

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

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

April 12, 2023

Abstract

This paper reveals the nonlinear dependence, measured by the between-squares correlation, between stock returns and earning risk exists and shows that incorporating the between-squares correlation in a life-cycle model can explain the households' portfolio choice puzzles. Empirical studies support model predictions that households with higher between-squares correlations are less likely to participate in the stock market and lower their risky asset holdings conditional on their participation. Our model suggests a novel “stock-like” theory that the nonlinear dependence could drive labor income to be “stock-like” via both skewness and kurtosis channels. Moreover, we show 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, Life-cycle Portfolio Choice and Consumption, Wealth Inequality, Equivalent Wealth Loss.

*We thank Stephan Siegel, Pengfei Sui (discussant), and seminar participants at the 2023 AEA poster session, SFS Cavalcade Asia-Pacific 2022, Bachelier World Congress 2022, and the PKU-NUS Annual International Conference in Quantitative Finance and Economics 2022.

[†]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 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 consistently predict all households should participate in the stock market and 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 These papers rely 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

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

²For 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.

³Besides 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)).

⁴See 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).

⁵Campbell 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.

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.

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. For example, monetary policy can have an immediate impact on the stock market but takes time to produce an effect in the labor market. Therefore, we can observe a significant positive jump in the stock market, while the labor market is still suffering. Moreover, assuming two variables follow mixed normal distributions under some regularity conditions, we show that between-squares correlation is a sufficient measure to capture all other nonlinear dependence.

Our study initially examines a simple one-period portfolio optimization problem that involves higher moments. We demonstrate that the between-squares correlations between stock returns and labor income can significantly impact both the skewness and kurtosis of portfolio return. We emphasize that the existence of nonlinear dependence highlights a novel version of the “stock-like” theory, wherein the nonlinear dependence between labor income and stock returns renders households' labor income more “stock-like” (i.e., Benzoni et al., 2007). Importantly, our research reveals that the “stock-like” nature of labor income is primarily driven by the skewness channel when the between-squares correlation is negative and by the kurtosis channel when the between-squares correlation is positive. This leads to 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. The main innovations are twofold. First, we introduce the nonlinear dependence between labor income and stock returns. Using the Panel Study of Income Dynamics Survey (PSID), we find that between-squares correlations are significant, whereas correlations are close to zero. 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 find heterogeneity in our results. On average, college graduates

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

Second, we allow higher moments in both labor income shocks and stock returns with the mixed 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 (Güvener 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. We first categorize households into two groups: workers without college degrees versus workers with college degrees. Then we estimate the dynamics of labor income, annual stock returns, and between-squares correlations for each group. 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.

In addition, we find that households with between-squares correlations hold less risky asset shares conditional on their participation in the stock market compared with those without between-squares correlations. Moreover, the wealth threshold of participation is mildly U-shaped with respect to age. The main drivers behind the U-shaped pattern of the participation threshold are the hump-shaped labor income process and the fact that young households seek to hold equity more aggressively than older households. We also observe between-squares correlations increase the wealth threshold across the ages, an increase that is higher for college graduates.

Our results hold for different levels of the household's risk aversion coefficient, such as $\gamma = 5$ and $\gamma = 3$. 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 do not accumulate significant wealth and live paycheck to paycheck. Hence, they start holding risky assets with higher normalized cash on hand.

We also find empirical evidence to support 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 households with higher between-squares correlations are less likely to participate in the stock market. Second, households do adjust their portfolio holdings of risky assets when facing higher between-squares correlations, a finding that is consistent with the model's prediction. A one-standard-deviation increase in between-squares correlations is associated with a 2.43% to 3.16% decline in equity shares. Moreover, in line with the model's predictions, college graduates' portfolio decisions 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 12.47% when the household is young and their wealth accumulation is low. On the other hand, via a simulation study, we show that ignoring between-squares correlation could increase the Gini index by more than 3% for different levels of the bequest ratio and lower households' median wealth by around 10%. 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 correlation helps to match the data well. However, the empirical evidence on this 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 covariance 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 between-squares correlation, and documents stylized facts regarding it. 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 provides a portfolio perspective of between-squares correlation, and Section 8 concludes.

2 Between-Squares Correlation

In this paper, we introduce the between-squares correlation (henceforth referred to as the BS-Corr) to capture the nonlinear dependence between stock market returns and labor income shocks. The BS-Corr is defined as the Pearson correlation between the de-meaned squares of two time series:

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

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

2.1 Properties of BS-Corr

The BS-Corr can be viewed as the normalized cokurtosis,⁶ 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 BS-Corr will tend to undergo both extremely positive and extremely negative deviations. Specifically, under a normality assumption, the BS-Corr is fully determined by the linear correlation, that is, $\text{Corr}^{\text{bs}}(Z_1, Z_2) = \text{Corr}^2(Z_1, Z_2)$ when both Z_1, Z_2 are normal.

However, when data are subject to risks from higher-order moments (i.e., Shen (2022)), the BS-Corr 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 mixed normal distribution, and thus both labor income and stock return are subject to higher-order moments and the BS-Corr between them exists.

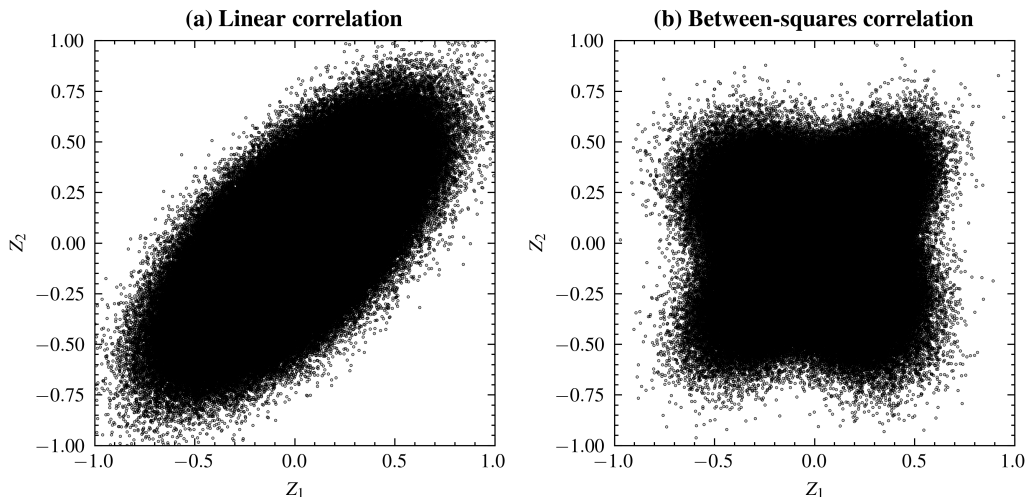
To illustrate the differences between the BS-Corr and the correlation, Figure 1 presents scatter plots for two random variables under both normal and mixed normal assumptions. In panel B, the two random variables are subject to mixed normal distribution such that higher-order moments and nonlinear dependence exist. The scatter plot in this panel shows that, when the BS-Corr is positive, two types of comovements are possible: two variables may move in the same *and* opposite directions. In contrast, as shown in panel A, when both variables satisfy normal distribution, a positive linear correlation implies that they only tend to move in the same direction. Apparently, our nonlinear dependence measure is able to simultaneously capture both comovements between the labor market and stock market as shown in empirical studies.⁷

⁶Cokurtosis of two random variables Z_1 and Z_2 is defined by $K(Z_1, Z_2) = \frac{\mathbb{E}[(Z_1 - \mathbb{E}[Z_1])^2(Z_2 - \mathbb{E}[Z_2])^2]}{\sigma_{Z_1}^2 \sigma_{Z_2}^2}$ and connects to the BS-corr by $K(Z_1, Z_2) = \sqrt{(\text{Kurt}[Z_1] - 1)(\text{Kurt}[Z_2] - 1)} \text{Corr}^{\text{bs}}(Z_1, Z_2)$.

⁷Michelacci and Quadrini (2009), Bronars and Famulari (2001), and Nickell and Wadhvani (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

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 $\mathbb{E}[Z_1^4]/\text{Var}^2[Z_1] = \mathbb{E}[Z_2^4]/\text{Var}^2[Z_2] = 3$. 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 mixed normal distribution with a linear correlation of 0.1, and the BS-Corr of 0.3.



On the other hand, compared with cointegration, introducing BS-Corr requires fewer assumption on the evolution of stock price and labor income. To establish cointegration, some linear combination of the two series shall be stationary, which is a strong constraint. In addition, it is difficult to test the existence of cointegration, as empirical studies only provide limited and mixed evidence (see Benzoni et al. (2007)). In contrast, the BS-Corr can be considered a correlation of higher moments and requires no further theoretical constraints.

More importantly, we find that the BS-Corr could be a sufficient statistic for nonlinear dependence under some proper assumptions. In Appendix B, we show that when two variables follow mixed normal distribution and satisfy some regularity conditions, the BS-Corr, together with linear correlation and other marginal moments, is sufficient to characterize the joint distribution of these two variables. In other words, this suggests that, given linear correlation and marginal moments, any other nonlinear dependence can be solely captured by the BS-Corr.

2.2 Stylized Facts of BS-Corr

Next, we estimate the BS-Corr between labor income and stock return using a large sample of household data. We use the time series of labor income shock from the Panel Study of Income Dynamics Survey (PSID) and excess market returns from Ken French's data library to estimate

value and labor income. Nickell and Wadhvani (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.

the BS-Corr.⁸ PSID was conducted annually from 1968 to 1997, but has been conducted biennially since 1997. Therefore, to make the best use of data after 1997, we consider two time series: annual data from 1968 to 1997 and biennial data from 1968 to 2017. For the biennial series, data after 1997 are directly taken from the PSID, and data between 1970 and 1997 are constructed from annual data. To get a more efficient estimator, when computing the correlations, we only consider households with a minimum of 20 waves of valid labor income data.⁹

Figure 2: BS-Corr Distribution

This graph plots the histogram and kernel density of BS-Corr estimations between labor income shock and excess stock return. Labor income and stock return data are taken from the PSID and Ken French’s data library. The annual data are from 1970 to 1997, and the biennial data are from 1970 to 2017. The sample is restricted to households with a minimum of 20 waves of valid labor income data.

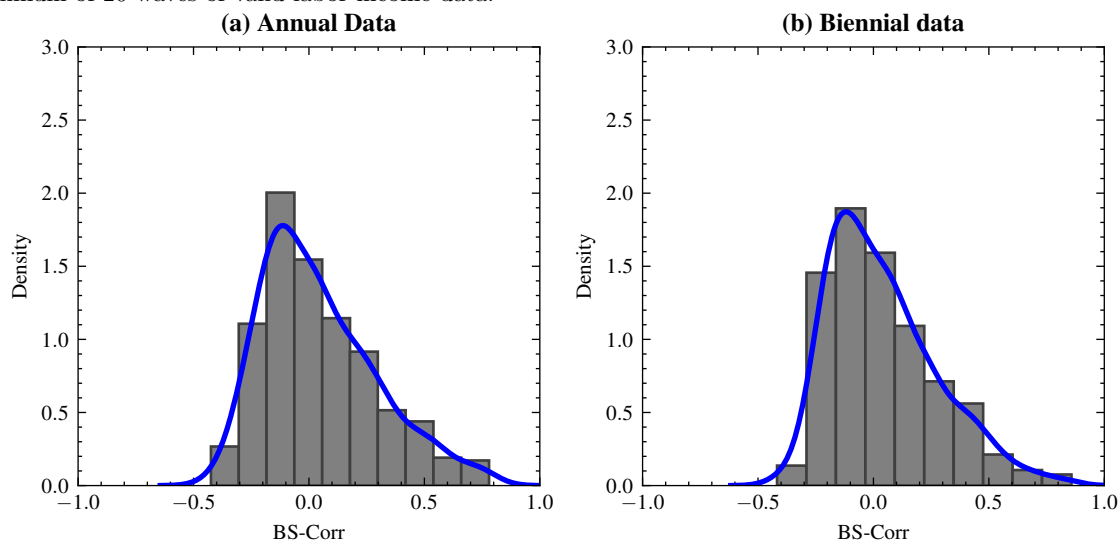


Figure 2 displays the histogram and kernel density of estimations over the full sample with different frequencies. The figure clearly reveals the high density of a positive BS-Corr. The long right tail also shows that many households earn income that is highly exposed to the stock market in the sense of the BS-Corr. Although few samples have a negative BS-Corr, we notice that these negative estimations are concentrated close to zero, meaning we can reasonably assume a positive BS-Corr in the model setup. Moreover, this assumption is also consistent with Table 1. Table 1 presents summary statistics for the BS-Corr. Results from data with different frequencies both support a significant and positive level of the BS-Corr. Panel A shows that, for the full sample, the average of the BS-Corr is both positive (0.044 and 0.041) and significantly nonzero with p -values

⁸See <https://simba.isr.umich.edu/data/data.aspx> and mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. See equation (8) for the strict definition of labor income shock $\delta_{i,t}$. Here, we estimate BS-Corr between excess stock return and the d -order difference of $\delta_{i,t}$ ($d = 1$ for annual data, $d = 2$ for biennial data). According to Ken French’s data library, the excess return on the market is defined as value-weight stock return minus the Treasury-bill rate.

