

Income Volatility and Hedging Motive: Depopulation in Stock Market

Tai Lo Yeung, USI

May 2023

What this paper is about

- Economic theory predicts that
 - [H1](#): The market participation propensity should **increase** as the correlation between income growth and stock market returns **decreases**.
 - [H2](#): The **higher** a worker's wage volatility, the **lower** her exposure to the stock market.
- However, empirical studies find limited support for
 - Why do so few hold stocks?
 - What is the main driver when people make decision on holding stocks or not? For the purpose of hedging income risk (risk averse) or reaching for yield (risk seeker)?
- Thus, this paper provides new empirical evidence by
 1. Showing that the SIPP is suitable for doing research in this field.
 2. Using individual level risk volatility and correlation estimated out of SIPP to test the hypothesis.

What this paper is (**not**) about

- I am mainly using
 - A binary variable to identify whether people participate in the stock market or not, but I do not have the intensity of the participation for the all the survey waves.
 - Hence, the empirical model I am using can only capture people that enter or leave the market but cannot observe when they are adjusting their portfolio. But I considered it already contained sufficient information to answer the research question.
- So, this paper is not going to tell the story about precise portfolio decisions, but rather trying to show what triggered people to enter or exit the market.
- I will still provide some tests related to the risky share in investors' portfolio using a truncated sample but that won't be the main take away from this research.

Nationally Representative Sample

- SIPP (Survey of Income and Program Participation)
 - Monthly data from 2004-2020. With 2004 Panel-(2004.02-2008.01), 2008 Panel (2008.09-2013.12), 2014 Panel (2013.01-2016.12), data are provided annually after 2014 Panel
 - Information about stock: Y/N question before 2013, numerical questions after 2013
 - Demographic: Age, Location, Education, Hours worked, Type of job...
 - **Why SIPP over PSID?** Monthly level captures extra information about wage, i.e., people who earned the same amount of income each year might have different level of volatility within the year. For example, someone earned 50,000 USD in 2008, but he earned it from the first 4 months at lost his job in May, we won't observed this in annual panel.
- SCF (Survey of Consumer Finances)
 - Data available triennially, 2004, 2007, 2010, 2013, 2016, 2019
 - Information about stock: Y/N question
- SOI (Statistics of Income) Tax Stats
 - Country level data available annually from 2011
 - Information about stock: Dividend tax

Data Selection

- SIPP (to select the representative working group in the U.S.)
 - Age from 18 to 66
 - Monthly labor income from 500 to 50,000
 - Only self-reported value (no imputed or proxy value)

