Heterogeneity of U.S. Household Labor Income:

Evidence from Stock Market

Zhixue Su

May 18, 2014

Abstract

The thesis is aimed to have a deeper understanding of heterogeneity of household labor income process across different educational levels and industrial categories. I utilize a fixed-effect model and risk decomposition methodology to analyze the differences of risk profile of labor income innovations and the correlation between labor income process and stock portfolio returns. The results from U.S. empirical data have shown quite different risk profiles for households within different educational and industrial cell. The study reinforces present research results and serves as a good explanation for the failure of mutual fund separate theorem.

Keywords: Labor Income Process, Heterogeneity, Optimal Portfolio Choice, Beta Decile Index Portfolio, Sector Index Portfolio

Acknowledgement: I would like to specially thank my supervisor, Professor Paolo Sodini, for providing me so many insights and advices, and for being patient and supportive in the whole period. I would also warmly thank my family and friends for their love and support throughout my academic life.

Contents

List of Tables

List of Graphs

1 Introduction

The privatization of the social security system has become a heating topic during the past few decades. There are two main reasons which have driven the debate. First of all, American households are becoming more sophisticated towards financial instruments. People get to know that there are a large group of alternative investment vehicles offering various combinations between risk and rate of returns. Secondly, the current social security system in US is facing unsustainable problems. The net present value of unfunded liabilities in the American social security system are around $9-10 trillion (in 1997 dollars) if the current one remains unchanged according to estimation in Geanakoplos, Mitchell and Zeldes (2000). They have pointed out that it is necessary either to cut down the unfunded liabilities (a burden on today's retirees) or to raise tax contributions (a burden on today's workers) to maintain the sustainability of the social security system.

In their thesis, they have pointed out that the implicit rates of return on tax contributions to the current system have gradually declined. What's more, the adverse trend would not stop as a result of unfavorable demographic change[1]. The authors have proven that the net present value of transfers to different generations has to sum to zero in an ongoing system. The current social security system has included the aged who were already eligible or close to retirement to receive benefits from the system. Soon after the program was launched, a portion of social security taxes went directly to the aged to pay the due liabilities instead of being invested, which led to a lower rate of return for the younger generations as the older generations have got high returns. This has been a generally used point by the supporters of the social security reform as they argue that privatization will allow households to chase higher rates of returns from investment, such as stocks.

As the social security privatization is such a complicating problem which has many factors needed to be considered, there are many methods to approach the social security privatization problem. In my thesis, I am going to focus on the relationship between household income process and riskiness of different types of stock portfolios. The rationale behind that is based on one of the findings in Campbell,

[1] The proportion of the aged in the total population is going up because of improving medical care and declining birth rate, which will create extra burden to today's workers.

Cocco, Gomes and Maenhout (2001). In their study, they have found out that for the self-employed college graduates, their labor income volatility is greater and it is also more correlated with returns of stocks and long-term bonds. Their simulation results have shown that it will be optimal for people with these kinds of risk to hold less risky portfolios from utility perspective.

In my study, I will utilize a similar approach and try to expand the range of stock portfolios available and incorporate the household industry into my analysis as well. Regardless of the fact that how the social security privatization reform should be conducted remains unresolved, my results can at least act as a general guidance for households to choose different investment vehicles for their personal savings. In addition, several studies have agreed on that the social security privatization will be detrimental to households that lack the essential financial knowledge to decide their own portfolio choices. Hopefully, the result can bring some insights into how the administrator should set up a portfolio recommendation list for those households with less financial sophistication assuming the privatized pension system has been in place. Besides that, the problem of unsustainable social security system does not only appear in US due to growing aging population and post-war generations. The study might be helpful for researchers in other countries as well.

My study shows that: 1) households with higher educational level are less vulnerable to transitory shock to labor income process; 2) the total variance of labor income process and the correlation between household labor income process and beta decile index returns varies greatly across different educational and industrial cells. These results, together with simulation results from previous studies by Cocco, Gomes and Maenhout (2005), provide very useful implications for optimal portfolio choice problem. A similar analysis on sector index returns does not show any strong relationship between labor income process and industrial sector stock returns.

The rest of the thesis is organized as follows. Section 2 provides some backgrounds in regard to the discussion on social security privatization and the development of study about labor income process. A literature review is in Section 3. Section 4 introduces both the income data and stock data I have used in the study. The methodology I utilized is described in Section 5. Section 6 presents the regression results, explores the heterogeneity of the labor income across different

educational levels and industrial categories and summarizes implications of my study. Limitations are discussed in Section 7. Finally, Section 8 concludes with unresolved problems and potential research topics.

2 Background

In this section, I mainly present two main themes related with my study, that is, the social security privatization together with surrounding controversies and the development in studies of optimal portfolio choice.

2.1 Social Security Privatization

As I mentioned before, social security privatization in US is a complicating problem with a lot of factors needed to be considered. It is Mitchell and Zeldes (1996) who have provided a good structure with a wide range of interesting factors to analyze their proposal of a two-pillar plan. In their study, they have brought up five key factors which should be taken into consideration regarding this topic. The first two of them have taken a household perspective in analyzing the problems while the rest three will catch more attention from policymakers. More details on these factors and related researches are provided below.

In the following two paragraphs, I will talk about the factors related with households' interests and behaviors. First of all, we want to know how the risks households face will change given that any reform scheme of the social security system is implemented. The topics include income uncertainty, inflationary shock, longevity risk, disability insurance, political risk and intergenerational risk-sharing. Burtless (2000) have evaluated performance of a privatized social security system between 1911 and 1999 by using the historical financial returns in US. It has demonstrated that the good returns of a private plan come at a price of weakening risk sharing across geography and generation.

Secondly, we need to analyze the potential household behavior changes as a result of social security privatization. For example, household will get more freedom in choosing portfolio but whether it is beneficiary or detrimental remains unclear. Besides, household saving habits and work incentives will also change due to the social security privatization. Nishiyama and Smetters (2007) have calculated efficiency gains from a stylized 50-percent privatization model under various assumptions of the risk. They have shown that there are efficiency gains from privatization due to incentive to increase labor supply. However, the weakening risk

sharing function of social security system will leave the households vulnerable when they face uninsurable risks.

Besides the influences on single household or a group of them with similar demographic profile, we can never emphasize too much on the potential macro effects from social security privatization, which will be of great interest to the policymakers. Firstly, it is important to investigate how the redistribution to the low earners will change since the government wants to balance between fairness across different income levels and minimum welfare received by the low earners. Secondly, there are several macroeconomic implications connected with social security privatization, such as transition, effects on national savings and rates of return comparison. These questions are very tricky and sometimes nearly impossible to be evaluated, nevertheless the success of social security privatization relies largely on how well they can be solved. Last but not the least, assuming the social security privatization will increase the actual implicit rates of returns, the government still needs to measure if it is still favorable after deducting potential increasing administrative cost.

It is no wonder that these questions have drawn as much attention as the first two factors. Conesa and Kruegar (1999) have asserted that it is difficult for privatization plan to get enough political support. Agents with idiosyncratic income risk would prefer to stay in the current system which has the feature of partial insurance and redistribution. Geanakoplos, Mitchell and Zeldes (2000) have tried to answer the question whether there will be increase in the rates of return from privatization. They have asserted the benefits will be less than the normal expectations from reformers if we take the riskiness of stocks and existing unfunded liabilities in the current social security system into consideration.

Besides all these, there are also a group of studies focused on the framework and design of the social security privatization. Gramlich (1996) have discussed on three different approaches and used money's worth ratio (the ratio of the present value of expected benefits to the present value of taxes) to measure the performance. He has come with a conclusion that a plan both lowers the existing accrued benefits and introduces partial privatization is needed to solve the unsustainable problem of the current system. Kotlikoff, Smetters and Walliser (1999) have used a dynamic

simulation model to investigate potential reform choices and recommended a progressive plan with low transition cost and favorable macroeconomic outcomes. They have found out that the long run gains from privatization can be substantial but costs welfare loss for the transition generations.

2.2 Studies of Optimal Portfolio Choice

As I mentioned before, I focus on the relationship between household labor income process and the riskiness of different stock portfolios in the thesis. The aim is: 1) to evaluate risk characteristics of alternative stock portfolios by considering potential income risk faced by ordinary households; 2) to discuss the related implications on portfolio choice problems for different households. An overview of the academic achievements in this field is presented below.

The pioneers in the theoretical study on multi-period optimal portfolio choice were Samuelson (1969) and Merton (1969). Samuelson (1969) has proved that so-called "businessman's risk"[2] will be invalid assuming frictionless market and independently and identically distributed (i.i.d.) rates of returns on both safe and risky assets. Under the same assumptions, Merton (1969) has concluded that long-term investors will make the same decision as single-period investors. Even though these studies have provided a lot of interesting implications about optimal portfolio choice, there are two main drawbacks in them. Firstly, as we all know, the assumption about frictionless market and i.i.d. rates of return does not always hold. Secondly, these discussions have totally ignored a very important factor, that is, labor income for household to make optimal portfolio choice.

In Merton (1971), he has tried to make up for that by introducing labor income into the model. He has made a thorough analysis under the assumption that the agents live in a complete market where man can capitalize his/her labor income and buy insurance against labor income uncertainty. However, the market is incomplete in reality, which means people have to face borrowing constraints so they cannot capitalize their labor income. What's more, it is impossible to perfectly hedge against the idiosyncratic labor income risk. In other words, a model with loosened conditions that can resemble the real market is needed to conduct empirical studies. After that,

[2] It is only advisable for young businessman instead of widows to hold risky investments.

the study on optimal portfolio choice problem over the life cycle has been stagnant for a long time.

Thanks to a raising concern about how an ordinary household makes a living and how different characteristics might influence individual income, saving, asset accumulation and consumption, a national survey called Panel Study of Income Dynamics (PSID) has been conducted annually since 1968 in US. It has gradually become a comprehensive dataset for studies related with household finance. Besides the availability of empirical data, another factor which has confined the study on optimal portfolio choice over the life cycle is that the complexity of the models will increase significantly even if only one market incompleteness has been introduced. Most times the complexity of the model makes it impossible to find analytical solutions. However, a lot of economists have come back to the topic of optimal portfolio choice over life cycle because of development of numerical solution and simulation techniques and massive use of computers.

Some of the interesting studies and their results are listed as follows.

