Essays in Consumer Debt, Personal Saving Rate, and Household Insolvency in Canada

by

Mohamad Ghaziaskar

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Economics

Waterloo, Ontario, Canada, 2017

Author's Declaration

"I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.
I understand that my thesis may be made electronically available to the public."

Abstract

This thesis consists of three essays attempting to determine the key determinants of spending-saving behaviour and financial stability of Canadian households from both micro and macro economics.

In the first chapter, we try to isolate and evaluate the socio-economic characteristics of households who accumulate debt by spending more than what they earn in a given year. In particular, with a focus on the right tail of spending distribution –households who tend to spend a larger fraction of their income– we use multivariate regression type analysis to isolate socio-economic factors that contribute to debt accumulation and lead to insolvency. We aim to highlight the micro level factors that have contributed to the increase in the proportion of spender households in the population. Specifically, what are the marginal effects of age, income level, education, and family structure on the probability of a given households spends more than its income?

Related to this question, we also consider the effect of budget allocation decisions on the probability of spending more than income and accumulating debt. We find that budget share of specific items in household consumption basket, has important information about the spending-saving behaviour of a household. Our analysis provides valuable information about what goods and services are the main outlays of expenditure for households in severe debt.

The second chapter is about evaluating the relationship between household's saving rate and its long-run income from a more technical perspective. This chapter is an attempt to address the possible endogeneity issue present in this relationship. In addition to the conventional and widely exercised methodology, three alternative approaches are considered, and in a Monte Carlo experiment, the performance of four approaches is tested in three different environments. Results of our analysis show that the conventional methodology outperforms only when there is a simple

linear type of endogeneity in the model. However, when more complicated types of endogeneity are present, it fails to predict saving rate unbiasedly. In the end, using FAMEX and SHS datasets from Canada, we re-evaluate the question with all different methods. Our empirical analysis suggests that more affluent households do save a larger fraction of their income, and the results are consistent across different years not sensitive to different instruments.

Finally, the third chapter looks at the household financial stability from a macroeconomic perspective. Using aggregate data, at the provincial level, on households insolvency rate, we try to point out important aggregate key factors in determining financial instability of households. In a series of panel regression analysis, we explore the effect of aggregate variables such as GPD, unemployment rate, housing prices, interest rate, and household debt level, on the insolvency rate of households. Moreover, in a panel vector autoregressive estimation, we attempt to investigate the interactive effects and consequences of insolvency rate and gross domestic product, while controlling for other related aggregate variables. The key finding is that higher levels of household debt are associated with higher insolvency rate, and insolvency rate has a negative impact on GDP.

Acknowledgements

Firstly, I would like to express my special appreciation to my supervisor Professor Anindya Sen for the continuous support of my Ph.D. study, for his patience, and motivation. It was a great experience to work with him. I learned a lot of things about economics and research from him, but most importantly he has been so inspiring for me and after each meeting with him, I found my energy and desire to do things has been doubled. Prof. Sen, your advice on both research as well as on my career have been priceless.

I also want to thank my co-supervisor Professor Thomas Parker. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study. Tom, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for helping me to grow as a young econometrician. I would also like to thank Professor Mikko Packalen, whose presence brought balance to my committee. Thanks for your insightful comments and encouragement, but also for the hard questions that helped me to widen my research perspectives from different angles.

I would also like to thank the faculty of Economics Department at the University of Waterloo, in particular, Professor Sue Horton, Margaret Insley, and Pierre Chauss, for their great support in many circumstances. I took many wonderful courses in my post-graduate studies. I thank all my instructors, and in particular, Pierre Chauss, Dinghy Xu, John Burbidge, Lutz Busch, and Jean-Paul Lam. I thank Pat Shaw, Pat Gruber, Kayla McKinnon, and Becky Moore, the fabulous staff of our department. I thank all Ph.D. students with whom I collaborated during my Ph.D. study, in particular, Farrokh Darvishi, Sonja Yu Chen, Behnoush Amery and many more.

University of Waterloo provides great services in many different ways for graduate students, among whom I benefited the most from people in the writing center and Career center. I would like to thank Marry McPherson for her great lessons on

my writing. Christine Kampen Robinson and Erica Refling, thank you for all your services, you are one the most valuable assets for graduate students in this university.

My special thanks goes to my consultant, mentor, success coach, and a great friend, Dave Logan in counselling services, who was the biggest help and support through all years. Thanks Dave for all the time you spent with me to encourage me, and show me the best direction and the best decision.

There is no doubt that completing my Ph.D. was impossible without the unconditional and consistent support of my beloved parents and siblings. My absence among them was the largest opportunity cost I bared for six years, and can not be replaced by anything ever. At the end, it all becomes to one person and it is my beloved wife, *Hanieh*, who stood beside me and never gave up in all ups and downs. Words cannot express how much I appreciate everything you have done for me, I love you Hanieh.

Table of Contents

List of Figures

List of Tables

Chapter 1

Household Spending and Consumer Debt in Canada

1.1 Introduction

Since the beginning of the "Great Moderation" in which the US economy gained a significant reduction in the variance of its GDP, consumers have overstretched themselves with spending and debt accumulation. From a global perspective, the US economy is not an exception and high national debt, in which consumer debt holds an outstanding share, has risen in most major economies. The case of Canada is an exception in comparison since it started in a more favourable situation in the early 1980s, but dropped steeply and currently stands in a worse situation.

The personal saving rate in Canada declined from about 19% in the early 1980s to less than 2% in 2005, and after small fluctuations, it was less than 4% in last quarter of 2014. In aggregate terms, household debt increased in Canada from $363 billion in 1990 to about $1 trillion in 2005 and $1.8 trillion in 2014 (6 fold in comparison to 1990).[1] Consumer spending and debt have been investigated extensively in the US, starting as early as the 1990s. However, for Canada, it has become a matter of importance in recent years as the level of household debt is approaching the same level as the US before the major recession in 2008.

While private consumption is an important factor in growing and stimulating the economy, accumulating debt in this sector increases the vulnerability of the household section, and the whole economy in general. In the case of an adverse income shock, with high levels of accumulated debt, the possibility of going to recession is higher and recovering from such recession would take longer. Accordingly, the faster growth of debt compared to disposable income has raised concerns. Consequently, policy-wise it is imperative to evaluate and monitor the relative effects of contributing factors to overspending and debt accumulation.

Moreover, the budget allocation among households who are accumulating debt has important policy implications. Whether a debt is created from spending on

[1] CANSIM table 378-0122.

daycare expenses, university tuition and health care costs have different policy implications than if a debt is created upon expensive vacations and other recreational activities. In this study, we attempt to isolate and evaluate the socio-economic determinants of households who spend more than their income and accumulate debt in a given year. We also consider the effect of the budget share of specific items in the overall budget of households on the probability of a household spending more than its earnings.

Age and income level are considered as the most important determinants of saving/spending decisions. With a focus on the right tail of spending distribution –who spend more than their income– we use multivariate regression type analysis to isolate socio-economic factors that contribute to debt accumulation and lead to insolvency. We aim to highlight the micro level factors that have contributed to the increase in the proportion of spender households in the population. Specifically, what are the marginal effects of age, income level, education, and family structure? Related to this, what are the differences in debt levels across age, income level, education, and family structure? And perhaps most importantly, what goods and services are the main outlays of expenditure for households in severe debt?

Our econometric estimates reveal interesting features of spending behaviour that, to the best of our knowledge, have not been widely documented. For example, households that spend more income than they report –constituting more than 30% of the population– devote a greater portion of their expenditures to mortgage and rent, recreation expenditures and private transport relative to other households. They also allocate a proportionately smaller amount on pension and insurance payments as well as to personal taxes relative to households that spend what they earn. Further, household income is the most significant predictor of whether a household is in debt. Interestingly, neither marital nor household type dummy variables are consistently significant.

The rest of this article is structured as follows. The next section, extensively reviews previous related studies. Section.1.3, explains the Survey of Household Spending, that is used as the main sample in this study. Section.1.4 provides a general analysis of spending behaviour along the distribution of Canadian households. Section.1.5 presents empirical results and concluding remarks are provided in Section.1.6.

1.2 Literature Review

Globally, extant literature has grown around the issue of declined personal saving rate and the increased amount of household debt. Researchers have been trying to address the issue from different aspects such as; (i) The heterogeneity of accumulated debt among different groups of households with different socio-demographic characteristics, trying to link the high amount of debt to specific groups of households. (ii) The causes and factors behind the declined in saving rate, also its effects and consequences from micro and macro economic perspectives. (iii) Sources and uses of household debt. In other words, it has been investigated that how much of household debt is backed up with houses and properties and how much unsecured debt is carried by households, while some others consider the composition of household debt, whether borrowed funds are invested in home equity, education and other types of investment, or they are used to finance consumption.

1.2.1 Heterogeneity in Household Debt

Many researchers have attempted to depict the distribution of households as it relates to saving and debt accumulation and explain the heterogeneity behind the aggregate trends. Some studies tried to find a potential relationship between certain household characteristics and the amount of debt carried. A household's socio-economic characteristics such as age, income, dwelling status and geographic location have been examined to try and answer "what type of households are more likely to save less and accumulate debt?". It would have significantly different policy implications if all household types are saving less proportionately rather than when certain groups of households do so.

For example, Allen and Damar (2012) using the administrative data of household bankruptcy data in Canada and show that bankrupt individuals in Canada are more likely to be unemployed, and are typically renters. A large percentage of Canadians do not owe any mortgage debt but instead maintain bank loans and credit cards from several different creditors.

Performing multivariate regression analysis, Hurst (2011) shows that even after controlling for income, age, and other socio-economic factors, family type has explanatory power. Results from Hurst (2011) show that the highest debt-to-income ratio belongs to younger families, and single parents have the highest debt-to-asset ratio, being more likely to hold a higher debt-service ratio. Couples without children and single individuals are also more likely to have higher debt-to-income ratios than couples with children. On the other hand, family type is not associated with debt-to-asset ratio and debt service ratio of above 40%.

Hurst (2011) also tests for a gap between Canadian-born residents and immigrants. He finds that after controlling for income, education and geographic location, immigrants have significantly higher debt-service ratio compared to people born in Canada.

Chawla and Uppal (2012) provides more detailed information about demographic features and spending patterns of Canadian households. They show that the likelihood of and the amount of debt is higher among younger home-owners, young families with children, the higher-educated, and those with high incomes. This study shows that individuals who are 40 years old and under carry 60% of household debt in Canada, and almost half of the household debt belongs to couples with young children. By assessing the distribution of household debt, they find that debt is more equally distributed among better-educated households and households with higher income. Conversely, debt distribution is substantially unequal among less-educated, unattached individuals and renters.

In another study Chawla (2011) investigates the distribution of mortgage debt among Canadian households and examine the financial burden of homeowners by estimation of mortgage-liability ratio as the share of mortgage payments in household disposable income. He shows the proportion of households who devote at least 20% of their disposable income on mortgage payments has increased faster among younger households.

Faruqui (2008) uses 1999 to 2007 waves of the Canadian Financial Monitor to examine the financial vulnerability of Canadian households. He evaluates the distribution of debt-service burden among Canadian households, and shows that from 1999 to 2007 the proportion of households in the lower tail of DSR distribution had decreased. His findings imply that the growth in debt-to-income ratio has been larger among upper quantiles of income rather than lower-income families. He concludes that "despite the increase in the debt-to-income ratio since the late 1990s, households remain well positioned to manage their increased debt levels". By doing a cross country comparison between US and Canadian households, based on microdata analysis, he finds that Canadians were standing at a better financial position in 2004 than US households.

Simone (2014) analyzes household debt from a different perspective. He explores household debt distribution in neighbourhood scale within Canada's major cities, in order to investigate the role of house prices and neighbourhood composition on the rising levels of household debt. He also tries to relate these factors to the level of indebtedness among immigrants and racialized neighbourhoods to see if concentrations of immigrants and visible minorities are associated with higher debt, after controlling for other demographic factors.

In Toronto, Montreal and Vancouver, Simone (2014) finds higher levels of indebtedness for suburban regions relative to central regions. He also determines a positive relationship between credit card debt and other consumer debts and proportion of visible minorities and immigrants. A significant trend is depicted in Simone (2014) that neighbourhoods concentrating Chinese in each of the three global cities present higher levels of credit card debt and other consumer debts, while neighbourhoods with Blacks and Latin Americans have lower total household debt levels. He also shows a negative relationship between proportion of seniors and renter households in neighbourhoods and the debt to disposable income ratio in that neighbourhood.

Walks (2013) also provides a cross-sectional mapping of the geography of household debt in Canada's major cities. Consistent with Simone (2014), he finds that the amount of debt is lower in smaller and slower-growing cities, while more dominated by unsecured forms of consumer debt. Also neighbourhoods housing younger families revealing higher debt loads. However Walks (2013) does not find any significant evidence supporting that concentration of immigrants in the global cities is a factor pushing up debt levels.

In a Bank of Canada discussion paper, Meh et al. (2009) uses Canadian Financial Security survey from 1999 to 2005 to investigate different measures of debt as debt to asset and debt to income ratios at different quantiles of income and wealth distribution. They identify an increasing trend in debt to income ratio at the lower end of income distribution that are more vulnerable to interest rate increases and negative income shocks. This is in contrast to the findings of Faruqui (2008). Meh et al. (2009) also provide evidence that shows households balance sheets are more sensitive to housing prices relative to past, as mortgage payments are taking up a larger share of household budget.

In another study Chawla and Wannell (2005) use two waves of the Family Expenditure Survey to compare the patterns of saving and spending between two years of 1982 and 2001. They categorize households into two groups; spenders and savers, that are characterized as those who spend more than their reported income and those who spend less than their reported income, respectively. Using summary statistics, they compare the two groups at two time periods and evaluate micro factors contributing to the declining saving rate in Canada.

For example, they find that on average, spender households actually spent more than saver households in 1982, even though they made about 28% less on average. In 2001, as the income gap rose to 35% between the two groups, the average expenditure of saver households exceeds that of spenders. They also investigate the composition of two household types in terms of income and other demographic variables, e.g., "in 1982, 57% of all households with incomes under $20,000 were spenders compared with 16% of those with incomes of $100,000 and over. By 2001, the proportions had risen to 66% and 23%, respectively."

The most striking finding of Chawla and Wannell (2005) is that, on average, spender households spend more than savers even in absolute dollars, respectively $39,000 vs. $33,900 in 1982, and $41,700 spent by spender households vs $37,900 spent by savers in 2001. They show that car purchases contribute the most in this regard.

Burbidge and Davies (1994) is an early and widely cited study that uses Canadian survey data to examine the age profile of income, consumption, saving and holding of wealth. They use the Family Expenditure Survey and Survey of Consumer Finances, and employ a "Kernel-Smoothed Quantiles" technique to estimate a smoothed distribution for saving over the life cycle. They show that saving rate peaks at age 45 and increases across higher quantiles of income. They also estimate an age distribution for all other indicators such as income, consumption and different measures of saving.

1.2.2 Sources and Uses of Household Debt

It is of primary importance to know how much of household debt is secured by properties and equities, and how much unsecured debt is carried by different groups of households. Also the evolution of this ratio holds valuable information.

Bailliu et al. (2012) use 1999 to 2010 waves of Canadian Financial Monitor survey data to examine the sources and uses of household borrowing. They categorize household debt into three categories of: i) unsecured consumer debt flows, ii) home-equity extraction (borrowing against equity in existing houses through increases in mortgage debt and draws on home-equity lines of credit); and iii) mortgage debt flows associated with the purchase of newly constructed houses. They find that increases in home-secured debt, particularly home-equity extraction, contributed the largest share to the rise in total household debt in Canada between 1999 and 2010.

Mian and Sufi (2009) use individual-level data on home ownership and defaults in the United States for the period of 1997 to 2008 and provide evidence that existing homeowners borrowing against the increase in their house prices stands for a great portion of increase in the household leverage from 2002 to 2006, and also for the increase in the number of defaults from 2006 to 2008. Their estimates show a one standard deviation increase in house prices results in 0.25 standard deviation increase in the debt to income ratio. They discovered that the average homeowner borrows 25 to 30 cents for every dollar increase in house prices.

Findings in Mian and Sufi (2009) suggest that borrowed funds against increasing home equity are not used for any new real estate investment or repayments of credit card balances, but more spending on consumption and home renovations. They show that younger households, those with low credit scores and households with high initial credit card usage, are more likely to borrow against increased home equity. According to Mian and Sufi (2009), and of defaults from 2006 to 2008 originated from home-equity extraction, standing for 2.8% of GDP.

Another aspect of Canadian household debt that has not been widely investigated is the uses of household borrowing. Borrowed funds by households can be used toward different purposes such as purchasing a new home or renovating one, purchasing goods and services, investment and financing current debts. Knowing how households have used borrowed funds can be profoundly important for estimating the effects of increase and decrease in the amount of household debt and also in estimating the effectiveness of fiscal and monetary policies. Whether borrowed funds are used by households to finance consumer spending or they are invested toward housing or any type of investment will have significantly different results.

For example, if households use home equity to borrow funds and use them toward consumption, a decline in house prices reduces their home equity and brings down their borrowing and spending power.

Bailliu et al. (2012) count five different uses that households' debt goes: consumption, home renovation, financial investment, non-financial investment and debt repayment. They analyze these categories for both secured and unsecured sources of debt. From secured debt flows, they show that 40% of home-equity extraction is used to finance consumption and home renovation, 34% is used for financial and non-financial investment and the rest, 26%, is used for debt repayments.

Also the majority of unsecured debt flows (more than 70%) is used toward consumption that has been fairly stable. Debt repayments hold a share of about 14%, financial investment 6% and home renovations about 5%.

1.2.3 Causes and Consequences of Household Indebtedness

Household debt has been rising at a faster pace relative to disposable income in many advanced economies as well as in Canada for decades. But it has been accelerated since the late 1990s. Several factors may have contributed to this international phenomenon. There have been competing theories from both macro and microeconomic level points of view, trying to explain causes of the declined personal saving rate and increasing levels of household debt.

On the other hand, this phenomenon has raised concerns regarding financial health of households, while many researchers has cast doubt about future economic growth. Accordingly, numerous studies have investigated the effects and consequences of the phenomenon on different aspects of households' financial health and the growth of the economy in general.

Causes

In the late 1980s, after the major decline in the personal saving rate that occurred in almost all major economies such as US, UK, Japan and as well as Canada, micro level investigation of the reasons of the decline in saving rate started. Almost all studies fail to explain the decline in personal saving rate by demographic and micro level factors, but valuable facts were discovered. The declining trend is universal and even though dramatic, not sudden. Furthermore, changes in demographic variables and population structure are very slow and gradual to have any explanatory power. An early documentation of theories and facts about the decline in saving rate can be found in Browning and Lusardi (1996). They count eleven potential explanations for the decline in saving rate in 1980s:

> *"Changes in the age structure of the population; changes in the saving propensities of different cohorts; changes in household structure; increased government insurance; changes in the distribution of income; the decline in aggregate growth; capital gains on housing; capital gains on stocks; the increased annuitization of wealth; cash payouts by firms to shareholders; and the development of financial markets"*

In the following a selection of these hypotheses are discussed.

Ageing Population

In the Life Cycle Model framework, elderly are supposed to not save and cut down assets, accordingly if the proportion of elderly people increases in the population. When the baby-boomer generation reaches their 60s, aggregate saving rate should fall. The above proposition can be questioned from two different aspects. i) Is the population truly aging? And ,is the movement in age profile of population have enough variation that have any explanatory power?, ii) Assuming that population is actually getting older, does saving rate vary significantly across age groups?

Sabelhaus et al. (1991) provide evidence from US Consumer Expenditure Surveys and show that movements in demographics have a modest contribution (if any at all) in the decline in saving rate. They specifically show that saving rate has decreased among all age groups. They show that calculating saving rate in 1983, while keeping the age profile the same as 1973, would result in an increase of only about 1%.

For the case of Canada in particular, Burbidge and Davies (1994) explicitly remark that "continued saving after retirement is more the rule than the exception and saving rate does not decrease among elderly people" in Canada. They use waves of Survey of Family Expenditure and show that saving rate changes more by income level than by age.

In a more recent study, Crawford and Faruqui (2012) consider the other side of the issue, which is the amount of debt households carry in different stages of life. As a consequence of an aging population, they show that in 1999 a households with older heads hold lower levels of debt compared to 2010. As a counter-fact for 2010, in the absence of population aging –keeping age distribution the same as in 1999, they find a similar distribution of debt for both years. This is while the proportion of older households has increased and older population tends to hold less debt. They explain that population aging phenomenon has a moderating effect on household indebtedness but it is offset by other aggregate factors. Crawford and

Faruqui (2012) finds that mean of debt level has increased for all age groups from 1999 to 2010 that suggests a systematic change in the environment of cohorts.

> *"For example, while a typical household aged 31 to 35 years in 1999 (i.e., born between 1964 and 1968) had total real debt of approximately $75,000, a representative household in the same age range in 2010 (i.e., born between 1975 and 1979) had a mean real debt of almost $120,000".*

Gokhale et al. (2009) consider three other factors besides age profile; (i) cohort specific effects, (ii) ratio of available resources to output in each age group and (iii) the ratio of total resources to output. Using the US Consumer Expenditure Surveys, they documented main factors of the decline in saving rate as; redistribution of income from current and future young generations to current old generation; and a sharp decrease in the propensity to consume among elderly. Comparing early 1960s to late 1980s, they show that elderly people have access to more resources relative to younger people and have dramatically higher propensity to consume.

There are few potential explanations in the literature for the lower saving rate or higher amount of debt that more recent cohorts are carrying relative to 1980s and early 1990s. First, shifts in the underlying preferences in favour of current consumption relative to future consumption. The cohort that saw the Great Depression and then the World War in their childhood learned that debt is a sin and saving is virtuous. Cross (2015) explains the etymology of debt in European languages roots to words such as "Fault", "Sin" and "Guilt".

On the other hand, the cohort who was raised in the post-war period, witnessed strong economic growth and prosperity. They had a different view toward current and future consumption and debt was significantly more accessible to them. Consequently, more recent generations have a higher demand for credit due to higher expectations of future income. Lower interest rates play an enhancing role while new financial innovations play the easing role to provide higher and easier access to credit for households.

Rise in Income Inequality

It is believed that income inequality –larger income growth in higher quantiles of income distribution– pushed people in the middle class to consume more and save less. After the seminal work of Piketty and Saez (2001) about the issue of income inequality in the United States, a vast literature base was established to further investigate the consequences of this increased inequality on the economy. One of these consequences is the phenomenon known as "*Trickle-Down Consumption*", "*Conspicuous Consumption*" or the phenomenon of "*Status Seeking*" among middle or lower income households. This phenomenon is motivated when income and consumption rises at the top of income distribution significantly faster than lower quantiles, induces households in lower quantiles of income to devote a larger share of their budget to consumption, which is often luxury-oriented, in order to signal a higher socio-economic status.

Another twist in the income distribution that has potential contribution in the decline in saving rate, is the conventional notion that higher quantiles of distribution save a larger proportion of their income. Usually this is taken as a basic principle, especially by non-economists. However, economists have been skeptical to this proposition and literature is very inconclusive. If households in higher quantiles of income save a larger proportion of their income, with increase in their share of income in household sector, aggregate household saving rate must rise, in other words, a lower saving rate must be the consequence of lower inequality.[2]

Bertrand and Morse (2012) explore the household level data to find evidence in favour of "*Trickle-Down Consumption*" and test the hypothesis of whether the increased income inequality in the last three decades is related to declined saving rate

[2] See Browning and Lusardi (1996).

that happened in the same timeframe. They use the panel of Consumer Expenditure Survey (CEX) and Panel Study of Income Dynamics (PSID) from the US to exploit state-year variation in income and consumption of the top quantiles of income distribution to find supporting evidence for their hypothesis.

They regress log consumption of households on the log of average consumption in 80^{th} quantile of income distribution –rich people– at state-year level, proxied by the log of average income. Their important finding is that a one percent increase in consumption of the rich is associated with 0.182 percent increase in the consumption of households in lower quantiles, holding income constant – the effect is even stronger among only 20^{th} to 80^{th} percentiles.

Their findings suggest that middle income households that are more exposed to more affluent people in their market would spend a greater share of income, when income is held constant. In their framework, they do not find any strong evidence in favour of traditional theories such as permanent income hypothesis, precautionary saving and wealth effect.

In their behavioural analysis they suggest that the increased consumption is more supply driven in the form of luxury goods provided in market that creates the desire to seek a higher social status through consuming more visible goods.

Moav and Neeman (2012) takes a game theory approach from macro perspective and develops an overlapping generations model in which households care about their economic status and signal their unobserved income higher falsely. In particular they show that if human capital is observable and correlated with income, individuals with lower human capital in a signalling equilibrium devote a greater share to conspicuous consumption –goods that are more visible and do not help with increasing human capital– while more educated people tend to spend less.

In their model, an individual's utility is a function of his consumption, investment in human capital of their offspring and their own socio-economic status. In this framework, Moav and Neeman (2012) provide an explanation for the increasing saving rate associated with higher levels of income, at the same time why poverty is persistent and individuals cannot escape the poverty trap.

Considering the relationship of saving rate and income, the most comprehensive empirical study is provided by Dynan (2004). In this study the hypothesis is strongly supported that higher permanent income –or life-time income– is associated with higher saving rate. She uses the Panel Study of Income Dynamics (PSID) and estimate saving rate for five quintiles of permanent income proxied by different instruments such lagged income and non-durable consumption.

Dynan (2004) show that not only saving rate increases with income, bequest and precautionary motives for saving are stronger in higher quantiles of households compared to lower quantiles. This implies increasing slope rather than a constant correlation between saving rate and income level.

On the other hand, in a more recent study by Alan et al. (2015), they follow Dynan (2004)'s methodology and apply it to Canadian Family Expenditure (FAMEX) survey data. The striking finding is that saving rate does not increase by permanent income in Canada, at least for the upper quantiles of income compared to middle quintiles.

Low Interest Rate and Relatively Lower Cost of Borrowing

Mortgage debt forms about 70% of household debt in Canada. Unlike many OECD countries, Canada's mortgage market did not experience a severe downturn in 2008, and monetary policy worked so well that it became internationally known as a successful financial regulation[3]. Even though mortgage borrowing decreased in many

[3] See Kiff (2009).

countries during and after the 2008 recession, the Canadian mortgage market continued to grow and consequently, the debt-to-disposable-income ratio reached for a new high year after year, in the winter of 2014 it was 163.3%.[4] This is even higher than what it was in the US and UK when the recession occurred in 2008.

Fortin (2014) is one of few studies that uses aggregate data to estimate a demand function for mortgage loans market in Canada. He estimates three equations for three different measures of mortgage debt, i) average real value of new mortgage loans, ii) number of new mortgage loans and iii) flow of real repayments of existing loans. The results from this study show that: first, interest rate influences mostly number of new loans –one percent lower interest rate increases number of mortgage loans by 13%. Second, house price is the main source of changes in average value of new loans –one percent increase in house prices, increases the average value of new mortgage loans by about 8%. Lower inflation – lower real interest rate– also increases household mortgage debt and reduces the rate of repayment of existing loans.

Findings in Fortin (2014) suggest that the stable rise in (real) house prices in addition to low inflation was the most forceful factor behind the growth of mortgage debt from 2000 to 2007 in Canada. Following that, sustained low interest rate increased the number of new loans.

Effects and Consequences

Many researchers have tried to investigate the effects and consequences of rising household debt on general financial health of households in the short run, while some others consider facts in the long run. Different perspectives on this issue propose different scenarios and have resulted in some disagreements among economists. While many economists are concerned with the vulnerability of households with respect to short run negative shocks, others try to highlight benefits of created investments and wealth in long-run.

[4] The Daily - National balance sheet and financial flow accounts

In a Bank of Canada study, Crawford and Faruqui (2012) use Canadian Financial Monitor survey data in order to investigate the trends in household indebtedness from both supply and demand perspectives. On the demand side, households enjoy a smoother consumption, and are able to allow their consumption to be different from income in different stages of life. Also in the event of a loss in income they can maintain a more stable consumption when they have access to the credit market. Finally, it makes them able to borrow and invest in housing, education and other types of financial and non-financial investments.

On the other hand, while individual households were enjoying the recent financial innovations in the credit market, the major recession in 2008 showed that excessively accumulating debt and the relaxed standards of lending is a threat to financial health of households and makes the economy vulnerable to negative shocks, which highlights the importance of considering supply factors.

Long-Run

Long-run consequences of rising household debt has been more controversial than those in the short-run. Some researchers highlight the role of investment and wealth created through borrowed funds, while others cast doubt about the adequacy of savings of the baby-boomer generation for their retirement.

Crossan et al. (2014) evaluates financial literacy of Canadian households through an internationally comparable survey data on financial literacy and retirement planning. The data is collected through a questionnaire designed to be comparable to surveys conducted in a number of other countries such as Australia, France, Germany, Italy, Japan, and the United States. It is also investigated what types of individuals have higher financial knowledge and are planning for retirement. The survey contains three simple questions about interest rate compounding, inflation and stock market and risk diversification. 42% percent of respondents answered three questions correctly. While it sounds low, in comparison with other countries, Canadians are doing fairly similar.

Findings from this survey show that Canadians are standing in a better position compared to Americans with score of 32%, while they do worse than Germans who score 52%. Among Canadians, young, old, women, minorities and individuals with lower education have less financial knowledge. Provincially, individuals in New Brunswick, Prince Edward Island, Nova Scotia and French-speaking individuals in Quebec have lower scores.

In a research bulletin, Cross (2015) discusses the long-term perspective of household debt in Canada and tries to provide evidence that Canadian households are spending more responsibly. He explains that credit is a recent innovation and it is too early to judge what is the right and optimal growth rate for household debt.

He describes that debt-to-income ratio is as high as in the US but lower than many other advanced economies. Also the problem with the US debt-to-income ratio was the flawed distribution of debt, not the high ratio. Cross (2015) emphasizes that the long-term increase in household debt has resulted a much larger gain in household wealth and assets, uplifting both incomes and consumption. He notes that the growth of credit has declined to one-third of what it was in the 2008 recession, even though the interest rate has been at its lowest in the post-recessions period.

Short-Run

The household debt issue became more important than any time, after the economic recession in 2008. Slow recovery of the US economy after the 2008 recession, raised more concerns around the levels of household debt. The issue in Canada is not as serious as in the US, however the declining saving rate trend in Canada has shown a much steeper slope than the US. Increasing levels of household debt is a prominent factor in the possibility of any recession in future, the vulnerability of households in case of recession and more difficult and longer recovery periods after any future recession.

Applying panel analysis on debt to income ratio of households in a list of advanced economies and over the period of major recession in 2008, Bailliu et al. (2012) suggest that household indebtedness constitutes an important source of risk to household spending. This risk makes households more vulnerable to consequences of economic downturns in the event of decline in house prices.

They show that countries with the largest increases in both house prices and their ratios of household debt to income in the decade leading up to the 2008 crisis tended to experience the largest contractions in consumption during the subsequent recession. Also countries with high ratios of household debt to income tend to experience more severe and prolonged recessions.

Mian and Sufi (2011) also use US data to explain the slow recovery of the US economy from the 2008 recession through the accumulated levels of household debt that was elevated in the housing boom era. Performing their analysis at county level, they focus on the two tails of debt-to-income ratio distribution, and show that counties that raised larger amounts of debt during the housing boom, experienced a much slower recovery in the post-recession period. In particular, their analysis reveals that, auto sales in counties with high debt-to-income ratio households began to decline in early 2006, long before sales began to fall in the low household debt counties. During the recession, all counties underwent a major decline in auto sales, however, counties with low debt households, faced a more robust recovery. In contrast, even after the official end of the recession, counties with high debt level households, stayed involved in a severe recessionary environment. They conclude by stating that ". . . Our view is that the depth and length of the current recession relative to previous recessions, is closely linked to the tremendous rise in household debt that preceded it. This view is supported by survey evidence that the main worry of businesses is sales, not financing."

Multiple studies use household level data on consumption and credit card debt to investigate a relationship between household debt and consumption. Ekici and Dunn (2010) use the Consumer Expenditure Survey and the Ohio Economic Survey from US to provide evidence on a negative correlation between credit card debt and consumption. As claimed by Ekici and Dunn (2010) , a $1000 increase in credit card debt results in almost 2% decrease in quarterly consumption growth of households. According to their estimations, the relationship is not sensitive to durable and non-durable consumption.

1.3 Data

Data used in this study are drawn from a famous Canadian survey, the Survey of Household Spending (SHS), which is a revised version of the older Family Expenditure Survey (FAMEX). This survey is conducted annually by Statistics Canada in the ten provinces and usually every other year in the territories and offers available data from 1997 onwards. However, the data are not based on a panel of the same households over time. SHS collects detailed information on household specific expenditures such as the annual income of household members (from administrative data files), demographic characteristics of the household, certain dwelling characteristics (e.g., type, age and tenure) and certain information on household equipment (e.g., electronics and communications equipment). Unfortunately, SHS lacks wealth information and does not include the specific household assets and liabilities, although it does record yearly net changes in assets and liabilities.

The coverage rate has been quite high in SHS, above 95% in most years. However the response rate has declined from about 80% in 1997 to about 65% in 2009[5].

[5] See Barrett et al. (2013)

Cross-section samples in SHS have a stratified, two-stage design. In the first stage, a sample of geographic areas is selected. Then, from a list of all the dwellings in selected areas –called clusters– a sample of dwellings is chosen. The ample size ranges between 12000 to 17000 households.

1.4 General Analysis of Spending Behaviour in Canada

In this section, in order to obtain a better view of the debt accumulation in different types of Canadian households, we first provide a broader picture of the aggregate trends obtained from our sample. We do this in two levels, in the first level we define over-spending groups broadly. More specifically, we divide our sample into three groups: Spenders, who spend more than 105% of income; Savers who spend less than 95% of their income; and finally, Balanced-Households, whose expenditure is equal to their income with a 5% margin of error.

In level two, we break down these categories further for deeper analysis. Specifically, summary statistics are calculated for; (1) households whose consumption is 80% or less than their income; (2) households whose consumption is 50% less than their income. The motivation is to study differences in consumption between households who do not spend a significant amount of their income. Similarly, we construct summary statistics for; (3) households spending more than 110% of their income; (4) households spending more than 120% of income; and (5) households spending 150% or more of their income. The aim is to identify substantial differences in consumption patterns between households with different spending tendencies.

We calculate the income distribution of different spending groups, as well as the proportion of each spending group at different income levels. At the end of this section, we use summary statistics to calculate the budget allocation of households in different spending categories.

1.4.1 Distribution of Households by Overall Spending

The general distribution of households by their propensity to spend is presented in Figure.1.1 from 1997 to 2009. The distribution of households is broken into eight groups by their tendency to spend.