⁹We also apply the data select principle from Nakajima and Smirnyagin (2019) to the PSID data. See Appendix A for details.

Table 1: Summary Statistics

This table reports the mean, different percentages, and some fractions over different samples, based on the data from the PSID and Ken French’s data library. Panel A shows the results over the full sample. To observe the BS-Corr levels of different educational groups, we split the sample into two groups: households without college degrees and households with college degrees. Panels B and C show the group results. The column “Annual data” (Biennial data) means the statistics below are based on an annual (biennial) series in [1970, 1997] ([1970, 2017]). “Fraction > k” represents the fraction of samples in which the BS-Corr is more than k (expressed as a percentage).

	<i>Panel A: Full sample</i>		<i>Panel B: No college</i>		<i>Panel C: College</i>	
	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
<i>p</i> -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%
No. of obs	435	517	306	348	129	169

< 0.2%. Although the medians are close to zero, substantial variability remains in the values of the BS-Corr. Panel A also provides the interquartile range of the BS-Corr for both data frequencies over the full sample. We find that the 75th and 90th percentiles are substantially positive at 0.2 and 0.4, respectively. Further, we calculate the fraction of estimations with a positive BS-Corr or a substantially positive BS-Corr. We find that more than half of the estimations are positive, and about one-third feature a substantially positive value of the BS-corr no less than 0.1. More remarkably, over one-fifth of the estimations exceed 0.2. Thus, we assume a positive BS-Corr and use the sample mean as our calibration target when we set up the model.

To control for education, we split all households in two groups: college graduates and a no-college group. This is according to a well-established fact that households exhibit different behavior patterns across education groups (see Attanasio (1995); Hubbard et al. (1995); Cocco et al. (2005)). Panels B and C in Table 1 report the statistics for two education groups. The results are qualitatively similar to those for the full sample. However, almost all statistics for the no-college group are somewhat lower than those for college graduates. Especially, the average of the BS-Corr for college graduates is about twice the value for the no-college group. In Section 4, we separate the

calibrations for both groups and study the difference between corresponding policy rules.

2.3 A One-period Model: Skewness Channel and Kurtosis Channel

To understand the effect of the BS-Corr on portfolio allocation, we first provide a parsimonious one-period portfolio optimization problem with higher moments (i.e., Harvey et al. (2010)). We will show that the BS-Corr has a nonlinear effect on portfolio allocation, and it affects households' portfolio allocation via two channels: 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 and consumes a constant value during this period. A household with initial wealth W_0 invests α share of wealth in stock and $1 - \alpha$ share in the bond. Meanwhile, the household faces labor income risk, captured by $W_0 R^L$, where R^L is the income shock. Then the household's wealth in the next period W_1 is given by

$$(2) \quad W_1 = W_0(1 + \alpha R^S + (1 - \alpha)R_f + R^L) = W_0(1 + R_f + \alpha R^E + R^L),$$

where $R^E = R^S - R_f$ denotes the stock's excess return.¹⁰ Thus, the household's one-period utility depends entirely on the portfolio return R^p :

$$(3) \quad R^p \triangleq 1 + R_f + \alpha R^E + R^L.$$

We find that the household's utility maximization problem can be approximated by the following portfolio optimization problem with higher moments (i.e., Harvey et al. 2010):

$$(4) \quad \max_{\alpha} \mathbb{E}(R^p) - \lambda_1 \text{Var}(R^p) - \lambda_2 \text{Skew}(R^p) - \lambda_3 \text{Kurt}(R^p),$$

where $\lambda_1, \lambda_2, \lambda_3$ are risk weights that depend on the household's utility. Appendix C.1 provides details on the approximation of this first model.

In line with our life-cycle model in Section 3, we further assume that the excess return, R^E , and labor income risk, R^L , both satisfy the mixed normal assumptions defined in Appendix B¹¹ and the nonlinear dependence between them exists. For a given risky asset share, α , the change in the BS-Corr only affects the skewness and kurtosis of portfolio returns, which suggests the BS-

¹⁰This portfolio return has a constant unit of labor income return. It importantly captures that labor income is nontradable and households can only adjust their risky asset holding to achieve a target portfolio return.

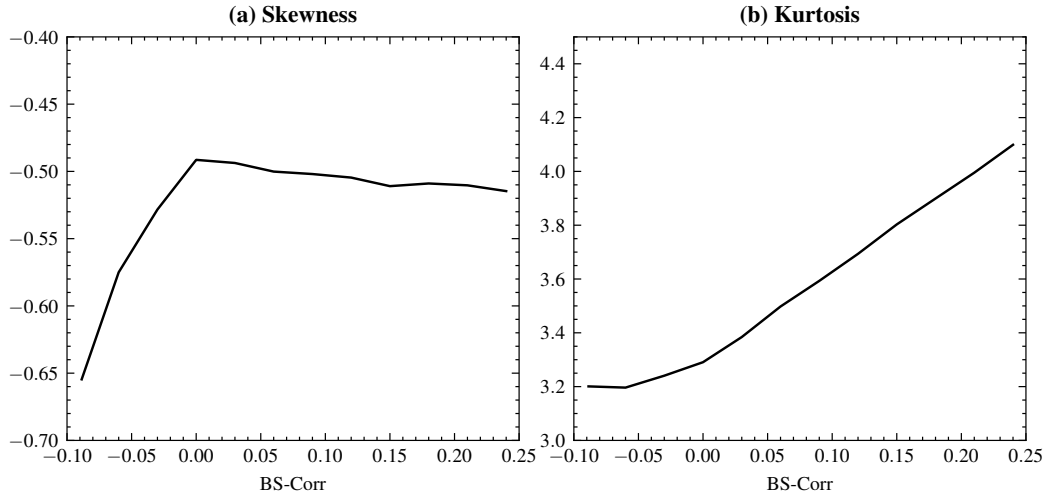
¹¹For simplicity, we also assume $\rho_{11} = \rho_{22}$, $\rho_{12} = \rho_{21}$, where $\{\rho_{ij}\}_{i=1,2,j=1,2}$ are correlation parameters (see Appendix B Definition 1).

Corr could affect the household’s portfolio risk profile through skewness and kurtosis channels. In following, we provide a numerical example to show the effect of the BS-Corr on optimal portfolio choices.

In Figure 3, we change the BS-Corr from -0.1 to 0.25 for a fixed $\alpha = 1$ and report the corresponding skewness and kurtosis of the portfolio returns. We find that the BS-Corr has quite a different effect on portfolio skewness and kurtosis. When the BS-Corr decreases from 0 to -0.10 , the skewness decreases substantially from -0.50 to -0.65 , while kurtosis stays stable. On the other hand, when the BS-Corr increases from 0 to 0.25 , the kurtosis increases significantly from 3.25 to 4.10 , while skewness stays stable with a very mild decline. This observation reveals that the BS-Corr affects portfolio risk mainly through the skewness channel when BS-Corr is negative and the kurtosis channel is stronger when BS-Corr is positive. To mitigate whole portfolio risk, given a nontradable labor income, a household will optimally lower its risky share to maximize the risk-adjusted objective in (4).

Figure 3: Changes in Skewness and Kurtosis

These two figures plot the variation of portfolio higher order moments with changing BS-Corr. We set the risky share $\alpha = 1$ and fix all other moments as: $\mathbb{E}[R^L] = 0$, $\text{Var}(R^L) = 0.25^2$, $\text{Skew}(R^L) = -0.5$, $\text{Kurt}(R^L) = 3.0$ for labor income risk, $\mathbb{E}[R^S] = 0$, $\text{Var}(R^S) = 0.2^2$, $\text{Skew}(R^S) = -1$, $\text{Kurt}(R^S) = 4.5$ for stock return risk and linear correlation of 0.1 between them.



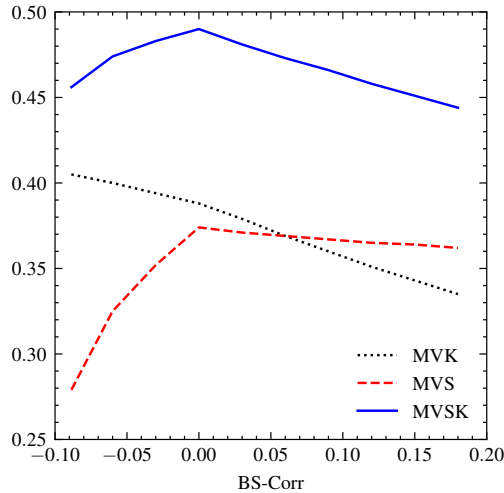
We then report the optimal risky share from this simple model. For simplicity, we denote model (4) by MVSK, which means households consider mean, variance, skewness and kurtosis in portfolio selection. To isolate the effect of the BS-Corr, we further consider two reduced models: the model that ignores kurtosis (MVS, $\lambda_3 = 0$) and the one that ignores skewness (MVK, $\lambda_2 = 0$).

Figure 4 shows the optimal risky share with varying BS-Corr. We find that the BS-Corr has a nonlinear effect on the optimal risky share. When the BS-Corr increases from -0.10 to 0 , the optimal risky share increases. When the BS-Corr increases from 0 to 0.20 , the optimal risky share decreases.

This can be explained with the help of the reduced models, MVK and MVS. When the BS-Corr is negative, the red dashed line (MVS) experiences a more drastic variation compared to the black dotted line (MVK), which suggests that negative BS-Corr has a greater impact on the portfolio's skewness than kurtosis. As a result, the MVSK exhibits a similar pattern to MVS when the BS-Corr is negative, indicating a skewness channel. On the other hand, when the BS-Corr turns positive, the effect on portfolio skewness is minimal since the red dashed line for MVS remains relatively flat. However, the impact of BS-Corr on portfolio kurtosis is significant, as demonstrated by the black dotted line for the MVK model under positive BS-Corr. Therefore, there is a comparable pattern between MVSK and MVK, as the impact of positive BS-Corr on kurtosis dominates the optimal risky share, suggesting a kurtosis channel.

Figure 4: Optimal Risky Share

This graph shows the comparison of optimal risky share with varying BS-Corr. The parameters 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 $\lambda_1, \lambda_2, \lambda_3$ is 0.9473, -0.0054 , 0.0011. To illustrate the effect of the BS-Corr, we compare three models (MVK, MVS, MVSK).