Brennan and Xia (1998) have considered about time varying interest rate and have found out there is a negative correlation between bonds and expected future interest rates. They have asserted that this important property makes multi-period portfolio choice different from Tobin separation theorem, which only considers a one-period optimal question. Wachter (2001) have reviewed the riskiness of a short-term bond, whose return is usually regarded as riskless. The study has come with the conclusion that a multi-period investor with high risk aversion will prefer long-term bond since the rates of return on short-term bond follow a stochastic process over lifetime.

Besides the riskless rate, economists are also very interested in time varying equity premiums. Campbell and Viceira (1999) have valued a model with constant riskless rate and a time varying equity premiums and concluded that an optimal amount of stock investment will greatly increase and timing the stock market will be important as well under this assumption. Brandt (1999) has done analysis about optimal portfolio choice assuming time varying equity returns. The results have shown that the time varying premiums, together with investor’s horizon and rebalancing frequency play an important role in deciding optimal portfolio. Brennan,

Schwartz and Lagnado (1997) have done a more complex model in which they allow for both time varying riskless rate and risky investments. The numerical solution shows there is a big difference in optimal portfolio choice due to investor's horizon. The study also shows it is with sufficient reliability to adjust investment strategy according to predication of asset returns.

As I mentioned before, this study is aimed at estimating the relationship between household labor income process and various stock investments and evaluating its implications on optimal portfolio choice. A literature review about how the labor income process will influence optimal portfolio choice will be presented in next section.

3 Literature Review

The introduction of labor income into optimal portfolio choice is very intuitive as most households rely greatly on their labor income. A concise summary of the whole idea is that we can regard the future labor incomes for a household as a nontradable asset with multi-period cash inflows. In other words, having a labor income means the household implicitly holds a bond-like asset in his portfolio. However, the risk profile of this asset can vary a lot depending on household's demographic characteristics, such as, educational level, industry of current job, and etc. Economic implications on related topics are listed below.

Firstly, the correlation between labor income and risky assets plays an important role in deciding optimal portfolio choice. Assuming that the labor income process is uncorrelated with other risky investments, then it will be optimal for the household to hold more risky assets in their explicit portfolio. But if there is a positive correlation between labor income and risky investments, the household will choose to hold less or even none risky assets.

Secondly, another possible situation is that the household cannot hold the optimal amount of risky assets because they are facing borrowing constraints. In other words, it is impossible to capitalize the future labor income. This will result in a corner solution where the household invest all their money into risky investments.

Sometimes, the fixed participating cost will deter especially young people or poor people from holding any risky investments. This situation may improve later in their life when the profits from risky investments can justify paying the cost.

Besides correlation, borrowing constraints and fixed participating cost, other interesting topics include tax distortion on optimal portfolio choice, flexible labor supply and its effect on optimal portfolio choice and so on.

Empirical studies in these areas have generated many interesting results. Bodie, Merton and Samuelson (1992) have shown that compared to the old, the young are willing to take more investment risk under the assumption that it is easier for the young to vary their labor supply to compensate for potential loss. Campbell, Cocco, Gomes and Maenhout (2001) have utilized a life-cycle model in their study and tested how the optimal portfolio choice will vary due to changes in the risk aversion,

correlation between labor income risk and other investments, and other market incompleteness. Gomes and Michaelides (2004) have justified the substitution effect between labor income and bonds/stocks depending on the correlation between labor income and various investment vehicles. Their results have served as a good answer to the asset allocation puzzle[3] discussed in Canner, Mankiw and Weil (1997). Davis and Willen (2012) have developed and applied a simple graphical approach to analyze occupation-level income risk with different stock portfolios. They have assumed partially insurable income process and provided evidence for significant deviation from two-fund separation principle as well.

My study will be concentrated on the uninsurable income risk and its correlation with various stock portfolios. It estimates labor income process for households with different backgrounds and then calculates the correlation between uninsurable income risk and various stock investments. It will be a good complement to previous studies by using up-to-date PSID empirical data. The expansion of alternative stock investments provides useful insights in regard to optimal portfolio choice problem. The comparison with current studies will also bring in new findings or strengthen the reliabilities of previous studies in regard of heterogeneous labor income process and its impact on optimal portfolio choice.

[3] The discrepancy between asset allocation advice and mutual fund separation theorem.

4 Data

The data used in my study was comprised of several different sources, including panel data of household labor income from Panel Study of Income Dynamics – PSID, sector indices data from Datastream and beta decile indices data from Center for Research in Security Prices – CRSP. The detailed descriptions are as follows.

4.1 Household Labor Income Process from PSID

PSID is the longest running longitudinal household survey in the world. It is a survey of over 18,000 individuals and their descendants living in 5,000 families in the United Stated, which firstly started since 1968. The survey covers topics such as employment, income, demographic characteristics including sex, education, age, and marital status and so on. Between 1968 and 1997, the survey was conducted on an annual basis. Since 1999, the frequency of survey has been changed to be biennial.

To make a full use of the survey, I utilized a sample between 1970 and 2011 in my study. I only included male heads in my analyses due to the limited sample size of the female head, plus the fact that household labor income process has different age profiles and income levels across head sex. In the appendix, I listed the tables and graphs describing the differences between male and female head household. I excluded the Survey of Economic Opportunity sub-sample because it focused on low-income families in US. The immigrant sub-sample which was gradually added in 1990s was also excluded to obtain a random sample. Besides all above, I limited the sample by only including head's age between 20 and 65[4] and currently working. The sample was unbalanced as family unit with shorter time span was remained.

In the following, I will describe how I calculated the household labor income and all the demographic characteristics used in the study.

Household Labor Income – The household labor income is defined to be in line with Cocco, Gomes and Maenhout (2005), which implicitly allows for self-insurance against pure labor income risk. Under the broad definition, the household labor income contains pure labor income, social security income, AFDC/TANF[5] income,

[4] For head with education level college or above, the lowest age is 22 instead.
[5] Aid to family with dependent children or temporary assistance for needy families.

SSI[6] income, unemployment compensation, workers compensation, child support, other welfare, help from relatives and help from others for both head and wife (if present). All the households still with zero annual income were excluded and the household labor income was deflated by using the Consumer Price Index[7], with 1982 as the base year. Then I took common logarithm of the deflated income before it can be used in the regression.

Demographic Characteristics

Educational Level – The whole sample was divided into three groups according to the educational level of the head: the first group with unfished high school level, the second with high school but unfinished college level, and the last with college and above. The sample split is based on findings of Hubbard, Skinner and Zeldes (1994) and Attanasio (1995). They have shown that the age profiles are different in shape across educational levels. For head whose education jumped from lower level to higher one, the original family unit was discontinued and the following observations were regarded as a new entity to avoid problem in the fixed-effect regression.

Industrial Category – In 1970 and 2000, there were two separate Census of Population conducted in US. Before 2001, PSID reported head industrial category by following the rule of 1970 census, which contained 12 different categories of industry. However, PSID reported head industrial category in a more detailed 19 groups starting from 2003 wave. I took the original 12 categories as a basis and there were 9 categories I could match before and after the change, which gave me a full-length sample with 9 categories and a shorter sample with the rest 3 categories. The additional categories which were only present after 2003 were dropped as they have 5 observations for each family unit at maximum. The detail about industrial category can be found in the appendix.

Others – I ran the regression for each combination of educational level and industrial category separately with a fixed-effect regression. Other explanatory variables include Head Age, Head Marital Status and Family Unit #, which equals the additional number of persons in the same family unit other than head and his spouse (if present). The model controls for family specific variations as well.

[6] Supplemental security income.
[7] The Consumer Price Index statistics are published monthly by the Bureau of Labor Statistics.

4.2 Sector Indices from Datastream

To investigate the relationship between stock market and personal income risk profile, I firstly used empirical data of sector indices from Datastream. The sector is defined according to Global Industry Classification Standard (GICS) and the data contains 10 different sectors of stocks in US market. The Datastream sector indices have been updated on a daily basis since 1972. The index price level at the end of each year was used to calculate annual log returns.

4.3 Beta Decile Indices from CRSP

Being different from the sector indices data, the beta decile indices provide a way to analyze how well a family unit with a specific combination of educational level and industrial category can hedge the income shock against stock portfolios with different risk levels. The CRSP beta decile indices are based on NYSE/AMEX stocks and have a full span between 1970 and 2011. Same methodology was implemented to get the annual log returns.

5 Methodology

To study the heterogeneity of household income across educational level and industrial category, the first step is to estimate the household income process. In this study, I employed a similar approach as Cocco, Gomes and Maenhout (2005) to estimate the household labor income before retirement. Investor *i*'s labor income at age *t*, Y_{it} is exogenously given by

(1) $$log(Y_{it}) = f(t, Z_{it}) + v_{it} + \varepsilon_{it} \quad for\ t \leq K,$$

In the formula above, $f(t, Z_{it})$ is a deterministic function of investor's age and other demographic characteristics, including Head Marital Status, Family Unit #, Family Unit Identifier. The residual can be regarded as labor income risk and it is composed of two separate risk sources. ϵ_{it} is an idiosyncratic transitory shock with a distribution of $N(0, \sigma_{\varepsilon}^2)$. And v_{it} is defined by

(2) $$v_{it} = v_{i,t-1} + u_{it},$$

u_{it} has a distribution of $N(0, \sigma_u^2)$ and it is uncorrelated with the transitory shock ϵ_{it}. In short, log labor income is the sum of a deterministic function and two random risk sources, one permanent and one transitory. The deterministic function can capture the hump-shaped labor income over investor's life.

After estimating the labor income, I create age profiles of the labor income process over life for different educational level. This step is to mimic Campbell, Cocco, Gomes and Maenhout (2001) and Cocco, Gomes and Maenhout (2005) and will act as a verification of PSID dataset as well.

