602 |
| | (0.644)** | (0.601)** | (0.601)** | (0.687)** | (0.621)** | (0.586)** |
| urban | 1.107 | 0.733 | 1.035 | 0.655 | 1.082 | 0.655 |
| | (0.513)* | (0.546) | (0.516)* | (0.565) | (0.541)* | (0.547) |
| city | 1.791 | 1.670 | 1.865 | 1.575 | 1.989 | 1.575 |
| | (0.626)** | (0.642)** | (0.626)** | (0.659)* | (0.638)** | (0.666)* |
| electricity | | | | -0.422 | | |
| | | | | (0.443) | | |
| television | | | | 1.730 | | |
| | | | | (0.520)** | | |
| refrigerator | | | | 1.361 | | |
| | | | | (0.441)** | | |
| furniture | | | | -0.406 | | |
| | | | | (0.585) | | |
| jewelery | | | | -0.698 | | |
| | | | | (0.349)* | | |
| electricalappl | | | | 0.001 | | |
| | | | | (0.455) | | |
| car | | | | 0.897 | | |
| | | | | (0.494)+ | | |
| bicycle | | | | 0.360 | | |
| | | | | (0.523) | | |
| cattle | | | | -0.347 | | |
| | | | | (0.462) | | |
| sheep | | | | -0.053 | | |
| | | | | (1.184) | | |
| Constant | 7.385 | 15.567 | 10.788 | 15.028 | 15.146 | 15.028 |
| | (2.346)** | (1.764)** | (2.032)** | (1.858)** | (1.766)** | (1.502)** |
| Observations | 1444 | 1444 | 1438 | 1444 | 1444 | 1444 |
| R-squared | 0.15 | 0.15 | 0.15 | 0.16 | 0.14 | 0.16 |
| turn age: | 55.0 | 55.8 | 55.2 | 55.0 | 56.4 | 55.0 |
| adjusted coeff: | | 0.293 | | | | |
| System test: Chi2 | | | | | | 893.1 |
| df | | | | | | 100 |
| P value: | | | | | | 0.000 |

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%
Standard errors in the LW regression have not been corrected for the two-step nature of the estimation procedure

**Table 6: Asset proxies in the KIDS survey, allowing for independent effects**

| Dependent variable: BMI | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | ln exp | FP Asset | ln income | LW2 |
| empl | 0.269 | 0.264 | 0.152 | 0.306 |
| | (0.376) | (0.378) | (0.382) | (0.348) |
| age | 0.357 | 0.368 | 0.364 | 0.362 |
| | (0.066)** | (0.066)** | (0.066)** | (0.058)** |
| agesq | -0.003 | -0.003 | -0.003 | -0.003 |
| | (0.001)** | (0.001)** | (0.001)** | (0.001)** |
| yrs educ | -0.033 | 0.006 | -0.016 | -0.003 |
| | (0.060) | (0.059) | (0.060) | (0.055) |
| female | 4.050 | 4.053 | 4.071 | 4.040 |
| | (0.355)** | (0.356)** | (0.356)** | (0.330)** |
| Exp/proxy | 0.810 | 0.070 | 0.447 | 0.103 |
| | (0.316)* | (0.457) | (0.202)* | (0.039)** |
| refrigerator | 0.972 | 1.238 | 1.123 | 1.217 |
| | (0.444)* | (0.570)* | (0.434)** | (0.438)** |
| television | 1.278 | 1.488 | 1.357 | 1.503 |
| | (0.519)* | (0.546)** | (0.515)** | (0.557)** |
| electricity | -0.262 | -0.318 | -0.325 | -0.259 |
| | (0.415) | (0.561) | (0.417) | (0.418) |
| car | 0.617 | 0.899 | 0.731 | 0.859 |
| | (0.506) | (0.533)+ | (0.499) | (0.465)+ |
| bicycle | 0.092 | 0.226 | 0.208 | 0.226 |
| | (0.518) | (0.521) | (0.518) | (0.523) |
| children | 0.111 | 0.148 | 0.150 | 0.172 |
| | (0.073) | (0.072)* | (0.072)* | (0.078)* |
| numadults | -0.021 | 0.029 | -0.026 | 0.025 |
| | (0.088) | (0.086) | (0.090) | (0.089) |
| indian | -4.717 | -4.193 | -4.185 | -3.837 |
| | (0.669)** | (0.639)** | (0.639)** | (0.642)** |
| urban | 0.913 | 0.950 | 0.781 | 0.705 |
| | (0.529)+ | (0.551)+ | (0.532) | (0.554) |
| city | 1.749 | 1.938 | 1.753 | 1.741 |
| | (0.636)** | (0.647)** | (0.635)** | (0.669)** |
| Constant | 9.847 | 14.461 | 11.978 | 15.028 |
| | (2.506)** | (1.976)** | (2.090)** | (1.570)** |
| Observations | 1444 | 1444 | 1438 | 1444 |
| R-squared | 0.16 | 0.16 | 0.16 | 0.16 |
| turn age: | 54.2 | 54.8 | 54.2 | 55.0 |
| System test: Chi2 | | | | 87.4 |
| Df | | | | 25 |
| P value: | | | | 0.000 |

