Nonlinear Dependence and Households' Portfolio Decisions over the Life Cycle *

Wei Jiang[†] Shize Li[‡] Jialu Shen[§]

August 25, 2023

Abstract

This paper uncovers the nonlinear relationship between earning risk and stock returns, as measured by the between-squares correlation. By incorporating this between-squares correlation into a life-cycle model, we demonstrate that it lowers households' participation rate and generates moderate risky asset holdings. We identify two pathways through which the between-squares correlation affects portfolio choices: the skewness and kurtosis channels. The extent to which these channels dominate each other depends on the level of between-squares correlation, leading to a nonlinear relationship between this variable and household decisions. Our empirical studies support the model's predictions. Moreover, we find that ignoring between-squares correlations leads to substantial welfare loss and contributes to increasing wealth inequality.

JEL D31, D63, D91, E21, E32, G11.

Keywords: Nonlinear Dependence, Between-squares Correlation, Life-cycle Portfolio Choice and Consumption, Wealth Inequality, Equivalent Wealth Loss.

1 Introduction

Microeconomic data on household portfolios in the United States show that fewer than 50% of households participate in the stock market, and, conditional on one's participation in asset markets, average equity holdings as a share of financial wealth total only 55%.[1] However, life-cycle models

*We thank Francisco Gomes, Alex Michaelides, Michael O'Doherty, Stephan Siegel, Pengfei Sui (discussant), and seminar participants at the 2023 AEA poster session, SFS Cavalcade Asia-Pacific 2022, Bachelier World Congress 2022, Silicon Prairie Finance Conference 2023, and the PKU-NUS Annual International Conference in Quantitative Finance and Economics 2022, for their helpful comments.

†Hong Kong University of Science and Technology. E-mail: weijiang@ust.hk

‡Hong Kong University of Science and Technology. E-mail: slidq@connect.ust.hk

§Trulaske College of Business, University of Missouri. E-mail:jshen@missouri.edu

[1]See Ameriks and Zeldes (2011), Faig and Shum (2002), Heaton and Lucas (2000), and Poterba et al. (2001).

consistently predict all households should participate in the stock market and young households should invest almost all of their financial wealth in stocks.[2]

A crucial element that has been widely considered in addressing these two puzzles is labor income and the risk associated with it.[3] One strand of literature discussing labor income risk uses micro-data to calibrate the individual labor income process,[4] while another strand of literature considers the dependence between labor income and stock returns.[5] The latter approach relies on the correlations or cointegration between labor income and stock returns in order to fit low participation rates and moderate life-cycle stock holdings conditional on one's participation while avoiding counterfactually high risk aversion estimators. However, empirical support for correlation and cointegration is limited and mixed. Davis and Willen (2013, 2000), Bonaparte et al. (2014), and Catherine (2022) demonstrate a relatively large correlation, suggesting a considerable part of income risk can be hedged using stocks, but Campbell et al. (2001), Vissing-Jørgensen (2002), and Cocco et al. (2005) find an extremely low and insignificant correlation. With respect to cointegration, because of the unavailability of data, it is difficult to assess the magnitude of cointegration, and thus the empirical support for cointegration is minimal.

To cut through the mixed evidence, this work offers a novel solution. In contrast to previous literature focusing on linear dependence, this paper aims to shed new light on the dependence between labor income and stock returns. Specifically, we can get rid of the correlation. This paper emphasizes nonlinear dependence and introduces the "between-squares correlation," which we define as the correlation between the squares of labor income shocks and stock returns. By doing so, we offer a new perspective on the risky labor income process and delve into how our measure affects households' welfare and wealth inequality. To the best of our knowledge, this is the first paper capturing the nonlinear dependence between labor income and stock returns in the form of the between-squares correlation.

[2] For instance, Heaton and Lucas (2000), Koo (1998), Viceira (2001), Cocco et al. (2005), Gomes and Michaelides (2005a), and Polkovnichenko (2007) are typical examples of how labor income can be viewed as an implicit riskless asset that is abundant early in life and, therefore, induces higher stock market exposure in that period.

[3] Besides labor income, some other factors have been widely discussed as contributing to household portfolio decisions, such as household preference (see, e.g., Gomes (2005); Cao et al. (2005); Peijnenburg (2018); Pagel (2018)), participation costs (Vissing-Jørgensen (2002); Gomes and Michaelides (2005a)), peer effects (Hong et al. (2004)), housing (Cocco (2005); Yao and Zhang (2005)), and borrowing constraints (Guiso et al. (1996); Haliassos and Michaelides (2003)).

[4] See Jagannathan et al. (1996), Davis and Willen (2013), Campbell et al. (2001), Viceira (2001), Haliassos and Michaelides (2003), Cocco et al. (2005), Guvenen et al. (2014), Shen (2022), and Catherine (2022).

[5] Campbell et al. (2001), Viceira (2001), Cocco et al. (2005), and Gomes and Michaelides (2005b) explore the effect of contemporaneous correlation between labor income and stock market innovations. A number of studies allow for long-run codependence between labor income and stock returns in the form of cointegration between these two processes (Campbell (1996); Baxter and Jermann (1997); Lucas and Zeldes (2006); Santos and Veronesi (2006); Benzoni et al. (2007); Huggett and Kaplan (2011)). Lynch and Tan (2011) and Huggett and Kaplan (2016) consider labor income and stock returns under the context of the VAR framework.

The between-squares correlation measures the likelihood that extreme values jointly occur in the labor and stock markets, and, therefore, the analysis focuses on tail behavior. Such a specification is independent of any linear dependence structure, such as correlation or cointegration, and thus is consistent with empirically observed low contemporaneous correlations and cointegration between market returns and labor income shocks. Specifically, between-square correlations can capture events in which labor income shocks move opposite to stock return shocks, even though two shocks could be positively correlated.[6] Moreover, assuming two variables follow mixture normal distributions under some regularity conditions, we show that between-squares correlation is a sufficient measure to capture all other nonlinear dependence.

We start by analyzing a basic one-period portfolio optimization problem that considers higher moments of portfolio returns. Our findings show that the between-squares correlations between stock returns and labor income can significantly impact both the skewness and kurtosis of portfolio returns. It is worth noting that the presence of nonlinear dependence between labor income and stock returns highlights a new interpretation of the "stock-like" theory, where households' labor income becomes more "stock-like" due to the nonlinear dependence. Our study uncovers that the "stock-like" nature of labor income is mainly influenced by the skewness channel when the between-squares correlation is negative and by the kurtosis channel when the between-squares correlation is positive. This results in a nonlinear relationship between the between-squares correlation and portfolio allocations.

Then, using between-squares correlations, we revisit the participation decisions and optimal portfolio choices of households with nontradable labor income over their life cycle. Our approach has two main innovations. First, we introduce the nonlinear dependence between labor income and stock returns. Using the Panel Study of Income Dynamics Survey (PSID) and Center for Research in Security Prices (CRSP), we find that between-squares correlations are significantly positive, whereas correlations are insignificant. More specifically, the cross-sectional distributions of the between-squares correlations are right-skewed, indicating a significant fraction of households have an extremely positive between-squares correlation between their income and return shocks. Moreover, we categorize households into two groups: workers without college degrees versus workers with college degrees. Our results reveal notable heterogeneity across education groups. On average, college graduates exhibit higher between-squares correlation, and the distribution of between-squares

[6] For example, macroeconomic policies aimed at boosting economic growth, such as increased government spending, can have negative distributional effects on labor income, leading to reduced real purchasing power for workers. Increased competition for labor and resources resulting from infrastructure spending can also drive up prices and reduce profits in other industries. Additionally, policies that increase inflation or interest rates may result in a decrease in job opportunities and labor income. Therefore, by utilizing between-square correlation, researchers can analyze the joint occurrence of extreme values in the labor and stock markets that may arise due to these policies, enabling a more comprehensive understanding of their effects on workers and the economy.

correlations is more positively skewed compared with that of their counterparts. Therefore, our specification of the between-squares correlation in the model is consistent with empirical evidence.

Second, we allow higher moments in both labor income shocks and stock returns with the mixture normal distributions. The assumption of nonlinear dependence relies on the empirical evidence of higher-order moments in both labor income and stock returns. Higher-order moments in the labor income process are well documented and discussed in the literature (Guvenen et al. (2014); Shen (2022); Catherine (2022)), while higher-order moments in annual stock returns are neglected. The existence of non-normalities in daily stock returns has been recognized for at least 50 years, but higher moments of long-horizon returns are difficult to measure accurately. The recent paper by Neuberger and Payne (2021) stands out from this dearth of evidence: the authors find strong evidence of higher-order moments of annual stock returns. Motivated by this new evidence, we allow stock return innovations to deviate from normality, which allows for more flexibility when we construct our measure of nonlinear dependence.

We find that between-squares correlation is important in simultaneously explaining low stock market participation rates and moderate equity holdings for stock market participants. For each education group, we estimate the dynamics of labor income, annual stock returns, and between-squares correlations. Next, we incorporate these processes into the life-cycle model, which takes into account not only the higher-order moments in both income process and stock returns but also the nonlinear dependence between these two. The model matches well with the U.S. data with a moderate coefficient of relative risk aversion $\gamma = 4.8$ and a small participation cost, compared with $\gamma = 5.6$ in Shen (2022), $\gamma = 7.0$ in Catherine (2022) and $\gamma = 11$ in Fagereng et al. (2017).

Specifically, we find that households with between-squares correlations hold less shares of financial wealth invested in stocks (risky shares) conditional on their participation in the stock market compared with those without between-squares correlations. Moreover, after introducing the between-squares correlation, the wealth threshold of participation still exhibits a mild U-shaped pattern with respect to age, which is consistent with previous literature.[7] We also observe that the between-squares correlations increase the wealth threshold of participation across all ages, with a higher increase for households without college degrees. This is likely because the between-squares correlations increase uncertainty in the labor market, which leads to households being more cautious about investing in the stock market. Households without college degrees tend to accumulate less wealth and are generally less willing to participate in the stock market, making them more cautious about investing and requiring a higher level of wealth before they are willing to do so.

Our findings are robust to different levels of the household's risk aversion coefficient, such as

[7] See Fagereng et al. (2017).

$\gamma = 5.3$ and $\gamma = 4.3$. We also find optimal risky asset holdings are very sensitive to risk aversion coefficients, a finding that is consistent with the empirical observation that asset holdings are highly heterogeneous. Moreover, we consider heterogeneity in the elasticity of intertemporal substitution (EIS) and the discount factor. The EIS has a limited impact on optimal risky asset holdings across age groups, while the discount factor has a stronger effect on risky asset holdings. Households with higher discount factors tend to accumulate less wealth and live paycheck to paycheck. Hence, they start holding risky assets when their normalized cash on hands reach a certain level.

In addition, we provide empirical support for our model's predictions. Using the PSID, we investigate the possibility of a relation between households' portfolio decisions (participation decisions and risky asset holdings) and between-squares correlations. First, we find that the more the between-squares correlation deviates from zero, the less likely households participate in the stock market. Second, households do adjust their portfolio holdings of risky assets when between-squares correlations present, a finding that is consistent with the model's prediction. A one-standard-deviation from zero in between-squares correlations is associated with a 2.43% to 3.16% decrease in equity shares. Moreover, in line with the model's predictions, households without college degrees tend to have portfolio decisions that are more correlated with their between-squares correlations.

Finally, we investigate the consequence for both households and society if they ignore nonlinear dependence. Measuring the equivalent wealth loss, we find that the cost of ignoring between-squares correlation could reach 3.32% of their wealth when the households are young, lack college degrees and accumulate low wealth. On the other hand, via a simulation study, we show that ignoring between-squares correlation could increase the Gini index by 2.73% for the overall society and by 3.63% for college graduates. Apparently, our finding shows that nonlinear dependence, as a new channel, plays an important role in the evolution of wealth inequality.

This paper builds on a large body of literature that studies the role of nondiversifiable labor income risk on life-cycle portfolio decisions. Research in this literature usually focuses on analyzing labor income shocks that follow a lognormal distribution (e.g., Carroll (1997); Cocco et al. (2005); Gomes and Michaelides (2005a); Fagereng et al. (2017)). Bagliano et al. (2019) consider personal disaster risk during an individual's working age period (20-65 years old), while Chang et al. (2018) have assessed age-dependent labor market uncertainty. Inspired by Guvenen et al. (2014), Catherine (2022) and Shen (2022) examine the importance of countercyclical earnings risk. However, what has not yet been well investigated is the nonlinear dependence between labor and stock market risk and how it affects households' portfolio decisions. Our paper aims to fill this gap.

Further, our examination of nonlinear dependence contributes to another strand of the literature analyzing the interaction between labor income risk and financial portfolio choice. Storesletten et al.

(2007), Benzoni et al. (2007), and Lynch and Tan (2011) show that labor income tends to move together with stock returns in the long run. Including this linear correlation helps to match the data well. However, the empirical evidence on the linear correlation is somewhat mixed (e.g., Campbell et al. (2001)). For example, Huggett and Kaplan (2016) find that human capital and stock returns have a smaller correlation than the one in Benzoni et al. (2007). Our model does not rely on the linear dependence between labor and stock market risk. Instead, we investigate the nonlinear dependence between labor and stock market, and its impact on life-cycle portfolio decisions.

According to our theory, workers with college degrees tend to have stronger between-squares correlation and take less risk with their financial investments. This could be because college graduates are more likely to receive performance-based compensation, making their income flow closely related to firms' market value. Our result is consistent with the results of Campbell et al. (2001), Angerer and Lam (2009), and Betermier et al. (2012), all of whom show that workers in an industry with higher income risk exhibit less risky asset shares.

This paper also contributes to another branch of the literature examining nonlinear dependence across or within financial assets. Harvey and Siddique (2000) show that coskewness between an asset and the market portfolio exists and can explain parts of the apparent nonsystematic components in cross-sectional variation in expected returns. Patton (2009) examines the nonlinear dependence between hedge fund performance and market factor via testing the nonlinearity between the conditional mean and variance with market factor. Mencía (2011) further develops a multivariate framework to investigate the nonlinear dependence in the conditional variance between the hedge fund and the market portfolio. Regarding nonlinear dependence within an asset, Ding et al. (1993) and Granger and Ding (1994) first propose using the powers of absolute returns to examine the "ARCH effect" and volatility clustering. Cont (2001) revisits this nonlinear dependence by showing the persistence of autocorrelation in squared returns for the S&P 500. Other papers that investigate nonlinear dependence include LeBaron (1988), Scheinkman and LeBaron (1989), Hsieh (1989, 1993), Edwards and Susmel (2001), Shephard (2010), and Madan and Wang (2020). Our analysis contributes to this literature by examining nonlinear dependence across different markets. Motivated by the measures proposed in Ding et al. (1993) and Granger and Ding (1994), we construct the between-squares correlation to capture nonlinear dependence and document evidence that the nonlinear dependence across labor market and stock market exists for a large sample of households.

The rest of the paper is organized as follows. Section 2 introduces the concept of between-squares correlation and documents stylized facts related to it. A first model that captures the essential ingredients for our analysis and thus highlights the main mechanisms is also provided in Section 2. Section 3 introduces the model's setup. Section 4 shows the calibration, and Section

5 presents the quantitative analysis. Section 6 conducts empirical analysis. Section 7 reveals economics implications of the between-squares correlation, and Section 8 concludes.

2 Between-Squares Correlation

In this paper, we introduce the between-squares correlation to capture the nonlinear dependence between stock market returns and labor income shocks. The between-squares correlation is defined as the Pearson correlation between the de-meaned squares of two random variables:

$$\text{Corr}^{\text{sq}}(Z_1, Z_2) := \text{Corr}\left((Z_1 - \mathbb{E}Z_1)^2, (Z_2 - \mathbb{E}Z_2)^2\right), \tag{1}$$

where $\text{Corr}(\cdot, \cdot)$ denotes the Pearson correlation coefficient function.

2.1 Properties of Between-squares Correlation

The between-squares correlation can be viewed as the normalized cokurtosis,[8] which takes the value within the range of $[-1, 1]$, and captures the common sensitivity to extreme states for the two variables. For example, two random variables with a high level of the between-squares correlation will tend to undergo both extremely positive and extremely negative deviations.

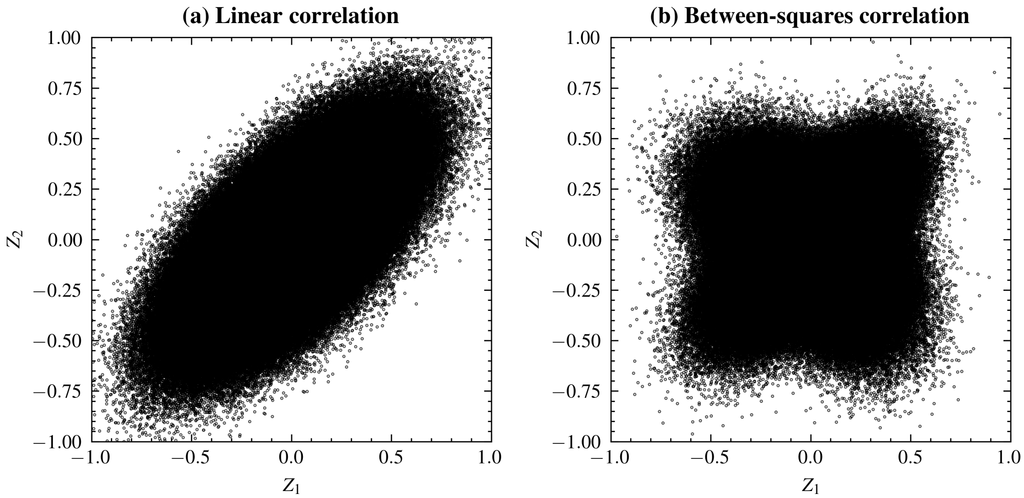
Under a normality assumption, the between-squares correlation is completely determined by the linear correlation, that is, $\text{Corr}^{\text{sq}}(Z_1, Z_2) = \text{Corr}^2(Z_1, Z_2)$ when (Z_1, Z_2) is binormal. However, when data are subject to risks from higher-order moments (i.e., Shen (2022)), the between-squares correlation emerges and cannot be determined by the linear correlation. For instance, in this paper, we assume both labor income shocks and stock return innovations follow a mixture normal distribution, and thus both labor income and stock return are subject to higher-order moments and the between-squares correlation between them exists.

To illustrate the differences between the between-squares correlation and the linear correlation, Figure 1 presents scatter plots for two random variables under both normal and mixture normal assumptions. In panel A, when both variables satisfy normal distribution, a positive linear correlation implies that they only tend to move in the same direction. In contrast, as shown in panel B, the two random variables are subject to mixture normal distribution such that higher-order moments and nonlinear dependence exist. The scatter plot in this panel shows that, when the between-squares correlation is positive, two types of comovements are possible: two variables may

[8] The cokurtosis of two random variables Z_1 and Z_2 is defined by $K(Z_1, Z_2) = \frac{\mathbb{E}[(Z_1 - \mathbb{E}[Z_1])^2 (Z_2 - \mathbb{E}[Z_2])^2]}{\sigma_{Z_1}^2 \sigma_{Z_2}^2}$, where σ_{Z_1} and σ_{Z_2} are the standard deviations of Z_1 and Z_2, respectively. The between-squares correlation can be related to the cokurtosis through the following equation: $K(Z_1, Z_2) = \sqrt{(\text{Kurt}[Z_1] - 1)(\text{Kurt}[Z_2] - 1)}\text{Corr}^{\text{sq}}(Z_1, Z_2)$.

Figure 1: Scatter Plots of Different Correlations

This graph compares the scatter plots of the linear correlation and the between-squares correlation, with the same mean and variance. We fix $\mathbb{E}[Z_1] = \mathbb{E}[Z_2] = 0$, $\mathbb{E}[Z_1^2] = \mathbb{E}[Z_2^2] = 0.04$, $\mathbb{E}[Z_1^3] = \mathbb{E}[Z_2^3] = 0$, and $\mathbb{E}[Z_1^4]/\mathrm{Var}^2[Z_1] = \mathbb{E}[Z_2^4]/\mathrm{Var}^2[Z_2] = 3$. In panel A, (Z_1, Z_2) follows the binomial normal distribution with a linear correlation of 0.3. In panel B, (Z_1, Z_2) follows a binomial mixture normal distribution with a linear correlation of 0.1, and the between-squares correlation of 0.3.



move in the same *and* opposite directions. Apparently, our nonlinear dependence measure is able to simultaneously capture both comovements between the labor market and stock market as shown in empirical studies.[9]

On the other hand, compared with cointegration, introducing between-squares correlation requires fewer assumptions on the evolution of stock price and labor income. Cointegration requires a linear combination of the two series to be stationary, which is a stringent constraint. In contrast, the between-squares correlation measures the correlation of higher moments and does not impose any additional theoretical constraints. Therefore, the between-squares correlation is a more flexible and convenient measure for capturing the nonlinear dependence between two variables without requiring a specific functional form.