To decompose the permanent and transitory shocks to labor income, I followed Campbell, Cocco, Gomes and Maenhout (2001) and Carrol and Samwick (1997). Firstly, I defined Y_{it}^* as follow

(3) $$log(Y_{it}^*) \equiv log(Y_{it}) - \hat{f}(t, Z_{it}),$$

which means $log(Y_{it}^*)$ is the residual from estimating formula (1). Then a d-year difference between the residuals will be

(4) $$log\left(Y_{i,t+d}^*\right) - log(Y_{it}^*) = v_{i,t+d} - v_{it} + \varepsilon_{i,t+d} - \varepsilon_{it},$$

Then by substituting (2) into (4) recursively, I could get

$$(5) \qquad \begin{aligned} log\left(Y_{i,t+d}^{*}\right) - log(Y_{it}^{*}) &= v_{i,t+d} - v_{it} + \varepsilon_{i,t+d} - \varepsilon_{it} \\ &= \left(u_{i,t+d} + u_{i,t+d-1} \cdots + u_{i,t+1}\right) + \varepsilon_{i,t+d} - \varepsilon_{it} \end{aligned}$$

Because of the assumption that u_{it} and ϵ_{it} are uncorrelated with each other at all leads and lags, (5) can yield a generalized formula for d-year variance

$$(6) \qquad Var\left[log\left(Y_{i,t+d}^{*}\right) - log(Y_{it}^{*})\right] = d\sigma_u^2 + 2\sigma_\varepsilon^2,$$

Then I ran an OLS regression of $Var\left[log\left(Y_{i,t+d}^{*}\right) - log(Y_{it}^{*})\right]$ on $[\boldsymbol{d} \quad \mathbf{2}]$, which stands for two vectors: $\{1 \quad \cdots \quad d\}'$ and $\{2 \quad \cdots \quad 2\}'$ respectively. The procedure is repeated household by household within each given group (educational level / industrial category / etc.). By doing this, I could get the estimations of permanent and transitory income shock for each specific group.

To estimate the potential gains from customized investment based on investor's demographic characteristics, I need to estimate the correlation coefficient between labor income shocks and stock returns. Similar to Campbell, Cocco, Gomes and Maenhout (2001), the permanent labor income shock u_{it} is assumed to have an aggregate component and idiosyncratic component.

$$(7) \qquad u_{it} = \theta_t + \omega_{it},$$

From above, we can see that the aggregate component θ_t is the same for all the households in the same group while the idiosyncratic component ω_{it} is uncorrelated across households. This definition indicates that the random component of labor income on aggregate level will follow a random walk, which is in alignment with assumption made by Fama and Schwert (1977)[8].

Since the change in $log(Y_{it}^{*})$ can be written as

$$(8) \qquad \Delta log(Y_{it}^{*}) = log(Y_{it}^{*}) - log\left(Y_{i,t-1}^{*}\right) = u_{it} + \varepsilon_{it} - \varepsilon_{it-1} = \theta_t + \omega_{it} + \varepsilon_{it} - \varepsilon_{it-1},$$

The average across household will yield

$$(9) \qquad \overline{\Delta log(Y_{it}^{*})} = E(\theta_t + \omega_{it} + \varepsilon_{it} - \varepsilon_{it-1}) = \theta_t,$$

[8] Cited from pp.6 Cocco, Gomes and Maenhout (2001).

Then an OLS regression of $\overline{\Delta log(Y_{it}^{*})}$ on the demeaned excess returns of stock portfolio will give the correlation coefficients desired.

(10) $$\overline{\Delta log(Y_{it}^{*})} = \beta\left(R_{t+1} - \overline{R_f} - \mu\right) + \varphi,$$

To analyze the potential difference in the correlation coefficients between labor income process and stock returns across educational level and industrial category for the sample, I ran the same regression on two different empirical datasets. One is the sector indices from Datastream and the other one is the beta decile indices from CRSP.

6 Empirical Results

6.1 Fixed-effect regression on different educational level

In this part, I will present the results from fixed-effect regression, which is a necessary step before risk decomposition of labor income. As I mentioned before, the common logarithm of household labor income is used as the explained variable. Then the households are grouped according to their educational level and are regressed by using Equation (1), which can assure us to capture the different age profiles of labor income for different education groups. Table 1 tabulates the estimated coefficients for 3 different educational levels and corresponding model fitting information.

Table 1: Labor Income Process: Fixed-effect Regression

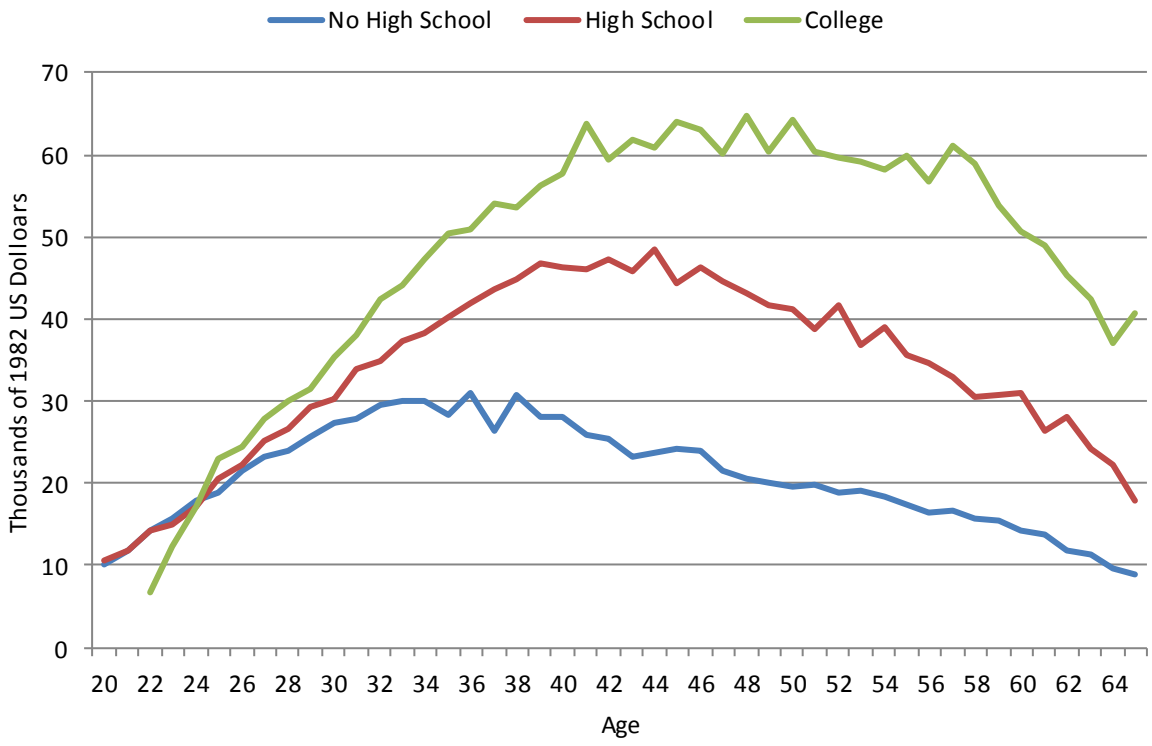
	No High School		High School		College	
	coefficient	t-statistic	coefficient	t-statistic	coefficient	t-statistic
Constant	3,9531	91,72	4,2531	119,38	4,6085	99,14
Family Size	-0,0630	-17,29	-0,0412	-19,48	-0,0488	-14,88
Marital Status	-0,0154	-1,15	0,0711	10,87	0,1770	17,89
N	12285		41205		20567	
Mean Square Error (MSE)	0,2868		0,2363		0,2378	
R2	0,0567		0,0941		0,1090	
F-Statistic	15,65		90,92		55,77	

Most of the age dummies across different educational level are statistically significant in the regression. However, the estimates are hidden for space. The only estimated coefficient which is not statistically significant is marital status for no high school education group. The family size has a negative correlation with log labor income, which means family unit of smaller size tends to have a higher labor income. At the same time, marital status has a positive influence on the household labor income and the effect is even stronger in high education group. In general, the regression results are in alignment with Cocco, Gomes and Maenhout (2005).

In graph 1, I plot the fitted labor income processes for agents from different education level. The plot starts at age 20 and ends at age 65 except for college

group, which starts at age 22 and ends at the same age. The graph clearly shows the hump shape labor income across ages for all three groups. There are also some interesting trends based on the results. It is obvious that the higher education received, the more labor income will be received by the household. Compared to other two groups, households from college group have a steeper growth curve and reach labor income peak in the later stage of their lives. What's more, they can enjoy the stable high labor income for a much longer period than the other two groups. On the contrary, the households from no high school group reach their income peak at mid-30s and their labor income starts to decrease as they are approaching age 40.

Graph 1: Fitted Labor Income Process across Ages



6.1.1 Risk decomposition across different educational level

Using the residuals from the regression above, I can estimate the permanent and transitory labor income risk for households with different educational backgrounds by using Equation (6). It follows a similar procedure as Carroll and Samwick (1997) and Cocco, Gomes and Maenhout (2005). The results are presented in table 2. The estimated coefficients are all statistically significant only with one exception, that is,

the permanent shock in the labor income process of high school group. In Cocco, Gomes and Maenhout (2005), they have investigated the influence of labor income risk on the optimal portfolio choice. One of their findings shows that lower total variance of shock to labor income[9] will result in more allocation to risky asset. However, only a large difference in total variance of shock to labor income can cause a significant deviation of optimal portfolio choice from the base case. From this table, we can see that the college group has a lower transitory shock to labor income and a lower total variance of shock to labor income than the rest two groups do in spite of the higher permanent shock.

Table 2: Labor Income Risk Decomposition on Educational Level

	No High School	t-statistic	High School	t-statistic	College	t-statistic	Overall	t-statistic
Permanent Shock	0,0043	2,27	0,0017	1,01	0,0075	4,10	0,0039	3,60
Transitory Shock	0,0172	4,51	0,0221	6,42	0,0109	2,67	0,0181	7,95

6.1.2 Correlation between labor income and beta decile indices

As we mentioned before, another important factor in optimal portfolio choice is the substitution effect between labor income and risky investments, which means that household will invest less in stock market if the shock to their labor income is highly correlated with stock returns. Table 3 shows both the correlation and beta between labor income and beta decile indices on different educational level. From this table, we can see that the labor income is showing strong positive correlation with most beta decile index of NYSE/AMEX stocks across all three educational levels. Besides, the correlation is more prominent in mid-beta decile index, such as beta decile group 6. Another interesting finding is that the correlation becomes even stronger for higher education group, which has also been mentioned in Davis and Willen (2012). Betas between labor income and the beta decile index returns have shown similar patterns as correlation while it is more prominent in high-beta decile group. The positive correlation will lower the benefits of investing in the stock market and thus result in less wealth accumulation in the whole life. This positive correlation between labor

[9] The total variance of shock to labor income is simply the sum of permanent and transitory shock for given group of households.

income innovations and stock returns has also been found in early studies, including Davis and Willen (2012) and Heaton and Lucas (1999).

