A Cohort Analysis of Household Income, Consumption and Saving

John Gibson
University of Waikato

and

Grant Scobie
The Treasury

Abstract

This paper presents a cohort analysis of household income, consumption and saving in New Zealand. It is based on an analysis of unit record data from March years 1984 to 1998 taken from the Household Economic Survey (HES). These data are a series of cross-sectional surveys rather than a true panel, so we construct synthetic cohorts and use a range of regression models to separate out the effect of age, birth-year cohort and survey year on income, consumption and saving rates. There appears to be a "V" shaped cohort pattern in household income and saving, such that the age profile of saving shifted down for the cohorts born between 1920 and 1939 relative to the younger and older cohorts studied.

JEL: D12, D91

Acknowledgements: We are grateful to Ivan Tuckwell for assistance with the data and to two anonymous referees and seminar participants at the Treasury, the University of Waikato and the New Zealand Association of Economists Conference in Christchurch for helpful comments.

I. Introduction

In many countries, concern has been raised by apparent falls in household saving rates. Low levels of household saving are linked to falling national saving and growing current account deficits on the one hand, and concern for the adequacy of retirement income on the other. Typically saving is measured from the national accounts as the difference between income and consumption and in New Zealand this approach shows an apparent decline in saving. But aggregate saving could be observed to be quite low (or even negative) without any "problem" for household retirement incomes. In an ageing population, households in their working lives could be saving while those in their retirement years were drawing down their saving and on balance there may be little net saving observed, even with high levels of saving among the saving households. Thus, if we really want to understand what is happening to saving we have to examine the dynamics of saving patterns using data on individual households.

To help study these issues, a Household Saving Survey (HSS) is being conducted in 2001, by Statistics New Zealand for the Office of the Retirement Commissioner. However, even when the results of this survey are available, they will simply provide a baseline for estimating net household worth, which is a stock variable. To get a flow measure of saving, the HSS will have to be repeated to give measures of the change in net worth over time. Therefore, some other source of data is needed if the dynamics of saving patterns for individual households are to be studied more immediately.

In this paper we use records for individual households from the 1984-98 Household Economic Surveys (HES) to examine patterns of household saving behaviour in New Zealand. The HES does not provide direct estimates of saving but it does provide details on income and expenditure from which saving can be derived as a residual. These data are a series of cross-sectional surveys rather than a true panel, so we construct synthetic cohorts using methods outlined by Paxson (1996) and Deaton (1997). The analysis pays particular attention to age (i.e., lifecycle) and cohort effects in saving rates. These are examined jointly with the age and cohort effects in household income and consumption to give some insights into the shifting patterns of income distribution in New Zealand.

Most of the analysis focuses on households rather than on the individuals who are the decision-making units in theoretical treatments of saving behaviour. This approach can be justified by the fact that many consumption and saving choices are taken at household level (Attanasio, 1998). But it is not completely satisfactory because it requires dating the household according to the birth year and age of the household head, yet households may contain members at many different stages of their lifecycles. Therefore, we also report some sensitivity analyses that are based on an individual version of an empirical lifecycle model, following the lead of Deaton and Paxson (2000).

The next section discusses the HES data and the way that household saving is constructed from these data. Basic descriptive statistics on the trends in the estimated saving rates and on the pattern of saving across income groups are also presented. Section 3 introduces the concept of synthetic cohorts, which form the basis of our analysis. The following section presents the age and cohort decompositions of income, consumption and saving. This section also reports results with controls for survey year added to the model. In Section 5, an attempt is made to relax the constraints that are imposed by using households as the unit of analysis, with an age and cohort decomposition based on estimates of individual's income, consumption and saving. Sensitivity analyses using different definitions of saving are reported in Section 6, while the summary and implications are in Section 7.

2. Data Description

The measure of household saving used here is based on the residual when current consumption expenditure is subtracted from household disposable income. The estimate of current consumption expenditure ("consumption" for short) removes from the HES expenditure estimates items that provide consumption benefits over more than one year. In total, 17 types of expenditure are removed, with the main categories being household durable goods and vehicles, payments for education, life and health insurance, and mortgage principal repayments and contributions to savings.[1]

It is important to note that the basic approach, of comparing household income and expenditure estimates to derive saving, is one that Statistics New Zealand warn against.[2] This concern partly reflects the fact that by taking a large value with a high sampling error (income), and subtracting another large value with a high sampling error (expenditure), the outcome is a small value (implicit savings) surrounded by a large sampling error. Our defence in disregarding this warning has several elements. First, analysts in other countries have successfully used similar data, particularly the U.K. Family Expenditure Survey (Attanasio and Banks, 1998), and the U.S. Consumer Expenditure Survey (Attanasio, 1998). Our results appear to be quite consistent with these previous studies, and if statistical chance was the only factor operating, it is odd that the same patterns would repeat across countries. Instead, because much of the variation in our estimated saving rates is explained by economic and social phenomena, we argue that the HES data do give a useful perspective on changes in savings though time, although there may be doubt about whether the overall level of savings is accurately measured by these data (see below). Our final defence is that there is currently no other source of micro data for examining household saving behaviour in New Zealand.

A total of 50,624 households are available in the HES sample over the 15 years from 1983/84 (referred to by the latter year). But the sample was truncated by removing households with negative disposable income (*n*=330) and also those whose head was less than 19 years old or greater than 74 at the time of the survey. Those younger than 19 were not considered important for studying lifetime saving patterns, while amongst the elderly, the HES does not cover those living in rest homes. Thus, without this age limit the elderly in the sample would be increasingly likely to be those with the means to maintain their own households, who are presumed to come from the wealthier part of the population. These sample deletion rules left 46,269 households.

Although most of the analysis will be based on cell averages and quantiles, with cells formed from the interaction of age and survey year (see Section 3), we also provide some basic statistics calculated from the individual household saving estimates. Rather than taking the ratio of saving to disposable income, we follow Attanasio (1998) and calculate saving rates by the ratio of saving to consumption (S/X). This has the advantage of being defined even when reported disposable income is zero and is a monotonic transformation of the more usual ratio of saving to income.[3] Table 1 contains the number of observations in each survey year together with the estimated household saving rate at the mean, median, 25th and 75th percentiles of the distribution of S/X. The final column reports the ratio of the averages of saving to consumption, as distinct from the average of the ratios.

(Table 1 about here)

The mean saving rate of 0.35 is, of course, not comparable with any other estimates because it reflects the items excluded from household expenditures.[4] But the estimates should be comparable across years. The trend in the mean saving rate, established by regressing the mean for each year in Table 1 on a time trend, is for the saving rate to rise by an average of

0.4 percentage points per year. Thus there is an important contradiction between the trend in the household data and the trend in the national accounts data, although no attempt is made here to resolve this discrepancy. In all survey years, households at the 25th percentile of the S/X distribution are dis-saving, with current consumption appearing to exceed disposable income. Moreover, the trend is for the saving rate at the 25th percentile to become more negative in latter survey years,[5] in contrast to the 75th percentile where the saving rate is rising. Indeed, it appears that the rise in the mean saving rate over time is being generated in the upper part of the distribution because there is no trend in the median of S/X.