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%
Standard errors in the LW regression have not been corrected for the two-step nature of the estimation procedure

**Table 7: Racial differences in the correlates of BMI in the DHS**

| Dependent variable: BMI | (1) LW | (2) FP | (3) LW Black | (4) FP Black | (5) LW White | (6) FP White |
|---|---|---|---|---|---|---|
| employed | 0.240 | 0.181 | 0.341 | 0.304 | -1.119 | -1.280 |
| | (0.142)+ | (0.143) | (0.164)* | (0.164)+ | (0.515)* | (0.522)* |
| age | 0.425 | 0.425 | 0.438 | 0.441 | 0.300 | 0.321 |
| | (0.023)** | (0.023)** | (0.027)** | (0.027)** | (0.049)** | (0.049)** |
| age^2 | -0.004 | -0.004 | -0.004 | -0.004 | -0.003 | -0.003 |
| | (0.000)** | (0.000)** | (0.000)** | (0.000)** | (0.001)** | (0.001)** |
| education | 0.067 | 0.079 | 0.092 | 0.100 | -0.086 | -0.136 |
| | (0.020)** | (0.020)** | (0.023)** | (0.023)** | (0.095) | (0.101) |
| asset proxy | 0.756 | 1.103 | 0.754 | 1.055 | -0.265 | -0.427 |
| | (0.060)** | (0.104)** | (0.077)** | (0.118)** | (0.064)** | (0.420) |
| children | 0.141 | 0.137 | 0.128 | 0.132 | -0.058 | -0.103 |
| | (0.044)** | (0.044)** | (0.049)** | (0.048)** | (0.199) | (0.201) |
| numadults | -0.087 | -0.065 | -0.100 | -0.089 | 0.211 | 0.281 |
| | (0.045)+ | (0.045) | (0.052)+ | (0.052)+ | (0.183) | (0.196) |
| smoker | -2.077 | -2.096 | -2.032 | -2.081 | -0.826 | -0.756 |
| | (0.140)** | (0.142)** | (0.173)** | (0.173)** | (0.359)* | (0.359)* |
| female | 2.962 | 2.958 | 3.550 | 3.522 | -0.617 | -0.556 |
| | (0.136)** | (0.136)** | (0.163)** | (0.163)** | (0.351)+ | (0.351) |
| black | 1.473 | 1.756 | | | | |
| | (0.277)** | (0.311)** | | | | |
| coloured | 0.898 | 1.235 | | | | |
| | (0.332)** | (0.355)** | | | | |
| asianind | -1.133 | -0.806 | | | | |
| | (0.377)** | (0.394)* | | | | |
| Constant | 11.361 | 12.471 | 12.127 | 13.409 | 20.580 | 22.648 |
| | (0.641)** | (0.645)** | (0.656)** | (0.701)** | (1.720)** | (1.724)** |
| Observations | 10299 | 10299 | 7557 | 7557 | 915 | 915 |
| R-squared | 0.18 | 0.17 | 0.21 | 0.20 | 0.07 | 0.05 |
| turn age: | 56.0 | 56.6 | 56.3 | 56.5 | 50.6 | 51.3 |
| System test: Chi2 | 3788.5 | | 925.3 | | 541.0 | |
| Df | 121 | | 88 | | 88 | |
| P value: | 0.000 | | 0.000 | | 0.000 | |

Robust standard errors in parentheses (corrected for clustering)
+ significant at 10%; * significant at 5%; ** significant at 1%
Standard errors in the LW regressions have not been corrected for the two-step nature of the estimation procedure