Table 3: Correlation and Beta between Labor Income and Beta Decile Index Returns

	No High School		High School		College	
	Correlation	Beta	Correlation	Beta	Correlation	Beta
Beta Decile Group 1	0,1508	0,0344	0,1962	0,0324	0,2816	0,0403
Beta Decile Group 2	0,1825	0,0403	0,2342	0,0388	0,3325	0,0495
Beta Decile Group 3	0,1955	0,0476	0,2540	0,0461	0,3356	0,0564
Beta Decile Group 4	0,2246	0,0542	0,2777	0,0521	0,3761	0,0654
Beta Decile Group 5	0,1730	0,0460	0,2937	0,0514	0,4099	0,0668
Beta Decile Group 6	0,2359	0,0550	0,3155	0,0561	0,4239	0,0714
Beta Decile Group 7	0,1833	0,0435	0,3106	0,0511	0,4424	0,0691
Beta Decile Group 8	0,1733	0,0549	0,2662	0,0574	0,3842	0,0747
Beta Decile Group 9	0,1406	0,0767	0,3182	0,0861	0,4229	0,1047
Beta Decile Group 10	0,0756	0,0718	0,1942	0,0723	0,3136	0,0900

The beta decile group ranges between 1 and 10. Group 1 contains stocks that have beta less than the 10th percentile of betas of stocks in the whole market while group 10 contains stocks that have beta larger than the 90th percentile of betas.

6.2 Fixed-effect regression on both education and industry

In the second part of my study, I investigated more about the heterogeneity of household labor income based on not only the educational level but also industrial category. The industrial dummies are added into the fixed-effect regression as a preparation for decomposition of labor income risk and estimation of the correlation between labor income innovation and returns of stock portfolio. In other words, the methodology implies that we will allow the industrial category to influence the level instead of the shape of the curve for household in each education and industry cell. Table 4 summarizes the results for coefficient estimates and the model fitting information.

As I mentioned in the data description sector, PSID dataset reports head industrial category by following Census of Population in US. However, there are changes in the definition for different industrial categories and new jobs/industries are included as a result of social change. I then limit the range of my study by only including the 9 industrial categories which have a full-length sample. The included industry names can be found in table 4. As I am going to calculate the correlation between labor

income innovations and industrial index returns from Datastream, which has not started until 1972, I implement a further cutoff on the dataset to only include observations after 1972.

Table 4: Labor Income Process: Fixed-effect Regression with Industrial Dummies

	No High School		High School		College	
	coefficient	t-statistic	coefficient	t-statistic	coefficient	t-statistic
Constant	4,0915	75,59	4,2531	119,38	4,6085	99,14
Family Size	-0,0494	-12,47	-0,0412	-19,48	-0,0488	-14,88
Marital Status	-0,0069	-0,49	0,0711	10,87	0,1770	17,89
Agriculture, Forestry, and Fisheries	-0,2921	-9,51	-0,2380	-17,18	-0,3202	-14,86
Mining	-0,0063	-0,13	0,0718	3,17	-0,0191	-0,45
Construction	0,0900	3,22	0,0262	2,61	-0,0312	-1,71
Manufacturing	0,0308	1,15	-0,0152	-1,70	-0,0586	-4,64
Wholesale and Retail Trade	0,0687	2,40	-0,0065	-0,69	-0,1028	-7,40
Finance, Insurance, and Real Estate	0,0862	1,97	-0,0109	-0,74	-0,0076	-0,50
Professional and Related Services	0,0037	0,10	-0,1045	-8,55	-0,1532	-12,85
Entertainment and Recreation Services	0,0163	0,24	0,0257	1,16	0,0017	0,06
Public Administration	-	-	-	-	-	-
N	9493		33323		17893	
Mean Square Error (MSE)	0,2470		0,2140		0,2174	
R2	0,0872		0,1151		0,1308	
F-Statistic	16,40		78,67		50,64	

The estimates for constants, family size and marital status remain stable compared to the fixed-effect regression without industry dummies. The estimation for Public Administration is missing since we only need 8 dummy variables for 9 different industrial categories. The significance of the estimated industrial dummies is mixed. However, we can see some interesting patterns in the labor income across educational level and industrial category. Firstly, household with head in agriculture always has a much lower labor income level than other industries. Compared with family unit from public administration, which is generally in the high labor income group, household in agriculture has roughly 40% to 50% less labor income if everything else is held equal. Secondly, household with head in finance or public administration tends to have relatively high labor income than other groups for household with head education level high school or above. Thirdly, the ideal industry varies depending on the household head educational level. For example, construction and trade industries are more desirable for households with no high school education than jobs in public administration sector.

6.2.1 Risk decomposition across both education and industry

Before using the residuals from the fixed-effect regression to investigate heterogeneity in labor income risk, I need to tabulate the distribution of the whole sample in different education and industry cells. As we can see from table 5, there is a huge variation across different industries. The number of household-year observations in some cells is very limited, such as mining and entertainment and recreation services. To avoid that our estimation for some cells will be determined by only a few households, I exclude cells where any PSID wave data has fewer than 15 observations from my following study. However, these cells are included when I report the results for a given educational level across different industries or for a given industrial category across different educational levels. The results for mining and entertainment industry will be completely omitted in the following section.

Table 5: Cell Sizes for Different Education and Industry Combination

Industry	No High School	High School	College	Total
Agriculture, Forestry, and Fisheries	882*	1632	605*	3119
Mining	155*	472*	130*	757*
Construction	2041	5318	961*	8320
Manufacturing	3736	10628	3920	18284
Wholesale and Retail Trade	1605	7415	2446	11466
Finance, Insurance, and Real Estate	197*	1386	1729	3312
Professional and Related Services	433*	2426	5705	8564
Entertainment and Recreation Services	66*	494*	299*	859*
Public Administration	378*	3552	2098	6028
Total	9493	33323	17893	60709

* Educational and industrial cell that has fewer than 15 observations in any PSID wave of survey.

Risk decomposition only across industrial category

Firstly, I conduct a risk decomposition of labor income process only across different industries. This can help us to understand how labor income risk varies for household with head working in different industries. The result can be found in table 6 and estimate for the variance with a negative sign is set to zero. As we can see from the results, people working in the agricultural industry have a much higher permanent labor income shock compared with households in other industries. In regard to transitory labor income shock, people in agricultural industry and wholesale/retail

trade industry suffer from huge labor income uncertainty. Households in manufacturing/public administration have the lowest total labor income variance and the rest industrial categories have a similar mid-level shock to labor income.

Table 6: Labor Income Risk Decomposition on Industrial Category

Industry	Permanent Shock	t-statistic	Transitory Shock	t-statistic
Agriculture, Forestry, and Fisheries	0,0340	2,67	0,0472	1,78
Construction	0,0024	1,90	0,0192	7,74
Manufacturing	0,0073	3,37	0,0044	1,00
Wholesale and Retail Trade	0,0000		0,0407	3,55
Finance, Insurance, and Real Estate	0,0052	2,22	0,0154	3,12
Professional and Related Services	0,0046	2,78	0,0144	3,95
Public Administration	0,0001	0,03	0,0155	2,72

Risk decomposition across both educational level and industrial category

Next, I use the same methodology but with smaller cells based both on educational level and industrial category for the household to estimate shocks to labor income process. We can see the results from table 7. All the results for cells that don't have enough observations for each wave of PSID survey are omitted from the table. The results are all available for households from high school educational level as the sample size is big enough. The table shows that the total variances of labor income process for households in manufacturing and public administration are consistently low across educational level. However, for households in wholesale/retail trade industry, the transitory shock to their labor income is mainly concentrated in mid-education group.

Table 7: Labor Income Risk Decomposition across Education and Industry

	No High School		High School		College	
Industry	Perm	Tran	Perm	Tran	Perm	Tran
Agriculture, Forestry, and Fisheries	-	-	0,0085	0,0676	-	-
Construction	0,0025	0,0155	0,0024	0,0219	-	-
Manufacturing	0,0081	0,0047	0,0074	0,0045	0,0042	0,0056
Wholesale and Retail Trade	0,0003	0,0140	0,0000	0,0553	0,0060	0,0093
Finance, Insurance, and Real Estate	-	-	0,0055	0,0103	0,0047	0,0211
Professional and Related Services	-	-	0,0030	0,0316	0,0058	0,0090
Public Administration	-	-	0,0033	0,0080	0,0000	0,0104

6.2.2 Correlation between labor income and beta/sector index

In this section, I will explore the correlation between labor income innovations and different stock portfolios. Besides the beta decile index portfolio I used before, I also utilize the sector index portfolio from Datastream. In this section, I will firstly present the results by only using industrial category. Then the results by using both educational level and industrial category will be discussed as well.

In table 8, we can see the correlation and beta between labor income process by industry and beta decile index returns from CRSP. We can see that in general, households in construction, manufacturing and professional and related services have strong positive correlation between their labor income process and beta decile index returns. In addition, the relationship is prominent regardless of different beta levels. As I mentioned before, a high positive correlation between labor income innovations and stock portfolio returns means that households implicitly hold risky assets in their overall portfolio. It is optimal for them to tilt towards riskless asset or assets that have very low correlation with stock portfolios returns under this situation. As construction and manufacturing are traditionally pro-cyclical industries, this finding will be in alignment with a pretty intuitive suggestion from a financial counselor, that is, people working in pro-cyclical industry should hedge their income risk by investing in safer investment products that have low correlation with stock market. For households from other industries in the table, one implication is that it is preferable for them to invest in low to medium beta level stock portfolios because of the low correlation between their labor income and stock returns. However, they would not

hold stock investment as much as an agent whose labor income process is completely uncorrelated with stock returns.

Then I compared the labor income process in different industries with sector index portfolio returns from Datastream. The result can be found in table 9. As we can see, there is not any single beta estimation is statistically significant this time. What's more, both the correlation and beta have quite low values compared to the results in previous table. The discrepancy between industry definition from the Census of Population and sector category from Global Industry Classification Standard (GICS) might be one of the reasons for the ambiguous relationship between labor income and stock portfolio returns.

Since households with different education level have very different labor income innovations, the last part of my study explored how the correlation varies across both educational level and industrial category. Again, sector index returns from Datastream have not shown any clear correlation with the labor income process and the results can be found in appendix. I will focus the discussion on the results from beta decile index in this paragraph. The results can be found in table 10, 11 and 12 for different educational level. The cells with larger sub-sample size are more likely to have statistically significant estimates for beta. From the result, we can see that all the construction and manufacturing cells, except for construction with a college education cell, have very high correlation between labor income process and beta decile index returns. The correlations are generally more than 0,15 and even around 0,40 under some cases. Besides that, the tables also show that for households in college education group, their labor income process will be positively correlated with beta decile index returns if they work in wholesale and retail trade, professional and related services and public administration. In other words, their explicit optimal portfolio will tilt towards safer assets as well.