How do saving rates vary with household income? Table 2 reports saving rates and total savings by household disposable income decile, where these deciles have been formed from the pooled sample across survey years after an initial deflation by the CPI.[6] Over one-half of total household savings over the 15 year period are generated by those in the richest decile of households, while the three lowest income deciles contribute negative savings. Of course, not all households in these lower income groups have negative saving rates, as shown by the saving rates at the 75th percentile within these three deciles. However, even within household income deciles 4-6 there is some negative saving towards the bottom of the distribution, and overall, 30 percent of households in the pooled sample appear to have negative saving rates.

(Table 2 about here)

It is also apparent that there is growing inequality in household saving rates. The lower three income deciles show a negative trend in their mean saving to consumption ratio across the survey years, while the upper income deciles show an increasingly positive trend. Because the saving rate is falling at the lower end of the distribution and rising at the upper end, it must be that saving rates are becoming more dispersed over time. Indeed, Lorenz curves estimated for the beginning and end years of the survey period indicate just such a shift (Gibson and Scobie, 2001). This increase in the inequality of saving parallels the increased inequality in the distribution of income which has been reported by other investigators using New Zealand data (Stephens *et. al*., 2000).

3. The Cohort Approach and Estimation Framework

Lifecycle saving profiles would ideally be studied with panel data, where the same people are tracked over time. But in the absence of such data, a time-series of cross-sectional surveys can be used to form *synthetic panels* (Deaton, 1985).[7] In the current context, synthetic panels require following birth-year cohorts across successive Household Economic Surveys. In the absence of significant immigration, emigration and mortality, each successive survey lets us track movements in the average behaviour of each cohort over time and the estimates from these synthetic panel data should be consistent with estimates from genuine panel data (Deaton, 1997). For example, the average saving rate of 30-year olds in the 1985 survey could be connected to the average saving rate of those who are 31-years old in the 1986 survey because both averages refer to the cohort born in 1955.

Synthetic panels avoid attrition bias because they are constructed from fresh samples each year. There also may be less bias due to measurement error problems because we are typically working with a cohort average (or some other quantile), which should reduce the impact of idiosyncratic variability that is a feature of data on individual units. However, the small size of the HES sample means that many of the cell averages represent small samples when they are formed by the interaction of birth-year and survey year (although similar results are obtained using broader cohorts defined by 5-year birth intervals). A further difficulty, when using data on households, is that issues of household dissolution and reformation matter. For example, the elderly may be absorbed into younger households (e.g., going to live with their children), if they lack sufficient wealth to maintain an independent household , so previously "old" households become "young" households in subsequent years.[8]

Estimation Framework

According to the lifecycle model, a person saves at one stage of his or her life to consume in another period.[9] Therefore, saving behaviour should differ for different individuals at different stages of their lifecycles and may also evolve over time and across (birth-year) cohorts as economies grow and as certain fluctuations affect individuals contemporaneously. Under this lifecycle model, consumption is proportional to lifetime wealth, with a factor of proportionality that depends on age and the real interest rate. Ignoring the effect of interest rates, for individual i who is observed in year t and was born in year b,

$$c_{ibt} = g_i(t-b)W_{ib} \tag{1}$$

where c is consumption and W is wealth, and noting that age, $a = t - b$. Adapting the model to households, lifetime resources are assumed to be set at the time of household formation and the task is to allocate consumption over time, according to the household's preferences as represented by the function g (Deaton, 1997). Taking logarithms and averaging over all households whose head is in the cohort born at time b and observed at t:

$$\overline{\ln c}_{bt} = \overline{\ln g}(t-b) + \overline{\ln W}_{b} \tag{2}$$

so that average consumption is the sum of two components, one of which depends only on age and one of which depends only on birth-year cohort.

Equation (2) can be estimated by regressing the sample average of the logarithm of consumption for each cohort in each survey year on a set of age and cohort dummy variables. But following Deaton (1997) the average number of elderly (n^e) and prime-age (n^a) adults and children (n^c) per household in each cohort×year cell is also included in the regression to control for differential consumption requirements.[10] Therefore, the equation to be estimated is:

$$\overline{\ln c_{bt}} = D^a a_c + D^b g_c + d_1 n^c + d_2 n^a + d_3 n^e + u_c \tag{3}$$

where D^a is a matrix of age dummies, D^b is a matrix of cohort (birth-year) dummies, the coefficients a_c and g_c are the age and cohort effects in consumption, the *d*'s control for demographic composition and u_c is the error surrounding the sample estimate of cell average log consumption. In addition to decomposing the mean of ln c_{bt} into age and cohort effects, equation (3) is also applied to the cell median and the 25th and the 75th percentile of (log) household consumption.[11]

The average income for different birth cohorts can also be tracked over time, where the underlying relationship assumes that income at any age can be expressed proportional to lifetime resources, with the proportionality factor depending on age (Deaton and Paxson, 2000). Thus, the equation for log household incomes, where cohorts are again based on the year of birth of the head, is:

$$\overline{\ln y_{bt}} = D^a a_y + D^b g_y + d_1 n^c + d_2 n^a + d_3 n^e + u_y \,. \tag{4}$$

The difference between the logarithm of income and the logarithm of consumption is a monotone increasing function of both the saving-to-income and the saving-to-consumption ratios and obeys the inequality:

$$s/y \leq \left[\overline{\ln y} - \overline{\ln c}\right] \leq s/x$$

so the cohort-age decomposition of log consumption and log income automatically gives a cohort-age decomposition of the savings ratio (Deaton, 1997). Subtracting equation (3) from (4) gives:

$$s/x \approx \overline{\ln y} - \overline{\ln c} = D^a\left(a_y - a_c\right) + D^b\left(g_y - g_c\right) + u_y - u_c \tag{5}$$

where $(a_y - a_c)$ is the estimated age effect in the saving rate and $(g_y - g_c)$ is the estimated cohort effect.

This age-cohort decomposition of log consumption and log income is based on samples with 840 observations, where those observations are themselves cell averages (or quantiles). These cells are formed by the interaction of 15 survey years (1984-1998) and birth-year cohorts that ranged from 1910 to 1979. Note however that the cohorts at either end of the range do not appear in all 15 surveys because the sample is restricted to households whose head was between the ages of 19-74 at the time of the survey. The smallest number of households used to construct a cohort×year cell is 11 (for 19-year-olds in 1987), while the median cell size is 53.[12] The March quarter CPI was used to convert all monetary values to December 1993 prices, prior to forming the cell averages of household incomes and consumption.

4. Results

Figure 1 shows the age effects in consumption and income for the mean household, and also at the median, and 25th and 75th percentiles. Average consumption appears to peak by the time household heads reach age 40, while the age profile of average household income is everywhere higher and appears to peak somewhat later, although there is some variability in the graph.[13] Consequently, the saving rate (shown by the graph with a bold line and using the scale on the right-hand side), peaks in the decade after the household head reaches age 50 and then declines somewhat in the 60's but still remains well above zero. This apparent lack of dissaving amongst households headed by the elderly may reflect bequest motives or differential mortality (Johnson and Stears, 1998; Attanasio and Hoynes, 2000), whereby the poor – who are likely to have low or negative saving rates – die earlier than the rich and so cease to be a downward drag on the average.