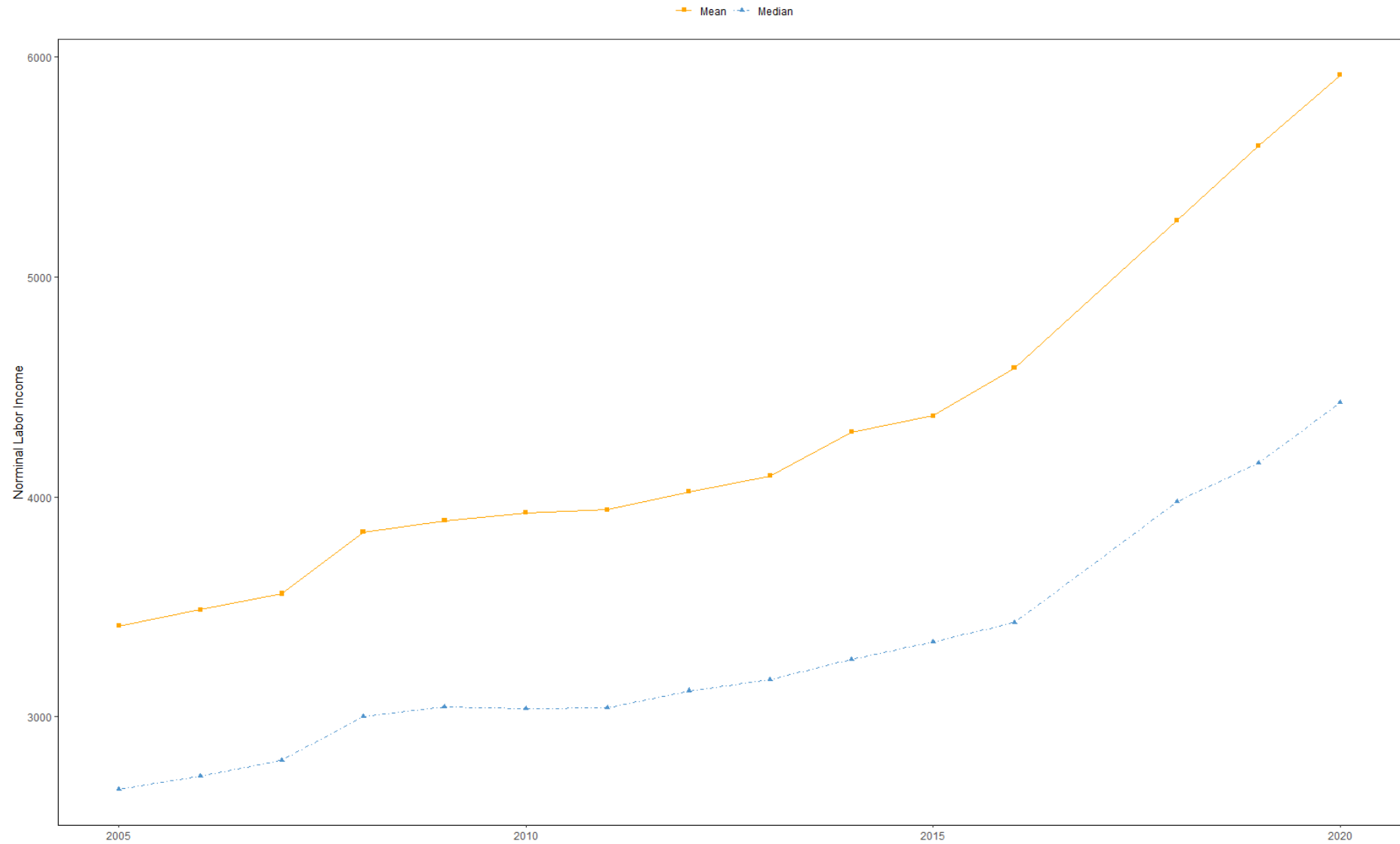
Table 2: Summary Statistic (SIPP)

	Mean	Std. dev	Median
Demographics			
Age	42.5	14.3	43
Male	0.51	0.5	1
Hispanic	0.13	0.33	0
Citizen	0.93	0.26	1
Education Level	41	2.82	41
Income & Wealth (\$)			
Salary	3616	4626	2598
Participation			
Participates in STMF	0.17	0.38	0
Observations ($N = n \times T$)	5763503		
Households (n)	216706		
Average Months (T)	26.6		

Table 20: Summary Statistic (This Paper)

	Mean	Std. dev	Median
Demographics			
Age	43.7	11.8	44
Male	0.46	0.5	0
Hispanic	0.11	0.31	0
Citizen	0.94	0.24	1
Education Level	41.5	2.69	41
Income & Wealth (\$)			
Salary	4138	3914	3157
Participation			
Participates in STMF	0.19	0.39	0
Observations ($N = n \times T$)	1845180		
Households (n)	93091		
Average Months (T)	19.8		

Trend of Labor Income 2004-2020



Trend of Labor Income 2004-2020

From Berthold (2023, EER)

