Households’ Propensity to Meet the Capital Accumulation Ratio Over Time:
Evidence from the 1992-2007 Surveys of Consumer Finance

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Jodi C. Letkiewicz, B.S.
Graduate Program in Human Ecology

******

The Ohio State University
2010

Thesis Committee:
Dr. Sherman Hanna, Advisor
Dr. Jonathan Fox
Dr. Catherine P. Montalto
ABSTRACT

With the burden of retirement planning shifting to the individual, individuals are responsible now more than ever to understand the complexities of retirement planning. Financial ratios can help simplify financial analysis and provide basic rule-of-thumb guidelines that can be applied to most households. The capital accumulation ratio (CAR), defined as the proportion of net worth held in investment assets, is intended to identify the share of assets held primarily for future consumption.

This thesis explores the time trends of the capital accumulation ratio and considers whether changes in stock indexes relative to housing indexes might have an impact on the percentage of households that meet the 25% CAR threshold. The components that make up the ratio, investment assets and net worth, are discussed. In addition, a logistic regression is used to ascertain which factors are related to whether households will meet the threshold.

The percentage of households meeting the 25% CAR threshold varies significantly between most of the survey years. In periods when the stock market increased more than housing prices (1992-1995, 1995-1998 and again from 2004-2007), the percentage of households meeting the 25% CAR threshold increases from the previous year. In periods when housing prices increased more than the stock market
(1998-2001 and 2001-2004), the percentage of households meeting the threshold decreases from previous periods. Based on the multivariate regression, the difference between the years is significant for every period except 1995-1998.

Education and income are positively related to meeting the guideline. Black, Hispanic and Asian/other households are less likely to meet the guideline than similar white households; unmarried couples and single households (male and female) are less likely to meet the guideline than married households; and households with a child under 19 at home are less likely to meet the guideline than households without a child under 19 at home. The likelihood of meeting the guideline increases with age until age 66.35 and then decreases.
ACKNOWLEDGEMENTS

I wish to thank my advisor, Dr. Sherman Hanna, for his support in helping me complete this. Without his unending patience, support, input, direction, thorough reviews, and responsiveness, this thesis would never have been completed.

I would also like to thank the other two members of my committee, Dr. Jonathan Fox and Dr. Catherine Montalto. Their thorough reviews, comments and suggestions helped strengthen my thesis and gave me confidence in my ability to complete it.
VITA

January 2, 1977……………………………………………… Born – Albany, NY, USA

May 1999…………………………………………………. B.S. Accounting, Miami University, Oxford, OH

September 2008 – August 2009………………………… Graduate Research Associate, The Ohio State University

September 2010 – August 2010………………………… Graduate Teaching Associate, The Ohio State University

August 2010……………………………………………… M.S. Family Resource Management, The Ohio State University

Honors and Awards…………………………………….. Department of Consumer Sciences
Graduate Teaching Award (May 2010)

Recipient of the Deacon-Firebaugh Fellowship for the 2010-2011 academic year

FIELDS OF STUDY

Major Field: Human Ecology

Minor Field: Statistical Data Analysis
# TABLE OF CONTENTS

Abstract.................................................................................................................. ii
Acknowledgements ................................................................................................ iv
Vita ........................................................................................................................... v
List of Tables ......................................................................................................... vii
List of Figures ....................................................................................................... ix
Introduction ........................................................................................................... 1
  Research Questions .............................................................................................. 3
Literature Review ................................................................................................... 5
  Retirement Adequacy .......................................................................................... 5
  Capital Accumulation Ratio ................................................................................. 7
  Economic Conditions .......................................................................................... 14
  Conclusion ............................................................................................................ 21
Theoretical Framework and Hypotheses ............................................................... 23
  Life Cycle Savings Theory .................................................................................. 23
  Herd Behavior .................................................................................................... 24
  Research Hypotheses ......................................................................................... 26
Methods .................................................................................................................. 29
  Data and Sample ............................................................................................... 29
  Weights ............................................................................................................... 30
  Multiple Imputation and RII Techniques .......................................................... 30
  Dependent variables .......................................................................................... 32
  Independent variables ......................................................................................... 33
Analysis .................................................................................................................. 35
  RII Means Test .................................................................................................. 35
  Logistic Regression ............................................................................................ 35
Results ..................................................................................................................... 37
Descriptive Characteristics ................................................................. 37
RII Procedures and Multivariate Results .............................................. 41
Discussion .......................................................................................... 46
Conclusion and Implications ............................................................... 50
List of References .............................................................................. 52
Appendix A: Thesis Code .................................................................... 57
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>20</td>
</tr>
<tr>
<td>4.1</td>
<td>34</td>
</tr>
<tr>
<td>5.1</td>
<td>38</td>
</tr>
<tr>
<td>5.2</td>
<td>39</td>
</tr>
<tr>
<td>5.3</td>
<td>40</td>
</tr>
<tr>
<td>5.4</td>
<td>41</td>
</tr>
<tr>
<td>5.5</td>
<td>43</td>
</tr>
<tr>
<td>5.6</td>
<td>45</td>
</tr>
<tr>
<td>6.1</td>
<td>49</td>
</tr>
</tbody>
</table>

- Stock Price Changes vs. Housing Price Changes
- Independent Variables
- Distribution of the Capital Accumulation Ratio
- Descriptive Characteristics of All Households
- Mean and Median Analysis of CAR Components
- Percentage of Households that Meet the 25% CAR Threshold
- Logistic Regression and RII Means Test for 3-Year Periods
- Logistic Regression Analysis of the Households that Meet the CAR 25% Threshold
- Research Hypotheses and Summary of Empirical Results
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Stock Price Changes vs. Housing Price Changes</td>
<td>21</td>
</tr>
<tr>
<td>4.1</td>
<td>Percentage of Households that Meet the 25% CAR Threshold</td>
<td>43</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Retirement planning is a scary and complex endeavor for many people. Retirement planning involves a lot of unknown variables and scenarios: How long will I live? When will I want to retire? How much money will I be making when I retire? How will my health be in retirement? All of these questions lead up to the main question: How much do I need to save today for retirement? Because of all the complexities and unknown variables involved, most people aren’t interested in learning the details or don’t have the patience to do so. Life is hectic, many people are overworked and retirement planning tends to take a backseat.

Research indicates that Americans are not saving enough for their retirement. About half of workers saving for retirement report that the total value of their investments, excluding their home and their defined benefit plan, was less than $25,000 (Helman, VanDerhei, & Copeland 2007). According to the 2007 Retirement Confidence Survey, only 66% of workers report that they have saved for retirement and only 60% report that they are currently saving.
The shift from defined benefit to defined contribution plans in recent years has transferred much of the burden for retirement planning to the individual. As such, individuals are now responsible for determining whether to save for retirement, how much to save and how much risk to take. In 2008, there were approximately 3.2 workers per every beneficiary in the U.S. Social Security system. While this ratio has been stable since about 1974, the baby-boom generation will have mostly retired by 2030 and the ratio of workers to retirees is projected to drop to 2.2 (OASDI Board of Trustees, 2009).

Retirement planning is growing in importance and complexity. Many rule of thumb guidelines have emerged in an effort to make financial decisions easier. One common retirement rule of thumb suggests setting aside 10% of gross income for savings. Many experts agree that this isn’t so much a rule as it is a starting point. A major financial news outlet suggests that people save “10% for basics, 15% for comfort, [and] 20% to escape.” Authors suggest that an individual’s highest priority should be saving for retirement because every dollar forgone could cost $10 or more in retirement income (Weston, 2009). Rules of thumb exist because people don’t have the knowledge, time or interest necessary to sort through all of the details in order to make a sound financial plan. Many people find rule-of-thumb guidelines useful because they simplify the financial planning process.

Financial ratios can help simplify financial analysis and provide basic rule-of-thumb guidelines that can be applied to most households. Financial ratios provide a convenient way to analyze the financial condition of households (Greninger, Hampton, Kitt, & Achascoso, 1996) and are a quantitative tool useful in financial decision-making.
(Yao, Hanna, & Montalto, 2002). The capital accumulation ratio (CAR), defined as the proportion of net worth held in investment assets, is intended to identify the share of assets held primarily for future consumption. The CAR has been used to assess household well-being (DeVaney, 1993), retirement adequacy (DeVaney, 1995; Yao, Hanna, & Montalto, 2003), financial strength over time (Garman & Forgue, 2008) and change in wealth over the lifecycle (Harness, Finke, & Chatterjee, 2009).

While there has been a fair amount of research on the capital accumulation ratio and retirement adequacy, there has not been significant work to assess if and how the number of households meeting the CAR threshold changes over time. There have been several economic events over recent decades that could impact personal savings rates and thus the CAR. With all of these changes in the market, it is possible that the CAR would fluctuate. A cross sectional review of the CAR is useful in understanding how economic events or trends impact household investments and investment decisions. Cross sectional data allows for a comprehensive view of how economic trends may cause CAR fluctuations for the general population rather than for a specific household.

Research Questions

This thesis will investigate the following research questions:

(1) Does the percentage of households meeting the 25% CAR threshold change in reaction to recent changes of stock indexes relative to housing indexes, or does it remain stable?

(2) Which household characteristics are related to whether households will meet the threshold?
This thesis will look at the capital accumulation ratio over a 15 year range using the Survey of Consumer Finances national dataset for 1992, 1995, 1998, 2001, 2004 and 2007. It will contribute to the existing literature by exploring cross sectional trends associated with the capital accumulation ratio. By using a nationally representative data set, the information provided will be representative of all U.S. households.

The remaining chapters in this thesis are organized as follows. Chapter 2 reviews previous research on retirement adequacy, the capital accumulation ratio, and economic trends. Chapter 3 presents the theoretical framework and hypotheses. The data set and method of analysis used in this study are covered in Chapter 4. The findings from the study are summarized and discussed in Chapter 5. Finally, Chapter 6 provides a discussion of the conclusions and implications.
CHAPTER 2

LITERATURE REVIEW

This chapter reviews and summarizes existing research on retirement adequacy, the capital accumulation ratio and economic conditions. Research hypotheses are presented and discussed at the end of the chapter.

Retirement Adequacy

Retirement adequacy is defined as having sufficient resources to meet retirement needs. A household is considered to have adequate retirement coverage if its assets are equal to or greater than its needs at retirement. An analysis of retirement adequacy involves a comprehensive review of all components of retirement income. Once there is a comprehensive measure of retirement resources, retirement adequacy is determined by comparing needs at retirement to retirement resources. Retirement wealth can be considered adequate if total retirement income is equal to or greater than the planned retirement consumption level (Yuh, Montalto, & Hanna, 1998). If households do not have adequate resources, they are considered to be inadequately prepared (Yao et al.,
Most of the recent research utilize data either from the Health and Retirement Study (Moore & Mitchell, 2000) or the Survey of Consumer Finances (Yuh, Hanna, & Montalto, 1998; Yuh, Montalto, et al., 1998; Yao et al., 2003).