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 age 20 and receive uncertain labor income exogenously until age 65. They retire at age 65. The probability that an adult is alive at age t conditional on being alive at age $t - 1$ is denoted as p_t .¹² 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 of 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

$$(5) \quad U_{i,t} = \left\{ (1 - \beta)C_{i,t}^{1-1/\psi} + \beta \left(E_t \left[p_{t+1}U_{i,t+1}^{1-\gamma} + b(1 - p_{t+1})X_{i,t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}},$$

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

$$(6) \quad R_{t+1}^S = R_f + \mu + \eta_{t+1},$$

where η_{t+1} is the shock to returns and is independently and identically distributed as

$$(7) \quad \eta_t \sim \begin{cases} N(\mu_{\eta,1}, \sigma_{\eta,1}) & \text{with prob. } p_\eta, \\ N(\mu_{\eta,2}, \sigma_{\eta,2}) & \text{with prob. } 1 - p_\eta, \end{cases}$$

Our assumption of a mixed normal distribution is mainly motivated by two reasons. First, as argued before, the mixed normal assumption allows us to deal with a more complicated dependence structure, such as the BS-Corr between stock return and labor income introduced in our model. Under normality, the BS-Corr is always the square of the correlation; however, in Section 2, we show that from the data, the BS-Corr is actually similar to the correlation in magnitude. Therefore, deviation from normality allows us to construct a more realistic BS-Corr. Second, the mixed 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 to be 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

¹² p_t is taken from the mortality tables of the National Center for Health Statistics.

level of negative skewness and excess kurtosis of annual U.S. equity market returns. Therefore, we introduce mixed normal shocks instead of normal shocks to capture such moments.

3.3 Labor Income Process

During its working period, at age t , household i 's labor income $Y_{i,t}$ is given by

$$(8) \quad \log Y_{i,t} = f(t, Z_{i,t}) + \delta_{i,t},$$

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. We further decompose the labor income shock $\delta_{i,t}$ into a persistent shock, $\nu_{i,t}$, and a transient shock, $\epsilon_{i,t}$:

$$(9) \quad \delta_{i,t} = \nu_{i,t} + \epsilon_{i,t}.$$

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

$$(10) \quad \nu_{i,t} = \lambda \nu_{i,t-1} + u_{i,t},$$

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

$$(11) \quad u_{it} = \begin{cases} N(\mu_{u,1}, \sigma_{u,1}^2) & \text{with prob. } p_u, \\ N(\mu_{u,2}, \sigma_{u,2}^2) & \text{with prob. } 1 - p_u. \end{cases}$$

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 BS-Corr. We define the correlations between components of $u_{i,t}$ and $\eta_{i,t}$ as

$$(12) \quad \rho_{a,b} = \text{Corr}(u_{i,t}^{(a)}, \eta_t^{(b)}), a = 1, 2, b = 1, 2.$$

The four correlations $\rho_{a,b}$ control for the value of the correlation and BS-Corr between $u_{i,t}$ and $\eta_{i,t}$.¹³

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

¹³See Appendix B for the explicit forms.

follows

$$(13) \quad Y_{i,t} = kP_{iK}, \quad t > K,$$

where it is a constant fraction k 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

$$(14) \quad X_{i,t+1} = (X_{i,t} - C_{i,t})R_{i,t+1}^p - FI_P P_{i,t} + Y_{i,t+1},$$

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

$$(15) \quad R_{i,t+1}^p = \alpha_{i,t}R_{t+1}^S + (1 - \alpha_{i,t})R_f.$$

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

$$(16) \quad 0 \leq \alpha_{i,t} \leq 1,$$

$$(17) \quad 0 < C_{i,t} \leq X_{i,t}.$$

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 $(I_{i,t}^p, C_{i,t}, \alpha_{i,t})$ based on cash-on-hand $X_{i,t}$ to maximize the expected

utility. The optimization problem can be stated as

$$(18) \quad V_{i,t} = \max_{\{I_{i,u}^p\}_{u=t}^T, \{\alpha_{i,u}\}_{u=t}^T, \{C_{i,u}\}_{u=t}^T} \mathbb{E}_t(U_{i,t}),$$

where $U_{i,t}$ is defined in equation (5) and is subject to the constraints given by equations (6) to (17).

Since analytical solutions do not exist for this problem, we use a numerical solution method. We standardize the entire problem with permanent labor income $P_{i,t}$ defined as $\ln P_{i,t} = f(t, Z_{i,t}) + \nu_{i,t}$, and introduce a new state, $w_{i,t}$, defined as $\ln w_{i,t} = (1 - \lambda)\nu_{i,t}$. 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

$$(19) \quad v_{i,t}(x_{i,t}, w_{i,t}) = \max_{\substack{\{\alpha_{i,u}\}_{u=t}^T \\ \{c_{i,u}\}_{u=t}^T}} \left\{ \begin{array}{l} (1 - \beta)c_{i,t}^{1-\frac{1}{\psi}} \\ + \beta \left(\mathbb{E}_t \left(\left(\frac{P_{i,t+1}}{P_{i,t}} \right)^{1-\gamma} \left[p_{t+1}v_{i,t+1}^{1-\gamma} + (1 - p_{t+1})bx_{i,t+1}^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right) \end{array} \right\}^{\frac{1}{1-\frac{1}{\psi}}},$$

subject to

$$(20) \quad x_{i,t+1} = \begin{cases} (x_{i,t} - c_{i,t})r_{i,t+1}^p \frac{P_{i,t}}{P_{i,t+1}} - FIP_{i,t} + e^{\varepsilon_{i,t+1}} & \text{for } t \leq K, \\ (x_{i,t} - c_{i,t})r_{i,t+1}^p \frac{P_{i,t}}{P_{i,t+1}} - FIP_{i,t} + k & \text{for } t > K. \end{cases}$$

$$(21) \quad \ln w_{i,t+1} = \begin{cases} \lambda \ln w_{i,t} + (1 - \lambda)u_{i,t+1} & \text{for } t \leq K, \\ 0. & \text{for } t > K. \end{cases}$$

$$(22) \quad \frac{P_{i,t}}{P_{i,t+1}} = \begin{cases} w_{i,t}e^{f(t, Z_{i,t}) - f(t+1, Z_{i,t+1}) - u_{i,t+1}} & \text{for } t \leq K, \\ 1 & \text{for } t > K. \end{cases}$$

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 A presents the details of the numerical solution method.

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 Ken French’s data library, 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 (8). Following the method of Cocco et al. (2005), we use the PSID from 1970 to 2017 and run a regression analysis to estimate $f(t, Z_{i,t})$.¹⁴ 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 Ken French’s data library between 1970 and 2017, for which the de-measured 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 conduct 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 BS-Corr between $\delta_{i,t} - \delta_{i,t-1}$ and η_t .¹⁵ 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 $\{25, 35, 45, 55\}$, and each group contains ages ± 5 years.¹⁶ Lastly, for the dependence structure between two markets, we use the average BS-Corr 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 $\rho_{11} = \rho_{22}$ 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 BS-Corr and calibrate the parameters. Overall, we apply GMM with 19 moments to

¹⁴See Table 9 for variable definitions and Appendix A for the data selection principle and other details.

¹⁵We use third- and fourth-order moments in GMM. For simplicity, we refer to such higher-order moments as skewness and kurtosis.

¹⁶Group $h = 25$ contains ages 22-35; group $h = 35$ contains ages 35-45; group $h = 45$ encompasses ages 45-55; and group $h = 55$ aggregates the remaining ages 55-65.

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.¹⁷ The SMM seeks the parameters that minimize

$$(23) \quad (\hat{m} - m)'W(\hat{m} - m),$$

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 reports the parameters of shocks. There are 17.1% chances 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

¹⁷We 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].

stock market becomes more volatile.¹⁸ Meanwhile, shocks to labor income also display a structure similar to the stock return innovations. For college graduates, the probability of the mixed normal distribution (p_ν) is 0.279. This implies college graduates are less likely to expect a negative mean growth (-15.6%) compared with 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 do 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) imply the linear correlation and BS-Corr are 0.0457 , 0.0692 for the college group, and 0.0382 , 0.0327 for households without college degrees. The results of the correlation are consistent with those of Cocco et al. (2005), who show that a correlation is not significantly different from 0. The BS-Corr is in the same order of magnitude as the linear correlation, but it is big enough to generate an obvious difference in policy function, as we will discuss in Section 5.

Panel D in Table 2 shows the calibration results of preference parameters and the fixed cost rate, and Figure 5 shows the fitness. 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 (over 10 in Fagereng et al. (2017), around 6 in Shen (2022) and 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 , which is within the range of existing empirical evidence and calibrations. Households without college degrees, compared with college graduates, 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 BS-Corr increases the labor income risk, and thus labor 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.

¹⁸Such 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 Ken French's data library and construct a two-step calibration. Panel A reports the parameters for the mixed normal shock, which controls the stock return process. Panel B reports the parameters for persistent shock $\nu_{i,t}$ (mixed 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 cost for different groups. We assume the two groups have same risk aversion and fixed cost.

<i>Panel A: Stock return shock, η_t</i>			
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	
<i>Panel B: Labor income shocks $\nu_{i,t}$ and $\epsilon_{i,t}$</i>			
		No college	College
Mixture weight of $\nu_{i,t}$	p_ν	0.271	0.279
Mean of 1st component of $\nu_{i,t}$	$\mu_{u,1}$	-0.124	-0.156
Mean of 2nd component of $\nu_{i,t}$	$\mu_{u,2}$	0.046	0.060
SD of 1st component of $\nu_{i,t}$	$\sigma_{u,1}$	0.172	0.231
SD of 2nd component of $\nu_{i,t}$	$\sigma_{u,2}$	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
<i>Panel C: Dependence parameters</i>			
		No college	College
Dependence parameter 1	ρ_1	0.832	0.769
Dependence parameter 2	ρ_2	-0.166	-0.222
<i>Panel D: Preferences and fixed cost</i>			
		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 BS-Corr. Moments from the model are computed with calibrated parameters under a mixed normal distribution. Empirical moments are from the PSID and Ken French's data library.

<i>Panel A: Moments of stock return shock</i>					
		Model	Data		
SD		0.216	0.216		
3rd moment		-1.412	-1.410		
4th moment		8.595	8.620		
<i>Panel B: Moments of labor income shock</i>					
		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
<i>Panel C: Dependence structure</i>					
		No college		College	
		Model	Data	Model	Data
Linear corr		0.0382	0.0377	0.0457	0.0456
BS-Corr		0.0327	0.0338	0.0692	0.0696

Figure 5: Calibration Results

The three graphs show the comparison of moments from the baseline model, the model without BS-Corr, and SCF data. The parameters for the baseline model have been taken from Table 2. For the main model without BS-Corr, we adjust only the two correlation parameters to decrease BS-Corr 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.