More importantly, we find that the between-squares correlation can serve as a sufficient statistic for nonlinear dependence under some proper assumptions. Specifically, in Appendix B, we show that when two variables follow a mixture normal distribution and satisfy certain regularity conditions, the between-squares correlation, together with linear correlation and other marginal moments, is sufficient to characterize the joint distribution of these two variables. This implies that, once the linear correlation and marginal moments are known, any additional nonlinear dependence can

[9] Michelacci and Quadrini (2009), Bronars and Famulari (2001), and Nickell and Wadhwani (1991), among others, document two types of comovements between shareholder value and wages. For instance, Bronars and Famulari (2001) show that firms in a growing environment pay lower wages, implying the opposite movement between firm value and labor income. Nickell and Wadhwani (1991) show that firms facing severer financial constraints tend to pay lower wages, implying firm value and labor income move in the same direction during a financial crisis.

be entirely captured by the between-squares correlation. Hence, the between-squares correlation is a valuable tool for characterizing the nonlinear dependence between two variables, even when knowledge of the joint distribution underlying them is limited.

2.2 Stylized Facts of Between-squares Correlation

We employ a rigorous empirical framework to investigate the between-squares correlation between labor income shocks and innovations to stock returns by using the PSID survey data and CRSP data.

The PSID is a comprehensive survey that encompasses a diverse range of information about the labor market and demographic variables, including age, education, household composition, and marital status. The survey was conducted annually from 1968 to 1997 and biennially after 1997. However, income-related variables were bracketed in the 1968 and 1969 waves, making them unusable for our analysis. To optimize the use of data after 1997, we consider two time series: one with annual data ($\Delta t = 1$) from 1970 to 1997 and another with biennial data ($\Delta t = 2$) from 1970 to 2017. To construct the biennial series, we can directly use the PSID data after 1997, while for the data before 1997, we only retained observations from every other year, despite the data being collected annually.

The age profile may differ among households with a female head of household, necessitating a separate estimation. Thus, the sample was divided based on the gender of the head of the household. However, the subsample with a female head of household had insufficient observations, and we were unable to perform a reliable estimation. Consequently, the analysis was restricted to households with a male head of household.

We defined labor income as total reported labor income plus unemployment compensation, workers' compensation, social security, supplemental social security, other welfare, child support, and total transfers, all this for both the head of household and if present his spouse. Then we deflate labor income using the 2016 Consumer Price Index from the Bureau of Labor Statistics.

We closely follow the sample selection strategy of Nakajima and Smirnyagin (2019).[10] We only include households where the head has a minimum of 20 waves of positive labor income data. If a household head retires in the current survey year, we delete all the information about this household.

Employing a methodology similar to that of Cocco et al. (2005), we estimate the logarithm of

[10] See Appendix A for details.

household i's labor income at age t during the working period (y_{it}) as follows:

$$y_{it} = f(t, Z_{it}) + \delta_{i,t}. \tag{2}$$

where $f(t, Z_{i,t})$ is a deterministic function of age t and a vector of other individual characteristics $Z_{i,t}$, and $\delta_{i,t}$ is the labor income shock. The function $f(t, Z_{it})$ is additively separable in t and Z_{it}. The vector Z_{it} of personal characteristics other than age and the fixed household effect includes marital status and household composition. Household composition is defined as the additional number of family members in the household besides the head and, if applicable, spouse.

To estimate the error structure of the labor income process, we first regress the logarithm of labor income on dummies for age, family, marital status, and household composition. Next, we define $\Delta y^*_{it,\Delta t}$ as

$$\Delta y^*_{it,\Delta t} = y^*_{i,t+\Delta t} - y^*_{it}, \tag{3}$$

where Δt indicates the sampling frequency , and y^*_{it} is given by

$$y^*_{it} = y_{it} - \hat{f}(t, Z_{it}). \tag{4}$$

By doing so, $\Delta y^*_{it,\Delta t}$ contains only shocks.

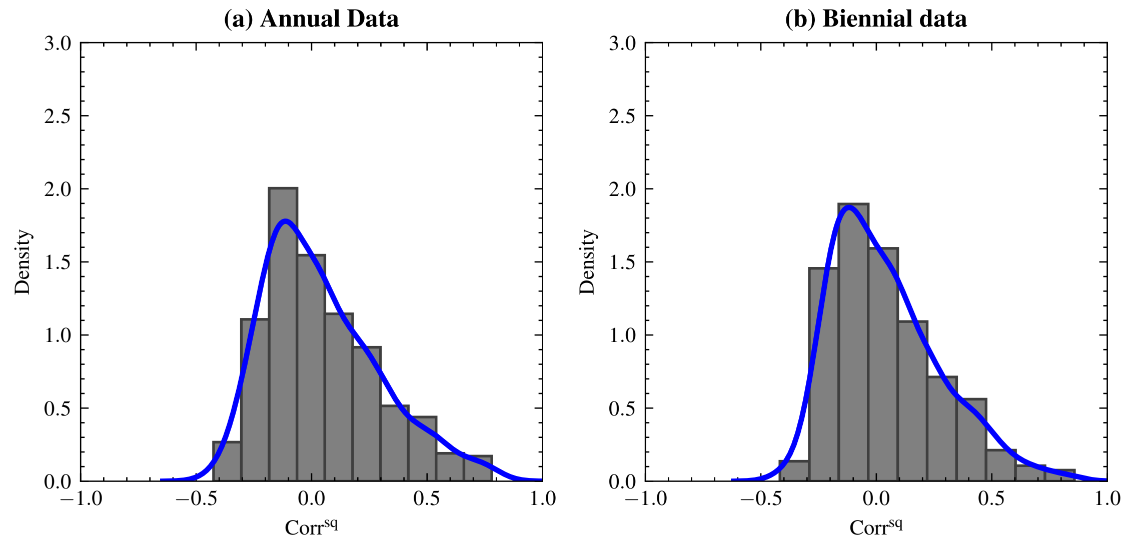
To quantitatively assess the excess return on our stylized risky asset, we use CRSP data on the New York Stock Exchange value-weighted stock return relative to the Treasury bill rate. The linear correlation and between-squares correlation are then easily computed.

We categorize all households into two groups, college graduates and a no-college group, to account for the impact of education. This is based on the well-established observation that households exhibit distinct behavioral patterns across different education levels (see Attanasio (1995); Hubbard et al. (1995); Cocco et al. (2005)). Subsequently, we conduct separate estimations of the linear correlation and between-squares correlation for each group to understand the effect of education on our results.

Figure 2 shows the histogram and kernel density of estimations across the full sample, which highlights a high density of positive between-squares correlation. The long right tail indicates that many households earn income that is highly exposed to the stock market concerning the between-squares correlation. Though there are few negative between-squares correlations, these negative estimates cluster around zero, suggesting that a positive between-squares correlation is a reasonable assumption for the model setup. Furthermore, this assumption is supported by the

Figure 2: between-squares correlation Distribution

This graph displays the histogram and kernel density of between-squares correlation estimates between labor income shock and excess stock return for the full sample. We use labor income and stock return data from the PSID and CRSP, respectively. Panel (a) uses the annual data from 1970 to 1997, while panel (b) uses the biennial data covering the period from 1970 to 2017.



findings presented in Table 1.

Table 1 provides summary statistics for the between-squares correlation, with results from different data frequencies supporting a significant and positive between-squares correlation. Panel A reveals that the full sample exhibits positive average values of 0.044 and 0.041 for annual and biennial data, respectively. These values are significantly different from zero, with p-values of less than $< 0.2\%$. While medians are close to zero, substantial variability exists in the values of the between-squares correlation. Panel A also provides the interquartile range of the between-squares correlation for both data frequencies over the full sample. We find that the 75^{th} and 90^{th} percentiles are substantially positive at 0.2 and 0.4, respectively. Moreover, more than half of the estimations are positive, with about one-third featuring a substantially positive between-squares correlation (≥ 0.1), and over one-fifth exceeding 0.2. Therefore, we assume a positive between-squares correlation and use the sample mean as our calibration target for model setup.

Panels B and C in Table 1 present statistics for two education groups, with results similar to those of the full sample. However, most statistics for the no-college group are lower than those for college graduates, with the average between-squares correlation for college graduates approximately twice that of the no-college group. Based on these findings, we perform separate calibrations for each group and investigate the disparities in corresponding policy rules in Section 4.

Table 1: Summary Statistics

This table presents summary statistics from various samples using PSID and CRSP data. Panel A displays results for the entire sample, while panels B and C show results for households without and with college degrees, respectively. The column labeled "Annual data" (or "Biennial data") indicates whether the statistics are derived from annual (or biennial) data from the period of 1970 to 1997 (or 2017). "Fraction > k" shows the percentage of samples in which the between-squares correlation exceeds k.

	Panel A: Full sample		*Panel B: No college*		*Panel C: College*	
	Annual data	Biennial data	Annual data	Biennial data	Annual data	Biennial data
Mean	0.044	0.041	0.034	0.032	0.070	0.060
SD	0.247	0.233	0.241	0.224	0.259	0.250
p-value	$< 0.2\%$	$< 0.1\%$	0.014	$< 1\%$	$< 0.3\%$	$< 0.2\%$
10th	−0.232	−0.208	−0.236	−0.214	−0.211	−0.186
25th	−0.142	−0.140	−0.152	−0.141	−0.115	−0.135
Median	0.006	0.001	0.007	0.000	0.005	0.004
75th	0.200	0.179	0.177	0.150	0.244	0.206
90th	0.400	0.378	0.357	0.366	0.433	0.406
Fraction > 0	51.26%	50.29%	51.63%	50.00%	50.39%	50.89%
Fraction > 0.1	34.71%	34.24%	34.64%	33.33%	34.88%	36.09%
Fraction > 0.2	24.83%	23.60%	22.55%	22.41%	30.23%	26.04%
Observations	435	517	306	348	129	169

2.3 A First Model: Skewness Channel and Kurtosis Channel

To understand the impact of between-squares correlation on portfolio allocation, we start with a parsimonious one-period portfolio optimization problem. Through this model, we show that the between-squares correlation has a nonlinear effect on portfolio allocation, and it affects households' portfolio allocation via two pathways: skewness channel and kurtosis channel.

We consider a one-period economy with two assets, a stock with a random return R^S and a bond with a constant gross return R_f. For simplicity, we assume the household is endowed with a fixed amount of initial wealth (W_0) and consumes a constant value during this period. Then she decides how many shares of wealth (α) are invested in stock and saves the rest. Meanwhile, she faces labor income risk, captured by W_0R^L, where R^L is the income shock. So her next-period-wealth (W_1) is given by

$$W_1 = W_0R^p = W_0(1 + \alpha R^S + (1 - \alpha)R_f + R^L) = W_0(1 + R_f + \alpha R^E + R^L), \tag{5}$$

where $R^E = R^S - R_f$ is excess return and $R^p = 1 + R_f + \alpha R^E + R^L$ is portfolio return.[11] In addition, we consider mixture normal distributions for both excess return, R^E, and labor income risk, R^L, and account for the existence of linear correlation and between-squares correlation between them.

Let $u(\cdot)$ be any utility function, and her one-period portfolio optimization problem is

$$\max_{\alpha} \mathbb{E}[u(W_1)] \quad \text{s.t.} \quad W_1 = W_0 R^p \,. \tag{6}$$

Given W_0 is known, her end-of-period utility entirely depends on the portfolio return R^p.

In accordance with Appendix C, the optimization problem (Equation 6) for this household can be approximated by the following optimization problem that takes into account the mean, variance, skewness, and kurtosis of portfolio returns (MVSK):

$$\max_{\alpha} \ \mathbb{E}(R^p) - \lambda_1 \mathrm{Var}(R^p) - \lambda_2 \mathrm{Skew}(R^p) - \lambda_3 \mathrm{Kurt}(R^p), \tag{7}$$

where $\lambda_1, \lambda_2, \lambda_3$ are risk weights contingent upon the household's utility.

The main focus of our investigation is to examine the influence of between-squares correlation on households' portfolio allocation, as determined by Equation (7). As demonstrated by the following theorem, given risky asset shares, between-squares correlation solely impacts the skewness and kurtosis of portfolio returns, while having no effect on variance.[12]

Theorem 1. *$(R^E, R^L)^T$ is a random vector. A portfolio return $\Pi(\alpha, R^E, R^L)$ is defined as the linear combination $\alpha R^E + R^L$. Then the standard deviation, skewness, and kurtosis of this portfolio can be represented as:*

$$\begin{aligned}
\sigma_\Pi =& \sqrt{\alpha^2 \sigma_E^2 + 2\alpha \sigma_E \sigma_L \mathrm{Corr}\,(R^E, R^L) + \sigma_L^2}, \\
\mathrm{Skew}_\Pi =& \frac{1}{\sigma_\Pi^3} \left[\alpha^3 \sigma_E^3 S_E + \sigma_L^3 S_L + 3\alpha^2 \sigma_E^2 \sigma_L S(R^E, R^E, R^L) + 3\alpha \sigma_E \sigma_L^2 S(R^E, R^L, R^L)\right], \\
\mathrm{Kurt}_\Pi =& \frac{1}{\sigma_\Pi^4} \left[\alpha^4 \sigma_E^4 K_E + \sigma_L^4 K_L + 4\alpha^3 K(R^E, R^E, R^E, R^L) + 6\alpha^2 K(R^E, R^E, R^L, R^L)\right. \\
& \left. + 4\alpha K(R^E, R^L, R^L, R^L)\right],
\end{aligned}$$

where $\sigma_a, S_a, K_a, a \in \{E, L\}$ are the standard deviation, skewness and kurtosis of R^E and R^L.

Furthermore, it is worth noting that variance is solely dependent on linear correlation and

[11] This portfolio return has a constant unit of labor income return. It captures that labor income is nontradable and households can only adjust their risky asset holding to achieve a target portfolio return.

[12] The proof is relegated to Appendix C.2.

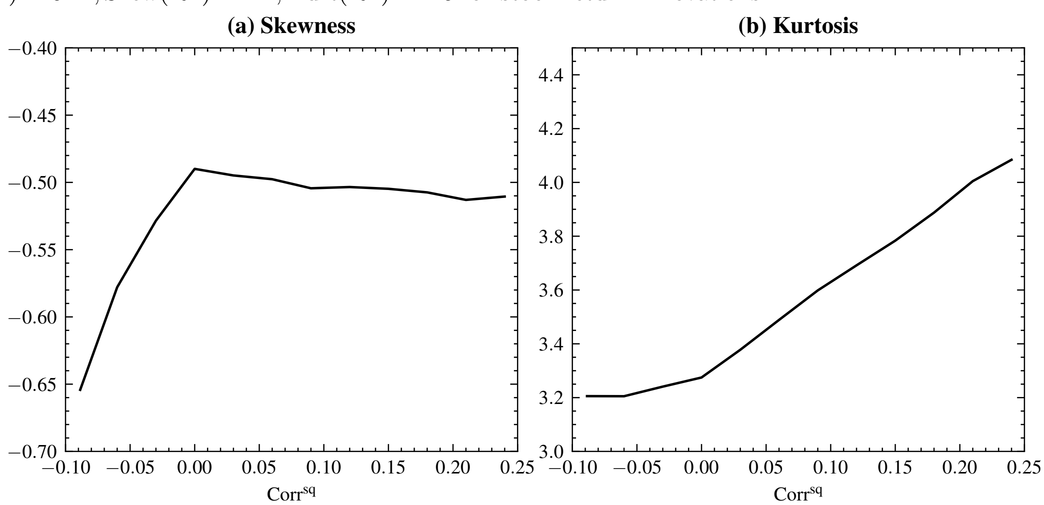
marginal distributions, and can not capture the risks associated with nonlinear dependence. Consequently, our results suggest that an alternative approach of using utility functions with higher-order moments, rather than the conventional Markowitz mean-variance portfolio optimization, is more effective in capturing and managing portfolio risks.

Figure 3 shows the skewness and kurtosis of portfolio returns as the between-squares correlation varies, while holding the linear correlation and other moments constant at $\alpha = 1$. We find that changes in the between-squares correlation have dissimilar impacts on portfolio skewness and kurtosis. More precisely, we note a substantial reduction in skewness from -0.50 to -0.65, while kurtosis remains unchanged, as the between-squares correlation falls from 0 to -0.10. In contrast, as the between-squares correlation rises from 0 to 0.25, we observe a marked increase in kurtosis from 3.25 to 4.10, while skewness remains relatively constant with a slight decrease.

Overall, these findings indicate that the impact of between-squares correlation on portfolio risk is primarily driven by the skewness channel when the between-squares correlation is negative, while the kurtosis channel becomes more prominent when it is positive. Therefore, it is expected that households with a nontradable labor income would strategically decrease their risky share to minimize portfolio risk and optimize the risk-adjusted objective when the between-squares correlation deviates from zero.

Figure 3: Changes in Skewness and Kurtosis

This figure shows the changes in higher-order moments of portfolio returns as the between-squares correlation varies while keeping all other parameters constant. Panel a shows the changes in skewness, and panel b shows the changes in kurtosis. The risky share of wealth (α) is set to 1 and linear correlation is set to 0.1. Other moments are fixed as follows: $\mathbb{E}[R^L] = 0, \mathrm{Var}(R^L) = 0.25^2, \mathrm{Skew}(R^L) = -0.5, \mathrm{Kurt}(R^L) = 3.0$ for labor income risk, and $\mathbb{E}[R^S] = 0, \mathrm{Var}(R^S) = 0.2^2, \mathrm{Skew}(R^S) = -1, \mathrm{Kurt}(R^S) = 4.5$ for stock return innovations.



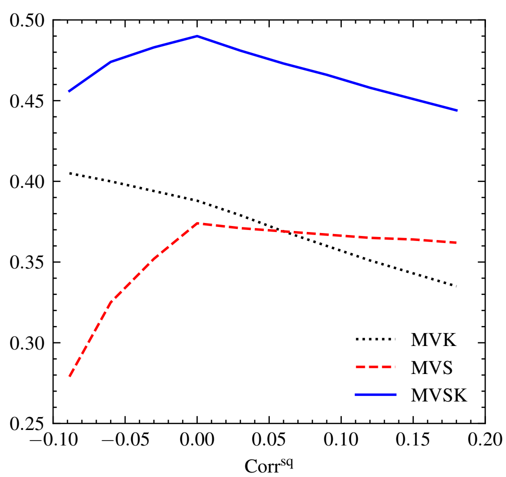
As anticipated, the results depicted in Figure 4 are consistent with our findings, displaying the optimal risky share as the between-squares correlation varies. To gain a deeper understanding of the asymmetric impact of between-squares correlation on skewness and kurtosis, we analyze another two model specifications. The first specification accounts for mean, variance, and skewness (MVS), while the second specification considers mean, variance, and kurtosis (MVK).

As the between-squares correlation increases from -0.10 to 0, the optimal risky share rises, but it decreases when the between-squares correlation increases from 0 to 0.20. These findings can be explained through the lens of the reduced models, MVK and MVS. When the between-squares correlation is negative, the MVS undergoes a more significant change compared to the MVK, indicating that negative between-squares correlation has a more substantial impact on portfolio skewness than kurtosis. Consequently, the MVSK displays a similar pattern to MVS when the between-squares correlation is negative, suggesting a skewness channel.

Conversely, when the between-squares correlation becomes positive, the impact on portfolio skewness is minor, as MVS remains relatively flat. However, the effect of between-squares correlation on portfolio kurtosis is significant, as depicted by the MVK under positive between-squares correlation. Therefore, a similar pattern is observed between MVSK and MVK, as the impact of positive between-squares correlation on kurtosis dominates the optimal risky share, indicating a kurtosis channel.

Figure 4: Optimal Risky Share

This graph compares the optimal risky share for varying between-squares correlation while holding all other parameters constant. The analysis is performed under three distinct model specifications: MVSK, MVS, and MVK. The MVSK model considers the mean, variance, skewness, and kurtosis of portfolio returns, while the MVS model only considers the mean, variance, and skewness of portfolio returns, and the MVK model only takes into account the mean, variance, and kurtosis of portfolio returns. The parameters used in the simulation are $\mathbb{E}[R^S] = 0.04$, $\text{Std}[R^S] = 0.2$, $\text{Skew}[R^S] = -0.5$, $\text{Kurt}[R^S] = 4.5$, $\mathbb{E}[R^L] = 0.01$, $\text{Std}[R^L] = 0.25$, $\text{Skew}[R^L] = -1$, $\text{Kurt}[R^L] = 3$, $\text{Corr}[R^S, R^L] = 0.1$, $W_0 = 1$. The corresponding values of $\lambda_1, \lambda_2, \lambda_3$ are 0.9473, -0.0054, 0.0011.



3 Model

3.1 Preferences

We follow the convention in life-cycle models and let adult age (t) correspond to effective age minus 19. Each period corresponds to one year and households live for a maximum of 81 periods (age 100). Households start working at the age of 20 and receive uncertain labor income exogenously until the age of 65. They retire at the age of 65. The probability that an adult is alive at age t conditional on being alive at age $t-1$ is denoted as p_t.[13] At each point in time there is a stationary age distribution of households in the economy with no population growth.

Households have Epstein-Zin (1989) preferences, a recursive preference in which the elasticity of intertemporal substitution is separated from the relative risk aversion. For household i, let $X_{i,t}$ denote the cash-on-hand at the beginning of age t, and the utility function is given by

$$U_{i,t} = \left\{ (1-\beta) C_{i,t}^{1-1/\psi} + \beta \left(E_t \left[p_{t+1} U_{i,t+1}^{1-\gamma} + b(1-p_{t+1}) X_{i,t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}}, \tag{8}$$

where β is the discount factor, b determines the strength of bequest motive, γ is the coefficient of relative risk aversion, and ψ is the elasticity of intertemporal substitution.