**Table 8: Correlates of BMI in the DHS by Race and Gender**

| Dependent variable: BMI | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Black women | | Black men | | White women | | White men | |
| | LW | FP | LW | FP | LW | FP | LW | FP |
| proxy | 0.888 | 1.280 | 0.600 | 0.774 | -1.062 | -1.084 | 0.561 | 0.289 |
| | (0.106)** | (0.164)** | (0.092)** | (0.150)** | (0.323)** | (0.652)+ | (0.116)** | (0.421) |
| employed | 0.225 | 0.169 | 0.578 | 0.532 | -1.090 | -1.309 | -0.539 | -0.991 |
| | (0.259) | (0.259) | (0.193)** | (0.196)** | (0.664) | (0.656)* | (0.713) | (0.754) |
| age | 0.595 | 0.597 | 0.230 | 0.237 | 0.277 | 0.295 | 0.299 | 0.323 |
| | (0.039)** | (0.039)** | (0.033)** | (0.034)** | (0.086)** | (0.088)** | (0.073)** | (0.070)** |
| age^2 | -0.005 | -0.005 | -0.002 | -0.002 | -0.003 | -0.003 | -0.003 | -0.003 |
| | (0.000)** | (0.000)** | (0.000)** | (0.000)** | (0.001)** | (0.001)** | (0.001)** | (0.001)** |
| educ | 0.118 | 0.125 | 0.072 | 0.078 | -0.284 | -0.380 | 0.126 | 0.114 |
| | (0.033)** | (0.034)** | (0.027)** | (0.028)** | (0.163)+ | (0.158)* | (0.103) | (0.114) |
| children | 0.149 | 0.144 | 0.141 | 0.152 | -0.155 | -0.254 | 0.064 | 0.085 |
| | (0.067)* | (0.065)* | (0.055)* | (0.055)** | (0.291) | (0.291) | (0.249) | (0.257) |
| numadults | -0.125 | -0.119 | -0.105 | -0.085 | 0.380 | 0.435 | 0.037 | 0.152 |
| | (0.076)+ | (0.076) | (0.059)+ | (0.060) | (0.235) | (0.252)+ | (0.223) | (0.243) |
| smoker | -2.272 | -2.312 | -1.909 | -1.975 | -0.745 | -0.535 | -1.199 | -1.086 |
| | (0.434)** | (0.435)** | (0.177)** | (0.179)** | (0.672) | (0.661) | (0.440)** | (0.464)* |
| Constant | 11.713 | 13.449 | 17.197 | 17.801 | 27.846 | 26.369 | 15.241 | 18.650 |
| | (0.927)** | (0.988)** | (0.796)** | (0.854)** | (2.766)** | (2.791)** | (1.941)** | (1.913)** |
| Observations | 4342 | 4342 | 3215 | 3215 | 497 | 497 | 418 | 418 |
| R-squared | 0.12 | 0.12 | 0.11 | 0.11 | 0.10 | 0.08 | 0.12 | 0.07 |
| turn age: | 55.3 | 55.4 | 60.8 | 61.5 | 51.5 | 51.1 | 49.6 | 50.2 |
| System test: Chi2 | 520.1 | | 500.7 | | 304.7 | | 285.8 | |
| df | 121 | | 121 | | 121 | | 121 | |
| P value: | 0.000 | | 0.000 | | 0.000 | | 0.000 | |

Robust standard errors in parentheses (corrected for clustering)
+ significant at 10%; * significant at 5%; ** significant at 1%
Standard errors in the LW regressions have not been corrected for the two-step nature of the estimation procedure.

**Table 9: Nonlinearities in the impact of assets**

| Dependent variable: BMI | (1) FP BF | (2) FP BM | (3) FP WF | (4) FP WM |
|---|---|---|---|---|
| employed | 0.216 | 0.545 | -1.273 | -0.815 |
| | (0.258) | (0.195)** | (0.655)+ | (0.746) |
| age | 0.599 | 0.235 | 0.292 | 0.323 |
| | (0.039)** | (0.034)** | (0.087)** | (0.072)** |
| age^2 | -0.005 | -0.002 | -0.003 | -0.003 |
| | (0.000)** | (0.000)** | (0.001)** | (0.001)** |
| educ | 0.123 | 0.078 | -0.378 | 0.096 |
| | (0.034)** | (0.028)** | (0.155)* | (0.116) |
| proxy | 1.243 | 0.799 | 0.537 | 3.209 |
| | (0.163)** | (0.152)** | (1.642) | (1.420)* |
| proxy^2 | -0.585 | 0.273 | -0.498 | -0.938 |
| | (0.157)** | (0.149)+ | (0.497) | (0.425)* |
| children | 0.159 | 0.147 | -0.249 | 0.066 |
| | (0.066)* | (0.055)** | (0.292) | (0.253) |
| numadults | -0.127 | -0.082 | 0.427 | 0.130 |
| | (0.075)+ | (0.059) | (0.252)+ | (0.242) |
| smoker | -2.331 | -1.963 | -0.535 | -1.153 |
| | (0.435)** | (0.179)** | (0.660) | (0.454)* |
| Constant | 13.784 | 17.671 | 25.235 | 16.981 |
| | (0.987)** | (0.852)** | (3.101)** | (2.080)** |
| Observations | 4342 | 3215 | 497 | 418 |
| R-squared | 0.12 | 0.11 | 0.08 | 0.08 |
| turn age: | 55.5 | 61.3 | 51.1 | 49.7 |
| turn assets: | 1.06 | -1.46 | 0.54 | 1.71 |
| % below asset turning point | 94 | 0 | 2 | 45 |