The general picture looks very similar for all survey years; however, if we examine each spending group, specifically in Figure.1.2, we can see that the proportion of households with different propensities to consume have different patterns since 1997. Between %34 to %40 of households spend more than %105 of their income in all survey years. The proportion of households who spend less than %80 of their total income, decreases slightly, from %14.5 in 1997 to %13 in 2002, and increases to %22 in 2009.[6] On the other hand, the proportion of households who spend more than %120 of their total income increases, from %17 in 1997 to %25 in 2005, and decreases again, to %19 in 2009. It worth mentioning that households with an MPC of greater than %150 make up about %7 of the sample each year, which is not trivial. Households with an MPC between %110 and %120 have a constant share of %10 in the distribution. Households who spend more than %95 and less than %110 of their total income have a monotone decreasing pattern, from %29 to about %20.

Understanding the relative effects of different socio-economic factors on the probability of falling into either the Spender or Saver category is of key policy importance. Sample means of key variables are calculated for the whole sample for various categories of spending, year by year. Once we compare the sample means with respect to household characteristics, some interesting differences between Saver and Spender households are revealed. Sample means are very similar across years. The propor-

[6] The spike in 2006 is due to the fact mentioned earlier: the "Balance Edit" is not applied to data in this year.

tion of households spending more than income increases in higher age categories, and reaches its maximum at category of households with a head of 40-49 years old, then decreases as the head of household approaches retirement. The group of low income households that actually made less than 50,000$ in the survey year, had accumulated the highest level of debt.

In terms of home ownership, less than 30% of Spender households own a dwelling with no mortgage, while from 30% to 36% possess a dwelling with a mortgage. Comparing the proportion of Spender and Saver households reveals that the number of households owning a dwelling without a mortgage increases quite significantly among the Saver households. The proportion of households with a head who is younger than 30 years old, is much higher than saver households in the same age group, which is not surprising. The proportion of households that live in apartments is higher among the Spender group compared to Saver households.

1.4.2 Income Distribution of Spender Households

Figure.1.3 shows the proportion of different income groups among Spender households –households who spend more than %105 of their income. Households with less than $50,000 yearly income form the majority of spender households; among them, the proportion of those with less than 30,000$ yearly income has declined slightly. Spender households whose income is between 50,000$ and 75,000$, take up about 20%. The key trend, however, is on the right tail of the income distribution: the proportion of Spender households with an income above 75,000 has increased. More specifically, households in debt, with an income that exceeds $100,000 and $150,000, do not account for more than %12 of the in-debt households at maximum in 2005, however, they show a constant increase from 1997 to 2009, at an average rate of %4 per year. Considering the magnitude of income and debt in these groups relative to households with less than $30,000 income, they are not negligible in aggregate terms.

Figure.1.4 breaks the category of spender households into three spending groups; i) those who spend between %105 and %110 of their income, ii) as well as households who spend between %110 and %120 of their reported income, and iii) households with an MPC of greater than %120. Even though we see some movements in the two categories of households whose MPC is between %105 and %120 the proportion of different income categories among households who spend more than %120, has been quite stable since 1997.

Figure.1.5 shows the flip side of the coin, which is the proportion of in-debt households in each income category. From 1997 to 2009, between 44.9% to 56.5% of the households with less than $30,000 income were in debt and have shown a gradual increase since 1997. The surprising fact is the magnitude and rapid increase of overspending among more well-off households with an income above $75,000 and $100,000. The proportion of overspending households in this income bracket, have increased from 22.2% in 1998 to more than 29.5% in 2008. Moreover, about 35.7% of the households with income above $100,000 spent more than their income in 2005. Figure.1.6 shows the proportion of households in debt for all income levels.

1.4.3 Changes in Income and Consumption 1997-2009

Income is widely considered as the main factor that affects the spending trends of low-income households.

Figure.1.3 shows the trends in the proportion of households in different income categories (in real terms). The proportion of households with less than $30,000 income has decreased from %34 in 1997 to about %27 in 2009. A reverse trend can be seen in category of households with income between $100,000 and $150,000, that increases %7 to %13. The proportion of households in categories of %30,000 to $75,000 do not show a significant change and remains around %20 to %24. The share of the category of highest income households also has increased by more than double, from %2.4 in 1997 to %5.7 in 2009.

Figure.1.7 illustrates a better picture of the income distribution. Ten deciles of income distribution are depicted from 1997 to 2009, and you can see a monotonic but not equal increase in all quantiles of income. While income has increased %12 in the first decile, there has been a %23 increase in income in the highest decile. Again more emphasis must be put on the magnitude of income in higher quantiles of income.

1.4.4 Trends in Specific Items

In this section, summary statistics of a broad variety of items are analyzed. The items that we study are: Shelter (G001), Personal Taxes (O201), Food (F001), Transport (K001), Household Operations (H001), Recreational (M101), Personal Insurance and Pension (O301), Clothing (J001), Health Expenditures (L101), Tobacco and Alcohol (N101), Household Furnishings and Equipment (I001), and Education (M301) plus Tuition(M308). Almost all households report positive expenditures with respect to these items.[7]

Furthermore, we analyze summary statistics of some sub-category items. One example includes the trends such as of relative share of home-made food expenditure, vs outdoor food expenditure. Another policy wise important item, that is categorized under of Total "Household Operations", includes childcare expenses that we specifically analyze its levels and trends. We discuss the movements in rent and mortgage payments under "Shelter" expenses. Similarly, we analyze main contributors to group of total "Transport Expenditures" - gasoline expenditure and expenses on operation of owned and leased vehicles. Operation of owned and leased vehicles includes: accessories, maintenance of used or lease vehicles, insurance premiums, and gasoline expenditures.

Personal taxes include: "income taxes paid in the reference year for that year

[7]The identifiers in the parentheses are the actual codes from the Surveys of Household Spending in 2009.

and on income from previous years if applicable. Also included are other personal taxes (e.g., gift taxes) minus income tax refunds received in the reference year, except for federal Child Tax Benefits, Goods and Services Tax Credits and Provincial Tax Credits. These tax credits are included in average household income before taxes".

Pension funds and insurance payments expenses are included of: "payments for life insurance, annuities, employment insurance, public and private pension plans, and similar items". In this category, further pension contributions are further assessed.

Tuition fee expenditures are also studied in isolation from other education items.

In Figure.1.8 and 1.9, expenditure share of broad categories of household budget are depicted. From this broad perspective there has been little change in the weights of different expenditure categories. It is obvious that the mass of the budget (about 60%) is allocated to shelter, personal taxes, food and transportation.

1.4.5 Trends in Item Specific Expenditures by Spending Category

We now present the expenditure share of specific items for households in different spending groups from 1997 to 2009. We categorize households into six spending groups. Households who spend exactly their income (with %5 margin of error), categories of Saver households; spending less than 80% and between %80 and %95 of income, and two categories of spender households; those who spend between 110% and 120% of their income, and those who spend more than %120 of their income.

Figure.1.10 and 1.11 show trends of expenditure share of specific items in a household budget from various spending categories. Household Operations includes expenditures on communications, child care, and cleaning supplies, and constitutes between 5% to 7% of total expenditures across all households. Total share of household operation shows a gradual increase about %1 of total expenditure from 1997 to 2009 in all types of households. The interesting point about this category of consumption is that no specific pattern is visible with respect to household spending categories. Spenders who overspend more than %120 of their income, spend nearly the same portion of their budget on Household Operation as Saver households who spend less than %80 of their income. However households who spend what they earn, and those with MPC between %105 and %110 spend higher portion of their budget on household operation expenses compare to other households.

Expenditures on education are relatively low and show a declining trend since 2003 for most households, It is the lowest for Saver households. Tuition and fees however, have had an almost constant share ranging from %0.3 to more than %1 for different household categories.

Food expenditures are unsurprisingly a significant portion of total expenditure for all households, ranging from 12% to 16%. Surprisingly households with the highest MPC have the smallest food expenditure share -quite stable around %12-, while other households have a declining share of food expenditure from %16 to less than %13. Excluding the highest MPC category of households, other categories show a converging trend in food expenditure share over the time. The highest budget share for food expenditure differs %3 from the lowest in 1997, and the difference decreases to %1.2 in 2008.

Figure.1.12 shows that almost all types of households have had a declining share on homemade food until 2005 and it starts to rise since after. Restaurant food expenditure has almost the reverse trend. Spender households have the lowest homemade food expenditure and the highest share in outdoor food expenditure share. Considering the weight of homemade food in total food expenditure -around %80-, it is obvious that the declining trend in total food expenditure share is driven by the home-made food component.

As we saw in Figure.1.8, the largest chunk of consumer budget is cut by shelter expenses. Figure.1.10 gives a better picture of share of shelter expenses in budget of households with different spending habits. Households with the lowest MPC, devote %19 of their expenditure to shelter in 1997 and it increases to more than %21 in 2009, while households with a MPC of greater than %110 devote about %24 of their budget in 2009. The most surprising trend belongs to Spender households with a MPC more than %120. From year 2000 to 2009, budget share of shelter expenses increases from %20 -lowest share among spender households - to %24 which is above all other types of households in 2009. This must be the result of a policy change that has helped low income households to devote more share to housing expenditures.

Figure.1.13 confirms that a significant portion of shelter expenditures, %50 on average, consists of mortgage and rent payments. Other shelter related expenditures include repairs and maintenance, condominium charges and property taxes, homeowner insurance premiums, and water, fuel, and electricity, and traveler accommodation.

Personal taxes have less than %10 of budget share in spender households while more than %25 in saver households. Since taxes are related to household income level, it reveals the relationship between spending behavior and income level. Similarly, pension contributions and insurance payments are higher for Savers relative to Spender households.

Furniture expenditures which consist of all indoor and outdoor furniture and household appliances, are a relatively small part of overall expenditures and a positive relationship can be identified between MPC and the Furniture expenditure share. The higher the MPC, the larger the furniture expenditure share. The same relationship can be seen for transportation, higher MPC households devote larger share of their budget to transportation expenditures. However, transportation share varies more between households with different propensity to consume. Saver households with a MPC of less than %80, have less than %10 expenditure share on transportation, while this is about %20 for households who spend more than %120 of their income.

Furthermore, most of the variation in expenditure share of transportation among different spending groups is attributable to corresponding differences in private transport expenditures that has about %90 weight in total transportation expenditures and %8 to %18 of total expenditures. In Figure.1.14, surprisingly we see that higher propensity to consume is correlated with higher share of private transportation and lower share of public transportation in total transportation expenditures. The fact that Spenders households devote a higher budget share to private transportation can be explained by lower income, however the fact that they spend less on public transportation could be caused by higher expenditure on flight expenses by Saver households that may not be significant in the Spenders budget.

Health care expenditures are a limited portion of total expenditures, about %3, however they show a significant and consistent upward trend over time. It is hard to recognize any relationship between expenditure share and spending behavior of household in this category. Health expenditures include household expenses on supplies such as first aid kits, bandages, hearing aids, thermometers, wheelchairs and other appliances, medicinal and pharmaceutical products, prescription medicines, physicians' care, eye-care goods and services, dental services, hospital care, health insurance premiums, and personal care.

Recreational expenditures consist of a wide range of goods and services that include: sports equipment, toys and electronic games, computer equipment and supplies, photographic goods and services, recreational vehicles (such as bicycles), home entertainment equipment and services, admission to movies, live sports events, live arts, cable-vision and satellite services, sports activities and children's camps, and reading materials. While expenditure share shows a growing path for Saver households, Spender households tend to devote a smaller share to recreational expenses in

2009 compare to 1997.

Expenditures on tobacco and alcohol constitute from %2 to %3.9 of total expenditures. Spender households incur a slightly higher proportion than saver households, but no specific pattern is recognizable between expenditure share and spending behavior.

Overall, these results indicate that in some categories household from different spending group have similar patterns. Food, shelter (net of mortgage and rent), mortgage and rent, private transport, and personal taxes and insurance occupy the majority of household budget, even though some subtleties are observable. Spender households, who add up to about 30% of the population, allocate larger share of their budget to mortgage and rent, recreation expenditures and private transport relative to other households. Consequently, there would be less room for pension and insurance payments, in addition to personal taxes relative to households in other spending groups.

1.5 Econometric Estimates

In this section, we draw our attention to identify a relationship between socio-economic characteristics of households, and the incidence of household consumption exceeding its income. In particular, we test if any of the observed characteristics can meaningfully predict the probability of this incident. Toward this end, we implement the following multivariate regression model to assess the relative effects of marital status, gender and age group of the head of household, household size, household total income and the ownership situation of the dwelling in which household resides in.

$$
\begin{aligned}
OS = \beta_0 + \beta_1 female + \beta_2 married \\
+ \beta_9 age_{30} + \beta_{10} age_{40} + \beta_{11} age_{50} + \beta_{12} age_{60} + \beta_{13} age_{70} \\
+ \beta_3 + INC_{30}^{50} + \beta_4 INC_{50}^{75} + \beta_5 INC_{75}^{100} + \beta_6 INC_{100}^{150} + \beta_7 INC_{150}^{200} + \beta_8 INC_{200} \\
+ \beta_9 HHsize_1 + \beta_{10} HHsize_2 + \beta_{11} HHsize_3 + \beta_{12} HHsize_4 + \beta_{13} HHsize_5 + \beta HHsize_6 \\
+ \beta_{14} OwnerNoMortgage + \beta_{15} Tenant + \beta_{16} weeksworked \quad (1.1)
\end{aligned}
$$

We use Ordinary Least Squares (OLS) to estimate this simple empirical specification . The dependent and all independent variables have the subscript "i" which refers to the ith observation -household- in the survey data. Most of these variables are dummy variables.

$female$ is a dummy variable to control any effect for households with female head. The marriage variable is equal to 1 if the individual is legally married or in a common law relationship, and 0 otherwise. age dummy variables represent the age of the respondent and take a value of 1 if the respondent falls in the following brackets: 30 to 39 years; 40 to 49 years; 50 to 59 years; 60 to 69 years; and 70 years and up, omitted age category is 20 to 29 years.

Income dummy variables INC_i^j takes value of 1 if the total household income falls in the interval of ($i \times 1000\$$ to $j \times 1000\$$), and 0 otherwise, omitted income category is the group of households with less than $30,000. We include dummy variables for different household sizes between 2 and 3, 4 and 5, and 6 and higher (omitted category is the households with size of 1).

In order to control for the effect of tenure type, we include dummy variables for each type. Variable $OwnernoMortgage$ represents households who own a dwelling while paying off their mortgage and $Tenant$ represents those who are living in a rented dwelling, the omitted category is households who own a dwelling without mortgage. *weeksworked* also represents total number of weeks worked by reference person and its spouse together. Lastly, province dummy variables are employed to capture, if there is any, unobserved heterogeneity at the province level.

We employ the same approach to estimate the effects of the composition of household expenditure on the likelihood of consumption exceeding earned income in that specific year. We perform this for households in different income groups. Toward this end, share of different items in total expenditure is added to Equation3.1 as separate explanatory variables. In particular, we estimate the effects of Homemade Food Expenditure, Outdoor Food Expenditure, Shelter (net of Mortgage and Rent), Mortgage and Rent; Household Operations, Household Furnishing and Equipment, Clothing, Car Operations, Gas Expenditures, Reading Materials, Tuition Fees; other Education expenses, Tobacco products and Alcoholic beverages, Personal Taxes, and Personal Insurance payments.

1.5.1 The Effects of socio-economic factors on Over5Spending

Tables.2.1 to 2.6, contain OLS estimates for the effect of different household characteristics on the likelihood of household spending more than their income, by different percentages, 5%, 10% and 20%. For the matter of convenience in comparison across years, coefficients are depicted in Figure.1.15 to 1.19 and $\times$ signs are for coefficients that are significant at 5% level.

Obtained coefficients of marital status are similar over the sample period and they are not sensitive to how we define the left-hand side variable. Gender of the head of household also does not show any significant effect on probability of spending more than income, regardless of the magnitude of overspending.

The strongest effect, belongs to income dummy variables, they are all statistically significant with negative signs, that is due to how they are defined in relation with omitted category, which is households with less than \$30,000 income. The higher income categories are associated with less likelihood of overspending and the relative likelihood decreases marginally by 0.4 to 1% from households who overspend by more than %105 to %110 and %120 of their total income. The interesting result about income effect is that as time goes on, the marginal effect of moving from 30,000\$–50,000\$ to next higher income group stays the same around 1%, however, the marginal effect of moving from 150,000\$–200,000\$ to next higher income category, increases over time from about 0.3% in 1997 to 1.2% in 2009.

Regarding age dummy variables, all categories are compared to households whose heads are between 20 to 29 years old, the omitted category. While the relative likelihood of overspending in middle ages is not consistently significant with respect to omitted category, it decreases and becomes more significant, with negative sign, as age increases, meaning that compared to the households with head of household in its 20's, older households are less likely to overspend.

Regarding other characteristics of households, it is interesting that having kids reduces the likelihood of spending more than income, and coefficients are statistically significant, implying that families with young children are less likely to accumulate debt, relative to families without kids. On the other hand, the coefficient estimates of the household size dummies are positive, statistically significant and their value increase in larger families. Considering the fact that we control for the number of younger children, positive effect of household size implies that higher number of adults is associated with higher probability of spending more than earnings.

The effect of number of weeks worked by both respondent and spouse is estimated consistently significant and negatively associated with a likelihood of spending exceeding income. The last but not least, with respect to dwelling status, there are three categories; households who own a dwelling without mortgage, those who own with mortgage, as well as households who reside in a rented dwelling. The omitted category is the owners without mortgage. We find that this category is more likely to spend more than income, compared to those who own a dwelling with mortgage and those who do not own any dwelling.

1.5.2 The Effects of The Composition of Household Expenditure - By Income

In this section, we examine the effect of the composition of household expenditure on the likelihood of the incidence that a household spends more than its income. We perform this analysis for households in different income groups and for different magnitudes of overspending. In order to avoid collinearity issues, not all variables employed in previous analysis, are included at this stage, since some of these variables are highly correlated. Moreover, we are mostly focused on the components of household expenses.

We find comparable results for all the survey years and with respect to different magnitudes of overspending. We also find interesting results within varying income groups. The most common result is the negative impact of Food expenditure, Shelter Net of Mortgage and Rent, Mortgage and Rent, and the Household Operations covariates. The negative sign is interpreted such that, households who devote a higher share to these categories in their overall expenditure, are less likely to spend more than their income. This can be explained by substitution effects associated with other non-essential goods or services that are replaced for the core items. Note that we are controlling for income level, and all these impacts are occurring within income groups.

Obtained coefficient of homemade food expenditure is 1.2 to 3 fold larger than expenditure on food from restaurant, and the difference tends to reduce when moving to higher income groups. Considering their negative sign, it conveys the fact that larger share on home-made food is associated with much lower likelihood of overspending. Tax, Insurance and Pension contributions are all, largely and negatively related to overspending. Not surprisingly, these findings express the fact that households that devote higher share to these expenditure categories are less likely to spend beyond their income.

Other covariates that their larger share in overall expenditure are associated with lower probability of overspending, are health expenditures, car operations and tobacco and alcohol expenses. Interestingly, the magnitude of the negative effect, increases with both magnitude of overspending and income level.

Positive coefficients belong to expenditure on Childcare, Household Equipment and Furnishings, Education (net of tuition fees) and tuition payment categories. Childcare expenses are not widely significant and the magnitude varies from year to year. Expenditure on household equipment and furnishings is significant until 2004 with marginal effects ranging from 0.3 to 1.8, and is more significant in higher levels of overspending. While share of the tuition payments is significantly related to probability of overspending, the share of education expenses is not broadly significant.

In general, therefore, it seems that the higher share of core items, such as, home-made food, mortgage and rent, household operations, tax, pension contributions and insurance payments possess the largest coefficients - and are all negatively associated with likelihood of overspending.

1.6 Summary

In this study we used a series of cross-section household survey data, from 1997 to 2009, to investigate the spending distribution of Canadian households and its trend during the sample period. We employed regression analysis to tie the spending behavior to certain household characteristics and their budget allocation decisions.

In order to calculate summary statistics, we performed our analysis in two phases.

First, we divided households into two general groups of Spenders –those who spend 105% and more beyond their income– and Saver households – that spend less than 95% of their annual income. In the second phase, we break each of these groups into three subgroups with different propensity to spend. According to summary statistics, a non-trivial portion of the population – between 35% to 39% of the population belong to spender group, and about 18% of the population spend more than 120% in excess of their annual income in 2009.

The key finding of our analysis from summary statistics is that there has been a shift in the income distribution of Spender households from lower income groups to higher income households from 1997. In particular in 1997, about 70% of Spender households belonged to the group of households with less than 50,000$ income. While in 2009, this number reduced to less than 60%. On the other hand, the proportion of more well off household with income of higher than 100,000$ has increased from less than 5% in 1997 to about 10% in 2009. Also the proportion of average households who overspend, with income between 75,000$ to 100,000$, has increased from 8% to 12% in the same period.

More emphasis must be given to this result. While the majority of Spender households are formed by low income households with less than 50,000$ income, the trend is in favor of households in higher income groups. Considering higher magnitudes of overspending –households who spend 110% or 120% more than their income– the proportion of high income households increases in the population.

By reviewing summary statistics, it becomes apparent that the mass of household budget is devoted to items such as: Shelter, Personal Taxes, Food and Transportation –about 60%. Yet there are some subtleties. Relative to Saver households, Spenders allocate a greater portion of their budget on Mortgage and Rent, Recreation and Private Transportation. In contrast, budget share of Pension Contributions, Insurance payments as well as Personal Taxes have a smaller budget share in the group of spender households.

We used regression analysis to predict the probability of household consumption exceeding its income –overspending– by two different types of predictors. First we use socio-economic characteristics of households to estimate the effects of variables such as; income, age, household structure and dwelling status. Furthermore, we used the budget share of different items in overall household expenditure to identify the linkage between budget allocation and debt accumulation.

Our findings show that household total income is the most powerful predictor of the probability of consumption exceeding income, i.e., higher income groups are less likely to overspend, compared to low income households. Statistically speaking, size of household also plays an important role in this regard and larger household size is associated with higher chance of falling into Spenders category. Gender and marital status of the head of households were found to be not significant.

The second class of regressions, also provides interesting results. Share of Food expenditure, Shelter net of Mortgage and Rent, Mortgage and Rent, Household Operation expenses, Taxes, Insurance Payments and Pension Contributions have all significant negative effect on the probability of household spending more than its income, consistently across all specifications. On the other hand, coefficient estimates of Childcare, Household Equipment and Furnishings, Education (net of tuition fees) and tuition payments possess positive signs and express that households that devote a larger share of their budget to these expenditure categories, are more likely to spend more than their income.

Overall, our key finding is the paradox that the probability of spending more than one's income increases with smaller share of core items such as shelter and household expenditures, as well as pension and insurance payments. Even though this can be due to our inability to control for other unobserved factors, there are alternative explanations. As mentioned above, larger share of core items might be associated with less room for other non-essential items and force households to have a more organized budget allocation plan. In addition, the tendency to spend more on non-core items is associated with crowding out effects on core items that raise the possibility of overspending.

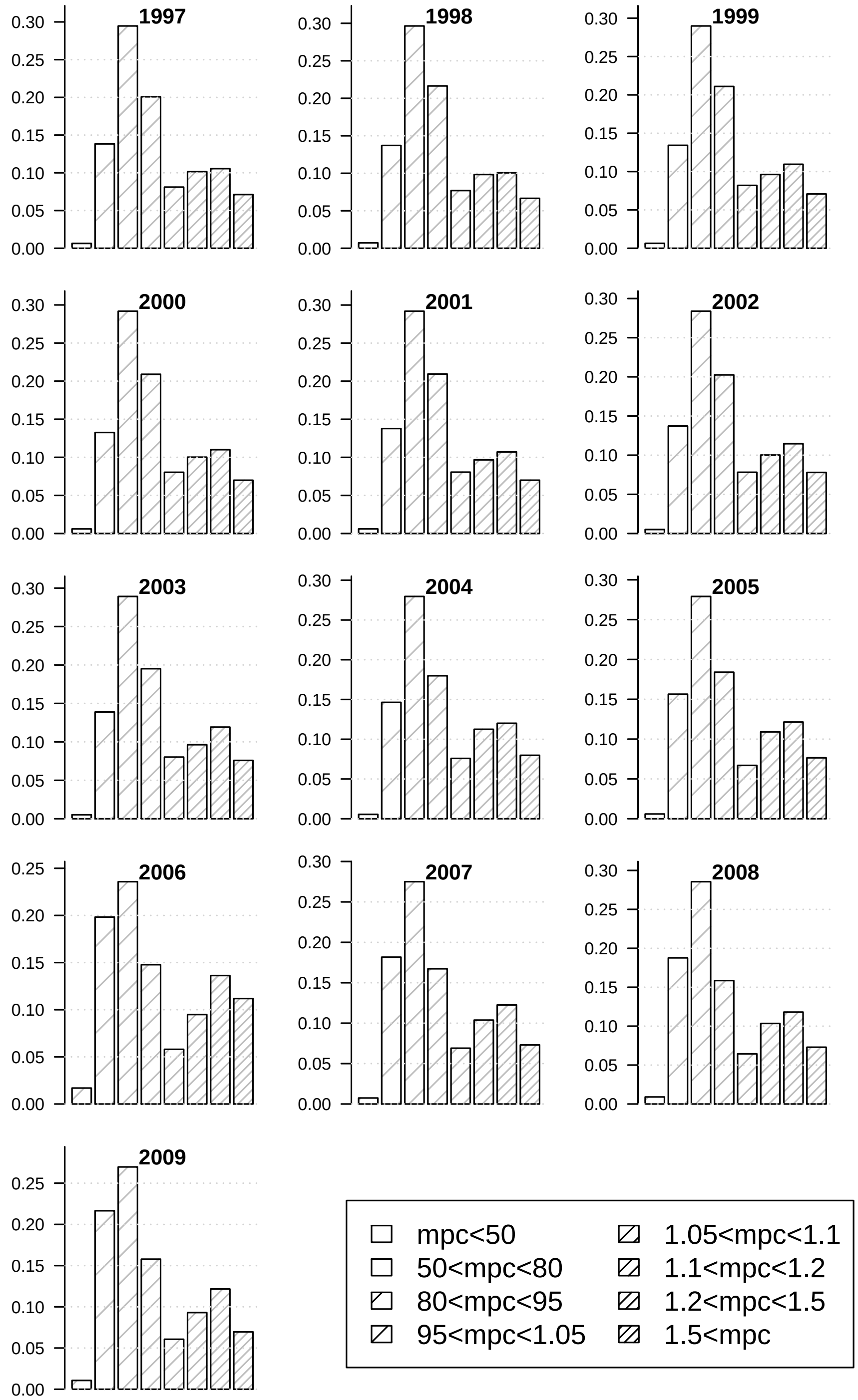

Figure 1.1: Distribution of Households by Propensity to Spend.

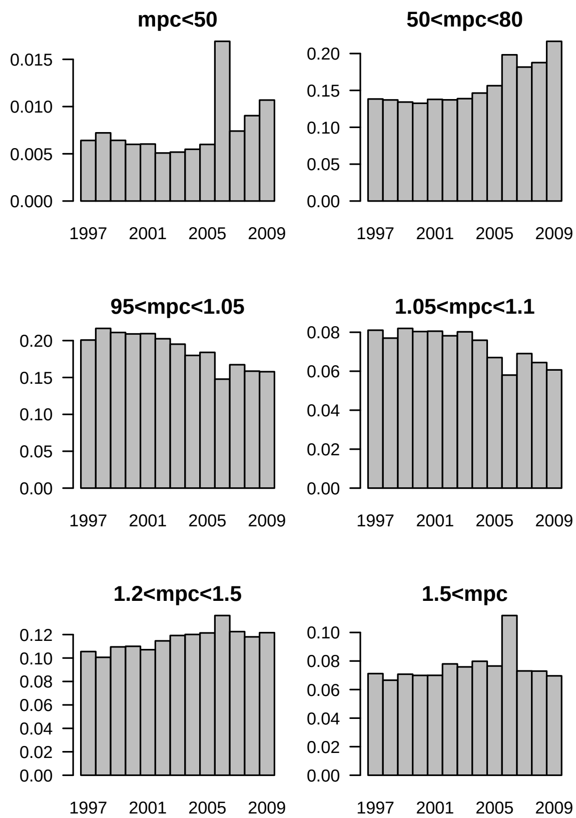

Figure 1.2: Trends in the proportion of different spending groups in samples over time.