3.2 Financial Assets Return

We assume financial market assets consist of two financial assets in which households can invest, one riskless and one risky asset. The riskless asset has a constant gross return R_f, and the return of the risky asset is given by

$$R_{t+1}^S = R_f + \mu + \eta_{t+1}, \tag{9}$$

where μ is the excess return of the risky asset and η_{t+1} is the shock to returns, independently and identically distributed as

$$\eta_t \sim \begin{cases} N(\mu_{\eta,1}, \sigma_{\eta,1}) & \text{with prob. } p_\eta, \\ N(\mu_{\eta,2}, \sigma_{\eta,2}) & \text{with prob. } 1-p_\eta, \end{cases} \tag{10}$$

Our assumption of a mixture normal distribution is mainly motivated by two reasons. First, as argued before, the mixture normal assumption allows us to deal with a more complicated dependence structure, such as the between-squares correlation between stock return and labor income

[13] p_t is taken from the mortality tables of the National Center for Health Statistics.

introduced in our model. Under normality, the between-squares correlation is always the square of the correlation; however, in Section 2, we show that from the data, the between-squares correlation is actually similar to the correlation in magnitude. Therefore, deviation from normality allows us to construct a more realistic between-squares correlation. Second, the mixture normal assumption enables us to incorporate higher moments in the stock return innovations. Higher moments in the high-frequency returns data have been discussed and considered important for many years, but the long-horizon higher moments are difficult to measure accurately using standard techniques. Neuberger and Payne (2021) show that short-horizon returns can be used to estimate the higher moments of long-horizon returns, and their empirical results identify a high level of negative skewness and excess kurtosis of annual U.S. equity market returns. Therefore, we introduce mixture normal shocks instead of normal shocks to capture such moments.

3.3 Labor Income Process

Recall that the labor income process during the working period is defined as (2) in section 2.2. We further decompose the labor income shock $\delta_{i,t}$ into a persistent shock, $\nu_{i,t}$, and a transient shock, $\epsilon_{i,t}$:

$$\delta_{i,t} = \nu_{i,t} + \epsilon_{i,t}. \tag{11}$$

We assume that $\epsilon_{i,t}$ is normally distributed as $N(0, \sigma_\epsilon^2)$ and independent of $\nu_{i,t}$ and the stock return shock, $\eta_{i,t}$. The persistent shock ν_{it} is given by

$$\nu_{i,t} = \lambda \nu_{i,t-1} + u_{i,t}, \tag{12}$$

where $u_{i,t}$ follows a mixture normal distribution:

$$u_{it} = \begin{cases} N(\mu_{u,1}, \sigma_{u,1}^2) & \text{with prob. } p_u, \\ N(\mu_{u,2}, \sigma_{u,2}^2) & \text{with prob. } 1 - p_u. \end{cases} \tag{13}$$

We assume $u_{i,t}$ is correlated with the stock return shock $\eta_{i,t}$, and the dependence structure is captured by both the Pearson correlation and between-squares correlation. We define the correlations between components of $u_{i,t}$ and $\eta_{i,t}$ as

$$\rho_{a,b} = \text{Corr}(u_{i,t}^{(a)}, \eta_t^{(b)}), a = 1, 2, b = 1, 2. \tag{14}$$

The four correlations $\rho_{a,b}$ control for the value of the correlation and between-squares correlation between $u_{i,t}$ and $\eta_{i,t}$.[14]

Income during retirement is assumed to be exogenous and deterministic, with all households retiring in time period K, corresponding to the retirement age of 65. Income during retirement follows

$$y_{it} = log(\lambda) + f(K, Z_{iK}) + v_{iK}, \quad t > K, \tag{15}$$

where it is a constant fraction (λ) of the permanent component of labor income in the last working period.

3.4 Wealth Accumulation

At each period t, households start with accumulated wealth, receive labor income, and decide whether they want to participate in the stock market. They consume C_t and invest the rest on a portfolio consisting of α_t shares of risky assets and $1 - \alpha_t$ of riskless assets if they choose to participate in the stock market. Otherwise, they will invest the rest on riskless assets. The participation decision is represented by a dummy variable I_P, which is one if households decide to participate and zero otherwise. To participate in the stock market, households need to pay a fixed cost, F, which represents the cost of acquiring information about the stock market and transaction fees. We write cash-on-hand $X_{i,t}$ as

$$X_{i,t+1} = (X_{i,t} - C_{i,t})R^p_{i,t+1} - FI_pP_{i,t} + Y_{i,t+1}, \tag{16}$$

where $R^p_{i,t+1}$ is the portfolio return given by

$$R^p_{i,t+1} = \alpha_{i,t}R^S_{t+1} + (1 - \alpha_{i,t})R_f. \tag{17}$$

Lastly, we restrict borrowing from riskless assets or future labor income, and short sales on risky assets. Specifically, households face the following constraints:

$$0 \leq \alpha_{i,t} \leq 1, \tag{18}$$

$$0 < C_{i,t} \leq X_{i,t}. \tag{19}$$

[14] See Appendix B for the explicit forms.

3.5 Households' Optimization Problem

In each period t, households choose whether or not they participate in the market, and decide their consumption and risky shares $(I^p_{i,t}, C_{i,t}, \alpha_{i,t})$ based on cash-on-hand $X_{i,t}$ to maximize the expected utility. The optimization problem can be stated as

$$V_{i,t} = \max_{\{I^p_{i,u}\}^T_{u=t}, \{\alpha_{i,u}\}^T_{u=t}, \{C_{i,u}\}^T_{u=t}} \mathbb{E}_t(U_{i,t}), \tag{20}$$

where $U_{i,t}$ is defined in equation (8) and is subject to the constraints given by equations (9) to (19).

Since analytical solutions do not exist for this problem, we use a numerical solution method. We standardize the entire problem with persistent labor income $P_{i,t}$ defined as $log(P_{it}) = f(t, Z_{i,t}) + \nu_{it}$, and introduce a new state, adjusted persistent income shock $w_{i,t}$, defined as $\log(w_{i,t}) = (1-\lambda)\nu_{i,t}$.[15] Specifically, let $x_{i,t} = \frac{X_{i,t}}{P_{i,t}}$ and $c_{i,t} = \frac{C_{i,t}}{P_{i,t}}$ be the normalized cash on hand and consumption. The normalized value function $v_{i,t} = \frac{V_{i,t}}{P_{i,t}}$ is

$$v_{i,t}(x_{i,t}, w_{i,t}) = \max_{\substack{\{\alpha_{i,u}\}^T_{u=t} \\ \{c_{i,u}\}^T_{u=t}}} \left\{ \begin{array}{l} (1-\beta) c_{i,t}^{1-\frac{1}{\psi}} \\ +\beta \left(\mathbb{E}_t \left(\frac{P_{i,t+1}}{P_{i,t}} \right)^{1-\gamma} \left[p_{t+1} v_{i,t+1}^{1-\gamma} + (1-p_{t+1}) b x_{i,t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \end{array} \right\}^{\frac{1}{1-1/\psi}}, \tag{21}$$

subject to

$$x_{i,t+1} = \begin{cases} (x_{i,t} - c_{i,t}) r^p_{i,t+1} \frac{P_{i,t}}{P_{i,t+1}} - F I^p i,t + e^{\varepsilon_{i,t+1}} & \text{for } t \le K \ , \\ (x_{i,t} - c_{i,t}) r^p_{i,t+1} \frac{P_{i,t}}{P_{i,t+1}} - F I^p i,t + k & \text{for } t > K \ . \end{cases} \tag{22}$$

$$\ln w_{i,t+1} = \begin{cases} \lambda \ln w_{i,t} + (1-\lambda) u_{i,t+1} & \text{for } t \le K \ , \\ 0. & \text{for } t > K \ . \end{cases} \tag{23}$$

$$\frac{P_{i,t}}{P_{i,t+1}} = \begin{cases} w_{i,t} e^{f(t, Z_{i,t}) - f(t+1, Z_{i,t+1}) - u_{i,t+1}} & \text{for } t \le K \ , \\ 1 & \text{for } t > K \ . \end{cases} \tag{24}$$

We use a numerical method to recursively solve the optimization problem above. In the last period, households predict a certain death, and therefore the policy functions are determined by the bequest motive. We start from this terminal condition and then iterate backward. Appendix D presents the details of the numerical solution method.

[15] We use $w_{i,t}$ as a state instead of $\nu_{i,t}$ for the convenience of calculation.

4 Calibration

In this section, we discuss the calibration of the model. We categorize the parameters into two groups. The first group includes the parameters shaping the dynamics of labor income and stock return. We use the PSID and CRSP, as well as generalized method of moments (GMM), to match the empirical moments. The second group includes the parameters of preference determining the decision rules of households' optimization problem. Using the simulated method of moments (SMM), we calibrate these parameters to match the average life-cycle profiles from SCF data.

4.1 Labor Income and Stock Return

We estimate the parameters of labor income and stock return jointly with the GMM. To estimate labor income shocks, we need to get the family-specific fixed effects $f(t, Z_{i,t})$ of labor income in equation (2). Following the method of Cocco et al. (2005), we use the PSID from 1970 to 2017 and run a regression analysis to estimate $f(t, Z_{i,t})$.[16] The residuals from the regression can be considered to be realizations of the labor income shock, $\delta_{i,t}$. Then we take the annual excess returns from CRSP between 1970 and 2017, for which the de-meaned series can be considered to be realizations of stock return shocks, η_t. Using these data, we can calibrate parameters of labor income and stock returns.

We apply GMM targeting (1) mean, variance, skewness, and kurtosis of stock return shocks η_t; (2) mean and variance, skewness, and kurtosis of $\delta_{i,t}$; and (3) correlation and between-squares correlation between $\delta_{i,t} - \delta_{i,t-1}$ and η_t.[17] For stock returns, Neuberger and Payne (2021) propose a new method using short-horizon stock returns to estimate long-horizon moments, and we use their results for the annual logarithm return with the standard deviation of 0.216, skewness of -1.41, and excess kurtosis of 5.62. For the labor income process, following Nakajima and Smirnyagin (2019), we target the moments of different age groups. We consider four age groups indexed by $\{30, 40, 50, 60\}$, and each group contains ages ± 5 years.[18] Lastly, for the dependence structure between two markets, we use the average between-squares correlation in Table 1 and the average correlation estimated through the same methodology.

In total, we calibrate five parameters $\{p_\eta, \mu_\eta^1, \mu_\eta^2, \sigma_\eta^1, \sigma_\eta^2\}$ that control for stock return dynamics, seven parameters $\{p_u, \mu_u^1, \mu_u^2, \sigma_u^1, \sigma_u^2, \lambda, \sigma_\epsilon\}$ that control for labor income dynamics, and four parameters $\{\rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}\}$ that determine the dependence structure. For simplicity, we assume that

[16] See Table 10 for variable definitions and Appendix A for the data selection principle and other details.

[17] We use third- and fourth-order moments in GMM. For simplicity, we refer to such higher-order moments as skewness and kurtosis. See Appendix C for details.

[18] Group 30 contains ages 22-35; group 40 contains ages 35-45; group 50 encompasses ages 45-55; and group 60 aggregates the remaining ages 55-65.

$\rho_{11} = \rho_{22}$ and $\rho_{12} = \rho_{21}$. Consistent with Section 2, we control for education and split households into two groups: households without college degrees and college graduates. For each group, we estimate the between-squares correlation and calibrate the parameters. Overall, we apply GMM with 19 moments to 14 parameters.

Panels A and B in Table 2 report the calibration results. For those parameters unrelated to shocks, we take their values from Cocco et al. (2005). Specifically, the riskless rate $R_f - 1$ is set to 2% and the equity premium μ is 4%. The replacement rate λ is calibrated as the ratio of the average of labor income for retirees in a given education group to the average of labor income in the last working year prior to retirement, which is 0.903 for the no-college group and 0.945 for college graduates.

4.2 Preference and Bequest Motive

We use the SCF data from 2007 to 2019 and calibrate the preference parameters and fixed cost rate to match the average participation rate, risky asset share, and normalized wealth for different age groups.

To capture the different behaviors between two educational groups, we calibrate the parameters separately, except that relative risk aversion γ and the fixed cost rate F are assumed to be the same between groups. To approximate the life-cycle profiles for an average household, we assume that households with college degrees make up 30% of the world's population, which is observed from the PSID.

We calculate the average participation rate, risky asset share, and normalized wealth for 15 age groups in $[20, 65]$ from the SCF data and obtain 45 moments in total.[19] The SMM seeks the parameters that minimize

$$(\hat{m} - m)' W (\hat{m} - m), \tag{25}$$

where $\hat{m}$ refers to the simulated moments, m refers to the targets, and W is the inverse of the covariance matrix of the empirical moments, which is estimated by bootstrapping the true data.

4.3 Results

In this section, we summarize the results of a two-step calibration. Table 2 presents the calibration parameters. Panel A describes the parameters for the stock return shock. There are 17.1% chances

[19] We construct nonoverlapping age groups every 3 ages. For example, the first group includes samples of ages $[20, 22]$, and the second includes samples of ages $[23, 25]$. The exception is the last group, which includes samples of ages $[62, 65]$.

to draw the innovation from N(-0.184, 0.394) and 82.9% chances to draw the innovation from N(0.038, 0.126). Although recessions or rare disasters happen less frequently when they happen, households are more likely to suffer negative stock returns and the stock market becomes more volatile.[20] Meanwhile, shocks to labor income also display a structure similar to the stock return innovations. Panel B shows that for college graduates, the probability of the mixture normal distribution (p_u) is 0.279. This implies college graduates are less likely to expect a negative mean growth (-15.6%) than a positive mean growth (6%). Households without college degrees have fewer opportunities to experience a negative growth rate (-12.4%), but on the other side, their positive growth rate is also smaller (4.6%) compared with college graduates. In addition, households without college degrees face much more transient risk than college graduates, with a standard deviation of 20.4% compared with 13.8% for college graduates. This can be explained by the fact that households with higher education levels are more likely to have stable income flows.

The parameters of the dependence structure (college graduates, 0.769 and -0.222; households without college degrees, 0.832 and -0.166) in panel C imply the linear correlation and between-squares correlation are 0.0457, 0.0692 for the college group, and 0.0382, 0.0327 for households without college degrees. The calibration results of the correlation are consistent with those of Cocco et al. (2005), which show that a correlation is not significantly different from 0. The between-squares correlation is in the same order of magnitude as the linear correlation. However, it is big enough to generate an obvious difference in policy function, as we will show in Section 5.

Panel D in Table 2 shows the calibration results of preference parameters and the fixed cost rate, and Table 3 reports the fitness of the calibration. Our model matches the data well with a moderate risk aversion of 4.8 and low fixed cost of 0.005 (0.5% of the household's expected annual income). Compared to the previous literature ($\gamma = 11$ in Fagereng et al. (2017), $\gamma = 5.6$ in Shen (2022) and $\gamma = 7$ in Catherine (2022)), we provide a possibility to fit the empirical data, especially the portfolio choice, with a lower but reasonable risk aversion estimate. In addition, the strength of the bequest motive (b) is 2.5, is within the range of existing empirical evidence and calibrations. Compared with college graduates, households without college degrees are more impatient with a discount factor of 0.92. No college group accumulates less wealth and has a weaker incentive to pay the fixed cost, leading to a low participation rate. College graduates have an EIS of $\psi = 0.9$, and households without college degrees have a lower EIS of $\psi = 0.3$. These measures are consistent with those of Vissing-Jørgensen (2002), Brav et al. (2002), Malloy et al. (2009), and Gomes and Michaelides (2008).

The existence of the between-squares correlation increases the labor income risk, and thus labor

[20] Such a structure of stock return shock has been studied by several researchers. See Fagereng et al. (2017), Shen (2022), and Catherine (2022).

Table 2: Baseline Calibration Parameters

This table reports the calibrated parameters of the life-cycle model. We use data from the PSID and CRSP and construct a two-step calibration. Panel A reports the parameters for the mixture normal shock, which controls the stock return process. Panel B reports the parameters for persistent shock $\nu_{i,t}$ (mixture normal distributed) and transient shock $\epsilon_{i,t}$ (normal distributed) which controls for the labor income process for different groups. The two correlation parameters are correlations between components of the stock return shock, η_t, and the persistent labor income shock, $\nu_{i,t}$. Panel C reports the preference parameters and fixed costs for different groups. We assume the two groups have the same risk aversion and fixed cost.

Panel A: Stock return shock, η_t			
Mixture weight	p_η	0.171	
Mean of 1st component	$\mu_{\eta,1}$	-0.184	
Mean of 2nd component	$\mu_{\eta,2}$	0.038	
SD of 1st component	$\sigma_{\eta,1}$	0.394	
SD of 2nd component	$\sigma_{\eta,2}$	0.126	
Panel B: Labor income shocks $u_{i,t}$ and $\epsilon_{i,t}$			
		No college	College
Mixture weight of $u_{i,t}$	p_u	0.271	0.279
Mean of 1st component of $u_{i,t}$	$\mu_{u,1}$	-0.124	-0.156
Mean of 2nd component of $u_{i,t}$	$\mu_{u,2}$	0.046	0.060
SD of 1st component of $u_{i,t}$	$\sigma_{u,1}$	0.172	0.231
SD of 2nd component of $u_{i,t}$	$\sigma_{u,2}$	0.010	0.013
SD of transient shock $\epsilon_{i,t}$	σ_ϵ	0.204	0.138
1st auto-correlation coefficient	λ	0.918	0.959
Replacement rate	k	0.903	0.945
Panel C: Dependence parameters			
		No college	College
Dependence parameter 1	ρ_1	0.832	0.769
Dependence parameter 2	ρ_2	-0.166	-0.222
Panel D: Preferences and fixed cost			
		No college	College
Relative risk aversion	γ	4.8	4.8
EIS	ϕ	0.3	0.9
Discount factor	β	0.92	0.98
Bequest motive	b	2.5	2.5
Fixed cost rate	F	0.005	0.005

Table 3: Moments of Labor Income and Stock Return

This table reports relative errors of the moments targeted in the calibration of labor income and stock return. Panels A and B report the mean, variance, and higher moments of stock return shock and labor income shock. Panel C reports the error in the dependence structure, including the Corr and between-squares correlation. Moments from the model are computed with calibrated parameters under a mixture normal distribution. Empirical moments are from the PSID and CRSP.

Panel A: Moments of stock return shock

	Model	Data
SD	0.216	0.216
3rd moment	-1.412	-1.410
4th moment	8.595	8.620

Panel B: Moments of labor income shock

		No college		College	
		Model	Data	Model	Data
SD:	25 – 35	0.378	0.408	0.386	0.405
	35 – 45	0.428	0.432	0.412	0.395
	45 – 55	0.449	0.443	0.416	0.392
	55 – 65	0.457	0.464	0.417	0.414
3rd moment:	25 – 35	-0.021	-0.021	-0.032	-0.026
	35 – 45	-0.027	-0.032	-0.034	-0.036
	45 – 55	-0.029	-0.032	-0.035	-0.040
	55 – 65	-0.029	-0.033	-0.035	-0.044
4th moment:	25 – 35	0.066	0.081	0.076	0.086
	35 – 45	0.107	0.106	0.097	0.094
	45 – 55	0.128	0.116	0.100	0.093
	55 – 65	0.138	0.130	0.101	0.105

Panel C: Dependence structure

	No college		College	
	Model	Data	Model	Data
Linear correlation	0.0382	0.0377	0.0457	0.0456
Between-squares correlation	0.0327	0.0338	0.0692	0.0696

income becomes more stock-like. Households need to balance their stock holdings in order to control for total risk. Introducing nonlinear dependence provides us with a better understanding of labor income risk without a counterfactually high-risk aversion, which may complicate an explanation of other economic behaviors.

Figure 5: Calibration Results

The three graphs show the comparison among the baseline model, the model without between-squares correlation, and SCF data for an average household. To align with the PSID, we assume that households with college degrees make up 30% of the world's population. The parameters for the baseline model have been taken from Table 2. For the main model without between-squares correlation, we adjust only the two correlation parameters to decrease between-squares correlation to zero. We use triennial SCF data from 2007 to 2019 to calculate empirical moments. We calculate moments using nonoverlapping age groups from 20 to 65. Each age group contains samples of three successive ages. Panels A, B, and C show the results of participation rate, conditional risky share, and normalized wealth, respectively.