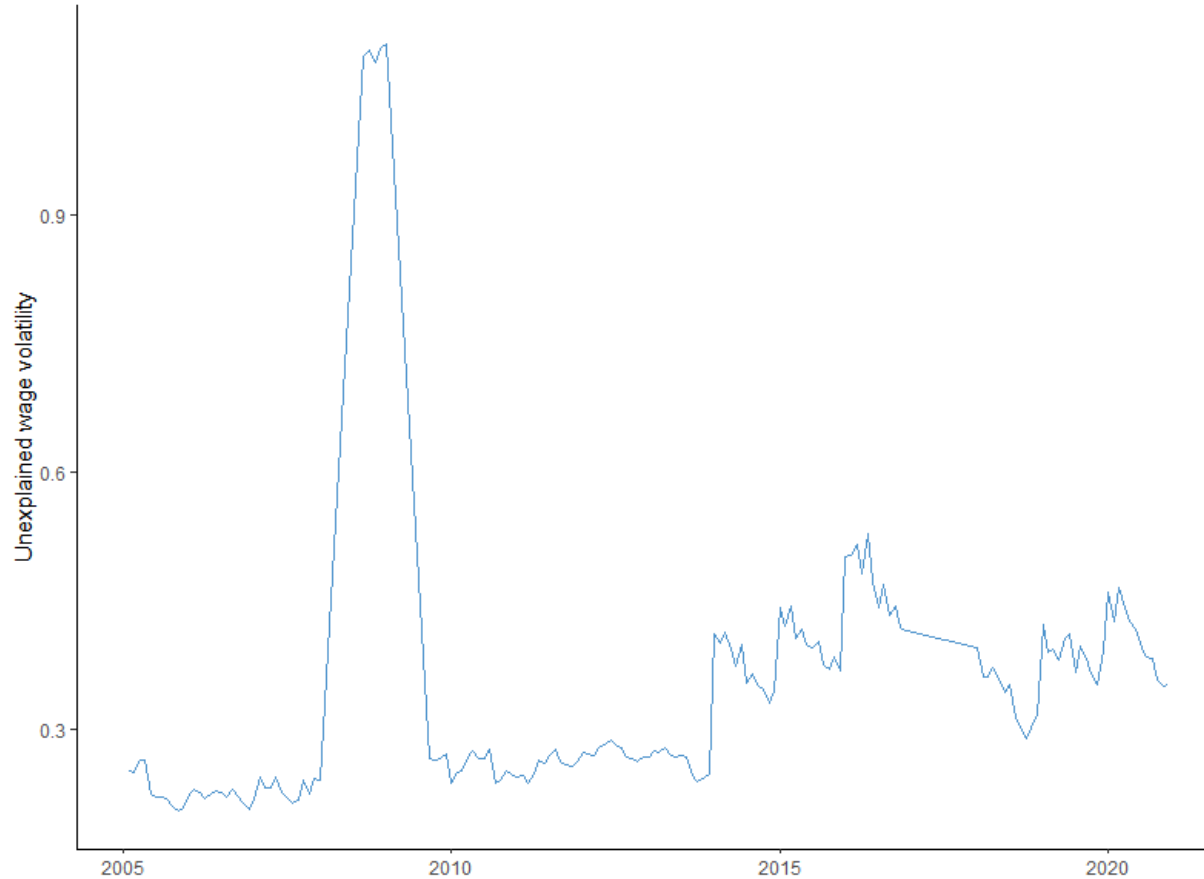


Figure 1 UNCERTAINTY AND RISK AVERSION MEASURES FROM BEKAERT ET AL. (2021)



NOTE. This Figure displays the Uncertainty and Risk Aversion variables from [Bekaert et al. \(2021\)](#) as well as the likely sources of the different spikes. While the raw correlation between the two series is high (0.81), there are also periods with distinctly different behaviour.

Measurements of Market Participation Rate

- From SCF

$$Participation Rate_t = \frac{\# \text{ Have a brokerage account (or have stock holding)}}{\# \text{ Returns}}$$

- From SOI

$$Participation Rate_t = \frac{\# \text{ Returns with Ordinary Divedends}}{\# \text{ Returns}}$$