(Figure 1 about here)

The age effects in median consumption, income and saving rates appear to be more favourable to the lifecycle hypothesis, with dissaving apparent at both young and old age. The 'hump' in savings for the decade beginning about age 50 is also more apparent than at the mean, with saving rates rising by about ten percentage points, which is equivalent to one-third of the amplitude shown over the lifecycle. It appears that as household heads advance past age 60 median household incomes fall more rapidly than do mean incomes and it is this, rather than differences in the age profiles of consumption at the mean and the median, which causes the differences in the age-saving profiles.

The bottom two panels of Figure 1 show the age profiles at the lower and upper end of the distribution. At any age, the income of households at the 25th percentile of the within-cell distribution is less than the current consumption expenditure of households at the 25th percentile of the expenditure distribution, while the reverse pattern holds at the 75th percentile. This pattern may also reflect the greater dispersion of household incomes compared to consumption. The 'hump' in savings is less apparent amongst the richer households, with a near-linear trend in the savings rate from age 25 to age 55. The drop in saving rates after age 60 is also less marked for these rich households.

The cohort effects are shown in Figure 2, where cohorts are labelled according to their age in the first year of the survey (1984), so older cohorts are further to the right. There is considerable year-to-year variability in these cohort graphs, but the cohort effect in mean, median and 25th percentile consumption appears to be rising slightly with age in 1984. In other words, at any given age, the real consumption level of households whose head comes from an earlier birth year is somewhat higher than the consumption level of the more recent cohorts. In contrast, there is no overall trend in the cohort effect for the consumption of rich households, with both younger and older cohorts appearing to have higher consumption than the cohorts whose household head was middle-aged in 1984. These cohort effects in real consumption may be consistent with the reports (see, for example, Stephens *et al*., 2000) that living standards for many households in the middle and bottom of the distribution may have been stagnant or even declining since the period of economic reforms beginning around 1984.[14]

(Figure 2 about here)

The cohort effects in mean and median household income also rise slightly with age at time of the first survey, so, for example, a household headed by someone in their 20's in 1984 (i.e., born 1955-64) has real income that is two percent lower than a household in its 20's had 40 years before (born 1915-24). But much more apparent than this slight trend is the 'dip' in average household incomes for those households headed by someone born between ca. 1925-1939. For example, a household headed by a 49-year old in 1984 has income that is about 12 percent lower than the income of a 49-year old household 11 years later, in 1995.[15]

One possible explanation for this 'dip' in incomes is that the household heads aged ca. 45-55 in 1984 were especially vulnerable to the effects of New Zealand's economic restructuring from the mid-1980s onwards, and so they suffered a permanent fall in their lifetime incomes. But for this explanation, it is puzzling that the dip in incomes for those born ca. 1925-1939 is much more apparent for rich households than the poor (see the bottom panels of Figure 2). Accounts of New Zealand's economic restructuring often suggest that it has been the poor who have borne the brunt, so one would expect to see the fall in incomes for the 1925-1939 cohorts to be more apparent at the 25th percentile and less apparent at the 75th percentile.

The combination of little change in average consumption across cohorts and a dip in the average incomes of certain birth-year cohorts produces a noticeable cohort effect in the saving rate. Both younger and older cohorts have higher average household saving rates than do those households whose head was between 45 and 60 years old in 1984 (i.e., born between 1924-1939). The fall in the saving rates for households with heads in this age range is about 10 percentage points at the mean and median and about 15 percentage points at the 75th percentile. The pattern of cohort effects at the 75th percentile is almost 'V' shaped, so it is possible that downward trends in aggregate saving rates might be temporary; older cohorts with higher saving rates have been replaced by middle-aged cohorts with low saving rates, but those middle-aged will, in turn, be replaced by younger cohorts with high saving rates.

This negative cohort saving pattern for the middle cohorts is remarkably similar to what Attanasio (1998) observes for the United States. A possible explanation for this pattern that is suggested by Attanasio is that it reflects the increases in social security entitlements enjoyed by those cohorts. The results presented here may also be consistent with the evidence reported by Thomson (1996) that the welfare state in New Zealand has been most generous to those born between about 1920 and 1945. These birth years correlate closely with the cohorts shown in Figure 2 to have lower lifetime saving rates.

Adding Year Effects

In a world without uncertainty, knowing the age and birth year of an individual may be sufficient information for predicting their saving behaviour: Variation in lifetime income across birth-year cohorts would shift their lifecycle saving profile, while their age would indicate their position along the profile. However, saving behaviour may also change over time as various macroeconomic fluctuations affect individuals contemporaneously. Hence, controls for time (i.e., survey year) may be needed to capture things like macroeconomic shocks that surprise all members of a cohort (e.g., faster economic growth than expected). However, these controls require more than simply adding a set of survey-year fixed effects, D^{a+b} ($a+b=t$, $t+1$, .., T, where t and T are the first and last available cross-sections) because the year in which each household is observed equals the age of the household head, a plus their year of birth, b, so the separate effect of a, b, and $a+b$ cannot be identified. Therefore, we follow a common practice in the literature on household saving of including year effects in a normalised form so that they sum to zero and are orthogonal to a time trend:[16]

$$\sum_{a+b=t}^{T} D^{a+b} = 0$$
$$\sum_{a+b=t}^{T} (a+b) \cdot D^{a+b} = 0$$

These time effects reflect additive macroeconomic shocks or the residual influence of non-systematic measurement error (Jappelli, 1999).

(Figure 3 about here)

Figure 3 plots the age, cohort, and survey-year effects in mean saving rates. The panel at the bottom right-hand side shows that all three effects are statistically significant at the $p<0.01$ level. Comparing Figure 3 with the saving rate graphs in Figures 1 and 2 shows that the addition of the time effects does not alter the basic shape of the age-profile of saving rates but the cohort effect becomes more pronounced. Specifically, the fall in saving rates for those households whose head was aged ca. 40-60 in 1984 is more apparent once the time effects are included. Gibson and Scobie (2001) also report the age, cohort and time effects at the medians and 25th and 75th percentiles of the saving rate and these are found to be similar to the patterns when year effects are excluded. This similarity may indicate some robustness in the age and cohort patterns described above.[17]

5. Sensitivity Analysis with "Individual" Saving Rates

Using households as the unit of analysis and dating them by the age of the household head may obscure some patterns in saving behaviour. For example, if low savers amongst the elderly move in with their children once they have exhausted their financial resources, while high savers continue as household heads, the age profile of saving will not decline with age in a way that reflects the average behaviour of the elderly.