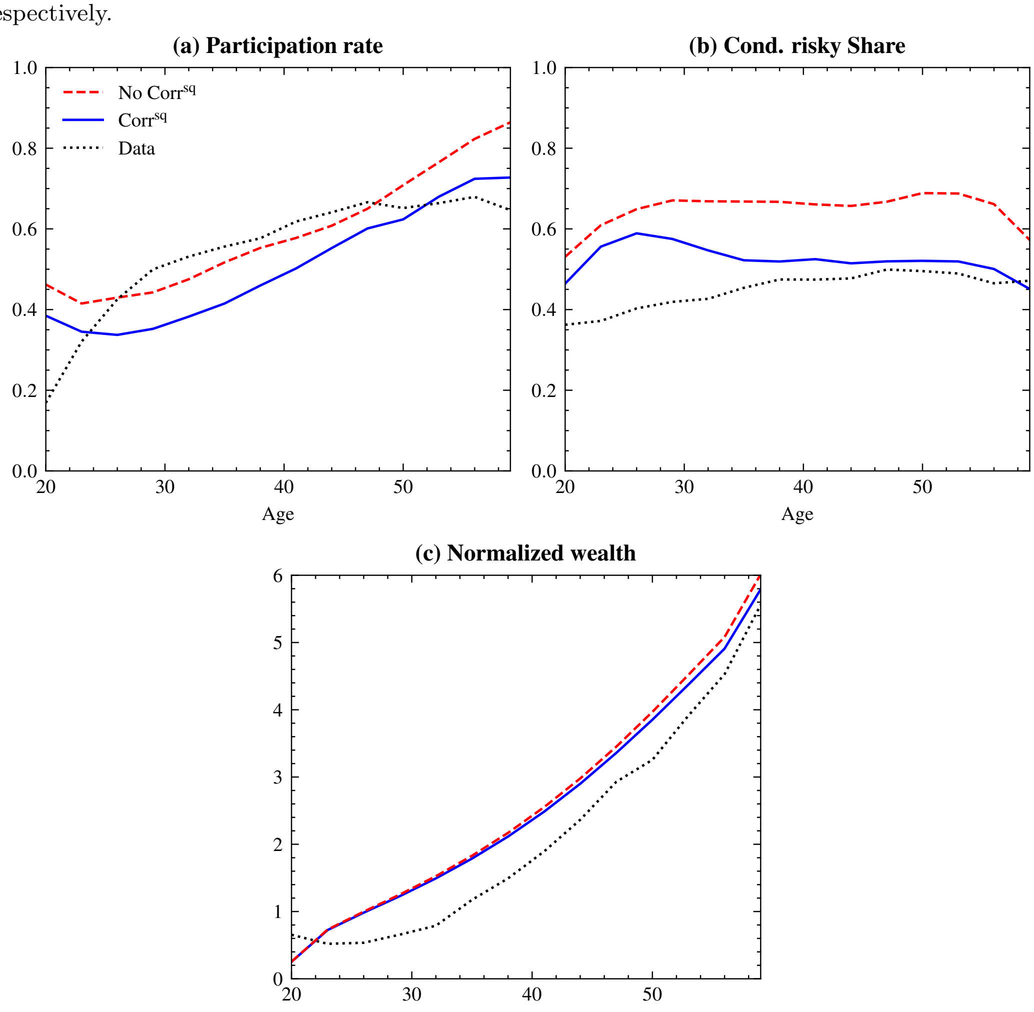


Figure 5 shows the life-cycle profiles of participation rate, conditional risky asset share and the normalized wealth for an average household. In order to gain a more comprehensive understanding of the impact of the between-squares correlation, we also simulate life-cycle profiles in a model that does not account for the between-squares correlation. This alternative model can be viewed as being similar to the ones presented in Catherine (2022) and Shen (2022), and serves as a useful point of comparison. To do so, we fix all parameters in Table 2, except these two dependence parameters. We adjust ρ_1 and ρ_2 to obtain a dependence structure with the same correlation as our main model, but with a zero between-squares correlation. Specifically, the between-squares correlation is manually adjusted from 0.069 (0.033 for the no-college group) to 0. We find that without the between-squares correlation, the model generates much more imprudent portfolio choices. The participation rate also increases by 10% to 20% across different age groups, and conditional risky asset shares increase by 20% on average. The model without the between-squares correlation fails to fit the data with moderate risk aversion and a low fixed cost, although the between-squares correlation only changes slightly.

5 Quantitative Analysis

The between-squares correlation, as a measure of nonlinear dependence, exposes households to a new risk source and makes labor income more stock-like. To better understand the implications of the between-squares correlation, we present the benchmark model results with calibrated parameters and compare them with those generated by the model without the between-squares correlation. We keep correlations the same for both models.

5.1 Policy Function

Figure 6 shows the policy functions for the share of wealth in stocks at different ages for both the no-college group and college graduates. Conditional on participation, introducing a small between-squares correlation lowers the optimal portfolio rule significantly, with the largest drop at 52% at age 40 for the college group.

We also observe that the optimal portfolio rule decreases with both financial wealth and age, which is consistent with the literature (e.g., Cocco et al. (2005)). This is mainly driven by the bond-like property of labor income. The present value of labor income can be considered to be a substitution for riskless assets. Households will evaluate the risk-free position from labor income and balance the portfolio choice. The effect is weaker for wealthier households since labor income becomes trivial compared to wealth. Thus, households with less wealth will hold more positions of riskless assets through labor income and make more aggressive portfolio choices. However, labor income is not a complete substitution for riskless assets since it has risks from different shocks. As stated in Benzoni et al. (2007), labor income also has stock-like features. The nonlinear dependence on stock returns will make labor income more stock-like when considering the between-squares correlation. Although the decreasing pattern remains, the model with the between-squares correlation narrows the gap between the rich and poor.

To take a closer look, we calculate the difference in the optimal portfolio rules between the benchmark model and the model without the between-squares correlation. Table 4 presents the results. We find the difference is much more significant for the no-college group with little wealth. For example, among all age-20 households without college degrees, the largest decrease is 14.8%; however, for the wealthiest age-20 households, this difference can be as small as 5.6%. This heterogeneity across different wealth levels is driven by the fact households with less wealth attaching more importance to the value of their labor income stream, and extra risk exposure from the between-squares correlation will make them tilt their financial portfolio more aggressively toward safe assets.

The negative correlation with age follows from the same logic. For older households, the optimal portfolio rule is less aggressive, since the present value of labor income decreases with age. To compensate for this drop in bond-like wealth, households reduce their relative holding of risky financial assets. Table 4 quantifies this heterogeneity across ages. The largest drop decreases from 22.80% (14.0%) at age 40 to 21.6% (5.6%) at age 60 for the no college group (college graduates).

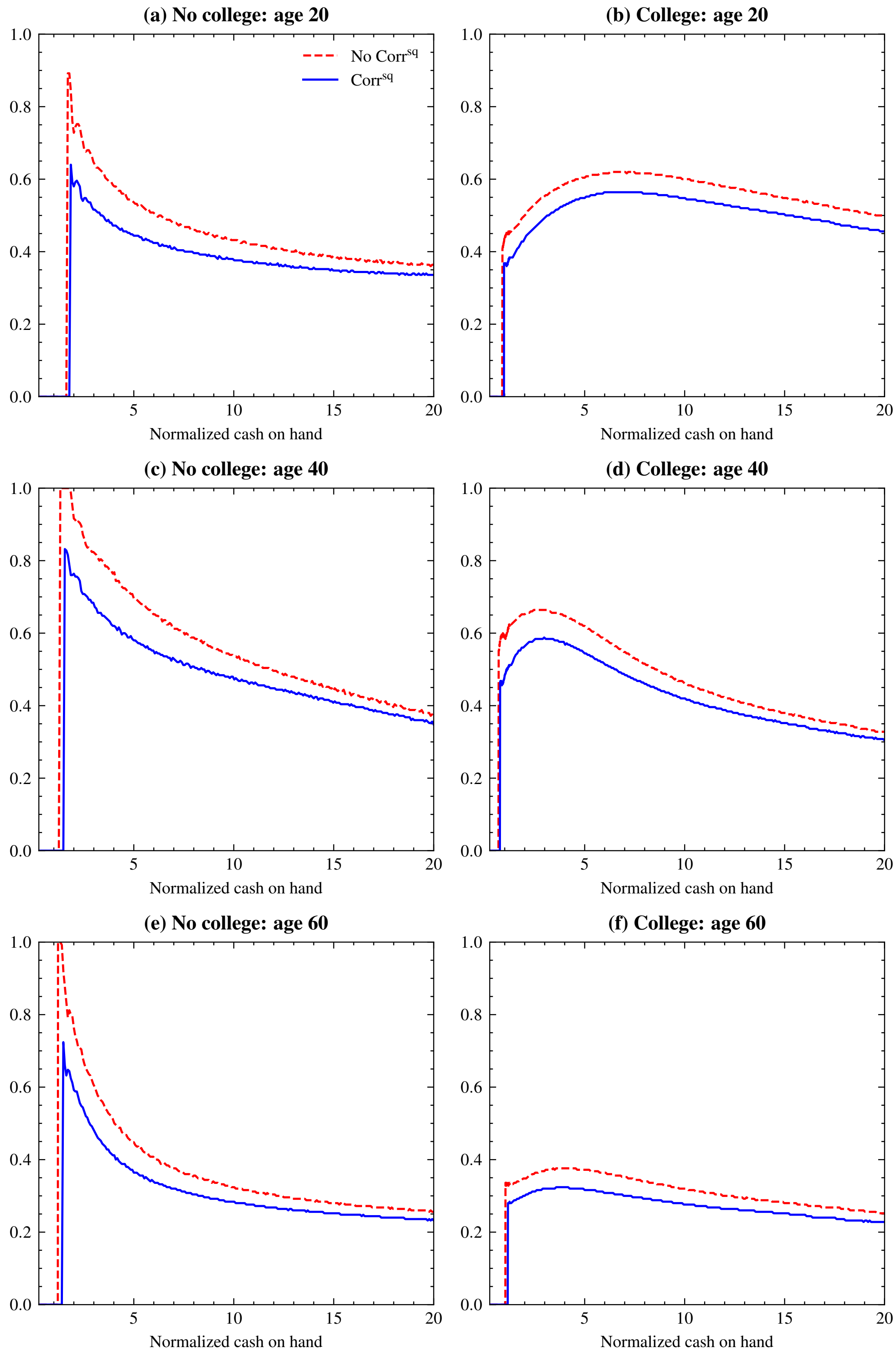
Table 4: Change in Optimal Risky Asset Shares with the between-squares correlation
This table reports the difference in optimal risky share α_t between models with and without the between-squares correlation for college graduates and the no-college group. The data are from the policy functions discussed in Section 5.1. We calculate the difference for each age and wealth ratio. For instance, the second percentage (-6.4%) corresponding with the age-20 college graduates is calculated as difference in the optimal share at a wealth ratio of 2.0 from model with the between-squares correlation (43.6%) from the model without the between-squares correlation (50%). We also report the largest difference for each age.

Group	Age	Largest diff.	Wealth ratio			
			2.0	4.0	8.0	10.0
No college	20	−20.4%	−14.8%	−10.8%	−6.4%	−5.6%
	40	−22.8%	−15.2%	−14.8%	−8.0%	−6.4%
	60	−21.6%	−17.2%	−8.8%	−5.2%	−4%
College	20	−8.8%	−6.4%	−6.0%	−6.0%	−5.2%
	40	−14.0%	−8.8%	−8.0%	−5.6%	−4.8%
	60	−5.6%	−4.8%	−5.2%	−4.8%	−4.4%

Including a per-period participation cost allows us to discuss the wealth-participation threshold. Two facts may prevent households from participating in the stock market. First, the participation cost might be too high relative to households' accumulated wealth. Then for households with less wealth, the optimal choice is not to pay for stock market participation. Second, labor income risk alone might already exceed households' risk tolerance. When households expect higher labor income

Figure 6: Optimal Risky Asset Shares Policy Function

The six panels plot the optimal risky asset shares policy functions for different ages and groups. The first column corresponds with the no-college group and the second corresponds with the college group. The policy functions are solved with calibrated parameters. We take the adjusted persistent income shock $w_{i,t}$ as 1.

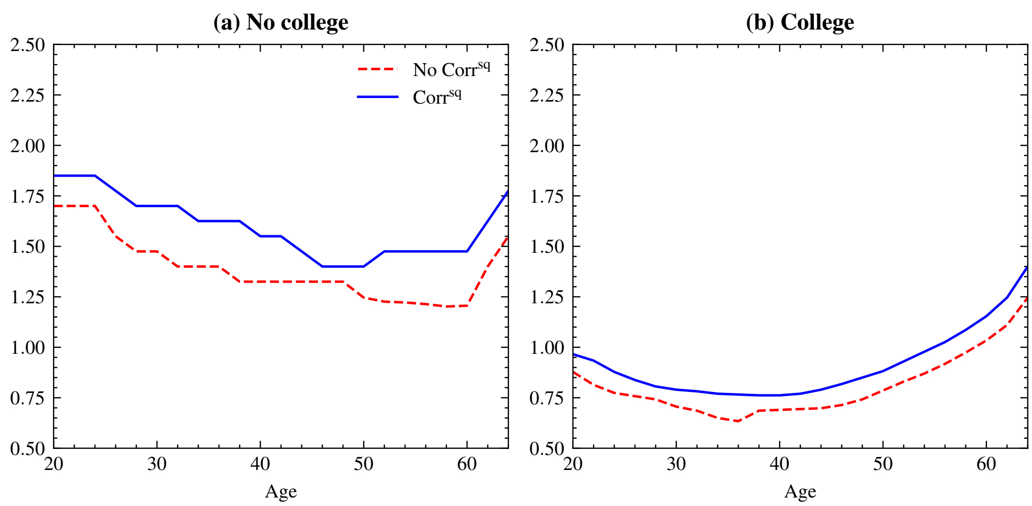


risk, they will not participate in the stock market as a way to avoid taking additional risks. Such a case is less likely for those wealthier households since labor income makes up a smaller proportion of their wealth. Introducing between-squares correlation makes labor income more stock-like, thus affecting households' participation decisions through the second channel.

Figure 7 presents the wealth-participation threshold with respect to age. For both groups, the inclusion of the between-squares correlation implies a higher threshold for participation. In general, the wealth threshold increases by 3.32% to 22.64% (5.66% to 17.26%) for the no-college group (college graduates). Moreover, the U-shaped pattern of the participation threshold is also consistent with the results in Fagereng et al. (2017). At early ages, households accumulate little wealth, thereby limiting the benefit of participating below the participation cost. When households age, they accumulate more wealth, resulting in a higher optimal risky asset share and making stock market participation more attractive. Therefore, the wealth threshold decreases. When households approach retirement age, the fall in the present value of future labor income flows gradually dominates and leads to the rebalancing of the portfolio. In addition, we also observe heterogeneity between college graduates and the no-college group. The U-shaped pattern of the participation threshold decays for both groups when the between-squares correlation is included, but the effect is much stronger for households without college degrees.

Figure 7: Participation Wealth Threshold

The two panels plot the participation wealth threshold across a range of ages for the two groups. The threshold is calculated from the corresponding equity share policy function. It takes the largest value of wealth ratios at which households will not participate in the stock market. We take the adjusted persistent income shock $w_{i,t}$ as 1.



5.2 Simulation Results

We simulate the participation decisions, risky asset shares, and wealth accumulation of $100,000$ households over their life cycle. Figure 8 shows the average profiles for college graduates and the no-college group.

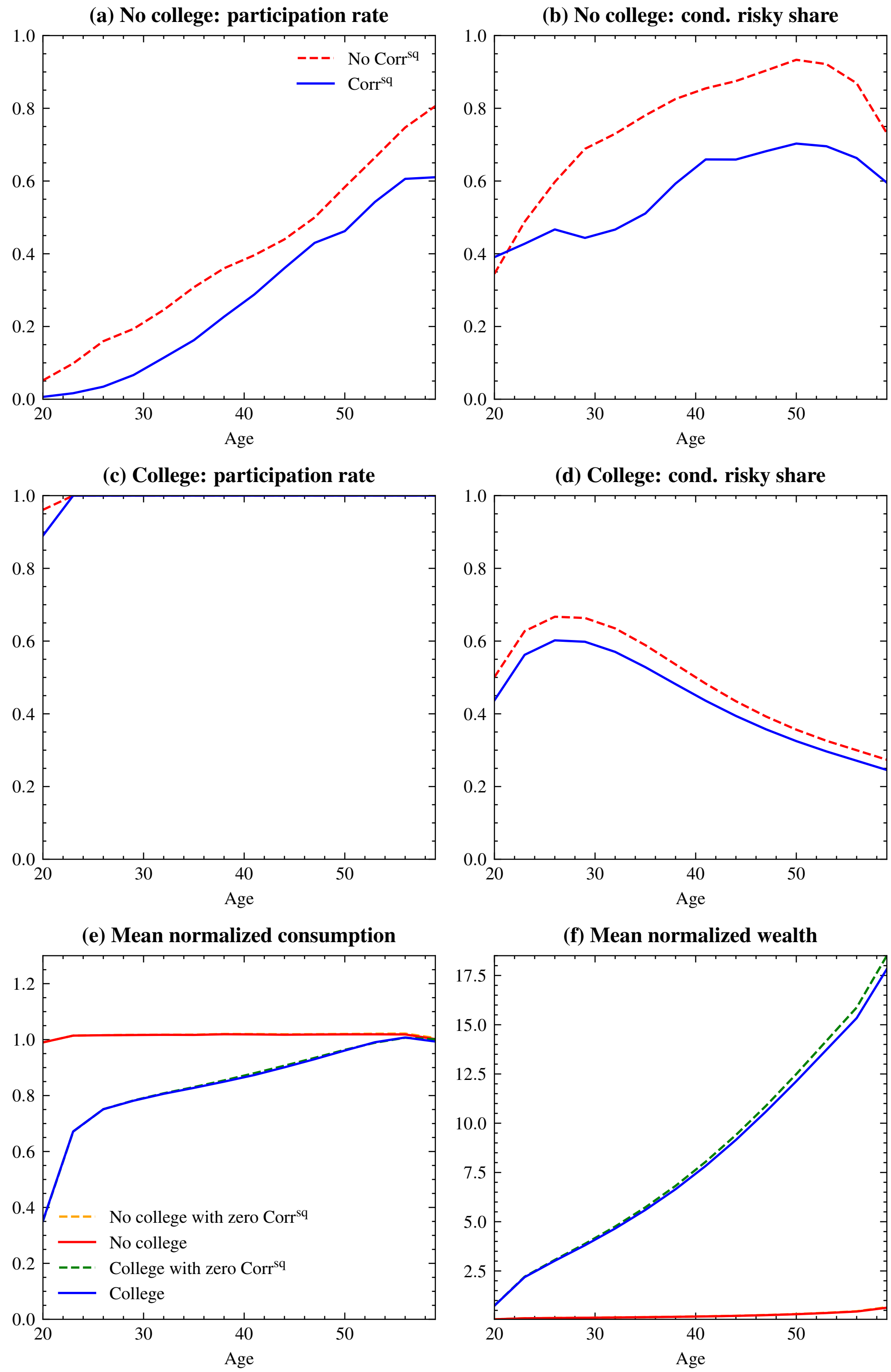
First, we find that the participation rate increases with age no matter whether or not the between-squares correlation is included (panels A and C). Since households typically accumulate more wealth as they age, it is much easier for older households to pass the participation threshold. Second, average conditional risky asset shares are roughly hump shaped with respect to age (panels B and D). This is consistent with Cocco et al. (2005); however, a large correlation, such as 0.2 and 0.4, is required in Cocco et al. (2005) to generate a very significant hump-shaped pattern of conditional risky share. Here, we do not need to impose a large correlation. Instead, a small between-squares correlation and higher-order moments can pull the trigger. Since optimal portfolio rules decrease with wealth, relatively poor investors will be very aggressive and almost invest fully in the stock market, whereas the richer investors will invest more prudently. At each period, some households will just cross the participation threshold and become aggressively poor investors.

On the contrary, once households enter the market, they will decrease their risky asset shares with more wealth accumulation. During early working periods, since most households have very low wealth accumulation, the effect of new investors will dominate and increase average conditional risky asset shares. When more and more households participate in the market, the rebalancing effect from those investors who stay in the market will dominate and decrease the average risky asset shares. Thus, the average optimal risky asset share is hump shaped. Its peak depends on the steepness of the participation rate. In our model, college graduates accumulate wealth much faster and are more likely to enter the stock market in their early twenties. Thus, conditional risky asset shares for college graduates reach the peak much earlier compared with that for the no-college group.

Then we find that the introduction of the between-squares correlation decreases the share of wealth in stocks to a large extent but leads to only a tiny reduction in average wealth (panel F) and average consumption (panel E). The existence of the between-squares correlation between stock returns innovations and labor income shocks indicates that large joint downward movements are more likely, thus making labor income more uncertain and underminingg the nature of income serving as a riskless asset. As a result, the average return on households' financial wealth is lower, thereby lowering households' the accumulated wealth over their life cycle.

Figure 8: Life-Cycle Profile of Portfolio Decisions

The six panels plot the life-cycle profiles for the two groups and compare the results between the baseline model and the model without the between-squares correlation. The profiles are calculated from a simulation of 100000 households.



5.3 Sensitivity Analysis

To test the robustness of the effect of the between-squares correlation, we perform a sensitivity analysis with respect to relative risk aversion, EIS, and the discount rate. For simplicity, we report only the results for the college group at age 20. Results for the no-college group and other ages are qualitatively similar and are available upon request.

Figure 9: Sensitivity Analysis

The four panels plot the mean conditional risky share across ages for different parameters. In each panel, we change only one parameter to test the robustness. The profiles are calculated from a simulation of 100000 households.