Table 8: Correlation and Beta between Labor Income and Beta Decile Index Returns: Industry Only

	Beta Decile Group 1		Beta Decile Group 2		Beta Decile Group 3		Beta Decile Group 4		Beta Decile Group 5	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	0,0222	0,0442	-0,0125	0,0355	-0,0724	0,0195	-0,0330	0,0386	-0,1028	0,0004
Construction	0,2955	0,0704**	0,3579	0,0851**	0,4438	0,1071**	0,4392	0,1196**	0,4204	0,1035**
Manufacturing	0,3180	0,0790**	0,3478	0,0894*	0,3714	0,1041**	0,4135	0,1222**	0,4127	0,1058*
Wholesale and Retail Trade	0,0954	0,0532	0,0939	0,0567	0,1362	0,0699	0,1344	0,0785	0,1656	0,0690
Finance, Insurance, and Real Estate	-0,0156	0,0464	0,0114	0,0525	-0,0150	0,0565	0,0541	0,0767	0,0412	0,0604
Professional and Related Services	0,2610	0,0732*	0,3128	0,0840*	0,3226	0,0967*	0,3521	0,1121**	0,4002	0,1008*
Public Administration	-0,0664	0,0434	0,0345	0,0523	0,0464	0,0616	0,0908	0,0734	0,1696	0,0665

	Beta Decile Group 6		Beta Decile Group 7		Beta Decile Group 8		Beta Decile Group 9		Beta Decile Group 10	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,0367	0,0302	-0,0678	0,0106	-0,0676	0,0149	-0,1043	0,0177	-0,0560	0,0473
Construction	0,4510	0,1162**	0,3755	0,0992*	0,3671	0,1055**	0,3541	0,1379**	0,1750	0,1057*
Manufacturing	0,4259	0,1163**	0,4213	0,1081*	0,3764	0,1112*	0,3928	0,1548**	0,2854	0,1372**
Wholesale and Retail Trade	0,1917	0,0786	0,2082	0,0738	0,1775	0,0769	0,2554	0,1215*	0,1469	0,1055*
Finance, Insurance, and Real Estate	0,0827	0,0742	0,0846	0,0671	0,0443	0,0662	0,0532	0,1015	-0,0548	0,0813
Professional and Related Services	0,3938	0,1087*	0,4108	0,1026*	0,3538	0,1049*	0,4100	0,1534**	0,2698	0,1341**
Public Administration	0,1737	0,0732	0,1964	0,0679	0,1410	0,0697	0,2367	0,1134*	0,0706	0,0974*

* Significance at the 10 percent level.

** Significance at the 5 percent level.

Table 9: Correlation and Beta between Labor Income and Sector Index Returns: Industry Only

	Market Index		Oil & Gas		Basic Mats		Industrials		Consumer Gds		Health Care	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,2542	-0,0698	-0,0907	-0,0470	-0,1828	-0,0741	-0,2526	-0,0777	-0,2910	-0,0759	-0,1924	-0,0333
Construction	-0,0555	0,0542	0,0832	0,0083	0,0905	0,0311	-0,0211	0,0353	-0,0569	0,0352	-0,1348	0,0424
Manufacturing	-0,1649	0,0565	0,0654	0,0024	0,1490	0,0414	-0,0315	0,0448	-0,0673	0,0467	-0,1697	0,0584
Wholesale and Retail Trade	-0,2270	0,0358	-0,1648	-0,0286	-0,1510	-0,0045	-0,1516	0,0220	-0,1142	0,0336	-0,2582	0,0340
Finance, Insurance, and Real Estate	-0,2203	0,0249	-0,0968	-0,0264	0,0248	0,0240	-0,1796	0,0090	-0,1299	0,0273	-0,1837	0,0364
Professional and Related Services	-0,1269	0,0665	-0,0413	-0,0118	-0,0142	0,0188	-0,0443	0,0451	-0,0211	0,0539	-0,1482	0,0660
Public Administration	-0,1894	0,0531	-0,2818	-0,0343	-0,2397	-0,0069	-0,1658	0,0285	-0,0534	0,0451	-0,1958	0,0547

	Consumer Svs		Telecom		Utilities		Financials		Technology	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,0903	-0,0019	-0,2261	-0,0361	-0,1742	-0,0638	-0,2506	-0,0487	-0,3056	-0,0733
Construction	0,0517	0,0415	-0,2306	0,0365	-0,0379	0,0292	0,0407	0,0617	-0,1825	0,0171
Manufacturing	0,0082	0,0450	-0,3801	0,0415	-0,0464	0,0387	-0,0782	0,0619	-0,3640	0,0120
Wholesale and Retail Trade	-0,0377	0,0338	-0,2925	0,0402	-0,0941	0,0243	-0,2277	0,0342	-0,1729	0,0252
Finance, Insurance, and Real Estate	-0,0519	0,0321	-0,2852	0,0272	-0,1434	0,0068	-0,2132	0,0274	-0,2587	0,0049
Professional and Related Services	0,0963	0,0548	-0,3324	0,0554	0,0398	0,0536	-0,0198	0,0710	-0,2140	0,0318
Public Administration	0,0286	0,0424	-0,0432	0,0803	0,0376	0,0467	-0,1228	0,0541	-0,1685	0,0330

Table 10: Correlation and Beta between Labor Income and Beta Decile Index Returns: Education and Industry – No High School

	Beta Decile Group 1		Beta Decile Group 2		Beta Decile Group 3		Beta Decile Group 4		Beta Decile Group 5	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	0,0858	0,0843	0,0198	0,0677	-0,0070	0,0675	0,0362	0,0957	-0,0184	0,0550
Construction	0,1729	0,0687	0,2424	0,0891*	0,2988	0,1121**	0,2843	0,1225**	0,2400	0,0989
Manufacturing	0,2290	0,0783*	0,3092	0,0911*	0,3206	0,1051*	0,3685	0,1234*	0,3122	0,0991
Wholesale and Retail Trade	-0,1153	0,0293	-0,1280	0,0269	-0,1393	0,0308	-0,0666	0,0480	-0,0342	0,0417
Finance, Insurance, and Real Estate	0,0466	0,0689	0,0722	0,0832	0,0647	0,0916	0,1115	0,1253	0,0598	0,0867
Professional and Related Services	0,1208	0,0817	0,0926	0,0830	0,1074	0,0983	0,0331	0,0946	0,0881	0,0880
Public Administration	-0,0476	0,0369	0,0199	0,0685	0,0133	0,0755	0,0058	0,0808	0,0071	0,0655

	Beta Decile Group 6		Beta Decile Group 7		Beta Decile Group 8		Beta Decile Group 9		Beta Decile Group 10	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	0,0181	0,0779	0,0124	0,0662	0,0269	0,0813	-0,0579	0,0749	0,0107	0,1152
Construction	0,2795	0,1163*	0,2170	0,0955	0,2546	0,1123*	0,1850	0,1281*	0,0639	0,0942
Manufacturing	0,3767	0,1148*	0,2982	0,0970	0,2760	0,1041	0,2737	0,1489*	0,1909	0,1394**
Wholesale and Retail Trade	-0,0299	0,0470	0,0331	0,0518	-0,0193	0,0478	0,0592	0,0933	0,1256	0,1074*
Finance, Insurance, and Real Estate	0,0889	0,1072	0,0521	0,0820	0,1003	0,1120	0,0012	0,1052	-0,0230	0,1007
Professional and Related Services	0,0398	0,0855	-0,0019	0,0661	0,0939	0,0972	0,0414	0,1283	-0,0357	0,1127
Public Administration	0,0163	0,0777	-0,0265	0,0427	-0,0305	0,0473	-0,0349	0,0789	-0,0418	0,0806

* Significance at the 10 percent level.

** Significance at the 5 percent level.

Table 11: Correlation and Beta between Labor Income and Beta Decile Index Returns: Education and Industry – High School

	Beta Decile Group 1		Beta Decile Group 2		Beta Decile Group 3		Beta Decile Group 4		Beta Decile Group 5	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	0,0734	0,0467	0,0156	0,0328	-0,0553	0,0126	0,0056	0,0386	-0,0418	0,0128
Construction	0,2801	0,0716**	0,3091	0,0834**	0,3976	0,1072**	0,3761	0,1168**	0,3693	0,1027**
Manufacturing	0,3359	0,0813**	0,3372	0,0899**	0,3613	0,1048**	0,3953	0,1228**	0,4021	0,1085**
Wholesale and Retail Trade	0,0830	0,0566	0,0831	0,0605	0,1188	0,0746	0,1029	0,0816	0,1087	0,0685
Finance, Insurance, and Real Estate	-0,0433	0,0417	-0,0474	0,0423	-0,0478	0,0496	-0,0090	0,0631	-0,0527	0,0422
Professional and Related Services	-0,0404	0,0550	-0,0459	0,0562	-0,0743	0,0617	-0,0666	0,0711	-0,0239	0,0612
Public Administration	-0,0900	0,0412	-0,0378	0,0469	-0,0420	0,0543	-0,0049	0,0654	0,1031	0,0633

	Beta Decile Group 6		Beta Decile Group 7		Beta Decile Group 8		Beta Decile Group 9		Beta Decile Group 10	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,0307	0,0193	-0,0013	0,0272	-0,0756	-0,0011	-0,0778	0,0112	0,0490	0,0716
Construction	0,3837	0,1132**	0,3065	0,0944*	0,2847	0,0977*	0,2944	0,1324**	0,1268	0,0979*
Manufacturing	0,4073	0,1181**	0,4253	0,1140**	0,3784	0,1154**	0,3949	0,1577**	0,2847	0,1356**
Wholesale and Retail Trade	0,1460	0,0808	0,1267	0,0704	0,1216	0,0768	0,1941	0,1248*	0,0659	0,1020
Finance, Insurance, and Real Estate	0,0138	0,0592	-0,0385	0,0420	-0,0328	0,0493	-0,0098	0,0847	-0,0791	0,0743
Professional and Related Services	-0,0159	0,0691	-0,0045	0,0615	-0,0613	0,0597	0,0169	0,1132	-0,1146	0,0933
Public Administration	0,0672	0,0653	0,1404	0,0661	0,0881	0,0671	0,1670	0,1107*	0,0558	0,0980*

* Significance at the 10 percent level.