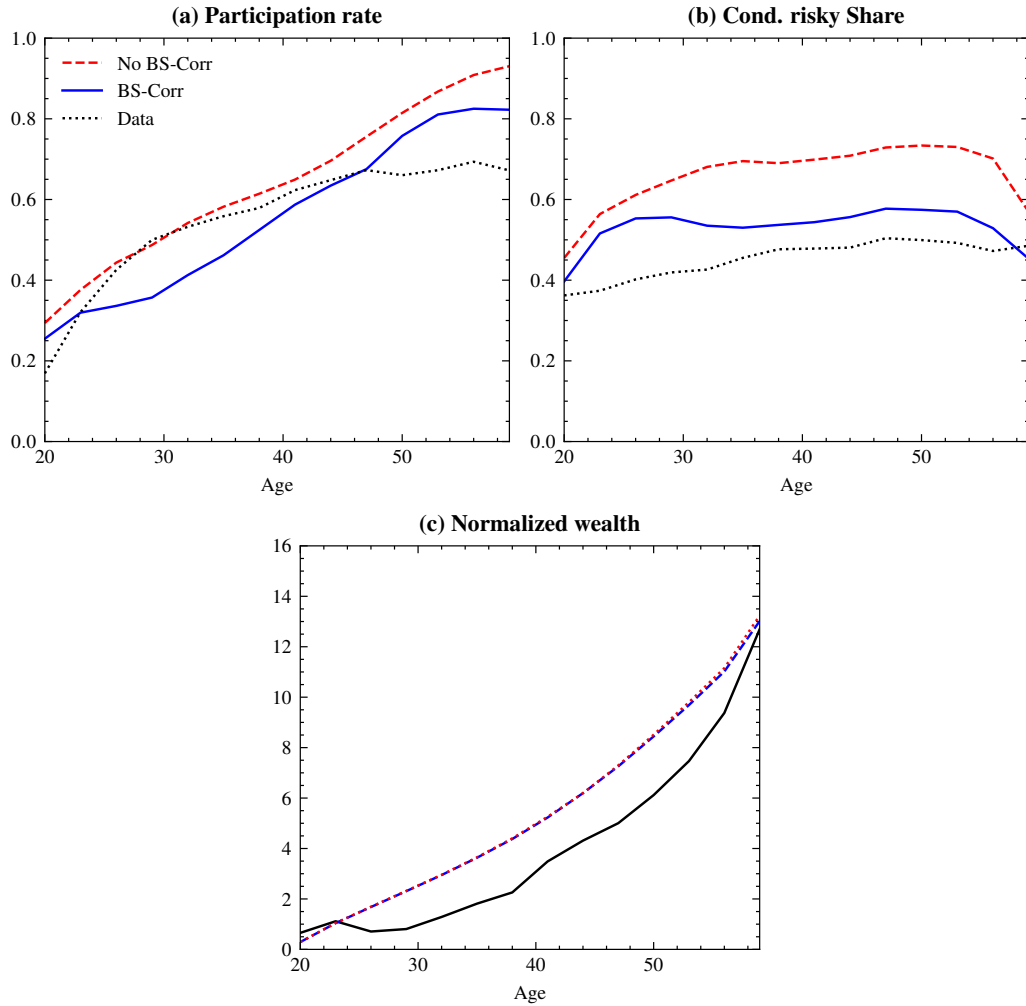


Table 3 reports the fitness of the calibration and shows that our model matches the data well. We obtain relative errors less than 20% for all moments except one, and over 50% of errors are less than 5%.¹⁹

To better understand the effect of the BS-Corr, we also simulate age profiles in a model without the BS-Corr for 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 BS-Corr. Specifically, the BS-Corr is manually adjusted from 0.069 (0.033 for the no-college group) to 0. Figure 5 shows the results. Without the BS-Corr, 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 BS-Corr fails to fit the data with moderate risk aversion and a low fixed cost, although the BS-Corr only changes slightly.

5 Quantitative Analysis

The BS-Corr, 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 BS-Corr, we present the benchmark model results with calibrated parameters and compare them with those generated by the model without the BS-Corr. We keep correlations the same for both models.

5.1 Policy Function

Figure 6 plots the portfolio rules at different ages for both the no-college and college groups. Conditional on participation, introducing a small BS-Corr 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

¹⁹The exception is the third moment of age group [25, 35]. The relative errors are 26.47% for the no-college group and 25% for college graduates.

on stock returns will make labor income more stock-like when considering the BS-Corr. Although the decreasing pattern remains, the model with the BS-Corr 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 BS-Corr. Table 4 presents the results. We find the difference is much more significant for households with little wealth. For example, among all age-20 households, the largest decrease is 39%; however, for the wealthiest age-20 households, this difference can be as small as 4%. 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 BS-Corr 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 39% at age 20 to 34% at age 60.

Table 4: Change in Optimal Risky Asset Shares with the BS-Corr

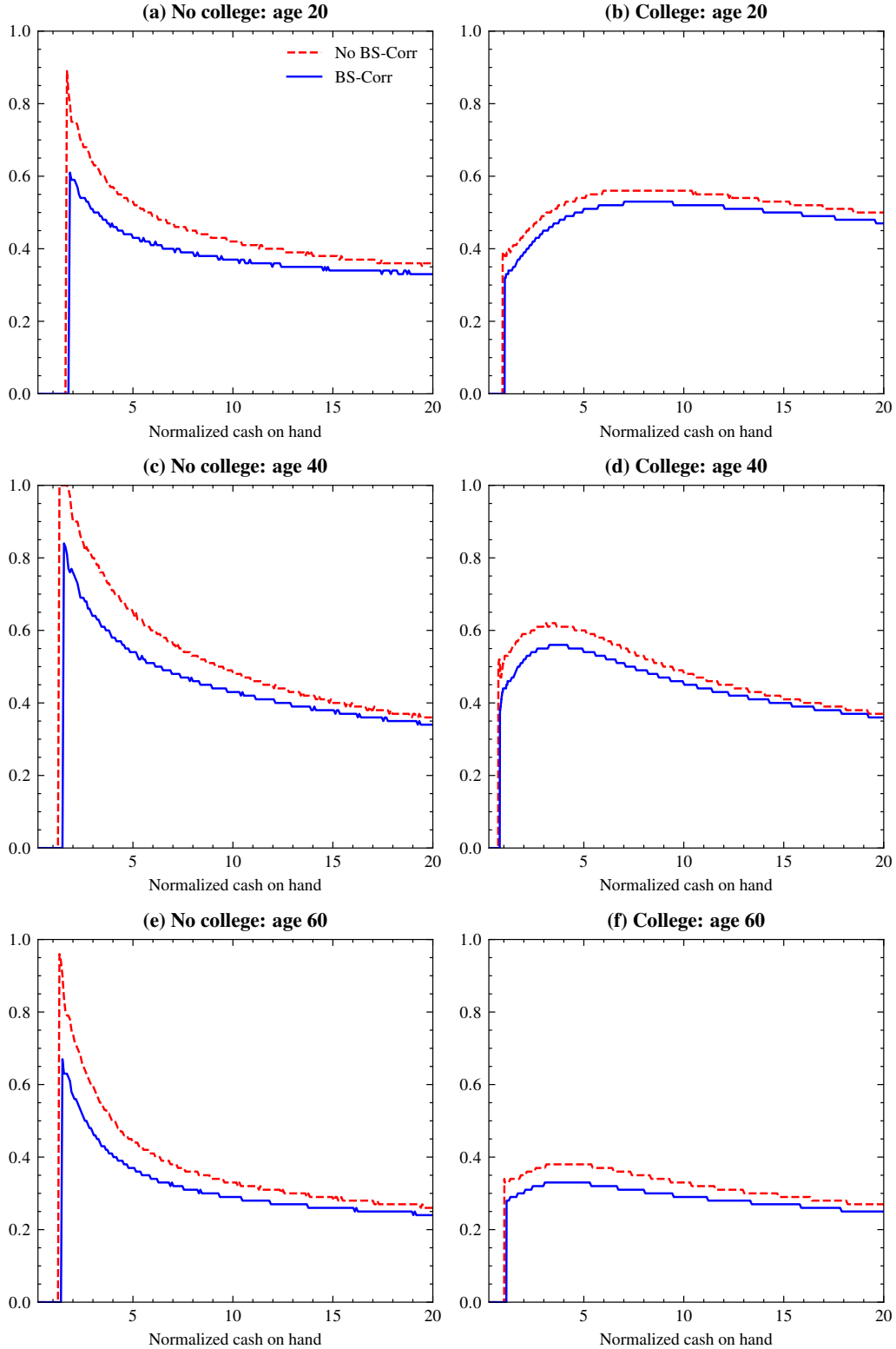
This table reports the difference in optimal risky share α_t between models with and without the BS-Corr. The data are from the policy functions discussed in Section 5.1. For simplicity, we report results for the college group only. We calculate the difference for each age and wealth ratio. For instance, the second percentage (−18%) corresponding with the age of 20 is calculated as the optimal share (67%) at a wealth ratio of 3.0 from model with the BS-Corr (−93%) from the model without the BS-Corr. We also report the largest difference for each age.

Age	Largest diff.	Wealth ratio			
		1.0	3.0	5.0	10
20	−39%	−39%	−4%	−3%	−4%
30	−51%	−7%	−6%	−5%	−4%
40	−52%	−8%	−7%	−6%	−3%
50	−43%	−7%	−6%	−5%	−2%
60	−34%	−5%	−4%	−5%	−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 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

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 college group and the second corresponds with the no-college group. The policy functions are solved with calibrated parameters. We take the permanent labor growth state $w_{i,t}$ as 1.

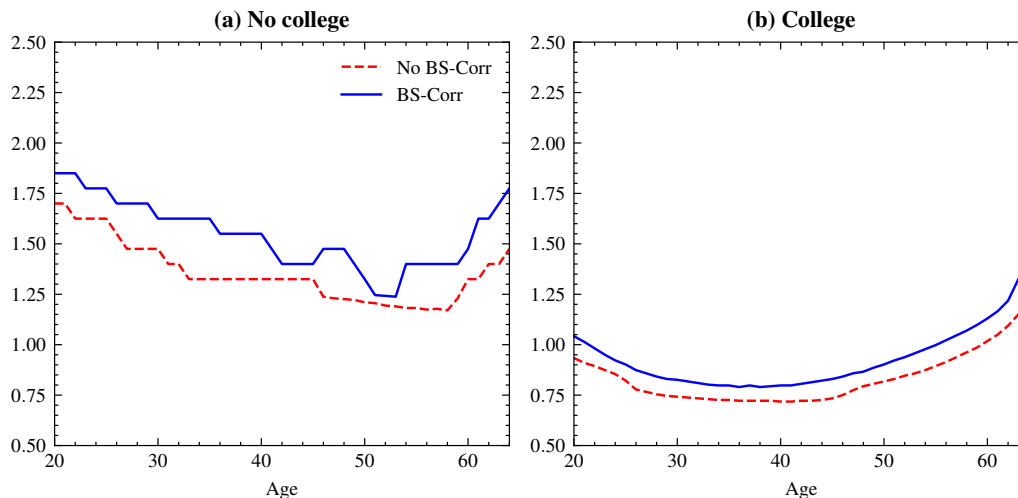


their wealth. Introducing BS-Corr makes labor income more stock-like, thus affecting households' participation decisions through the second channel.

Figure 7 presents the wealth-participation threshold with respect to age. For both groups, the inclusion of the BS-Corr 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 BS-Corr is included, but the effect is much stronger for college graduates.

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 permanent labor growth state $w_{i,t}$ as 1.



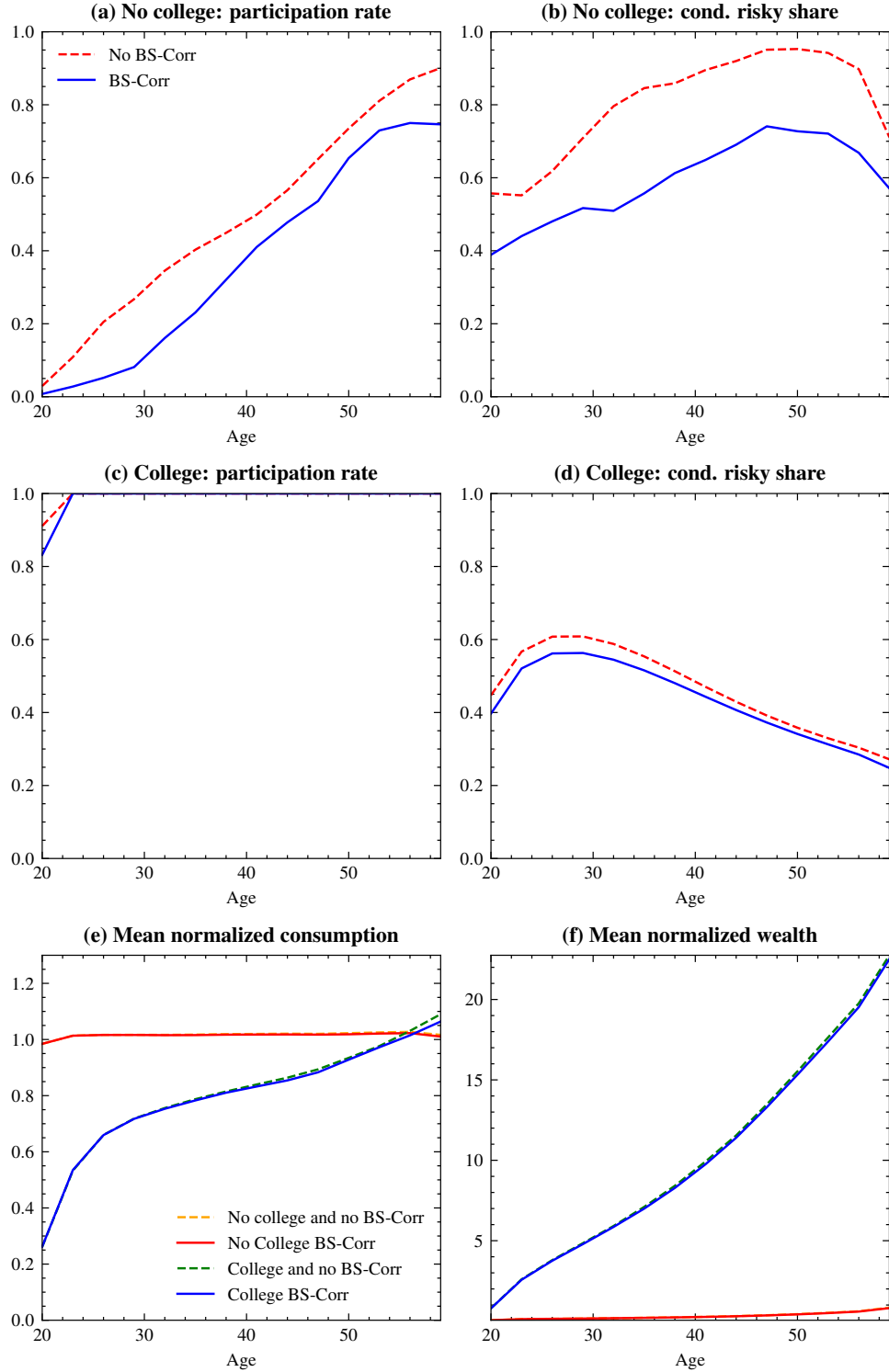
5.2 Simulation Results

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 both college graduates and the no-college group.