Yuh, Montalto, et al. (1998) used the 1995 Survey of Consumer Finances to estimate the adequacy of retirement wealth. Retirement wealth was estimated using data from the SCF that includes financial assets, nonfinancial assets (including housing wealth), and retirement income from defined contribution plans, pension plans and Social Security. Retirement wealth was then projected using a combination of planned retirement age and portfolio allocation. The household’s level of consumption prior to retirement was used as a proxy for the household’s desired level of consumption during retirement. The amount of retirement income needed was calculated using expenditure functions which were estimated using data from the 1993-1994 Consumer Expenditure Survey. A bivariate analysis was conducted which indicated that about 52% of the households in the sample had adequate wealth for retirement given their planned retirement age. Further analysis indicated that households that spent less than income were much more likely to be prepared for retirement. Planned retirement age also had a significant effect on adequacy.

Yuh, Hanna, et al. (1998) analyzed households’ ability to be adequately prepared for retirement based on mean and pessimistic projections of investments until planned retirement ages. This study used the 1995 Survey of Consumer Finances. In order to project levels of retirement wealth, the future value of current assets was projected based on planned retirement age and portfolio performance. Using a lognormal forecasting
model, rates of return at the 50th percentile and the 5th percentile were selected for portfolio performance. Based on the analysis, the researchers found that approximately 52% of the households in the sample had adequate wealth for retirement under the mean case projection. Using the pessimistic projections, only 42% of the households in the sample had adequate wealth for retirement.

Moore & Mitchell (2000) investigated retirement adequacy by comparing the projected savings rates with the optimal savings rates under the life cycle model. When they assumed that all workers would retire at 65, 40% of households would have adequate retirement wealth, but when they assumed that all workers would retire at 62, only 31% of households would have adequate retirement wealth. Their analysis showed that females were less likely to be prepared for retirement because they have a longer life expectancy. Both having a pension and owning a home were positively correlated with retirement adequacy.

*Capital Accumulation Ratio*

The capital accumulation ratio (CAR) is an indicator of how well an individual or household is advancing toward financial goals for capital accumulation. The CAR is defined as investment assets-to-net worth and compares the value of investment assets with net worth. Several studies have been conducted assessing the capital accumulation ratio. Most studies of the capital accumulation ratio have sought to identify specific characteristics of households meeting certain, specified guidelines. Other studies have focused on finding the ratio of investment assets to net worth that most accurately
predicts retirement adequacy. Another subset of studies analyzes how the CAR predicts different aspects of household financial well-being. However, there has been only a limited amount of research showing patterns of financial ratios and if the proportion of households that meet these ratios change over time.

The capital accumulation ratio is calculated from information on investment assets and net worth. As defined in past studies, (Yao et al., 2002; DeVaney, 1997) investment assets consist of stocks, bonds, mutual funds, retirement accounts (including IRAs, thrift accounts and future pensions), certificates of deposit, cash value of life insurance, other managed assets, other nonresidential assets (such as loans owed to the household, art work, antiques), and other real estate excluding the home and net business assets. Net worth is the sum of liquid assets, investment assets, and nonfinancial assets minus consumer debt and property debt (Kennickell, 2000).

There is some debate as to what should be considered the optimal CAR. Lytton, Garman, & Porter (1991) suggested that the CAR should be at least 25%. DeVaney (1993) also proposed that the CAR should be at least 25%. Greninger et al. (1996) conducted a Delphi study which resulted in the suggested mean value CAR to be just over 50%. Yao et al. (2003) examined retirement adequacy in relation to the CAR and found less of an error rate with households meeting the 25% guideline. They concluded that the 25% guideline was more appropriate than the 50% guideline.

Lytton et al. (1991) identified the ratio of investment assets within a portfolio as a benchmark for a satisfactory household portfolio. They used hypothetical case studies to suggest that the CAR should be at least 25%, although suggested that it should vary
across the life cycle. The researchers suggest that younger households should have a CAR less than 20% and that the ratio should increase over time. Garman & Forgue (2008) also asserted that the CAR should increase as households’ progress through the lifecycle. During the initial stages of the life cycle, young people typically have little money to save. The money they do accumulate is often used for large purchases, such as a house or a car rather than investing for retirement.

DeVaney (1993) conducted a longitudinal study that looked at changes in financial ratios of households to determine whether households improved their financial status over a three year period. DeVaney used the 1983 and 1986 panel study of the Survey of Consumer Finances. She found that American households during that time showed an increase in the ratio of investment assets to net worth (CAR), though she did not indicate whether the increase was statistically significant. The percent of those meeting the 25% guideline increased from 39% to 41% from 1983 to 1986. DeVaney asserted that the growth in personal debt and decline in household savings rates during the 1980s made exploration of analyzing family financial status more important than ever.

Greninger et al. (1996) used a panel of 156 experts for a Delphi study designed to identify ratios and provide refinement of the ratios and benchmarks that could be used to measure financial wellbeing. The panel was comprised of 85 financial planners and 71 educators. The main objectives of the study were to determine the extent to which the panel members could come to a consensus regarding the measures and to develop general
guidelines for financially “healthy” households. The mean of the experts’ recommendations for the capital accumulation ratio was slightly over 50 percent.

Harness et al. (2009) conducted a study to examine whether an initially high CAR impacts change in wealth across a decade for households in the accumulation stage of the life cycle. The data for this study came from the 1979 National Longitudinal Survey of Youth (NLYS) 1979 cohort. The NLYS is a nationally representative panel data set made up of youth who were between the ages of 14 and 21 as of December 31, 1979. This particular cohort of individuals is considered to be the young baby boom generation. The researchers ran an ordinary least squares regression that tested the impact of meeting the 25% CAR guideline on the log change of net worth between 1994 and 2004. They found that meeting the 25% guideline resulted in a 28.1% increase in net worth over the ten years, but did not report how this compared to households that did not meet the guideline. They also found that among those who met the 25% threshold, an increased CAR percentage came with increased wealth dispersion of over 35%. The risk-return that comes with financial investments can result in a higher variance of net worth.

Baek and DeVaney (2004) conducted a study to examine the financial wellness of the baby boomers using both objective and subjective financial wellness. Objective financial wellness was evaluated using three financial ratios: liquid assets-to-income, debt-to-assets and investments assets-to-net worth. Data for the study came from the 2001 Survey of Consumer Finances. A chi-square analysis was conducted to compare the percentage of baby boomers meeting the three financial ratios. Baek and DeVaney (2004) found that college education, higher incomes, a higher tolerance for risk, paying
off credit card balances, saving regularly, and spending less than one’s income were all positively associated with meeting a CAR guideline of 25%. Higher levels of risk aversion were associated with a lower chance of meeting the CAR threshold.

Yao et al. (2002) ascertained household factors related to meeting the CAR guidelines using the 1998 Survey of Consumer Finances. A chi-square bivariate analysis was used to look at meeting the CAR thresholds by various demographic and financial characteristics. When certificates of deposit were not included as investment assets, 56% of households met the 25% CAR guideline while only 40% met the 50% CAR guideline. Results from the bivariate analysis indicate that education, income, and financial risk tolerance all have a positive effect on meeting the CAR guideline. Both variables for whether a household overspent or under-spent were positive and significant at both the 25% and 50% level. The number of years until retirement had a negative effect on meeting the guideline. Two logistic regression models were used to test various rule of thumb guidelines, including meeting the 25% and 50% CAR threshold, where the CAR thresholds were modeled as dependent variables.

DeVaney (1995) assumed that meeting the 25% CAR guideline was related to having adequate retirement preparation. She analyzed the 1989 Survey of Consumer Finances to examine differences between baby boomer cohorts. DeVaney was specifically interested in the differences between the older baby boomers (ages 34-42) and the younger baby boomers (ages 24-33). The cohorts were determined by the age of the household head at the time of the 1989 Survey. A logit estimation method was used to examine the specific factors that were associated with a household’s tendency to meet
the 25% CAR threshold. The analysis indicated that being white and expecting a large
inheritance were positively correlated with meeting the threshold for younger baby
boomers. The older cohort of the baby boomers had increased likelihood of meeting the
guideline if the household head was in good health, was male, and had pension coverage.
For both cohorts, as age and education increased, it was more likely that households
would meet the 25% CAR guideline.

Moon, Yuh, & Hanna (2002) evaluated six financial ratios developed in the
United States for their usefulness in assessing the financial situation of households in
South Korea. The six ratios investigated in the study include debt safety, debt service,
solvency, liquidity, savings, and capital accumulation. The researchers used the 1996
National Survey of Family Income and Expenditures in Korea as the data set. This
survey is conducted every 5 years and is a nationally representative data set sampling
30,000 South Korean households. The financial ratios were used as dependent variables
in the logistic regression models. An ANOVA $F$ test was used to test whether each
financial ratio was significantly different based on age of the household head.
Additionally, logistic regression models were employed to identify the significant factors
that affect whether households met the financial ratios. This study indicated that less
than 15% of Korean households met the 25% CAR guideline. They found that the CAR
increased up to age 39, and then started a decline through the rest of the life cycle,
suggesting likely cultural differences and preferences for investment assets.

Bae, Hanna, & Baek (2005) analyzed financial ratios before and after the Korean
economic crisis (1997-1998) and found that the percentage of households meeting a 20%
CAR guideline decreased by almost 6%. They attributed the decrease in the CAR to the decline of the stock market after the economic crisis.

Yao et al. (2003) used the Yuh, Montalto, et al. (1998) approach to estimating retirement adequacy, and analyzed the relationship between retirement adequacy and meeting CAR guidelines. The model included dummy variables for meeting both the 25% ratio guideline and the 50% ratio guideline. They found that 63% of households had a consistent relationship between retirement adequacy and meeting the 25% ratio guideline -- 46% of all households met the 25% ratio guideline and were prepared for retirement, and 17% of households did not meet the 25% ratio guideline and were not prepared for retirement. While the researchers concluded that this was an unacceptable error rate, they found that the 50% ratio guideline had an even higher error rate and supported DeVaney’s (1997) guideline of 25% rather than the Greninger et al. (1996) guideline of 50%.

For the purposes of this study, we will assess factors related to the likelihood of meeting the CAR guideline of 25%. We are interested in whether the proportion of households meeting the ratio changes over time. The 25% threshold is selected over the 50% threshold because a high CAR threshold of 50% would likely exclude many of the younger households that may be in the initial stages of the lifecycle. The 25% guideline may provide more insight than the 50% guideline into changes over time and effects of household characteristics. Furthermore, Yao et al. (2003) found that there was a stronger relationship between retirement adequacy and the 25% CAR guideline than between retirement adequacy and the 50% guideline.
Economic Conditions

Economic conditions and economic events can have an impact on a household’s ability to meet the CAR threshold of 25%. Given the construction of the CAR, there are factors that could logically affect the ratio. Since the numerator of the CAR is comprised of investment assets, will CAR increase in response to changes in the stock market? The denominator of the CAR is comprised of net worth, which includes both investment assets and non-financial assets, such as housing values. When net worth increases through home values, will the CAR decrease? Will the percentage of households that meet the 25% CAR threshold stay fairly consistent over time? If households do not respond to market changes, the CAR may be affected by these market forces and the percentage of households meeting the threshold may change over time.

Net Worth

Net worth is defined as households’ total assets minus liabilities. Net worth has a skewed distribution due extreme high net worth of a small portion of the population and smaller amount of net worth held by a larger portion of the population. In line with this, a rise in median net worth that is larger than a rise in the mean suggests relatively less growth for wealthy families than for families in the middle of the wealth distribution. The opposite is also the case. When change in mean net worth is greater than the change in median net worth there is more growth for wealthy families.