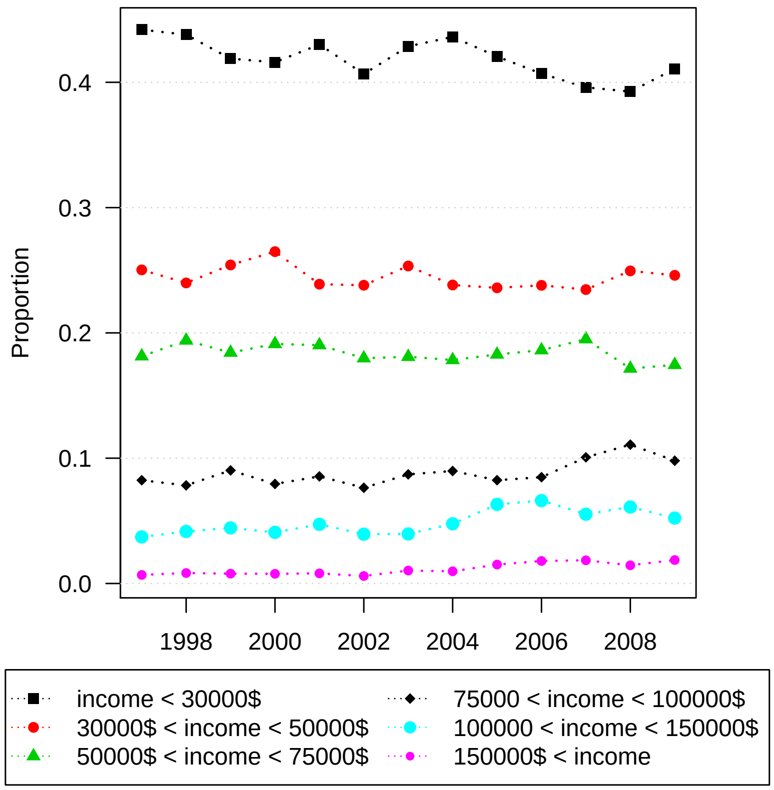

Figure 1.3: Income Distribution of Spender Households - MPC>105%

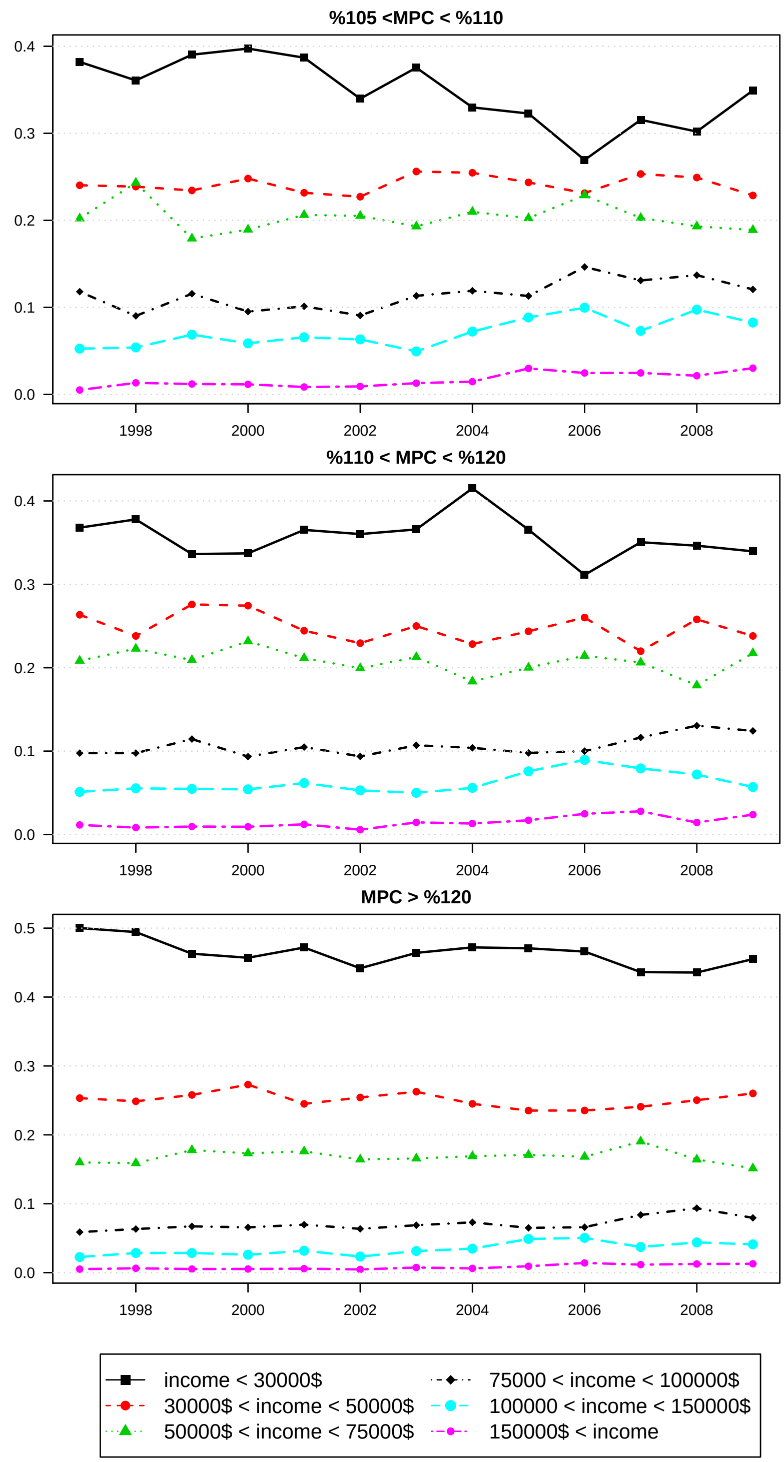

Figure 1.4: Income Distribution of Different Spending Groups

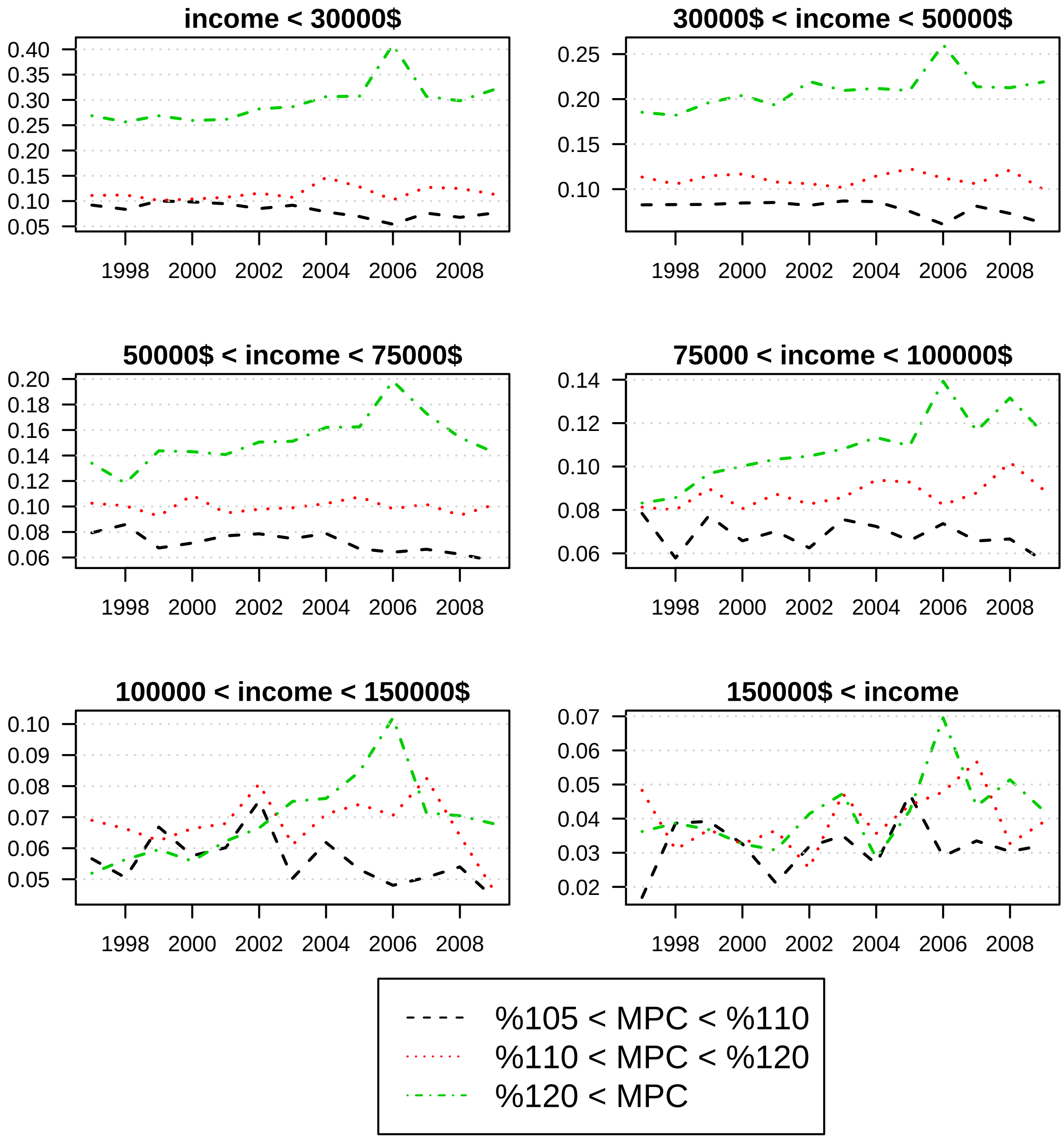

Figure 1.5: Proportion of Spender Groups at Different Income Levels

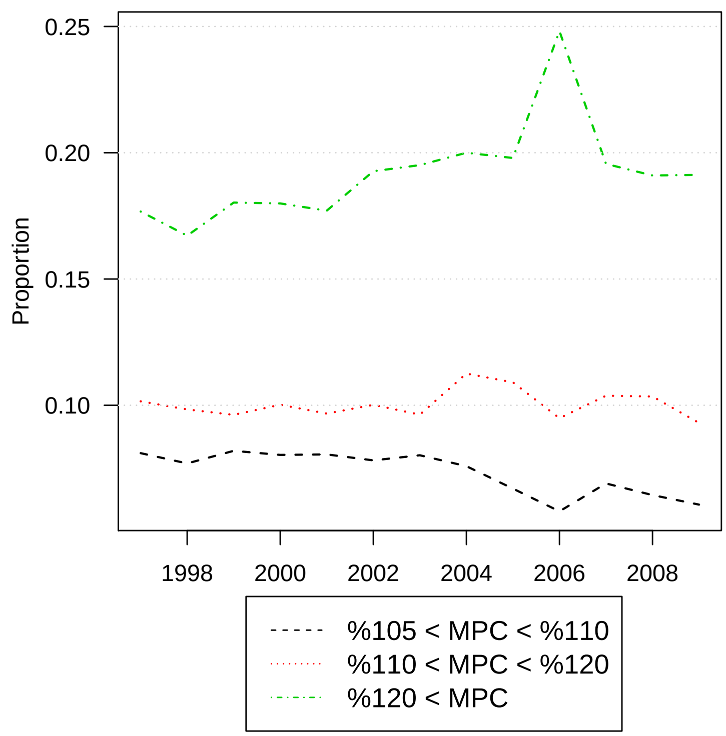

Figure 1.6: Proportion of Spender Groups for All Income Levels

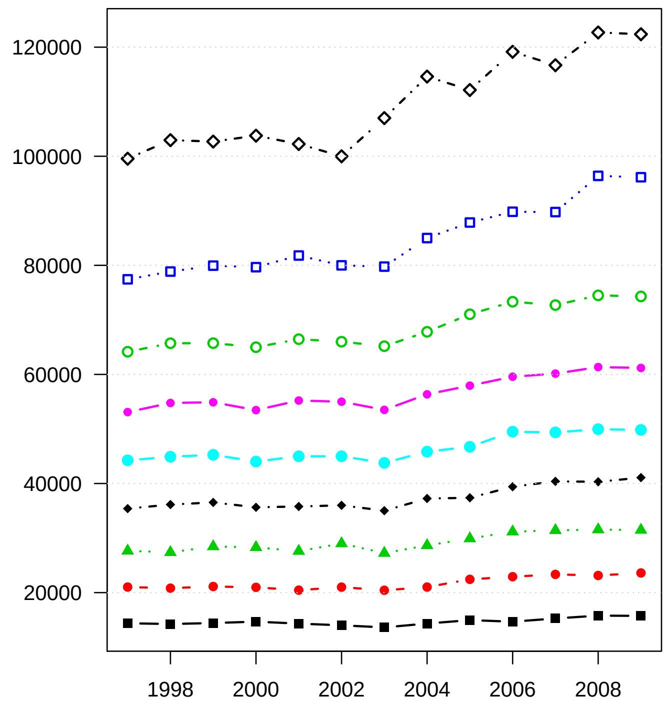

Figure 1.7: Trends in Deciles of Income

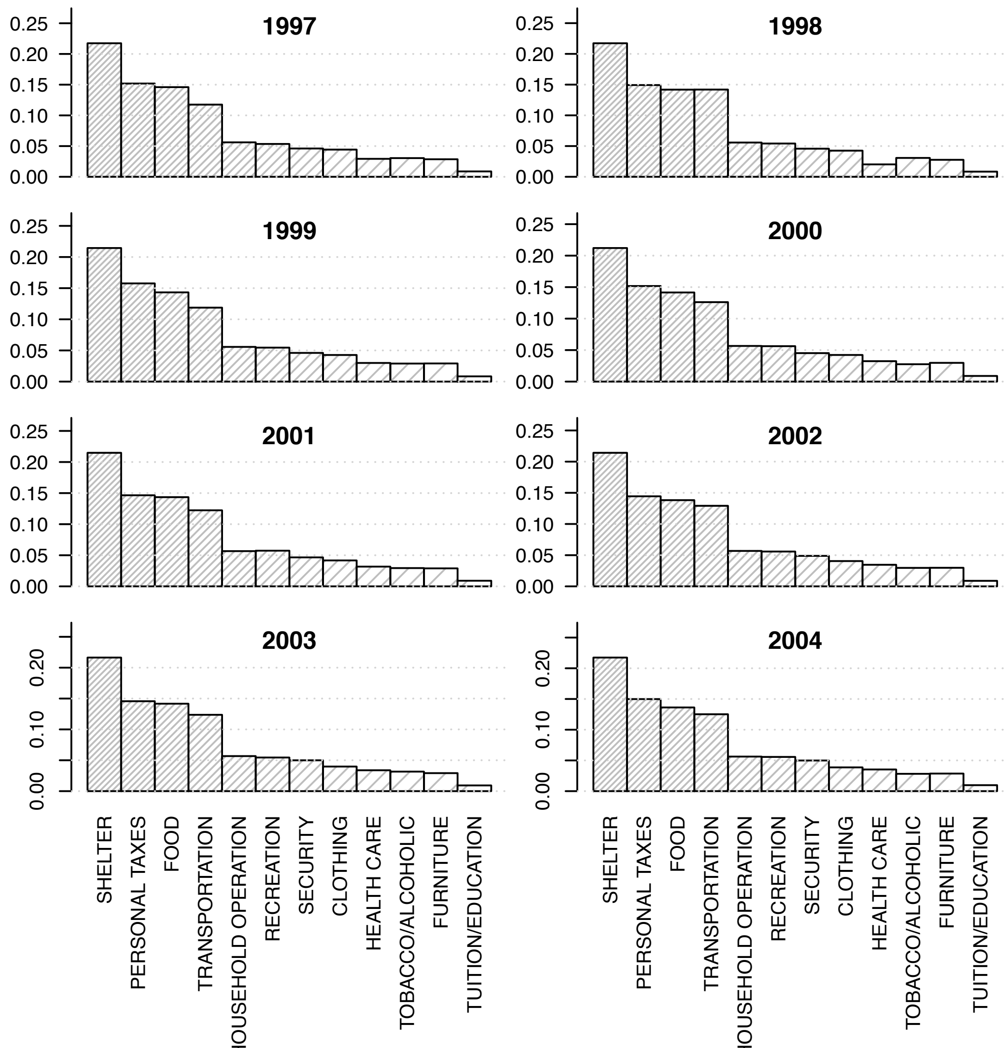

Figure 1.8: Expenditure Share of Specific Items in Households Budget 1997-2004

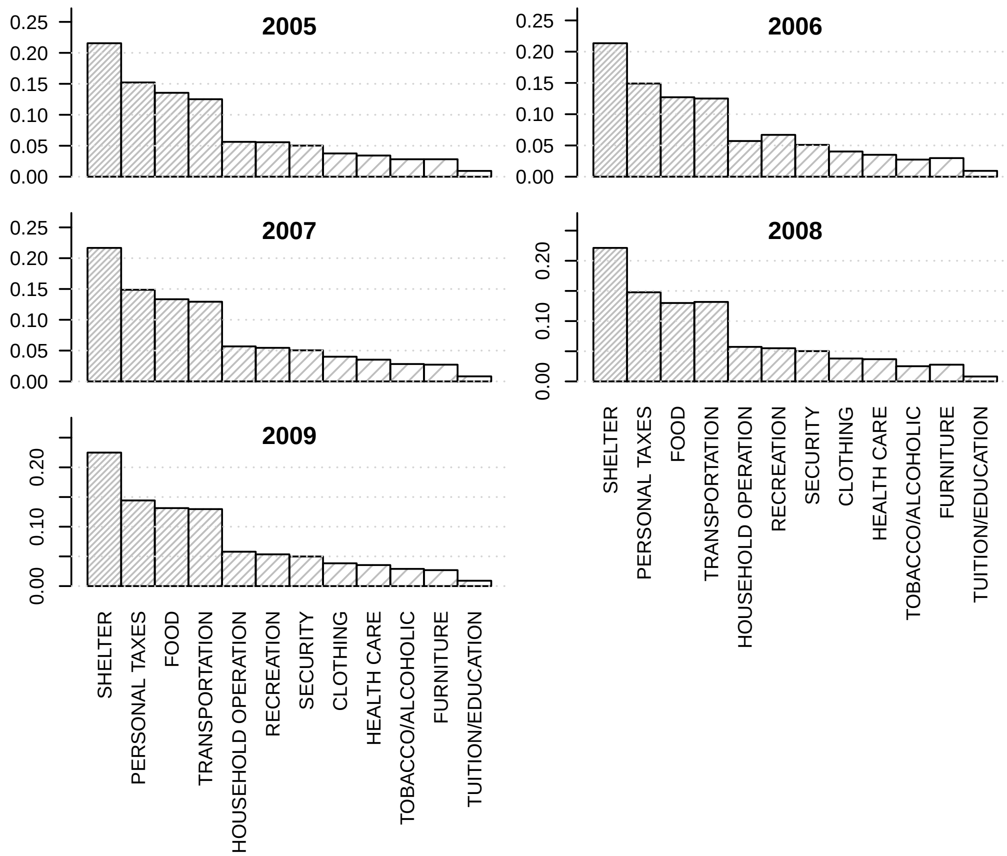

Figure 1.9: Expenditure Share of Specific Items in Households Budget 2005-2009

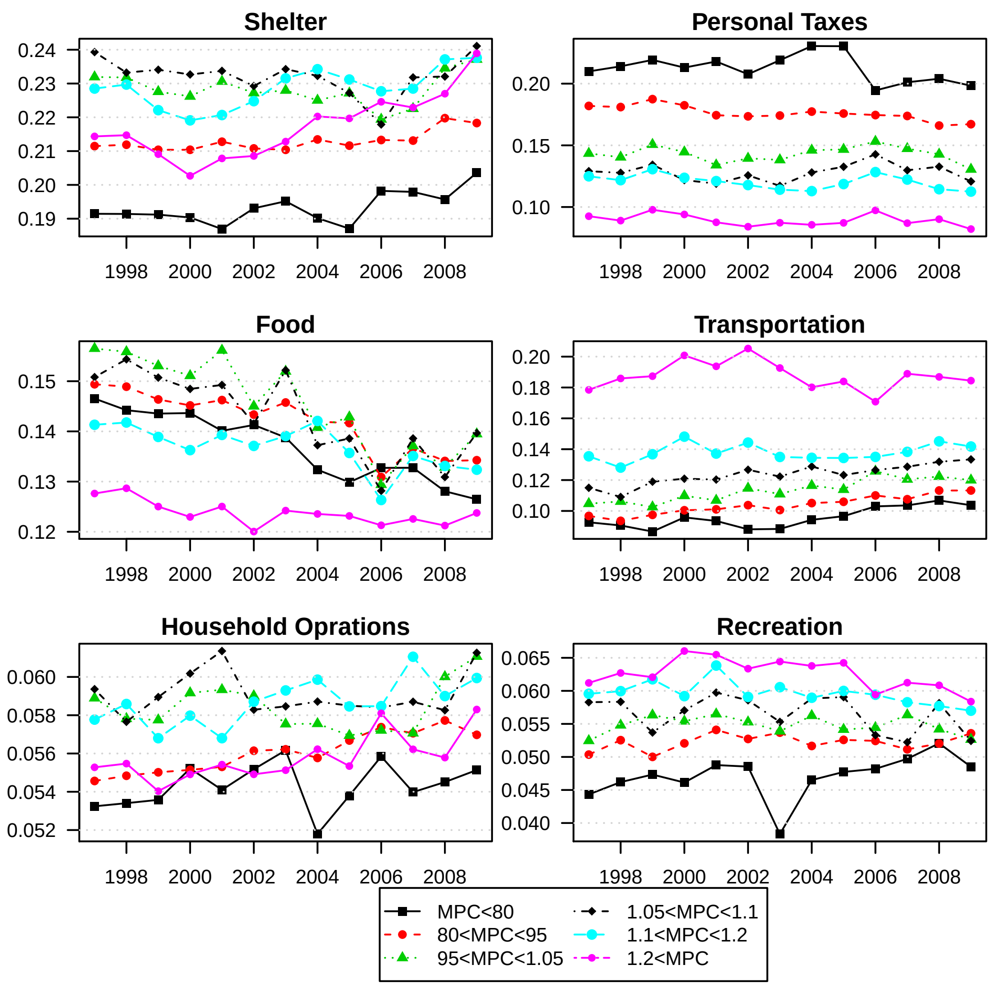

Figure 1.10: Expenditure Share of Specific Items in Households Budget

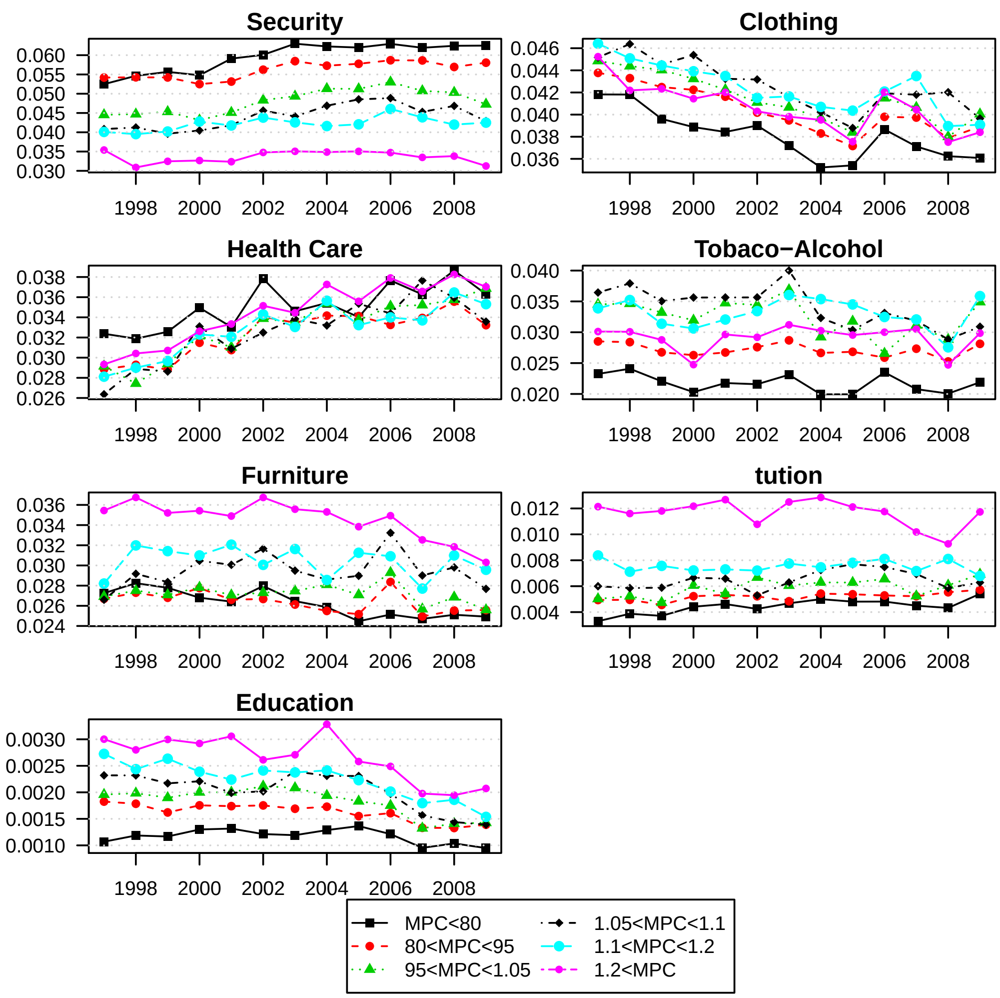

Figure 1.11: Expenditure Share of Specific Items in Households Budget

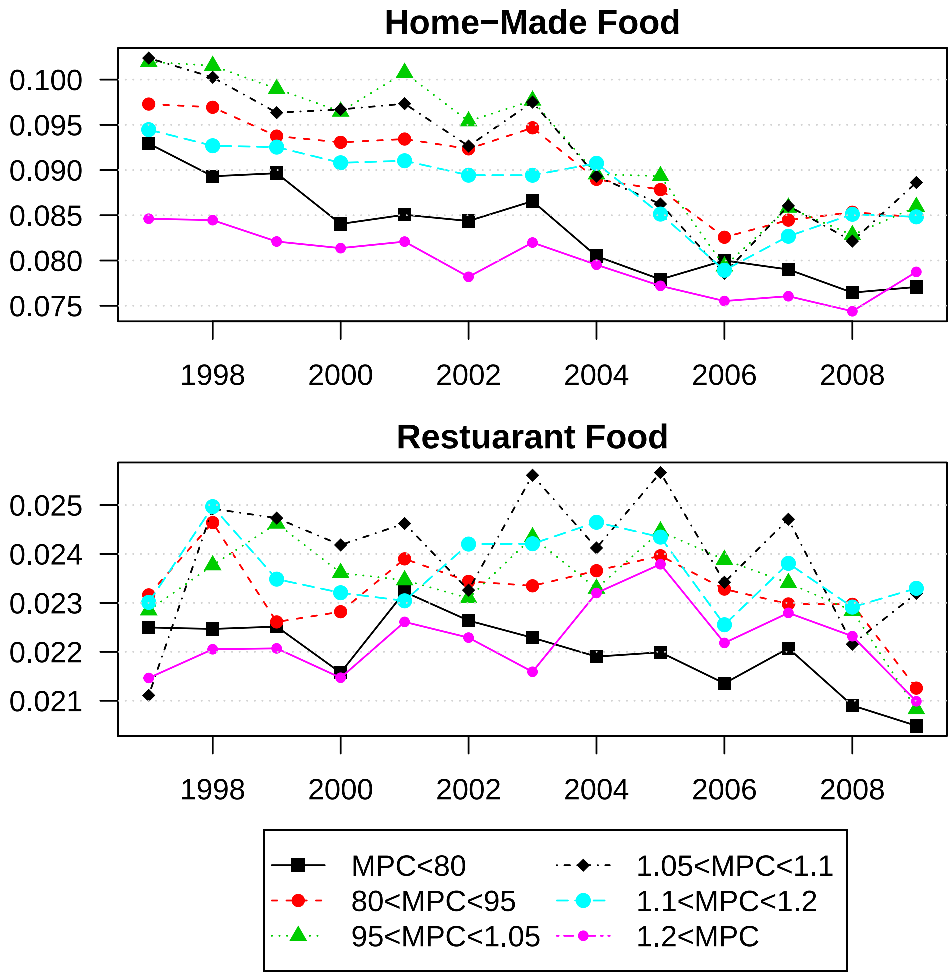

Figure 1.12: Expenditure Share of Specific Items in Households Budget

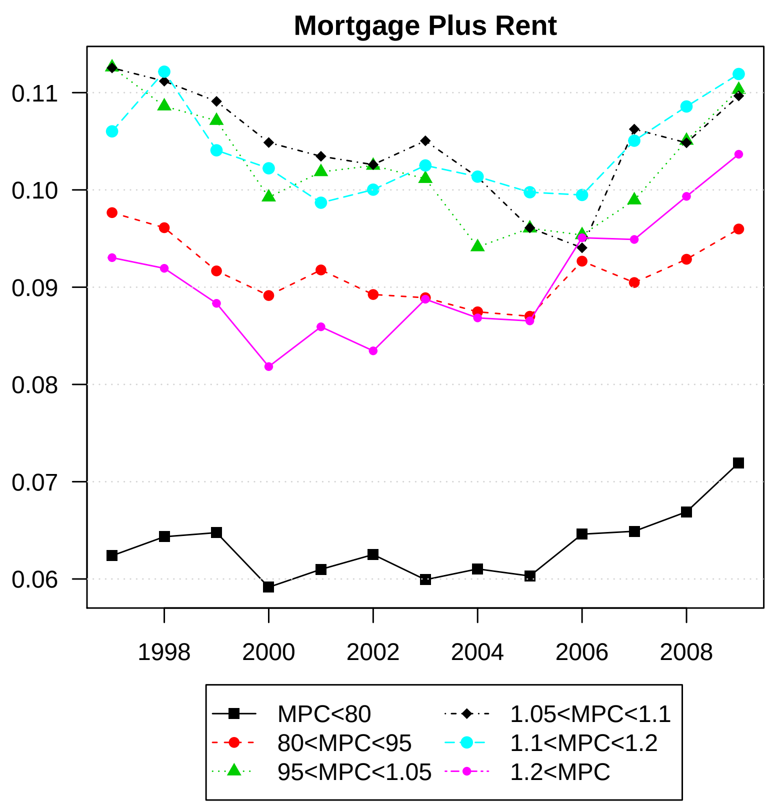

Figure 1.13: Expenditure Share of Mortgage plus Rent in total Shelter Expenditures

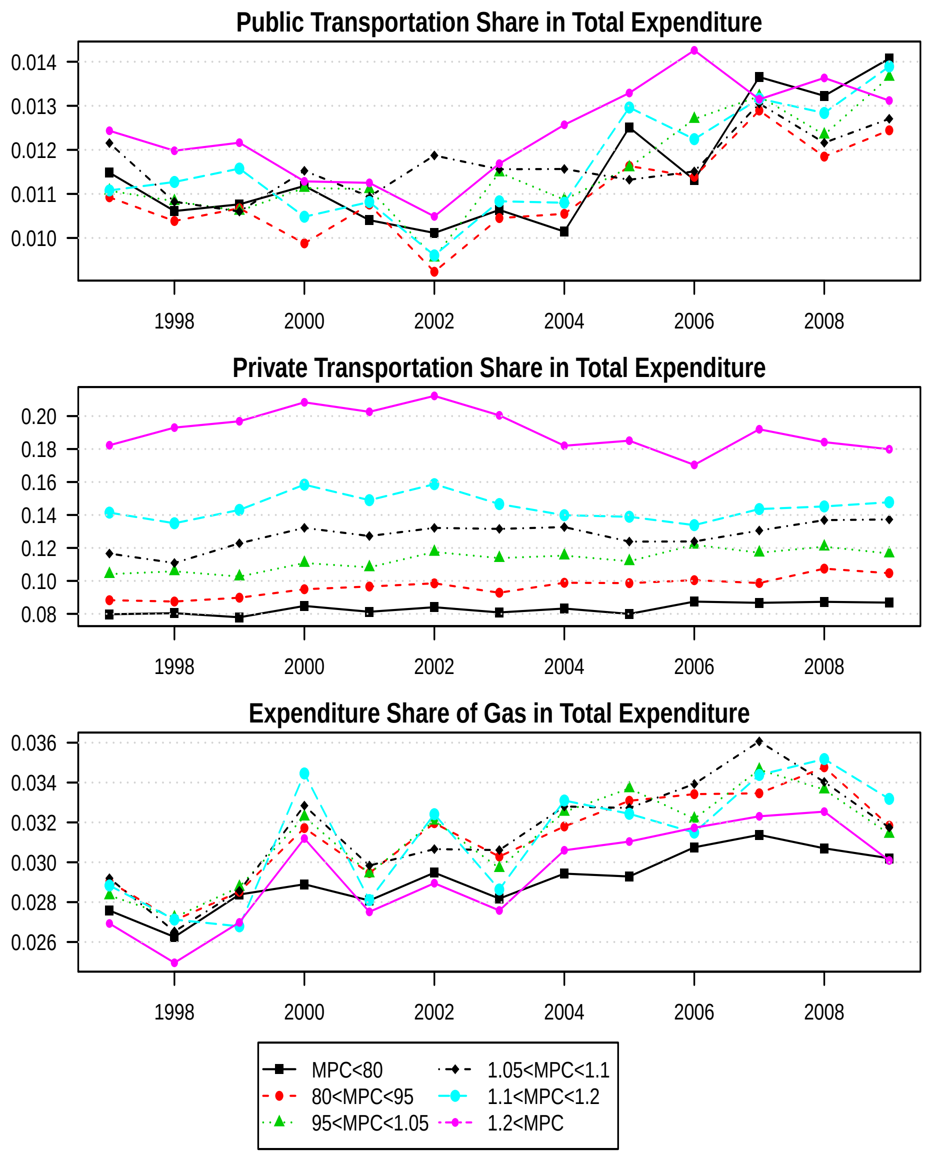

Figure 1.14: Expenditure Share of specific items for Households with Different Propensity to Consume

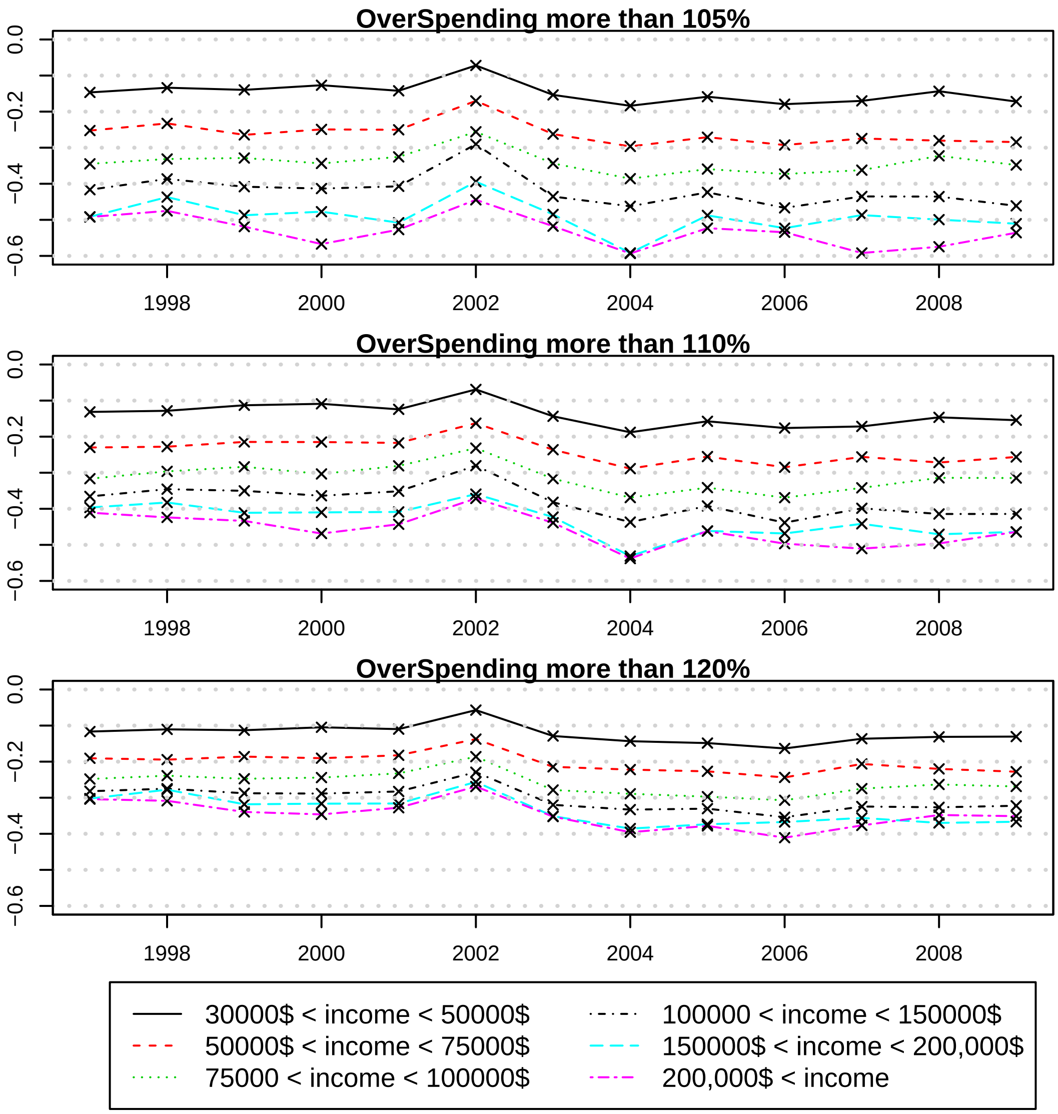

Figure 1.15: The effect of being a specific income group on the probability of spending more than income, in comparison with households with less than 30,000$ income. × signs indicate if the coefficient is significant at 5% confidence level.

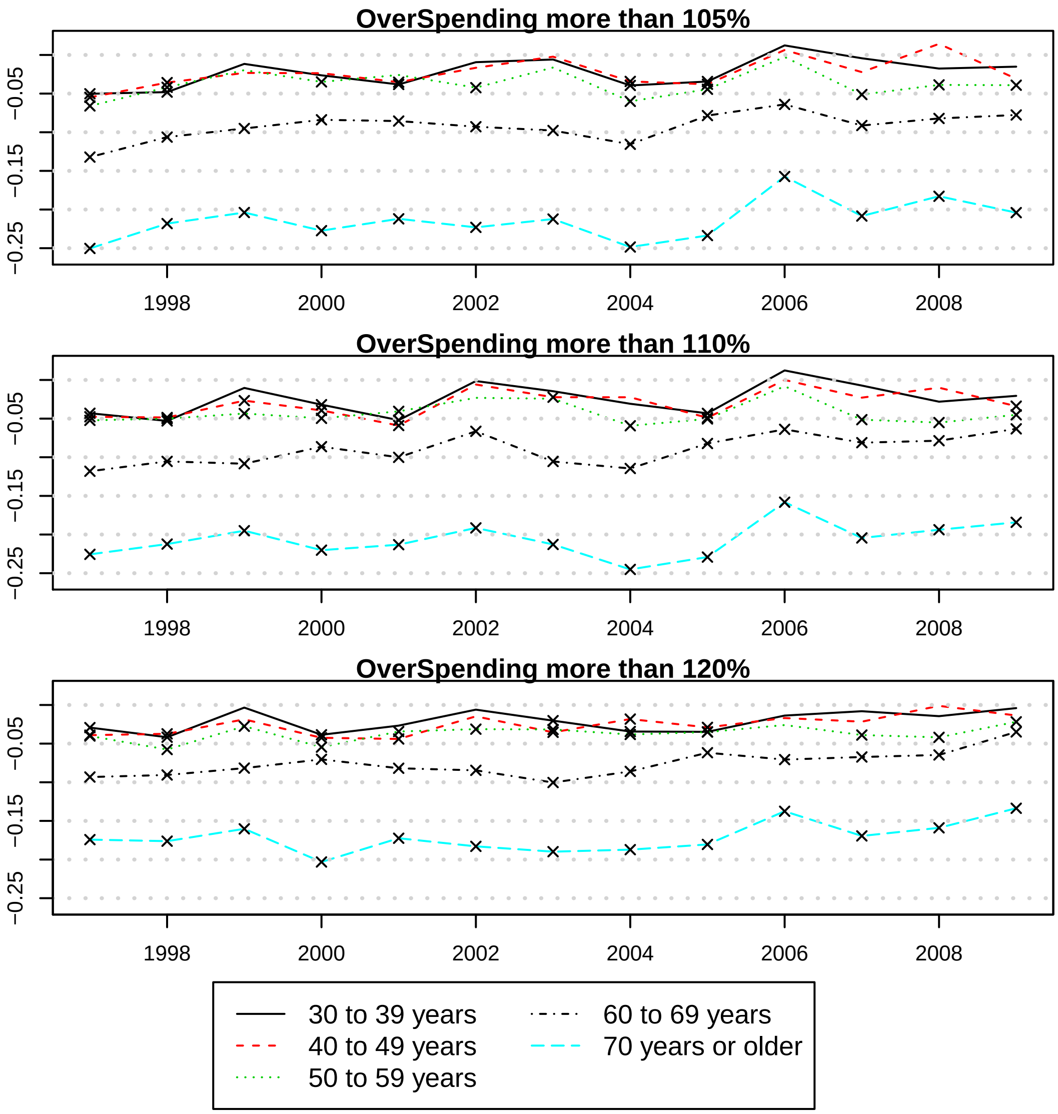

Figure 1.16: The effect of being a specific age group on the probability of spending more than income, in comparison with households with head younger than 30 years old. × signs indicate if the coefficient is significant at 5% confidence level.