First, we find that the participation rate increases with age no matter whether or not the BS-

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 BS-Corr. The profiles are calculated from a simulation of 100000 households.



Corr 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 BS-Corr 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 BS-Corr 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 D). The existence of the BS-Corr between stock returns innovations and labor income shocks indicates that large joint downward movements are more likely, thus making labor income more uncertain and undermining 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.

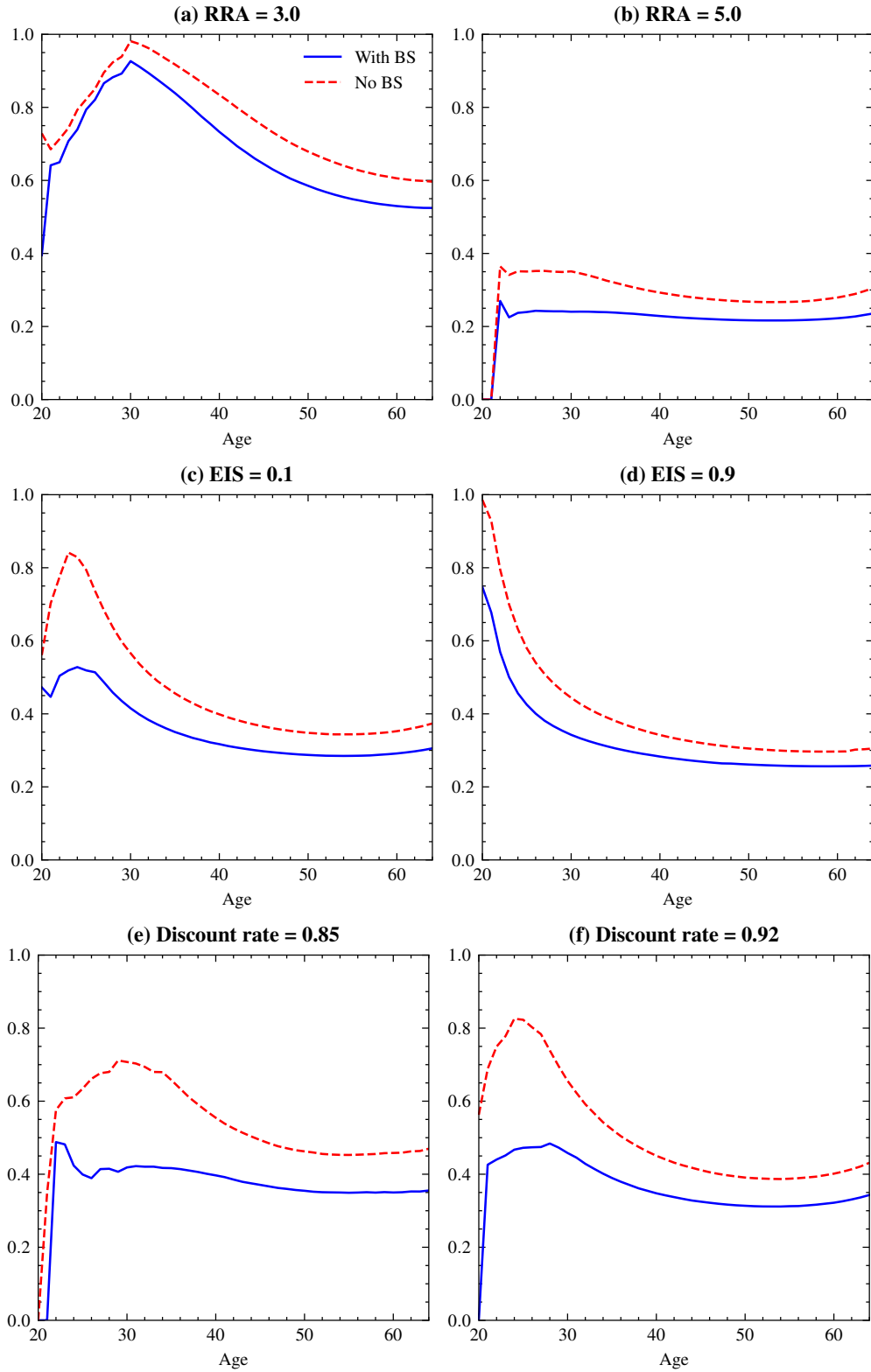
5.3 Sensitivity Analysis

To test the robustness of the effect of the BS-Corr, 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.

Figure 9 plots the life-cycle profile of risky asset shares with different parameters. We find that the BS-Corr has a very robust effect on portfolio choices. Although the change of parameters may

Figure 9: Sensitivity Analysis

The three 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.



reshape the risky asset shares curve a lot, the effect of the BS-Corr always exists and stays strong. Among the different experiments, only when risk aversion is quite low, is the effect of the BS-Corr relatively small. For example, panel a shows that the effect of the BS-Corr is not as strong as that in panels b and c. 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 BS-Corr 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 BS-Corr.

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 BS-Corr. To do this, we use data from the PSID survey.

The PSID was an annual survey from 1968 to 1997 and became a biennial survey after 1997. We only use the PSID from 1997 to 2019, as before 1997, PSID did not report the value of stockholdings.²⁰ 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 BS-Corr between labor income risk and stock returns risk for each household, we use annual income level and the Center for Research in Security Prices (CRSP) index. Besides, we observe a nonlinear relationship between the BS-Corr and households' portfolio decisions from the one-period model in Section 2.3 and Appendix E. More specifically, the BS-Corr has an almost opposite relationship with portfolio decisions at the tipping point of zero. Therefore, we use the absolute value of the BS-Corr, instead of the BS-Corr, to better capture this nonlinear relationship.

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

²⁰Before 1997, only the 1984 and 1989 waves report the data of stock holdings.

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

$$(24) \quad I_{it} = \begin{cases} 1 & \text{if household participates, or} \\ 0 & \text{if household does not participate,} \end{cases}$$

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 BS-Corr, which is inspired by its demonstrated nonlinear impact on portfolio decisions, as discussed in the first model in section 2.3. We investigate whether a link exists between the absolute value of BS-Corr 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:

$$(25) \quad \mathbb{P}(\alpha_i > 0) = \Phi(\beta_0 + \beta_1|BS - Corr|_i + \beta_2Corr_i + \beta_3 \ln(y_i) + \beta_4 age_i + \beta_5 age_i^2 + \beta_6 w_i + \beta_7 s_i + \beta_8 sk_i + \beta_9 k_i + D_m I_M + \sum D_i^{(year)} I_i^{(year)}),$$

$$(26) \quad \alpha_i^* = \beta_0 + \beta_1|BS - Corr|_i + \beta_2Corr_i + \beta_3 \ln(y_i) + \beta_4 age_i + \beta_5 age_i^2 + \beta_6 w_i + \beta_7 s_i + \beta_8 sk_i + \beta_9 k_i + D_m I_M + \sum D_j^{(year)} I_{ij}^{(year)},$$

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

$$(27) \quad \alpha_i = \begin{cases} 0, & \alpha_i^* \leq 0 \\ \alpha_i^*, & 0 < \alpha_i^* < 1 \\ 1, & \alpha_i^* \geq 1. \end{cases}$$

Table 5 shows the marginal effects from probit market participation regressions. Our results indicate that the absolute value of the BS-Corr is significantly negatively correlated with households' participation decisions. In particular, we find that when the BS-Corr is low in absolute value, the market participation propensity decreases. This result still holds under different model speci-

cations. Moreover, the coefficients of the BS-Corr remain statistically significant for both college

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)
$ BS-Corr $	-0.541*** (0.056)	-0.318*** (0.059)	-0.251*** (0.068)	-0.576*** (0.074)	-0.387*** (0.077)	-0.246*** (0.089)	-0.423*** (0.091)	-0.255*** (0.094)	-0.247*** (0.107)
$Corr$		-0.020 (0.036)	-0.016 (0.037)		-0.060 (0.048)	-0.043 (0.049)		-0.008 (0.055)	-0.027 (0.057)
$\ln(y)$		0.926*** (0.033)	0.660*** (0.040)		0.852*** (0.047)	0.543*** (0.055)		0.687*** (0.051)	0.522*** (0.066)
$Age/10$			-0.282** (0.140)			-0.062 (0.182)			-0.351 (0.227)
$Age^2/100$			0.027 (0.017)			0.006 (0.022)			0.039 (0.027)
$Marriage$			-0.336*** (0.051)			-0.154*** (0.069)			-0.324*** (0.080)
$FWealth$			0.244*** (0.009)			0.253*** (0.014)			0.167*** (0.011)
$Std(d\delta)$			-0.015 (0.097)			-0.037 (0.129)			0.0422 (0.152)
$Skew(d\delta)$			-0.022 (0.022)			-0.016 (0.030)			-0.025 (0.035)
$Kurt(d\delta)$			-0.025* (0.014)			-0.030 (0.018)			-0.035* (0.021)
N	7,954			4,794			3,160		

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

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 BS-Corr 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 BS-Corr. 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 BS-Corr. 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 BS-Corr (-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

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)
$ BS-Corr $	-0.318*** (0.031)	-0.191*** (0.030)	-0.110*** (0.033)	-0.370*** (0.047)	-0.246*** (0.046)	-0.131*** (0.049)	-0.221*** (0.040)	-0.151*** (0.039)	-0.094** (0.042)
$Corr$		-0.015 (0.018)	-0.008 (0.018)		-0.036 (0.029)	-0.032 (0.027)		-0.012 (0.023)	-0.005 (0.023)
$\ln(y)$		0.470*** (0.017)	0.456*** (0.019)		0.498*** (0.028)	0.325*** (0.030)		0.286*** (0.021)	0.282*** (0.025)
$Age/10$			-0.514 (0.675)			0.466 (1.001)			0.314 (0.885)
$Age^2/100$			1.309 (0.796)			-0.231 (1.186)			0.332 (1.037)
$Marriage$			-0.186*** (0.025)			-0.089** (0.038)			-0.140*** (0.031)
$FWealth$			0.008*** (0.000)			0.076*** (0.005)			0.004*** (0.001)
$Std(d\delta)$			0.063 (0.047)			-0.020 (0.071)			0.114* (0.060)
$Skew(d\delta)$			-0.012 (0.010)			-0.018 (0.016)			-0.004 (0.013)
$Kurt(d\delta)$			-0.004 (0.006)			-0.004 (0.010)			-0.013* (0.007)
N	7,954			4,794			3,160		

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

find that the empirical evidence supports our model's predictions: for all households, the absolute value of the BS-Corr 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.