** Significance at the 5 percent level.

Table 12: Correlation and Beta between Labor Income and Beta Decile Index Returns: Education and Industry – College

	Beta Decile Group 1		Beta Decile Group 2		Beta Decile Group 3		Beta Decile Group 4		Beta Decile Group 5	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,1287	-0,0869	-0,1237	-0,1002	-0,1796	-0,1653	-0,1353	-0,1344	-0,2126	-0,2187
Construction	-0,1931	0,0207	-0,1443	0,0259	-0,1601	0,0289	-0,0974	0,0454	-0,0713	0,0377
Manufacturing	0,1916	0,0730*	0,2606	0,0871*	0,2917	0,1031*	0,3371	0,1226**	0,3866	0,1127*
Wholesale and Retail Trade	0,2140	0,0690*	0,2207	0,0765*	0,2727	0,0945*	0,2237	0,0983*	0,3155	0,0991*
Finance, Insurance, and Real Estate	0,0336	0,0541	0,0752	0,0656	0,0393	0,0675	0,1092	0,0929	0,1336	0,0840
Professional and Related Services	0,3539	0,0802**	0,4208	0,0945**	0,4388	0,1089**	0,4817	0,1276**	0,5179	0,1156**
Public Administration	0,1256	0,0537	0,2131	0,0657	0,2442	0,0781*	0,2857	0,0927*	0,2749	0,0781*

	Beta Decile Group 6		Beta Decile Group 7		Beta Decile Group 8		Beta Decile Group 9		Beta Decile Group 10	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,1423	-0,1493	-0,1791	-0,1970	-0,1384	-0,1527	-0,1811	-0,2247	-0,1469	-0,1541
Construction	-0,0460	0,0473	-0,0619	0,0365	-0,1293	0,0281	-0,0140	0,0835	-0,0068	0,0882
Manufacturing	0,3637	0,1187*	0,3957	0,1156*	0,3551	0,1187*	0,3728	0,1648**	0,2696	0,1448**
Wholesale and Retail Trade	0,2907	0,1030*	0,3250	0,1024*	0,2680	0,1004*	0,2988	0,1424**	0,1854	0,1183*
Finance, Insurance, and Real Estate	0,1475	0,0948	0,1932	0,0990	0,1115	0,0862	0,1284	0,1271	-0,0121	0,0894
Professional and Related Services	0,5190	0,1251**	0,5441	0,1209**	0,4819	0,1223**	0,5291	0,1711***	0,4138	0,1507**
Public Administration	0,3440	0,0929*	0,2843	0,0794	0,2369	0,0808	0,3180	0,1249**	0,2050	0,1088**

* Significance at the 10 percent level.

** Significance at the 5 percent level.

*** Significance at the 1 percent level.

7 Limitations

In this section, I will discuss about shortcomings in the methodology and pitfalls of the datasets regarding the whole study.

One of the first concerns regarding my study is about the data. As I focus my study on the labor income heterogeneity across educational levels and industrial categories, how these two demographic metrics are processed will have great influence on the reliability of final results. One might argue that there should be some continuity in the labor income even though the household gets to a new educational level, which may cause a problem as I regard household with a new educational level as a new separate household. One remedy to that problem is that we can always keep only the last educational level labor income information for each household. However, this would not be a huge problem for PSID data as I only find few households who have educational level change in the sample.

In regard to the industrial categories, there are several issues popping out in my study. Firstly, it ignored the influence from the wife's industry while the wife's labor income is a major component in household total labor income. The difference between head's and wife's labor income risk will diminish the reliability of results across industries. Secondly, the labor income reported in PSID for each wave is for the previous calendar year but the industry reported is for head's current job. There might be some timing issues if the head has changed his job in between. Thirdly, as the grouping criteria of industrial categories have been changed since 2003 PSID wave, I have to use the industry name to match the data before and after the change. The risk profile of household labor income process might be changed due to the fact that some occupations are added into or deleted from the 9 industries I investigate. This will be problematic for the accuracy of the analysis.

Besides, even though PSID is the longest and the most comprehensive survey in this area, the limited sample size in some cells across educational level and industrial category makes it hard for us to draw a general picture of their risk profile of labor income process, which means the results can only be used to solve the optimal portfolio choice problem for some sub-groups in the whole labor force.

8 Conclusion

In this thesis, I investigated the heterogeneity of labor income process across both educational level and industrial category and draw the implications about how the heterogeneity will influence optimal portfolio choice for different households. By breaking down the PSID data into small cells, I have found there are major differences in different cells regarding labor income risk profile. The result of this study is relevant in the context of current literature. It can help to justify the needs for customized investment portfolio in a defined contribution pension plan[10], which means the risk profile of household labor income will play an important role in deciding household risk appetite. Besides, household that are exposed to high riskiness of labor income should adjust their portfolio choice for their own savings accounts accordingly.

All in all, my studies have strengthened several findings in previous studies of labor income process uncertainty and provided more insights in the heterogeneity of labor income process. It has shown that the effect of transitory shock to labor income process is gradually decreasing as the household education level increases while in the same time, households from college group have a lower total income variance compared with the rest two educational groups. In regard to heterogeneity of labor income by industrial category, the study reveals that manufacturing and construction have shown persistent positive correlation with beta decile index returns across different educational levels. On the contrary, industries, such as public administration, professional services and so on, have mixed correlations with beta decile index returns for different educational levels.

What's more, the results generate some opposite implications regarding optimal portfolio choice problem. For example, when we look the household labor income process at industrial level, a pair of interesting findings in my study is that: 1) households from manufacturing industry have the lowest total variance of labor income process across all three educational levels; 2) for households from manufacturing industry, their labor income innovations are highly positively correlated

[10] Different from defined benefit plan, contributions are paid into an individual account for each member in a defined contribution plan. The contributions are invested based on individual choices and the returns on the investment (which may be positive or negative) are credited to the individual's account.

with beta decile index portfolio returns based on historical data. The simulation results in Cocco, Gomes and Maenhout (2005) have shown that higher total variance of labor income innovations or higher positive correlation between labor income process and stock returns will incur a lower proportion of asset to be invested in stock market, which means our findings regarding these two factors will have opposite influence on the optimal portfolio choice problem.

For further studies, one can use simulation methodology to analyze the sensitivity of asset allocation choices on these two factors. One can also use the empirical data of actual asset allocation for households from different educational and industrial cells to test these findings and their influences.

Reference

Attanasio, O.P. 1995, "The intertemporal allocation of consumption: theory and evidence", Carnegie-Rochester Conference Series on Public PolicyElsevier, , pp. 39.

Bodie, Z., Merton, R.C. & Samuelson, W.F. 1992, "Labor supply flexibility and portfolio choice in a life cycle model", Journal of Economic Dynamics and Control, vol. 16, no. 3, pp. 427-449.

Brennan, M.J., Schwartz, E.S. & Lagnado, R. 1997, "Strategic asset allocation", Journal of Economic Dynamics and Control, vol. 21, no. 8, pp. 1377-1403.

Brennan, M.J. & Xia, Y. 1998, "Resolution of a financial puzzle", .

Burtless, G. 2000, "Social security privatization and financial market risk", .

Campbell, J.Y., Cocco, J.F., Gomes, F.J. & Maenhout, P.J. 2001, "Investing retirement wealth: A life-cycle model" in Risk aspects of investment-based Social Security reform University of Chicago Press, , pp. 439-482.

Campbell, J.Y. & Viceira, L.M. 1999, "Consumption and portfolio decisions when expected returns are time varying", The Quarterly Journal of Economics, vol. 114, no. 2, pp. 433-495.

Canner, N., Mankiw, N.G. & Weil, D.N. 1997, "An asset allocation puzzle", The American Economic Review, vol. 87, no. 1, pp. 181-191.

Carroll, C.D. & Samwick, A.A. 1997, "The nature of precautionary wealth", Journal of Monetary Economics, vol. 40, no. 1, pp. 41-71.

Cocco, J.F., Gomes, F.J. & Maenhout, P.J. 2005, "Consumption and portfolio choice over the life cycle", Review of Financial Studies, vol. 18, no. 2, pp. 491-533.

Conesa, J.C. & Krueger, D. 1999, "Social security reform with heterogeneous agents", Review of Economic Dynamics, vol. 2, no. 4, pp. 757-795.

Davis, S.J. & Willen, P. 2012, "Occupation-level income shocks and asset returns: their covariance and implications for portfolio choice", The Quarterly Journal of Finance, .

Davis, S.J. & Willen, P. 2000, "Using financial assets to hedge labor income risks: estimating the benefits", manuscript, University of Chicago, .

Davis, S.J. & Willen, P. 2000, "Using financial assets to hedge labor income risks: estimating the benefits", manuscript, University of Chicago, .

Fehr, H. & Habermann, C. 2008, "Risk Sharing and Efficiency Implications of Progressive Pension Arrangements*", The Scandinavian Journal of Economics, vol. 110, no. 2, pp. 419-443.

Genakoplos, J., Mitchell, O.S. & Zeldes, S.P. 2000, Would a privatized social security system really pay a higher rate of return, .

Gomes, F. & Michaelides, A.G. 2004, A human capital explanation for an asset allocation puzzle? Financial Markets Group, London School of Economics and Political Science.

Gramlich, E.M. 1996, "Different approaches for dealing with Social Security", The American Economic Review, , pp. 358-362.

Heaton, J. & Lucas, D. 2000, "Portfolio choice and asset prices: The importance of entrepreneurial risk", The journal of finance, vol. 55, no. 3, pp. 1163-1198.

Hubbard, R.G., Skinner, J. & Zeldes, S.P. 1994, "The importance of precautionary motives in explaining individual and aggregate saving", Carnegie-Rochester Conference Series on Public PolicyElsevier, , pp. 59.

Kotlikoff, L.J. 1996, "Privatization of social security: how it works and why it matters" in Tax Policy and the Economy, Volume 10 MIT Press, , pp. 1-32.

Kotlikoff, L.J., Smetters, K. & Walliser, J. 1999, "Privatizing social security in the United States—comparing the options", Review of Economic Dynamics, vol. 2, no. 3, pp. 532-574.

Merton, R.C. 1971, "Optimum consumption and portfolio rules in a continuous-time model", Journal of Economic Theory, vol. 3, no. 4, pp. 373-413.

Merton, R.C. 1969, "Lifetime portfolio selection under uncertainty: The continuous-time case", Review of Economics and statistics, vol. 51, no. 3, pp. 247-257.

Mitchell, O.S. & Zeldes, S.P. 1996, Social security privatization: a structure for analysis, .

Nishiyama, S. & Smetters, K. 2007, "Does social security privatization produce efficiency gains?", The Quarterly Journal of Economics, vol. 122, no. 4, pp. 1677-1719.