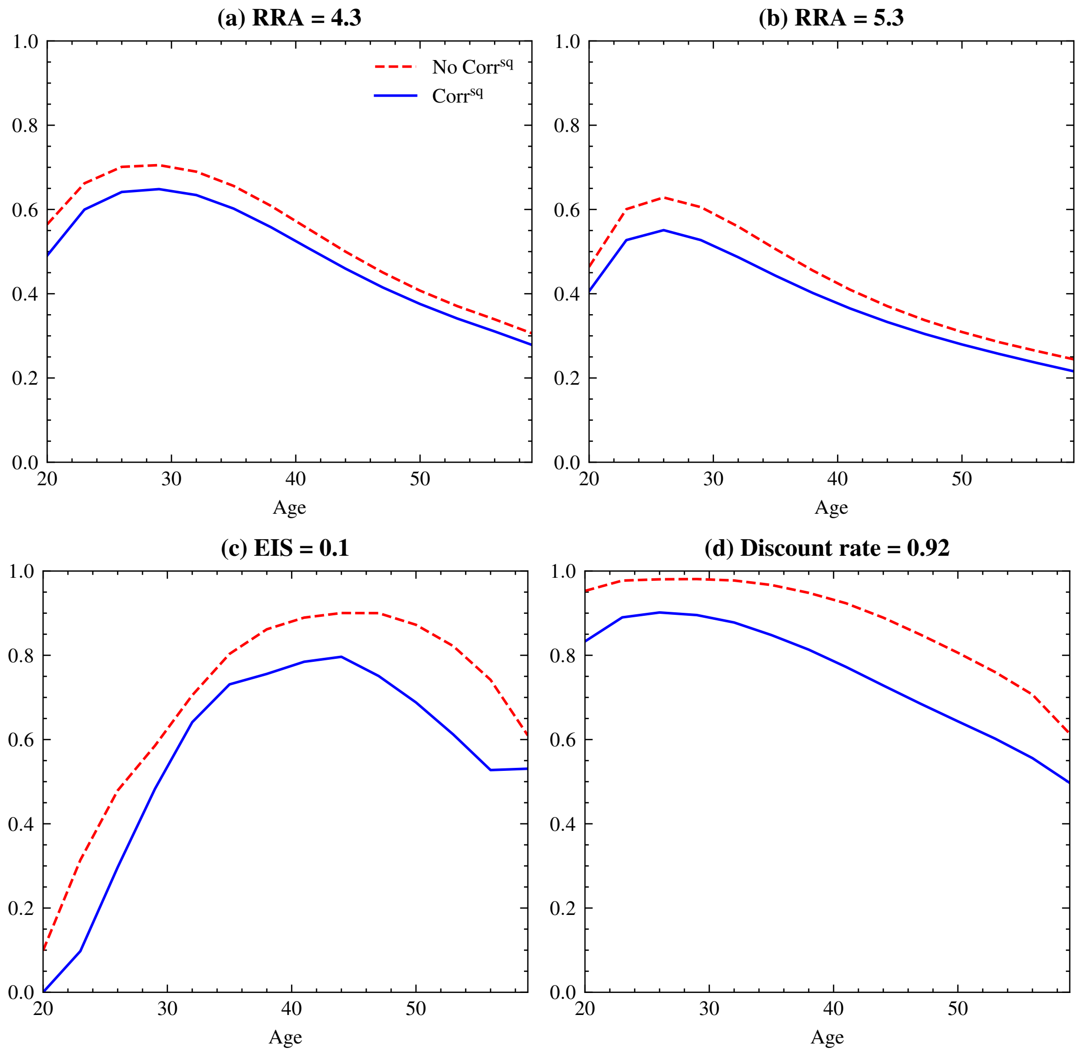


Figure 9 plots the life-cycle profile of risky asset shares with different parameters.[21] We find that the between-squares correlation has a very robust effect on portfolio choices. Although the change of parameters may reshape the risky asset shares curve a lot, the effect of the between-

[21] Throughout this section, we assume that the state variable $w_{i,t}$ takes a value of 1. In Appendix F, we conduct another sensitivity analysis for the case where $w_{i,t}$ is not fixed at 1.

squares correlation always exists and stays strong. Among the different experiments, only when risk aversion is quite low, is the effect of the between-squares correlation relatively small. For example, panel A shows that the effect of the between-squares correlation is not as strong as that in panels B, C and D. Risk-loving households tend to ignore risk and invest most of their wealth in stocks.

6 Empirical Evidence

Calibrated to the U.S. data, our model shows that the presence of the between-squares correlation between stock and labor income significantly affects households' portfolio decisions. In particular, households may choose not to participate in the stock market or delay their participation. Even when households do participate in the stock market, their risky asset holdings are significantly reduced because of the between-squares correlation.

To look for any empirical support for these predictions, we conduct a regression analysis concerning how participation decisions and risky asset holdings respond to the between-squares correlation. To do this, we use data from the PSID.

We only use the PSID from 1997 to 2019, as before 1997, PSID did not report the value of stockholdings.[22] The detailed information enables us to explore the empirical link between labor income and portfolio decisions. We define the variables following Brunnermeier and Nagel (2008). Financial wealth is calculated as the sum of equity in stocks and the value in safe accounts, where the value in safe accounts is the money amount in checking and savings accounts, money market funds, certificates of deposit, government bonds, or Treasury bills. We calculate the ratio of financial wealth invested in stocks and assign a value of one if households participate in the stock market and zero if not. To estimate the between-squares correlation between labor income risk and stock returns risk for each household, we use the annual income level and the CRSP index. Moreover, we observe a nonlinear relationship between the between-squares correlation and households' portfolio decisions from the model.[23] Specifically, the between-squares correlation has an almost opposite relationship with portfolio decisions at the tipping point of zero. Therefore, we use the absolute value of the between-squares correlation, instead of the between-squares correlation, to better capture this nonlinear relationship.

We begin our empirical analysis by estimating a probit stock market participation regression.

[22] Before 1997, only the 1984 and 1989 waves report the data of stock holdings.

[23] See Section 2.3 for details.

In this analysis, we define the dummy variable for participation as the dependent variable:

$$I_{it} = \begin{cases} 1 & \text{if household participates, or} \\ 0 & \text{if household does not participate,} \end{cases} \tag{26}$$

which takes the value of one if households own stocks and the value of zero if households do not participate in the market. We estimate a household's propensity to participate in the stock market, which is represented by $p(I = 1)$.

Our key independent variable is the absolute value of the between-squares correlation, which is inspired by its demonstrated nonlinear impact on portfolio decisions, as discussed in the first model in Section 2.3 and Appendix E.[24] We investigate whether a link exists between the absolute value of between-squares correlation and participation decisions and if there is any heterogeneity between college graduates and households without college degrees. Following the related literature, we control for households' characteristics, such as the level of income and wealth, income risk, marriage status, age, and education. These household-level variables serve as a reasonable proxy for real and perceived market participation costs:

$$\begin{aligned} \mathbb{P}(\alpha_i > 0) =& \Phi(\beta_0 + \beta_1 |Corr^{sq}|_i + \beta_2 Corr_i + \beta_3 \ln(y_i) + \beta_4 age_i + \beta_5 age_i^2 + \\ & \beta_6 w_i + \beta_7 s_i + \beta_8 sk_i + \beta_9 k_i + D_m I_M + \sum D_i^{(year)} I_i^{(year)}), \end{aligned} \tag{27}$$

$$\begin{aligned} \alpha_i^* =& \beta_0 + \beta_1 |Corr^{sq}|_i + \beta_2 Corr_i + \beta_3 \ln(y_i) + \beta_4 age_i + \beta_5 age_i^2 + \\ & \beta_6 w_i + \beta_7 s_i + \beta_8 sk_i + \beta_9 k_i + D_m I_M + \sum D_j^{(year)} I_{ij}^{(year)}, \end{aligned} \tag{28}$$

where α_i is the risky share, Φ is the cumulative distribution function of the standard normal distribution, y_i is the labor income, w_i is the financial wealth, s_i, sk_i, k_i are the standard deviation, skewness, and kurtosis of biennial log growth rate of labor income, I_m is the dummy variable for marriage status, $\{I_{ij}^{(year)}\}$ are the year dummies, and α_i^* is defined by

$$\alpha_i = \begin{cases} 0, & \alpha_i^* \leq 0 \\ \alpha_i^*, & 0 < \alpha_i^* < 1 \\ 1, & \alpha_i^* \geq 1. \end{cases} \tag{29}$$

Table 5 shows the marginal effects from probit market participation regressions. Our results indicate that the absolute value of the between-squares correlation is significantly negatively correlated

[24] Appendix E provides evidence of nonlinear effect of between-squares correlation for our full model.

with households' participation decisions. In particular, we find that when the between-squares correlation is low in absolute value, the market participation propensity increases. This result still holds under different model specifications. Moreover, the coefficients of the between-squares

Table 5: Probit Participation Regression Estimates

This table reports marginal effects from probit regressions. The dependent variable is a dummy variable that denotes whether an individual participates in the stock market. We construct probit models using different sample populations. We omit the constants and year dummies.

	Full sample			No college			College		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\lvert Corr^{sq} \rvert$	-0.541^{***}	-0.318^{***}	-0.251^{***}	-0.576^{***}	-0.387^{***}	-0.246^{***}	-0.423^{***}	-0.255^{***}	-0.247^{***}
	(0.056)	(0.059)	(0.068)	(0.074)	(0.077)	(0.089)	(0.091)	(0.094)	(0.107)
Corr		-0.020	-0.016		-0.060	-0.043		-0.008	-0.027
		(0.036)	(0.037)		(0.048)	(0.049)		(0.055)	(0.057)
ln(y)		0.926^{***}	0.660^{***}		0.852^{***}	0.543^{***}		0.687^{***}	0.522^{***}
		(0.033)	(0.040)		(0.047)	(0.055)		(0.051)	(0.066)
Age/10			-0.282^{**}			-0.062			-0.351
			(0.140)			(0.182)			(0.227)
$Age^2/100$			0.027			0.006			0.039
			(0.017)			(0.022)			(0.027)
Marriage			-0.336^{***}			-0.154^{***}			-0.324^{***}
			(0.051)			(0.069)			(0.080)
FWealth			0.244^{***}			0.253^{***}			0.167^{***}
			(0.009)			(0.014)			(0.011)
$Std(d\delta)$			-0.015			-0.037			0.0422
			(0.097)			(0.129)			(0.152)
$Skew(d\delta)$			-0.022			-0.016			-0.025
			(0.022)			(0.030)			(0.035)
$Kurt(d\delta)$			-0.025^{*}			-0.030			-0.035^{*}
			(0.014)			(0.018)			(0.021)
N		7,954			4,794			3,160	

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

correlation remain statistically significant for both college graduates and the no-college group. Except for the regressions with full controls, the correlation is a little stronger for the no-college group than the college graduates (-0.576 from model (4) vs. -0.423 from model (7), -0.387 from model (5) vs. -0.255 from model (8)). The difference in estimates suggest that households with college degrees are more resilient.

Next, we look for a link between the absolute value of the between-squares correlation and risky asset shares. We run tobit regressions for market participants. In this exercise, the dependent variable is the share of wealth in stocks and the independent variable of interest is the absolute value of the between-squares correlation. Additionally, we control for a set of variables that may cause movement in risky asset holdings.

Table 6 presents the regression results for all households, the no-college group, and college graduates. Our focus is on the coefficient for the absolute value of the between-squares correlation. We find it is consistently negative for all households, the no-college group, and college graduates, with a high statistical significance. For example, in model specification (3), where we consider a broad set of control variables, the coefficient estimate of the absolute value of the between-squares correlation (-0.110) remains significant. For college graduates, the coefficient estimate is -0.094, which is smaller than that (-0.131) for the no-college group, although both estimates are statistically significant. Overall, we find that the empirical evidence supports our model's predictions: for all households, the absolute value of the between-squares correlation is negatively correlated with participation in the stock market and risky asset holdings. This result is stronger for the no-college group compared with college graduates.

Table 6: Tobit Model of the Risky Share

This table reports marginal effects from tobit regressions. The dependent variable is the risky share of households. We construct tobit models using different sample populations. We omit the constants and year dummies.

	Full sample			No college			College		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\lvert Corr^{sq}\rvert$	-0.318^{***}	-0.191^{***}	-0.110^{***}	-0.370^{***}	-0.246^{***}	-0.131^{***}	-0.221^{***}	-0.151^{***}	-0.094^{**}
	(0.031)	(0.030)	(0.033)	(0.047)	(0.046)	(0.049)	(0.040)	(0.039)	(0.042)
$Corr$		-0.015	-0.008		-0.036	-0.032		-0.012	-0.005
		(0.018)	(0.018)		(0.029)	(0.027)		(0.023)	(0.023)
$ln(y)$		0.470^{***}	0.456^{***}		0.498^{***}	0.325^{***}		0.286^{***}	0.282^{***}
		(0.017)	(0.019)		(0.028)	(0.030)		(0.021)	(0.025)
$Age/10$			-0.514			0.466			0.314
			(0.675)			(1.001)			(0.885)
$Age^2/100$			1.309			-0.231			0.332
			(0.796)			(1.186)			(1.037)
$Marriage$			-0.186^{***}			-0.089^{**}			-0.140^{***}
			(0.025)			(0.038)			(0.031)
$FWealth$			0.008^{***}			0.076^{***}			0.004^{***}
			(0.000)			(0.005)			(0.001)
$Std(d\delta)$			0.063			-0.020			0.114^{*}
			(0.047)			(0.071)			(0.060)
$Skew(d\delta)$			-0.012			-0.018			-0.004
			(0.010)			(0.016)			(0.013)
$Kurt(d\delta)$			-0.004			-0.004			-0.013^{*}
			(0.006)			(0.010)			(0.007)
N		7,954			4,794			3,160	

$^{*}\ p < 0.1$; $^{**}\ p < 0.05$; $^{***}\ p < 0.01$

7 Eonomics Implications of the between-squares correlation

7.1 Cost of Ignoring the between-squares correlation

So far, we have shown that the between-squares correlation has a significant impact on households' participation decisions and portfolio choices. However, it is also important to understand how the between-squares correlation affects households' welfare. To answer this question, we consider the equivalent wealth loss from ignoring the between-squares correlation between the labor income market and the stock market. We denote v, v^0 as the value functions of models with the between-squares correlation and without the between-squares correlation, respectively. For other parameters, we take the value from Section 4.3. Then the equivalent wealth loss $\Delta x(t, w)$ at age t with adjusted persistent income shock w can be defined as

$$v(t, x, w) = v^0(t, x + \Delta x, w). \tag{30}$$

Table 7: Cost of Ignoring the between-squares correlation

This table reports the relative equivalent wealth loss, $\Delta x/x$ for two education groups at various ages and levels of wealth. The values reported in the table were obtained using the parameter values calibrated in Section 4, Tables 2 and 3, with the adjusted persistent income shock as $w = 1$.

Group	Age	Normalized wealth					
		1	2	4	6	8	10
No college	20	3.32%	2.86%	2.54%	2.05%	1.87%	1.74%
	40	3.14%	2.67%	1.97%	1.83%	1.58%	1.44%
	60	0.28%	0.18%	0.14%	0.15%	0.12%	0.11%
College	20	1.69%	1.57%	1.25%	1.00%	0.84%	0.72%
	40	1.24%	0.89%	0.53%	0.34%	0.23%	0.13%
	60	0.05%	0.09%	0.10%	0.09%	0.08%	0.07%

Table 7 reports the relative equivalent wealth loss, $\Delta x/x$, against different ages and wealth with an adjusted persistent income shock as $w = 1$. The results are reported separately for households with college degrees and households without college degrees. Our findings suggest that ignoring the between-squares correlation between labor income and stock markets can lead to substantial welfare losses for households. This effect is especially pronounced for households without college degrees. In particular, the welfare losses for households without college degrees can reach up to 3.32% of equivalent wealth. These losses are economically significant, given that the between-squares correlation is estimated to be 0.033 for the no-college group. Furthermore, we find that the equivalent wealth loss decreases with age and wealth, which is consistent with the diminishing effect of the between-squares correlation on optimal portfolio rules and the participation threshold.

7.2 Wealth Inequality

Given the importance of the between-squares correlation for household welfare, it is natural to consider how it may affect macroeconomics, particularly with respect to wealth inequality. Recent studies have made major inroads into documenting trends in wealth inequality in the United States (see Piketty and Saez (2003); Saez and Zucman (2020)). The causes and consequences of these trends in wealth inequality have become a prominent topic of debate in recent years, underscoring the importance of understanding the role of the between-squares correlation in shaping broader economic outcomes.

To investigate the potential impact of the between-squares correlation on wealth inequality, we simulate 100,000 households over a period of 500 years. We then calculate the Gini index, a widely used measure of wealth inequality, and compare the Gini indexes obtained when the between-squares correlation is recognized to those obtained when it is ignored. We recognize that dividing households into only two groups may not capture the full range of heterogeneity in their preferences and attitudes towards risk. Therefore, we explore various combinations of discount factors and risk aversion to better capture the diversity of household behaviors in their saving and investment decisions (see MaCurdy (1982); Gourinchas and Parker (2002)). Furthermore, external shocks can cause households to change their preferences over time, as demonstrated by Guiso et al. (2018).

Table 8 presents the difference in Gini indexes between ignoring and recognizing the between-squares correlation. The between-squares correlation for households with college degrees is 0.069 and the between-squares correlation for households without college degrees is 0.033. Panel A of Table 8 shows that when the between-squares correlation is ignored, the Gini index increases inefficiently for the full sample of households. Specifically, with different combinations of discount rates and risk aversion parameters, the difference in Gini indexes increases by 0.62% to 2.72%, indicating a significant impact of ignoring the between-squares correlation on wealth inequality.[25]

When households ignore the between-squares correlation between the labor income and stock markets, they overlook the interconnections between these two markets. This can result in a misallocation of resources, where households may invest more heavily in the stock market. Moreover, such misallocation can have a heterogeneous effect on households with different levels of wealth.

[25] Considering the between-squares correlation yields a Gini index of 0.203, while ignoring it generates a higher Gini index of 0.210. The models generate lower Gini coefficients than empirical data due to two limitations of our baseline model. Firstly, the model fails to capture the behavior of the wealthiest households, who accumulate wealth through means that differ significantly from labor income. Secondly, the model does not account for portfolio heterogeneity, with rich households having portfolios dominated by corporate and noncorporate equities while middle-class households have highly leveraged portfolios heavily concentrated in residential real estate. Despite these limitations, the models still exhibit a valid trend, and incorporating these features could result in even greater wealth inequality.

Households with lower levels of wealth may increase their risky asset shares more than those with higher levels of wealth, even though the latter group may invest more in dollar value terms. Therefore, the misallocation of resources resulting from ignoring the between-squares correlation can exacerbate wealth inequality. Households with higher levels of wealth may be better equipped to diversify their portfolios, manage their risk exposure and benefit more from investing in the stock market, while households with lower levels of wealth may be more vulnerable to changes in the labor and stock markets.

Meanwhile, Table 8 panel B suggests that the difference between ignoring and recognizing the between-squares correlation is even wider for households with college degrees compared to those without. Specifically, the difference in Gini index can be as high as 3.63% for households with college degrees, whereas it is only 2.28% for households without college degrees. This indicates a greater concentration of wealth among households with higher levels of education, which may contribute to overall inequality in the society.

Table 8: Changes in Gini Index

This table examines the impact of the between-squares correlation on wealth inequality, comparing the Gini index obtained when this correlation is ignored to those obtained when it is recognized. Panel A shows the results for full samples of households considering different combinations of discount rates and risk aversion parameters. Panel B shows the results separately for households with and without college degrees using the same combinations of discount rates and risk aversion parameters as in panel A. The notation (β^c, β^{nc}) indicates that we use β^c as the discount rate for the college group and β^{nc} as the discount rate for the no college group. This notation allows us to distinguish between the two education groups and analyze their respective impacts on wealth inequality.

Panel A: Full samples

	Risk aversion		
Discount rate	4.3	4.8	5.3
(92%, 92%)	1.47%	2.21%	2.72%
(94%, 92%)	1.66%	2.16%	2.19%
(96%, 92%)	1.50%	1.60%	1.47%
(98%, 92%)	0.84%	0.85%	0.62%

Panel B: Changes in Gini Index for each education group

	Risk aversion		
Discount rate	4.3	4.8	5.3
(92%, 92%): College	1.20%	2.42%	3.63%
(94%, 92%): College	1.52%	2.28%	2.78%
(96%, 92%): College	1.32%	1.59%	1.64%
(98%, 92%): College	1.25%	1.32%	1.01%
(92%, 92%): No college	1.28%	1.80%	2.28%

Moreover, given that we observe a right-skewed distribution of the between-squares correlation in Section 2.2, it is important to investigate the effect of higher correlation levels to fully understand the potential impact on wealth inequality. This motivates our investigation into the differences in Gini index, as shown in Table 9, particularly in relation to higher levels of between-squares correlation. To achieve this, we raise the between-squares correlation by 50% for both education groups. Our findings indicate that as the between-squares correlation increases, the difference in the Gini index between recognizing and ignoring the between-squares correlation for all samples reaches a maximum of 3.72%, increasing from the initial 2.72%. Specifically, for households with college degrees, the difference is even wider at 4.88%.

While the magnitude of these differences may not seem significant at first glance, it is important to note that the Gini index has increased by less than 3.0% from 1970 to 2017, and the maximum change in the US over time since 1969 has been 6.76%.[26] This highlights the importance of accounting for the between-squares correlation, as ignoring it can have a considerable impact on wealth inequality.

Overall, recognizing the importance of the between-squares correlation between the labor income and stock markets can help households make more informed decisions about their investment and risk management strategies. This can lead to a more efficient allocation of resources and a potentially more equitable distribution of wealth. Our findings highlight the importance of a channel that has received relatively little attention to date: the between-squares correlation between labor income and stock markets induces large shifts in the wealth inequality.