There are enormous difficulties in retrieving estimates of individual income, consumption and saving from the household data that we have available. However, treating the results just as a sensitivity analysis, we follow the example of Deaton and Paxson (2000) in estimating an individual version of an empirical lifecycle model. The key assumption of this approach is that households are simply a veil for the individuals within them, with no effect on anyone's income, consumption, or saving (that is, scale economies and specialisation between market and home production are ignored). Thus, the β's recovered from a cross-sectional regression of household consumption c_{ht} on the numbers of people in the household in each age n_{aht}, serve as estimates of "individual" consumption:

$$c_{ht} = \sum_{a=0}^{N} n_{aht} b_{at} + v_{ht}, \qquad t = 1, .., T \qquad (6)$$

A similar approach can be used to measure "individual" income, and the individuals within households are assumed to each follow their own lifecycle trajectory, in which both consumption and income at each age is the product of an age effect and a lifecycle wealth effect. Hence, a counterpart to equation (1) is:

$$b_{at} = f(a) W_{t-a} \qquad (7)$$

where *t-a* gives the year of birth. These measures of individual consumption (and income) can be decomposed into age effects and cohort (wealth) effects. As Deaton and Paxson (2000) point out, provided that the β's are positive, this decomposition can be performed by pooling the β's estimated from all of the survey years, taking logarithms and regressing on a set of age and cohort dummies. This corresponds to using equations (3) and (4), which is the first and only stage when working with household-level data, because the β's serve as an estimate of the average consumption (or income) for people of age *a*, observed in year *t*.

The major practical problem, however, is that the regression in equation (6) does not always result in positive values for the β's, especially for those at very young and very old ages, so the decomposition is initially not possible because it needs the logarithm of negative values. Therefore we use an inequality constrained estimator to ensure the non-negativity of the estimated β's for individual consumption and income. By using non-linear least squares to estimate equation (6), but with the β's squared, we implicitly find the set of age-specific individual consumptions that minimise the residual sum of squares subject to the constraint that they be non-negative. For example, if we write a fragment of equation (6):

$$c_{ht} = \ldots + b_{19ht} n_{19ht} + b_{20ht} n_{20ht} + \ldots + v_{ht}$$

what we estimate is:

$$c_{ht} = \ldots + (b_{19ht} \cdot b_{19ht}) n_{19ht} + (b_{20ht} \cdot b_{20ht}) n_{20ht} + \ldots + v_{ht}$$

so that regardless of whether the estimates produced by the non-linear least squares algorithm, say, $\hat{b}_{20ht}$, are positive or negative, they must be squared (and hence, become positive) in order to restore the underlying linear regression in equation (6).[18] We also impose some further structure, by requiring the β's to be equal within the following age groups: 0-4, 5-9, 10-14, 15-19 and 80-99 years. This helps with the convergence of the non-linear least squares algorithm and should not affect our results, because we focus on the saving behaviour of individuals between ages 20-75.

Figure 4 plots the derived saving rates (more precisely, the difference between the logarithms of income and consumption) for individuals and households.[19] The age profiles for the individual saving rates are considerably more hump-shaped than are the profiles using the household data and the age of the household head. This follows from the fact that both approaches should have the same mean saving rate (under the assumption that households have no effect on anyone's income or consumption) but the individual approach allows greater dissaving amongst the young and the old. Deaton and Paxson (2000) also find that the individual approach gives age profiles that are more in accord with the lifecycle hypothesis. Although the individual approach shows a greater rise in the saving rate, from approximately age 40 onwards, once the age effects are rescaled (the bottom panels of Figure 4) the age patterns are similar except in the retirement phase. The household-level saving rate rises as the household head's age rises from the early 60s to the mid 70's while the individual saving rate declines. This discrepancy may reflect the tendency of the less wealthy elderly to be absorbed into younger households (e.g., moving in with their children), causing the household data to reflect a shifting wealth distribution.

(Figure 4 about here)

The cohort effects in the individual and household saving rates follow the same pattern as the age effects, in the sense that using the individual approach uncovers much greater variability. But once the cohort effects are rescaled, in the bottom panel of Figure 4, a similar group emerge as having the lowest saving rates. This group is comprised of those who were aged between ca. 45 and 60 years in 1984 (or equivalently, were born in ca. 1924-39). Hence, using the available methods for deriving individual saving rates shows that the pattern of cohort effects discovered in the household data also is reproduced in the individual data.

6. Sensitivity Analyses with Different Definitions of Saving

The definition of consumption used when deriving the saving rate excludes items that are more properly considered as forms of investment and hence are a type of saving. In this section, we assess whether the cohort effects that we have found previously are sensitive to removing some of these items from the savings estimate. In particular, we remove from saving and add back to "current" consumption expenditure: purchases of durable goods; expenses on education, medicine, life and health insurance; and repayments of mortgage principal and contributions to savings.

When spending on durable goods is considered as a form of consumption rather than investment, the average ratio of saving to consumption is halved, from 0.353 to 0.177. But this change in the definition of saving does not appear to alter the basic cohort pattern, of lower saving rates for those households whose head was ca. 50-60 years old in 1984. In fact, comparing the first two panels of Figure 5, it appears that the cohort effect is slightly amplified when expenses on durables are treated as consumption rather than saving.

(Figure 5 about here)

Including either education, medical, and insurance payments (or mortgage principal repayments and contribution to savings) as forms of consumption reduces the average saving to consumption ratio to 0.293 (0.263). The lower two panels of Figure 5 show that these adjustments also have less impact on the size of the cohort effects than did the inclusion of durables spending in consumption. But despite this flattening in the graphs for the cohort effects, the basic shape is unaltered. Thus, the overall pattern shown in Figure 5 highlights the robustness of the relative cohort effects to the definition of saving used, and lower lifetime saving rates consistently located amongst those households whose head was born somewhere between 1924 and 1939.

7. Summary and Implications

This paper has used Household Economic Survey data and synthetic panel techniques to examine household saving behavior in New Zealand. Amongst the findings to emerge from the analysis are that a very large share of total household saving comes from a small number of high-income households. Approximately 30 percent of all households used in the analysis reporting negative saving. The distribution of household saving also appears to be becoming more unequal. The trends in the household data show saving rates rising rather the drop observed in the aggregate data, so a reconciliation of these differences is a task for future research.

By imposing some identifying restrictions on the data we are able to separate out the effect of age, cohort and year on consumption, income and saving rates. This decomposition suggests that consumption largely tracks income over the lifecycle but appears to peak earlier. Consequently there is somewhat of a 'hump' in saving rates for households whose head is between 50-60 years old. This 'hump' pattern is especially apparent when using cell medians.

There is a slight linear trend in consumption and income across cohorts, with the later born appearing to have lower lifetime wealth but the more noticeable effect is the decline in incomes for those cohorts aged ca. 45-55 in 1984. This produces a corresponding fall in saving rates for the cohorts whose household head was born somewhere in the decades between 1920 and 1939. This pattern also persists when controls for survey year are introduced and when an attempt is made to estimate individual consumption, income and savings.

Why these groups should have significantly lower lifetime saving rates than those born before or after this period remains an unanswered question. Some tentative attempts to correlate these cohort effects with policy variables are reported in Gibson and Scobie (2001), where the evidence seems consistent with the hypothesis that the provision of higher state benefits or greater certainty dampens the incentive for private saving. Perhaps stronger evidence comes, however, from the similarity of the cohort pattern found here with what Attanasio (1998) has observed in the United States, where lifetime saving rates are lowest for the cohorts born between 1920 and 1939. The hypothesis advanced by Attanasio, that this cohort pattern reflects increases in social security entitlements, remains a subject that requires on-going research.