After falling between 1989 and 1992, median and mean net worth have steadily been on the rise. From 1992 to 1995, mean net worth rose 2.7% while median net worth
rose 6.8% (Kennickell, Starr-McCluer, & Sundén, 1997). Continuing the acceleration trend from 1992, both mean and median net worth increased robustly between 1995 and 1998. Mean net worth rose 25.7% and the median rose 17.6%. The levels of both of these measures exceeded the levels observed in 1989, which was toward the end of the last expansion, representing another market expansion. Compared with 1989, both mean and median net worth were approximately 20% higher in 1998 (Kennickell, Starr-McCluer, & Surette, 2000).

From 1998 to 2001, median net worth rose 10.4% and mean net worth increased 28.7%. Median net worth, at this point, had increased 40.5% over the period from 1992 to 2001 (Aizcorbe, Kennickell, & Moore, 2003). The pace of net worth growth slowed down, though still increased, from 2001 to 2004. The median rose 1.5% and the mean rose 6.3% (Bucks, Kennickell, & Moore, 2006).

From 2004 to 2007, net worth growth was once again accelerated. Median net worth rose 17.7%, and the mean rose 13%. Both median and mean net worth have risen consistently since 1998, but overall the mean has gained more (54.7%), compared with the median (31.8%). Unrealized capital gains were an important factor in the increase in net worth over the 2004–2007 period; the share of total assets attributable to unrealized capital gains from real estate, businesses, stocks, or mutual funds rose 5.1%, to 35.8% in 2007 (Bucks, Kennickell, Mach, & Moore, 2009).

The rising values of primary residences over the 2004-2007 period outpaced the increases in home-secured debt and therefore, the average amount of home equity held by
households increased. Median home equity among homeowners rose from $76,900 to $91,000 over the period, an 18.3% increase (Bucks et al., 2009).

Towards the end of 2007, the housing crisis had begun and both housing prices and equity prices started to decline. The SCF estimates that since the 2007 survey was released there have been large drops in median and mean net worth since the 2007 survey, an estimated -17.8% and -22.7% percent, respectively (Bucks et al., 2009). While net worth has fluctuated over time, the impact this may have on the CAR depends on the more granular components that make up net worth, such as housing and investment assets.

**Investment and Financial Assets**

For the purposes of this thesis, financial assets are classified as either investment assets or liquid assets. Financial assets are considered to be investment assets if they are intended for some longer-term future use. Investment assets are comprised of assets such as stocks, bonds, pension plan assets, certificates of deposit, etc. Liquid assets are not considered investment assets because they are generally held in non-interest bearing accounts and are used in the short-term. Liquid assets are often referred to as transaction accounts because of the nature of the accounts (check writing, ATM withdrawals, etc).

Composition of financial assets has changed over time. There was a considerable shift from 1989 to 1995 where the share of traditional financial assets, such as transaction accounts and certificates of deposit, declined sharply: from 30% in 1989 to only 19% in 1995. At the same time, the share of tax-deferred retirement accounts and equities started to rise almost proportionately, from 38% in 1989 to 56% in 1995 (Kennickell et al.,
Continuing with this trend, the share of financial assets in transaction accounts decreased another 15.7% from 1995 to 1998. From 1989 to 1998, growth was concentrated among stocks, mutual funds, tax-deferred retirement accounts, and other managed assets. Combined, these assets accounted for 71.3% of financial assets in 1998, up from 48.4% in 1989 (Kennickell et al., 2000). After the acceleration of financial assets, the share of financial assets in households’ total assets only increased by 1.3 percentage points between 1998 and 2001 (Aizcorbe et al., 2003). The peak of financial assets held as a percentage of assets occurred in 2001. Financial assets (as a percentage of total assets) fell 6.3 percentage points from 2001 to 2004 (Bucks et al., 2009).

**Stock Market Index**

Standard & Poor’s 500 stock-market index, the S&P 500, is useful for evaluating the past performance of stock investments because the index is well regarded as a proxy for the large-cap stock market. Companies that are chosen for the index are the foremost companies in leading industries within the U.S economy. The Ibbotson SBBI Data Series for Large Company Stocks is constructed from the S&P 500 Composite, but includes dividends reinvested. The components that make up the index include total return, income return and capital appreciation return (Ibbotson Associates, 2009, pp. 5). The Ibbotson SBBI Data Series for Large Company Stocks is used in this analysis. The Ibbotson Large Stock Index was initialized in December, 1925 at $1.00 and closed in 2008 at $2,049.45 for a compound annual growth rate of 9.62% (Ibbotson Associates, 2009, pp. 19). Throughout the 1990s, the Ibbotson Large Stock Index experienced significant growth. From 1992 to 1995, the index increased 53.44%. During the period
from 1995 to 1998, the index increased 110.85%. This was followed by a brief dip from 1998 to 2001 of -3.06%. The index then rebounded and from 2001 to 2004 and increased 11.15% and again increased during the period of 2004 to 2007 by 28.16% (Ibbotson Associates, 2009).

**Savings Rates**

In 1960, Americans were saving 5.4% of their total personal income. Savings rates reached a high of 14.6% in 1975 and by 1982 it leveled off at 10.9%. Since the early 1980s, savings rates in the United States have been plummeting, dropping from averages of around 9% in the 1980s and reaching a low of -1% in 2006 (Bureau of Economic Analysis [BEA], 2009). The high savings rates represented a time before credit became abundant. Once credit was more easily obtained, savings rates started dropping. Just before the stock market crash in 1987, savings rates had dipped to a low of 3.0%. The dot-com bubble in the late 1990s created a lot of wealth on paper and brought soaring prices to the stock market indexes. The savings rate again dipped to a low of 0.9% in October of 2001. The credit crisis of 2007-2009, which was brought on by housing bubbles and sub-prime mortgages, has brought renewed interest in savings with rates in May of 2009 reaching 6.0% (BEA, 2009).

**Housing Values**

Housing values are a large part of the CAR because they are excluded from the investment asset component (the numerator) of the ratio. Therefore, as housing values rise CAR is expected to decrease. In order to balance out the CAR, this increase in housing values should be offset by an increase in investment assets. If a household tries
to maintain its CAR, it will have to adjust its investment asset levels in response to changes in relative changes in housing and stock prices.

The S&P/Case-Shiller Home Price Indices are the leading measure of United States residential real estate prices. The index includes indices for different metropolitan areas capturing approximately 75% of residential housing stock in the U.S. The methodology measures the movement in the price of single-family homes by collecting data on actual sale prices of single-family homes in their specific regions.

The 10 City Composite Case-Shiller Index indicates that real estate prices dropped slightly each year from 1989 to 1993. After a decline of -4.57% from January of 1989 to January 1990, prices fell less drastically and declined approximately 1-1.5% until 1993. Starting in 1996, real estate prices started to climb, beginning with a 2.05% increase in 1996 and climbing to double digit increases starting in 1999 (11.06%). This pace continued until 2006, when real estate prices began declining. From 2005 to 2006, prices went down -0.56%. From 2006 to 2007, real estate prices dropped -11.44% and from 2007 to 2008 there was almost a 20% drop (Standard and Poor’s [S&P 500], 2010).
Table 2.1:

Stock Price Changes vs. Housing Price Changes

<table>
<thead>
<tr>
<th>Years</th>
<th>S&amp;P 500 Index(^1)</th>
<th>Case-Shiller Index(^2)</th>
<th>Relative Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989 - 1992</td>
<td>36.05%</td>
<td>-6.80%</td>
<td>1.46</td>
</tr>
<tr>
<td>1992 - 1995</td>
<td>53.44%</td>
<td>0.08%</td>
<td>1.53</td>
</tr>
<tr>
<td>1995 - 1998</td>
<td>110.85%</td>
<td>17.19%</td>
<td>1.80</td>
</tr>
<tr>
<td>1998 - 2001</td>
<td>-3.06%</td>
<td>37.59%</td>
<td>0.70</td>
</tr>
<tr>
<td>2001 - 2004</td>
<td>11.15%</td>
<td>54.62%</td>
<td>0.72</td>
</tr>
<tr>
<td>2004 - 2007</td>
<td>28.16%</td>
<td>4.64%</td>
<td>1.22</td>
</tr>
</tbody>
</table>

\(^1\)The figures were calculated based on data obtained from the Ibbotson SBBI Data Series for Large Company Stocks (Ibbotson Associates, 2009, pp. 215).

\(^2\)The figures were calculated based on data obtained from the April 2010 Seasonally Adjusted Tables (S&P 500, 2010).

Table 2.1 shows the stock price changes (S&P 500 Index), the housing price changes (Case-Shiller Index) and the relative changes in each time period. For instance, in the middle of 1992 the S&P index was 1.36 times as high as it was in the middle of 1989, while the Case-Shiller index was only 0.93 the level of 1989. The last column is calculated by dividing these ratios, so, for the first period, 1.36/0.93 = 1.46. The changes in stock prices and housing prices are depicted graphically in Figure 2.1.
Conclusion

There have been many economic changes over the periods covered in this research. Beginning in 1989, there was a shift out of traditional financial assets, such as certificates of deposit, and into investment assets as people started saving independently for retirement. During the late 1990s, specifically from 1995-1999 the stock market increased substantially. This is evidenced by both a growth in median net worth of 17.6% as well as triple digit stock market growth (110.85%). During the same period housing prices increased, but not at the same accelerated rate as stock prices. Housing values began to climb in 1996 and continued a steady climb until 2006. This is represented by an increase in home equity from 2004 to 2007 of 18.3% and increased home values of 37.59% from 1998-2001 and 54.62% from 2001 to 2004.
These changes and patterns may have an impact on a household’s propensity to meet the 25% CAR threshold. If households are passive investors, then they will be more likely to meet the CAR threshold of 25% during periods when stock prices increase more than housing prices, specifically from 1992-1995, 1995-1998 and 2004-2007. In the same token, households will be less likely to meet the CAR threshold of 25% during periods when housing prices increase more than stock prices, specifically from 1998-2001 and 2001-2004. (See Table 2.1 for further reference.)
CHAPTER 3

THEORETICAL FRAMEWORK AND HYPOTHESES

*Life Cycle Savings Theory*

Classic economic theory assumes that people will behave rationally, which in the case of saving for retirement, means that individuals will attempt to smooth their consumption over time. The life cycle savings theory states that people will save when income is high and dissave when income is low in order to smooth consumption over one’s lifetime. The life cycle savings theory includes the assumption that households seek to maximize utility from consumption over their lifetimes (Ando & Modigliani, 1963).

The life cycle savings theory provides for some explanation of larger debt loads among younger households. Younger households are likely to take on substantial debt in order to finance a house, purchase a new car or pay off student loans. These debts will slowly decrease over time while investment assets and home values will increase. The same can be said for the accumulation of assets over time. Assets are accumulated during an individual’s work life in order to finance consumption after retirement. Meeting the
CAR guideline can be considered, to some extent, a choice to defer consumption. Median and mean net worth generally show a “hump” pattern that peaks in the 55–64 age range. This pattern reflects both lifecycle saving behavior and growth in real wages over time (Bucks et al, 2009). In line with this, CAR should be lower among younger households, increase as households’ age and then decrease after retirement.