8 Conclusion

In this paper, we consider the optimal portfolio decision of a household when stock returns and income shocks are nonlinearly dependent. We show that in the presence of the between-squares correlation, households are less willing to participate in the market and significantly reduce their stock investments. Moreover, our paper shows that the between-squares correlation is independent of the linear correlation, which the literature provides mixed evidence about. Therefore, our model complements existing studies and can potentially help explain both the limited participation puzzle and moderate risky asset holdings observed in the data. Using a simple one-period model, we show that the impact of the between-squares correlation on portfolio choices can be attributed to two pathways: the skewness channel and the kurtosis channel. The degree to which one channel dominates over the other is contingent on the level of between-squares correlation. We further provide empirical evidence that is supportive of the model's prediction. In addition, we find that ignoring the between-squares correlation reduces households' welfare and increases wealth inequality.

[26]Figure 11 in Appendix F compares the Gini coefficient in each year to the benchmark year of 1969 and displays the relative changes over time.

Table 9: Changes in Gini Index with Higher Between-squares Correlation

This table examines the impact of the between-squares correlation on wealth inequality, comparing the Gini index obtained when this correlation is ignored to those obtained when it is recognized, with a focus on higher levels of between-squares correlation. Panel A shows the results for full samples of households considering different combinations of discount rates and risk aversion parameters. Panel B shows the results separately for households with and without college degrees using the same combinations of discount rates and risk aversion parameters as in panel A. The notation (β^c, β^{nc}) indicates that we use β^c as the discount rate for the college group and β^{nc} as the discount rate for the no college group. This notation allows us to distinguish between the two education groups and analyze their respective impacts on wealth inequality.

Panel A: Full samples

	Risk aversion		
Discount rate	4.3	4.8	5.3
(92%, 92%)	2.00%	3.04%	3.72%
(94%, 92%)	2.33%	2.96%	3.07%
(96%, 92%)	2.12%	2.25%	2.15%
(98%, 92%)	0.87%	0.92%	0.90%

Panel B: Changes in Gini Index for each education group

	Risk Aversion		
Discount rate	4.3	4.8	5.3
(92%, 92%): College	1.59%	3.27%	4.88%
(94%, 92%): College	2.08%	3.08%	3.75%
(96%, 92%): College	1.84%	2.21%	2.50%
(98%, 92%): College	1.09%	1.22%	1.24%
(92%, 92%): No college	1.47%	2.01%	2.48%

Although our model accounts for between-squares correlation between labor and stock markets, it does not endogenize this nonlinear dependence, nor does it address the asset pricing implications. To further explore this relationship, it would be valuable to study the between-squares correlation using a general equilibrium framework, as in Guvenen (2009) for a macroeconomic model. Such an extension would offer a characterization of the risk premium, taking into account the complex and interdependent nature of labor and stock markets.

References

Ameriks, J. and Zeldes, S. P. (2011). *How do household portfolio shares vary with age?* Columbia University.

Angerer, X. and Lam, P.-S. (2009). Income risk and portfolio choice: An empirical study. *The Journal of Finance*, 64(2):1037–1055.

Attanasio, O. P. (1995). The intertemporal allocation of consumption: theory and evidence. In *Carnegie-Rochester conference series on public policy*, volume 42, pages 39–89. Elsevier.

Bagliano, F. C., Fugazza, C., and Nicodano, G. (2019). Life-cycle portfolios, unemployment and human capital loss. *Journal of Macroeconomics*, 60:325–340.

Baxter, M. and Jermann, U. J. (1997). The international diversification puzzle is worse than you think. *The American Economic Review*, 87(1):170–180.

Benzoni, L., Collin-Dufresne, P., and Goldstein, R. S. (2007). Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. *The Journal of Finance*, 62(5):2123–2167.

Betermier, S., Jansson, T., Parlour, C., and Walden, J. (2012). Hedging labor income risk. *Journal of Financial Economics*, 105(3):622–639.

Bonaparte, Y., Korniotis, G. M., and Kumar, A. (2014). Income hedging and portfolio decisions. *Journal of Financial Economics*, 113(2):300–324.

Brav, A., Constantinides, G. M., and Geczy, C. C. (2002). Asset pricing with heterogeneous consumers and limited participation: Empirical evidence. *Journal of Political Economy*, 110(4):793–824.

Bronars, S. G. and Famulari, M. (2001). Shareholder wealth and wages: Evidence for white-collar workers. *Journal of Political Economy*, 109(2):328–354.

Brunnermeier, M. K. and Nagel, S. (2008). Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals. *American Economic Review*, 98(3):713–36.

Campbell, J. Y. (1996). Understanding risk and return. *Journal of Political economy*, 104(2):298–345.

Campbell, J. Y., Lettau, M., Malkiel, B. G., and Xu, Y. (2001). Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk. *The Journal of Finance*, 56(1):1–43.

Cao, H. H., Wang, T., and Zhang, H. H. (2005). Model uncertainty, limited market participation, and asset prices. *The Review of Financial Studies*, 18(4):1219–1251.

Carroll, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *The Quarterly journal of economics*, 112(1):1–55.

Catherine, S. (2022). Countercyclical labor income risk and portfolio choices over the life cycle. *The Review of Financial Studies*, 35(9):4016–4054.

Chang, Y., Hong, J. H., and Karabarbounis, M. (2018). Labor market uncertainty and portfolio choice puzzles. *American Economic Journal: Macroeconomics*, 10(2):222–62.

Cocco, J. F. (2005). Portfolio choice in the presence of housing. *The Review of Financial Studies*, 18(2):535–567.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *The Review of Financial Studies*, 18(2):491–533.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223.

Davis, S. J. and Willen, P. (2000). Using financial assets to hedge labor income risks: estimating the benefits. Technical report, Citeseer.

Davis, S. J. and Willen, P. S. (2013). Occupation-level income shocks and asset returns: Their covariance and implications for portfolio choice. *The Quarterly Journal of Finance*, 3(3-4):1–53.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.

Edwards, S. and Susmel, R. (2001). Volatility dependence and contagion in emerging equity markets. *Journal of Development Economics*, 66(2):505–532.

Fagereng, A., Gottlieb, C., and Guiso, L. (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance*, 72(2):705–750.

Faig, M. and Shum, P. (2002). Portfolio choice in the presence of personal illiquid projects. *The Journal of Finance*, 57(1):303–328.

Gomes, F. and Michaelides, A. (2005a). Optimal life-cycle asset allocation: Understanding the empirical evidence. *The Journal of Finance*, 60(2):869–904.

Gomes, F. and Michaelides, A. (2005b). Optimal life-cycle asset allocation: Understanding the empirical evidence. *The Journal of Finance*, 60(2):869–904.

Gomes, F. and Michaelides, A. (2008). Asset pricing with limited risk sharing and heterogeneous agents. *The Review of Financial Studies*, 21(1):415–448.

Gomes, F. J. (2005). Portfolio choice and trading volume with loss-averse investors. *The Journal of Business*, 78(2):675–706.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. econometrica. *Econometrica*, 70(1):41–89.

Granger, C. W. and Ding, Z. (1994). Stylized facts on the temporal and distributional properties of daily data from speculative markets. *UCSD Department of Economics Discussion Paper*, pages 94–19.

Guiso, L., Jappelli, T., and Terlizzese, D. (1996). Income risk, borrowing constraints, and portfolio choice. *The American Economic Review*, pages 158–172.

Guiso, L., Sapienza, P., and Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421.

Guvenen, F. (2009). A parsimonious macroeconomic model for asset pricing. *Econometrica*, 77(6):1711–1750.

Guvenen, F., Ozkan, S., and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660.

Haliassos, M. and Michaelides, A. (2003). Portfolio choice and liquidity constraints. *International Economic Review*, 44(1):143–177.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, 55(3):1263–1295.

Heaton, J. and Lucas, D. (2000). Portfolio choice and asset prices: The importance of entrepreneurial risk. *The Journal of Finance*, 55(3):1163–1198.

Hong, H., Kubik, J. D., and Stein, J. C. (2004). Social interaction and stock-market participation. *The Journal of Finance*, 59(1):137–163.

Hsieh, D. A. (1989). Testing for nonlinear dependence in daily foreign exchange rates. *Journal of Business*, pages 339–368.

Hsieh, D. A. (1993). Implications of nonlinear dynamics for financial risk management. *Journal of Financial and Quantitative Analysis*, 28(1):41–64.

Hubbard, R. G., Skinner, J., and Zeldes, S. P. (1995). Precautionary saving and social insurance. *Journal of political Economy*, 103(2):360–399.

Huggett, M. and Kaplan, G. (2011). Human capital values and returns: Bounds implied by earnings and asset returns data. *Journal of Economic Theory*, 146(3):897–919.

Huggett, M. and Kaplan, G. (2016). How large is the stock component of human capital? *Review of Economic Dynamics*, 22:21–51.

Jagannathan, R., Kocherlakota, N. R., et al. (1996). Why should older people invest less in stocks than younger people. *Federal Reserve Bank of Minneapolis Quarterly Review*, 20:11–20.

Koo, H. K. (1998). Consumption and portfolio selection with labor income: a continuous time approach. *Mathematical Finance*, 8(1):49–65.

LeBaron, B. D. (1988). *Stock return nonlinearities: Some initial tests and findings*. PhD thesis, University of Chicago, Department of Economics.

Lucas, D. and Zeldes, S. P. (2006). Valuing and hedging defined benefit pension obligations-the role of stocks revisited. *Northwestern University and Columbia University, working paper, September*.

Lynch, A. W. and Tan, S. (2011). Labor income dynamics at business-cycle frequencies: Implications for portfolio choice. *Journal of Financial Economics*, 101(2):333–359.

MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics*, 18(1):83–114.

Madan, D. B. and Wang, K. (2020). Correlated squared returns. *Available at SSRN 3560226*.

Malloy, C. J., Moskowitz, T. J., and Vissing-Jørgensen, A. (2009). Long-run stockholder consumption risk and asset returns. *The Journal of Finance*, 64(6):2427–2479.

Mencía, J. (2011). Testing Nonlinear Dependence in the Hedge Fund Industry. *Journal of Financial Econometrics*, 10(3):545–587.

Michelacci, C. and Quadrini, V. (2009). Financial markets and wages. *The Review of Economic Studies*, 76(2):795–827.

Nakajima, M. and Smirnyagin, V. (2019). Cyclical labor income risk. *Available at SSRN 3432213*.

Neuberger, A. and Payne, R. (2021). The skewness of the stock market over long horizons. *The Review of Financial Studies*, 34(3):1572–1616.

Nickell, S. and Wadhwani, S. (1991). Employment determination in british industry: investigations using micro-data. *The Review of Economic Studies*, 58(5):955–969.

Pagel, M. (2018). A news-utility theory for inattention and delegation in portfolio choice. *Econometrica*, 86(2):491–522.

Patton, A. J. (2009). Are "market neutral" hedge funds really market neutral? *The Review of Financial Studies*, 22(7):2495–2530.

Peijnenburg, K. (2018). Life-cycle asset allocation with ambiguity aversion and learning. *Journal of Financial and Quantitative Analysis*, 53(5):1963–1994.

Piketty, T. and Saez, E. (2003). Income inequality in the united states, 1913–1998. *The Quarterly journal of economics*, 118(1):1–41.

Polkovnichenko, V. (2007). Life-cycle portfolio choice with additive habit formation preferences and uninsurable labor income risk. *The Review of Financial Studies*, 20(1):83–124.

Poterba, J. M., Samwick, A., et al. (2001). Household portfolio allocation over the life cycle. *Aging issues in the United States and Japan*, 2.

Saez, E. and Zucman, G. (2020). The rise of income and wealth inequality in america: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives*, 34(4):3–26.

Santos, T. and Veronesi, P. (2006). Labor income and predictable stock returns. *The Review of Financial Studies*, 19(1):1–44.

Scheinkman, J. A. and LeBaron, B. (1989). Nonlinear dynamics and stock returns. *Journal of Business*, pages 311–337.

Shen, J. (2022). Countercyclical risks and optimal life-cycle profile: Theory and evidence. *Management Science*.

Shephard, N. (2010). Stochastic volatility models. *Macroeconometrics and Time Series Analysis*, pages 276–287.

Storesletten, K., Telmer, C. I., and Yaron, A. (2007). Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics*, 10(4):519–548.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *The Journal of Finance*, 56(2):433–470.

Vissing-Jørgensen, A. (2002). Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy*, 110(4):825–853.

Yao, R. and Zhang, H. H. (2005). Optimal consumption and portfolio choices with risky housing and borrowing constraints. *The Review of Financial Studies*, 18(1):197–239.

Internet Appendix

Nonlinear Dependence and Households' Portfolio Decisions over the Life Cycle

Wei Jiang Shize Li Jialu Shen

Appendix A Data

A.1 Panel Study of Income Dynamics

The Panel Study of Income Dynamics (PSID) is a longitudinal household survey spanning from 1968 to 2019, with data from over 18,000 individuals residing in 5,000 families in the United States. The survey covers a wide range of topics, including family information and financial situations, which are essential in most life cycle household models. PSID is particularly suitable for our model, as we require time series data on labor income to estimate its correlations with stock returns. However, PSID has some well-know issues for researchers and many papers have discussed the approaches to deal with them[1].

In this paper, we follow the sample selection principle of Nakajima and Smirnyagin (2019). First, we exclude years prior to 1970, as the 1968 and 1969 waves lack some data and are inconsistent with subsequent waves. Second, we drop SEO and Latino samples, which are collected with unequal selection probabilities. PSID includes approximately 2,000 low-income SEO samples and 2,000 Latino samples, and we exclude them from our analysis. Third, we drop observations with missing or non-positive head/spousal labor incomes. The top 1% of head's and wife's labor incomes are then trimmed, as PSID brackets many finance-related variables (such as labor income) with an upper boundary. This trimming selection is intended to exclude bracketed or extreme samples. Finally, households with income growth anomalies (annual log growth rate outside the range of 1/20 to 20) are also dropped from our analysis. We construct variables following Brunnermeier and Nagel (2008), with variable definitions summarized in the following table.

A.2 SCF Data

The Survey of Consumer Finances (SCF) is a triennial cross-sectional survey of U.S. families that collects information on families' balance sheets, pensions, income, and demographic characteristics. The survey also includes data from related surveys of pension providers and earlier surveys conducted by the Federal Reserve Board. We utilize data from five waves of the survey, spanning from

[1]Such as Cocco et al. (2005), Gomes and Michaelides (2005b) and Nakajima and Smirnyagin (2019).

Table 10: Varible definitions

Variable	Definition
Labor Income Y_{it}	Includes the labor income of both reference person and spouse, labor part of farm income and business income, all transfer income of the family, and social securities.
Riskless Assets	Checking and savings accounts, money market funds, certificates of deposits, savings bonds, and treasury bills
	Bonds, bond funds, cash value in a life insurance, valuable collection for investment purposes, and rights in a trust or estate
Risky Assets	the combined value of shares of stock in publicly held corporations, mutual funds, and investment trusts
IRA Assets	Value of private annuities or Individual Retirement Accounts
Other Debts	Credit card debt, student loans, medical or legal bills, and loans from relatives
Home Equity	Value of the home minus remaining mortgage principal
Liquid Assets	Riskless Assets + Risky Assets
Financial Wealth	Liquid Assets - Other Debts + Home Equity

2007 to 2019, and construct variables using the code-book and macro-variable definitions from the Federal Reserve website.

Following Gomes and Michaelides (2005b), we construct variables for SCF. Labor income is defined as the sum of wages and salaries (X5702), unemployment or worker's compensation (X5716), and Social Security or other pensions, annuities, or other disability or retirement programs (X5722). Financial wealth is then constructed in the same way as variable FIN in the publicly available SCF data set, which is comprised of LIQ (all types of transaction accounts – checking, saving, money market, and call accounts), CDS (certificates of deposit), total directly held mutual funds, stocks, bonds, total quasi-liquid financial assets (the sum of IRAs, thrift accounts, and future pensions), savings bonds, the cash value of whole life insurance, other managed assets (trusts, annuities, and managed investment accounts in which the household has equity interest), and other financial assets (includes loans from the household to someone else, future proceeds, royalties, futures, nonpublic stock, and deferred compensation). Furthermore, data on financial assets invested in the risky asset is obtained from variable EQUITY in the publicly available SCF data set, which consists of directly held stock, stock mutual funds, or amounts of stock in retirement accounts. We calculate the conditional risky share as (EQUITY)/(FIN) conditional on EQUITY being positive.

Appendix B Bivariate Mixture Normal Distribution

This section begins by defining the mixture normal distribution and describing its first four moments. Then, we establish that the between-squares correlation can function as an adequate measure of nonlinear dependence, subject to certain regularity conditions.

Definition 1. *A random vector $(X_1, X_2)^T$ follows a bivariate mixture normal distribution with parameters $p_1, \mu_{11}, \mu_{12}, \sigma_{11}, \sigma_{12}, p_2, \mu_{21}, \mu_{22}, \sigma_{21}, \sigma_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}$ if:*

$$X_1 = I_1 Z_{11} + (1 - I_1) Z_{12}, \tag{31}$$

$$X_2 = I_2 Z_{21} + (1 - I_2) Z_{22}, \tag{32}$$

where $I_i \sim \mathrm{B}(1, p_i)$ and $Z = (Z_{11}, Z_{12}, Z_{21}, Z_{22})^T$ is normally distributed subject to:

$$\mathbb{E}(Z) = (\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22})',$$

$$\mathrm{Cov}(Z) = \begin{bmatrix} \sigma_{11}^2 & 0 & \rho_{11}\sigma_{11}\sigma_{21} & \rho_{12}\sigma_{11}\sigma_{22} \\ 0 & \sigma_{12}^2 & \rho_{21}\sigma_{12}\sigma_{21} & \rho_{22}\sigma_{12}\sigma_{22} \\ \rho_{11}\sigma_{11}\sigma_{21} & \rho_{21}\sigma_{12}\sigma_{21} & \sigma_{21}^2 & 0 \\ \rho_{12}\sigma_{11}\sigma_{22} & \rho_{22}\sigma_{12}\sigma_{22} & 0 & \sigma_{22}^2 \end{bmatrix}.$$

For convenience, in the rest of the appendix, we use the shorthand notation $p_{11} = p_1, p_{12} = 1 - p_1, p_{21} = p_2, p_{22} = 1 - p_2$.

B.1 Central Moments and Correlations

Denote $\mu_i, \sigma_i, s_i, k_i$ as the mean, standard deviation, skewness and kurtosis of $X_i, i = 1, 2$. The first four moments of X_i are given by

$$\mu_i = \mathbb{E}[X_i] = p_{i1}\mu_{i1} + p_{i2}\mu_{i2}, \tag{33}$$

$$\sigma_i^2 = \mathrm{Var}[X_i] = p_{i1}\sigma_{i1}^2 + p_{i2}\sigma_{i2}^2 + p_{i1}p_{i2}(\mu_{i1} - \mu_{i2})^2, \tag{34}$$

$$s_i = \mathrm{Skew}[X_i] = \sigma_i^{-3} p_{i1}p_{i2}(\mu_{i1} - \mu_{i2}) \left(3(\sigma_{i1}^2 - \sigma_{i2}^2) + (1 - 2p_{i1})(\mu_{i1} - \mu_{i2})^2\right), \tag{35}$$

$$\begin{aligned} k_i = \mathrm{kurt}[X_i] = \sigma_i^{-4} \big[&3p_{i2}\sigma_{i1}^4 + 3p_{i1}\sigma_{i2}^4 + p_{i1}p_{i2}(\mu_{i1} - \mu_{i2})^2 \big(6(p_{i1}\sigma_{i2}^2 + p_{i2}\sigma_{i1}^2) + \\ &(3p_{i1}^2 - 3p_{i1} + 1)(\mu_{i1} - \mu_{i2})^2\big)\big]. \end{aligned} \tag{36}$$

The Pearson correlation and between-squares correlations between X_1 and X_2 are given by

$$\mathrm{Corr}(X_1, X_2) = \frac{1}{\sigma_1\sigma_2}\sum_{i,j=1,2} p_{1i}p_{2j}\sigma_{1i}\sigma_{2j}\rho_{ij} \tag{37}$$

$$\begin{aligned}\mathrm{Corr}^{\mathrm{sq}}(X_1, X_2) =&\frac{1}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}\Big(2\sum_{i,j=1,2} p_{1i}p_{2j}\sigma_{1i}^2\sigma_{2j}^2\rho_{ij}^2\\ &+4(\mu_{11}-\mu_{12})(\mu_{21}-\mu_{22})p_{11}p_{12}p_{21}p_{22}\sum_{i,j=1,2}(-1)^{i+j}\sigma_{1i}\sigma_{2j}\rho_{ij}\Big)\,.\end{aligned} \tag{38}$$

The moments and correlations can be calculated straightforwardly using their definitions. However, we elaborate on the process of characterizing the between-squares correlation. Denote $y_{ij} = \dfrac{x_{ij}-\mu_{ij}}{\sigma_{ij}}$, $i, j = 1, 2$, and we construct Schmidt orthogonalization as follows

$$\begin{aligned}y_{11}^* =&y_{11},\\ y_{12}^* =&y_{12},\\ y_{21}^* =&y_{21} - \rho_{11}y_{11}^* - \rho_{21}y_{12}^*,\\ y_{22}^* =&y_{22} - \rho_{12}y_{11}^* - \rho_{22}y_{12}^* + \frac{\rho_{11}\rho_{12}+\rho_{21}\rho_{22}}{1-\rho_{11}^2-\rho_{21}^2}y_{21}^*.\end{aligned}$$

Subsequently, we obtain

$$\begin{aligned}\mathrm{Corr}^{\mathrm{sq}}(X_1, X_2) =&\mathrm{Corr}\big((\sigma_{11}I_1y_{11} + \sigma_{12}(1-I_1)y_{12} + (I_1-p_1)(\mu_{11}-\mu_{12}))^2,\\ &(\sigma_{21}I_2y_{21} + \sigma_{22}(1-I_2)y_{22} + (I_2-p_2)(\mu_{21}-\mu_{22}))^2\big)\\ =&\frac{1}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}\mathrm{Cov}\big((\sigma_{11}I_1y_{11}^* + \sigma_{12}(1-I_1)y_{12}^* + (I_1-p_1)(\mu_{11}-\mu_{12}))^2,\\ &(\sigma_{21}I_2(\rho_{11}y_{11}^* + \rho_{21}y_{12}^* + y_{21}^*) + \sigma_{22}(1-I_2)(\rho_{12}y_{11}^* + \rho_{22}y_{12}^*\\ &-\frac{\rho_{11}\rho_{12}+\rho_{21}\rho_{22}}{1-\rho_{11}^2-\rho_{21}^2}y_{21}^* + y_{22}^*) + (I_2-p_2)(\mu_{21}-\mu_{22}))^2\big)\\ =&\frac{1}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}\mathrm{Cov}\big((\sigma_{11}I_1y_{11}^* + \sigma_{12}(1-I_1)y_{12}^* + (I_1-p_1)(\mu_{11}-\mu_{12}))^2,\\ &((\sigma_{21}I_2\rho_{11} + \sigma_{22}(1-I_2)\rho_{12})y_{11}^* + (\sigma_{21}I_2\rho_{21} + \sigma_{22}(1-I_2)\rho_{22})y_{12}^* +\\ &(I_2-p_2)(\mu_{21}-\mu_{22}))^2\big).\end{aligned}$$

By exploiting the mutual independence of y_{ij}, we derive (38).