Samuelson, P. 1969, "Lifetime portfolio selection by dynamic stochastic programming", .

Wachter, J.A. 2003, "Risk aversion and allocation to long-term bonds", Journal of Economic Theory, vol. 112, no. 2, pp. 325-333.

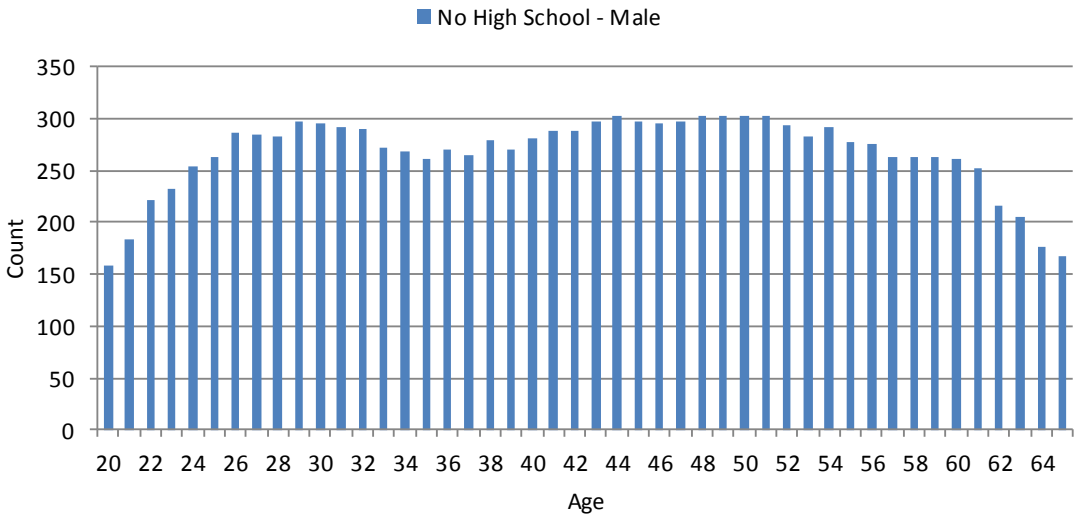
Appendix

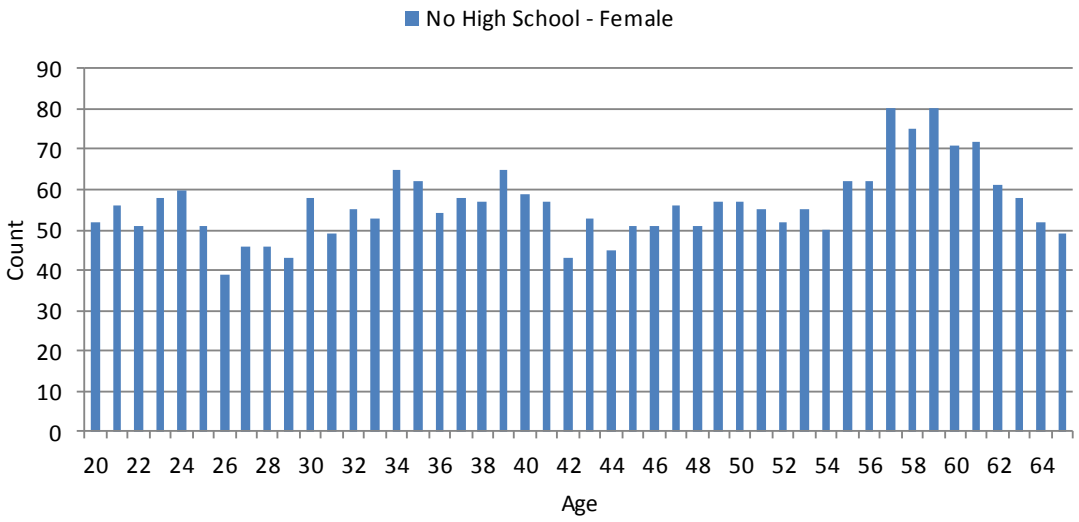
Table 13: Cell Size for Different Gender and Educational Level

	Male	Female
No High School	12285	2592
High School	41205	8904
College	20567	3357

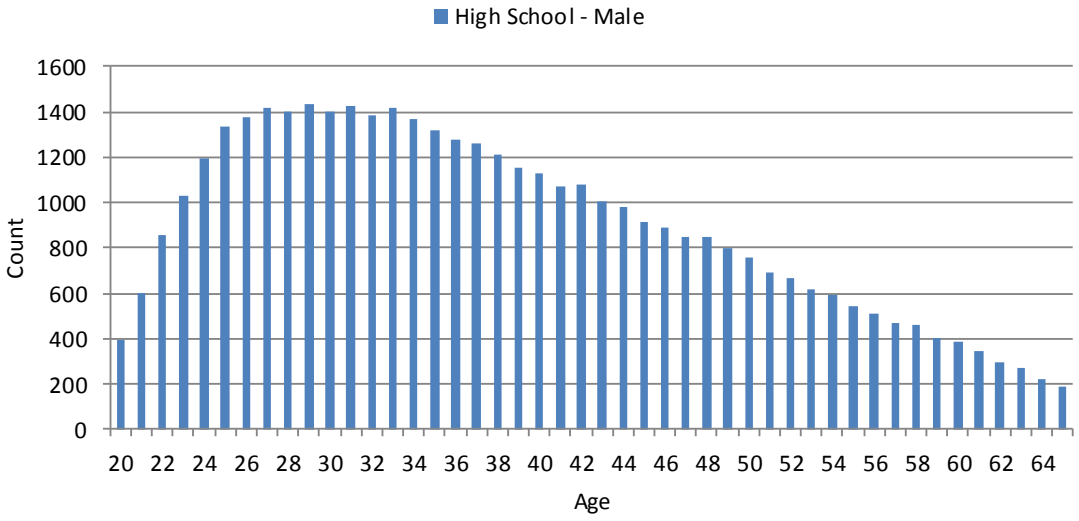
This table has shown the family-year counts for each cell with different gender and educational level combination, from which we can see that there are very limited size for household with female head.

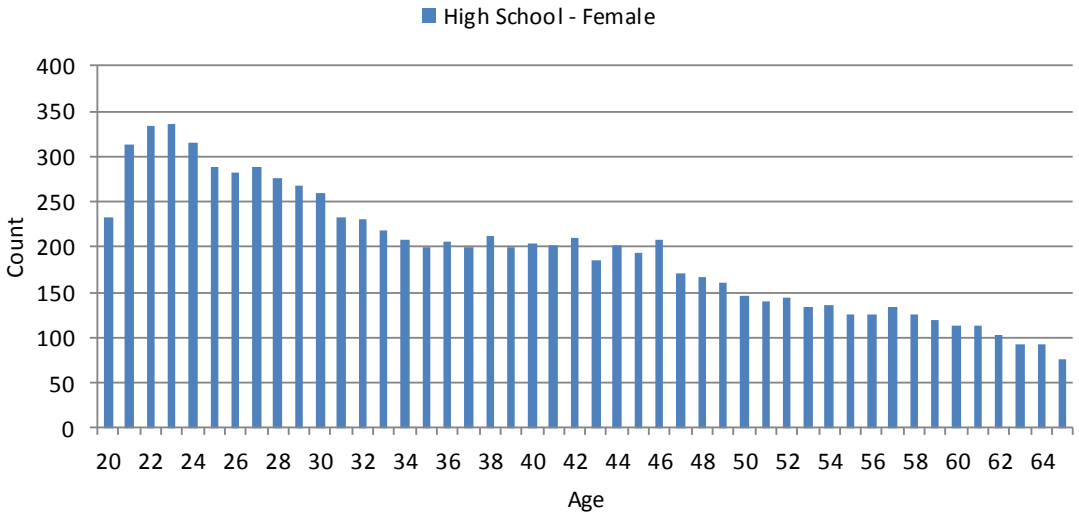
Graph 2: Age Distribution for Different Gender – No High School



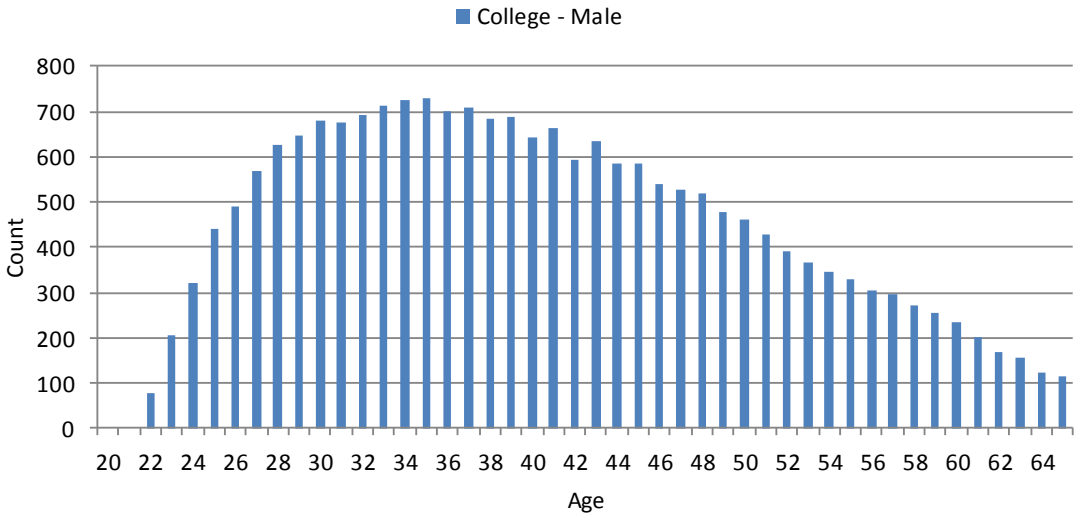


Graph 3: Age Distribution for Different Gender – High School





Graph 4: Age Distribution for Different Gender – College



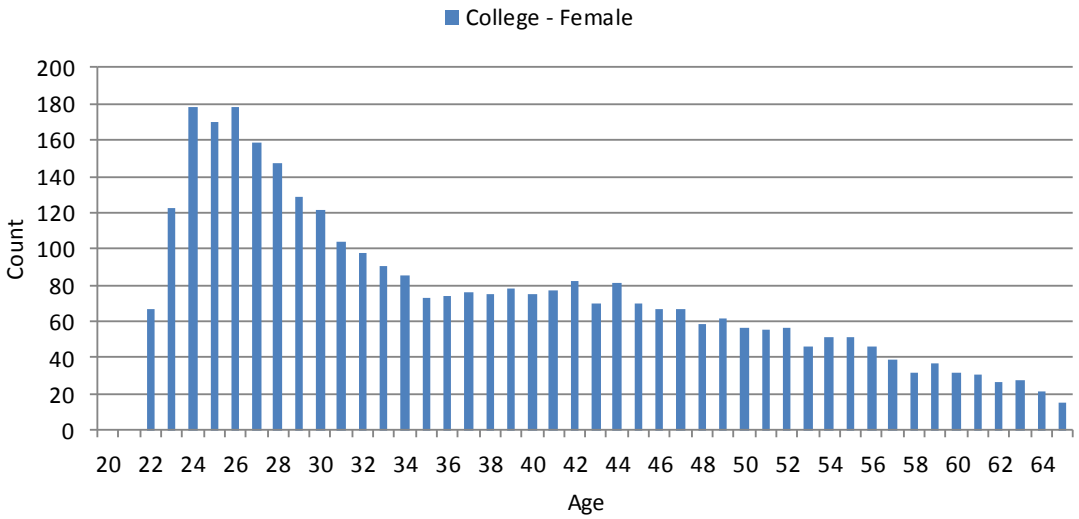


Table 14: Industrial Categories in Census 1970 and 2000 with Corresponding Scores