Hanna, Fan, & Chang (1995) developed various lifecycle scenarios and applied those scenarios to the life cycle savings model to provide implications for financial planning. For each of the scenarios, the amount to save each year depends on the difference between income and consumption. If income is greater than consumption, then the consumer is a saver. If income is less than consumption, then the consumer is a borrower (or dissaves). One clear implication from this analysis is that anticipated future income is very important in establishing consumption patterns. Consumers who are confident that their future income will increase do not need to save as much as consumers who expect to have either constant or declining income. The authors suggest that saving the same percentage of income each year is not necessarily valid. Households who are sure that their incomes will increase considerably in the future may be rational in not starting to save for retirement until 25 years before retirement. Households with uncertain future incomes and retirement ages may be rational in starting to save early.

**Herd Behavior**

Many situations exist in which people are influenced by other people’s decisions and actions. There are numerous examples where this behavior has been observed.
Existing studies have provided some evidence that asset managers follow the decisions of other asset managers (Scharfstein & Stein, 1990). This behavior can also be seen in everyday situations in the form of restaurant decisions (Banerjee, 1992) and voting (Cukierman, 1989). People’s decisions are influenced, either because they believe they lack the information necessary to make the decision or they assume if everyone else is doing it, then it must be right. Avery and Zemsky (1998) suggest that herd behavior, along with a certain level of uncertainty, can lead to stock bubbles.

Banerjee (1992) uses a simple model and example to explain the herd effect. In this study, Banerjee offers subjects the choice of two restaurants and provides imperfect information about the quality of the restaurants. The first individual follows his or her own signal, possibly setting the trend for others’ decisions. People who follow make their decisions based on their own signal or by following the decision of the people before them. If the first decision is bad, then everyone may end up choosing the poorer restaurant. This creates a bad equilibrium. This bad equilibrium stems from a “herd externality,” of imitating others and discarding one’s own information (Banerjee, 1992). This effect is a combination of social pressure, imperfect information and information interpretation.

Herd behavior might be able to explain the technology stock bubble in the late nineties. It is possible that people either react passively to changes in the marketplace (i.e. not balancing their portfolios regularly) or respond in line with other people’s decisions. Therefore, when stock prices increase more people see investing as lucrative and either invest more or begin investing for the first time. In this case, we may see the
value of the capital accumulation ratio increase. In periods when housing values are increasing, more people purchase homes, home values increase and the value of the CAR may decrease.

**Research Hypotheses**

If households attempt to meet the 25% CAR guideline, then despite the changes in stock and housing prices over time, the proportion meeting the guideline is expected to remain unchanged over time. On the other hand, if households react passively to changes in stock and housing prices or follow the trends of housing and stock prices, then the proportion meeting the guideline is expected to decrease in periods when housing prices increased more than stock prices (i.e., 1998-2001 and 2001-2004) and increase in periods when stock prices increased more than housing prices (the other periods shown in Table 2.1). Therefore, the following hypotheses are presented related to the percentage of households meeting the 25% CAR threshold.

H1: The percentage of households meeting the 25% CAR threshold is not expected to fluctuate significantly between survey years.

H1A: The percentage of households meeting the 25% CAR threshold will increase from 1992 to 1995.

H1B: The percentage of households meeting the 25% CAR threshold will increase from 1995 to 1998.

H1C: The percentage of households meeting the 25% CAR threshold will decrease from 1998 to 2001.
H1D: The percentage of households meeting the 25% CAR threshold will decrease from 2001 to 2004.

H1E: The percentage of households meeting the 25% CAR threshold will increase from 2004 to 2007.

The life cycle model predicts that a typical household will steadily accumulate investments until retirement, so age should have an influence on meeting the CAR threshold. There has been a fair amount of research on the variables that are significant in predicting whether a household will meet the 25% CAR threshold. Yao, et al. (2002) note that for some household characteristics there are not obvious hypotheses related to meeting the CAR guidelines. Being in a racial/ethnic group other than White is expected to have a negative effect on meeting the threshold. This is for various reasons, but generally speaking, Whites may have more information or experience investing. Education should have a positive effect on meeting the threshold because as education increases, information about investments should also increase. Presence of children under 19 is expected to have a negative effect on meeting the threshold. This may be related to the lifecycle or may simply be because households with children may need to spend money, which would otherwise be invested, on their children. Single males and females are expected to be less likely to meet the threshold when compared to married or unmarried couples. Single males and females may not be similarly motivated to save or invest as their married or couples counterparts and are therefore expected to be less likely to meet the threshold. DeVaney (1995) found households are more likely to meet the threshold if they are white, expecting an inheritance, headed by a male, in good health.
and covered by a pension. Homeownership is expected to have a negative relationship with meeting the threshold because a large portion of their investments are likely in the home, which is not considered an investment asset for the purposes of CAR and retirement investments.

Based upon previous research, the following hypotheses are suggested.

H2: Age will be positively associated with meeting the 25% CAR threshold until retirement.

H3: Racial/ethnic groups other than White will be less likely to meet the 25% CAR threshold.

H4: Education will be positively associated with meeting the 25% CAR threshold.

H5: Households with children under 19 will be less likely to meet the 25% CAR threshold.

H6: Marriage will be positively associated with meeting the 25% CAR threshold.

H7: Homeownership will be negatively associated with meeting the 25% CAR threshold.
CHAPTER 4

METHODS

In this chapter the data set and analysis used in this study are presented. The selection of variables, including both independent and dependent, are discussed.

Data and Sample

The data analyzed in this study are from the six most recent datasets (1992, 1995, 1998, 2001, 2004 and 2007) of the Survey of Consumer Finances (SCF). The SCF is sponsored by the Federal Reserve Board in conjunction with the U.S. Department of the Treasury. The SCF is a triennial survey of U.S. families and is designed to provide detailed financial information on American households. These data include information on households’ assets and liabilities, income, pensions, labor force participation, use of financial services and standard demographic characteristics. The interviews are conducted between May and December of each survey year (Aizcorbe et al., 2003).

The SCF has two techniques for sampling. A geographically-based random sample is employed to provide coverage of characteristics that are broadly distributed
across the U.S. population. The second sample is drawn from records provided by the Internal Revenue Service and results in households that are likely to be relatively wealthy (Bucks et al., 2009). These households hold a large fraction of sparsely held assets, such as non-corporate businesses and tax-exempt bonds. This group is selected from a list derived from tax returns and is referred to as the ‘list sample.’ (Kennickell, 1998)

Weights

The SCF does not have an equal-probability design, so weights are an important component of interpreting the survey data (Board of Governors of the Federal Reserve System, 1996). The core data set contains the final nonresponse-adjusted sampling weights. These weights are intended to compensate for the unequal probability selection in the original design and also for non-response.

Multiple Imputation and RII Techniques

Non-response rates in the SCF tend to be sizeable given the sensitive nature of the survey. To deal with the problem of missing responses, the SCF has employed multiple imputation techniques (Kennickell, 1998). The goal of multiple imputation is to provide data that are the best possible estimate of the missing data. Rather than providing one estimate for each missing variable, multiple responses are provided to represent the likely range of possible responses. In addition to addressing the concern of missing values, though secondary, multiple imputation also helps address privacy concerns. Fairly unusual observations in the population are relatively common in the SCF, and this aspect
of the survey can have privacy implications for the participating households. Each survey year used in this analysis consists of the five complete implicate and thus the number of observations for each survey year is five times the number of respondents (Board of Governors of the Federal Reserve System, 1996). All five implicate are used in this study.

The imputation techniques used to fill in the missing data result in variability appearing to be less than it would be in the absence of missing data, and thus estimates of variance will be lower than the “true” variance. The use of “repeated-imputation-inference” (RII) techniques can provide estimates of the true variance (Rubin, 1987). The coefficients and estimates of variance derived by RII techniques allow for more valid inference and tests of significance. RII techniques use the five implicate and average the point estimates of each of these to produce the best point estimate of the sample (Montalto & Sung, 1996). The variances from the separate implicate are then averaged and adjustments are made for both between and within implicate variances. The square root of the variance is then used to create the standard error. These are the estimates that are used to calculate the t-statistic.

These six years of datasets are combined to test for time trends. The five implicate are combined for each survey year in all analyses and weighted to represent the actual number of households in the survey each year. The sample sizes were 3,906 in 1992, 4,299 in 1995, 4,305 in 1998, 4,442 in 2001, 4,519 in 2004, and 4,418 in 2007, for a total sample size of 25,889 households.


**Dependent variables**

The Capital Accumulation Ratio (CAR) is defined as the ratio of invested assets to net worth and is intended to reflect the share of assets held primarily for future consumption. The CAR reveals how well an individual or family is progressing toward financial goals other than home ownership because it compares the actual accumulated value of investment assets to total net worth. The value of the ratio for each household will be compared to the criterion of holding invested assets which are at least 25\% of net worth. The dependent variable will be coded as one if the household meets the criterion and coded as a zero, if otherwise. If net worth is zero or negative, then the CAR will be defined as equal to the value of investments, in other words, the denominator will be assumed to be equal to 1. This approach is consistent with similar studies (Yao et al., 2002; DeVaney, 1997). If a household has positive investment assets and zero or negative net worth, the CAR will either be undefined (in the case where net worth is zero) or will be very high. DeVaney (1997) suggests that it is reasonable to define a ratio as equal to the numerator if the denominator is zero or negative. The rationale is that if a household has zero or negative net worth and positive investment assets, assets should be growing and therefore the household should be considered to have met the guideline. This paper focuses on whether households meet the threshold and not on exact values of the ratio. Therefore, the extreme values will not inappropriately influence the results.

Investment assets consist of stocks, bonds, mutual funds, retirement accounts (including IRAs, thrift accounts and future pensions), certificates of deposit, cash value of life insurance, other managed assets, other nonresidential assets (such as loans owed to
the household, art work, antiques), other real estate but not the home and net business assets. Net worth is the sum of liquid assets, investment assets, and nonfinancial assets minus consumer debt and property debt. Nonfinancial assets consist of the value of the residence, vehicles, business assets, land and other property, minus any outstanding debt against these assets.

Total debt includes housing debt (including mortgages and home equity loans), credit card debt, installment loans (including student loan debt and vehicle loans), outstanding line of credit loans, home improvement debt, amounts borrowed from life insurance, pension loans, and other consumer debt. In the SCF, net worth is inflation-adjusted and represents the difference between households’ gross assets and their liabilities.

Independent variables

The independent variables in the model act as predictors of the dichotomous dependent variable (25% CAR). Demographic variables include age, education, number of children, marital status and race/ethnicity of the household respondent. Age, education and income are all coded as continuous variables. Age is the age of the household head. Age-squared is used to capture any nonlinear changes in the age effect. Education is the highest year of education completed by the head of the household. The income variable is the annual household pre-tax income. The income variable is transformed using the logarithmic function because of its highly skewed distribution. A variable for the log of income is set to the log of 0.01 if the level is zero or negative. The
variable for whether there is at least one child under 19 living at home is based on the number of related children under age 19 living in the home. Marital status is recorded as either married, unmarried couple, single male or single female. Race/ethnicity is coded as White, Black, Hispanic or Asian/Other. Home ownership is coded as 1 if the respondent is a homeowner and zero, if otherwise. The independent variables are summarized in Table 4.1.