Robust standard errors in parentheses (corrected for clustering)
+ significant at 10%; * significant at 5%; ** significant at 1%
Standard errors in the LW regressions have not been corrected for the two-step nature of the estimation procedure

**Table 10: Testing for independent effects of the asset proxies among Black South Africans**

| Dependent variable: BMI | (1) Men - LW1 | (2) Men - LW2 | (3) Women - LW2 | (4) Women - LW3 |
|---|---|---|---|---|
| proxy | 0.611 (0.097)** | 0.375 (0.097)** | 0.364 (0.098)** | 0.163 (0.080)* |
| employed | 0.607 (0.193)** | 0.608 (0.194)** | 0.234 (0.259) | 0.217 (0.259) |
| educ | 0.068 (0.027)* | 0.072 (0.028)** | 0.122 (0.033)** | 0.118 (0.034)** |
| age | 0.232 (0.033)** | 0.231 (0.034)** | 0.596 (0.039)** | 0.596 (0.039)** |
| age^2 | -0.002 (0.000)** | -0.002 (0.000)** | -0.005 (0.000)** | -0.005 (0.000)** |
| children | 0.137 (0.055)* | 0.137 (0.056)* | 0.149 (0.067)* | 0.149 (0.067)* |
| numadults | -0.103 (0.060)+ | -0.099 (0.060)+ | -0.126 (0.076)+ | -0.126 (0.076) |
| smoker | -1.908 (0.177)** | -1.907 (0.179)** | -2.270 (0.434)** | -2.270 (0.435)** |
| Sheep/cattle | | 0.405 (0.248) | -0.181 (0.303) | -0.068 (0.310) |
| Car | | 1.026 (0.297)** | 0.891 (0.360)* | 0.969 (0.359)** |
| Bicycle | | -0.266 (0.284) | -0.353 (0.344) | -0.378 (0.345) |
| Television | | 0.176 (0.201) | 1.208 (0.252)** | 0.960 (0.281)** |
| Electricity | | | | 0.539 (0.285)+ |
| Refrigerator | | | | 0.407 (0.302) |
| Constant | 17.144 (0.796)** | 17.156 (0.803)** | 11.713 (0.929)** | 11.713 (0.945)** |
| Observations | 3215 | 3215 | 4342 | 4342 |
| R-squared | 0.11 | 0.11 | 0.12 | 0.12 |
| System test: Chi2 | 293.9 | 32.7 | 56.2 | 19.3 |
| df | 50 | 25 | 25 | 15 |
| P value: | 0.000 | 0.139 | 0.000 | 0.202 |

Robust standard errors in parentheses (corrected for clustering)
+ significant at 10%; * significant at 5%; ** significant at 1%

**Specification 1:**
Asset proxies: telephone (rescaled) electricity radio television refrigerator computer washmachine bicycle motorcycle car sheepcattle
Covariates not used as instruments: employed educ
Instruments: age agesq children numadults smoker

**Specification 2:**
Asset proxies: telephone (rescaled) electricity radio refrigerator washmachine motorcycle
Covariates not used as instruments: employed educ sheepcattle car bicycle television
Instruments: age agesq children numadults smoker

**Specification 3:**
Asset proxies: telephone (rescaled) radio washmachine motorcycle
Covariates not used as instruments: electricity refrigerator television car bicycle sheepcattle employed educ
Instruments: age agesq children numadults smoker

Standard errors in the LW regressions have not been corrected for the two-step nature of the estimation