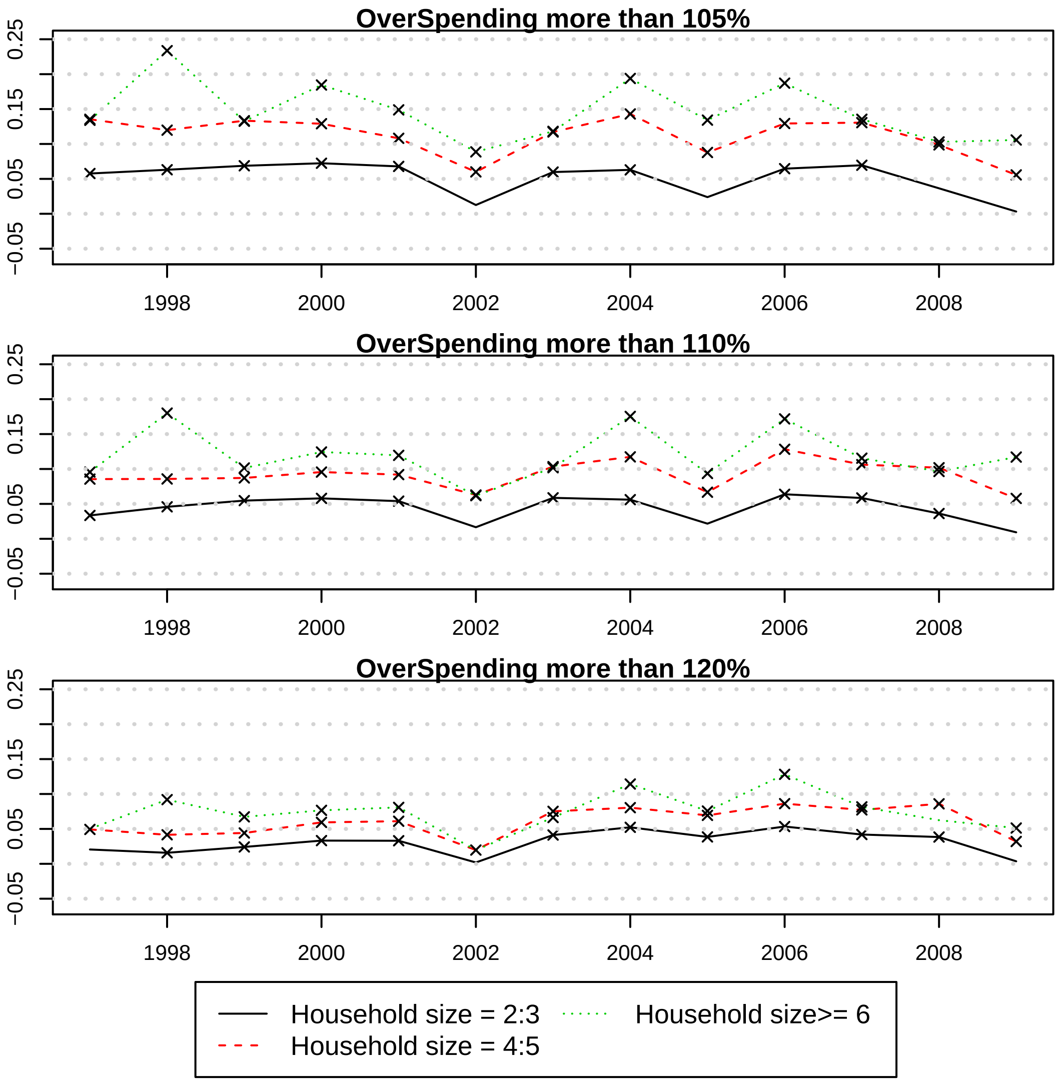

Figure 1.17: The effect of household size on the probability of spending more than income. × signs indicate if the coefficient is significant at 5% confidence level.

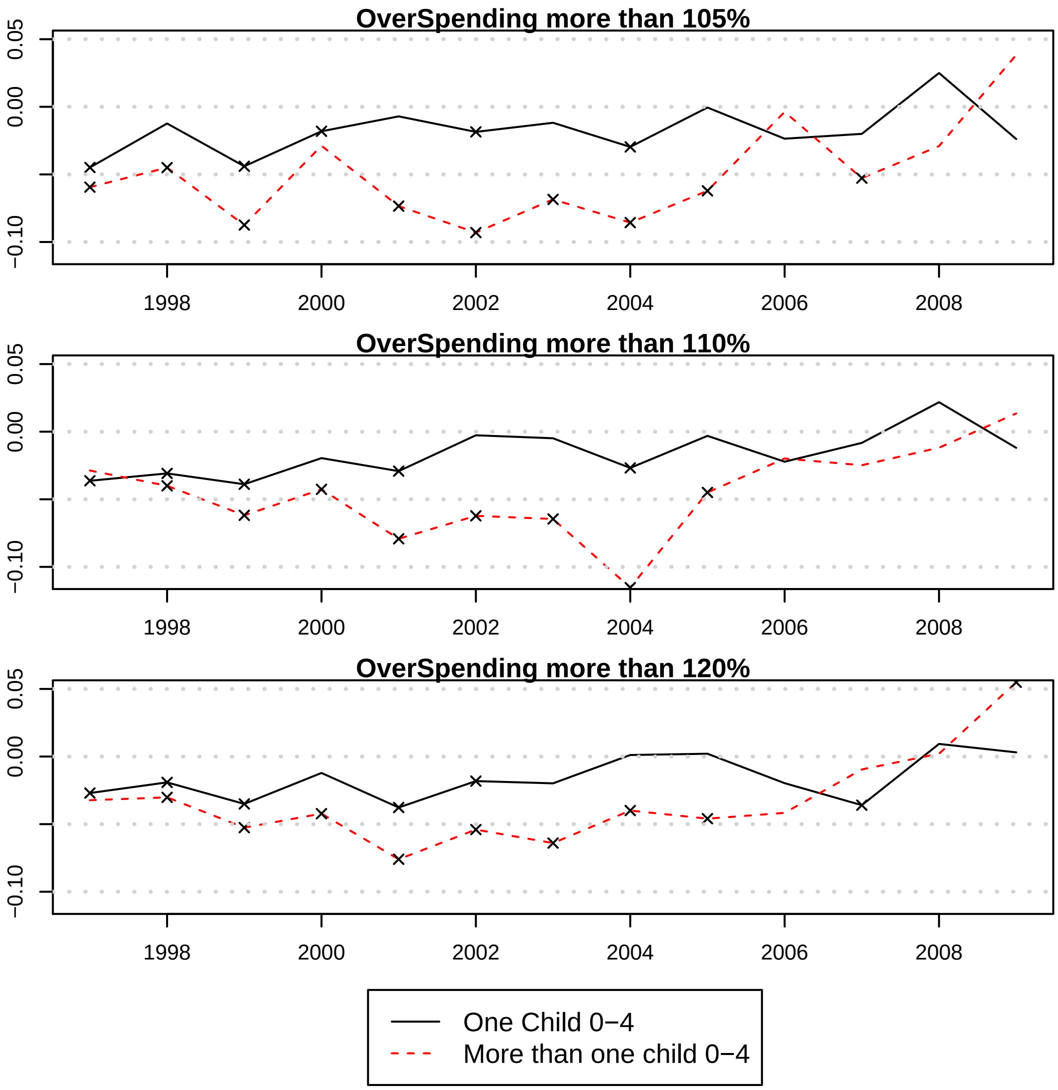

Figure 1.18: The effect of number of children younger than 4 years old on the probability of spending more than income, in comparison with households without children in this age range. × signs indicate if the coefficient is significant at 5% confidence level.

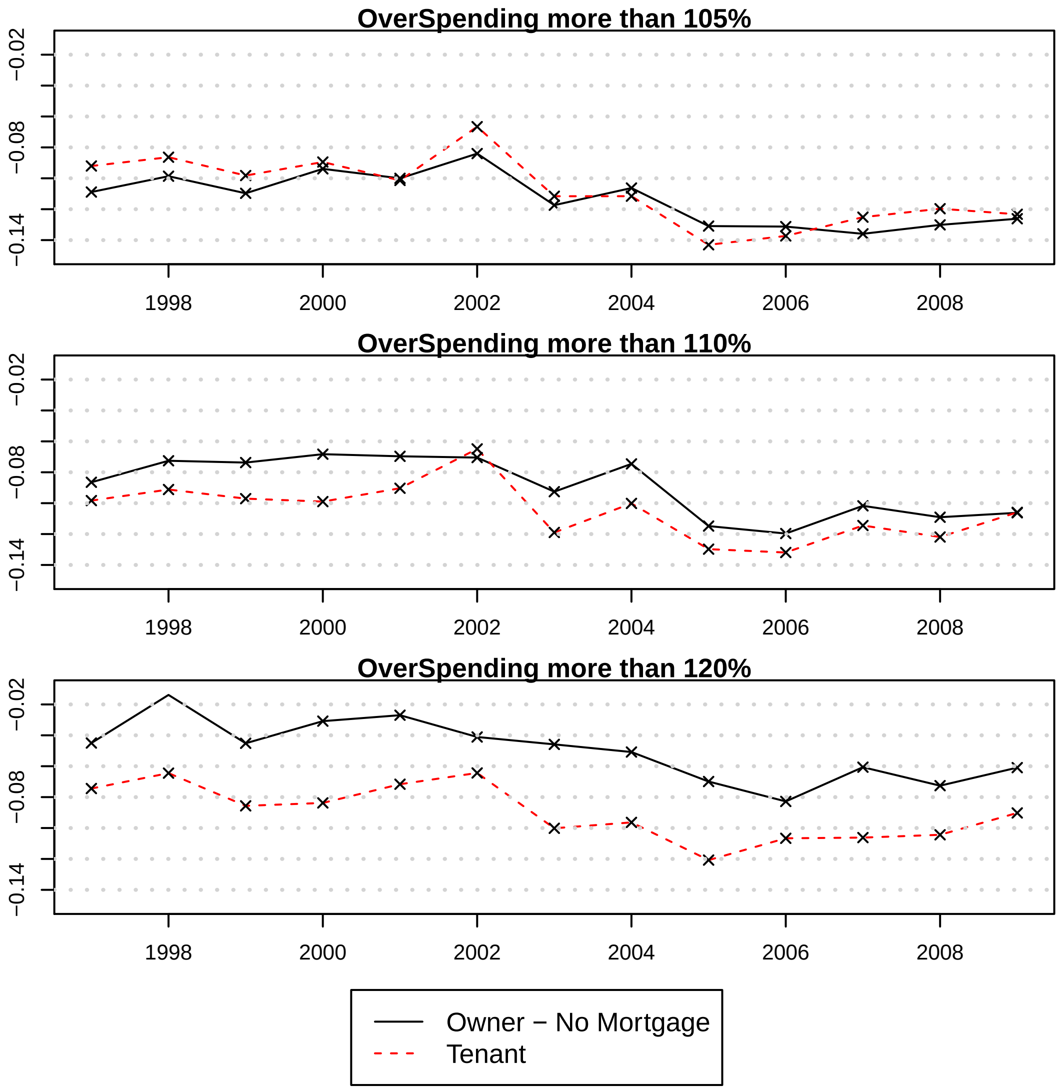

Figure 1.19: The effect of dwelling status on the probability of spending more than income, in comparison with households who are holding mortgage debt.

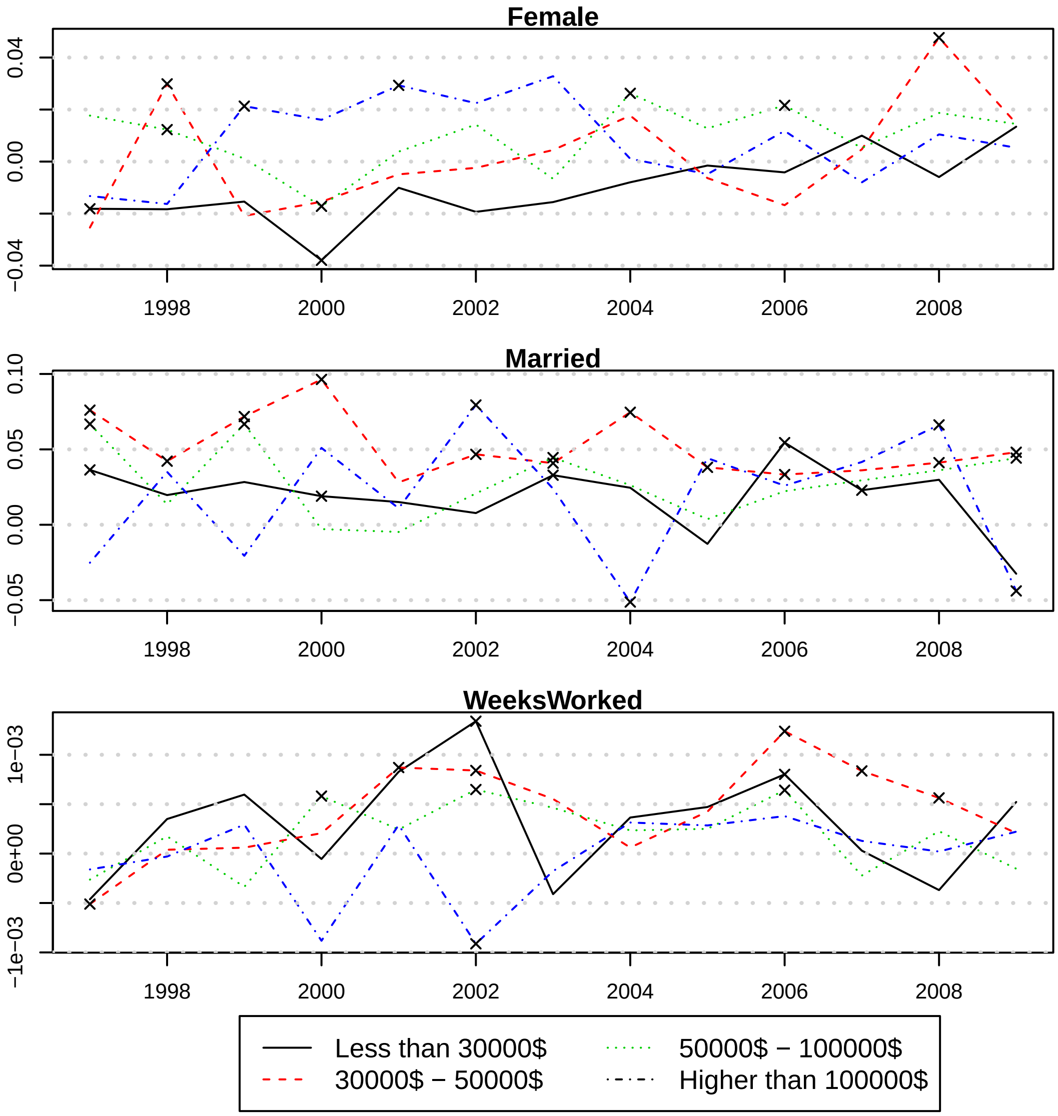

Figure 1.20: Top Panel: The effect of head of household being female on spending more than income by 5% or more, in different income groups. Middle Panel: The effect of head of household being married on the probability of spending more than income by 5% or more in different income groups. Bottom panel: The effect of number of weeks worked full-time and part-time on spending more than income by 5% or more in different income groups. × signs show if the coefficient is significant at 5% level.

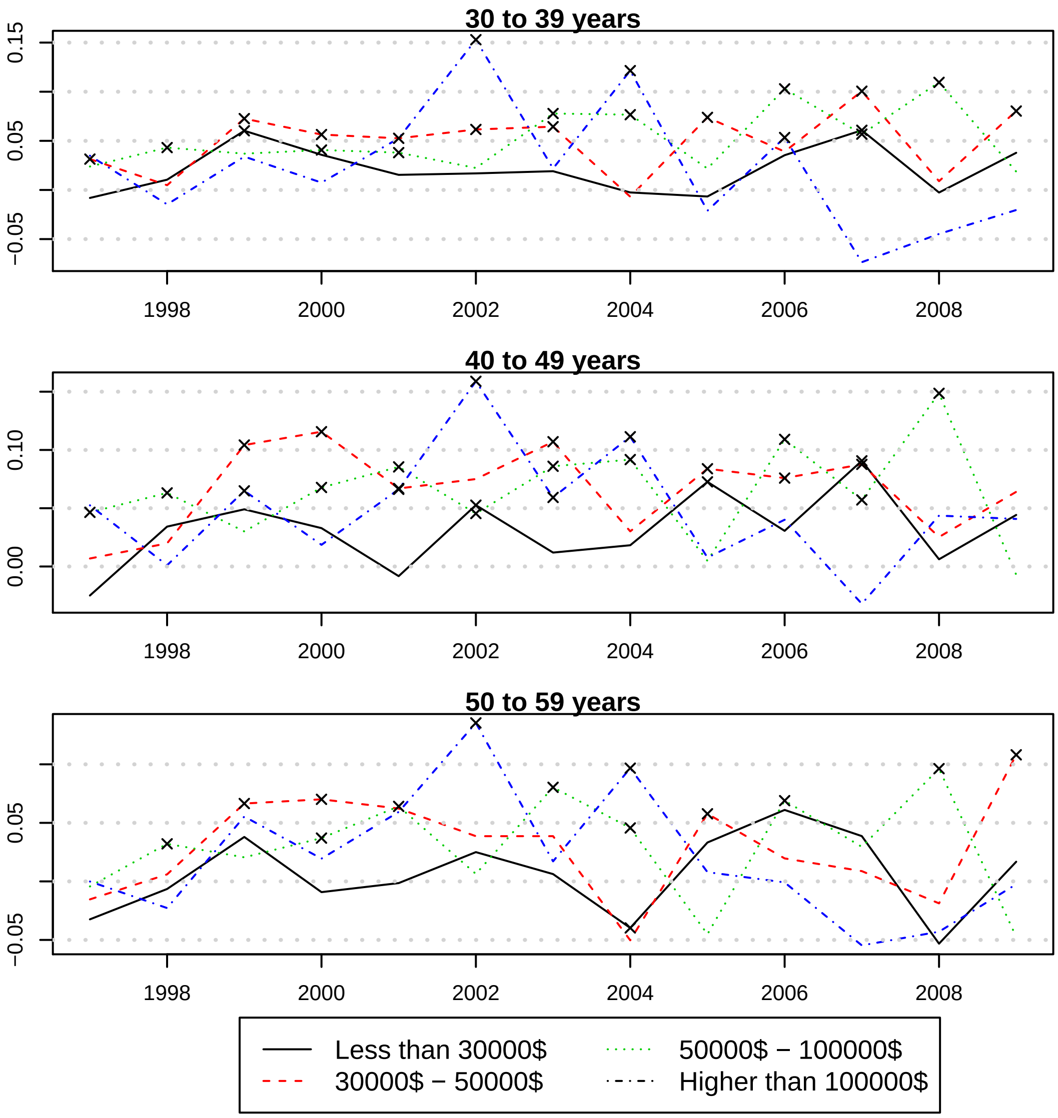

Figure 1.21: The effect of being different age groups on the probability of spending more than income by 5% or more, in comparison to households whose heads are younger than 30 years old. × signs show if the coefficient is significant at 5% level.

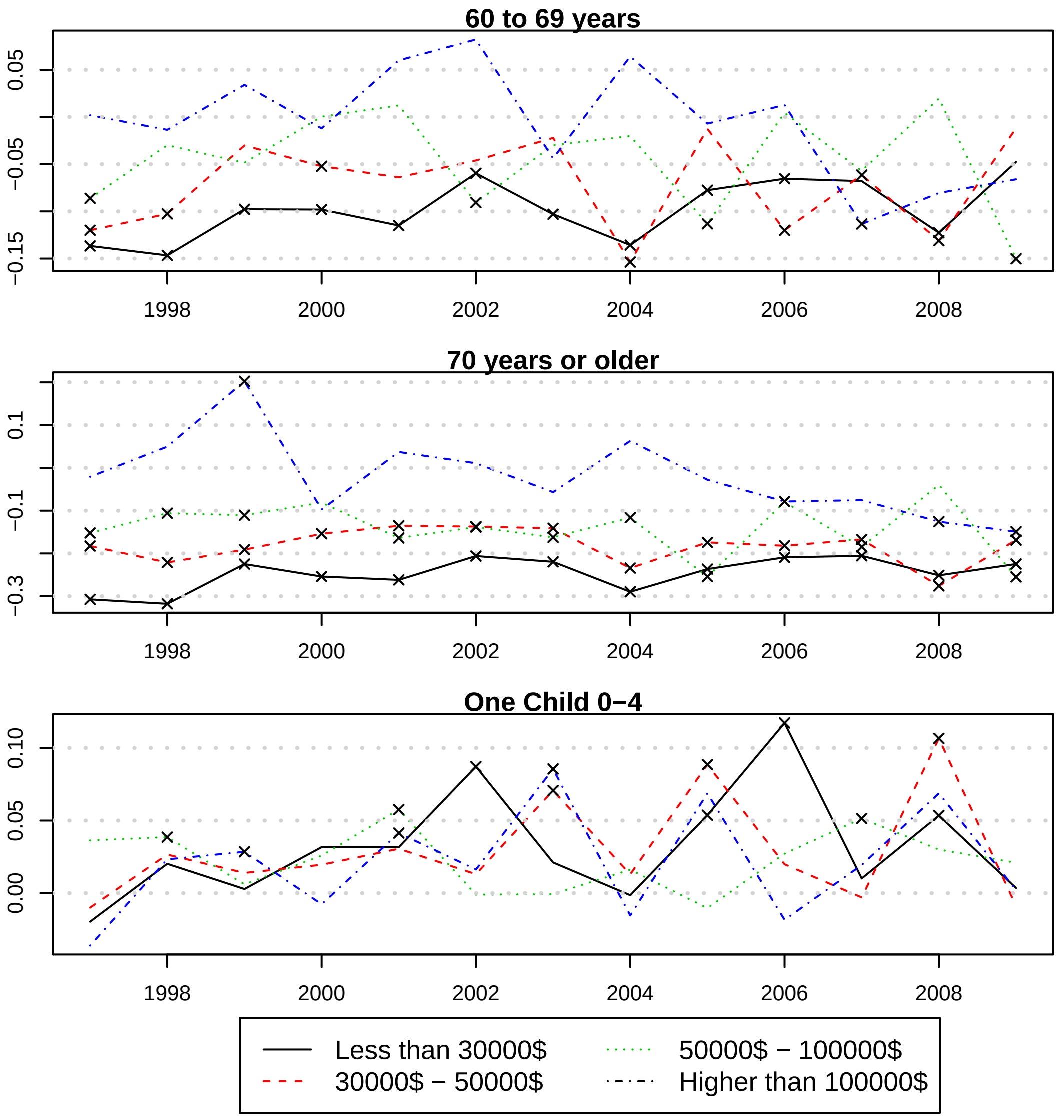

Figure 1.22: Top Two Panels: The effect of being different age groups on the probability of spending more than income by 5% or more, in comparison to households whose heads are younger than 30 years old. Bottom Panel: The effect of having a child less than four years old, on spending more than income by more than 5% in comparison with households without any child in that age range.

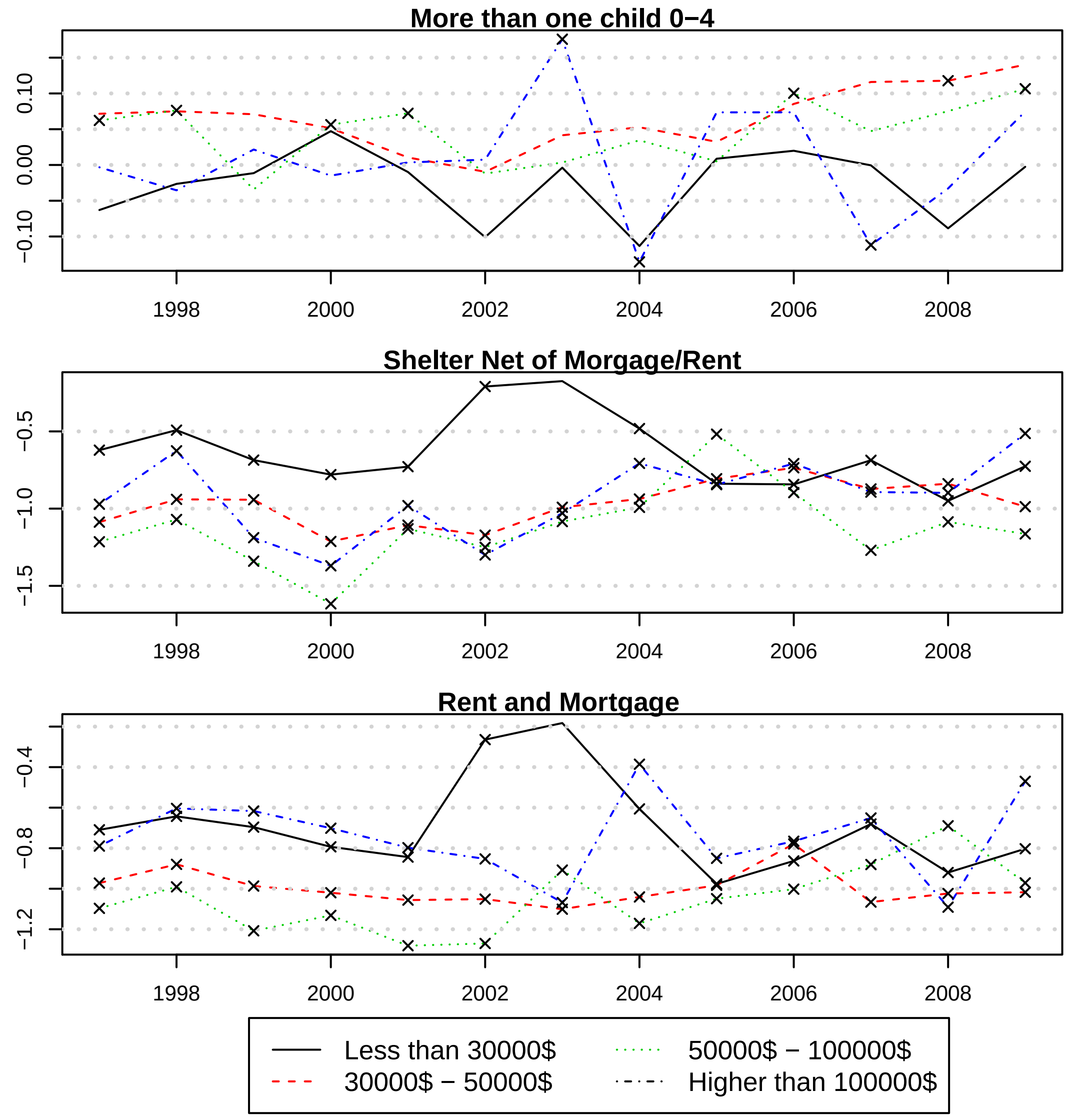

Figure 1.23: Top Panel: The effect of having more than one child younger than four years old, on spending more than income by more than 5%, in comparison with households without any child in that age range. Bottom Two Panels: The effect of dwelling status on spending more than income, in comparison to household who hold mortgage debt.

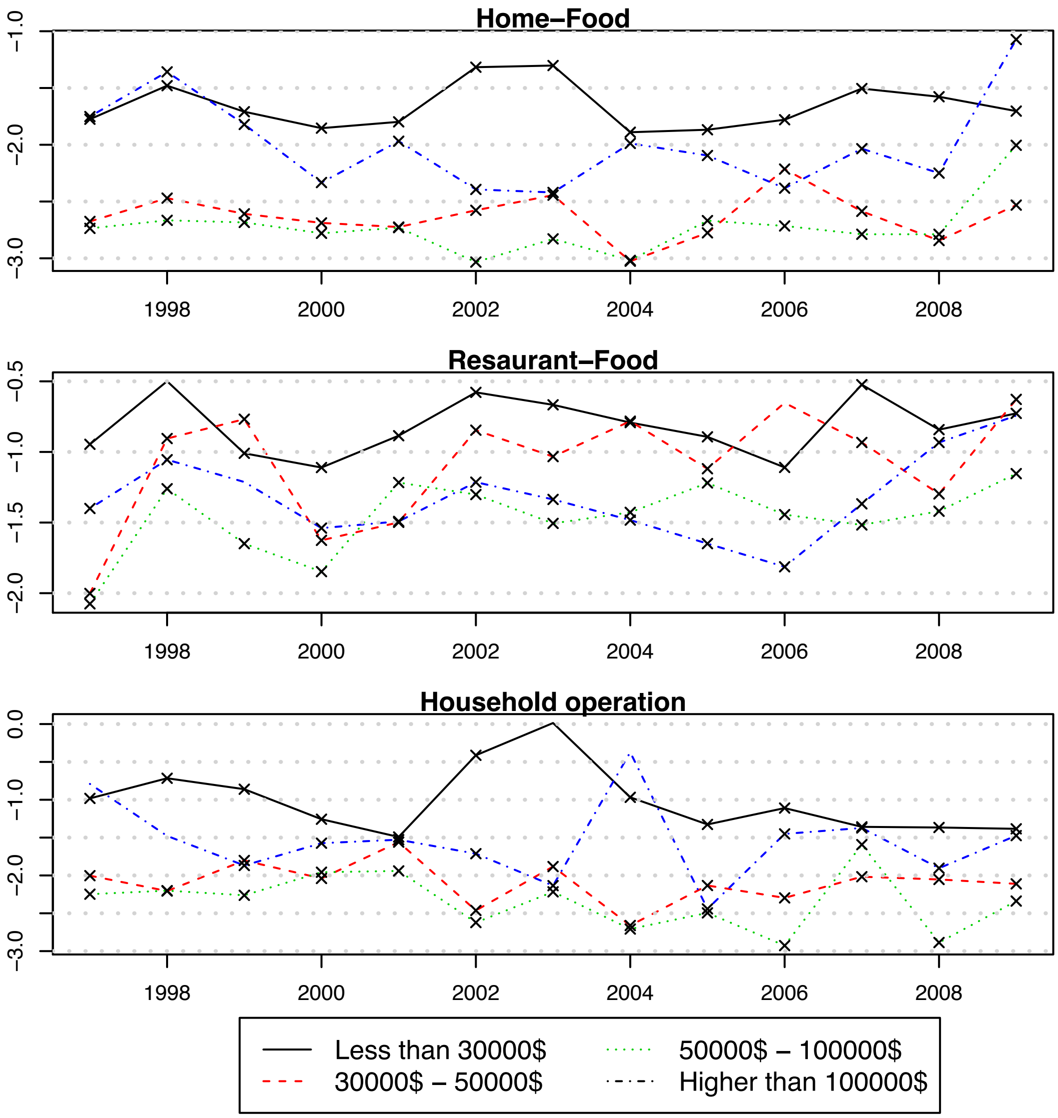

Figure 1.24: The effect of share of specific items in household budget in different income groups. × signs show if the coefficient is significant at 5% level.

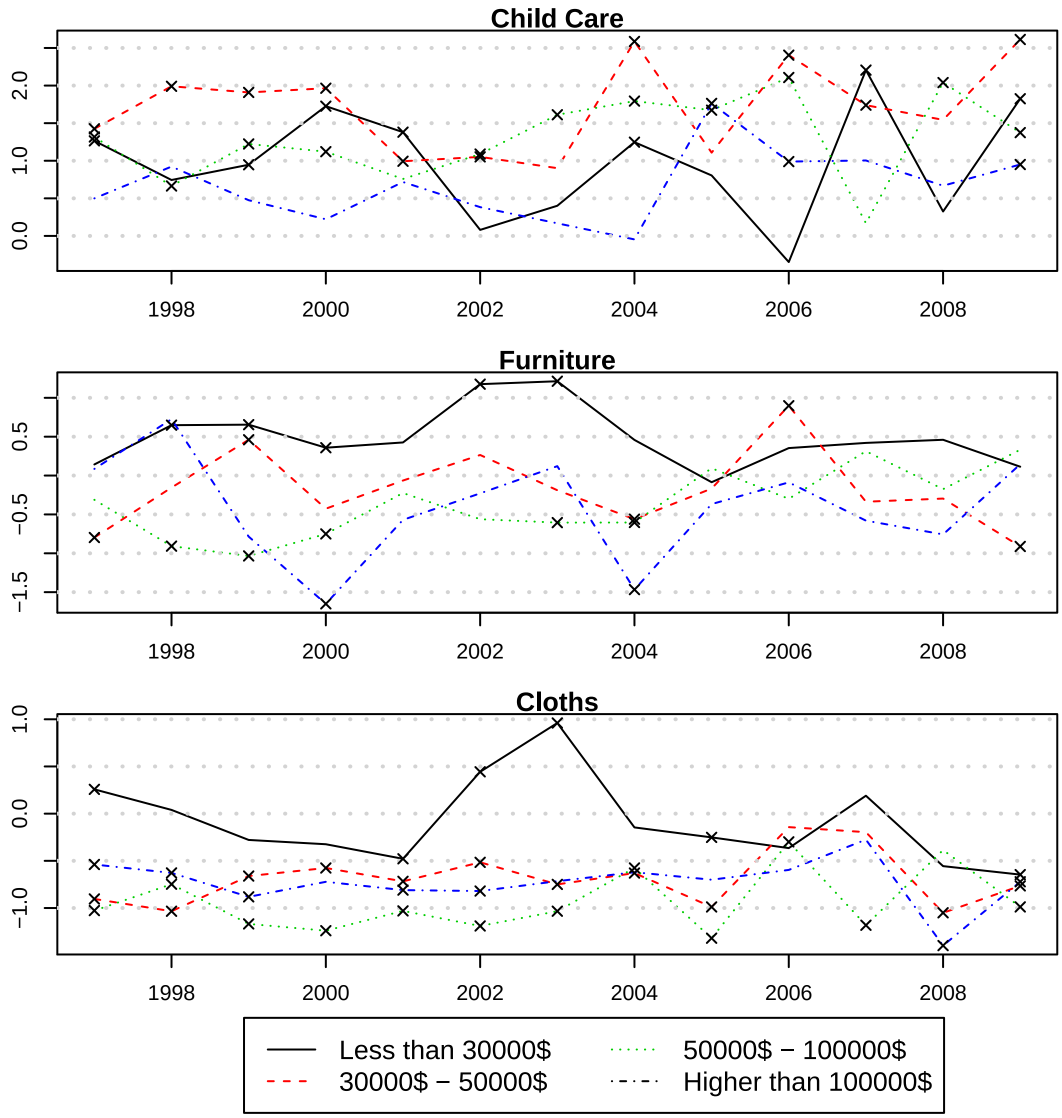

Figure 1.25: The effect of share of specific items in household budget in different income groups. × signs show if the coefficient is significant at 5% level.