References

Attanasio, O. 1998. A cohort analysis of saving behavior by U.S. households. *Journal of Human Resources* 33(3): 575-609.

Attanasio, O. and Banks, J. 1998. Trends in household saving: a tale of two countries. *Working Paper* No 98/15, The Institute of Fiscal Studies, London.

Attanasio, O., and Hoynes, H. 2000. Differential mortality and wealth accumulation. *Journal of Human Resources* 35(1): 1-29.

Browing, M., and Crossley, T. 2001. The life-cycle model of consumption and saving. *Journal of Economic Perspectives* 15(3): 3-22.

Deaton, A. 1985. Panel data from time series of cross sections. *Journal of Econometrics* 30(1): 109-126.

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

Deaton, A., and Paxson, C. 2000. Growth and saving among individuals and households. *Review of Economics and Statistics* 82(2): 212-225.

Gibson, J. and Scobie, G. 2001. Household saving behaviour in New Zealand: a cohort analysis. *Treasury Working Paper 01/18*.

Jappelli, T. The age-wealth profile and the life-cycle hypothesis: a cohort analysis with a time series of cross-sections of Italian households. *Review of Income and Wealth* 45(1): 57-75.

Johnson, P., and Stears, G. 1998. Why are older pensioners poorer? *Oxford Bulletin of Economics and Statistics* 60(3): 271-290.

Paxson, C. 1996. Saving and growth: evidence from micro data. *European Economic Review* 40(2): 255-288.

Shorrocks, A. 1975. The age-wealth relationship: a cross-section and cohort analysis. *Review of Economics and Statistics* 57(1): 155-163.

Stephens, R., Frater, P. and Waldegrave, C. 2000. Below the line: an analysis of income poverty in New Zealand, 1984 - 1998", GSBGM Discussion Paper 2/00, Victoria University of Wellington.

Thomson, D. 1996. *Selfish Generations? How Welfare States Grow Old* Bridget Williams Books: Wellington.

White, K. 1993. *Shazam User's Reference Manual Volume 7.0* McGraw-Hill, New York.

Table 1: Sample Size and Savings Rates by Survey Year

	Total Sample	Savings Rate (S/X) Mean	25th Percentile	Median	75th Percentile	$\frac{\bar{S}}{\bar{X}}$
1984	3331	0.376	-0.058	0.227	0.613	0.189
1985	3295	0.287	-0.106	0.168	0.498	0.139
1986	3174	0.318	-0.078	0.201	0.551	0.177
1987	3210	0.341	-0.084	0.209	0.581	0.204
1988	4021	0.347	-0.043	0.212	0.563	0.200
1989	3142	0.358	-0.080	0.207	0.601	0.200
1990	3047	0.313	-0.110	0.188	0.560	0.166
1991	2674	0.340	-0.099	0.193	0.575	0.227
1992	2712	0.380	-0.062	0.217	0.609	0.251
1993	4244	0.415	-0.057	0.222	0.621	0.270
1994	2839	0.338	-0.096	0.197	0.546	0.235
1995	2695	0.336	-0.110	0.191	0.607	0.234
1996	2621	0.355	-0.111	0.186	0.572	0.246
1997	2642	0.359	-0.091	0.193	0.578	0.259
1998	2622	0.428	-0.081	0.238	0.676	0.320
Total	**46269**	**0.353**	**-0.086**	**0.202**	**0.584**	**0.222**

Note: Year is the year the survey ended, so 1984 is the survey year from April 1983 – March 1984.

Table 2: Savings by Household Disposable Income Decile

	Mean Income	Total Savings ($m)	Trend in mean S/X	Savings Rate (S/X) Mean	25th Percentile	Median	75th Percentile	$\frac{\bar{S}}{\bar{X}}$
1	8540	-8720	-0.005	-0.115	-0.539	-0.166	0.184	-0.404
2	14188	-2870	-0.005	0.088	-0.247	-0.004	0.292	-0.118
3	17698	-1150	-0.001	0.153	-0.171	0.056	0.339	-0.041
4	21807	96	-0.006	0.190	-0.150	0.091	0.397	0.003
5	26428	2753	0.002	0.275	-0.091	0.146	0.479	0.074
6	31400	5885	0.003	0.343	-0.031	0.224	0.547	0.142
7	36933	9547	0.002	0.424	0.035	0.287	0.629	0.208
8	43686	13790	0.005	0.493	0.083	0.348	0.692	0.266
9	53357	21280	0.007	0.614	0.164	0.448	0.851	0.361
10	86465	52390	0.036	1.069	0.294	0.669	1.299	0.675
ALL	34049	93001	0.004	0.353	-0.086	0.202	0.584	0.222

Note: Decile 1 is poorest and decile 10 is richest and the deciles are formed from the combined sample of households across all 15 survey years. All monetary figures are in December 1993 values, with the March Qtr CPI used as the deflator. Total savings are across all 15 survey years. The time trend in mean S/X comes from a linear regression on a time index of the savings rate for each decile in each year.

Appendix Table A: Wald tests of joint statistical significance of age and cohort effects in the decompositions of income and consumption

	ln (disposable income)		ln (current consumption)	
	Age	Cohort	Age	Cohort
Mean	$\chi^2_{(55)}$= 347.55 *p*<0.000	$\chi^2_{(69)}$= 124.49 *p*<0.000	$\chi^2_{(55)}$= 229.33 *p*<0.000	$\chi^2_{(69)}$= 139.64 *p*<0.000
25th percentile	$\chi^2_{(55)}$= 207.94 *p*<0.000	$\chi^2_{(69)}$= 155.45 *p*<0.000	$\chi^2_{(55)}$= 166.96 *p*<0.000	$\chi^2_{(69)}$= 118.46 *p*<0.000
50th percentile	$\chi^2_{(55)}$= 400.08 *p*<0.000	$\chi^2_{(69)}$= 152.86 *p*<0.000	$\chi^2_{(55)}$= 209.08 *p*<0.000	$\chi^2_{(69)}$= 112.93 *p*<0.001
75th percentile	$\chi^2_{(55)}$= 668.65 *p*<0.000	$\chi^2_{(69)}$= 181.68 *p*<0.000	$\chi^2_{(55)}$= 191.04 *p*<0.000	$\chi^2_{(69)}$= 111.68 *p*<0.001

Note: Results are Wald tests of the hypothesis that coefficients on either age or cohort dummy variables are jointly zero in a seeming-unrelated regression model where ln Y and ln X are each regressed on age and cohort dummies and on cell averages of the number of children, adults and old people per household. The sample is 840 cell averages (medians or means as noted), where cells are based on the interaction of age (19-74 inclusive) and survey year (1984-98).