7 Microeconomics and Macroeconomics Implications of the BS-Corr

7.1 Cost of Ignoring the BS-Corr

So far, we have shown the significant effect of the BS-Corr on households' participation decisions and portfolio choices, but how does the BS-Corr affect households' welfare? To answer this question,

we consider the equivalent wealth loss from ignoring the BS-Corr between the labor income market and the stock market. We denote v, v^0 as the value functions of models with the BS-Corr and without the BS-Corr, 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 labor growth state w can be defined as

$$(28) \quad v(t, x, w) = v^0(t, x + \Delta x, w).$$

Table 7: Cost of Ignoring the BS-Corr

This table reports the relative equivalent wealth loss, $\Delta x/x$, against different ages and wealth for the college group, with the calibrated parameters in Section 4 and a labor growth rate, w , of 1.

Age	Normalized wealth					
	1	2	4	8	10	15
25	12.16%	9.11%	8.91%	6.16%	5.83%	5.11%
35	9.96%	12.47%	7.92%	4.75%	3.44%	2.91%
45	4.78%	7.81%	2.69%	2.77%	2.06%	1.18%
55	1.58%	2.62%	0.86%	0.22%	0.13%	0.05%

Table 7 reports the relative equivalent wealth loss, $\Delta x/x$, against different ages and wealth for college graduates, with the labor growth rate of w as one.²¹ We find that ignoring the BS-Corr between labor income and stock markets can be costly to households. With the BS-Corr being 0.069, the welfare losses are economically significant: households can reach 12.47% of equivalent wealth loss, which represents a substantial welfare loss. Losses tend to be maximized at age 35 with a relatively low wealth level. Meanwhile, we find that the equivalent wealth loss decreases with age and wealth in general, which is consistent with the diminishing effect of the BS-Corr on optimal portfolio rules and the participation threshold.

7.2 Wealth Inequality

Since we know that the BS-Corr matters for each household and ignoring it reduces household welfare, it is natural to ask how the BS-Corr will affect macroeconomics, more specifically, 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)), and the causes and consequences of widening disparities in wealth have become a defining debate of 2020's.

To answer this question, we simulate 100,000 households for 500 periods and compare the value function and wealth distribution between the cases recognizing and ignoring the BS-Corr. Table

²¹The results for the no-college degree group are similar. To save on space, we don't report them here.

Table 8: Welfare Loss, Wealth Distribution, and Gini Index

This table summarizes the effect of BS-Corr on welfare and wealth inequality, computed by simulating 100,000 households for 500 periods. We take parameters from Section 4.3 for the main model and compare results with the situation ignoring the BS-Corr. Panel A reports the welfare loss by ignoring the BS-Corr over different normalized wealth levels. Panel B depicts the wealth distribution by some wealth percentages in the last 200 periods. Panel C shows the Gini index with different bequest ratios.

<i>Panel A: Welfare loss</i>						
Normalized wealth	[1, 2)	[2, 4)	[6, 8)	[8,10)	[12,14)	[16,18)
Ignoring BS-Corr	1.820	1.744	2.307	2.500	2.721	3.081
Recognizing BS-Corr	1.820	1.751	2.319	2.519	2.775	3.249
Change	-0.01%	-0.37%	-0.49%	-0.76%	-1.92%	-5.17%
<i>Panel B: Wealth distribution</i>						
Wealth percentage	1%	20%	40%	60%	80%	99%
Ignoring BS-Corr	0.866	1.908	2.668	3.558	4.887	12.300
Recognizing BS-Corr	0.865	1.888	2.631	3.495	4.795	11.689
Change	0.12%	0.90%	1.30%	1.73%	2.03%	7.16%
<i>Panel C: Gini index</i>						
Bequest ratio	0	0.2	0.4	0.6	0.8	1
Ignoring BS-Corr	0.316	0.352	0.344	0.321	0.332	0.316
Recognizing BS-Corr	0.306	0.340	0.330	0.310	0.319	0.305
Change	-3.47%	-4.07%	-3.87%	-3.44%	-3.58%	-3.60%

8 shows the results. In panels A and B, we find that ignoring the BS-Corr inefficiently increases households' wealth across all percentiles and reduces their value. The reason is that ignoring the BS-Corr will mislead the household to inefficiently invest more in risky asset, which causes increases in wealth accumulation. Meanwhile, heterogeneity across the poor and rich exists. Richer households hold more positions in risky assets when ignoring the BS-Corr, are more exposed to the risk from the BS-Corr, and face a higher increase in wealth accumulation. As a consequence, the inefficient portfolio allocation from ignoring the BS-Corr causes welfare losses from 0.01% for the poor and more than 5% for the rich.

This is because the composition and leverage of household portfolios differ systematically along wealth distributions. These portfolio differences are persistent over time. We document this stylized fact and expose its consequences for the dynamics of the wealth distribution. A second consequence of portfolio heterogeneity is that asset price movements can create a wedge between the evolution of the income and wealth inequality.

We compute a commonly used measure of wealth inequality, the Gini index, for models ignoring and recognizing the BS-Corr. Table 8, panel C, shows the Gini coefficients for different bequest

ratios.²² We find that ignoring the BS-Corr increases wealth inequality by 3.47% to 4.07%.

Gini coefficients generated by the models are lower than their empirical counterparts because our baseline model has two limitations. First, our baseline model fails to capture the behavior of the richest households, such as Bill Gates, Jeff Bezos, or Warren Buffett. Those on the *Forbes* magazine list of the nation's 400 richest barely rely on their income to accumulate wealth, and their labor income process will be very different. Second, we do not consider the portfolio heterogeneity. For example, the portfolios of rich households are dominated by corporate and noncorporate equities, while the portfolio of a typical middle-class household is highly concentrated in residential real estate and, at the same time, highly leveraged. Because of the failure to capture these two features, our model generates wealth inequality not as large as the data suggest. However, the trend implied by the models is still valid and could be even larger if these two features are considered.

Overall, we find that a channel that has attracted little scrutiny so far has played an important role in the evolution of wealth inequality: the BS-Corr between labor income and stock markets induces large shifts in the wealth distribution and eventually 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 BS-Corr, households are less willing to participate in the market and significantly reduce their stock investments. Moreover, our paper shows that the BS-Corr 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. We further show empirical evidence that is supportive of the model's prediction and provide a portfolio selection perspective to understand households' market participation decisions and risky asset holdings relative to their labor income. In addition, we find that ignoring the BS-Corr reduces households' welfare and increases wealth inequality.

²²The bequest ratio describes how much wealth the new generation can inherit; that is, zero means none and one means complete inheritance.

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.
- 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.
- 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., Liechty, J. C., Liechty, M. W., and Müller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, 10(5):469–485.
- 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.
- 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 Wadhvani, 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.

Appendix

Nonlinear Dependence and Households' Portfolio Decisions over the Life Cycle

Wei Jiang Shize Li Jialu Shen

A Data

A.1 Panel Study of Income Dynamics

The Panel Study of Income Dynamics (PSID) is the longest running longitudinal household survey in the world, collecting data of over 18,000 individuals living in 5,000 families in the United States from 1968 to 2019. PSID interviewed on a broad range of topics, including the family information and financial situation which is required in most life cycle household models. PSID is even more appropriate for our model since we need time series of labor income to estimate its correlations with stock return. However, PSID has some well-know issues for researchers and many papers have discussed the approaches to deal with them²³. In this paper, we roughly follow the sample selection principle of Nakajima and Smirnyagin (2019). First, years prior to 1970 are dropped. Waves of 1968 and 1969 lack some data and are not completely consistent with the waves afterwards. Second, drop SEO and Latino samples. PSID includes about 2000 low income SEO samples and 2000 Latino samples. We drop all those samples since they are colleted with unequal selection probabilities. Third, observations with missing or non-positive head/spousal labor incomes are dropped. Then, the top 1% with respect to head's and wife's labor incomes are trimmed. PSID bracket lots of variables about finance, such as labor income, with an upper boundary. This trim selection is to drop bracketed or extreme samples. Finally, households with income growth anomalies (annual log growth rate must be between 1/20 and 20) are dropped.

We construct variables following Brunnermeier and Nagel (2008). The following table summarize the variable definitions.

A.2 SCF Data

The Survey of Consumer Finances (SCF) is normally a triennial cross-sectional survey of U.S. families. The survey data include information on families' balance sheets, pensions, income, and demographic characteristics. Information is also included from related surveys of pension providers and the earlier such surveys conducted by the Federal Reserve Board. We use data from the 2007 to

²³Such as Cocco et al. (2005), Gomes and Michaelides (2005b) and Nakajima and Smirnyagin (2019).

Table 9: Variable 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

2019, 5 waves in total. Variables are constructed using the code-book and macro-variable definitions from the Federal Reserve website.

We construct variables of SCF following Gomes and Michaelides (2005b). 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). Then, financial wealth is constructed in a same way as variable FIN in the publicly available SCF data set, made up 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). Further, data of financial assets invested in the risky asset is 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.

B Bivariate Mixed Normal Distribution

In this section, we first define the mixed normal distribution and characterize it using the first four moments. We then demonstrate that the BS-Corr can serve as a sufficient measure of nonlinear dependence under specific regularity conditions.

Definition 1. We say a random vector $(X_1, X_2)^T$ follows a bivariate mixed 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:

$$(29) \quad X_1 = I_1 Z_{11} + (1 - I_1) Z_{12},$$

$$(30) \quad X_2 = I_2 Z_{21} + (1 - I_2) Z_{22},$$

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

$$\begin{aligned} \mathbb{E}(Z) &= (\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22})', \\ \text{Cov}(Z) &= \begin{bmatrix} \sigma_{11}^2 & 0 & \rho_{11}\sigma_{11}\sigma_{21} & \rho_{12}\sigma_{11}\sigma_{22} \\ 0 & \sigma_{12}^2 & \rho_{21}\sigma_{12}\sigma_{21} & \rho_{22}\sigma_{12}\sigma_{22} \\ \rho_{11}\sigma_{11}\sigma_{21} & \rho_{21}\sigma_{12}\sigma_{21} & \sigma_{21}^2 & 0 \\ \rho_{12}\sigma_{11}\sigma_{22} & \rho_{22}\sigma_{12}\sigma_{22} & 0 & \sigma_{22}^2 \end{bmatrix}. \end{aligned}$$

In the following part of the appendix, we denote $p_{11} = p_1, p_{12} = 1 - p_1, p_{21} = p_2, p_{22} = 1 - p_2$ for convenience.

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

$$(31) \quad \mu_i = \mathbb{E}[X_i] = p_{i1}\mu_{i1} + p_{i2}\mu_{i2},$$

$$(32) \quad \sigma_i^2 = \text{Var}[X_i] = p_{i1}\sigma_{i1}^2 + p_{i2}\sigma_{i2}^2 + p_{i1}p_{i2}(\mu_{i1} - \mu_{i2})^2,$$

$$(33) \quad s_i = \text{Skew}[X_i] = \sigma_i^{-3} p_{i1}p_{i2}(\mu_{i1} - \mu_{i2}) (3(\sigma_{i1}^2 - \sigma_{i2}^2) + (1 - 2p_{i1})(\mu_{i1} - \mu_{i2})^2),$$

$$(34) \quad k_i = \text{kurt}[X_i] = \sigma_i^{-4} [3p_{i2}\sigma_{i1}^4 + 3p_{i1}\sigma_{i2}^4 + p_{i1}p_{i2}(\mu_{i1} - \mu_{i2})^2 (6(p_{i1}\sigma_{i2}^2 + p_{i2}\sigma_{i1}^2) + (3p_{i1}^2 - 3p_{i1} + 1)(\mu_{i1} - \mu_{i2})^2)].$$

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

$$\begin{aligned}
(35) \quad \text{Corr}(X_1, X_2) &= \frac{1}{\sigma_1 \sigma_2} \sum_{i,j=1,2} p_{1i} p_{2j} \sigma_{1i} \sigma_{2j} \rho_{ij} \\
\text{Corr}^{\text{bs}}(X_1, X_2) &= \frac{1}{\sigma_1^2 \sigma_2^2 \sqrt{(k_1 - 1)(k_2 - 1)}} \left(2 \sum_{i,j=1,2} p_{1i} p_{2j} \sigma_{1i}^2 \sigma_{2j}^2 \rho_{ij}^2 \right. \\
(36) \quad &\quad \left. + 4(\mu_{11} - \mu_{12})(\mu_{21} - \mu_{22}) p_{11} p_{12} p_{21} p_{22} \sum_{i,j=1,2} (-1)^{i+j} \sigma_{1i} \sigma_{2j} \rho_{ij} \right).
\end{aligned}$$

One can compute the moments and correlations simply through definitions. Here we provide the details of characterizing the BS-Corr. Denote $y_{ij} = \frac{x_{ij} - \mu_{ij}}{\sigma_{ij}}$, $i, j = 1, 2$, and we construct Schmidt orthogonalization as follows

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

Then we have:

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

Using the fact that y_{ij} are mutually independent, we obtain (36).