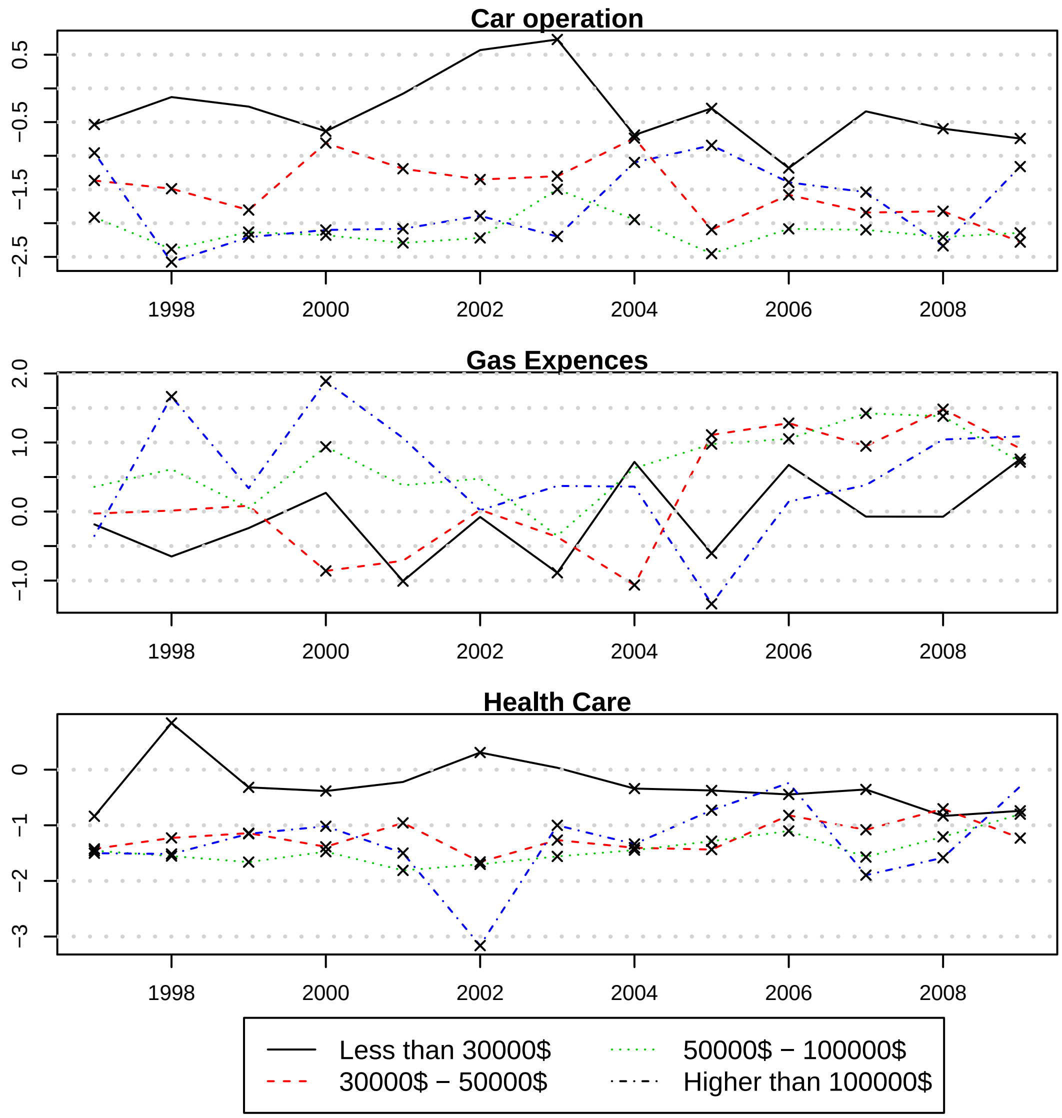

Figure 1.26: The effect of share of specific items in household budget in different income groups. × signs show if the coefficient is significant at 5% level.

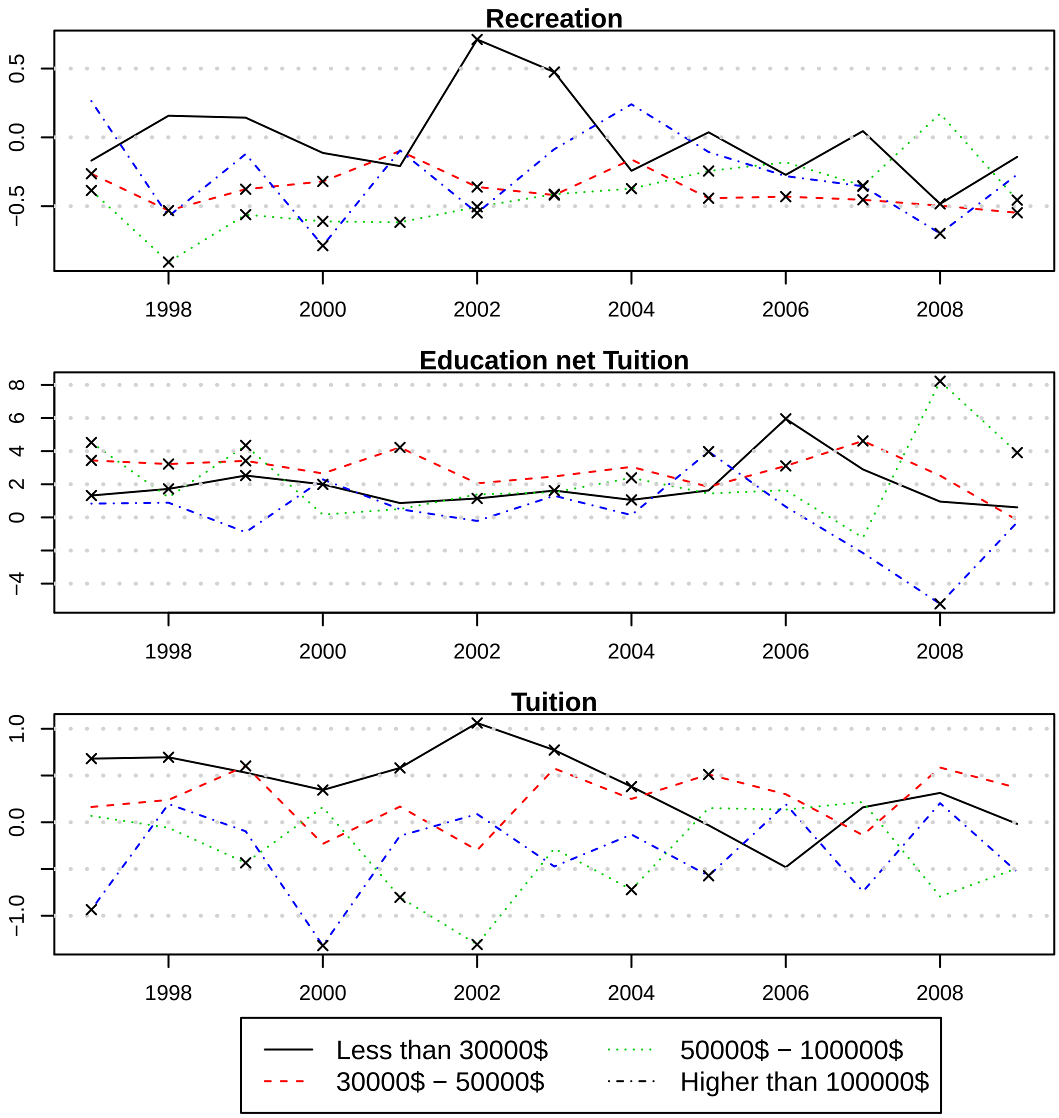

Figure 1.27: The effect of share of specific items in household budget in different income groups. × signs show if the coefficient is significant at 5% level.

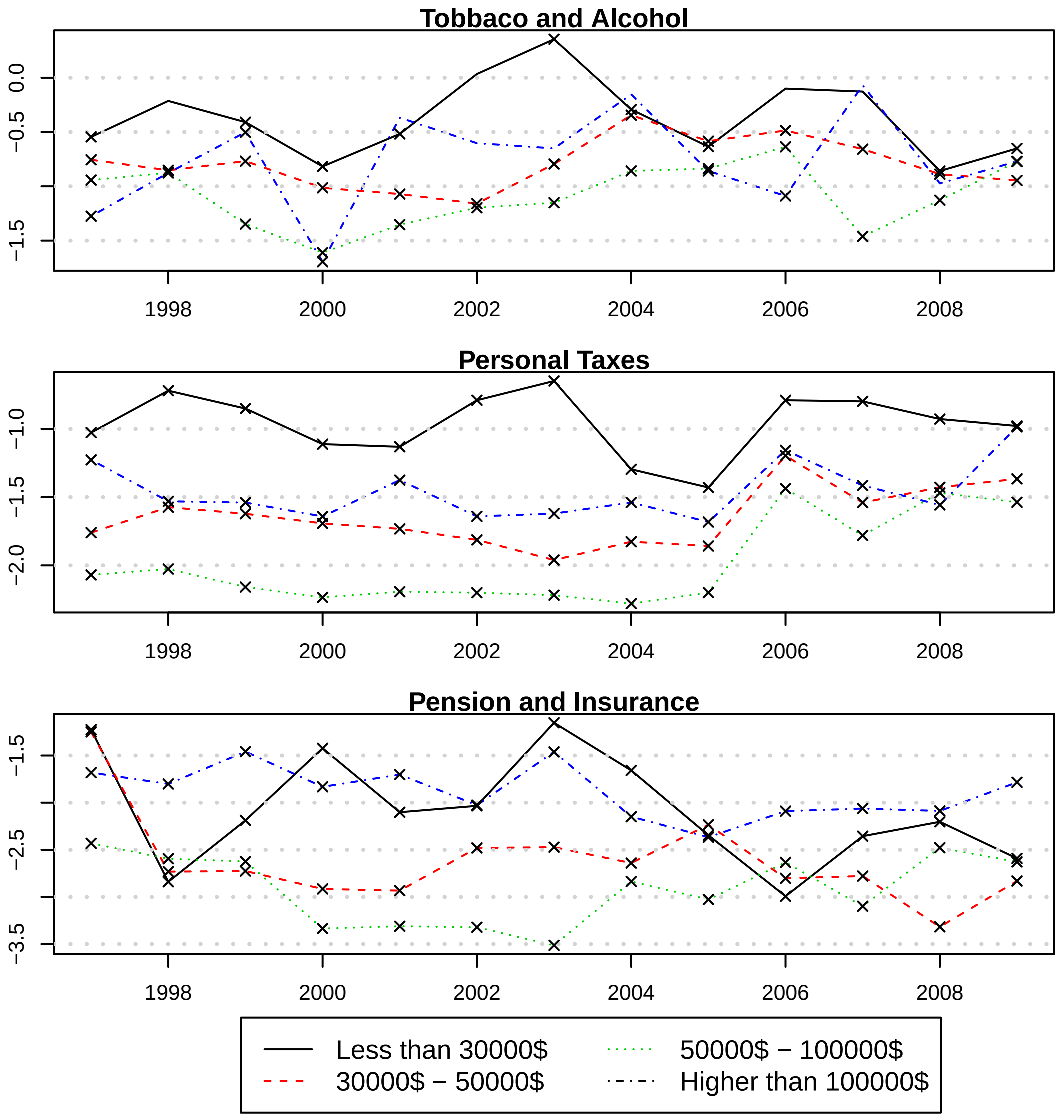

Figure 1.28: The effect of share of specific items in household budget in different income groups. $\times$ signs show if the coefficient is significant at 5% level.

Chapter 2

Do the Rich Save More? Revisiting an Old Question with New Approaches

2.1 Introduction

The purpose of this study is to shed more light on the long-debated question of whether more affluent households save a larger fraction of their income. In other words, the problem is whether the saving rate is flat across the income distribution or whether it increases at higher income brackets. Even though common sense suggests that more-affluent families save a larger fraction of their income, compared to middle-class and low-income households, historical evidence does not support this proposition.

Looking at aggregate data, there is no significant correlation between per capita income and the aggregate saving rate in time-series or cross-sectional datasets. In spite of ever increasing per capita income, the aggregate saving rate was fairly constant until the early 1980s and then continued to fall in the 1990s and 2000s. A comparison of the cross-section of countries also fails to suggest any meaningful relationship between per capita income and the aggregate saving rate.

Also, as income inequality rises in an economy, higher income households hold a larger share of total income. If these households at the top of the income distribution do save a larger fraction of their earnings, a higher aggregate saving rate should have been observed in the past few decades. That has not been the case in any of the advanced economies.

The empirical issue has remained cloudy, and there has not been a firm conclusion on the matter since the beginning of the debate in the 1950s. However, the evolution of macroeconomic models toward "representative agent" type models that assume a constant saving rate for heterogeneous people shows the dominance of one side in the development of economic models.

Policy-wise, the question bears a great deal of importance for several reasons. Optimality of the tax system greatly depends on the answer to the question of whether the propensity to save/consume varies over the income distribution, or whether high-income households save proportionately the same as middle-class households. The effectiveness of shifting tax policy towards consumption and income from saving accounts depends highly on the relative amount of saving in various income groups. If the saving rate does increase with long-run income, redistributing income from rich to poor within an age group would be more efficient than redistributing from young to old households within income groups.[1]

Moreover, finding heterogeneous trends in the decline of household saving rates, that have resulted in the decline of the aggregate household saving rate from the early 1980s to recent years, can help us understand the responsiveness of aggregate consumption and saving to different policy reforms and economic shocks. Different saving patterns among households also show relative degrees of preparedness for the post-retirement period.

To empirically answer the question of whether the rich save more, several challenges must first be tackled, making this issue even more important in the empirical literature. Both variables of interest are prone to suffer from measurement error. While measurement error in the dependent variable, which is the saving rate, only reduces the efficiency of estimates, a polluted independent variable, which is income, causes biased estimates and inconsistent estimates.

Depending on the area of interest, different definitions of saving might be used. They range from the most general definitions that include financial assets, owner-occupied businesses, realized and unrealized capital gains on housing, and other components of wealth, to a simple definition of saving that focuses on the difference between income, exclusive of capital gains, and current consumption. For the purpose of this particular question, two measures of saving have been used in the literature. The first is the change in household wealth that is not always available in all micro-level datasets. The second is the household disposable income minus its

[1] See for example Burbidge and Davies (1994)

consumption expenditures, excluding durables, which is more obtainable from survey data. Fortunately, both measures are available in the FAMEX and SHS datasets. We focus on the latter definition and use the former for sensitivity analysis.

The more-important concern is the definition and measurement of long-run income. The current measure of income reported in surveys is a weak proxy for the affluence of households and has two sources of pollution, i) transitory components of the income process and, ii) measurement error in self-reported income. According to Friedman (1957), the two are indissolubly merged into the correct measure of income. As long as saving is defined as the difference between income and consumption, if we consider saving rate as the dependent variable on the left-hand side, the unobserved components described above in current income would appear on both sides causing upward bias in estimates.

Restarting the long-standing debate after a period of quiescence, Dynan (2004) proposed a method to address the question and brought attention back to this issue. In their influential work, they proposed a simple two-step method to estimate the median saving rate at five different quintiles of income. Comparing five median saving rate at five different income brackets, they conclude that the saving rate increases in higher brackets of long-run income, under some circumstances even with increasing rate.

The approach proposed by Dynan (2004) is still widely used in the literature to address the question in different countries and varying data sets. Bozio and Dea (2011) use both administrative and survey data on British household wealth and consumption. Following Dynan's methodology, they find a positive relationship between long-run income and saving rates for households in the United Kingdom.

In another study, Hori et al. (2015) employ Dynan's methodology to evaluate the relationship in Japanese Income and Consumption Survey data. Their results are sensitive to the choice of instrument. When long-run income is proxied by education and occupation, a significant positive relationship is found between saving rate and long-run income. However, the positive correlation disappears when they use consumption as the long-run proxy.

Alan et al. (2015) apply this methodology to the Canadian FAMEX dataset and surprisingly reach a different conclusion. They conclude that the rich do not save more, at least compared to households in the middle of the income distribution. They show that saving rate is flat above the median long-run income, and under some circumstances, even falls after the median income level.

The controversial findings of Alan et al. (2015) made it appealing to revisit the question and re-evaluate Dynan's methodology with more advanced methods that address the endogeneity problem in a quantile regression context. Only a few years after Dynan (2004), more advanced methods were developed to handle different types of endogeneity in Quantile Regression literature, in particular, Ma and Koenker (2006), Lee (2004) and Horowitz and Lee (2007). Not only do these methods deal with the endogeneity problem more efficiently, they also provide more-detailed information about the joint distribution of the saving rate and income.

In this study, in a set of Monte Carlo simulations, we compare the performance of three different Instrumental Quantile Regression approaches, in addition to the methodology proposed in Dynan (2004). We define three different environments to highlight advantages and limitations of these methods in different frameworks. These methods are specifically chosen since each addresses certain characteristics of the data and their structure is more applicable to the problem. Also, they are more comparable to what Dynan (2004) propose, theoretically and empirically.

The rest of this paper is formatted as following; the next section discusses four different approaches that address the endogeneity issue with different presumptions and structure in the context of Instrumental Variable Quantile Regression. In Section.2.3, three different Monte Carlo environments are presented and then the performance of four methods are compared. The data is discussed in Section.2.4, and the result of implementing four approaches are reported in Section.2.5.

2.2 Alternative Approaches

In this section, we compare Dynan's methodology with other two-step estimators that have been developed in the Instrumental Variable Quantile Regression literature. We consider the parametric approach of Ma and Koenker (2006), the semi-parametric method proposed in Lee (2004), and finally, the fully non-parametric approach of Horowitz and Lee (2007). We further address the advantages and limitations of each method. The general framework is defined as:

$$s = f(Y^*) + Z_2'\alpha_2 + \epsilon \tag{2.1}$$

In the context of the saving rate and long-run income, s represents the household's saving rate, Y^* stands for the long-run income, and Z_2 stands for possible exogenous variables such as age and other demographic variables. The measurement error and transitory income shocks inherent in reported income make it a poor proxy for long-run income. If we define saving as income minus consumption expenditures, the presence of measurement error in current income on both sides of the equation (2.1), creates a positive bias in the estimation, even when there is no relationship between the two variables.

If the final goal is to model the relationship of long-run income and the saving rate, by replacing long-run income with the observed income, equation (2.1) will become:

$$s = f(Y) + Z_2'\alpha_2 + U \tag{2.2}$$

Assuming that $f(Y) - f(Y^*) = \phi(Y - Y^*)$, We can write:

$$f(Y) - f(Y^*) = \phi(V)$$

where $V = Y - Y*$, then the error term in the equation (2.2) is defined as

$$U = \phi(V) + \epsilon, \tag{2.3}$$

We can define V as the measurement error and transitory components of income. Then U is a composite error term which is a combination of measurement errors and transitory shocks in income and saving rate. The presence of V in equation (2.2) through U, either linearly or non-linearly, causes endogeneity, and any independent estimate of the relationship of income and the saving rate will be biased and inconsistent.

In equation (2.3), Dynan (2004) parameterize $f(.)$ by five dummy variables and assume a linear format for $\phi(.)$. By running a median regression on five dummy variables, they estimate the median saving rate in five different brackets of long-run income and compare them to see if the saving rate is increasing in higher income brackets.

In our first alternative approach, a specification proposed in Ma and Koenker (2006), V enters in equation (2.2) linearly while it is interacted with income. This interaction term introduces nice features that we will discuss in the next sections.

The second approach is a proposed framework in Lee (2004) which considers a linear format for $f(.)$, however, in order to capture any potential non-linearity between s and V, $\phi(.)$ is estimated non-parametrically. This semi-parametric specification assumes a linear relationship between income and saving rate, which varies at different quantiles of income distribution. Also, it does not make any assumption about the relationship of V and s.

Finally, we use the non-parametric instrumental variable approach developed by Horowitz and Lee (2007), which estimates $f(.)$ non-parametrically and makes no assumption about the underlying format of $f(.)$ and $\phi(.)$. Each framework is discussed in details in the following sections.

2.2.1 Dynan's Methodology

The widely used methodology, proposed by Dynan (2004), to assess the relationship of households saving rate and long-run income is as follows:

$$s = f(Y^*) + Z_2'\beta + U$$

where Y^* represents long-run income purged of transitory income and measurement errors. In order to allow for non-linearities in the relationship and, at the same time, avoid a complex estimation method, Dynan uses five dummy variables instead of $f(.)$, standing for five different brackets of long-run income.

In the first step, in order to control for transitory income shocks and the measurement error problem in the observed income, a proxy for long-run income is constructed by regressing observed income on some individual instruments such as education, lagged and lead income, and total expenditure. Then fitted values of income from this regression, representing true long-run income, are used to assign households to five brackets of long-run income.

$$D_{ij} = \begin{cases} 1 & \text{if } \hat{Y}_j \in i^{th}\text{Income Bracket,} \qquad i = \{1, ..., 5\} \\ 0 & \text{otherwise} \end{cases}$$

where

$$\hat{Y} = \hat{\beta}_0 + Z_1'\hat{\beta}_1 + Z_2'\hat{\beta}_2 \tag{2.4}$$

Here Z_1 is a vector of instruments and Z_2 includes other exogenous variables such as age group. In the second stage, a conditional median saving rate is estimated by running a median regression on the five dummy variables defined above and other exogenous variables while the intercept is suppressed.

$$Q_s(0.5|D, Z_2) = \sum_{i=1}^{5} D_i \alpha_i + Z_2' \beta$$

In this approach α_i's represent the median saving rate in five different brackets of long-run income, where long-run income is proxied by some instruments.

Due to its simple framework, this method has attracted a significant attention in the literature. Even though simple, this method works very well in some circumstances and predicts the saving rate very well. However, it has some drawbacks and we will show in Section.(2.3) that in more complicated situations, it fails to predict saving rate when we move toward the tails of the income distribution. This methodology works similar to a low-resolution non-parametric approach, and it does not provide an accurate picture of the entire joint distribution of income and saving rate.

2.2.2 Parametric Control Variate Model

The first alternative method that we implement is the Control variate Estimation method proposed by Ma and Koenker (2006). This specification has the following triangular specification.

$$s = \alpha_0 + \alpha_1 Y + Z_2' \alpha_2 + \alpha_3 (\underbrace{\epsilon + \lambda V}_{U}) Y \tag{2.5}$$

$$Y = \beta_0 + Z_1' \beta_1 + Z_2' \beta_2 + V \tag{2.6}$$

where

$$V \sim F_v, \epsilon \sim F_\epsilon$$

In this framework, V and ϵ are interacted with Y to capture any scale-effect that Y may introduce into s and that can be translated into the heteroskedasticity observed between income and saving/consumption. Assuming $\lambda > 0$, it captures the upward bias due to measurement error and transitory changes in income, i.e. V.

Consistent estimating of α_1, requires V and ϵ be i.i.d and independent of Z_1 and Y, with $V \sim F_v$ and $\epsilon \sim F_\epsilon$. This approach also requires Y to be predetermined with respect to s, this means that households observe income first and then decide how much to save. All these assumptions can be summarized in (2.7).

$$Q_U(\tau_1|Z_1, Z_2) = Q_U(\tau_1|V, Z_2) = Q_U(\tau_1|V) \equiv F_\epsilon^{-1}(\tau_1) + \lambda V \tag{2.7}$$

What equation (2.7) means is that the dependency of U on Z_1 can be explained through V, and they do not depend on Z_2. In other words, conditional distribution of U can be explained by conditioning on V.

The conventional 2SLS estimator of α_1 is

$$\alpha_1 = (\hat{Y}' M_{Z_2} \hat{Y})^{-1} \hat{Y}' M_{Z_2} s$$

where

$$\hat{Y} = \hat{\alpha_2} Z_1 + Z_2' \hat{\beta}_2$$

However, in the Control variate approach, in equation (2.2) instead of replacing Y by $\hat{Y}$, $Y - \hat{V}$ is used, which is equal to $\hat{Y}$. This basically means including $\hat{V}$ in equation (2.2) and estimates of α_1 and α_2 remain the same as the estimates of 2SLS, when ordinary least square is used. Now, quantile regression provides us with features that allows us to estimate the relationship of income and saving rate at different points of the joint distribution. To apply this method, we first estimate the conditional quantile function of Y (conditional on Z_1 and Z_2), in order to calculate $V(\tau_2) = V - F_v^{-1}(\tau_2)$. This quantile regression is done at different quantiles, in this study $\tau_2 = (0.1, 0.3, 0.5, 0.7, 0.9)$, to obtain a set of $\hat{V}(\tau_2)$s. In the second stage another quantile regression is done with each of these $\hat{V}(\tau_2)$s, in order to evaluate the relationship at different quantiles of income, τ_2.

$$Q_Y(\tau_2|Z_1, Z_2) = \beta_0 + Z_1'\beta_1 + Z_2'\beta_2 + F_v^{-1}(\tau_2) \tag{2.8}$$

By adding and subtracting $F_v^{-1}(\tau_2) = V - V(\tau_2)$ in the equation (2.5), which is the τ_2^{th} quantile of V, we have

$$
\begin{aligned}
s &= \alpha_0 + \alpha_1 Y + Z_2'\alpha_2 + \alpha_3\Big(U + \lambda V - \lambda F_v^{-1}(\tau_2) + \lambda F_v^{-1}(\tau_2)\Big)Y \\
&= \alpha_0 + \alpha_1 Y + Z_2'\alpha_2 + \alpha_3\Big(U + \lambda V(\tau_2) + \lambda F_v^{-1}(\tau_2)\Big)Y \\
&= \alpha_0 + Z_1\alpha_2 + Y\Big(\alpha_1 + \alpha_3 U + \alpha_3\lambda F_v^{-1}(\tau_2)\Big) + \alpha_3\lambda V(\tau_2)Y
\end{aligned}
$$

Therefore the conditional quantile of the saving rate at τ_1 is

$$
Q_s(\tau_1|Y, Z_2, V(\tau_2)) = \alpha_0 + Z_2\alpha_2 + Y\Big(\alpha_1 + F_u^{-1}(\tau_1) + \lambda F_v^{-1}(\tau_2)\Big) + \lambda\alpha_3 V(\tau_2)Y \tag{2.9}
$$

By estimating (2.9) with conventional quantile regression, we have:

$$
Q_s(\tau_1|Y, Z_2, V(\tau_2)) = \delta_0(\tau_1) + \delta_1(\tau_1, \tau_2)Y + \delta_2(\tau_1)Z_2 + \delta_3(\tau_1)\hat{V}(\tau_2)Y \tag{2.10}
$$

This implies the following equalities:

$$
\begin{aligned}
\delta_0(\tau_1) &= \alpha_0 \\
\delta_1(\tau_1, \tau_2) &= \alpha_1 + \alpha_3\Big(F_u^{-1}(\tau_1) + \frac{\lambda}{\beta_3}F_v^{-1}(\tau_2)\Big) \\
\delta_2(\tau_1) &= \alpha_2 \\
\delta_3(\tau_1) &= \lambda\alpha_3
\end{aligned} \tag{2.11}
$$

In equation (2.10), $\hat{V}(\tau_2)$ is estimated consistently as the residuals of a linear quantile regression of Y on instrument Z_1, exogenous variables Z_2 and 1 in equation (2.8), at different τ_2s. The fact that δ_1 is a function of τ_2 implies that, the effect of Y on s is different at different quantiles of income. Unlike $2SLS$, which provides a single slope coefficient for Y, at different quantiles of income distribution, we will have a different coefficient. This structural model helps to evaluate the relationship between saving rate and income at different quantiles of income distribution, i.e. estimating different quantiles of the saving rate at different points of income distribution. It is proved in Ma and Koenker (2006) that the 2SLS estimate is equal to the average of $\delta(\tau_1, \tau_2)$ over the range of $\tau 2$.

Including $\hat{V}$ in (2.10) helps us to estimate the relationship at different quantiles of income such that, by changing τ_1 and τ_2, we can estimate the relationship at different points of the joint distribution. This is comparable to 2SLS, which estimates a single coefficient for the entire income distribution, and the Dynan methodology which categorize the observations into a few dummy variables in order to predict the saving rate in each of these categories. Moreover interaction of $\hat{V}$ with income also provides the feature to control for the heteroskedastic relationship between income and the saving rate.

2.2.3 Semi-Parametric Control Function Model

The Control function approach, an extension to the Control variate method in a semi-parametric fashion, requires less restrictive assumptions and lowers the risk of mis-specification. It offers the same triangular structure without imposing any structure on $\hat{V}(\tau_2)$ in equation (2.5). In the same structural model as in the Control variate method,

$$s = Y\alpha_1(\tau_1) + Z_2\alpha_2(\tau_1) + U \tag{2.12}$$

$$Y = \beta_0 + Z_1\beta_1(\tau_2) + Z_2\beta_2(\tau_2) + V \tag{2.13}$$

The composite error term is defined as $U = \phi(V) + \epsilon$, where $\phi(.)$ is a real-valued function of V. A consistent estimation of α_1 and α_2 requires

$$Q_U(\tau_1|Z_1, Z_2) = Q_U(\tau_1|V, Z_2) = Q_U(\tau_1|V) \equiv \phi(V) \tag{2.14}$$

and

$$Q_V(\tau_2|Z_1, Z_2) = 0 \tag{2.15}$$

The equivalency implies that conditional quantiles of U are a function of V, i.e. $U = \phi(V) + \epsilon$. The Control function method imposes a weaker restriction and does not assume any functional form for the dependency of U on V. The two step estimator is as follows. In the first step, $\hat{V}(\tau)$ is obtained through a linear quantile regression of Y on the instrument Z_1, exogenous variables Z_2 and 1, the same as in the Control variate approach.

$$Q_Y(\tau_2|Z_1, Z_2)) = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 \tag{2.16}$$

The second step involves a semi-parametric regression of s on Y, Z_2 and a polynomial approximation of $\phi(.)$.

$$Q_s(\tau_1|Y, Z_2, \hat{V}(\tau_2)) = \alpha_0 + \alpha_1 Y + \alpha_2 Z_2 + \phi(\hat{V}) \tag{2.17}$$

Root MSPE is used at $\tau_2 = \tau_1 = 0.5$ to choose the optimal order of polynomial. Then equation (2.17) is estimated at five different quantiles. Again the different coefficient of Y would be different at different τ_2's, and they are being interpreted as the marginal effect of long-run income on the saving rate.

2.3 Monte Carlo Simulations

In this section, in a set of Monte Carlo experiments, we investigate the performance of the four approaches to estimate an endogenous relationship in three different environments. First, a regular type of endogeneity is introduced in a nonlinear DGP with different parameterizations to test the prediction power of all four approaches.

Also in a linear DGP, we introduce a non-linear type of endogeneity to test the performance of these approaches in correcting for the endogeneity problem, when error terms are non-linearly related to regressors. Finally, in a more restrictive structural model, a third type of endogeneity is used that introduces a scale-effect in the model while the DGP remains linear. To investigate the performance of the four methods discussed above, we compare the root MSPE of predicted values for the saving rate, $\hat{s}$, at five different quantiles of income.

Dynan's approach and the non-parametric IV method can provide us with only the predicted values of the saving rate $\hat{s}$. In the Control variate methods, the slope coefficients are also estimated so that we can compare their performances in a more detailed level. In all experiments, the sample size and number of iterations are set to 1000.

2.3.1 Non-Linear DGP with Linear Endogeneity

In this environment, a regular linear correlation between error term and the regressor is considered, while the regressor enters into the equation in a non-linear fashion. In the context of the relationship between saving rate and income, in equation (2.1), $f(.)$ is considered non-linearly, and $\phi(.)$ has a linear form.

Samples are generated through three standard normal random variables to create the error term, the endogenous regressor, and the IV, all of which are mutually correlated.[2]

$$U_1 \sim N(0,1), U_2 \sim N(0,1), U_3 \sim N(0,1) \tag{2.18}$$

The instrument, Z, is basically a standard normal random variable that is trans-

[2]This is a modified version of the specification in Horowitz and Lee (2009)

formed into $[0, 1]$ interval through its cumulative distribution function.

$$Z \sim CDF(U_1) \tag{2.19}$$

The endogenous variable (Income) is defined as a weighted average of U_1 and U_3 so it is correlated with both the instrument and the error term.

$$Y \sim CDF(\rho U_1 + \sqrt{(1-\rho^2)}U_3) \tag{2.20}$$

and the error term that is correlated with the endogenous variable through U_3 is defined below.

$$U \sim \sigma CDF(\eta U_2 + \sqrt{(1-\eta^2)}U_3) \tag{2.21}$$

A non-linear DGP is defined as a third degree polynomial as in (2.22). a non-linear DGP helps us avoid any assumption or restriction on the relationship of income and saving rate.[3]

$$f(Y) = \theta_0 + \theta_1 Y + \theta_2 Y^2 + \theta_3 Y^3 \tag{2.22}$$

And finally the saving rate is defined as the sum of the $f(.)$ and the error term.

$$s = f(Y) + U \tag{2.23}$$

Model parameters are defined as $(\rho, \eta, \sigma) = (0.8, 0.5, 0.1)$. A sample of 200 observations from data is depicted in Figure.**??**, where the numbers in the lower panel represent the correlation between variables. Bias, standard deviation, and the root mean squared error (MSPE) for all four methods are reported in Table.2.1. The NPQR row represents the results of a non-parametric quantile regression of the saving rate on income, without considering the endogeneity issue. We also evaluate the performance of the parametric Control variate method and the semi-parametric

[3]For the purpose of sensitivity analysis, we consider four different parameterizations for this non-linear DGP to make sure that our results are robust to different shapes of $f(Y)$. The results for the other three parameterizations are reported in the Appendix

Control function approaches in estimating the slope of $f(Y)$ at five quantiles of Y, which are reported in Table.2.4.

In this framework, Dynan's method outperforms the other methods in terms of both bias and variance, even though the other methods are doing a fine job. Table.2.4 shows the result of estimating the slope of $f(.)$ by the Control variate and the Control function methods at five quintiles of income. Lee's method has a better performance especially off the median, while the bias increases with higher quintiles in the Control variate approach.

Table 2.1: Monte Carlo Experiment, N = 1000, R = 1000.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Bias					
Control Variate	-0.0209	-0.0494	-0.0094	0.0414	0.0513
Control Function	-0.0372	-0.0447	-0.0009	0.0451	0.0356
Horowitz & Lee	0.1335	0.0316	-0.0666	-0.0169	0.1352
Dynan	-0.0085	-0.0235	0.0004	0.0229	0.0076
NPQR (No IV)	-0.0282	-0.0208	0.0326	0.0861	0.0930
Standard Deviation					
Control Variate	0.0002	0.0005	0.0011	0.0018	0.0012
Control Function	0.0011	0.0018	0.0017	0.0016	0.0010
Horowitz & Lee	0.0037	0.0043	0.0022	0.0045	0.0044
Dynan	0.0003	0.0002	0.0002	0.0002	0.0003
NPQR (No IV)	0.0006	0.0007	0.0007	0.0008	0.0010
Root MSPE					
Control Variate	0.0006	0.0029	0.0012	0.0036	0.0039
Control Function	0.0025	0.0038	0.0017	0.0036	0.0023
Horowitz & Lee	0.0215	0.0053	0.0066	0.0048	0.0227
Dynan	0.0003	0.0008	0.0002	0.0008	0.0003
NPQR (No IV)	0.0014	0.0011	0.0018	0.0082	0.0096

Note: Entries are the Bias, Standard Deviation, and Root MSPE for predictions of conditional median saving rate at five different quantiles of income, with the DGP described in Section.2.3.1.