Figure 1: Age Effects in (log) Income and Consumption and in the Derived Saving Rate [ln(y)-ln(c)]

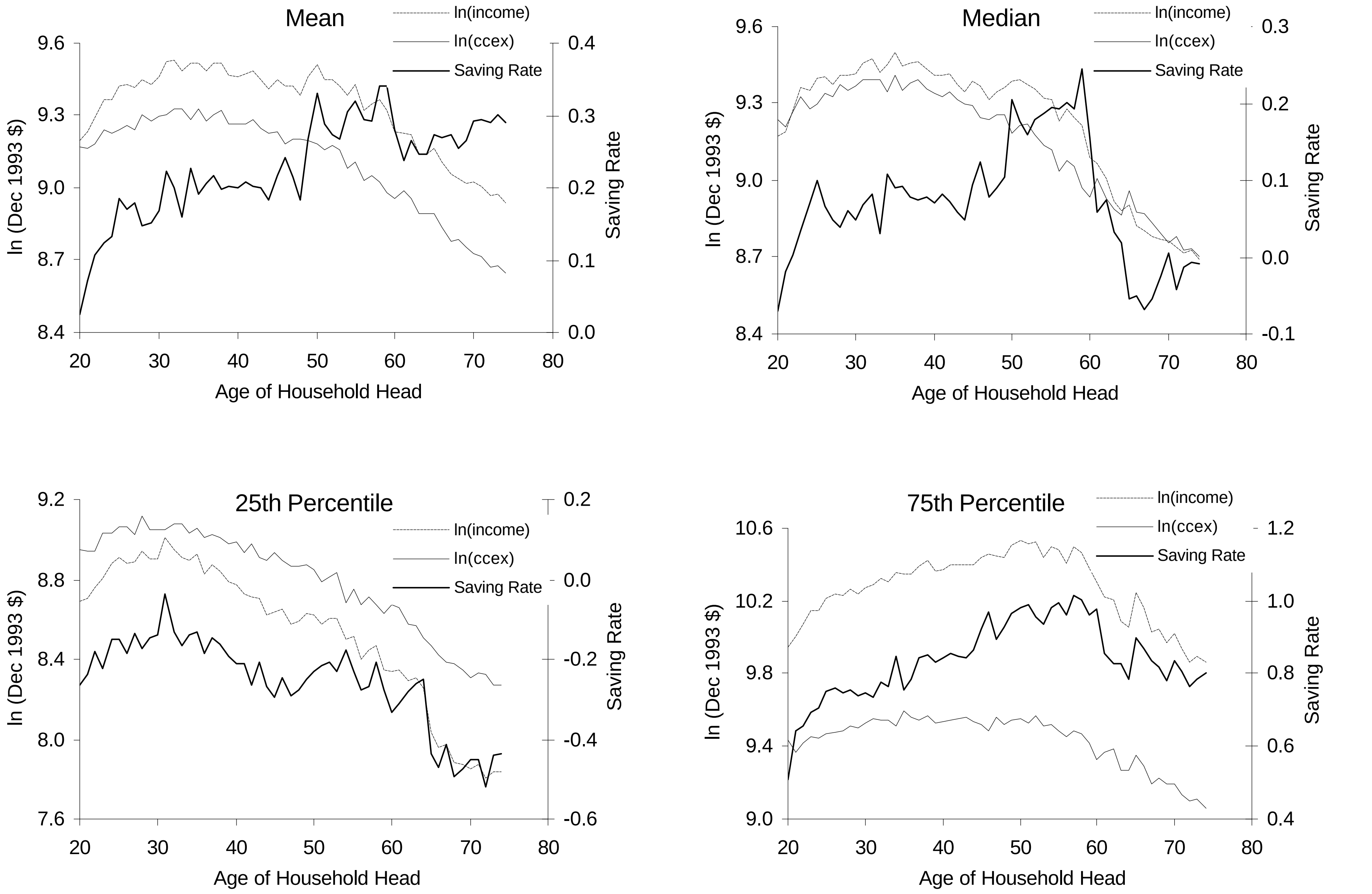


Figure 2: Cohort Effects in (log) Income and Consumption and in the Derived Saving Rate [ln(y)-ln(c)]

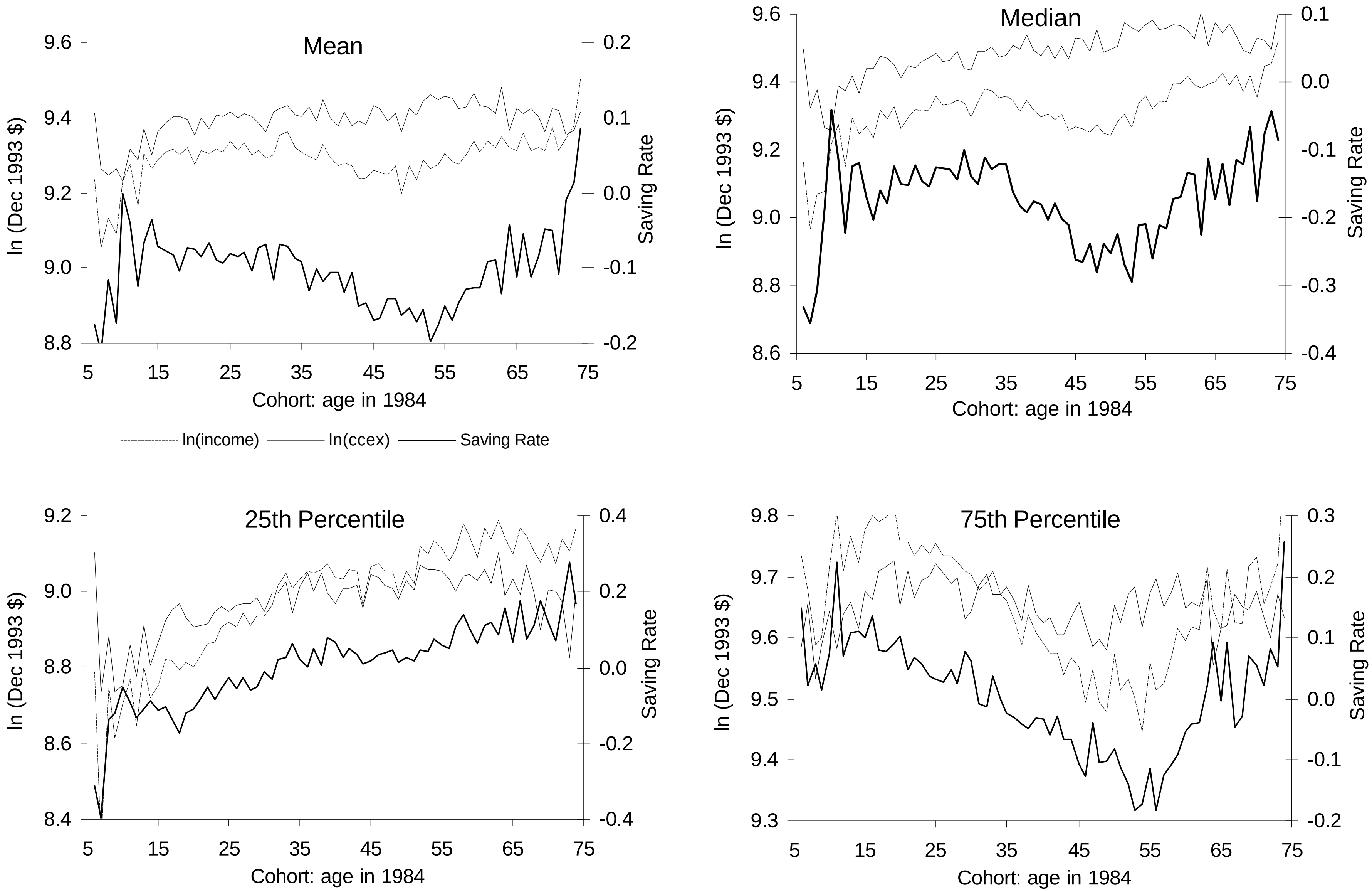


Figure 4: Age and Cohort Effects in Saving Rates [ln (y) - ln (c)] for Individuals and Households