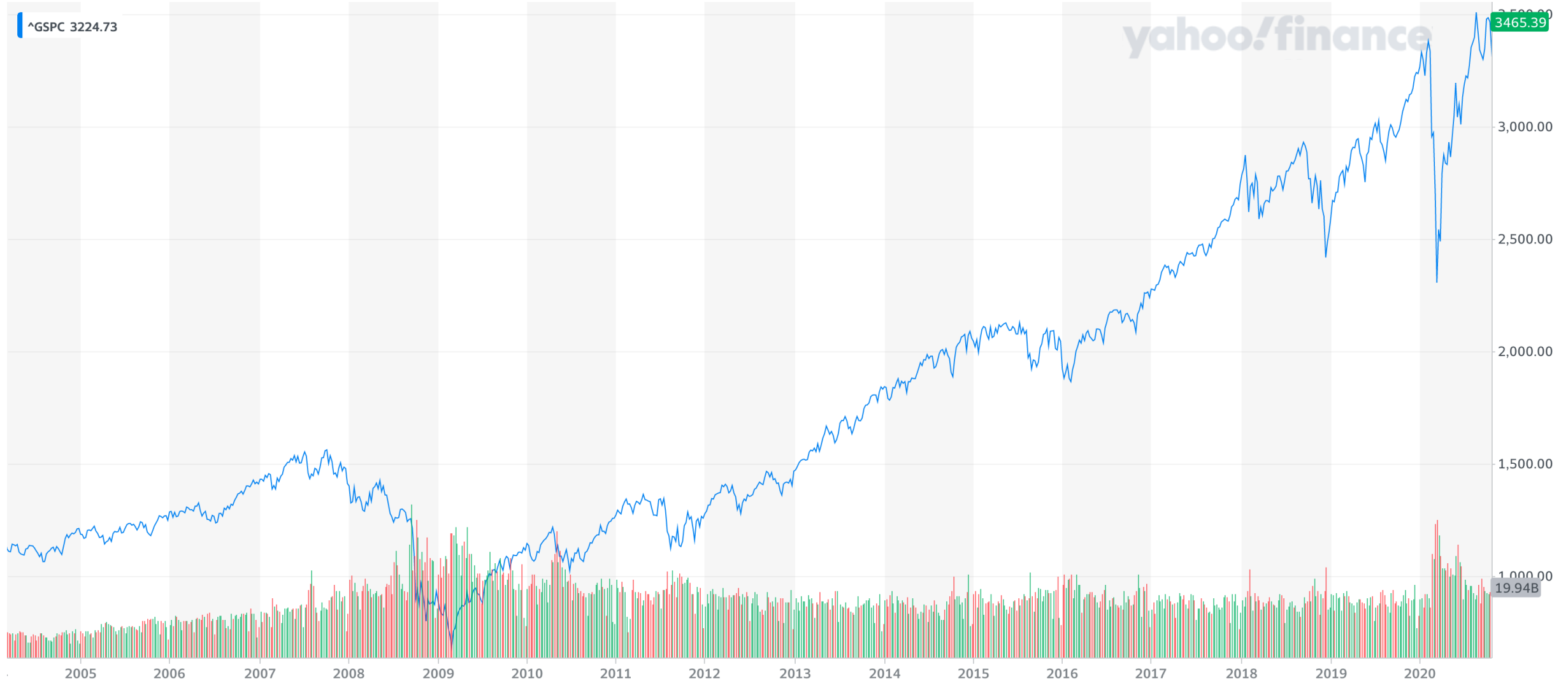
- From SIPP (After 2013)

$$Participation Rate_t = \frac{\# \text{ Value Invested in Stock and Fund Market} > 0}{\# \text{ Returns}}$$