Table 4.1: Independent Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of Household Head</td>
<td>Continuous</td>
</tr>
<tr>
<td>Years of Education (household head)</td>
<td>Continuous</td>
</tr>
<tr>
<td>Income (Log of household income)</td>
<td>Continuous</td>
</tr>
</tbody>
</table>
| Marriage Status (reference category = married) | 1 = Married  
                                           | 2 = Unmarried couple  
                                           | 3 = Single Male  
                                           | 4 = Single Female |
| Have Child < 19 at home               | 1 = Children < 19 in the home  
                                           | 0 = Other                                      |
| Race/Ethnicity of respondent (reference category=White) | 1= White  
                                           | 2 = Black  
                                           | 3 = Hispanic  
                                           | 4 = Asian/Other |
| Homeowner                             | 1 = Homeowner  
                                           | 0 = Other                                      |
Analysis

To determine whether the percentage of households meeting the 25% threshold changed across the survey years, an RII means test is performed. A logistic regression is used to examine the factors associated with a household’s propensity to meet the 25% CAR threshold. Code for the logistic regression and variable definition can be found in Appendix A.

RII Means Test

An RII means test is performed to compare the 25% CAR threshold across the six survey years of the Survey of Consumer Finances. The mean percentage of households meeting the criteria is compared for each year of the survey. The point estimate of the mean is the average value of five implicates. The point estimate of variance is the sum of “between” implicate variance and “within” implicate variance. Compared to non-RII methods, the RII method has a larger standard error; thus, has smaller t-statistics. The reduction of t-statistics reduces the likelihood of rejecting null hypothesis. The code used for the RII means test is provided in Appendix A.

Logistic Regression

A logistic regression (logit) is used to examine the factors associated with a household’s propensity to meet the 25% CAR threshold. Since this study focuses on whether households meet the 25% CAR threshold, the dependent variable for the logistic
regression is dichotomous. For a dichotomous variable, Y, with values of 1 and 0, the model focuses on how the natural log of the odds that Y=1 varies as a function of the linear predictor. Separate analyses are performed for each survey year, with that survey year serving as the reference category for the series of survey year dummy variables. This method enables a comparison of each survey year to the previous survey year to determine whether any change in the percentage of households meeting the CAR 25% threshold is significant when the other independent variables are controlled.
CHAPTER 5

RESULTS

This chapter presents the results from the statistical analysis of the data. The first section summarizes demographic characteristics of the households included in the survey. Results of the RII means test and multivariate analysis are summarized in the following, respective sections.

**Descriptive Characteristics**

The median CAR for all households is 35%, which is above the 25% threshold (Table 5.1). The extremely high maximum value of the ratio (over one billion percent) results in the mean of the ratio being very high, about 252%, supporting the approach of analyzing whether households meet the threshold rather than analyzing the actual value of the CAR. The CAR is negative for some households. This could be because both investment assets (possibly margins) and also net business equity have some negative reported values. Table 5.1 shows the distribution of the CAR.
Table 5.1:
Distribution of the Capital Accumulation Ratio

<table>
<thead>
<tr>
<th>Quantile</th>
<th>CAR Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% Max</td>
<td>13,100,000</td>
</tr>
<tr>
<td>99%</td>
<td>4,680</td>
</tr>
<tr>
<td>95%</td>
<td>1.51</td>
</tr>
<tr>
<td>90%</td>
<td>0.977</td>
</tr>
<tr>
<td>75% Q3</td>
<td>0.732</td>
</tr>
<tr>
<td>50% Median</td>
<td>0.352</td>
</tr>
<tr>
<td>25% Q1</td>
<td>0.00103</td>
</tr>
<tr>
<td>10%</td>
<td>0</td>
</tr>
<tr>
<td>5%</td>
<td>0</td>
</tr>
<tr>
<td>1%</td>
<td>&lt; 0</td>
</tr>
<tr>
<td>0% Min</td>
<td>&lt; 0</td>
</tr>
<tr>
<td>Mean</td>
<td>2.515</td>
</tr>
</tbody>
</table>


More than half of households (57%) meet the 25% CAR threshold (Table 5.2). Of the marriage status categories, married couples meet the threshold at a higher rate than other categories (69%). Households with a child at home have a slightly lower rate (56%) than do households without a child at home (57%). Homeowners meet the threshold at a higher rate (61%) than renters (48%). The rate at which households meet the threshold increases with age, reaching a maximum for the 51 to 60 age group, and then decreases. As expected, the rate at which households meet the threshold increases with level of education attained. Table 5.2 shows the descriptive characteristics of households.
Table 5.2:
Descriptive Characteristics of All Households

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total Sample</th>
<th>Percentage Meeting CAR 25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>100.0</td>
<td>56.51%</td>
</tr>
<tr>
<td>Marriage Status (reference category = married)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>52.13%</td>
<td>69.23%</td>
</tr>
<tr>
<td>Unmarried (couple)</td>
<td>6.50%</td>
<td>50.53%</td>
</tr>
<tr>
<td>Single Male</td>
<td>14.12%</td>
<td>54.72%</td>
</tr>
<tr>
<td>Single Female</td>
<td>27.26%</td>
<td>44.69%</td>
</tr>
<tr>
<td>Have Child &lt; 19 at home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have a child at home</td>
<td>43.67%</td>
<td>55.58%</td>
</tr>
<tr>
<td>Do not have a child at home</td>
<td>56.33%</td>
<td>57.23%</td>
</tr>
<tr>
<td>Race/Ethnicity of respondent (reference category=White)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>75.71%</td>
<td>61.78%</td>
</tr>
<tr>
<td>Black</td>
<td>12.76%</td>
<td>42.16%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>7.84%</td>
<td>30.37%</td>
</tr>
<tr>
<td>Asian/other</td>
<td>3.69%</td>
<td>53.52%</td>
</tr>
<tr>
<td>Homeowner</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own a home</td>
<td>66.78%</td>
<td>60.75%</td>
</tr>
<tr>
<td>Do not own a home/Rent</td>
<td>33.22%</td>
<td>47.98%</td>
</tr>
<tr>
<td>Age of Household Head</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 30</td>
<td>12.13%</td>
<td>46.21%</td>
</tr>
<tr>
<td>31-40</td>
<td>20.46%</td>
<td>56.57%</td>
</tr>
<tr>
<td>41-50</td>
<td>23.68%</td>
<td>61.50%</td>
</tr>
<tr>
<td>51-60</td>
<td>18.45%</td>
<td>63.04%</td>
</tr>
<tr>
<td>60-70</td>
<td>12.54%</td>
<td>59.03%</td>
</tr>
<tr>
<td>Greater than 71</td>
<td>12.72%</td>
<td>49.02%</td>
</tr>
<tr>
<td>Years of Education (Household Head)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>8.1%</td>
<td>27.80%</td>
</tr>
<tr>
<td>High school graduate</td>
<td>27.72%</td>
<td>49.75%</td>
</tr>
<tr>
<td>Some college</td>
<td>25.89%</td>
<td>61.23%</td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>22.84%</td>
<td>74.31%</td>
</tr>
<tr>
<td>Graduate Degree</td>
<td>15.44%</td>
<td>80.89%</td>
</tr>
</tbody>
</table>

The mean and median of the components that make up the capital accumulation ratio are provided in Table 5.3. The mean for investment assets is $295,795 and the median is $21,706. The mean for equity (value of stocks, bonds and mutual funds) is $89,422. The median is zero, indicating that many households do not hold equities. This is also the case with investment property and business assets. The largest component of non-financial assets is housing and is likely the most important component in the CAR, as well as the most important asset held by many households.

Table 5.3:
Mean and Median of CAR Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Assets</td>
<td>$480,455</td>
<td>$155,264</td>
</tr>
<tr>
<td>Financial Assets</td>
<td>$176,747</td>
<td>$19,000</td>
</tr>
<tr>
<td>Investments</td>
<td>$295,795</td>
<td>$21,706</td>
</tr>
<tr>
<td>Equity</td>
<td>$89,422</td>
<td>-</td>
</tr>
<tr>
<td>Non-Financial</td>
<td>$303,708</td>
<td>$114,586</td>
</tr>
<tr>
<td>House</td>
<td>$145,664</td>
<td>$87,720</td>
</tr>
<tr>
<td>Vehicles</td>
<td>$16,808</td>
<td>$11,085</td>
</tr>
<tr>
<td>Investment Property/Other Residential</td>
<td>$51,097</td>
<td>-</td>
</tr>
<tr>
<td>Business</td>
<td>$84,545</td>
<td>-</td>
</tr>
<tr>
<td>Net Worth</td>
<td>$412,146</td>
<td>$92,421</td>
</tr>
<tr>
<td>CAR</td>
<td>251.50%</td>
<td>35.24%</td>
</tr>
</tbody>
</table>

RII Procedures and Multivariate Results

An RII means test is run to determine if there are any significant changes in households meeting the 25% CAR threshold over the SCF survey years (Table 5.4). The percentage of households meeting the 25% CAR threshold is significantly different across all survey years, except for the years 1995 and 2007 (p=0.073). Percentage of households meeting the 25% CAR threshold is the lowest in 1992 (mean of 52.7%), then gradually increases until the highest point in 1998 (60.24%). The percentage of households meeting the 25% CAR threshold dips in 2004 (54.51%) before climbing slightly to 56.59% in 2007 (Figure 5.1).

Table 5.4:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of households</td>
<td>52.74%</td>
<td>55.73%</td>
<td>60.24%</td>
<td>58.91%</td>
<td>54.51%</td>
<td>56.59%</td>
<td>56.51%</td>
</tr>
</tbody>
</table>

Note. Based on RII means tests, the percentage for each survey year starting in 1995 is significantly different from the previous survey year.

The effects of the survey year dummy variables in Table 5.6 are relative to the reference category, 2004. The coefficient for 2001 indicates that at the mean value of other independent variables, the odds that households in 2001 meet the guideline are 1.283 times as high as the odds in 2004. The odds that households in 2007 meet the threshold are 1.31 times as high as the odds in 2004. Refer to Table 5.6 for results from the logistic regression.

In order to test the other hypotheses related to changes over time, separate logits (not shown) are run with other reference years in order to test whether changes between 1992 and 1995, 1995 and 1998, and 1998 and 2001 were significant. The change between 1992 and 1995 is positive (Table 5.5). Households in 1992 are only 87% as likely to meet the threshold as households in 1995. The change between 1995 and 1998
is not significantly different from zero (p=0.093). The change between 1998 and 2001 is positive. The odds that households in 1998 meet the threshold are 1.16 times as high as the odds in 2001. Table 5.5 shows the results from both the logistic regression (comparing survey years) and the RII means test.