B.2 Sufficiency of the First Four Moments

In Appendix B.1, we have presented the first four moments of a mixed normal distribution. In this section, we investigate whether these moments are sufficient to uniquely characterize such a distribution. To accomplish this, we will outline specific regularity conditions that establish the uniqueness of a mixed normal distribution based solely on its first four moments.

Theorem 2. *Consider a mixture normal distributed random variable X_1 defined as (29). The Denote mean, standard deviation, skewness and kurtosis of X_1 as μ, σ, s, k . For convenience, denote 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*

$$(37) \quad \frac{1}{2} + \frac{\sqrt{15}\sqrt{13 - 2\sqrt{31}}}{30} < p < 1,$$

$$(38) \quad -\frac{2p(1-p)(2p-1)}{(p^2 - p + \frac{1}{3})} \left(\frac{1}{3}t^{3/2} - t^{1/2} \right) \leq s < 0,$$

$$(39) \quad 0 < u_d \leq t\sigma^2,$$

where $t = \left(\sqrt{\frac{90k - 158}{31}} + 1.9 \right)$. Then $(\mu, \sigma, s, k)^T$ are sufficient statistics to uniquely characterize a mixed normal distribution.

Proof. From (32) and (33),

$$(40) \quad \sigma_1^2 = \frac{p^2 - 1}{3}u_d + \frac{\sigma^3 s}{3p} \frac{1}{\sqrt{u_d}} + \sigma^2,$$

$$(41) \quad \sigma_2^2 = \frac{p(1-2p)}{3}u_d - \frac{\sigma^3 s}{3(1-p)} \frac{1}{\sqrt{u_d}} + \sigma^2.$$

Institute to (34), we have:

$$(42) \quad f(u_d) = 0,$$

where,

$$(43) \quad f(u_d) = \left(-10(p^2 - p + \frac{1}{30})^2 + \frac{31}{90} \right) u_d^2 - 2\sigma^2 (2p-1)^2 u_d + \frac{2s\sigma^3 \cdot (1-2p)}{3p(p-1)} u_d^{1/2} \\ + \frac{2s\sigma^5 (2p-1)}{p(p-1)} u_d^{-1/2} + \frac{s^2\sigma^6 (p^2 - p + \frac{1}{3})}{p^2(p^2 - 2p + 1)} u_d^{-1} - \sigma^4(k-3).$$

We will prove that (42) only has one root under the three constrains.

1) $f(u_d)$ is convex.

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

The first three terms are positive since $p \in (1/2, 1)$, $s < 0$, and using the root formula of quartic equations, one can easily verify

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

is equivalent to (37). Then $f''(u_d) > 0$ i.e. $f(u_d)$ is convex.

2) Since the coefficient of u_d^{-1} in $f(u_d)$ is strictly positive, we have $\lim_{u_d \rightarrow 0} f(u_d) = +\infty$.

3) $f(t\sigma^2) < 0$. First,

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

where we use²⁴

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

when $\frac{1}{2} + \frac{\sqrt{15}\sqrt{13-2\sqrt{31}}}{30} < p < 1$. Second,

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

²⁴One can verify the inequality using the first derivative. The exact lower bound is complicated and here we round it to two significant digits.

With 1), 2) and 3), $f(u_d) = 0$ have and only have one root under the three constrains. According to (31), (40) and (41), u_d can uniquely determine the values of other parameters. Then we prove that $(\mu, \sigma, s, k)^T$ is a sufficient statistics. \square

Theorem 2 allows us to establish the sufficiency of the first four moments with p satisfying (37) and (38). The approximate range of (37) is $(0.6763, 1)$ and condition (38), (39) also imply ranges wide enough to cover parameter values estimated from data. For example, if we take $p = 0.7, \sigma = 0.2, k = 6$ and the two constrains become:

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

Note the calibrated parameters in Table 2 satisfy these three conditions. We thus introduce the following assumption, which ensures that the first four moments can serve as sufficient statistics to uniquely characterize a mixed normal distribution:

- **Assumption 1:** The probability parameter and first four moments of X_1, X_2 satisfy conditions (37), (38) and (39).

B.3 Sufficiency of BS-Corr

There are different covariances that have been used in literature to capture nonlinear dependences. For instance, Harvey and Siddique (2000) uses coskewness to measure the dependence between individual asset and market portfolio. In this part, we establish that, under some regularity conditions, BS-Corr is a sufficient measure to capture all nonlinear dependences.

Let's first consider the commonly used nonlinear dependence in 3rd and 4th degrees. We can define 5 possible coskewness and cokurtosis between two random variables X_1 and X_2 :

$$(44) \quad S(X_1, X_1, X_2) \triangleq \frac{\mathbb{E}[(X_1 - \mathbb{E}X_1)^2(X_2 - \mathbb{E}X_2)]}{\sigma_1^2 \sigma_2},$$

$$(45) \quad S(X_1, X_2, X_2) \triangleq \frac{\mathbb{E}[(X_1 - \mathbb{E}X_1)(X_2 - \mathbb{E}X_2)^2]}{\sigma_1^2 \sigma_2},$$

$$(46) \quad K(X_1, X_1, X_1, X_2) \triangleq \frac{\mathbb{E}[(X_1 - \mathbb{E}X_1)^3(X_2 - \mathbb{E}X_2)]}{\sigma_1^3 \sigma_2},$$

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

$$(48) \quad K(X_1, X_2, X_2, X_3) \triangleq \frac{\mathbb{E}[(X_1 - \mathbb{E}X_1)(X_2 - \mathbb{E}X_2)^3]}{\sigma_1 \sigma_2^3},$$

where σ_i is the standard deviation of X_i .

Is it sufficient to use BS-Corr to measure all the coskewness and cokurtosis above, at least under the mixture normal distribution? We will prove that the sufficiency only requires a mild condition. Since we have established the sufficiency of the first four moments of single mixture normal distributions in B.2, the moments except for Corr and Corr^{bs} will be considered fixed in this section.

Based on the properties of bivariate normal distribution²⁵, we can compute coskewness and cokurtosis for bivariate mixed normal distributions.

Lemma 3. *If $(X_1, X_2)^T$ follows a bivariate mixed normal distribution with parameters $p_1, \mu_{11}, \mu_{12}, \sigma_{11}, \sigma_{12}, p_2, \mu_{21}, \mu_{22}, \sigma_{21}, \sigma_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}$ and $\mu_i, \sigma_i, s_i, k_i$ are defined as B.1, then:*

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

Proof. We have:

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

where Z_{11}, Z_{12} are the two normal components of X_1 (i.e. $X_1 = I_1Z_{11} + (1 - I_1)Z_{12}, I_1 \sim B(1, p_1)$) and Z_{21}, Z_{22} are components of X_2 . According to the property of normal distributions, $S(Z_{1i}, Z_{1i}, Z_{2j}) = 0$. Thus, we have:

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

Institute $\mathbb{E}Z_{1i} = \mu_{1i}, \mathbb{E}X = p_1\mu_{11} + (1 - p_1)\mu_{12}$ and we obtain the equation of $S(X_1, X_1, X_2)$. The calculations for cokurtosis are similar so we omit the details here. \square

²⁵If $(Z_1, Z_2)^T$ 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:

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

It is clear that the cokurtosis $K(X_1, X_1, X_2, X_2)$ can be determined by $\text{Corr}^{\text{bs}}(X_1, X_2)$. As for the others, we need the following lemma.

Lemma 4. $(X_1, X_2)^T$ follows a bivariate mixture normal distribution with parameters $p_1, \mu_{11}, \mu_{12}, \sigma_{11}, \sigma_{12}, p_2, \mu_{21}, \mu_{22}, \sigma_{21}, \sigma_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}$ and $\rho_{11} = \rho_{22} = \rho_1, \rho_{12} = \rho_{21} = \rho_2$. Denote $\iota_{ij} = p_{1i}p_{2j}\sigma_{1i}\sigma_{2j}, i, j = 1, 2$. We have:

$$(49) \quad \rho_2 = k\rho_1 + b,$$

$$(50) \quad \text{Corr}^{\text{bs}}(X_1, X_2) = a_0 + a_1\rho_1 + a_2\rho_1^2, a_2 > 0,$$

where

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

Proof. With (35) and (36), the lemma can be easily verified. \square

Solve (50) and we get

$$(51) \quad \rho_1^+ = -\frac{a_1}{2a_2} + \left(\frac{1}{\sqrt{a_2}} \sqrt{\text{Corr}_{xy}^{\text{bs}} - \frac{4a_0a_2 - a_1^2}{4a_2}} \right), \rho_1^- = -\frac{a_1}{2a_2} - \left(\frac{1}{\sqrt{a_2}} \sqrt{\text{Corr}_{xy}^{\text{bs}} - \frac{4a_0a_2 - a_1^2}{4a_2}} \right)$$

According to Lemma 4 and (49), given marginal distributions and linear correlation, $S(X_1, X_1, X_2)$ and $K(X_1, X_1, X_1, X_2)$ are linear functions of ρ_1 , and thus can be uniquely determined by $\text{Corr}^{\text{bs}}(X_1, X_2)$ with the following assumption:

- **Assumption 2:** $(X_1, X_2)^T$ follows a bivariate mixture normal distribution satisfies $\rho_{11} = \rho_{22} = \rho_1, \rho_{12} = \rho_{21} = \rho_2$, and only one of ρ_1^+, ρ_1^- falls into $[-1, 1]$, which means (50) only solves one root in $[-1, 1]$.

This condition can be easily satisfied when $\text{Corr}^{\text{bs}}(X_1, X_2)$ takes positive values, which is in line with estimation from data. In the following theorem, we establish that the BS-Corr is a sufficient measure of nonlinear dependence under mixture normal distributions.

Theorem 5. $(X_1, X_2)^T$ follows 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 . With Assumption 1 and 2, $(\mu_1, \sigma_1, s_1, k_1, \mu_2, \sigma_2, s_2, k_2, \text{Corr}(X_1, X_2), \text{Corr}^{\text{bs}}(X_1, X_2))^T$ are sufficient statistics to uniquely determine the joint distribution of $(X_1, X_2)^T$.

Proof. According to Theorem 2, with assumption 1, we know that $(\mu_i, \sigma_i, s_i, k_i)^T$ is sufficient to X_i and determines the underlying parameters $(\mu_{i1}, \mu_{i2}, \sigma_{i1}, \sigma_{i2})^T$. Furthermore, assumption 2 assumes that (49) and (50) only solves one root, which implies that, given $(\mu_1, \sigma_1, s_1, k_1, \mu_2, \sigma_2, s_2, k_2, \text{Corr}(X_1, X_2), \text{Corr}^{\text{bs}}(X_1, X_2))^T$, one can uniquely determine ρ_1, ρ_2 . Thus, they are sufficient statistics. \square

The theorem suggests that, subject to certain mild conditions, the BS-Corr can serve as a sufficient measure of all forms of nonlinear dependence between two random variables that follow a bi-variate mixed normal distribution.