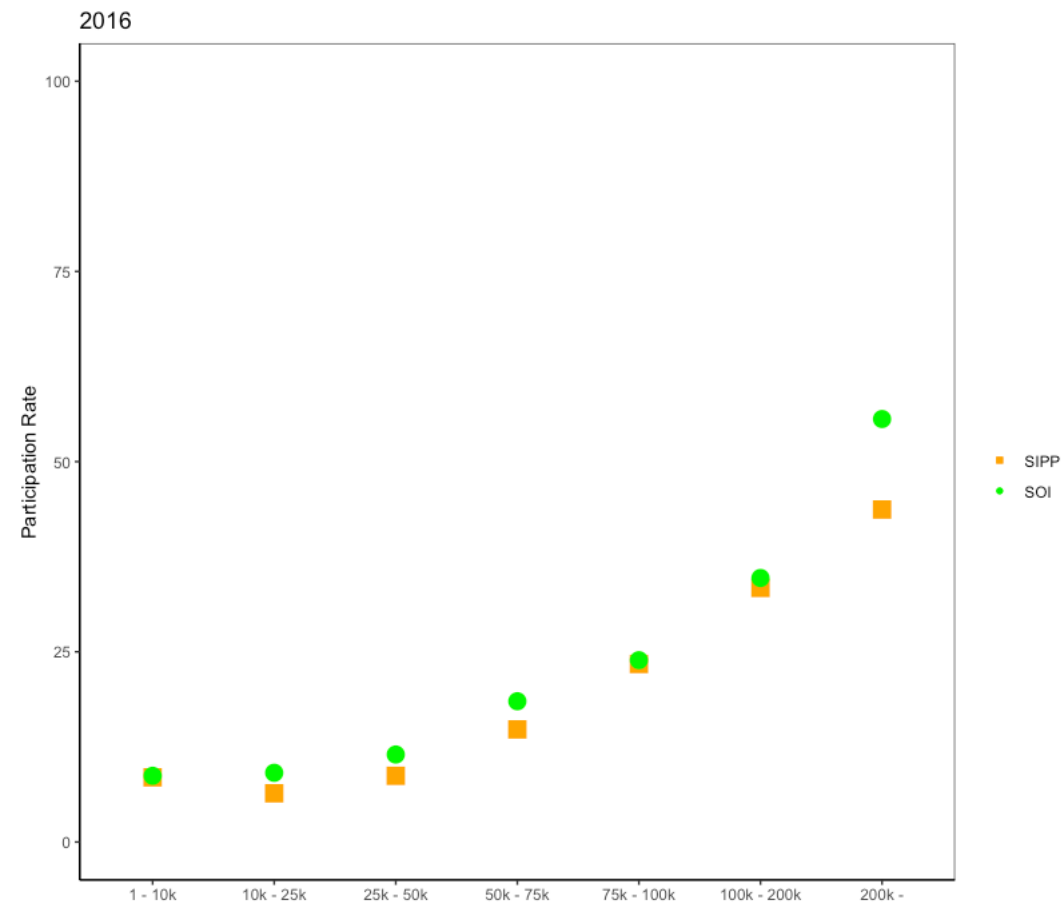
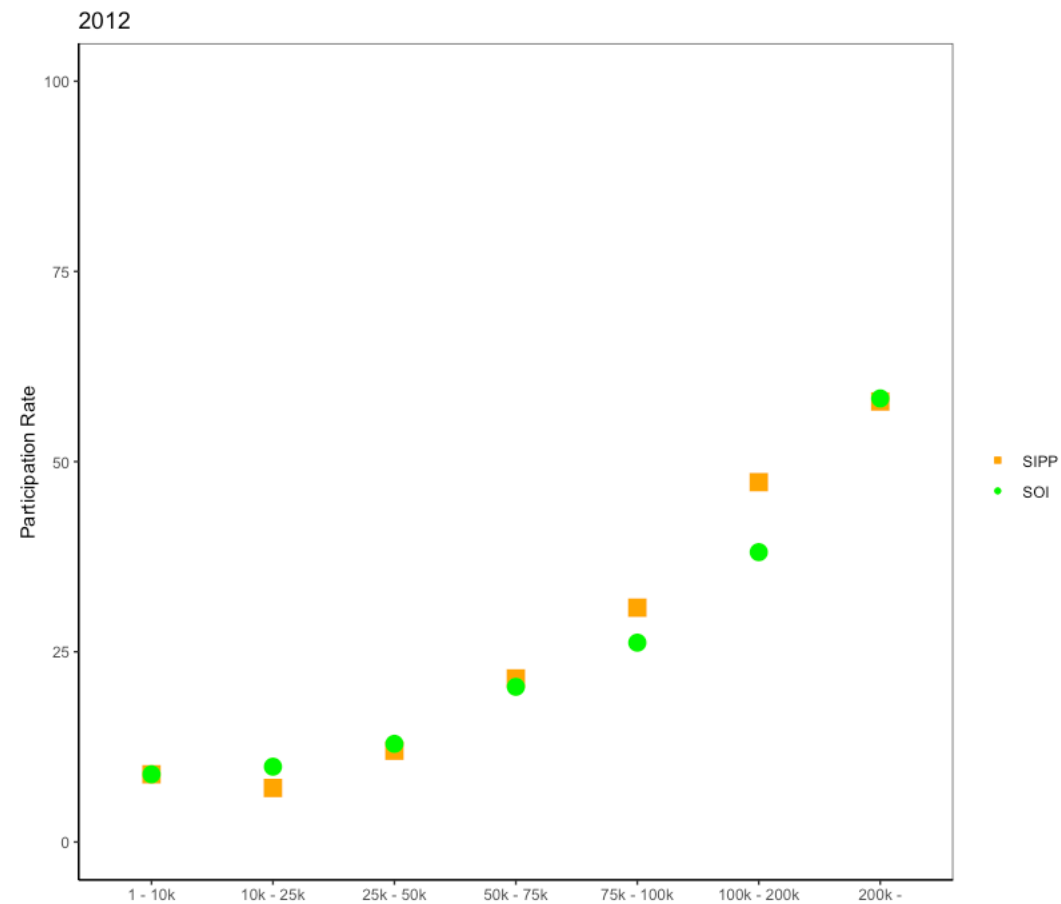
Measurements of Market Participation Rate