Table 5.5:
Logistic Regression and RII Means Test for 3-Year Periods

<table>
<thead>
<tr>
<th>Years</th>
<th>Logit P-value</th>
<th>RII Means P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992 - 1995</td>
<td>.0110*</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>1995 - 1998</td>
<td>.0933</td>
<td>&lt;.0001*</td>
</tr>
<tr>
<td>1998 - 2001</td>
<td>.0069*</td>
<td>.0045*</td>
</tr>
<tr>
<td>2001 - 2004</td>
<td>&lt;.0001*</td>
<td>0.000*</td>
</tr>
<tr>
<td>2004 - 2007</td>
<td>.0203*</td>
<td>&lt;.0001*</td>
</tr>
</tbody>
</table>


The logit is used to predict whether households will meet the 25% CAR threshold (Table 5.6). There is a high consistency between the predictions of whether a household will meet the 25% threshold as evidenced by the 82.5% concordance rate. The concordance rate indicates that the logit is successful at correctly classifying whether a household will or will not meet the threshold in 82.5% of cases.

Age, education, and income are all positively related to meeting the threshold. Unmarried couples, single males and single females are all negatively related to meeting
the threshold. The odds of an unmarried couple meeting the 25% CAR threshold is 80% that of a married couple. Similarly, a single male is only 79% as likely and single females are 59% as likely as married couples are to meet the threshold.

All three racial/ethnic group variables are negatively related to meeting the threshold. Households where the respondent is Black are 62% as likely to meet the threshold as a household where the respondent is white, and households where the respondent is Hispanic are less than 50% as likely to meet the 25% threshold. Households where the respondent is Asian (or other) are slightly more likely, but still only have a 64% chance of meeting the threshold when compared to households where the respondent is White.

Having a child under 19 at home is negatively associated with meeting the threshold. Compared to households without children, households with at least one child under 19 are 85% as likely to meet the threshold. Homeownership is not significant. The coefficient estimates for age and age squared are 0.0637 and -0.00048, respectively, and since the coefficient for age is positive, based on calculus\(^1\), we can conclude that the likelihood of meeting the guideline increases with age until age 66.35, then decreases.

\(^1\) The derivation of the combined effect of age and age-squared can be calculated as follows: Where \(y\) is the log odds: \(y = ax + bx^2\); slope is \(\frac{dy}{dx} = a + 2bx\);

At extreme, slope = 0, therefore \(a + 2bx = 0\); \(-0.0637/2(-0.00048)b = x\); \(x = 66.35\)
Table 5.6:

Logistic Regression Analysis of the Households that Meet the CAR 25% Threshold

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of Household Head</td>
<td>0.0637*</td>
<td>1.066</td>
</tr>
<tr>
<td>Age Squared</td>
<td>-0.00048*</td>
<td>1.000</td>
</tr>
<tr>
<td>Years of Education (household head)</td>
<td>0.2265*</td>
<td>1.254</td>
</tr>
<tr>
<td>Marriage Status (reference category = married)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmarried (couple)</td>
<td>-0.2224*</td>
<td>0.801</td>
</tr>
<tr>
<td>Single Male</td>
<td>-0.2362*</td>
<td>0.790</td>
</tr>
<tr>
<td>Single Female</td>
<td>-0.5329*</td>
<td>0.587</td>
</tr>
<tr>
<td>Have Child &lt; 19 at home</td>
<td>-0.1589*</td>
<td>0.853</td>
</tr>
<tr>
<td>Income (Log of household income)</td>
<td>0.3829*</td>
<td>1.467</td>
</tr>
<tr>
<td>Race/Ethnicity of respondent (reference category=White)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.4798*</td>
<td>0.619</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.7239*</td>
<td>0.485</td>
</tr>
<tr>
<td>Asian/other</td>
<td>-0.4491*</td>
<td>0.638</td>
</tr>
<tr>
<td>Homeowner</td>
<td>-0.0538</td>
<td>0.948</td>
</tr>
<tr>
<td>Year of survey: reference category = 2004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 1992</td>
<td>0.1652*</td>
<td>1.180</td>
</tr>
<tr>
<td>Year 1995</td>
<td>0.3051*</td>
<td>1.357</td>
</tr>
<tr>
<td>Year 1998</td>
<td>0.3969*</td>
<td>1.487</td>
</tr>
<tr>
<td>Year 2001</td>
<td>0.2493*</td>
<td>1.283</td>
</tr>
<tr>
<td>Year 2007</td>
<td>0.1231*</td>
<td>1.131</td>
</tr>
<tr>
<td>Intercept</td>
<td>-8.1062*</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05

Concordance Rate: 82.5%

The research hypotheses are summarized in Table 6.1. There are six hypotheses concerning the change in percentage of households meeting the CAR over time. Of these six hypotheses, two have results different from what is expected. The first is the null hypothesis (H1) that the percentage of households meeting the 25% CAR threshold is not expected to change. The expected result of the null hypothesis is that there will not be significant changes between pairs of survey years. The second (alternative) hypothesis is that the percentage of households meeting the 25% CAR threshold will increase from 1995 to 1998 (H1B). It is hypothesized that given the more extreme growth in the stock market as compared to housing values during the period from 1995 to 1998, that the percentage of households meeting the 25% CAR threshold will increase. Both of these hypotheses are rejected. The percentage of households meeting the 25% CAR threshold is significantly different between every pair of survey years based on the bivariate RII means test. When controlling for other factors in the multivariate analysis, such as income and homeownership, the percentage of households meeting the 25% CAR
threshold is significantly different between each pair of survey years, except 1995 and 1998. Income is the only variable likely to have an impact between these years. When controlling for all other variables except for income, the difference between 1995 and 1998 becomes significant. While income did not increase considerably between 1995 and 1998, there is some evidence that incomes in 1995 were still recovering from the 1989 recession (Kennickell et al., 1997) and therefore income growth in the 1995-1998 time period may offset the difference. The growth income during these periods may be why the difference between 1995 and 1998, when controlling for income, is not significant. There was also growth in ownership of stocks from 1995-1998 (up 20.7% from the prior period), which could also help explain the relative differences between these years.

In periods when the stock market increased more than housing prices (1992-1995, 1995-1998 and again from 2004-2007), the percentage of households meeting the 25% CAR threshold increased from the previous year. In periods when housing prices increased more than the stock market (1998-2001 and 2001-2004), the percentage of households meeting the threshold decreased from previous periods. Based on the multivariate regression, the difference between the years was significant for every period except 1995-1998. Based on these findings, it appears that market and economic climate play some role in households’ ability to meet the 25% threshold.

Of the six hypotheses concerning which variables are significant predictors, only one has a different result from what is expected. Homeownership is expected to have a negative effect on a household’s propensity to meet the 25% CAR threshold. The
rationale is that home owners may have more wealth in their home and have less in investment assets. This hypothesis is not supported. The outcomes for age, racial/ethnic group, education, children in the home, and marriage are all as expected.

The results of the logistic regression support findings in similar studies. Both marriage status and race/ethnic group are significant in this analysis. The results from the logistic regression indicate that having a child under the age of 19 in the home has a negative effect on meeting the threshold. This differs from Yao et al. (2002), who found that having a child under the age of 18 at home did not have a significant effect on meeting the 25% CAR guideline. Households with a child at home are more likely to spend money raising the children and may not have as much to invest as households without children at home. Alternatively, depending on the child’s age, households may be investing for college or to help their children with large purchases in the future. Therefore, it’s logical that this variable may fluctuate in terms of significance over time. Accordingly, the factors that increase households’ chances of meeting the threshold include; increasing age (until age 66), education, and income. These findings are similar to findings in other studies (Baek & DeVaney, 2004; Yao et al., 2002; DeVaney, 1995).
<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Expected Effect</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: The percentage of households meeting the 25% CAR threshold is not expected to fluctuate significantly from survey year to the next survey year.</td>
<td>Not significant</td>
<td>REJECT Significant between all years, except 1995 and 1998</td>
</tr>
<tr>
<td>H1A: The percentage of households meeting the 25% CAR threshold will increase from 1992 to 1995.</td>
<td>+</td>
<td>FAIL TO REJECT +*</td>
</tr>
<tr>
<td>H1B: The percentage of households meeting the 25% CAR threshold will decrease from 1995 to 1998.</td>
<td>+</td>
<td>REJECT +</td>
</tr>
<tr>
<td>H1C: The percentage of households meeting the 25% CAR threshold will decrease from 1998 to 2001.</td>
<td>-</td>
<td>FAIL TO REJECT −*</td>
</tr>
<tr>
<td>H1D: The percentage of households meeting the 25% CAR threshold will decrease from 2001 to 2004.</td>
<td>+</td>
<td>FAIL TO REJECT +*</td>
</tr>
<tr>
<td>H1E: The percentage of households meeting the 25% CAR threshold will increase from 2004 to 2007.</td>
<td>+</td>
<td>FAIL TO REJECT +*</td>
</tr>
<tr>
<td>H2: Age will be positively related to meeting the 25% CAR threshold until retirement.</td>
<td>+</td>
<td>FAIL TO REJECT +*</td>
</tr>
<tr>
<td>H3: Racial/ethnic groups other than White will be negatively related to meeting the 25% CAR threshold.</td>
<td>-</td>
<td>FAIL TO REJECT −*</td>
</tr>
<tr>
<td>H4: Education will be positively related to meeting the 25% CAR threshold.</td>
<td>+</td>
<td>FAIL TO REJECT +*</td>
</tr>
<tr>
<td>H5: Households with children under 19 will be less likely to meet the 25% CAR threshold</td>
<td>−</td>
<td>FAIL TO REJECT −*</td>
</tr>
<tr>
<td>H6: Marriage will be positively related to meeting the 25% CAR threshold.</td>
<td>+</td>
<td>FAIL TO REJECT +*</td>
</tr>
<tr>
<td>H7: Homeownership will be negatively related to meeting the 25% CAR threshold.</td>
<td>−</td>
<td>REJECT</td>
</tr>
</tbody>
</table>

* Indicates significant effect of variable or groups of variables, based on logit shown in Table 5.6, or similar logits using reference years other than 2004.
Conclusion and Implications

The capital accumulation ratio, which measures the proportion of net worth held in investment assets, is often used to measure financial well-being, retirement adequacy and financial strength. Despite its frequent use, there has been little analysis concerning the consistency of households meeting the ratio over time. This study attempts to look at the CAR over time and assess the consistency of the percentage of households that meet the 25% threshold.

The actual changes in the proportion of households meeting the 25% CAR guideline, and the changes based on the logit analyses, imply that households do not attempt to maintain their CAR levels. Rather the changes in the percentage of households meeting the 25% CAR threshold move along with the relative change in housing and stock prices. In periods when the stock market increases at a greater rate than housing prices, the percentage of households meeting the threshold increases. The opposite is also true. When housing prices increase more than stock prices the percentage of households meeting the threshold decreases. This could be because households react passively to changes in the economy or that households attempt to chase returns and follow trends. Homeownership increased only slightly from 1992 (64%) to 2007 (69%) so there is little evidence of this in homeownership. Stock ownership, however, increased substantially over this period. In 1992, stock ownership was at 37% and by 1998 it was at 48%. Stock ownership peaked in 2001 at 52% and leveled off just above 50%. While this isn’t conclusive evidence, there is some support for either passive
behavior or herd behavior (in the form of chasing returns and increased investments in the market).