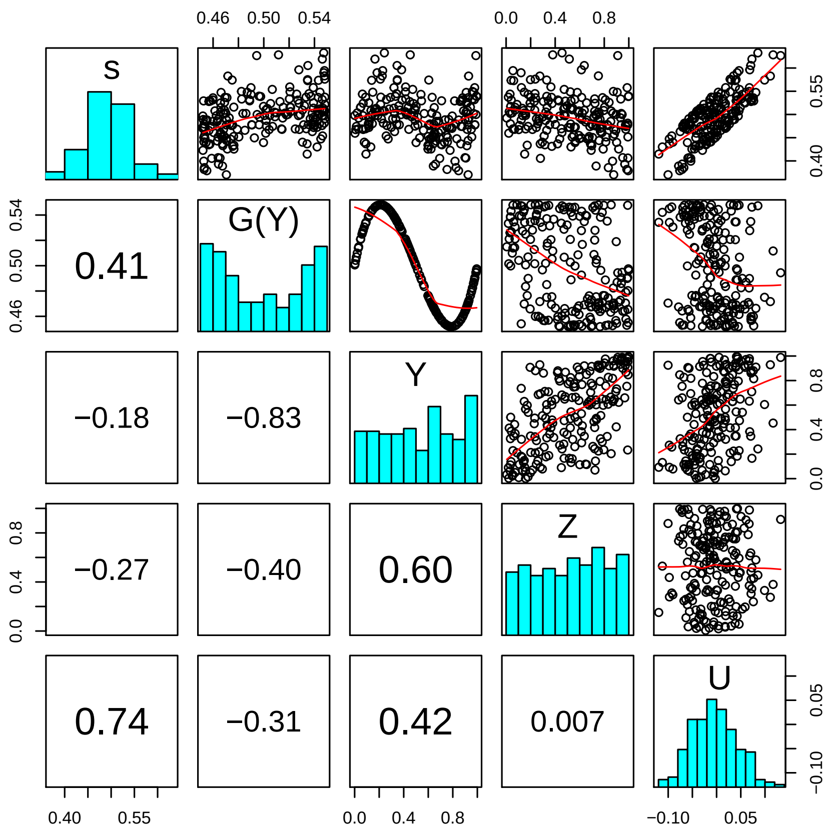

Figure 2.1: A sample of 200 observations from the DGP described in Section.2.3.1, Numbers in lower panel are the correlations between variables.

2.3.2 Linear DGP with Non-Linear Endogeneity

The environment in this section aims to examine and compare the performance of the four approaches when a non-linear type of endogeneity is present in the model. For the sake of simplicity, the relationship between Y and s is considered to be a linear one in order to focus on the effect of the new type of endogeneity.[4] It can be interpreted that in equation (2.1), $f(.)$ takes a linear form while $\phi(.)$ appears in a non-linear form as in (2.25).

Consider a recursive specification as indicated below,

$$\begin{aligned} s &= Y\beta + \epsilon + \phi(V) \\ Y &= Z\alpha + V \end{aligned} \tag{2.24}$$

where

$$\phi(V) = V + 4\exp[-(V-1)^2] \tag{2.25}$$

and

$$V \sim N(0,1), \epsilon \sim N(0,1) \tag{2.26}$$

In this environment, Y is related to the error term in a non-linear form. $\phi(V)$ produces a hump around one and is the source of non-linearity in the endogenous relationship. The IV is generated from a standard normal distribution while linearly normalized to $[0,1]$ interval. Figure.**??** shows a sample of 200 observations from one realization of the data. The numbers in the lower panels represent the correlations between variables.

Table.2.2 shows the bias, standard deviation and root MSPE of predicted values for $\hat{s}$ at five different quantiles of Y. In this specification, with non-linear endogeneity, Dynan does a poor job, even though it has a lower variance, due to its linear nature. The magnitude of the bias is not negligible and increases as we move away from the median. The Non-Parametric IV method does an even worse job in esti-

[4] This is a modified version of the specification in Lee (2004)

mating the true saving rate.

The Control function approach outperforms other methods in this environment as it is expected, while the Control variate method has slightly larger root MSPEs. From Table.2.5 we can see that the Control function approach has a trivially small bias in estimating the slope coefficient β. This means that small biases in estimates of $\hat{s}$ come mostly from the estimates of the intercept.

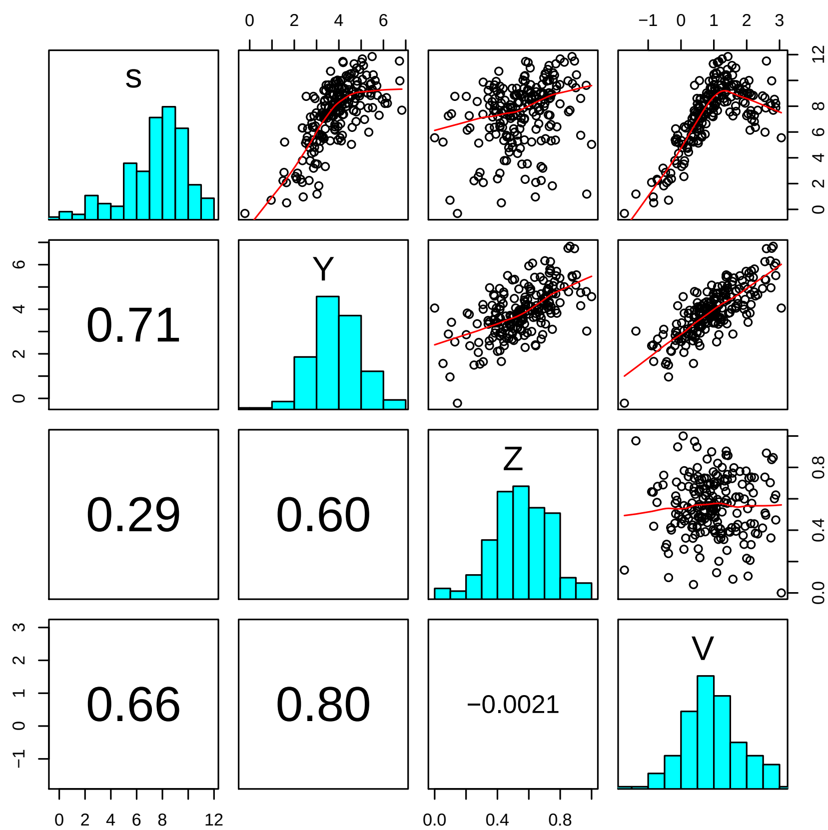

Figure 2.2: A sample of 200 observations from the DGP described in Section.2.3.2, Numbers in lower panel are the correlations between variables.

Table 2.2: Monte Carlo Experiment, N = 1000, R = 1000.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Bias					
Control Variate	0.2051	-0.0354	-0.1119	-0.1593	-0.0129
Control Function	-0.0338	-0.0645	-0.0433	0.0037	0.1153
Horowitz & Lee	-1.2226	-1.9465	-2.5604	-3.1777	-3.9477
Dynan	1.2553	0.4993	-0.0337	-0.5736	-1.3405
NPQR (No IV)	6.1890	6.0539	6.0078	5.9859	5.9926
Standard Deviation					
Control Variate	0.3462	0.4077	0.4440	0.5128	0.5461
Control Function	0.1302	0.1009	0.0948	0.0974	0.1299
Horowitz & Lee	0.0947	0.1170	0.1081	0.1206	0.0949
Dynan	0.1506	0.1575	0.1634	0.1570	0.1636
NPQR (No IV)	164.7090	164.7430	164.7453	164.7566	164.8593
Root MSPE					
Control Variate	0.4024	0.4092	0.4579	0.5370	0.5462
Control Function	0.1345	0.1198	0.1042	0.0975	0.1737
Horowitz & Lee	1.2263	1.9500	2.5626	3.1800	3.9489
Dynan	1.2643	0.5235	0.1669	0.5947	1.3505
NPQR (No IV)	164.8252	164.8542	164.8549	164.8653	164.9681

Note: Entries are the Bias, Standard Deviation, and Root MSPE for predictions of conditional median saving rate at five different quantiles of income, with the DGP described in Section.2.3.2.

2.3.3 Linear DGP with Heteroskedastic Endogeneity

In this simulation experiment we consider a structural model that combines a linear DGP with a scale-shift type of endogeneity as below.[5]

$$\begin{aligned} s &= \beta_0 + \beta_1 Y + \delta(\epsilon + \lambda V)Y \\ Y &= \alpha_0 + Z'\alpha_1 + V \end{aligned} \tag{2.27}$$

where

$$V \sim N(0,1), \epsilon \sim N(0, 0.25)$$

Again in the context of equation (2.1), both $f(.)$ and $\phi(.)$ are linear, however, the interaction between Y and $U = \epsilon + \lambda V$ creates a scale-shift effects that can be translated into the heteroskedasticity observed in the relationship between saving rate and income. Scale-shift effect means that larger values of V, are associated with not only the larger values of Y and s but also affect the conditional variance of s, and it increases at higher quantiles of Y. As it was shown in (2.11), the slope coefficient for income, $\delta_1(\tau_1, \tau_2)$ depends on τ_2, the income quantile.

The model is parameterized as $(\alpha_0, \alpha_1, \delta, \lambda, \beta_0, \beta_1) = (1, 1, 1.5, 5, 1, 1)$ and Z is generated from a normal distribution, $Z \sim N(1, 4)$ that is normalized to $[0, 1]$ interval. Figure.?? shows a sample of 200 observations from this DGP.

Results of this Monte Carlo experiment are reported in Table.2.3, Dynan's estimates are completely biased when we move away from the median, and the magnitude of the bias is not comparable with the results from Control variate and Control function approaches. Control function approach has a fairly better performance than the Control variate, regarding predicted values of saving rate, but not in estimating the slope coefficient of the income. While we saw that the performance of Control variate method was worse compared to Control function's performance in the previous section, we had a non-linear endogeneity.

[5] This is a modified version of the specification designed in Ma and Koenker (2006).

To summarize the results of three Monte Carlo simulations, we can say that in the first environment with linear endogeneity, all four methods do a fine job with Dynan slightly outperforming the rest in terms of both bias and variance. In the second environment, not surprisingly Dynan and Nonparametric IV methods show significant bias in predicting the saving rate. Control variate method does a more efficient job in estimating the slope coefficient than the predicting the saving rate.

In the third environment, the performance of Dynan and Nonparametric IV approaches are very poor and show significantly large biases. Regarding the prediction of saving rate, Control function method outperforms Control variate approach in this environment. However, Control variate approach estimates the slope coefficient with smaller bias and variance.

2.4 Data

Data used in this study are drawn from two famous Canadian surveys: the Family Expenditure survey (FAMEX), conducted by Statistics Canada in 1982, 1984, 1986, 1990, 1992 and 1996; and the Survey of Household Spending (SHS), a revised version of FAMEX, which is conducted on a yearly basis from 1997 to 2009. For several reasons these data sets, in particular, FAMEX, are known for possessing high quality[6]. Unlike most surveys, these are not diary-based surveys, instead, households are interviewed in person in January, February and March for the previous calendar year. Detailed information on expenditures is collected by recall and reference to bills, receipts, and records from the previous year. In order to gather thorough information, multiple lengthy interviews are conducted. Another feature contributing to the credibility of these surveys is the "Balance Edit" that is a quality control measure. Households whose report of expenditure are more than 20% different from

[6]See for example Brzozowski and Crossley (2011) and Barrett et al. (2013)

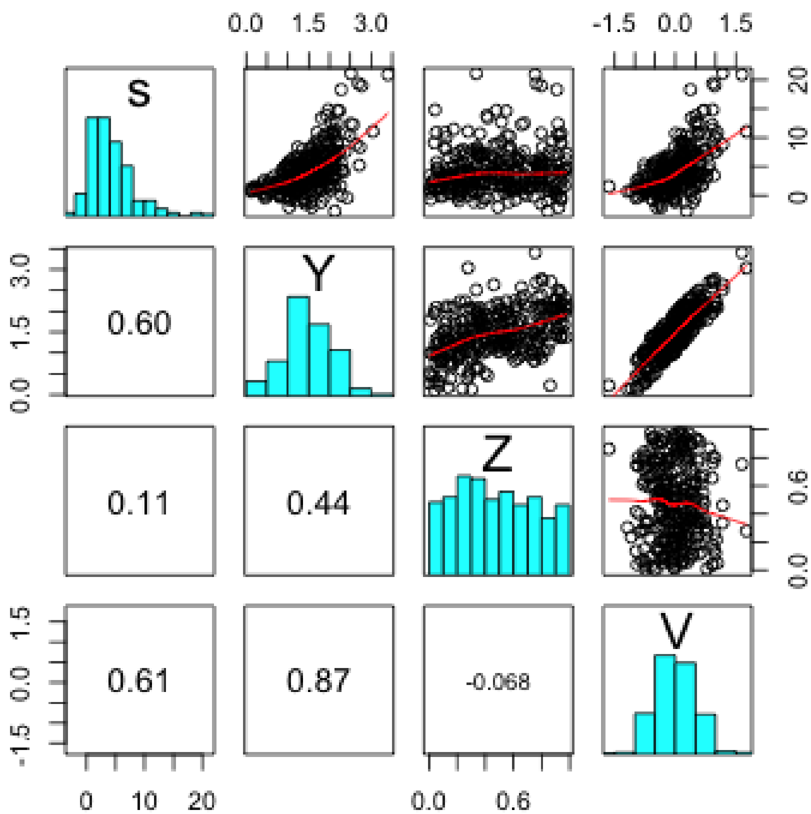

Figure 2.3: A sample of 200 observations from the DGP designed in Section.2.3.3, Numbers in lower panel are correlations between variables.

Table 2.3: Monte Carlo Experiment, N = 1000, R = 1000.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Bias					
Control Variate	0.2291	0.1950	0.0656	-0.1790	-0.6192
Control Function	0.2102	0.2007	0.0783	-0.1787	-0.6349
Horowitz & Lee	2.3664	0.3773	-2.3317	-6.0926	-13.2484
Dynan	4.4950	2.6029	-0.0614	-3.8368	-11.2127
NPQR (No IV)	-1.0785	-1.2985	-1.4735	-1.7806	-2.2615
Variance					
Control Variate	0.9607	1.3623	1.6905	1.9241	2.4266
Control Function	0.9874	1.3694	1.6871	1.9169	2.4368
Horowitz & Lee	0.0618	0.0832	0.1310	0.1381	0.0917
Dynan	0.4798	0.5375	0.5500	0.6254	0.6777
NPQR (No IV)	8.8607	8.8865	8.8878	9.0031	8.9332
MSPE					
Control Variate	0.9877	1.3762	1.6918	1.9324	2.5043
Control Function	1.0095	1.3840	1.6889	1.9252	2.5181
Horowitz & Lee	2.3672	0.3864	2.3354	6.0941	13.2487
Dynan	4.5205	2.6578	0.5534	3.8875	11.2332
NPQR (No IV)	8.9260	8.9808	9.0091	9.1775	9.2150

Note: Entries are the Bias, Standard Deviation, and Root MSPE for predictions of conditional median saving rate at five different quantiles of income, with the DGP designed in Section.2.3.3.

the sum of income and the net increase in assets, are asked to review their report in the data collecting stage. In the processing stage, if the disparity is still more than 20%, observation is removed from data.[7]

For the matter of comparability with previous works in the literature, we follow the same rules for sample selection and variable definitions. Even though we estimate this relationship at all survey years, we focus on the year of 1996 to compare our results with Alan et al. (2015) who chose the same year, due to data quality

[7]Except for the year of 2006, due to adaptation of Computer Assisted Personal Interviewing (CAPI) for the first time, instead of paper-pencil based interview, Balance Edit is not applied.

concerns[8]. In our sub-sample, the age profile of the head of households is restricted to 29 to 59, whose income is greater than 1000$.

In order to evaluate the relationship between the saving rate and household income, Dynan (2004) define the concept of "Active Saving", which is the fraction of current income not spent. FAMEX and SHS datasets provide us with two measures of saving in this context. The first one is the common measure of saving as the difference between a household's disposable income and their total expenditure. Elaborate information on the household expenditures in FAMEX dataset gives us the advantage of obtaining an accurate measure of saving, that is not available in other datasets used in the literature.

Another measure of saving that is reported in FAMEX and SHS dataset is the yearly change in the household's wealth. This includes the total net change in assets, less total net change in debts. The difference between the two measures is used to perform the "Balance Edit" test explained above. The two measures are highly correlated and if there is any, they suffer from the same source of measurement error. Findings of Alan et al. (2015) is very consistent across the two measures and the choice of saving measure makes little difference to our central question. Estimates based on net changes in wealth give lower saving rates than those based on the difference between income and consumption.

Disposable income, net household income after taxes, including wages and salaries, investment income, self-employment earnings, government transfers, and income from other sources. Total consumption is defined as household's total expenditure and includes expenditures for housing, food, clothing, household operations, personal care, transportation, recreation, education, tobacco and alcoholic beverages, reading materials and miscellaneous expenses. Principal payments of mortgage are considered as saving; and interest payments are treated as consumption. Our definition of

[8] e.g. after 1996, interest payments on the mortgage are not reported separately from the principal payments

non-durable consumption excludes expenditures on durables, i.e., vehicles, housing, and household furnishings and equipment.

2.5 Empirical Results

In this section, the results of estimating Equation.(2.1) with four different methods are reported. Even though we estimate this relationship at all survey years, for the matter of comparability with Alan et al. (2015), we also focus on the year of 1996. Due to data quality concerns, Alan et al. (2015) choose this year. For the relationship between saving rate and Current Income, Burbidge and Davies (1994) and Alan et al. (2015), find strong evidence for a positive relationship in Canada and we avoid to investigate this any further, as it is not our primary interest in this study. We start our analyses by documenting the results from different approaches we discussed earlier in this article.

We use Non-Durable Consumption and its components, as instruments in these regressions. Education of the head of the household and his/her spouse is another instrument that is used in the literature for long-run income. However, use of this instrument when saving is the left-hand side variable, have always been followed by strong critiques. The argument is that people with higher education are more patient and tend to consume less to invest for a higher consumption in future. For this reason, it was preferred to withdraw education from the set of instruments.

Table.2.7 to Table.2.11 show the predicted values for saving rate at different quintiles of long-run income. Our results are very consistent across different instruments. But the trends are not very similar among the different estimators. In Dynan's approach, five saving rates are obtained as the coefficients of five dummy variables from a single median regression. In control variate and control function methods, each saving rates is a predicted value of a separate regression. For the nonparametric approach, saving rates are the predicted values at the same percentiles as other approaches. The results from control variate and control function methods are close, and both show an increasing saving rate across higher quintiles of income. However, the results of the other two methods show a different behavior.

Table.2.7 to Table.2.11 show the predicted values for saving rate at different quintiles of long-run income. Our results are very consistent across different instruments. But the trends are not very similar between different estimators. In Dynan's approach, five saving rates are obtained from a single regression as the coefficients of five dummy variables. In control variate and control function approaches, each saving rate is a predicted value of a separate regression. For the non-parametric approach, saving rates are the predicted values at the same percentiles as other approaches. The results from control variate and control function methods are close and both show an increasing saving rate across higher quintiles of income. However, the results of the other two methods show a different behavior.

Standard errors are reported in parenthesis and the significance of coefficients are illustrated with asterisks. An age dummy variable for households with the head of younger than 45 years old is included in all regressions and coefficients are reported. In Dynan methodology, a single coefficient is estimated for age groups at all quintiles of income. However, since a separate regression is estimated at each quintile of income in Control variate and Control function methods, we have different estimates for the age dummy variable coefficient. In Nonparametric IV method, the function is estimated separately for the two age groups, and the difference is reported. Age dummy variables are not significant generally. Except in Dynan's method, they are positive and show different trends across estimators(but consistent across instruments).

Regression results are illustrated in Figure.2.4 to Figure.2.7. Each figure shows predicted saving rates estimated at different quintiles of income by one of the four methods using various instruments. Results are considerably consistent among the various instruments for all estimators. In Figure.2.4 and Figure.2.5 Control variate and Control function methods show a monotonically rising saving rate at different quintiles of long-run income. Control variate method shows a more convex relationship, meaning saving rate increases at a higher rate at higher quintiles of income.

Results from Dynan method are depicted in Figure.2.6. Consistent with findings of Alan et al. (2015), the saving rate is not increasing with quintiles of long-run income in this approach. It ranges from 5% to 25%, but peaks at the fourth quintile of income when "Non-Durable Consumption", "Food Expenditure" and "Restaurant" expenses are used as instruments for long-run income. When "Total Consumption" and "Food and Clothing" costs are used as the instrument, saving rate is at highest in the third quintile of income and then falls.

Figure.2.7 shows the non-parametric estimation of the relationship across the long-run income distribution when it is proxied by different instruments. Generally across all instruments, in the first quintile of income, the saving rate is increasing with a peak at the 20^{th} percentile of income and the group of younger households tend to have a higher saving rate. Generally speaking, the saving rate does not change significantly after the 20^{th} percentile of income, especially for the case that Restaurant expenses are used as the instrument. When total consumption is the instrument, the saving rate drops for both age groups, while younger households save more. When other components of consumption are used as the instrument, saving rate is decreasing until the 60th percentile of income and then it increases while again younger households tend to have a higher saving rate.

We can have a better comparison of the results of four methods in Figure.2.8 that shows the result of all four methods in one plot. These are predictions of saving rate for households with a head of older than 45 years old. Results from few other survey years are also reported in Figure.2.9 to Figure.2.11. The results are quite consistent across the years. Across all instruments, Dynan method and the Non-parametric method both show a bell-curved shape, that peaks in about 50^{th} percentile of income in the former and around 25^{th} percentile of income in the latter approach. In contrast, in Control variate method and Control function methods saving rate increases monotonically and Control function method estimates a steeper slope compared to Control variate approach. Estimates of the saving rate from different approaches are somehow close around the median income level. However in the bottom quantiles, Control variate, and Control function methods estimate a much lower saving rate and a higher saving rate at higher quantiles of income.

2.6 Conclusion

This paper revisited an old question with new approaches and tried to highlight limitations and advantages of these approaches in a series of Monte Carlo experiments. The long debated question is whether household saving rate increases over the long-run income distribution or more well-off households at higher quantiles of long-run income save the same proportion of their income as middle class and low income households.

We consider three alternative approaches in addition to the methodology proposed by Dynan (2004). In a series of Monte Carlo experiments, the performance of four approaches is tested in three different environments. Results show that Dynan's methodology outperforms only when there is a simple linear type of endogeneity in the model and with more complicated types of endogeneity, it fails to predict true saving rate, and its bias increases as we move away from median towards tails of distribution.

In the end, using FAMEX and SHS datasets from Canada from 1982 to 2009 with a focus on 1996. We re-evaluated the question with the alternative methods. Our findings varied from those of Alan et al. (2015) that use the same dataset. Our empirical analysis suggest that more affluent households do save a larger fraction of their income. The results are not sensitive to different instruments and does not change significantly over the years.

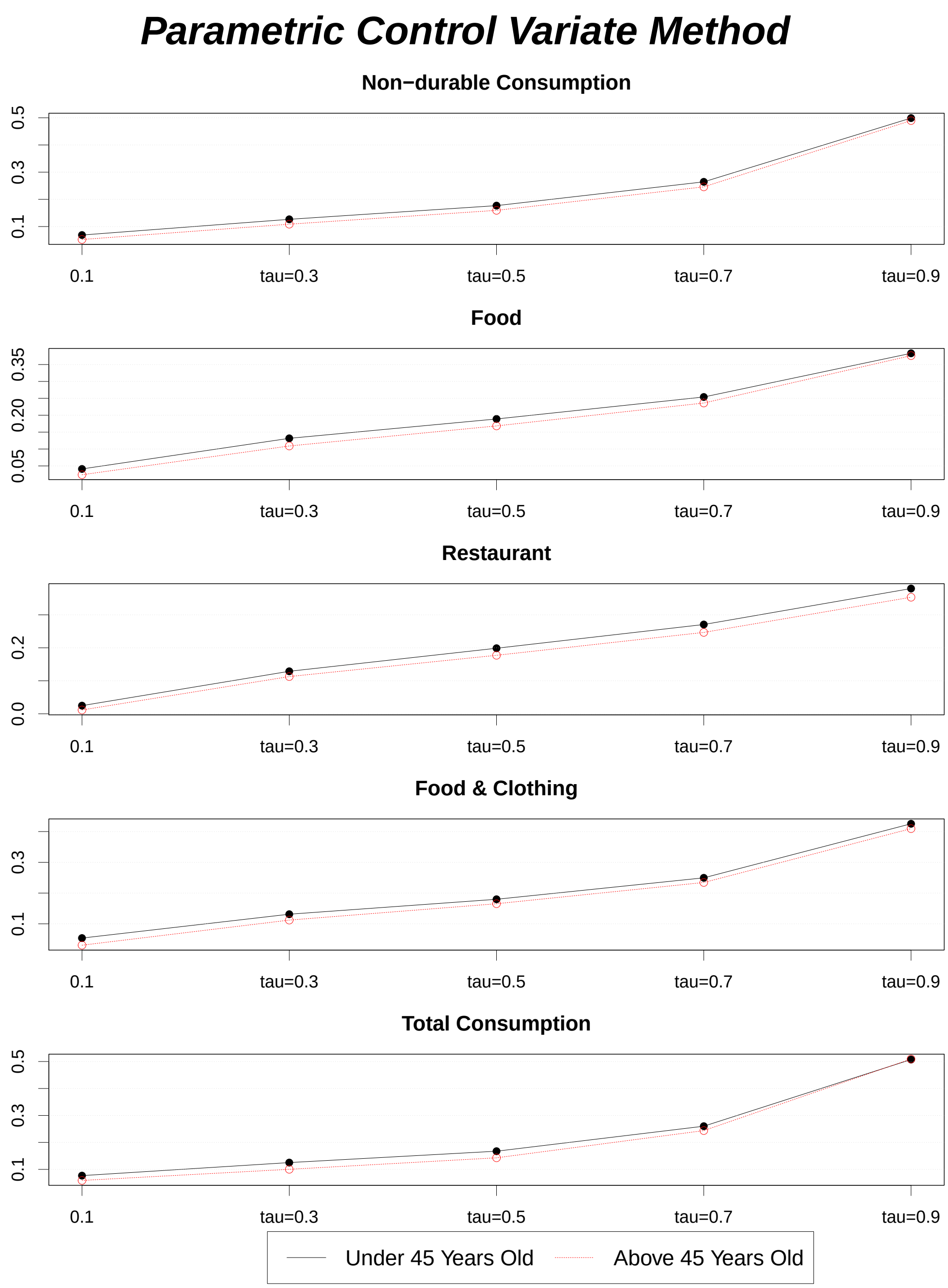

Figure 2.4: Median Saving Rates at Different Quantiles of Predicted Long-Run Income, Estimated By Control Variate Model, 1996.

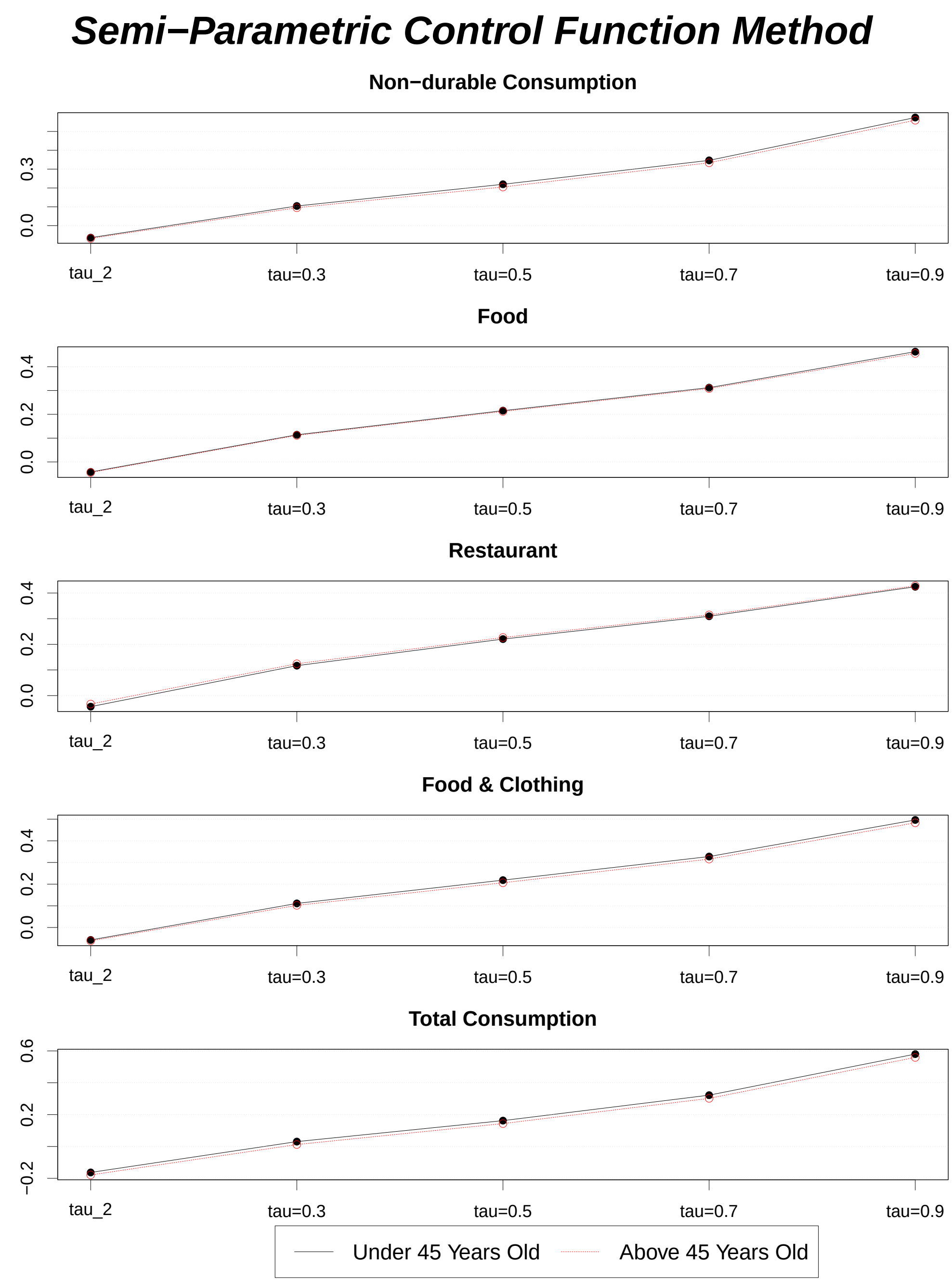

Figure 2.5: Median Saving Rates at Different Quantiles of Predicted Long-Run Income, Estimated By Control Function Model, 1996.

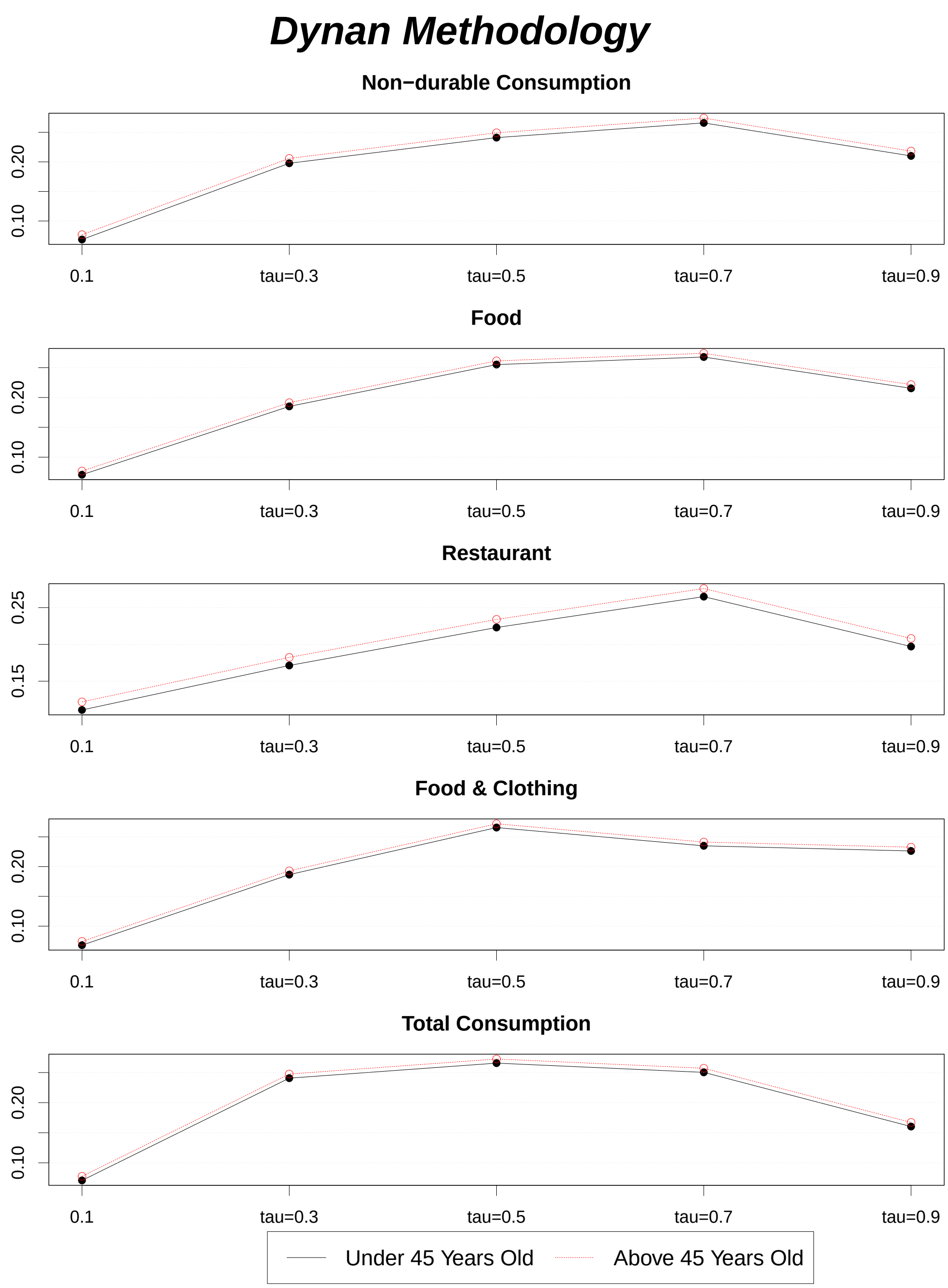

Figure 2.6: Median Saving Rates at Different Quantiles of Predicted Long-Run Income, Estimated By Dynan's Method, 1996.