- From SCF
 - Advantage: Tested and used by numerous research before.
 - Drawback: Only available every 3 years.
- From SOI
 - Advantage: It's tax report.
 - Drawback: It only contains those who buy and hold stocks with dividends payment (However 80% of the stocks out of S&P 500 pay dividends), so it can be treated as of a proxy of the lower bound of stock market participation rate.
- From SIPP
 - Advantage: Have detailed labor market microdata, getting more attention recently, 2 paper using the SIPP market participation data from JFQA, 2022 December.
 - Drawback: The questionnaires are different across waves (especially before/after 2013).

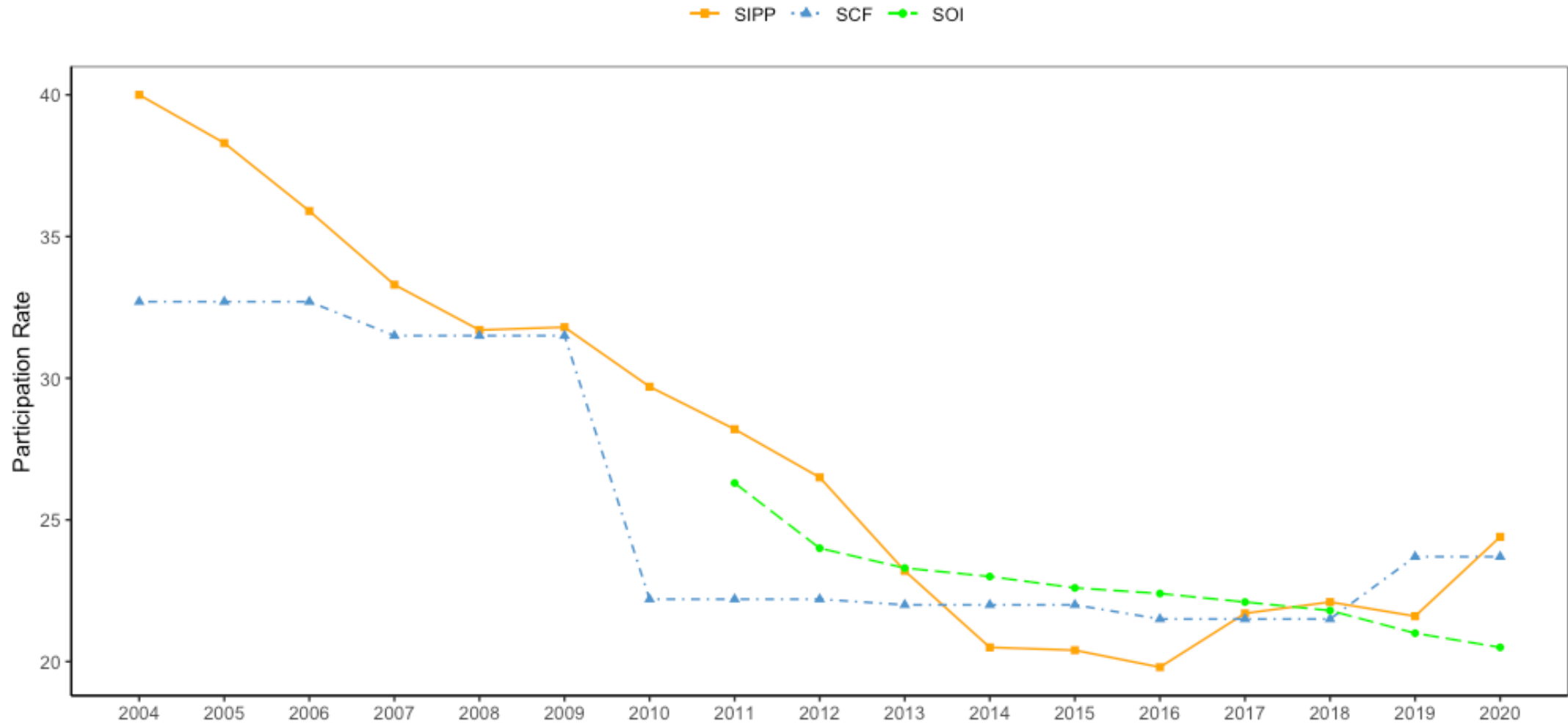
Market Participation Rate 2004 - 2020



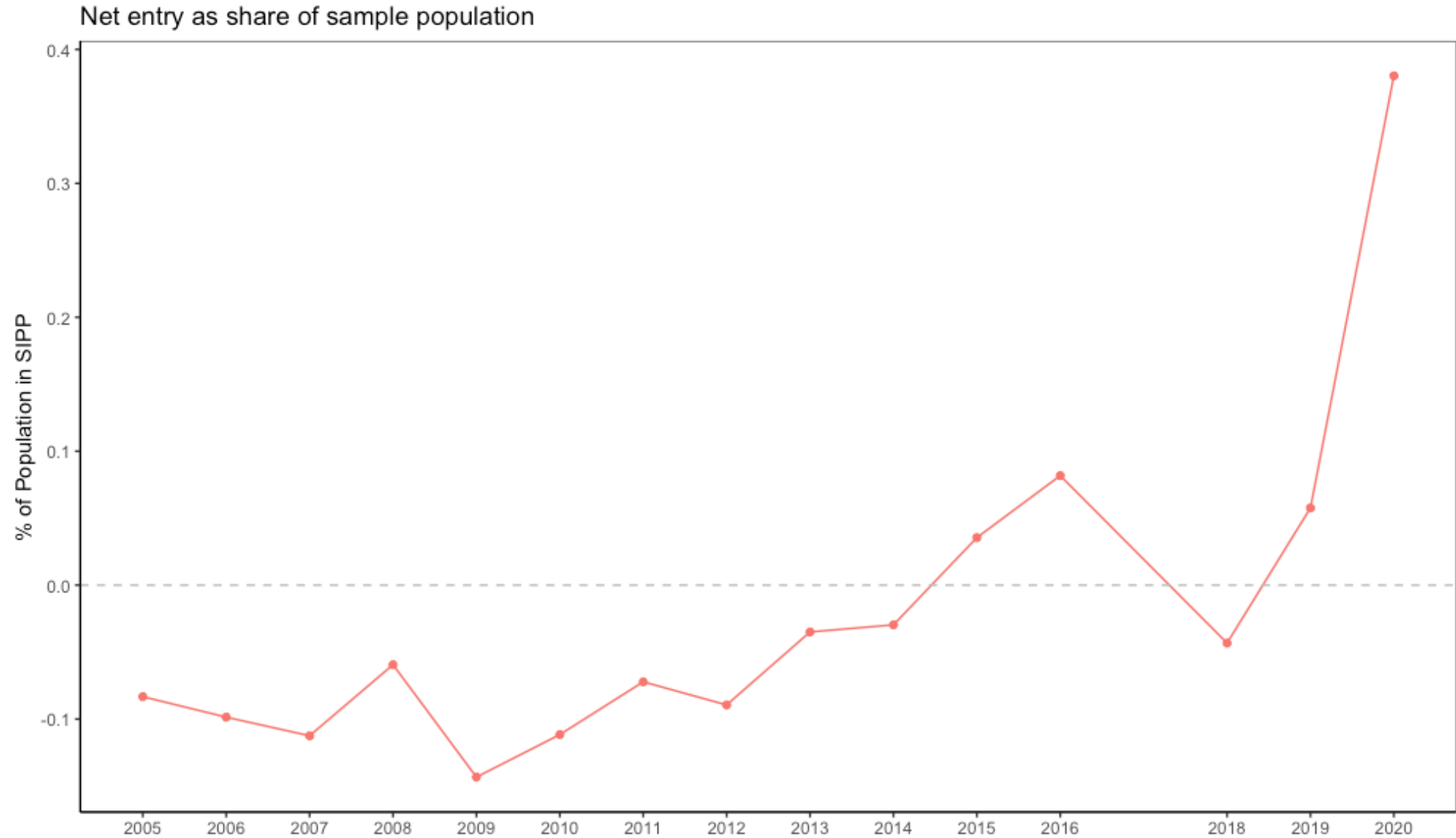
Market Participation Rate by Income



Market Participation Rate 2004 - 2020



Market Participation Rate 2004 - 2020



Literature: Income Risk & Portfolio Choice

- *Income Risk, Borrowing Constraints, and Portfolio Choice*, Guiso et al. (1996 AER)
 - Data: 1989 Italian Survey of Household Income and Wealth (SHIW)
 - Method: Income risk as subjective variance of real income growth
 - Conclusion: When confronted with **uninsurable** income risk, proportion of risky assets gets lower