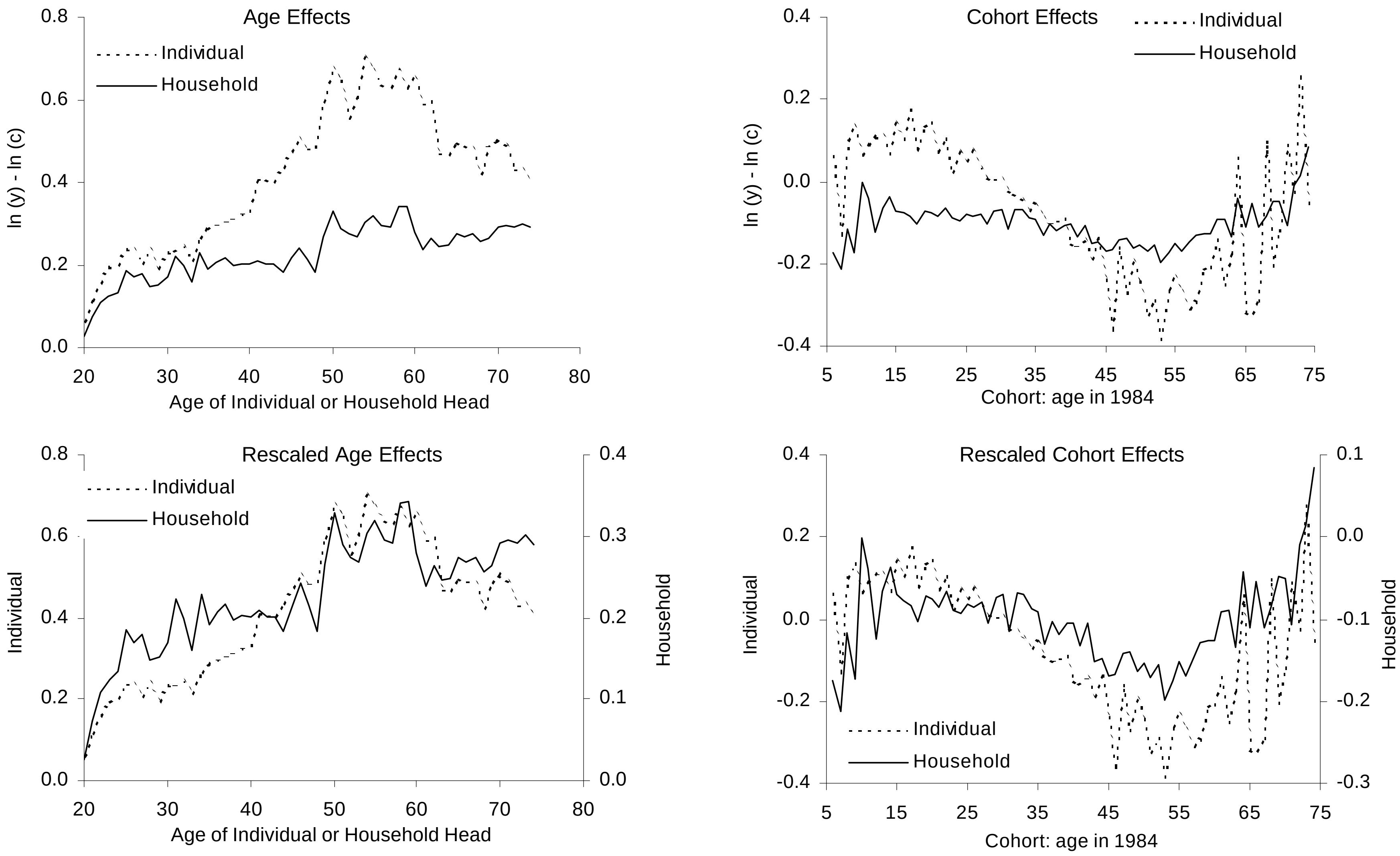


Figure 3: Age, Cohort and Time Decomposition of Mean Savings Rate (Age×Year Cells)