The fact that the percentage of households meeting the 25% CAR threshold changes over time is an indication for the need of households and financial advisors to monitor asset portfolios more actively. While the 25% CAR threshold is generally only used as a guideline, the strength of the guideline is the ability for a household to understand where they may be investing too heavily (such as a house) or too little (such as a retirement plan). While rebalancing portfolios is common in terms of stocks, bonds and mutual funds, it’s less often considered in the context of full asset management. Additionally, it is important to understand where a household is in the lifecycle. Younger households will likely have a lower CAR. The CAR is expected to increase as households’ age, until retirement when the ratio will begin to decrease. A meaningful extension of this study would be to study the composition of investment assets and examine how that changes over time. Another possible extension would be to assess what the value of capital accumulation ratios should be across the lifecycle to help households understand how their portfolios should change over time.
LIST OF REFERENCES


APPENDIX A: THESIS CODE
*categorical & dummy variables for respondent's age;

if respage<30 then age=25;
else if respage<40 then age=35;
else if respage<50 then age=45;
else if respage<60 then age=55;
else if respage<70 then age=65;
else age=75;
if age=25 then age_25=1; else age_25=0;
if age=35 then age_35=1; else age_35=0;
if age=45 then age_45=1; else age_45=0;
if age=55 then age_55=1; else age_55=0;
if age=65 then age_65=1; else age_65=0;
if age=75 then age_75=1; else age_75=0;

if year=1992 then y92=1;else y92=0;
if year=1995 then y95=1;else y95=0;
if year=1998 then y98=1;else y98=0;
if year=2001 then y01=1;else y01=0;
if year=2004 then y04=1;else y04=0;
if year=2007 then y07=1;else y07=0;

*head's education with 5 categories;
if X5905 = 3 or X5905 = 4 then h_grad=1;else h_grad=0;
if h_grad=1 then ed5cl=5;
else if X5904=1 and x5905^=1 then ed5cl=4;
else if (x5901 ge 13) then ed5cl=3;
else if (x5902 in (1 2)) then ed5cl=2;
else ed5cl=1;
if ed5cl =1 then hed_LTHS=1; else hed_LTHS=0;  *less than HS diploma;
if ed5cl =2 then hedHS=1; else hedHS=0;  *high school diploma;
if ed5cl =3 then hedSC=1; else hedSC=0;  *some college or AA degree;
if ed5cl =4 then hedBSng=1; else hedBSng=0;  *bach but not grad degree;
if ed5cl =5 then hedgrad=1; else hedgrad =0;  *grad degree;

* marital status of the HH head: 1=married/living with partner,
2=neither married nor living with partner  FRB approach - note - combines married with unmarried couples as lists as married;
IF X8023=1 then marcouple=1; else marcouple=0;
IF X8023=2 then unmarcouple=1; else unmarcouple=0;
IF X8023 in (3,4,5,6,0) and x8021=1 then singlemale=1; else singlemale=0;
IF X8023 in (3,4,5,6,0) and x8021^=1 then singlefemale=1;
else singlefemale=0;
* number of children (including natural children/step-children/foster children of head/spouse/partner) from SCF macro code;
  * NOTE: from 1995 forward, household listing information collected
  for one fewer HH member;
  ARRAY REL(*) X108 X114 X120 X126 X132 X202 X208 X214 X220 X226;
  KIDS=0;
  DO I=1 TO DIM(REL);
    KIDS=KIDS+(REL{I}=4|REL{I}=13|REL{I}=36);
  END;
  * our OLD code for counting number of related children under 18
  - need to check against old codebooks as x228 and x226 might not
    be available before 2007;
  *Using array statements to create variable for number of related
  children under age 19 in the household;
  array one{10} x110 x116 x122 x128 x134 x204 x210 x216 x222 X228;
  array two{10} x108 x114 x120 x126 x132 x202 x208 x214 x220 x226;
  *array 2 code 4. *Child (in-law) (of R or Spouse/Partner);
  rchild=0;
  do i=1 to 10;
    if one{i}<=18 and two{i}=4 then rchild=rchild+1;
  end;
  *Indicator variable for presence of related children under age
  19;
  if rchild>0 then child=1; else child=0;

* race/ethnicity;
  * NOTE: prior to 1998, the SCF only asked for one response.
  In 1998, respondents were allowed to give multiple responses,
  but they were asked to give first the category they
  identified
  with most strongly. Few people gave more than one response. For
  purposes of continuity with prior data, define the 1998+ variable in terms of the strongest identification;
  * beginning in 2004, respondents were also asked a question to
determine whether they were of Hispanic/Latino culture or
origin;
  IF YEAR GE 2004 THEN DO;
  * 1=white non-Hispanic, 2=nonwhite or Hispanic;
  * For the public data, we only keep the first two race
variables
  and only code x6810 with a 1 if there is any response or a
  5 if
  there is no response;
    RACECL=1+(X6809 ^= 1 | X6810 ^=5);
    H_RACECL=1+(X6809 ^= 1 | X6810 ^=5 | X7004=1);
  * 1=white non-Hispanic, 2=black/African-American, 3=Hispanic,
  4=Asian (only available in internal data set, see
  codebook),
    5=other;
  IF X6809=1 THEN RACE=1;
ELSE IF X6809=2 THEN RACE=2;
ELSE IF X6809=3 THEN RACE=3;
    ELSE RACE=4;

IF X6809=1 & X7004^=1 THEN H_RACE=1;
ELSE IF X6809=2 & X7004^=1 THEN H_RACE=2;
ELSE IF X6809=3 | X7004=1 THEN H_RACE=3;
    ELSE H_RACE=4;
END;
ELSE IF YEAR GE 1998 THEN DO;  * 1998 or 2001;
   * 1=white non-Hispanic, 2=nonwhite or Hispanic;
   * For the public data, we only keep the first two race
variables
    and only code x6810 with a 1 if there is any response or a
5 if
    there is no response;
    IF X6809=1 THEN RACE=1;*white;
ELSE IF X6809=2 THEN RACE=2; *Black;
ELSE IF X6809=3 THEN RACE=3; *Hispanic;
        ELSE RACE=4;  *asian or other;
    END;
ELSE DO;  *before 1998;
    RACECL=1+(X5909 ^= 5);
    IF X5909=5 THEN RACE=1;
    ELSE IF X5909=4 THEN RACE=2;
    ELSE IF X5909=3 THEN RACE=3;
        ELSE RACE=4;
    END;

*Dummy variables for race;
    if race=1 then white=1; else white=0;       *white;
    if race=2 then black=1; else black=0;       *black;
    if race=3 then hispanic=1; else hispanic=0;  *hispanic;
    if race=4 then asianother=1; else asianother=0;
    *asianother;

if x723 in (1,2,5) then homeowner=1; else homeowner=0;

* all types of transactions accounts (liquid assets);
    LIQ=CHECKING+SAVING+MMA+CALL;

* total financial assets;
    FIN=LIQ+CDS+NMMF+STOCKS+BOND+RETQLIQ+SAVBND+CASHLI+OTHMA+OTHFIN;

* total nonfinancial assets;
    NFIN=VEHIC+HOUSES+ORESRE+NNRESRE+BUS+OTHNFIN;

* total nonfinancial assets excluding principal residences;
    NHNFIN=NFIN-HOUSES;

* total assets;
    ASSET=FIN+NFIN;
* total debt;
  DEBT=MRTHEL+RESDBT+OTHLOC+CCBAL+INSTALL+ODEBT;

* total net worth;
  NETWORTH=ASSET-DEBT;

* Investment Assets;
  NHNFIN_INV=NFIN-HOUSES-vehic;
  fin_inv=FIN-LIQ;
  Invest=NHNFIN_INV+fin_inv;

* Investment Asset to Net Worth Ratio;
  If networth>0 then CAR_SH=invest/networth; else CAR_SH=invest;
  If CAR_SH>=0.25 then CAR25_SH=1; else CAR25_SH=0;

  place=CAR25_SH;
  cat=year;

proc univariate;weight nwgt;
var Invest CAR_SH income Networth equity;
title '1992-2007 CAR=invest/NW nwgt';
run;
ods rtf close;

proc logistic descending;weight wt5;
model CAR25_SH = X14 X14*X14 x5901 unmarcouple singlemale singlefemale child loginc black Hispanic asianother homeowner y92 y95 y98 y01 y07;
run;

proc logistic descending;weight wt5;
model CAR25_SH = X14 X14*X14 x5901 unmarcouple singlemale singlefemale child loginc black Hispanic asianother homeowner y92 y95 y98 y04 y07;
run;

proc logistic descending;weight wt5;
model CAR25_SH = X14 X14*X14 x5901 unmarcouple singlemale singlefemale child loginc black Hispanic asianother homeowner y92 y95 y98 y04 y07;
run;
ods rtf close;

proc freq;weight nwgt;
tables CAR25_SH*year;
title 'car25 by year(NetWorth<0, then CAR25=invest)';
run;

proc freq;weight nwgt;
tables CAR25_SH;
title 'car25 (NetWorth<0, then CAR25=invest)';
run;

proc univariate;weight nwgt;
var houses invest networth vehic NNRESRE oresre bus nfin fin equity;
run;

data data.logits;
set data.scf;
wgt=x42001/5;
nwgt1=x42001/24556.47902/5;
proc sort data=data.logits;
by CAR25_SH;
ods rtf file='G:\CODE\descriptives.rtf';
proc freq;weight nwgt;
tables child homeowner black Hispanic asianother white unmarcouple singlemale singlefemale marcouple age_25 age_35 age_45 age_55 age_65 age_75 hed_LTHS hedHS hedSC hedBSng hedgrad;
by CAR25_SH;
title 'Independent Variables';
run;
ods rtf close;

ods rtf file='C:\Users\jodilet\Documents\OSU\2009-2010\Winter 2010\Thesis\CODE\age.rtf';
proc freq;weight nwgt;
tables CAR25_SH*year*age;
run;

proc univariate trimmed = 0.10 ;
var CAR_SH;
run;
ods rtf close;

*RII Means Test
libname data 'C:\Users\Jodilet\Documents\OSU\SCF Dataset';
data data.RII;
set data.scf;
wgt=x42001/5;
nwgt=x42001/26283.8674/5; *mean value of X42001 is 26,283.8674;
if cat=1992 then x1=place;
if cat=1995 then x2=place;
if cat=1998 then x3=place;
if cat=2001 then x4=place;
if cat=2004 then x5=place;
if cat=2007 then x6=place;
keep wgt nwgt place CAR25_SH cat year implic x1 x2 x3 x4 x5 x6;
run;
proc freq;weight nwgt;
tables CAR25_SH*year;
title 'car25 by year ';
run;
PROC SORT DATA=data.RII;
BY Implic;
run;

PROC UNIVARIATE DATA=data.RII VARDEF=WDF;
   VAR X1 X2 X3 X4 X5 X6;
   WEIGHT nwgt;
   BY Implic;
   OUTPUT OUT=RII MEAN=QX1 QX2 QX3 QX4 QX5 QX6
           STD=UX1 UX2 UX3 UX4 UX5 UX6
           N=NX1 NX2 NX3 NX4 NX5 NX6;

PROC PRINT DATA=RII;

*RII TECHNIQUE FOR A SCALAR;