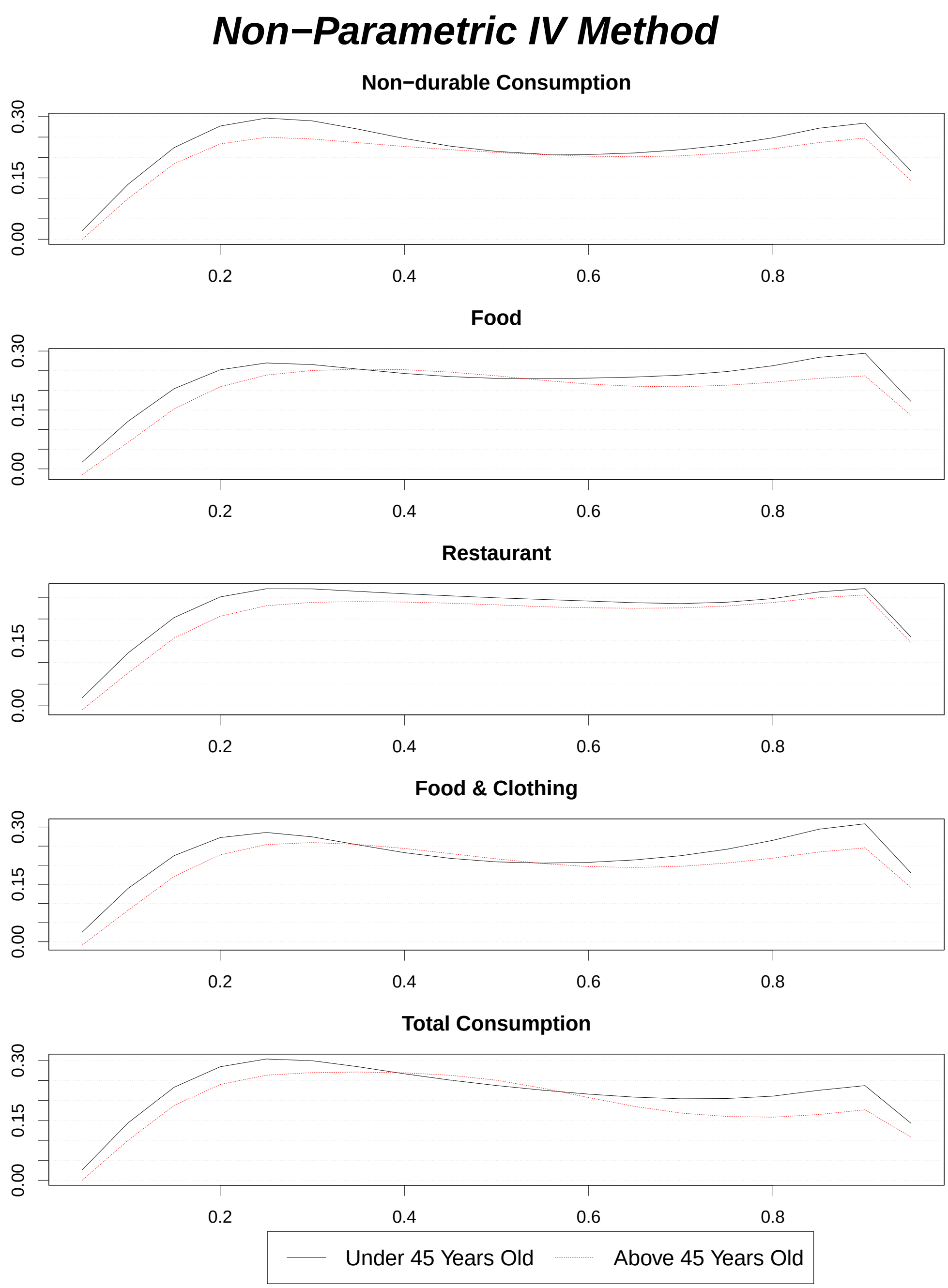

Figure 2.7: Non-Parametric Quantile IV Estimation of Median Saving Rate Through Predicted Long-Run Income Distribution, 1996.

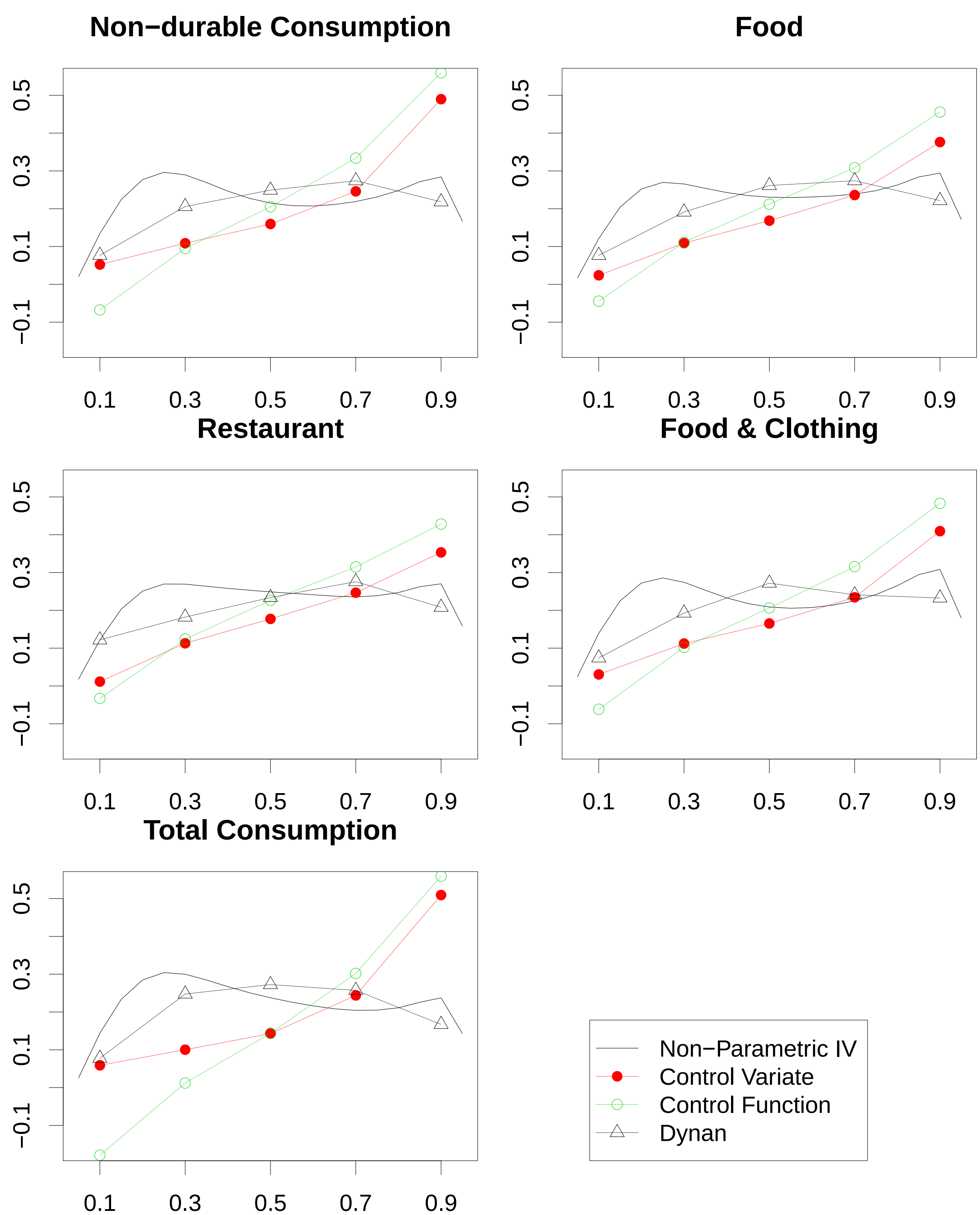

Figure 2.8: Median Saving Rates Estimated at Five Quintiles of Predicted Log-Run Income, Under 45 Years Old, 1996.

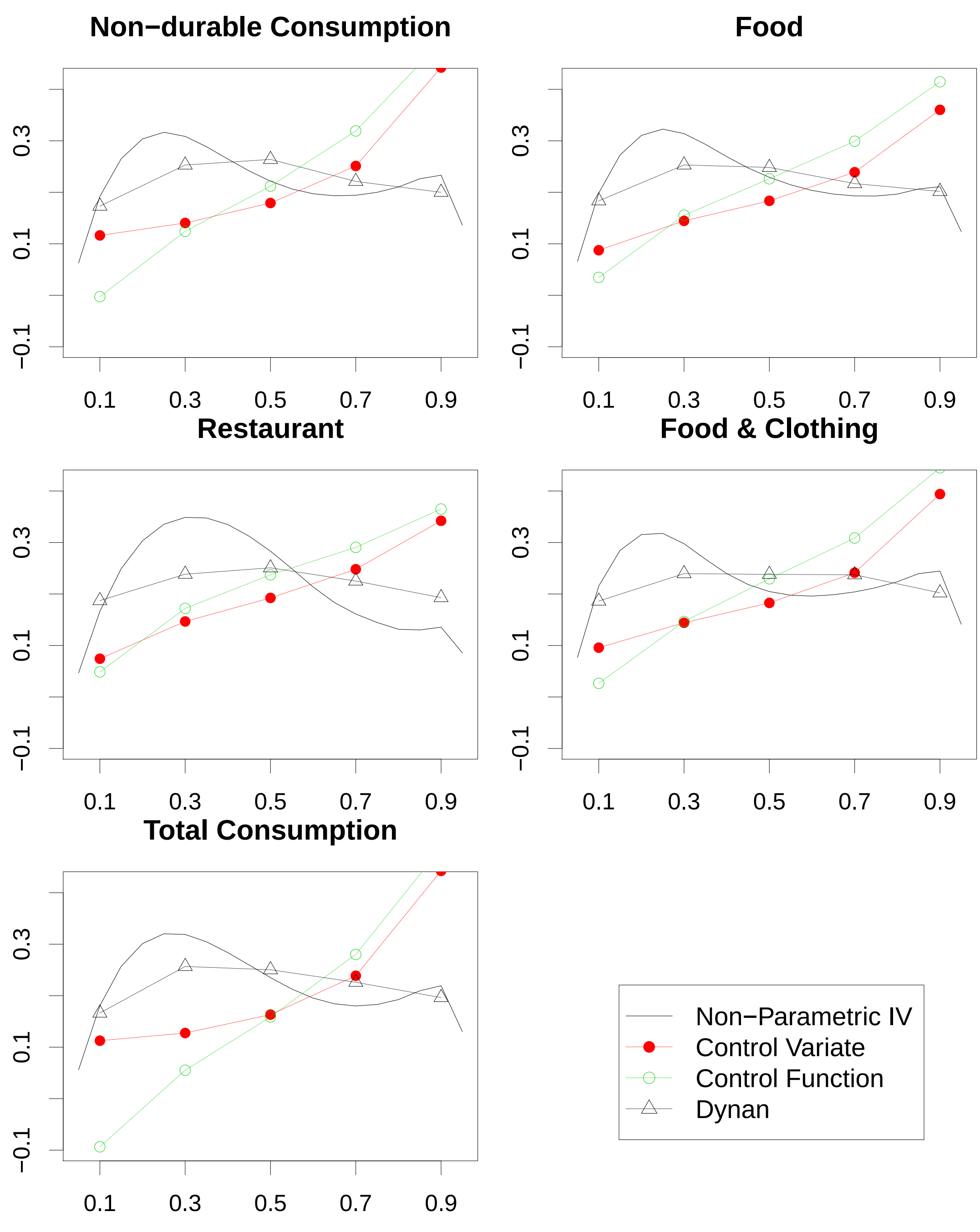

Figure 2.9: Median Saving Rates Estimated at Five Quintiles of Predicted Log-Run Income, Under 45 Years Old, 1982.

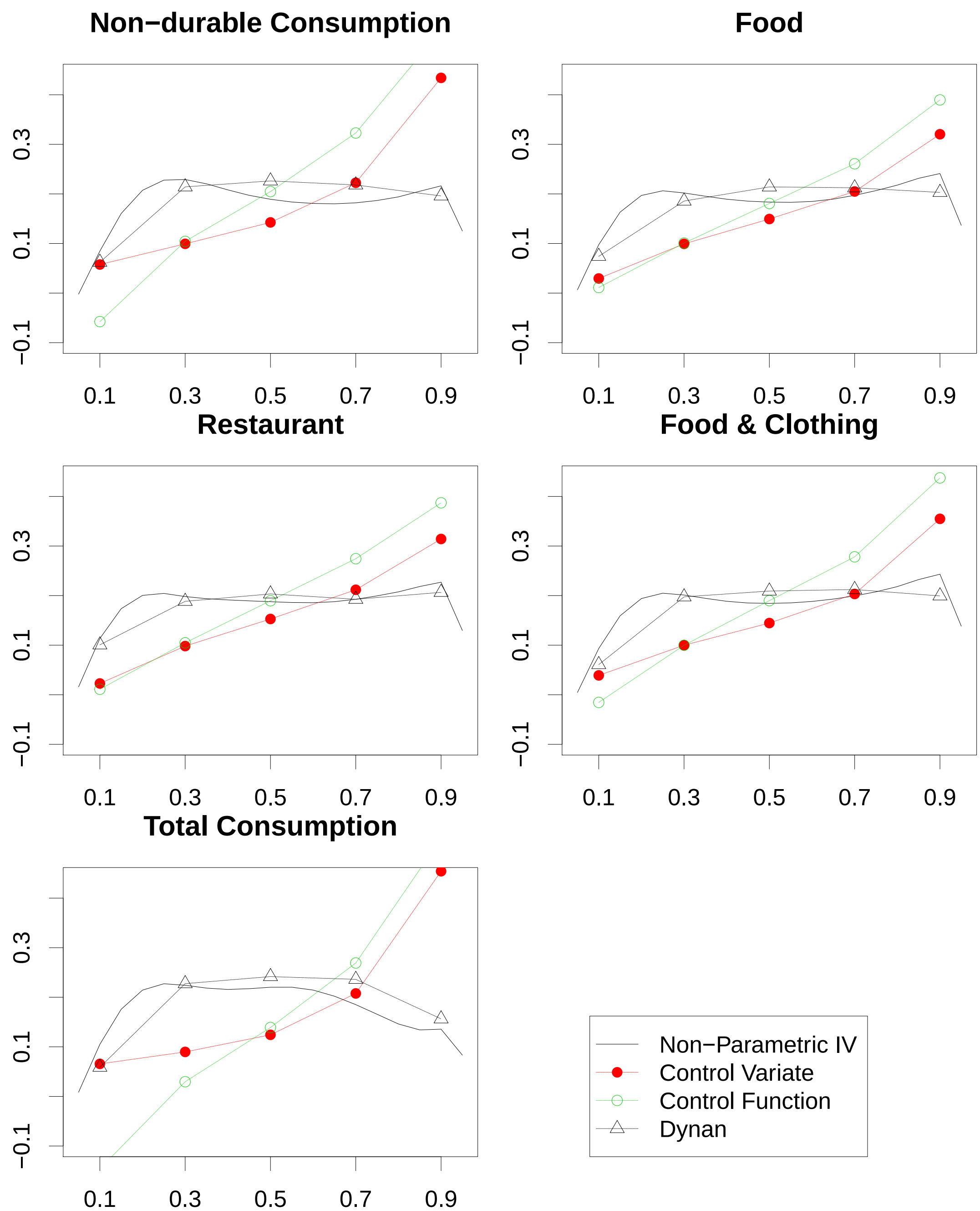

Figure 2.10: Median Saving Rates Estimated at Five Quintiles of Predicted Log-Run Income, Under 45 Years Old, 2001.

Table 2.4: Monte Carlo Experiment, N = 1000, R = 1000.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Bias					
Control Variate	0.0591	0.1532	0.1984	0.2188	0.2409
Control Function	0.0930	0.1649	0.1838	0.1647	0.0920
Variance					
Control Variate	0.0032	0.0012	0.0008	0.0007	0.0007
Control Function	0.0020	0.0009	0.0008	0.0010	0.0020
MSPE					
Control Variate	0.0067	0.0247	0.0402	0.0486	0.0587
Control Function	0.0107	0.0281	0.0346	0.0281	0.0105

Note: Entries are the Bias, Standard Deviation, and Root MSPE for estimation of the slope of $f(Y)$, the function defining the relationship between income and saving rate.

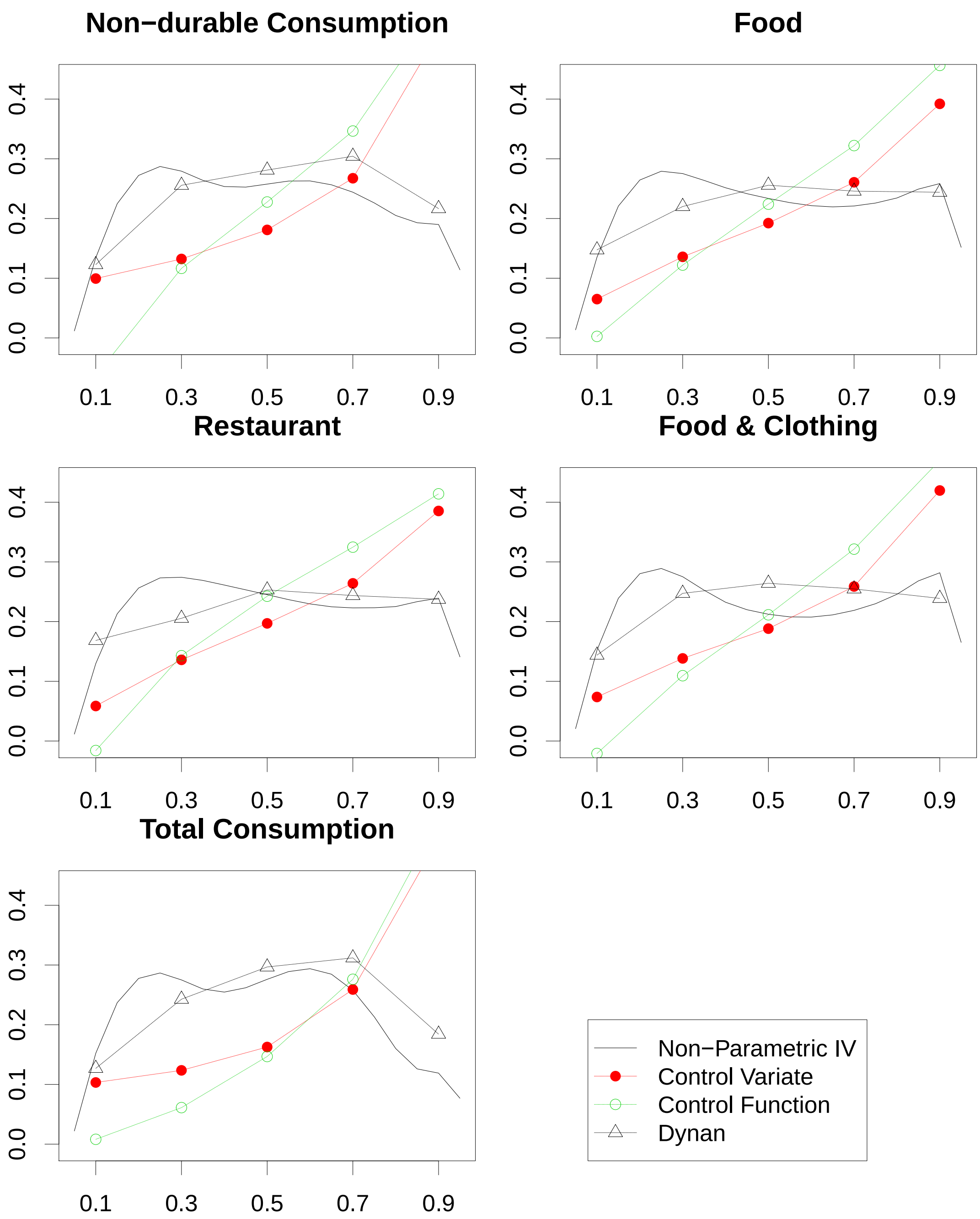

Figure 2.11: Median Saving Rates Estimated at Five Quintiles of Predicted Log-Run Income, Under 45 Years Old, 2008.

Table 2.5: Monte Carlo Experiment, N = 1000, R = 1000.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Bias					
Control Variate	-0.3944	0.0317	0.3175	0.6117	1.0272
Control Function	-0.0344	-0.0227	-0.0281	-0.0258	-0.0344
Standard Deviation					
Control Variate	0.4630	0.3822	0.3503	0.3266	0.3031
Control Function	0.4688	0.4162	0.4086	0.4227	0.4749
Root MSPE					
Control Variate	0.6082	0.3835	0.4728	0.6934	1.0710
Control Function	0.4700	0.4168	0.4096	0.4235	0.4762

Note: Entries are the Bias, Standard Deviation, and Root MSPE for estimation of the slope coefficient β in Equation.2.24.

Table 2.6: Monte Carlo Experiment, N = 1000, R = 1000.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Bias					
Control Variate	0.0533	0.0053	0.0100	0.0154	0.0097
Control Function	-0.1050	-0.0613	-0.0388	-0.0241	-0.0469
Variance					
Control Variate	0.5462	0.3450	0.3007	0.2691	0.2935
Control Function	1.0615	0.5809	0.5023	0.4170	0.4784
MSPE					
Control Variate	0.5488	0.3450	0.3008	0.2695	0.2937
Control Function	1.0666	0.5841	0.5038	0.4177	0.4807

Note: Entries are the Bias, Standard Deviation, and Root MSPE for estimation of the slope coefficient β_1 in Equation.2.27.

Table 2.7: Predicted Median Saving Rate at Different Quintiles of Predicted Long-Run Income, proxied by "**Non-Durable Consumption**", 1996.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Control Variate	5.23***	10.87***	15.96***	24.6***	48.99***
	(1.27)	(0.97)	(0.87)	(0.88)	(1.54)
Age < 45	1.64	1.79*	1.73*	1.84*	0.91
	(1.09)	(1.04)	(1.03)	(1.03)	(0.99)
Control Function	-7.09***	9.72***	21.4***	33.86***	57.33***
	(1.43)	(0.78)	(0.77)	(0.96)	(1.81)
Age < 45	1.28	0.35	0.14	0.27	-0.9
	(0.97)	(0.76)	(0.82)	(1.03)	(1.31)
Non-Parametric IV	9.93***	24.56***	21.23***	20.38***	24.78***
	(2.28)	(2.15)	(2.93)	(1.5)	(3.26)
Age < 45	3.5	4.39*	0.25	1.52	3.63
	(3.16)	(2.63)	(3.24)	(1.78)	(3.57)
Dynan	7.69***	20.58***	24.92***	27.39***	21.82***
	(1.26)	(1.64)	(1.74)	(1.28)	(1.24)
Age < 45	-0.81	-0.81	-0.81	-0.81	-0.81
	(1.05)	(1.05)	(1.05)	(1.05)	(1.05)

1. Sample Size: 6444 Households

2. Standard errors based on 999 bootstrap replications

3. ***p ¡ 0.01, **p ¡ 0.05, *p ¡ 0.1

Table 2.8: Predicted Median Saving Rate at Different Quintiles of Predicted Long-Run Income, proxied by "**Food Expenditures**", 1996.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Control Variate	2.4**	10.91***	16.84***	23.62***	37.64***
	(1.06)	(0.76)	(0.79)	(0.87)	(1.4)
Age $<$ 45	1.72	2.26**	2.06*	1.78	0.69
	(1.08)	(1.07)	(1.08)	(1.12)	(1.13)
Control Function	-4.83***	11.22***	21.71***	31.76***	46.68***
	(1.69)	(1.02)	(0.85)	(0.9)	(1.69)
Age $<$ 45	1.01	0.05	-0.41	-1.12	-1.42
	(1.03)	(0.94)	(0.78)	(0.96)	(1.09)
Non-Parametric IV	6.71**	25.1***	23.7***	20.88***	23.67***
	(2.72)	(1.75)	(2.96)	(1.8)	(3.47)
Age $<$ 45	5.4	1.45	-0.66	2.99	5.76
	(3.49)	(2.64)	(3.43)	(2.34)	(4.04)
Dynan	7.67***	19.13***	26.24***	27.42***	22.18***
	(1.45)	(1.48)	(1.45)	(1.4)	(1.37)
Age $<$ 45	-0.64	-0.64	-0.64	-0.64	-0.64
	(1.17)	(1.17)	(1.17)	(1.17)	(1.17)

1. Sample Size: 6444 Households
2. Standard errors based on 999 bootstrap replications
3. ***p ¡ 0.01, **p ¡ 0.05, *p ¡ 0.1

Table 2.9: Predicted Median Saving Rate at Different Quintiles of Predicted Long-Run Income, proxied by "**Restaurant Expenditures**", 1996.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Control Variate	1.16	11.29***	17.73***	24.68***	35.35***
	(1.04)	(0.85)	(0.81)	(0.83)	(1.06)
Age < 45	1.31	1.6	2.14**	2.39**	2.6**
	(1)	(1.04)	(1.07)	(1.05)	(1.08)
Control Function	-4.28***	11.62***	22.82***	32.73***	44.33***
	(1.38)	(1.12)	(0.97)	(1.06)	(1.56)
Age < 45	0.25	0.08	-0.88	-2.6**	-3.26**
	(1.09)	(1.01)	(1.01)	(1.08)	(1.49)
Non-Parametric IV	7.52***	23.89***	23.29***	22.55***	25.53***
	(2.69)	(1.62)	(2.26)	(1.83)	(3.73)
Age < 45	4.66	3.05	1.6	1.01	1.49
	(3.62)	(2.31)	(2.72)	(2.51)	(4.19)
Dynan	12.19***	18.23***	23.4***	27.61***	20.81***
	(1.86)	(1.66)	(1.49)	(1.68)	(1.54)
Age < 45	-1.1	-1.1	-1.1	-1.1	-1.1
	(1.21)	(1.21)	(1.21)	(1.21)	(1.21)

1. Sample Size: 6444 Households
2. Standard errors based on 999 bootstrap replications
3. ***p ¡ 0.01, **p ¡ 0.05, *p ¡ 0.1

Table 2.10: Predicted Median Saving Rate at Different Quintiles of Predicted Long-Run Income, proxied by "**Food and Clothing Expenditures**", 1996.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Control Variate	3.06***	11.22***	16.54***	23.45***	40.96***
	(1.07)	(0.99)	(0.94)	(0.89)	(1.24)
Age $<$ 45	2.32**	1.91	1.43	1.51	1.58
	(1.15)	(1.23)	(1.25)	(1.29)	(1.2)
Control Function	-6.18***	10.22***	21.33***	32.93***	49.08***
	(1.94)	(0.85)	(0.91)	(0.88)	(1.64)
Age $<$ 45	-0.1	0.84	0.05	-0.9	-0.75
	(1.08)	(0.93)	(0.94)	(1.02)	(1.12)
Non-Parametric IV	8.18***	25.97***	21.68***	19.69***	24.54***
	(2.18)	(1.92)	(1.84)	(1.49)	(1.96)
Age $<$ 45	5.76*	1.44	-0.8	2.8	6.31*
	(3.25)	(2.51)	(2.65)	(2.05)	(3.35)
Dynan	7.41***	19.24***	27.31***	24.11***	23.27***
	(1.63)	(1.55)	(1.51)	(1.34)	(1.62)
Age $<$ 45	-0.62	-0.62	-0.62	-0.62	-0.62
	(1.26)	(1.26)	(1.26)	(1.26)	(1.26)

1. Sample Size: 6444 Households
2. Standard errors based on 999 bootstrap replications
3. ***p ¡ 0.01, **p ¡ 0.05, *p ¡ 0.1

Table 2.11: Predicted Median Saving Rate at Different Quintiles of Predicted Long-Run Income, proxied by "**Total Consumption**", 1996.

	$\tau = 0.1$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.9$
Control Variate	5.65***	10.03***	14.29***	24.39***	50.94***
	(1.21)	(1)	(0.78)	(0.78)	(1.32)
Age < 45	1.96*	2.5**	2.45***	1.63*	-0.2
	(1.04)	(1.09)	(0.94)	(0.88)	(0.98)
Control Function	-18.39***	1.68	16.01***	32.31***	60.07***
	(1.82)	(1.28)	(0.88)	(1.24)	(2.32)
Age < 45	2.51**	1.16	-0.39	-1.82	-5.18***
	(1.17)	(1.12)	(0.99)	(1.15)	(1.77)
Non-Parametric IV	10.04***	27.02***	25.08***	16.85***	17.69***
	(1.95)	(2.5)	(3.19)	(2.03)	(3.62)
Age < 45	4.36	2.96	-1.33	3.59	6.05
	(2.92)	(3.33)	(3.79)	(2.81)	(4.38)
Dynan	7.85***	24.63***	27.25***	25.72***	16.71***
	(1.37)	(1.85)	(1.45)	(1.7)	(1.81)
Age < 45	-0.69	-0.69	-0.69	-0.69	-0.69
	(1.15)	(1.15)	(1.15)	(1.15)	(1.15)

1. Sample Size: 6444 Households
2. Standard errors based on 999 bootstrap replications
3. ***p ¡ 0.01, **p ¡ 0.05, *p ¡ 0.1

Chapter 3

Aggregate Determinants of Household Insolvency across Canadian Provinces, A Panel-VAR Approach

3.1 Introduction

Almost every quarter since the early 1990s, Canadians have experienced a new high record in their debt levels. It became more concerning when the debt-to-disposable income ratio approached the level similar to the level in the US when the recent great recession started in 2007. Debt-to-disposable income in that time was about 160% in the US. In 1990, the debt-to-disposable income ratio for Canadian households was 98% , while in the first quarter of 2016 the ratio was 165.3%.[1] This means that for every dollar of disposable income, households owed 1.65 dollars in debt. Accordingly, various policy institutions have raised concerns about increased levels of household debt. While some researchers claim that Canada is in the same amount of risk as the US in pre-recession period, others argue that Canadian households' net worth has risen and higher levels of debt are associated with higher gains in assets and income.

Overall, credit is a new tool and with the expansion of financial markets, borrowing against future income has become easier and easier. There is not enough historical observation to determine how much debt is considered good, or what the optimized level of debt is for households based on their demographic and financial characteristics. Confronting this issue is beyond the scope of this article. However, we try to address one of the side-effects of this phenomenon which is the increased number of insolvent households.

The ability to borrow against the future flow of income can boost current aggregate demand and lead to faster growth through higher current consumption expenditure. However, increased levels of debt make households more vulnerable and in the case of an adverse income shock, households will fall into bankruptcy easier. Historical data from Canada depicted in Figure.3.1 shows the cyclical co-movements –Deviation from the long-run trend– of the insolvency rate and debt to disposable income ratio. This can be seen in Figure.3.1 in national level and in Figure.3.6 in provincial levels. As levels of debt have increased in the past two decades, the cor-

[1] CANSIM Table 380-0073.

relation between unemployment rate and insolvency rate has also increased.

In Figure.3.7 We can see the negative correlation of insolvency rate and log GDP at provincial levels. Using national, provincial data and household level data in Canada, Fieldhouse et al. (2012) investigates the key factors that affect cyclical movements of consumer insolvency rate. They explore both demand and supply side and count adverse income shocks, and unemployment fluctuations, the main factor on the demand side. On the supply side, tightened standards for lending as well as the higher interest rate for riskier borrowers, makes it more costly for households to access credit and rollover existing debts. Using micro-level data, they provide demographic information on households who file bankruptcy or proposal such as age, marital status, household size, and income level.

Similarly, Garrett and Wall (2010) documents the relationship between the unemployment rate and personal bankruptcy using state-level data from the US. They find a countercyclical pattern in bankruptcy, and bankruptcy filing peaks close to the end of the recession and tends to decline slowly after the recession.

Using the Survey of Financial Security in Canada, MacGee (2012) draws comparisons between insolvency rate in the United States and Canada and concludes that "current levels of consumer debt offer cause for concern, but not panic". He suggests that in the most likely scenario, household debt levels must remain manageable, yet these high levels of debt make Canadian households susceptible to significant negative shocks that might result in a higher interest rate or lower income.

Meh et al. (2009) also uses Survey of Financial Security and investigates household debt in Canada. One of the key findings of this study is that between 1999 to 2005, the sensitivity of households to negative income shocks, interest rate changes and especially housing market movements, due to higher share of mortgages in the household budget, has increased significantly. It is suggested that due to higher debt to income ratio and generally higher debt burden, households in lower income brackets must be monitored more cautiously.

Livshits et al. (2015) in a DSGE model with a heterogeneous agent life-cycle an competitive financial intermediaries, finds that income uncertainty and adverse income shock alone cannot explain the rise in the bankruptcy rate in the US. Instead, they purport that this is a reflection of changes in financial market environments. They show that a lower cost of lending and decline in the cost of bankruptcy has had a large contribution to the rise in consumer bankruptcy rate.

The effects of economic recessions on household financial stability is studied by Jakubík (2014). In a macro model with a small open economy, representing a European economy, he asserts that a significant additional decline in consumer consumption is observed in recessions, which is a result of higher household insolvency and default rate. The key finding is that the insolvency rate greatly matters in policy implications for the financial section and cyclical fluctuations of the macroeconomy.

In this study, our final goal is to investigate the effects and consequences of higher financial vulnerability of households –as measured by insolvency rate– on the cyclical behavior of GDP. In other words we want to draw connections between the cyclical behavior of Insolvency rate and GDP across provinces, while controlling for other related variables.

As Figure.3.2 shows cyclical movements of insolvency rate and unemployment rate are strongly correlated. This correlation is shown at provincial level in Figure.3.6, and it can be seen that the insolvency rate rise and fall behind the unemployment rate. On the other hand we know that both variables are highly correlated with GDP. Common sense suggests that GDP growth reduces both the unemployment rate and the insolvency rate. Our main hypothesis is that higher levels of debt that are associated with increased financial instability of households, have a negative impact on the GDP.

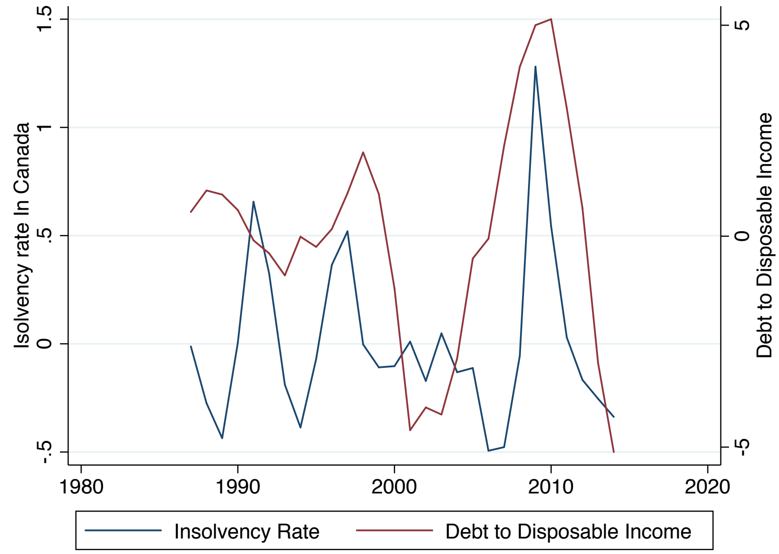

Figure 3.1: National Insolvency rate and Debt to Disposable Income Ratio, Deviation from Long-run Trend.