B.2 Sufficiency of the First Four Moments

In Appendix B.1, we provided the first four moments of a mixture normal distribution. In this section, we explore whether these moments can uniquely characterize such a distribution. To achieve this, we will outline particular regularity conditions that establish the uniqueness of a mixture normal distribution based on its first four moments alone.

Theorem 2. *Let X_1 be a random variable that follows a mixture normal distribution as defined in* (31). *Denote the mean, standard deviation, skewness, and kurtosis of X_1 as $\mu, \sigma, s,$ and k, respectively. For convenience, we refer to the parameters of X_1 as $p, \mu_1, \mu_2, \sigma_1, \sigma_2,$ and $u_d(X_1) = \mu_1 - \mu_2$. If the following conditions hold:*

$$\frac{1}{2} + \frac{\sqrt{15}\sqrt{13-2\sqrt{31}}}{30} < p < 1, \tag{39}$$

$$-\frac{2p(1-p)\,(2p-1)}{\left(p^2-p+\frac{1}{3}\right)}\left(\frac{1}{3}t^{3/2} - t^{1/2}\right) \leq s < 0, \tag{40}$$

$$0 < u_d \leq t\sigma^2, \tag{41}$$

where $t = \left(\sqrt{\frac{90k-158}{31}} + 1.9\right)$, then the first four moments ($(\mu, \sigma, s, k)^T$) are sufficient statistics for uniquely characterizing a mixture normal distribution, implying that no other mixture normal distributed random variable $\widetilde{X}_1$ exists with different parameters but the same first four moments as X_1.

Proof. From (34) and (35), we obtain

$$\sigma_1^2 = \frac{p^2-1}{3}u_d + \frac{\sigma^3 s}{3p}\frac{1}{\sqrt{u_d}} + \sigma^2, \tag{42}$$

$$\sigma_2^2 = \frac{p(1-2p)}{3}u_d - \frac{\sigma^3 s}{3(1-p)}\frac{1}{\sqrt{u_d}} + \sigma^2. \tag{43}$$

Then substitute (42) and (43) into (36), we get:

$$f(u_d) = 0, \tag{44}$$

where,

$$f(u_d) = \left(-10(p^2-p+\frac{1}{30})^2 + \frac{31}{90}\right)u_d^2 - 2\sigma^2\,(2p-1)^2\,u_d + \frac{2s\sigma^3\cdot(1-2p)}{3p\,(p-1)}u_d^{1/2} + \frac{2s\sigma^5\,(2p-1)}{p\,(p-1)}u_d^{-1/2} + \frac{s^2\sigma^6\left(p^2-p+\frac{1}{3}\right)}{p^2\,(p^2-2p+1)}u_d^{-1} - \sigma^4(k-3). \tag{45}$$

We can then demonstrate that under the following three constraints, (44) has a unique solution:

1) $f(u_d)$ is convex. Examining its second derivative,

$$\begin{aligned} f''(u_d) =& \frac{2\left(p^2-p+\frac{1}{3}\right)}{p^2\left(p^2-2p+1\right)} s^2\sigma^6 x_3^{-3} + \frac{3\left(2p-1\right)}{2p\left(p-1\right)} s\sigma^5 x_3^{-\frac{5}{2}} - \frac{\left(1-2p\right)}{6p\left(p-1\right)} s\sigma^3 x_3^{-\frac{3}{2}} \\ & - 20p^4 + 40p^3 - \frac{64p^2}{3} + \frac{4p}{3} + \frac{2}{3}. \end{aligned}$$

we can find that the first three terms are positive since $p \in (1/2, 1)$ and $s < 0$. By applying the root formula of quartic equations, we can easily verify that

$$-10p^4 + 20p^3 - \frac{32p^2}{3} + \frac{2p}{3} + \frac{1}{3} > 0, p \in [1/2, 1]$$

is equivalent to (39). Therefore, $f''(u_d) > 0$ i.e. $f(u_d)$ is convex.

2) As the coefficient of u_d^{-1} in $f(u_d)$ is strictly positive, we can conclude that $\lim_{u_d \to 0} f(u_d) = +\infty$.

3) $f(t\sigma^2) < 0$. First, we can rewrite the inequality as follows:

$$\begin{aligned} & \left(-10(p^2-p+\frac{1}{30})^2+\frac{31}{90}\right) u_d^2 - 2\sigma^2\left(2p-1\right)^2 u_d - \sigma^4(k-3) < 0 \\ \Leftrightarrow & u_d < \left(\frac{2(2p-1)^2}{-10(p^2-p+\frac{1}{30})^2+\frac{31}{90}} + \frac{(k-3)\sigma^2}{(-10(p^2-p+\frac{1}{30})^2+\frac{31}{90})}\frac{1}{u_d}\right)\sigma^2 \\ \Leftarrow & u_d \leq \left(3.8 + \frac{90}{31}(k-3)\frac{\sigma^2}{u_d}\right)\sigma^2 \\ \Leftarrow & u_d \leq \left(\sqrt{\frac{90k-158}{31}} + 1.9\right)\sigma^2, \end{aligned}$$

where we have used the following inequality:[2]

$$\frac{(2p-1)^2}{-10(p^2-p+\frac{1}{30})^2+\frac{31}{90}} = 1/\left(-\frac{5p^2}{2}+\frac{5p}{2}+\frac{11}{24}-\frac{1}{8\left(2p-1\right)^2}\right) > 1.9$$

which holds for $\frac{1}{2} + \frac{\sqrt{15}\sqrt{13-2\sqrt{31}}}{30} < p < 1$. Second, we have

$$\begin{aligned} & \left[\frac{2s\sigma^3\left(1-2p\right)}{3p\left(p-1\right)} u_d^{1/2} + \frac{2s\sigma^5\cdot\left(2p-1\right)}{p\left(p-1\right)} u_d^{-1/2} + \frac{s^2\sigma^6\left(p^2-p+\frac{1}{3}\right)}{p^2\left(p^2-2p+1\right)} u_d^{-1}\right]_{u_d = t\sigma^2} < 0 \\ \Leftrightarrow & \frac{2\left(1-2p\right)}{p\left(p-1\right)}\left(\frac{1}{3}t^{3/2} - t^{1/2}\right) + \frac{s\left(p^2-p+\frac{1}{3}\right)}{p^2\left(p^2-2p+1\right)} > 0 \end{aligned}$$

[2] This inequality can be verified using the first derivative. However, the exact lower bound is quite complicated, and here we have rounded it to two significant digits for simplicity.

$$\Leftrightarrow s > -\frac{2p(1-p)\,(2p-1)}{\left(p^2-p+\frac{1}{3}\right)}\left(\frac{1}{3}t^{3/2}-t^{1/2}\right).$$

Given constraints 1), 2), and 3), the equation $f(u_d) = 0$ has one and only one root. Using equations (33), (42), and (43), we can conclude that the value of u_d uniquely determines the values of the other parameters. Therefore, we can say that $(\mu, \sigma, s, k)^T$ is a sufficient statistic. □

Theorem 2 establishes that the first four moments are sufficient statistics for characterizing a mixture normal distribution, provided that p satisfies both (39) and (40). The range of values for p satisfying (39) is approximately $(0.6763, 1)$, and conditions (40) and (41) also imply that this range is wide enough to cover parameter values estimated from data. For example, if we consider $p = 0.7, \sigma = 0.2, k = 6$, the two constraints become:

$$\begin{aligned} -2.5457 \leq & s < 0, \\ 0 < & u_d \leq 0.2164, \end{aligned}$$

It is worth mentioning that the calibrated parameters in Table 2 satisfy these three conditions. Therefore, we introduce the following assumption, which guarantees that the first four moments can serve as sufficient statistics for uniquely characterizing a mixture normal distribution:

- **Assumption 1:** The probability parameter and first four moments of X_1, X_2 satisfy conditions (39), (40) and (41).

B.3 Sufficiency of Between-squares Correlation

Various measures of nonlinear dependence have been employed in the literature to capture nonlinear dependencies. For example, Harvey and Siddique (2000) use coskewness as a measure of dependence between an individual asset and a market portfolio. In this section, we demonstrate that, subject to certain regularity conditions, the between-squares correlation is a sufficient measure to capture all nonlinear dependencies.

First, let us examine the frequently used nonlinear dependences in the 3rd and 4th degrees. We can define five possible coskewness and cokurtosis measures between two random variables X_1 and X_2:

$$S(X_1, X_1, X_2) \triangleq \frac{\mathbb{E}[(X_1 - \mathbb{E}X_1)^2(X_2 - \mathbb{E}X_2)]}{\sigma_1^2\sigma_2}, \tag{46}$$

$$S(X_1, X_2, X_2) \triangleq \frac{\mathbb{E}[(X_1 - \mathbb{E}X_1)(X_2 - \mathbb{E}X_2)^2]}{\sigma_1^2\sigma_2}, \tag{47}$$

$$K(X_1,X_1,X_1,X_2) \triangleq \frac{\mathbb{E}[(X_1-\mathbb{E}X_1)^3(X_2-\mathbb{E}X_2)]}{\sigma_1^3\sigma_2}, \tag{48}$$

$$K(X_1,X_1,X_2,X_2) \triangleq \frac{\mathbb{E}[(X_1-\mathbb{E}X_1)^2(X_2-\mathbb{E}X_2)^2]}{\sigma_1^2\sigma_2^2}. \tag{49}$$

$$K(X_1,X_2,X_2,X_2) \triangleq \frac{\mathbb{E}[(X_1-\mathbb{E}X_1)(X_2-\mathbb{E}X_2)^3]}{\sigma_1\sigma_2^3}, \tag{50}$$

where σ_i is the standard deviation of X_i.

Assuming that X_1 and X_2 follow the mixture normal distribution, we can demonstrate that it is sufficient to use between-squares correlation to measure all of the coskewness and cokurtosis mentioned above, subject to a mild condition. As we have established the sufficiency of the first four moments of one mixture normal distribution in Appendix B.2, we will assume that the moments, except for Corr and Corr$^{\text{sq}}$, are fixed in this section.

Using the properties of the bivariate normal distribution,[3] Lemma 1 provides a method for calculating coskewness and cokurtosis for bivariate mixture normal distributions.

Lemma 1. *If $(X_1,X_2)^T$ follows a bivariate mixture normal distribution with parameters p_1,μ_{11},μ_{12}, $\sigma_{11},\sigma_{12},p_2,\mu_{21},\mu_{22},\sigma_{21},\sigma_{22},\rho_{11},\rho_{12},\rho_{21},\rho_{22}$ and μ_i,σ_i,s_i,k_i defined in Appendix B.1, then:*

$$\begin{aligned}
S(X_1,X_1,X_2) =& \frac{2(\mu_{11}-\mu_{12})p_{11}p_{12}}{\sigma_1^2\sigma_2}\left[\sigma_{11}(p_{21}\sigma_{21}\rho_{11}+p_{22}\sigma_{22}\rho_{12})-\sigma_{12}(p_{21}\sigma_{21}\rho_{21}+p_{22}\sigma_{22}\rho_{22})\right],\\
K(X_1,X_1,X_1,X_2) =& \frac{3}{\sigma_1^3\sigma_2}\left[p_{11}\sigma_{11}(\sigma_{11}^2+p_{12}^2(\mu_{11}-\mu_{12})^2)(p_{21}\sigma_{21}\rho_{11}+p_{22}\sigma_{22}\rho_{12})\right.\\
&\left.+p_{21}\sigma_{12}(\sigma_{12}^2+p_{11}^2(\mu_{11}-\mu_{12})^2)(p_{21}\sigma_{21}\rho_{21}+p_{22}\sigma_{22}\rho_{22})\right]\\
K(X_1,X_1,X_2,X_2) =& \sqrt{(k_1-1)(k_2-1)}\text{Corr}^{\text{sq}}(X_1,X_2)+1.
\end{aligned}$$

Proof. Let Z_{11} and Z_{12} be the two normal components of X_1, where $X_1 = I_1Z_{11}+(1-I_1)Z_{12}$ and $I_1 \sim B(1,p_1)$. Similarly, let Z_{21} and Z_{22} be the components of X_2. Using the properties of normal distributions, we have $S(Z_{1i},Z_{1i},Z_{2j})=0$. Therefore, we can rewrite the expression as follows:

$$\begin{aligned}
\mathbb{E}[(X_1-\mathbb{E}X_1)^2(X_2-\mathbb{E}X_2)] &= \sum_{i,j=1,2} p_{1i}p_{2j}\mathbb{E}[(Z_{1i}-\mathbb{E}Z_{1i})+\mathbb{E}Z_{1i}-\mathbb{E}X_1)^2((Z_{2j}-\mathbb{E}Z_{2j})+\mathbb{E}Z_{2j}-\mathbb{E}X_2)]\\
&= \sum_{i,j=1,2} p_{1i}p_{2j}\left[\sigma_{1i}^2\sigma_{2j}S(Z_{1i},Z_{1i},Z_{2j})+2(\mathbb{E}Z_{1i}-\mathbb{E}X)\rho_{ij}\right],
\end{aligned}$$

[3] If $(Z_1,Z_2)^T$ follows a bivariate normal distribution with mean $(\mu_1,\mu_2)^T$ and covariance $\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}$, then:

$$S(Z_1,Z_1,Z_2)=0, K(Z_1,Z_1,Z_1,Z_2)=K(Z_1,Z_2,Z_{21},Z_2)=3\rho, K(Z_1,Z_1,Z_2,Z_2)=1+2\rho^2.$$

$$=2\sum_{i,j=1,2} p_{1i}p_{2j}\rho_{ij}(\mathbb{E}Z_{1i}-\mathbb{E}X).$$

where σ_{1i}^2 and σ_{2j}^2 are the variances of Z_{1i} and Z_{2j}, respectively, and ρ_{ij} is the correlation between Z_{1i} and Z_{2j}. Since $\mathbb{E}Z_{1i}=\mu_{1i}$ and $\mathbb{E}X=p_1\mu_{11}+(1-p_1)\mu_{12}$, we can derive the equation of $S(X_1,X_1,X_2)$. The calculation for cokurtosis is similar, and thus, we will not provide the details here. □

It is evident that the cokurtosis $K(X_1,X_1,X_2,X_2)$ can be calculated using $\mathrm{Corr}^{\mathrm{sq}}(X_1,X_2)$. However, for the remaining cases, we require Lemma 2.

Lemma 2. *The bivariate mixture normal distribution with parameters p_1, μ_{11}, μ_{12}, σ_{11}, σ_{12}, p_2, μ_{21}, μ_{22}, σ_{21}, σ_{22}, ρ_{11}, ρ_{12}, ρ_{21}, ρ_{22} is represented by the random vector $(X_1,X_2)^T$. Here, $\rho_{11}=\rho_{22}=\rho_1$, $\rho_{12}=\rho_{21}=\rho_2$, and we define $\iota_{ij}=p_{1i}p_{2j}\sigma_{1i}\sigma_{2j}$, for $i,j=1,2.$. Using these notations, we can express the following equations:*

$$\rho_2=k\rho_1+b, \tag{51}$$

$$\mathrm{Corr}^{\mathrm{sq}}(\mathrm{X}_1,\mathrm{X}_2)=a_0+a_1\rho_1+a_2\rho_1^2, a_2>0, \tag{52}$$

where

$$\begin{aligned}
k=&-\frac{\iota_{11}+\iota_{22}}{\iota_{12}+\iota_{21}},\\
b=&\frac{\sigma_1\sigma_2\mathrm{Corr}(X_1,X_2)}{\iota_{12}+\iota_{21}},\\
a_2=&\frac{2}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}\left(\frac{\iota_{11}^2}{p_{11}p_{21}}+\frac{k^2\iota_{12}^2}{p_{11}p_{22}}+\frac{k^2\iota_{21}^2}{p_{12}p_{21}}+\frac{\iota_{22}^2}{p_{12}p_{22}}\right),\\
a_1=&\frac{4(\mu_{11}-\mu_{12})(\mu_{21}-\mu_{22})}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}(p_{12}p_{22}\iota_{11}-kp_{12}p_{21}\iota_{12}-kp_{11}p_{22}\iota_{21}+p_{11}p_{21}\iota_{22})\\
&+\frac{4kb}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}\left(\frac{\iota_{12}^2}{p_{11}p_{22}}+\frac{\iota_{21}^2}{p_{12}p_{21}}\right)\\
a_0=&\frac{2}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}\left(\frac{\iota_{12}^2}{p_{11}p_{22}}+\frac{\iota_{21}^2}{p_{12}p_{21}}\right)b^2-\frac{4(\mu_{11}-\mu_{12})(\mu_{21}-\mu_{22})}{\sigma_1^2\sigma_2^2\sqrt{(k_1-1)(k_2-1)}}\left(p_{12}p_{21}\iota_{12}+p_{11}p_{22}\iota_{21}\right)b.
\end{aligned}$$

Proof. We can easily verify Lemma 2 using equations (37) and (38). □

Solving (52), we obtain the following solutions

(53)
$$\rho_1^+ = -\frac{a_1}{2a_2} + \left(\frac{1}{\sqrt{a_2}}\sqrt{\mathrm{Corr}^{\mathrm{sq}}_{\mathrm{xy}} - \frac{4a_0a_2 - a_1^2}{4a_2}}\right), \rho_1^- = -\frac{a_1}{2a_2} - \left(\frac{1}{\sqrt{a_2}}\sqrt{\mathrm{Corr}^{\mathrm{sq}}_{\mathrm{xy}} - \frac{4a_0a_2 - a_1^2}{4a_2}}\right)$$

Assuming the marginal distributions and linear correlation are given, Lemma 2 and (51) imply that $S(X_1, X_1, X_2)$ and $K(X_1, X_1, X_1, X_2)$ are linearly dependent on ρ_1. Therefore, they can be determined uniquely by $\mathrm{Corr}^{\mathrm{sq}}(\mathrm{X}_1, \mathrm{X}_2)$, subject to the following assumption:

- **Assumption 2:** Suppose that $(X_1, X_2)^T$ follows a bivariate mixture normal distribution, where the parameters satisfy $\rho_{11} = \rho_{22} = \rho_1$, $\rho_{12} = \rho_{21} = \rho_2$, and only one of ρ_1^+ and ρ_1^- falls within the interval $[-1, 1]$. In this case, Equation (52) provides a solution for only one root in the interval $[-1, 1]$.

When $\mathrm{Corr}^{\mathrm{sq}}(\mathrm{X}_1, \mathrm{X}_2)$ is positive, the condition in Assumption 2 can be easily satisfied, which is consistent with data estimation. The following theorem establishes that the between-squares correlation is a sufficient measure of nonlinear dependence under mixture normal distributions.