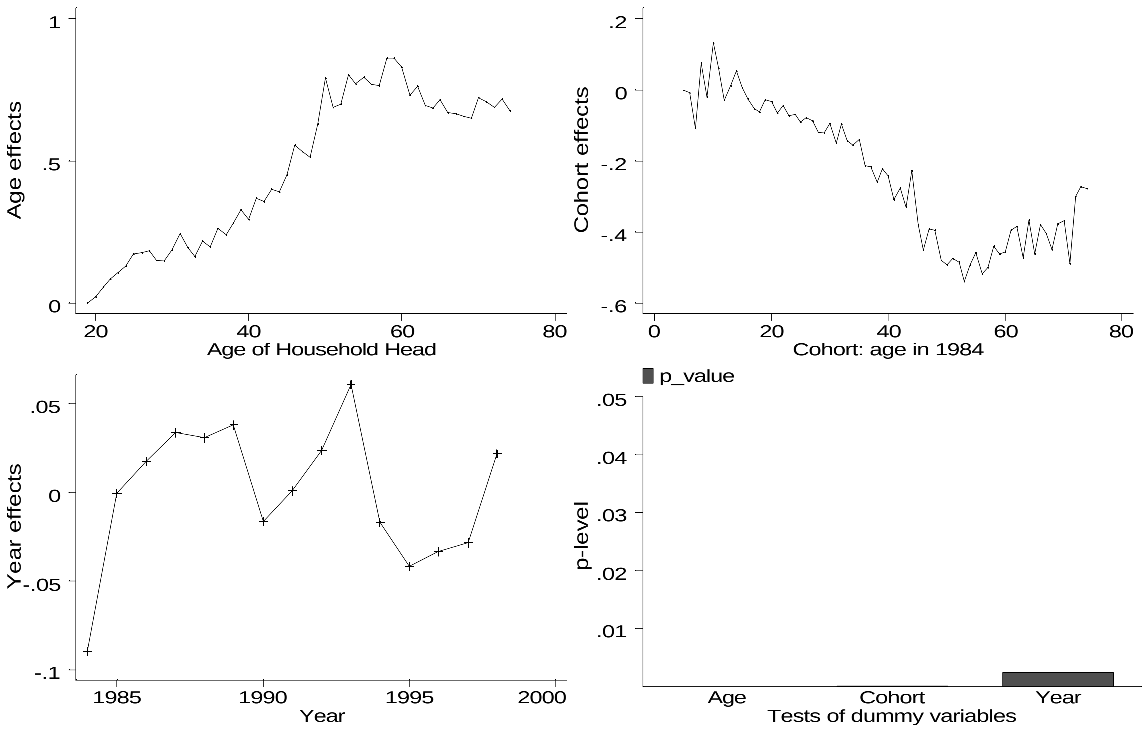
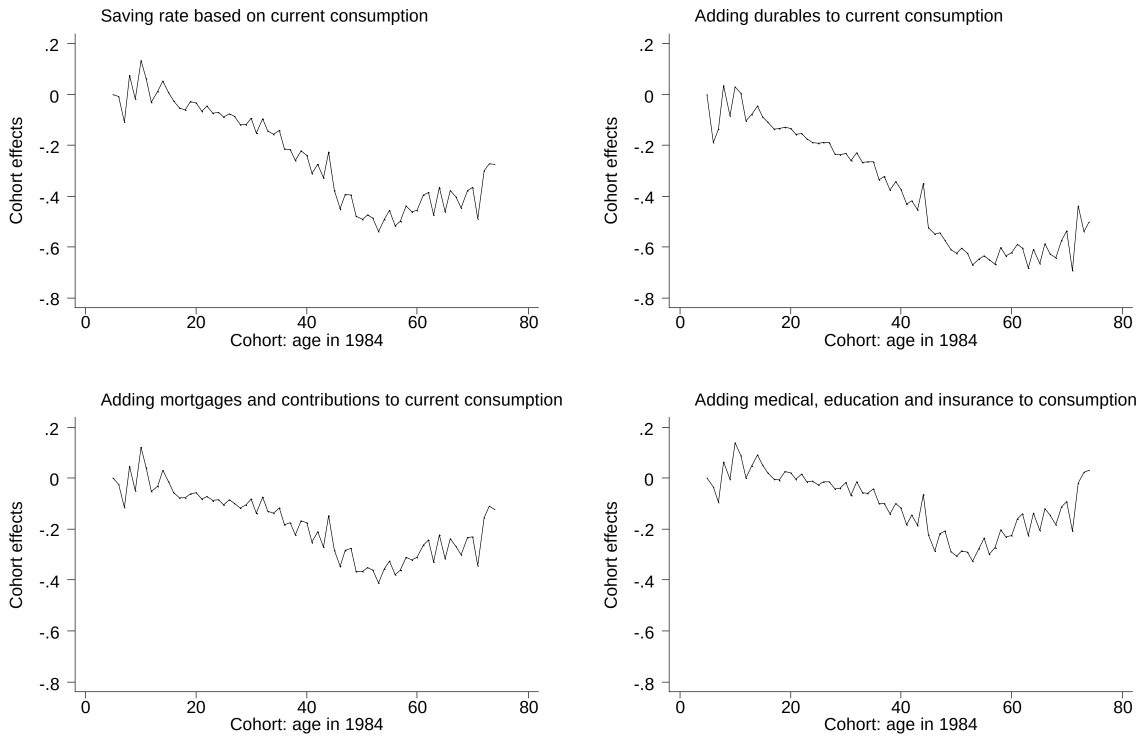


Figure 5: Cohort Effects with Different Definitions of Consumption and Saving



Endnotes

---

[1] Details are provided in Gibson and Scobie (2001), p.7-8.

[2] Household Economic Survey Background Notes 1996/97, p. 17.

[3] Denote saving by S, consumption by X and disposable income by Y. Then: S/Y =(S/X).(X/Y) so the saving ratios reported in this section can be converted by multiplying by the average propensity to consume.

[4] If education, health, insurance, durables and mortgage repayments are put back into the expenditure estimates, the estimated mean of S/X falls to 0.07 (Gibson and Scobie, 2001, Table 7.1).

[5] This is also established from a regression of the 25th percentile saving rate in each year on a time trend. However, this trend is not statistically significant (t=1.14), unlike the trend in the mean.

[6] A referee asked whether there was any allowance for the inflation portion of interest income, which reflects capital maintenance rather than interest income. The only adjustment for inflation was the initial deflation of income and consumption using the CPI. However, the estimate saving rates using the deflated values were very highly correlated (r=0.97) with those derived from nominal income and nominal expenditures, so we assume that the adjustment to interest income would not have a large impact on our results.

[7] The “snapshot” offered by a single cross-section is also unsuitable for observing life-cycle patterns because although a variety of ages are observed in a cross-section, they also represent different birth cohorts. If there are strong cohort effects, a cross-section age profile may be very different from the age profile of any individual, as noted by Shorrocks (1975).

[8] There is a related problem of the negative correlation between wealth and mortality (Attanasio and Hoynes, 2000) which causes samples of the elderly to become progressively richer as the poorer members die younger. This is a further reason for imposing an upper age limit on the sample.

[9] Browning and Crossley (2001) suggest the term ‘lifecycle framework’ because there are several different models consistent with this framework. They also discuss some of the implications of these models for saving behaviour.

[10] The transition from child to adult is assumed to occur at age 15, and from adult to elderly at age 65.

[11] All cell averages and quantiles formed for those households with a head born in b and observed in t are weighted by the household sampling weights recorded in the HES datafiles. A referee has pointed out that the HES sample changed between 1984 and 1998, with two-income households becoming underrepresented and the unemployed more likely to be measured. We did not have available the *ex post* weights that Statistics New Zealand have created to standardise this sample. However, to test the robustness of our results to weighting assumptions, we recalculated the cell averages with the sampling weights ignored. The correlation between the unweighted and the weighted mean saving rates is 0.963, while the correlation between the cohort effects in the saving rates is 0.997. Because of these high correlations, we assume that the patterns reported here may not be too sensitive to changes in the HES sample.

[12] Paxson (1996) carries out a similar analysis where the smallest cell size is only five households, although her median cell size is approximately 200.

[13] Each point on the income and consumption graphs corresponds to a coefficient from a regression. We have not applied any smoothing to these points and we do not show the standard error that surrounds each point. However, the age and cohort effects are both jointly statistically significant (see Appendix Table A).

[14] This claim also relies upon the CPI being the correct measure for putting the consumption standard of living observed in different survey years onto a consistent basis. To the extent that there is unmeasured quality change, living standards of younger cohorts may in fact exceed those of older cohorts, but this is not a topic

that we consider further. Moreover, there is a significant downward trend in household size across the survey years, so the trend in living standards for individuals may not be the same as the trend in living standards for households.

[15] This is the gap from the income graph in the first panel of Figure 2 between ca. 9.3 on the log scale for a 38-year old in 1984 and 9.2 on the log scale for a 49-year old in 1984.

[16] In addition to introducing these constrained year effects, we examine the ratio of household saving to consumption directly rather than studying the age, cohort and time patterns in income and consumption as was done for the earlier results. However, our data are still the sample of 840 cell averages.

[17] The association between the time effects in Figure 3 and various macroeconomic variables is also discussed by Gibson and Scobie (2001), with the closest links appearing with the real growth rate and the public sector saving rate.

[18] This approach to imposing inequality restrictions is noted by White (1993, p.107).

[19] The household estimates are obtained directly from Figures 1 and 2.