In our analysis, we also consider the effect of income inequality on financial vulnerability of households. If higher levels of inequality are associated with higher insolvency rates, it can be evidence for trickle down consumption theory and conspicuous consumption hypothesis. This can be explained when middle class and low income families seek a higher economic or social status through borrowing and spending money on luxury goods and services.

Insolvency rate in Canada is depicted in Figure.3.3. It had been quite constant from 1997 to 2007 just less than 4 per 1000 population, and increased by 50% during the 2008 recession to nearly 6 per 1000 population. We use panel regression analysis and Panel Vector Autoregression analysis to test our hypotheses and investigate the relationship between financial stability of households and GDP. We use panel regression analysis to find out the related variables to insolvency rate, in order to control for them in our panel-VAR model. Our findings show that unemployment rate, debt-

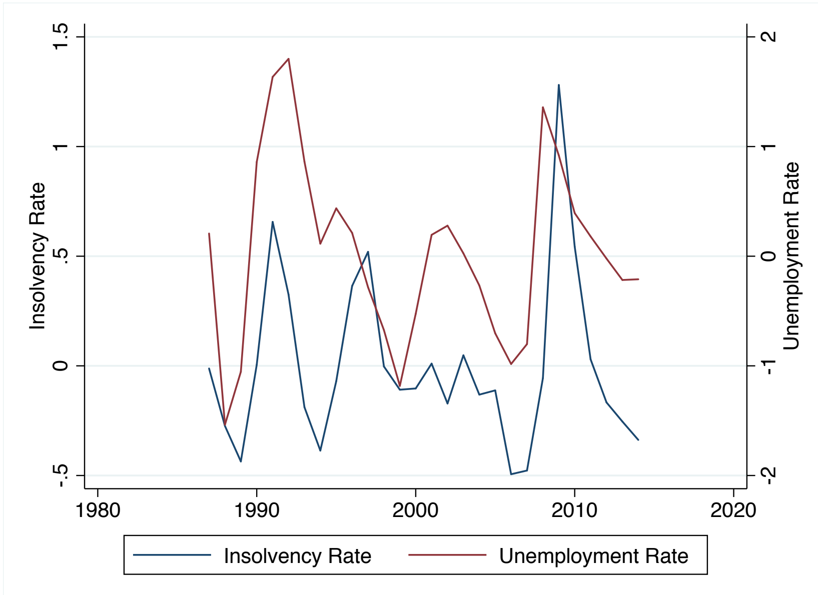

Figure 3.2: National Insolvency Rate and Unemployment Rate, Deviation from Long-run Trend

Figure 3.3: National Insolvency Rate, per 1000 people in population.

to-disposable income ratio, housing price index and consumer loan interest rate have positive effect on insolvency rate. on the other hand, GDP and non-housing prices are negatively correlated with insolvency rate. From the results of panel vector auto regression, we find that insolvency rate has a small, yet significant effect on cyclical movements of GDP.

The rest of this paper is organized as follows. The ext section provides a brief description of the datasets used in this study. Following the dataset description section.4 provides empirical methodology and results of our analysis and finally Section.5 will conclude with a discussion.

3.2 Aggregate Data

In the case of insolvency, Canadian Households have two options. First, they can file for bankruptcy that requires liquidating assets and as a result, unsecured debt will be written off. Secured debt remains in place, and they have to pay their mortgage payments. Other options for filing bankruptcy include, filing a proposal to destruct debt, called devision II proposal. If the majority of creditors approve the proposal, repayment of debt will be restructured. We investigate cyclical movements of the insolvency rate at the provincial level, that is define as the number of insolvent households per 1000 population in that Province. These data span the time period from 1987 to 2014 and are available on an annual basis.[2]

In a series of analyses, several variables are taken into account that are considered to have an effect on insolvency rate or are affected by insolvency rate fluctuations. In particular, we consider cyclical movements of GDP, its growth rate, unemployment rate, housing prices and non-housing prices. Since we intend to focus on the cyclical behavior of these variables, the Hodrick-Prescott filter is used to detrend aggregate

[2]Available at Office of the Superintendent of Bankruptcy Canada

variables and calculate the deviation from long-run trend. All the data are collected from the Statistics Canada website.

Different panel unit root tests are performed to ensure the stationarity of variables. Based on Im-Pesran-Shin, Levin-Lin-Shin and Bruiting unit root tests that test for stationarity of panel data, our detrended series are not integrated and all are stationarity.

3.3 Empirical Results

3.3.1 Simple Correlation Analysis

Before starting regression analysis, we report the contemporaneous and lagged correlation between variables of interest to have a better sense of the shape of the relationship between these variables. Table.3.1 shows the contemporaneous and lagged correlation between all the variables. Unemployment rate has a strong positive correlation with the insolvency rate. The first lag of the unemployment rate also has a positive and smaller correlation, however, the second lag has a strongly negative correlation with the insolvency rate. This can be interpreted as higher unemployment has strong positive effect on insolvency rate at the current period, while the effects of shocks to unemployment rate have negative effect on insolvency rate in next two periods.

Contemporaneous GDP and its first lag are negatively correlated, and the second lag is weakly positively correlated. The first lag of the housing prices has a stronger positive correlation with the insolvency rate than the contemporaneous and other lags. Meaning shocks to housing prices from last period have a stronger effect on the insolvency rate in the current period.

Non-housing prices are negatively correlated, and higher lags have a stronger negative correlation. At the end, we also consider the relationship between the Gini coefficient and the insolvency rate. Surprisingly it has a fairly strong positive correlation with the insolvency rate. This means that higher levels of inequality are associated with a higher insolvency rate.

3.3.2 Panel Regression Analysis

We begin our empirical analysis by a series of random effect panel regressions. The general specification is shown in Equation.3.1. We regress the provincial insolvency rate on the unemployment rate, log GDP, housing prices, non-housing prices, and the debt-to-disposable income ratio as well as provincial Gini coefficient. Based on the results of the Hauman test, the random effect model is the appropriate approach. Using Hedrick-Prescott filter, all variables are detrended and defined in deviation from long-run trend due to our focus on cyclical movements in these variables.

$$INSc_{it} = \alpha_0 + \alpha_1 URc_{it} + \alpha_2 logGDPc_{it} + \alpha_3 HPc_{it} + \alpha_4 NHPc_{it} + \alpha_5 GINIc_{it} + \alpha_6 DDIc_{it} + \epsilon_{it} \tag{3.1}$$

Results of a selection of specifications are reported in Table.3.2. For the matter of convenience of interpretation, Gini Coefficient is defined in a scale of zero to 100, instead of zero to one. It has a positive significant effect on the insolvency rate. Any point increase in Gini coefficient, is associated with an 0.1 increase in the insolvency rate. As we add a second lag of unemployment and log GDP, its effect reduces to almost half.

The effect of debt-to-disposable income ratio is quite consistent across different specifications. It has a strong and consistently significant positive effect. With every dollar increase in the debt-to-disposable income ratio, the insolvency rate increases by 0.03. Since 1990, the debt-to-disposable income has increased from 89.3 to 165.3

Table 3.1: Contemporaneous and Lagged Correlation between Variables.

	INSc	GINIc	URc	URc (-1)	URc (-2)	GDPc	GDPc (-1)	GDPc (-2)	Hpc	Hpc (-1)	Hpc (-2)	NHPc	NHPc (-1)
INSc	1												
GINIc	0.0857	1											
URc	0.353	-0.0383	1										
RUC (-1)	0.0216	-0.0767	0.5343	1									
RUC (-2)	-0.2333	-0.0665	0.0514	0.5312	1								
GDPc	-0.3355	-0.0321	-0.5535	-0.3969	-0.2201	1							
GDPc (-1)	-0.1468	-0.0102	-0.2307	-0.5577	-0.3929	0.5214	1						
GDPc (-2)	0.0995	-0.028	0.2229	-0.2294	-0.5577	0.047	0.5225	1					
HPc	0.0164	-0.0651	0.1811	-0.1056	-0.2332	0.1298	0.4529	0.487	1				
HPc (-1)	0.2283	-0.0882	0.5722	0.1833	-0.1076	-0.2776	0.1223	0.4542	0.5958	1			
HPc (-2)	0.1485	-0.103	0.5695	0.5728	0.1815	-0.3562	-0.2833	0.1245	0.1316	0.5962	1		
NHPc	-0.0691	-0.0387	0.271	0.0906	-0.0771	-0.0469	0.1788	0.266	0.4769	0.2662	0.0896	1	
NHPc (-1)	-0.0937	-0.0493	0.4166	0.2748	0.0883	-0.2699	-0.0504	0.1774	0.3151	0.4807	0.2704	0.6106	1
NHPc (-2)	-0.1534	-0.1046	0.3317	0.4182	0.2727	-0.2538	-0.2734	-0.05	0.1218	0.317	0.482	0.1367	0.6123

in 2016. According to our analyses, the pure effect of this increase has been about 2.3 increase in the insolvency rate.

Housing price index and non-housing prices have different effects on the insolvency rate. Even though common sense suggests that higher house prices have a positive wealth effect and should reduce insolvency rate, our findings prove otherwise and shows a positive relationship between housing prices and insolvency rate. This is consistent with findings of other researchers in the literature.[3] The effect of non-housing price index is significant and negative. This is consistent with the findings in the first chapter. Higher prices of core items, tightens households budget and forces more careful spending.

The coefficient of provincial unemployment rate is very consistent and significant across different specifications. Log GDP has the largest effect on the insolvency rate. Consistent with our expectations, log GDP has a negative effect on insolvency rate and with every 1% growth in gross domestic product, the insolvency rate decreases by about 5.5 units. Adding the consumer loan rate to the model, does not have any significant effect on other coefficients, while it increases the R^2 significantly. However the housing price index coefficient decreases and becomes insignificant. Higher interest rates are associated with a higher insolvency rate.

In various specifications that we consider, in addition to contemporaneous unemployment rate and contemporaneous log GDP, based on overall R^2 and significance level, we consider the second lag of the unemployment rate and log GDP as well. They are both statistically significant and help improve the prediction power of the model. At the end, we consider the contemporaneous effect of housing prices and non-housing prices, and by adding them to the model, even though they are both insignificant, the first lag of housing prices turns significant.

In Table.3.3 we add two dummy variables for two major recessions that occurred in Canada in early 1991 and in 2008. Across different specifications, two dummies are signifiant, and positive. The key result is the coefficient of 2008 recession is much larger, specially in specifications (1) and (2), that is before controlling for debt level. This difference expresses the effect of higher levels of debt on financial stability of households in recessions.

[3] See Fieldhouse et al. (2012).

Table 3.2: Dependent Variable: Provincial Insolvency Rate.

	(1)	(2)	(3)	(4)
Gini Coefficient	0.100***	0.0917***	0.0526*	0.0547**
	(3.62)	(3.35)	(1.9)	(1.97)
Debt to Disposable Income Ratio	0.0285***	0.0308***	0.0320***	0.0340***
	(3.31)	(3.62)	(3.87)	(3.95)
Unemployment Rate	0.131***	0.132***	0.135***	0.125***
	(3.5)	(3.59)	(3.79)	(3.41)
log GDP	-4.890***	-5.549***	-6.423***	-6.097***
	(-4.31)	(-4.88)	(-5.67)	(-5.20)
House Price Index L1	0.0357**	0.0249	0.0222	0.0402*
	(2.01)	(1.4)	(1.19)	(1.82)
Non Housing Prices L1	-0.113***	-0.110***	-0.0959***	-0.0998***
	(-5.04)	(-4.97)	(-4.45)	(-3.71)
Consumer Loan Rate		0.0325***	0.0236**	0.0290**
		(3.08)	(2.25)	(2.39)
Unemployment Rate L2			-0.159***	-0.155***
			(-4.79)	(-4.64)
log GDP L2			-2.462**	-2.097*
			(-2.03)	(-1.69)
Housing Prices				-0.0322
				(-1.56)
Non-Housing Prices				0.016
				(0.56)
Overall R^2	0.307	0.331	0.384	0.389
Within R^2	0.307	0.331	0.384	0.39
Between R^2	0.190	0.111	0.0293	0.0048

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 3.3: Dependent Variable: Provincial Insolvency Rate.

	(1)	(2)	(3)	(3)
Gini Coefficient	0.0713***	0.0709***	0.0451*	0.0477*
	(2.8)	(2.79)	(1.74)	(1.83)
Unemployment Rate	0.126***	0.131***	0.130***	0.118***
	(3.71)	(3.84)	(3.89)	(3.46)
log GDP	-4.080***	-4.270***	-5.389***	-5.048***
	(-4.08)	(-4.25)	(-5.04)	(-4.58)
House Price Index L1	0.00314	-0.00185	-0.000541	0.0194
	(0.19)	(-0.11)	(-0.03)	(0.92)
Non Housing Prices L1	-0.101***	-0.0986***	-0.0850***	-0.0915***
	(-4.96)	(-4.87)	(-4.20)	(-3.55)
1991 Recession	0.470***	0.357***	0.348***	0.345**
	(4.57)	(2.82)	(2.79)	(2.49)
2008 Recession	0.925***	0.940***	0.692***	0.700***
	(7.5)	(7.63)	(5.96)	(5.87)
Consumer Loan Rate		0.0183	0.0135	0.0191
		(1.52)	(1.13)	(1.53)
Debt to Disposable Income Ratio			0.0159*	0.0182**
			(1.94)	(2.18)
Unemployment Rate L2			-0.118***	-0.113***
			(-3.72)	(-3.56)
log GDP L2			-2.496**	-2.109*
			(-2.20)	(-1.82)
Housing Prices				-0.0357*
				(-1.85)
Non-Housing Prices				0.0218
				(0.73)
OverallR^2	0.430	0.435	0.468	0.475
Within R^2	0.430	0.435	0.468	0.475
BetweenR^2	0.0528	0.0318	0.0988	0.059

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

3.3.3 Panel Vector Autoregressive Analysis

In this section, using a Panel-VAR approach, we investigate the relationship between the insolvency rate and GDP in Canada. In panel regression analysis, in the previous section, we saw the effect of several variables on the insolvency rate. In a more dynamic approach, we want to explore the interrelation of the insolvency rate and GDP to see if insolvency rate has any effect on cyclical movements of GDP. The general framework of the model is as in Equation.3.2.

$$\begin{bmatrix} INSc_{it} \\ lGDPc_{it} \end{bmatrix} = \Gamma \begin{bmatrix} INSc_{it-1} \\ lGDPc_{it-1} \end{bmatrix} + \Psi X_{it} + \begin{bmatrix} \epsilon_{1it} \\ \epsilon_{2it} \end{bmatrix} \tag{3.2}$$

This specification represents a panel vector autoregressive of order one, PVAR(1), Γ is a 2x2 matrix of autoregressive coefficients, and Ψ is a 2xk matrix, where k is the number of exogenous variables. X_t is a kx1 vector of exogenous variables and includes; unemployment rate, housing prices, non-housing prices, Gini coefficient, and consumer loan rate. We investigate their effect on both insolvency rate and log GDP.

All variables are detrended and are defined as deviation from long-run trend. In order to avoid any co-integration issue, and to make sure the system is stationary, different panel unit root tests such as “Levin-Lin-Chu”, “Im-Pesaran-Shin”, and “Harris-Tzavalis” unit-root test have been applied. All variables in their detrended version are proved to be stationary and no evidence of unit root was found. Summary of unit root tests are reported in Appendix.3.4.

Results of the estimation of (3.2) are reported in Table.3.4. The optimal lag order is chosen based on BIC and AIC criteria. As it is expected, log GDP has a negative effect on the insolvency rate and the same results for the unemployment rate. Consistent with our analysis in the panel regression section, housing prices have a positive effect on the insolvency rate, while higher non-housing prices reduce insolvency rate. However, Gini coefficient and debt-to-disposable income ratio do not have any significant effect on the insolvency rate.

The main result from this set of analyses is the negative coefficient of the lag insolvency rate in the GDP equation. It is small, yet significant and conveys the message that higher insolvency rate –higher rates of the vulnerability of households– lowers the GDP growth rate. This is best seen in the impulse response function derived from this estimation in Figure.3.4. The IRF shows that with a one-time shock

Table 3.4: Panel Vector Autoregressive Estimation

		Coef.	Std.Err.	z	P<z	[95% Conf.	Interval]
Insolv Rate							
	Insolv Rate						
	L1.	0.3876***	0.0638	6.08	0.000	0.2625	0.5127
	log GDP						
	L1.	-2.5417*	1.5143	-1.68	0.093	-5.509	0.4264
	Unemployment	0.1323***	0.0436	3.03	0.002	0.04682	0.2178
	Unemployment L2	-0.1500***	0.0410	-3.65	0.000	-0.2305	-0.0695
	HP L1	0.0500**	0.0246	2.03	0.043	0.0016	0.0984
	NHP L1	-0.0816***	0.0217	-3.75	0.000	-0.1243	-0.0389
	Gini Coeff	0.0392	0.0300	1.31	0.191	-0.0196	0.0980
	Debt to Income	0.0167	0.0104	1.59	0.111	-0.0038	0.0372
	Cons Loan Rate	0.0015	0.0163	0.09	0.925	-0.0305	0.0336
log GDP							
	Insolv Rate						
	L1.	-0.0123***	0.0036	-3.34	0.001	-0.0195	-0.0050
	log GDP						
	L1.	0.5736***	0.0755	7.59	0.000	0.42554	0.7218
	Unemployment	-0.0084***	0.0017	-4.79	0.000	-0.0119	-0.0049
	Unemployment L2	0.0001	0.0017	0.08	0.937	-0.0033	0.0036
	HP L1	0.0004	0.0011	0.40	0.691	-0.0018	0.0020
	NHP L1	-0.0003	0.001	-0.27	0.784	-0.0025	0.0019
	Gini Coeff	0.0014	0.0017	0.86	0.390	-0.0018	0.0048
	Debt to Income	0.0008	0.0004	1.62	0.106	-0.0001	0.0017
	Cons Loan Rate	0.0008	0.0008	1.04	0.296	-0.0007	0.0020

to insolvency rate, one unit increase only for one period, reduces the log GDP for about 1.2% in the following period and it takes the system about 15 periods that the effect of this shock is fully eliminated.

On the other hand, the effect of a one-time shock to log GDP on insolvency rate is much larger and it reduces the insolvency rate by about 2.5 units and the effect diminishes and disappears after about 15 periods. This IRF is shown in Figure.3.5. Other than unemployment rate, none of the exogenous variables have any significant effect on the log GDP.

3.4 Summary

In this article, we studied the relationship of several aggregate variables with financial stability of households, at the provincial level. We used panel data of the provincial consumer insolvency rate as a measure of financial vulnerability, and investigated the effect of income inequality, household debt-to-income ratio, log GDP and unemployment rate on the number of insolvent households per year.

Consistently from both panel regression analysis and panel vector autoregression analysis, we observed a negative effect from log GDP and housing price index on Households' insolvency rate, while the effect of unemployment rate, income inequality, consumer loan interest rate and household debt-to-income ratio is positive. The key finding of this study is the negative impact of higher insolvency rate on log GDP. Our empirical results show that a higher insolvency rate has a negative impact on GDP and it lowers the growth rate of GDP in future periods.

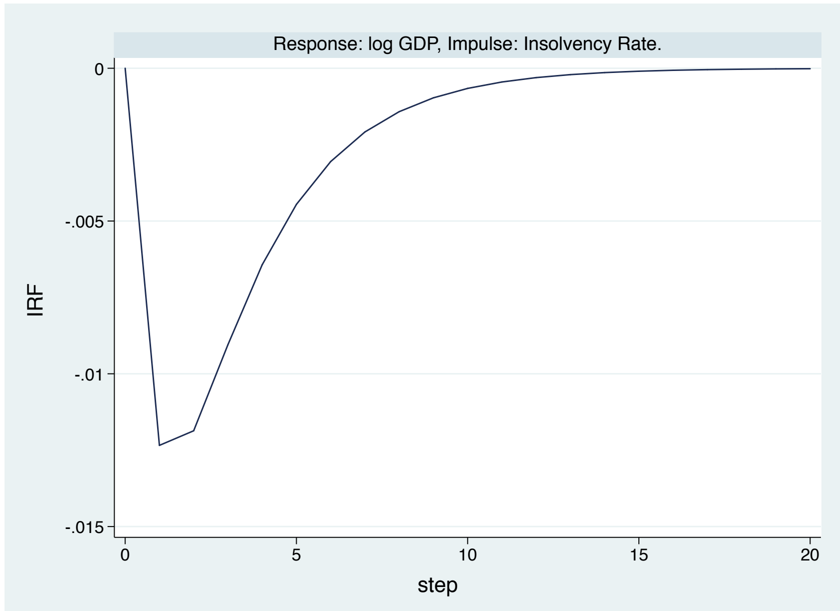

Figure 3.4: Response of Insolvency Rate to a one time positive shock to GDP

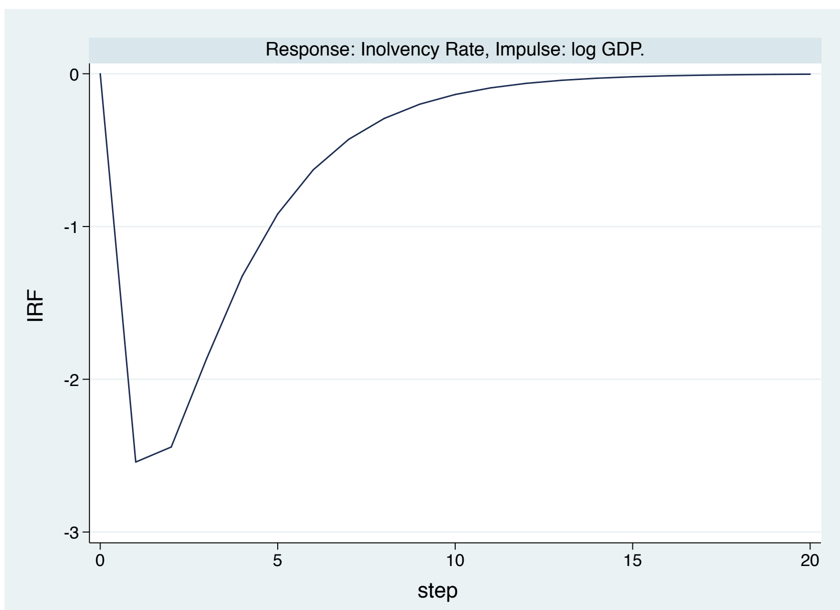

Figure 3.5: Response of GDP to a one time positive shock to Insolvency Rate

Figure 3.6: Provincial Insolvency Rate and unemployment Rate, Deviation from Long-Run Trend.

Figure 3.7: Provincial Insolvency Rate and log GDP, Deviation from Long-Run Trend.

Bibliography

Alan, S., Atalay, K., and Crossley, T. F. (2015). Do the rich save more? evidence from Canada. *Review of Income and Wealth 61.4 (2015): 739-758.*

Allen, J. and Damar, E. (2012). Household Insolvency in Canada. *Bank of Canada Review*, pages 43–55.

Bailliu, J., Kartashova, K., and Meh, C. (2012). Household borrowing and spending in Canada. *Bank of Canada Review*, (2011):16–29.

Barrett, G., Levell, P., and Milligan, K. (2013). A comparison of micro and macro expenditure measures across countries using differing survey methods. *National Bureau of Economic Research.*

Bertrand, M. and Morse, A. (2012). Trickle-Down Consumption. *NBER Working Papers.*

Bozio, A. and Dea, C. O. (2011). Do the rich really save more? Evidence from lifetime earnings and consumption data in the United Kingdom . *Institute for Fiscal Studies,*, pages 1–49.

Browning, M. and Lusardi, A. (1996). Household Saving : Micro Theories and Micro Facts. *Journal of Economic Literature*, 34(4):1797–1855.

Brzozowski, M. and Crossley, T. F. (2011). Viewpoint: Measuring the well-being of the poor with income or consumption: a Canadian perspective. *Canadian Journal of Economics*, 44(1):88–106.

Burbidge, J. and Davies, J. B. (1994). Household Data on Saving Behavior in Canada. *International Comparisons of Household Saving*, pages 11–56.

Chawla, R. K. (2011). The distribution of mortgage debt in Canada. *Perspectives on Labour and Income*, 3(2):40–47.

Chawla, R. K. and Uppal, S. (2012). Household debt in Canada. *Perspectives on Labour and Income*, 13(3):1–15.

Chawla, R. K. and Wannell, T. (2005). Spenders and Savers. *Statistics Canada*, (75):5–13.

Crawford, A. and Faruqui, U. (2012). What explains trends in household debt in Canada ? *Bank of Canada Review*, pages 3–15.

Cross, P. (2015). A Longer-term Perspective on Canada's Household Debt. *Fraser Research Bulletin*, (May):1–10.

Crossan, D., Feslier, D., and Hurnard, R. (2014). Financial literacy and retirement planning in Canada. *Journal of Pension Economics and Finance*, 10(04):619–635.

Dynan, K. E. (2004). How Prudent are Consumers? *Journal of Political Economy*, 101(6):1104.

Ekici, T. and Dunn, L. (2010). Credit card debt and consumption: evidence from household-level data. *Applied Economics*, 42(4):455–462.

Faruqui, U. (2008). Indebtedness and the Household Financial Health : An Examination of the Canadian Debt Service Ratio Distribution. *Bank of Canada Working Paper*.

Fieldhouse, D., Livshits, I., and Macgee, J. (2012). Income Loss and Bankruptcies over the Business Cycle.

Fortin, M. (2014). Why has the mortgage debt increased by so much in Canada? *Working Paper*.

Friedman, M. (1957). *Introduction to" A Theory of the Consumption Function"*, volume I.

Garrett, T. A. and Wall, H. J. (2010). Personal-Bankruptcy Cycles. *Federal Reserve Bank of St. Louis Working Paper Series*, (010A).

Gokhale, J., Kotlikoff, L. J., Sabelhaus, J., and Bosworth, B. (2009). Understanding the Postwar Decline in U.S. Saving: A Cohort Analysis. *Brookings Papers on Economic Activity.*, 1996(1):315–407.

Hori, M., Iwamoto, K., Niizeki, T., and Suga, F. (2015). Do the Rich Save More in Japan? Evidence Based on Two Micro Datasets for the 2000s Masahiro. *Economic and Social Research Institute*.

Horowitz, J. L. and Lee, S. (2007). Nonparametric Instrumental Variables Estimation of a Quantile Regression Model. *Econometrica*, 75(4):1191–1208.

Horowitz, J. L. and Lee, S. (2009). Testing a Parametric Quantile-Regression Model with an Endogenous Explanatory Variable against a Nonparametric Alternative. *Journal of Econometrics*, 152(2):141–152.

Hurst, M. (2011). Debt and family type in Canada. *Canadian Social Trends*, (91):40–47.

Jakubík, P. (2014). Household balance sheets and economic crisis. *Ekonomická revue - Central European Review of Economic Issues*, 17(2):55–66.

Kiff, J. (2009). Canadian Residential Mortgage Markets: Boring But Effective? *IMF working paper*, (June).

Lee, S. (2004). Endogeneity in Quantile regression Models: A Control Function Approach. *The Institue for Fiscal Studies*.

Livshits, I., MacGee, J., and Tertilt, M. (2015). Accounting for the Rise in Consumer Bankruptcies. *NBER Working Paper Series*, 1(0519324).

Ma, L. and Koenker, R. (2006). Quantile regression methods for recursive structural equation models. *Journal of Econometrics*, 134(2):471–506.

MacGee, J. (2012). The Rise in Consumer Credit and Bankruptcy: Cause for Concern? *Available at SSRN 2046574*, (346).

Meh, C., Terajima, Y., Chen, D. X., and Carter, T. (2009). Household Debt, Assets, and Income in Canada: A Microdata Study. *Bank of Canada Review*.

Mian, A. and Sufi, A. (2009). House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis. *American Economic Review*, 101(5):2132–56.

Mian, A. and Sufi, A. (2011). Household Debt and the Weak U.S. Recovery. *Perspectives on Labour and Income*, 6(12):17–28.

Moav, O. and Neeman, Z. (2012). Saving Rates and Poverty: The Role of Conspicuous Consumption and Human Capital. *Economic Journal*, 122(563):933–956.

Piketty, T. and Saez, E. (2001). Income Inequality in the United States,1913-1998. *NBER Working Paper Series*, 8467:1–121.

Sabelhaus, J., Butless, G., and Sabelhaus, J. (1991). Decline in Saving: Evidence from Household Surveys. *Brookings Papers on Economic Activity*.

Simone, D. (2014). *Household Indebtedness and Socio-Spatial Polarization among Immigrant and Visible Minority Neighbourhoods in Canada's Global Cities*. Masters, University of Toronto.

Walks, A. (2013). Mapping the Urban Debtscape: The Geography of Household Debt in Canadian Cities. *Urban Geography*, 34(2):153–187.

Appendices

Unit Root Tests

log GDP

Table 5: Im-Pesaran-Shin unit-root test for detrended *log* **GDP**

Ho: All panels contain unit roots		Number of panels = 10			
Ha: Some panels are stationary		Number of periods = 28			
AR parameter: Panel-specific		Asymptotics: T,N - Infinity sequentially			
Panel means: Included					
Time trend: Not included					
			Fixed-N exact critical values		
	Statistic	p-value	1%	5%	10%
t-bar	-2.6718		-2.18	-1.99	-1.88
t-tilde-bar	-2.3868				
Z-t-tilde-bar	-3.8402	0.0001			

Table 6: Levin-Lin-Chu unit-root test for detrended *log* **GDP**

Ho: Panels contain unit roots		Number of panels =10	
Ha: Panels are stationary		Number of periods = 28	
AR parameter: Common	Asymptotics: N/T = 0		
Panel means:	Included		
Time trend:	Not included		
ADF regressions:	1 lag		
LR variance:	Bartlett kernel,	9.00 lags	average (chosen by LLC)
	Statistic	p-value	
Unadjusted t	-9.8594		
Adjusted t*	-4.7822	0.0000	

Insolvency Rate

Table 7: Levin-Lin-Chu unit-root test for dtrended **Insolvency Rate**

Ho: Panels contain unit roots	Number of panels = 10	
Ha: Panels are stationary	Number of periods = 28	
AR parameter: Common	Asymptotics: N/T - 0	
Panel means: Included		
Time trend: Not included		
ADF regressions:	1 lag	
LR variance:	Bartlett kernel, 9.00 lags	average (chosen by LLC)
	Statistic	p-value
Unadjusted t	-15.7209	
Adjusted t*	-11.5975	0.0000

Table 8: Im-Pesaran-Shin unit-root test for detrended **Insolvency Rate**

Ho: All panels contain unit roots		Number of panels = 10			
Ha: Some panels are stationary		Number of periods = 28			
AR parameter: Panel-specific		Asymptotics: T,N - Infinity sequentially			
Panel means: Included					
Time trend: Not included					
			Fixed-N exact critical values		
	Statistic	p-value	1%	5%	10%
t-bar	-3.1126		-2.18	-1.99	-1.88
t-tilde-bar	-2.6839				
Z-t-tilde-bar	-5.0361	0.0000			

Unemployment Rate

Table 9: Levin-Lin-Chu unit-root test for dtrended **Unemployment Rate**

Ho: Panels contain unit roots	Number of panels = 10	
Ha: Panels are stationary	Number of periods = 28	
AR parameter: Common	Asymptotics: N/T - 0	
Panel means: Included		
Time trend: Not included		
ADF regressions:	1 lag	
LR variance:	Bartlett kernel, 9.00 lags	average (chosen by LLC)
	Statistic	p-value
Unadjusted t	-11.8527	
Adjusted t*	-7.7927	0.0000

Table 10: Im-Pesaran-Shin unit-root test for detrended **Unemployment Rate**

Ho: All panels contain unit roots		Number of panels = 10			
Ha: Some panels are stationary		Number of periods = 28			
AR parameter: Panel-specific		Asymptotics: T,N - Infinity sequentially			
Panel means: Included					
Time trend: Not included					
			Fixed-N exact critical values		
	Statistic	p-value	1%	5%	10%
t-bar	-2.9324		-2.18	-1.99	-1.88
t-tilde-bar	-2.5574				
Z-t-tilde-bar	-4.5269	0.0000			

Housing Price Index

Table 11: Levin-Lin-Chu unit-root test for dtrended **Housing Price Index**

Ho: Panels contain unit roots	Number of panels = 10	
Ha: Panels are stationary	Number of periods = 28	
AR parameter: Common	Asymptotics: N/T - 0	
Panel means: Included		
Time trend: Not included		
ADF regressions:	1 lag	
LR variance:	Bartlett kernel, 9.00 lags	average (chosen by LLC)
	Statistic	p-value
Unadjusted t	-11.1771	
Adjusted t*	-6.1495	0.0000

Table 12: Im-Pesaran-Shin unit-root test for detrended **Housing Price Index**

Ho: All panels contain unit roots		Number of panels = 10			
Ha: Some panels are stationary		Number of periods = 28			
AR parameter: Panel-specific		Asymptotics: T,N - Infinity sequentially			
Panel means: Included					
Time trend: Not included					
			Fixed-N exact critical values		
	Statistic	p-value	1%	5%	10%
t-bar	-2.7875		-2.18	-1.99	-1.88
t-tilde-bar	-2.4685				
Z-t-tilde-bar	-4.1690	0.0000			

Debt to Disposable Income Ratio

Table 13: Levin-Lin-Chu unit-root test for dtrended **Debt to Disposable Income Ratio**

Ho: Panels contain unit roots	Number of panels = 10	
Ha: Panels are stationary	Number of periods = 28	
AR parameter: Common	Asymptotics: N/T - 0	
Panel means: Included		
Time trend: Not included		
ADF regressions:	1 lag	
LR variance:	Bartlett kernel, 9.00 lags	average (chosen by LLC)
	Statistic	p-value
Unadjusted t	-10.9708	
Adjusted t*	-5.3215	0.0000

Table 14: Im-Pesaran-Shin unit-root test for detrended **Debt to Disposable Income Ratio**

Ho: All panels contain unit roots		Number of panels = 10			
Ha: Some panels are stationary		Number of periods = 28			
AR parameter: Panel-specific		Asymptotics: T,N - Infinity sequentially			
Panel means: Included					
Time trend: Not included					
			Fixed-N exact critical values		
	Statistic	p-value	1%	5%	10%
t-bar	-2.7875		-2.18	-1.99	-1.88
t-tilde-bar	-2.4685				
Z-t-tilde-bar	-6.0135	0.0000			