Theorem 3. *Let $(X_1, X_2)^T$ follow a bivariate mixture normal distribution. Denote $\mu_i, \sigma_i, s_i, k_i$ as the mean, standard deviation, skewness and kurtosis of X_i. Assuming Assumptions 1 and 2 hold, the vector $(\mu_1, \sigma_1, s_1, k_1, \mu_2,\ \sigma_2, s_2, k_2, \mathrm{Corr}(X_1, X_2), \mathrm{Corr}^{\mathrm{sq}}(X_1, X_2))^T$ constutues sufficient statistics that can uniquely determine the joint distribution of $(X_1, X_2)^T$.*

Proof. Based on Theorem 2 and assuming Assumption 1 holds, we can conclude that the vector $(\mu_i, \sigma_i, s_i, k_i)^T$ is sufficient for X_i and can be used to determine the underlying parameters $(\mu_{i1}, \mu_{i2}, \sigma_{i1}, \sigma_{i2})^T$. Additionally, Assumption 2 implies that Equations (51) and (52) have only one solution. Therefore, this vector $(\mu_1, \sigma_1, s_1, k_1, \mu_2, \sigma_2, s_2, k_2, \mathrm{Corr}(X_1, X_2), \mathrm{Corr}^{\mathrm{sq}}(X_1, X_2))^T$ constitutes sufficient statistics to uniquely determine ρ_1, ρ_2. $\square$

Theorem 3 establishes that, under Assumptions 1 and 2, the between-squares correlation can act as a sufficient measure of all types of nonlinear dependence between two random variables that follow a bivariate mixture normal distribution.

Appendix C A First Model

In this section, we first show that the household's one-period utility optimization problem can be simplified to a one-period model that incorporates higher-order moments of portfolio returns. We then explore the higher-order moments of the portfolio returns in this simplified model.

C.1 A First Model as an Approximation to the Full model

As per the assumptions outlined in Section 2.3, W_0 represents the initial wealth, $W_1 = R^p W_0$ represents the wealth after one period, and $w_1 = \mathbb{E}[W_1]$ represents the expected wealth. Let $u(\cdot)$ be the utility function for a representative household. The portfolio optimization problem based on utility maximization can be expressed as follows:

$$
\begin{aligned}
&\max_{\alpha} \mathbb{E}[u(W_1)] \\
&\text{s.t.} \quad W_1 = W_0 R^p = W_0(1 + R_f + \alpha R^E + R^L)\,.
\end{aligned} \tag{54}
$$

where R_f represents the risk-free rate, R^E represents the excess return of the risky asset, R^L represents the income shock, and α represents the allocation to the risky asset.

We can approximate $u(W_1)$ in terms of its Taylor series expansion around w_1. Specifically, we have:

$$
u(W_1) = u(w_1) + \sum_{n=1}^{4} \frac{1}{n!} u^{(n)}(w_1)(W_1 - w_1)^n + O[(W_1 - w_1)^5]\,.
$$

where $u^{(n)}(w_1)$ denotes the n-th derivative of $u(w_1)$, and $O[(W_1 - w_1)^5]$ represents the error term of the approximation.

Taking the expectation of both sides and keeping up to the fourth order terms, we obtain:

$$
\begin{aligned}
\mathbb{E}[u(W_1)] \approx & \frac{u(w_1)}{w_1}\mathbb{E}[W_1] + \frac{1}{2}u^{(2)}(w_1)\mathrm{Var}(W_1) + \frac{1}{6}u^{(3)}(w_1)\mathrm{Var}^{1.5}(W_1)\mathrm{Skew}(W_1) \\
& + \frac{1}{24}u^{(4)}(w_1)\mathrm{Var}^2(W_1)\mathrm{Kurt}(W_1) \\
\approx & \frac{u(w_1)}{w_1} W_0 \Big(\mathbb{E}(R^p) - \lambda_1 \mathrm{Var}(R^p) - \lambda_2 \mathrm{Skew}(R^p) - \lambda_3 \mathrm{Kurt}(R^p) \Big)\,,
\end{aligned}
$$

where $\mathrm{Var}(R^p)$, $\mathrm{Skew}(R^p)$, and $\mathrm{Kurt}(R^p)$ denote the portfolio's variance, skewness, and kurtosis, respectively, and the risk weights $\lambda_1, \lambda_2, \lambda_3$ satisfy

$$
\lambda_1 = -\frac{1}{2}\frac{u^{(2)}(w_1)w_1}{u(w_1)}W_0, \lambda_2 = -\frac{1}{6}\frac{u^{(3)}(w_1)w_1}{u(w_1)}Var^{1.5}(R^p)W_0^3, \lambda_3 = -\frac{1}{24}\frac{u^{(4)}(w_1)w_1}{u(w_1)}Var^2(R^p)W_0^4\,.
$$

The risk weights $\lambda_1, \lambda_2, \lambda_3$ can be estimated using the household's utility function and its conjectured wealth expectation and variance. For instance, if we denote the predicted waelth mean and portfolio return variance as $\widehat{w}_1$ and $\widehat{\mathrm{Var}}(R^p)$, respectively, we can approximate the objective function of (54) as follows:

$$
\mathbb{E}[u(W_1)] \approx \frac{u(\widehat{w}_1)}{\widehat{w}_1} W_0 \Big(\mathbb{E}(R^p) - \lambda_1 \mathrm{Var}(R^p) - \lambda_2 \mathrm{Skew}(R^p) - \lambda_3 \mathrm{Kurt}(R^p) \Big)\,, \tag{55}
$$

where the risk weights can be approximated as follows

$$\lambda_1 \approx -\frac{1}{2}\frac{u^{(2)}(\widehat{w}_1)\widehat{w}_1}{u(\widehat{w}_1)}W_0, \lambda_2 \approx -\frac{1}{6}\frac{u^{(3)}(\widehat{w}_1)\widehat{w}_1}{u(\widehat{w}_1)}\widehat{\mathrm{Var}}(R^p)^{1.5}W_0^3, \lambda_3 \approx -\frac{1}{24}\frac{u^{(4)}(\widehat{w}_1)w_1}{u(\widehat{w}_1)}\widehat{\mathrm{Var}}(R^p)^2 W_0^4 .$$

Hence, we can estimate the one-period utility optimization problem using Equation (7), and the risk weights λ_1, λ_2, and λ_3 are related to the utility of the household.

C.2 Higher Order Moments of the Portfolio Return

Using Lemma 1, we can calculate the skewness and kurtosis of the portfolio return in the first model, as presented in Theorem 1. In the following, we provide a detailed proof.

As shown in Appendix B.2, assuming that $(R^E, R^L)^T$ follows a bivariate mixture normal distribution and satisfies Assumptions 1 and 2 with dependence parameters ρ_1 and ρ_2, we can express $S(R^E, R^E, R^L)$ and $K(R^E, R^E, R^E, R^L)$ as linear functions of ρ_1, given marginal distributions and linear correlation. Thus we can represent them as

$$\begin{aligned}
S(R^E, R^E, R^L) =& a_{s,1}\rho_1 + b_{s,1}, \\
S(R^E, R^L, R^L) =& a_{s,2}\rho_1 + b_{s,2}, \\
K(R^E, R^E, R^E, R^L) =& a_{k,1}\rho_1 + b_{s,1}, \\
K(R^E, R^L, R^L, R^L) =& a_{k,2}\rho_1 + b_{s,2}.
\end{aligned}$$

The coefficients a's and b's can be calculated using Lemma 1 and 2.

Combining the above equations with Equation (53), we can rewrite the portfolio's (defined in Theorem 1) skewness and kurtosis as follows:

$$\text{Skew}_\Pi = \text{sgn}(\rho_1 + \frac{a_1}{2a_2})k_s\sqrt{\text{Corr}_{\text{EL}}^{\text{sq}} - \frac{4a_0a_2 - a_1^2}{4a_2}} + b_s, \tag{56}$$

$$\text{Kurt}_\Pi = a_k\text{Corr}_{\text{EL}}^{\text{sq}} + \text{sgn}(\rho_1 + \frac{a_1}{2a_2})b_k\sqrt{\text{Corr}_{\text{EL}}^{\text{sq}} - \frac{4a_0a_2 - a_1^2}{4a_2}} + c_k, \tag{57}$$

where $\text{sgn}(\cdot)$ represents the sign function and

$$\begin{aligned}
k_s &= \frac{3}{\sqrt{a_2}\sigma_\Pi^3}\left[\alpha^2\sigma_E^2\sigma_L a_{s,1} + \alpha\sigma_E\sigma_L^2 a_{s,2}\right], \\
b_s &= \frac{1}{\sigma_\Pi^3}\left[\alpha^3\sigma_E^3 S_E + \alpha^2\sigma_E^2\sigma_L b_{s,1} + 3\alpha\sigma_E\sigma_L^2 b_{s,2} + 3\sigma_L^3 S_L\right], \\
a_k &= \frac{6}{\sigma_\Pi^4}\alpha^2\sigma_E^2\sigma_L^2\sqrt{(K_E - 1)(K_L - 1)},
\end{aligned}$$

$$
\begin{aligned}
b_k &= \frac{4}{\sqrt{a_2}\sigma_\Pi^4}\left[\alpha^3\sigma_E^3\sigma_L a_{k,1} + \alpha\sigma_E\sigma_L^3 a_{k,2}\right], \\
c_k &= \frac{1}{\sigma_\Pi^4}\left[\alpha^4\sigma_E^4 K_E + 4\alpha^3\sigma_E^3\sigma_L b_{k,1} + 4\alpha\sigma_E\sigma_L^3 b_{k,2} + \sigma_L^4 K_L\right].
\end{aligned}
$$

We can use Equations (56) and (57) to solve the portfolio optimization problem (54) with the approximation form (55). This reveals that the nonlinear impact of the between-squares correlation is transmitted through the skewness channel and kurtosis channel.

Appendix D Numerical Methods

D.1 Numerical Solutions to the Household Optimization

The model can be solved numerically using a backward induction method. In this approach, the value function for each period depends on a continuous variable, the normalized cash on hand x_t, which needs to be discretized since we use a grid search to optimize the value function. The terminal condition in the last period is determined by the bequest motive, and the value function corresponds to the bequest function.

To solve the model, we iterate backward for each period t prior to T and for each point in the state space. We use a grid search to compute the value associated with each consumption and risky share grid, and select the optimal grids that achieve the maximum value as the policy rules. This iterative procedure is repeated for each period t until we obtain the optimal policy rules for all periods.

In order to approximate the expected value in the backward induction process, we use the Gauss-Hermite method for a mixture normal distribution. Specifically, we consider the expectation $\mathbb{E}[h(\gamma_{t+1}, \eta_{t+1}, \epsilon_{t+1})]$ with respect to $(\gamma_{t+1}, \epsilon_{t+1}, \eta_{t+1})$, where $\gamma_{t+1} = I_1 x_{11} + (1 - I_1)x_{12}$, $\eta_{t+1} = I_2 x_{21} + (1 - I_2)x_{22}$, and $\epsilon_{t+1} \sim N(0, \sigma_\epsilon^2)$.

To apply the Gauss-Hermite method, we can use it on the four normal distributions embedded in the mixture normal. Specifically, we obtain:

$$
\mathbb{E}[h] = \sum_{i,j=1,2} p_{1i}p_{2j}\mathbb{E}[h(x_{1i}, x_{2j}, \epsilon_{t+1})],
$$

where $p_{n1} = 1 - p_{n2} = \mathbb{P}[I_n = 1], n = 1, 2$. For each component of the summation, the vector $(x_{1i}, x_{2j}, \epsilon_{t+1})^T$ is normally distributed, and we can apply the Gauss-Hermite method to approximate the expected value.

D.2 Calibration

In this section, we present the following explicit form of moments that can be utilized in the GMM method discussed in Section 4.1.

Based on Equations (11) and (12), we can obtain

$$\delta_{it} = \sum_{s=0}^{t} \lambda^{t-s} u_{is} + \epsilon_{it}.$$

We then proceed to compute the moments:

$$\begin{aligned}
\mathrm{Corr}(\eta_{t+1}, \delta_{it+1} - \delta_{it}) =& \mathrm{Cor}(\eta_{t+1}, u_{it+1}) \frac{\sigma_u}{\sqrt{\sigma_u^2 + 2\sigma_\epsilon^2}} = \frac{\sigma_u}{\sqrt{\sigma_u^2 + 2\sigma_\epsilon^2}} \mathrm{Cor}(\eta_{t+1}, u_{t+1}), \\
\mathrm{Corr}^{\mathrm{bs}}(\eta_{t+1}, \delta_{it+1} - \delta_{it}) =& \mathrm{Cor}^{\mathrm{bs}}(\eta_{t+1}, u_{it+1}) \frac{\sqrt{\mathrm{Var}(u_{it+1}^2)}}{\sqrt{\mathrm{Var}((u_{it+1} + \epsilon_{it+1} - \epsilon_{it})^2)}} \\
=& \frac{\sqrt{\mathrm{kurt}_u - 1}\sigma_u^2}{\sqrt{(\mathrm{kurt}_u - 1)\sigma_u^4 + 8\sigma_\epsilon^4 + 8\sigma_u^2\sigma_\epsilon^2}} \mathrm{Cor}^{\mathrm{bs}}(\eta_{t+1}, u_{t+1}) \\
\mathrm{Var}[\delta_{i,t}] =& \frac{1 - \lambda^{2(t-s+1)}}{1 - \lambda^2} \sigma_u^2 + \sigma_\epsilon^2 \\
\mathrm{Skew}[\delta_{i,t}] =& \frac{\sigma_u^3}{(\mathrm{Var}[\delta_{i,t}])^{3/2}} \frac{1 - \lambda^{3(t-s+1)}}{1 - \lambda^3} \mathrm{Skew}_u, \\
\mathrm{Kurt}[\delta_{i,t}] =& \frac{1}{(\mathrm{Var}[\delta_{i,t}])^2} \left[3\sigma_\epsilon^4 + 6 \frac{1 - \lambda^{2(t-s+1)}}{1 - \lambda^2} \sigma_\epsilon^2 \sigma_u^2 + \right. \\
& \left. 3 \left(\left(\frac{1 - \lambda^{2(t-s+1)}}{1 - \lambda^2} \right)^2 - \frac{1 - \lambda^{4(t-s+1)}}{1 - \lambda^4} \right) \sigma_u^4 + \frac{1 - \lambda^{4(t-s+1)}}{1 - \lambda^4} \sigma_u^4 \mathrm{Kurt}_u \right],
\end{aligned}$$

where $\mathrm{Cor}(\eta_{t+1}, u_{t+1}), \mathrm{Cor}^{\mathrm{bs}}(\eta_{t+1}, u_{t+1}), \sigma_u, \mathrm{Skew}_u$ and Kurt_u can be computed using the results presented in Appendix B.1.

Appendix E Nonlinear Effect of the Between-squares Correlation

This section demonstrates that households facing a nonzero between-squares correlation between labor income and stock markets may choose to delay their investment and opt for less risky shares. Our numerical simulations confirm that the nonlinear impact of between-squares correlation on household policy functions remains consistent in our full model, aligning with the findings of the first model presented in Section 2.3.

Figure 10 illustrates the policy functions of households at ages 20, 40, and 60 in our full model, with different levels of between-squares correlation, using the parameter values of the first model in Section 2.3. Figure 10 panel A displays the participation wealth threshold, which indicates that households are more inclined to invest in the stock market when the between-squares correlation approaches a turning point near zero.

In addition, panel B demonstrates that households that participate in the stock market are more willing to hold a larger proportion of risky assets when the between-squares correlation approaches the turning point near zero. However, when the between-squares correlation deviates from the turning point, households tend to reduce their allocation to risky assets to decrease their portfolio risk.

These findings suggest that between-squares correlation has a nonlinear effect on household's portfolio decisions, as shown in the first model. This effect occurs through two channels: skewness and kurtosis. Negative correlation increases skewness, leading to delayed participation and reduced risky asset exposure, while positive correlation increases kurtosis, resulting in similar decisions. It is worth noting that The effects of the two channels are asymmetric, with the kurtosis channel having a greater impact. In conclusion, the nonlinear relationship between labor income and stock returns, as observed in the presence of between-squares correlation, supports a new interpretation of the "stock-like" theory (i.e., Benzoni et al. (2007)), where households' labor income becomes more similar to stocks. This finding highlights the importance of accounting for the nonlinear dynamics of asset returns and their influence on portfolio decisions.

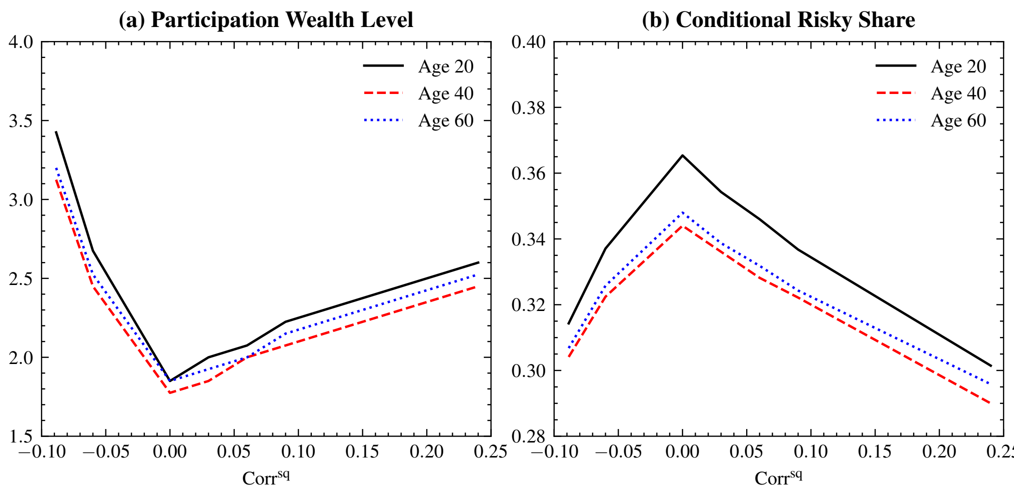
Appendix F Gini Coefficient Evolution in the USA

Figure 11 displays the evolution trend of the Gini coefficient in the United States from 1970 to 2017, using data from The World Inequality Database.

To analyze the evolution of income inequality over time, we have chosen 1970 as the benchmark year and calculated the relative changes in the Gini coefficient over the 48-year period. Our findings indicate that, except for the years surrounding the economic crises in 1987 and 2007, the changes in the Gini coefficient compared to 1969 remained within the range of $[-5\%, 5\%]$.

Figure 10: Changes in Portfolio Decisions

This figure shows the participation rate and conditional risky share generated by the policy functions with between-squares correlation ranging from -0.1 to 0.25. We set the parameters for labor income risk as $\mu^L = 0, \sigma^L = 0.25, \text{Skew}^L = -0.5, \text{Kurt}^L = 3.0$, and for stock return as $\mu^S = 0, \sigma^S = 0.2, \text{Skew}^L = -1, \text{Kurt}^L = 4.5$, with a linear correlation of 0.1 between them. The preference parameters are set as $\gamma = 4$ (relative risk aversion), $\phi = 0.4$ (EIS), $\beta = 0.9$ (discount rate), $b = 2.5$ (bequest motive), $F = 0.01$ (fixed cost rate). We assume the normalized persistent income shock as $w_{i,t} = 1$.



Appendix G Sensitivity Analysis of Adjusted Persistent Income Shock

In this paper, we assumed an adjusted persistent income shock as 1 when analyzing policy functions. However, different persistent income shocks can significantly impact households' portfolio decisions. To explore this effect, we present the policy functions for various adjusted persistent income shocks in Figure 12. The optimal risky share pattern is similar across different $w_{i,t}$, but the level differs. Generally, larger $w_{i,t}$ values indicate more aggressive portfolio decisions as households expect higher future labor income. Wealthier and older households are less sensitive to $w_{i,t}$ as labor income becomes more trivial for them. Specifically, wealthier households have accumulated a significant amount of wealth that can provide a buffer against income shocks and they may prioritize maintaining their level of cash on hand or other financial goals rather than adjusting their portfolio decisions based on changes in labor income. Similarly, older households may have already accumulated sufficient savings for retirement or have other sources of income, such as pensions or social security. Additionally, older households may be more concerned with preserving their wealth and managing risks associated with potential events such as death shocks, rather than making significant changes to their portfolio decisions based on changes in labor income.

Figure 11: Gini Coefficient Evolution in the US

This figure shows the Gini coefficient evolution trend in the US from 1970 to 2017, using data from The World Inequality Database. We have chosen 1962 as the benchmark year and calculated the relative changes in the Gini coefficient.

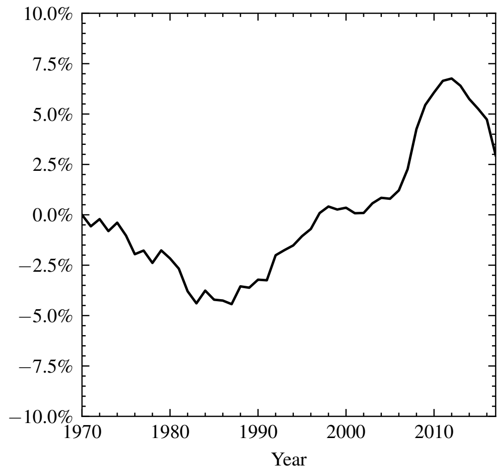


Figure 12: Sensitivity Analysis of adjusted persistent income shock

The figure presents six panels displaying the policy functions for optimal risky asset share across different age and education groups. The first column corresponds to the households without college degrees, and the second column corresponds to the households with college degrees. The policy functions are generated using calibrated parameters, with the adjusted persistent income shock $w_{i,t}$ set to $0.95, 1, 1.05$ for each panel.