PROC IML;
   RESET AUTONAME;
   USE RII VAR{QX1 QX2 QX3 QX4 QX5 QX6}; READ ALL INTO QI;
   USE RII VAR{UX1 UX2 UX3 UX4 UX5 UX6}; READ ALL INTO UI;
   USE RII VAR{NX1 NX2 NX3 NX4 NX5 NX6}; READ ALL INTO NI;
   MM= NROW(QI); JMAT = J(MM,1);
   NN= NCOL(QI); NAMES={X1 X2 X3 X4 X5 X6};

   *AVERAGE OF THE 6 POINT ESTIMATES OF THE MEAN (Eq. 1);
   QMBAR=QI(+,1)/MM;
   QMBAR12=(QI(+,1)-QI(+,2))/MM;
   QMBAR13=(QI(+,1)-QI(+,3))/MM;
   QMBAR14=(QI(+,1)-QI(+,4))/MM;
   QMBAR15=(QI(+,1)-QI(+,5))/MM;
   QMBAR16=(QI(+,1)-QI(+,6))/MM;
   QMBAR23=(QI(+,2)-QI(+,3))/MM;
   QMBAR24=(QI(+,2)-QI(+,4))/MM;
   QMBAR25=(QI(+,2)-QI(+,5))/MM;
   QMBAR26=(QI(+,2)-QI(+,6))/MM;
   QMBAR34=(QI(+,3)-QI(+,4))/MM;
   QMBAR35=(QI(+,3)-QI(+,5))/MM;
   QMBAR36=(QI(+,3)-QI(+,6))/MM;
   QMBAR45=(QI(+,4)-QI(+,5))/MM;
   QMBAR46=(QI(+,4)-QI(+,6))/MM;
   QMBAR56=(QI(+,5)-QI(+,6))/MM;

   *VARIANCE OF THE MEAN (SQUARE OF THE STANDARD ERROR);
   UI=UI#UI;

   *AVERAGE WITHIN IMPUTATION VARIANCE (Eq. 2);
   UMBAR1=UI(+,1)/MM;
   UMBAR2=UI(+,2)/MM;
   UMBAR3=UI(+,3)/MM;
   UMBAR4=UI(+,4)/MM;
   UMBAR5=UI(+,5)/MM;
   UMBAR6=UI(+,6)/MM;

   *INTERMEDIATE STEPS FOR CALCULATING BETWEEN IMPUTATION VARIANCE;
   QMBARX=QMBAR@JMAT;
   QDIF = QI-QMBARX;
   QDIFSQ = QDIF#QDIF;

   *BETWEEN IMPUTATION VARIANCE (Eq. 3);
   BM1 = QDIFSQ(+,1)/(MM-1);
   BM2 = QDIFSQ(+,2)/(MM-1);
BM3 = QDIFSQ[+3]/(MM-1);
BM4 = QDIFSQ[+4]/(MM-1);
BM5 = QDIFSQ[+5]/(MM-1);
BM6 = QDIFSQ[+6]/(MM-1);

*RII TOTAL VARIANCE OF THE MEAN (Eq. 4);
TM1 = UMBAR1+(1+1/MM)*BM1;
TM2 = UMBAR2+(1+1/MM)*BM2;
TM3 = UMBAR3+(1+1/MM)*BM3;
TM4 = UMBAR4+(1+1/MM)*BM4;
TM5 = UMBAR5+(1+1/MM)*BM5;
TM6 = UMBAR6+(1+1/MM)*BM6;

*Pooled standard deviation;
SP12=SQRT(((NI[+,1]-1)*TM1+(NI[+,2]-1)*TM2)/(NI[+,1]+NI[+,2]-2));
SP13=SQRT(((NI[+,1]-1)*TM1+(NI[+,3]-1)*TM3)/(NI[+,1]+NI[+,3]-2));
SP14=SQRT(((NI[+,1]-1)*TM1+(NI[+,4]-1)*TM4)/(NI[+,1]+NI[+,4]-2));
SP15=SQRT(((NI[+,1]-1)*TM1+(NI[+,5]-1)*TM5)/(NI[+,1]+NI[+,5]-2));
SP16=SQRT(((NI[+,1]-1)*TM1+(NI[+,6]-1)*TM6)/(NI[+,1]+NI[+,6]-2));
SP23=SQRT(((NI[+,2]-1)*TM2+(NI[+,3]-1)*TM3)/(NI[+,2]+NI[+,3]-2));
SP24=SQRT(((NI[+,2]-1)*TM2+(NI[+,4]-1)*TM4)/(NI[+,2]+NI[+,4]-2));
SP25=SQRT(((NI[+,2]-1)*TM2+(NI[+,5]-1)*TM5)/(NI[+,2]+NI[+,5]-2));
SP26=SQRT(((NI[+,2]-1)*TM2+(NI[+,6]-1)*TM6)/(NI[+,2]+NI[+,6]-2));
SP34=SQRT(((NI[+,3]-1)*TM3+(NI[+,4]-1)*TM4)/(NI[+,3]+NI[+,4]-2));
SP35=SQRT(((NI[+,3]-1)*TM3+(NI[+,5]-1)*TM5)/(NI[+,3]+NI[+,5]-2));
SP36=SQRT(((NI[+,3]-1)*TM3+(NI[+,6]-1)*TM6)/(NI[+,3]+NI[+,6]-2));
SP45=SQRT(((NI[+,4]-1)*TM4+(NI[+,5]-1)*TM5)/(NI[+,4]+NI[+,5]-2));
SP46=SQRT(((NI[+,4]-1)*TM4+(NI[+,6]-1)*TM6)/(NI[+,4]+NI[+,6]-2));
SP56=SQRT(((NI[+,4]-1)*TM5+(NI[+,6]-1)*TM6)/(NI[+,5]+NI[+,6]-2));

*Pooled standard error;
SE12=SP12*SQRT(1/NI[+,1]+1/NI[+,2]);
SE13=SP13*SQRT(1/NI[+,1]+1/NI[+,3]);
SE14=SP14*SQRT(1/NI[+,1]+1/NI[+,4]);
SE15=SP15*SQRT(1/NI[+,1]+1/NI[+,5]);
SE16=SP16*SQRT(1/NI[+,1]+1/NI[+,6]);
SE23=SP23*SQRT(1/NI[+,2]+1/NI[+,3]);
SE24=SP24*SQRT(1/NI[+,2]+1/NI[+,4]);
SE25=SP25*SQRT(1/NI[+,2]+1/NI[+,5]);
SE26=SP26*SQRT(1/NI[+,2]+1/NI[+,6]);
SE34=SP34*SQRT(1/NI[+,3]+1/NI[+,4]);
SE35=SP35*SQRT(1/NI[+,3]+1/NI[+,5]);
SE36=SP36*SQRT(1/NI[+,3]+1/NI[+,6]);
SE45=SP45*SQRT(1/NI[+,4]+1/NI[+,5]);
SE46=SP46*SQRT(1/NI[+,4]+1/NI[+,6]);
SE56=SP56*SQRT(1/NI[+,5]+1/NI[+,6]);

*T statistic;
T12=QMBAR12/SE12;
T13=QMBAR13/SE13;
T14=QMBAR14/SE14;
T15=QMBAR15/SE15;
T16=QMBAR16/SE16;
T23=QMBAR23/SE23;
T24=QMBAR24/SE24;
T25 = QMBAR25/SE25;
T26 = QMBAR26/SE26;
T34 = QMBAR34/SE34;
T35 = QMBAR35/SE35;
T36 = QMBAR36/SE36;
T45 = QMBAR45/SE45;
T46 = QMBAR46/SE46;
T56 = QMBAR56/SE56;
DF12 = NI[+,1] + NI[+,2] - 2;
DF13 = NI[+,1] + NI[+,3] - 2;
DF14 = NI[+,1] + NI[+,4] - 2;
DF15 = NI[+,1] + NI[+,5] - 2;
DF16 = NI[+,1] + NI[+,6] - 2;
DF23 = NI[+,2] + NI[+,3] - 2;
DF24 = NI[+,2] + NI[+,4] - 2;
DF25 = NI[+,2] + NI[+,5] - 2;
DF26 = NI[+,2] + NI[+,6] - 2;
DF34 = NI[+,3] + NI[+,4] - 2;
DF35 = NI[+,3] + NI[+,5] - 2;
DF36 = NI[+,3] + NI[+,6] - 2;
DF45 = NI[+,4] + NI[+,5] - 2;
DF46 = NI[+,4] + NI[+,6] - 2;
DF56 = NI[+,5] + NI[+,6] - 2;
* P-value;
P12 = PROBT(T12, DF12) * 2;
if P12 > 1 then P12 = (1 - PROBT(T12, DF12)) * 2;
P13 = PROBT(T13, DF13) * 2;
if P13 > 1 then P13 = (1 - PROBT(T13, DF13)) * 2;
P14 = PROBT(T14, DF14) * 2;
if P14 > 1 then P14 = (1 - PROBT(T14, DF14)) * 2;
P15 = PROBT(T15, DF15) * 2;
if P15 > 1 then P15 = (1 - PROBT(T15, DF15)) * 2;
P16 = PROBT(T16, DF16) * 2;
if P16 > 1 then P16 = (1 - PROBT(T16, DF16)) * 2;
P23 = PROBT(T23, DF23) * 2;
if P23 > 1 then P23 = (1 - PROBT(T23, DF23)) * 2;
P24 = PROBT(T24, DF24) * 2;
if P24 > 1 then P24 = (1 - PROBT(T24, DF24)) * 2;
P25 = PROBT(T25, DF25) * 2;
if P25 > 1 then P25 = (1 - PROBT(T25, DF25)) * 2;
P26 = PROBT(T26, DF26) * 2;
if P26 > 1 then P26 = (1 - PROBT(T26, DF26)) * 2;
P34 = PROBT(T34, DF34) * 2;
if P34 > 1 then P34 = (1 - PROBT(T34, DF34)) * 2;
P35=PROBT(T35,DF35)*2;
if P35>1 then P35=(1-PROBT(T35,DF35))*2;
P36=PROBT(T36,DF36)*2;
if P36>1 then P36=(1-PROBT(T36,DF36))*2;
P45=PROBT(T45,DF45)*2;
if P45>1 then P45=(1-PROBT(T45,DF45))*2;
P46=PROBT(T46,DF46)*2;
if P46>1 then P46=(1-PROBT(T46,DF46))*2;
P56=PROBT(T56,DF56)*2;
if P56>1 then P56=(1-PROBT(T56,DF56))*2;

*RII STANDARD ERROR OF THE MEAN (Eq. 5);
SDTM = SQRT(TM);

*RELATIVE INCREASE IN VARIANCE DUE TO NONRESPONSE (Eq. 8);
RM = (1+1/MM)*BM/UMBAR;

*DEGREES OF FREEDOM (Eq. 7);
VUI = (MM-1)*(1+1/RM)##2;

*FRACTION OF INFORMATION ABOUT PARAMETER Q WHICH IS MISSING (Eq. 9);
GAMMA = (RM+2/(VUI+3))/(RM+1);

COMMANDS TO PRINT RESULTS;
PRINT T12 T13 T14 T15 T16 T23 T24 T25 T26 T34 T35 T36 T45 T46 T56;
PRINT DF12 DF13 DF14 DF15 DF16 DF23 DF24 DF25 DF26 DF34 DF35 DF36 DF45 DF46 DF56;
run;