C The One-period Model

In this section, we first show that the household's utility optimization problem can be approximated by a simple one-period model with higher order moments in its objective. Then we examine the higher order moments for the portfolio in this first model.

C.1 The One-period Model as an Approximation to the Full model

According to the assumptions in Section 2.3, W_0 is the initial wealth, $W_1 = R^p W_0$ is the wealth after one period, and $w_1 = \mathbb{E}[W_1]$ is the expected wealth. Let $u(\cdot)$ be the utility function for a representative household. The utility based portfolio optimization problem is

$$(52) \quad \begin{aligned} & \max_{\alpha} \mathbb{E}[u(W_1)] \\ \text{s.t. } & W_1 = W_0 R^p = W_0(1 + R_f + \alpha R^E + R^L). \end{aligned}$$

Consider the Taylor Series expansion of $u(W)$ at w_1 :

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

Taking expectations on both sides suggests

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

where the risk weights $\lambda_1, \lambda_2, \lambda_3$ satisfy

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

The risk weights $\lambda_1, \lambda_2, \lambda_3$ can be approximated by the household's utility and its conjectured wealth expectation and variance. For instance, let $\widehat{w}_1, \widehat{\text{Var}}(R^p)$ be the predicted wealth mean and portfolio return variance, then the objective of (52) can be approximated by

$$(53) \quad \mathbb{E}[u(W_1)] \approx \frac{u(\widehat{w}_1)}{\widehat{w}_1} W_0 \left(\mathbb{E}(R^p) - \lambda_1 \text{Var}(R^p) - \lambda_2 \text{Skew}(R^p) - \lambda_3 \text{Kurt}(R^p) \right),$$

where the risk weights can be approximated as follows

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

Thus the one-period utility optimization problem can be approximated by (4) and the risk weights $\lambda_1, \lambda_2, \lambda_3$ are associated with the household's utility.

C.2 Higher Order Moments of the Portfolio Returns

Based on Lemma 3, we can compute the skenewss and kurtosis of the portfolio returns in the first model.

Theorem 6. $(R^E, R^L)^T$ is a random vector. A portfolio $\Pi(c)$ of $(R^E, R^L)^T$ is defined as the linear combination $cR^E + (1-c)R^L$. Then the standard deviation, skewness and kurtosis of Π can be computed by:

$$\begin{aligned}\sigma_{\Pi} &= \sqrt{c^2 \sigma_E^2 + 2c(1-c) \sigma_E \sigma_L \text{Corr}_{EL} + (1-c)^2 \sigma_L^2}, \\ \text{Skew}_{\Pi} &= \frac{1}{\sigma_{\Pi}^3} [c^3 \sigma_E^3 S_E + (1-c)^3 \sigma_L^3 S_L + 3c^2(1-c) \sigma_E^2 \sigma_L S(R^E, R^E, R^L) \\ &\quad + 3c(1-c)^2 \sigma_E \sigma_L^2 S(R^E, R^L, R^L)],\end{aligned}$$

$$\begin{aligned} \text{Kurt}_{\Pi} = & \frac{1}{\sigma_{\Pi}^4} [c^4 \sigma_E^4 K_E + (1-c)^4 \sigma_L^4 K_L + 4c^3(1-c)K(R^E, R^E, R^E, R^L) \\ & + 4c(1-c)^3 K(R^E, R^L, R^L, R^L) + 6c^2(1-c)^2 K(R^E, R^E, R^L, R^L)], \end{aligned}$$

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

According to Appendix B.2, if $(R^E, R^L)^T$ follows a bivariate mixture normal distribution and satisfies assumption 1 and 2 with dependence parameters ρ_1 and ρ_2 , given marginal distributions and linear correlation, $S(R^E, R^E, R^L)$ and $K(R^E, R^E, R^E, R^L)$ are linear functions of ρ_1 . Thus we can denote

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

Values of the coefficients can be computed by Lemma 3 and 4. Combine the notations with (51), we can rewrite the portfolio's skewness and kurtosis as follow:

$$(54) \quad \text{Skew}_{\Pi} = \text{sgn}(\rho_1 + \frac{a_1}{2a_2}) k_s \sqrt{\text{Corr}_{EL}^{\text{bs}} - \frac{4a_0 a_2 - a_1^2}{4a_2}} + b_s,$$

$$(55) \quad \text{Kurt}_{\Pi} = a_k \text{Corr}_{EL}^{\text{bs}} + \text{sgn}(\rho_1 + \frac{a_1}{2a_2}) b_k \sqrt{\text{Corr}_{EL}^{\text{bs}} - \frac{4a_0 a_2 - a_1^2}{4a_2}} + c_k,$$

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

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

With $c = \frac{\alpha}{1+\alpha}$, (54) and (55) allow us to solve the portfolio optimization problem (52) with the

approximation form (53), and show that the nonlinear effect of the BS-Corr is through the skewness channel and kurtosis channel.

D Numerical Methods

D.1 Numerical Solutions to the Household Optimization

The model can be numerically solved using backward induction. The value function for each period depends on normalized cash on hand x_t , which is a continuous variable and thus need to be discretized. The terminal condition in the last period is determined by the bequest motive, and the value function corresponds to the bequest function. For each period, we use a grid search to optimize the value function. We compute the value associated with each grid of consumption and risky share, and choose optimal grids achieving the maximum value as the policy rules. For each period t prior to T , and for each point in the state space, this procedure is iterated backward.

To approximate the expected value at next period which is required in the backward induction, we use Gauss-Hermite for mixed normal distributed. Generally, 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 $\epsilon_{t+1} \sim N(0, \sigma_\epsilon^2)$. We can apply Gauss-Hermite method to the four normal distributions embedded in the mixture normal. Specially, we have:

$$\mathbb{E}[h] = \sum_{i,j=0,1} \mathbb{E}[h|I_1 = i, I_2 = j] \mathbb{P}(I_1 = i, I_2 = j) = \sum_{i,j=1,2} \mathbb{E}[h(x_{1i}, x_{2j}, \epsilon_{t+1})] \mathbb{P}(I_1 = i) \mathbb{P}(I_2 = j).$$

Now for each component of the summation, the vector $(x_{1i}, x_{2j}, \epsilon_{t+1})^T$ is normally distributed and then we can apply the Gauss-Hermite method.

D.2 Calibration

We provide explicit form of moments used by GMM method in section 4.1. According to equation (9) and (10), we have

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

Then we calculate the moments:

$$\text{Corr}(\eta_{t+1}, \delta_{it+1} - \delta_{it}) = \text{Corr}(\eta_{t+1}, u_{it+1}) \frac{\sigma_u}{\sqrt{\sigma_u^2 + 2\sigma_\epsilon^2}} = \frac{\sigma_u}{\sqrt{\sigma_u^2 + 2\sigma_\epsilon^2}} \text{Corr}(\eta_{t+1}, u_{t+1}),$$

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

where $\text{Cor}(\eta_{t+1}, u_{t+1})$, $\text{Cor}^{\text{bs}}(\eta_{t+1}, u_{t+1})$, σ_u , Skew_u and Kurt_u can be computed using the results about in Appendix B.1.1, B.1.2 and B.1.3.

E Nonlinear Effect of the BS-Corr

In this section, we will show that a household that faces a negative enough BS-Corr may optimally delay its participation and hold less risky shares. We show numerically that the nonlinear effects of BS-Corr on household's policy functions still hold in our full model, in line with the first model in Section 2.3.

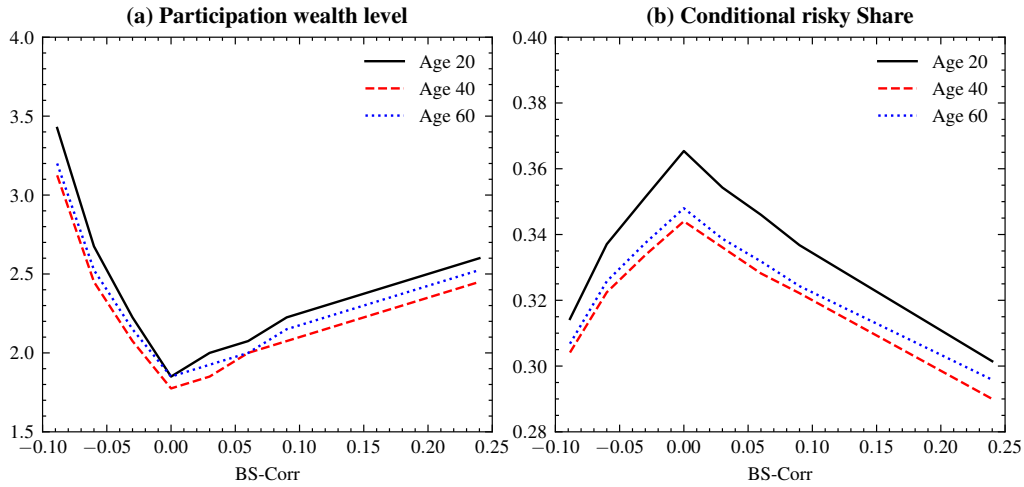
Figure 10 displays the policy functions from our full model for households at age 20, 40, and 60, with varying BS-Corr based on the parameter values in the first model in Section 2.3. Panel A depicts the participation wealth threshold, revealing that households are more likely to participate in the stock market when the BS-Corr approaches a turning point around zero. Furthermore, as shown in panel B, we observe that participating households exhibit a greater willingness to hold a higher proportion of risky assets when the BS-Corr approaches a turning point around zero. Whenever the BS-Corr deviates from the turning point around zero, households exhibit a tendency to delay their stock market participation and decrease their allocation to risky assets in order to lower their overall portfolio risk. This result suggests that BS-Corr has a nonlinear effect on household's portfolio decisions.

This finding can be attributed to the two channels through which BS-Corr influences household portfolio decisions, as demonstrated in the first model. When the BS-Corr is negative enough, the nonlinear dependence between labor income and stock returns significantly enhances the skewness

of household wealth, leading them to delay participation and reduce their exposure to risky assets. Conversely, when the BS-Corr is positive enough, the nonlinear dependence enhances the kurtosis of household wealth, contributing to the decision to delay participation and reduce risky asset holdings. It is worth noting that the effects of the two channels are asymmetric, with the kurtosis channel in this numerical case appearing to have a greater impact. Consequently, this finding suggests a novel version of the “stock-like” theory (i.e., Benzoni et al. (2007)), where the nonlinear dependence between labor income and stock returns renders households’ labor income more “stock-like”.

Figure 10: Changes in Portfolio Decisions

This figure plots the participation rate and conditional risky share from the policy functions with BS-Corr from -0.1 to 0.25. We take $\mu^L = 0, \sigma^L = 0.25, \text{Skew}^L = -0.5, \text{Kurt}^L = 3.0$ for labor income risk, $\mu^S = 0, \sigma^S = 0.2, \text{Skew}^S = -1, \text{Kurt}^S = 4.5$ for stock return risk and linear correlation of 0.1 between them and the preference parameters 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 take the permanent labor growth state $w_{i,t}$ as 1.



F Sensitivity Analysis of Permanent Labor Growth State

In this paper, we take the permanent labor growth state as 1 in analysis of policy function. However, different permanent labor growth state can have huge impact on households’ portfolio decisions. To understand this effect better, we show the policy functions of different labor growth state in Figure 11. The pattern of optimal risky share is similar with different $w_{i,t}$, but the level is different. In general, a larger $w_{i,t}$ indicates more aggressive portfolio decisions, since households predict more future labor income. In addition, wealthier and older households are less sensitive to $w_{i,t}$, since labor income will more trivial to them. More specifically, those wealthier households will pay more attention to their cash on hand and those older households are more likely to receive a death shock.

Figure 11: Sensitivity Analysis of Permanent Labor Growth State

The six panels plot the optimal risky asset shares policy functions for different ages and groups. The first column corresponds with the college group and the second corresponds with the no-college group. The policy functions are solved with calibrated parameters. We take the permanent labor growth state $w_{i,t}$ as 0.95, 1, 1.05, respectively.