	2000 Census Industrial Name (since 2003 wave)	2000 Score	1970 Census Industrial Name	1970 Score
1	Agriculture, Forestry, Fishing, and Hunting	17 - 29	Agriculture, Forestry, and Fisheries	17 - 28
2	Mining	37 - 49	Mining	47 - 57
3	Construction	77	Construction	67 - 77
4	Manufacturing	107 - 399	Manufacturing	107 - 398
5	Wholesale Trade	407 - 459	Wholesale and Retail Trade	507 - 698
	Retail Trade	467 - 579		
6	Finance and Insurance	687 - 699	Finance, Insurance, and Real Estate	707 - 718
7	Professional, Scientific, and Technical Services	727 - 749	Professional and Related Services	828 - 897
8	Arts, Entertainment, and Recreation	856 - 859	Entertainment and Recreation Services	807 - 809
9	Public Administration and Active Duty Military	937 - 987	Public Administration	907 - 937
10			Transportation, Communications, and Other Public Utilities	407 - 479
11			Business and Repair Services	727 - 759
12			Personal Services	769 - 798

In the raw data, head industry of a household is recorded as a three-digit number depending on his occupations. This table only tabulates those industries that have been investigated in my study and the first column is the final sequence number in the dataset.

Table 15: Correlation and Beta between Labor Income and Sector Index Returns: Education and Industry – No High School

	Market Index		Oil & Gas		Basic Mats		Industrials		Consumer Gds		Health Care	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,1336	0,0160	0,0012	-0,0066	0,0145	0,0298	-0,1171	-0,0031	-0,2494	-0,0476	0,0343	0,1094
Construction	-0,0539	0,0482	-0,0288	-0,0127	0,1397	0,0540	-0,0466	0,0266	0,0076	0,0445	-0,1265	0,0301
Manufacturing	-0,1451	0,0776	-0,0041	-0,0083	0,1366	0,0405	-0,0390	0,0536	-0,0675	0,0580	-0,1727	0,0768
Wholesale and Retail Trade	0,0800	0,0853	-0,1355	-0,0329	-0,0839	-0,0009	0,0642	0,0534	0,1395	0,0696	-0,0321	0,0622
Finance, Insurance, and Real Estate	0,1821	0,1805	0,0220	0,0021	0,1815	0,1063	0,0585	0,0789	-0,0755	0,0200	0,3390	0,2594**
Professional and Related Services	-0,2137	0,0415	-0,1288	-0,0397	-0,1769	-0,0210	-0,2308	0,0048	-0,3180	-0,0003	-0,1391	0,0648
Public Administration	0,1111	0,1677	-0,1502	-0,1069	-0,0231	0,0056	-0,0053	0,0494	0,1278	0,1312	0,0560	0,1277

	Consumer Svs		Telecom		Utilities		Financials		Technology	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,0922	0,0110	-0,0927	0,0564	0,0377	0,0743	-0,0554	0,0544	-0,3391	-0,0771
Construction	0,0976	0,0552	-0,2307	0,0131	-0,1482	-0,0105	-0,0835	0,0368	-0,0860	0,0222
Manufacturing	0,0574	0,0576	-0,4467	0,0585	-0,2233	0,0235	-0,1320	0,0709	-0,1805	0,0437
Wholesale and Retail Trade	0,2037	0,0699	0,2205	0,1213*	0,1274	0,0696	0,0427	0,0667	0,2688	0,0843*
Finance, Insurance, and Real Estate	0,2042	0,1256	0,2700	0,2265*	0,1475	0,1270	0,2291	0,1716	-0,0425	0,0485
Professional and Related Services	-0,1743	0,0190	-0,1604	0,0742	-0,1338	0,0168	-0,0080	0,0848	-0,3057	0,0034
Public Administration	0,1869	0,1474	0,2962	0,2935*	0,0838	0,1168	-0,0120	0,0697	0,1330	0,1230

* Significance at the 10 percent level.

** Significance at the 5 percent level.

Table 16: Correlation and Beta between Labor Income and Sector Index Returns: Education and Industry – High School

	Market Index		Oil & Gas		Basic Mats		Industrials		Consumer Gds		Health Care	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,0998	-0,0077	0,0710	0,0287	-0,0743	-0,0247	-0,0702	-0,0056	-0,1788	-0,0410	-0,0972	-0,0037
Construction	-0,0752	0,0458	0,1291	0,0194	-0,0162	0,0119	-0,0301	0,0315	-0,1291	0,0204	-0,1681	0,0293
Manufacturing	-0,1477	0,0480	0,1192	0,0135	0,1622	0,0459	0,0058	0,0457	-0,0671	0,0406	-0,1420	0,0514
Wholesale and Retail Trade	-0,2457	0,0274	-0,1457	-0,0315	-0,1969	-0,0176	-0,1698	0,0162	-0,1492	0,0277	-0,2890	0,0228
Finance, Insurance, and Real Estate	-0,3697	-0,0084	-0,3135	-0,0695	-0,0889	-0,0007	-0,3175	-0,0179	-0,2137	0,0122	-0,3536	-0,0003
Professional and Related Services	-0,4611	0,0008	-0,1723	-0,0390	-0,4839	-0,0693	-0,3394	-0,0040	-0,4215	-0,0035	-0,4422	0,0099
Public Administration	-0,2175	0,0454	-0,2459	-0,0363	-0,2422	-0,0126	-0,1829	0,0233	-0,1132	0,0383	-0,1794	0,0529

	Consumer Svs		Telecom		Utilities		Financials		Technology	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	0,0282	0,0326	-0,1505	-0,0195	-0,1385	-0,0514	-0,2501	-0,0608	-0,0782	-0,0015
Construction	-0,0404	0,0271	-0,2438	0,0249	-0,0360	0,0268	0,0531	0,0622	-0,1634	0,0144
Manufacturing	-0,0271	0,0367	-0,3414	0,0301	0,0050	0,0434	-0,0503	0,0579	-0,3775	-0,0014
Wholesale and Retail Trade	-0,0827	0,0286	-0,3512	0,0245	-0,1631	0,0081	-0,2481	0,0277	-0,1783	0,0222
Finance, Insurance, and Real Estate	-0,1333	0,0179	-0,4926	-0,0146	-0,3461	-0,0426	-0,3776	-0,0018	-0,2529	0,0060
Professional and Related Services	-0,3107	0,0044	-0,3936	0,0351	-0,2506	-0,0022	-0,3697	0,0229	-0,3853	0,0033
Public Administration	-0,0232	0,0381	-0,1112	0,0719	0,0327	0,0474	-0,1367	0,0508	-0,1836	0,0291

Table 17: Correlation and Beta between Labor Income and Sector Index Returns: Education and Industry – College

	Market Index		Oil & Gas		Basic Mats		Industrials		Consumer Gds		Health Care	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,2010	-0,2587	-0,2153	-0,2777	-0,2298	-0,3014	-0,2316	-0,2729	-0,2492	-0,2579	-0,0798	-0,0848
Construction	-0,1640	0,0332	-0,2607	-0,0642	-0,2687	-0,0445	-0,1742	0,0066	-0,1213	0,0260	0,0065	0,0759
Manufacturing	0,0110	0,0859	-0,0021	-0,0071	0,1722	0,0498	0,0017	0,0510	0,0992	0,0702	-0,0350	0,0796
Wholesale and Retail Trade	-0,1764	0,0340	0,0037	-0,0049	0,0109	0,0193	-0,1014	0,0239	-0,0766	0,0347	-0,1687	0,0383
Finance, Insurance, and Real Estate	-0,0959	0,0473	0,1005	0,0211	0,1014	0,0472	-0,0458	0,0336	-0,0144	0,0475	-0,1059	0,0470
Professional and Related Services	0,0749	0,0904	0,0563	0,0012	0,2180	0,0502	0,1336	0,0652	0,2064	0,0771	0,0142	0,0828
Public Administration	-0,0732	0,0580	-0,1212	-0,0210	-0,0406	0,0114	0,0062	0,0419	0,0394	0,0502	-0,1249	0,0528

	Consumer Svs		Telecom		Utilities		Financials		Technology	
Industry	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta	Correlation	Beta
Agriculture, Forestry, and Fisheries	-0,0190	-0,0080	-0,1988	-0,2261	-0,1505	-0,2196	-0,1183	-0,1116	-0,2150	-0,1997
Construction	-0,1261	0,0170	0,2017	0,1280*	-0,0270	0,0345	0,0124	0,0664	-0,2327	0,0054
Manufacturing	0,2423	0,0771	-0,0944	0,0811	0,1109	0,0696	0,0675	0,0830	-0,1715	0,0310
Wholesale and Retail Trade	-0,1542	0,0147	-0,1514	0,0521	0,0222	0,0447	-0,1169	0,0423	-0,2689	0,0049
Finance, Insurance, and Real Estate	0,0094	0,0430	-0,1380	0,0497	0,0001	0,0432	-0,0751	0,0486	-0,2191	0,0013
Professional and Related Services	0,2903	0,0745	-0,2126	0,0628	0,1976	0,0773	0,1352	0,0852	-0,0590	0,0435
Public Administration	0,0423	0,0414	-0,0072	0,0772	0,0042	0,0396	0,0215	0,0621	-0,1120	0,0315

* Significance at the 10 percent level.