- *Income Risk and Portfolio Choice: An Empirical Study*, Angerer & Lam (2009 JF)
 - Data: National Longitudinal Survey on Youth (NLSY79)
 - Method: Construct a group-specific risk using income variance
 - Conclusion: First paper attempts to measure **permanent** and **transitory** income risks and estimate their effects on risky asset holding and shows that **permanent** risk shifts household's portfolio toward risk-free asset.
- *Household Portfolio Choices under (Non)linear Income Risk: An Empirical framework*, Galvez (2022 WP)
 - Data: Panel Study of Income Dynamics (PSID) from 1999 to 2009
 - Method: Developed a semi-structural LC framework to empirically quantify the transmission, portfolio choice is an age-dependent functions of **persistent** and **transitory** components of income and assets.

Literature: Hedging & Portfolio Choice

- *Hedging Labor Income Risk*, Betermier et al. (2012 JFE)
 - Data: Longitudinal Individual Data Base (LINDA, Swedish) from 1999 to 2002.
 - Method: Wage volatility by industry as a proxy for all households working in that industry.
 - Conclusion: Household going from a **low** wage volatility industry to a **high** one would **decrease** its exposure to risky asset.
- *Income Hedging and Portfolio Decisions*, Bonaparte et al. (2014 JFE)
 - Data: DNB Household Survey 1993 to 2011 (+ NLSY79 for robustness)
 - Method: The correlation between market return and labor income, $\text{Corr}(R_m, dy)$.
 - Conclusion: When income-return correlation is **low**, individuals exhibit a **greater** propensity to participate in the market and with **larger** proportion of resources allocate to risky assets.
- *This paper* (2023 WP)
 - Data: Survey of Income and Program Participation (SIPP) from 2004 to 2020
 - Method: Monthly level income volatility and new measures for the correlation.
 - Conclusion: Household with **high** wage volatility (both occupation and individual level) has **less** exposure to the stock market; Provided extra empirical evidence to support the correlation hypothesis.

Construction of Income Volatility

- Generic structure of labor income process

$$\begin{aligned}w_{i,t} &= z_{i,t} + e_{i,t} \\ \Delta w_{i,t} &= \Delta z_{i,t} + \Delta e_{i,t}\end{aligned}$$

- Where $w_{i,t}$ is the log labor income for household i at time t , $z_{i,t}$ and $e_{i,t}$ represent the predictable and stochastic component, respectively.
- $\Delta e_{i,t}$ can be simply imputed by running a regression to remove the demographic factors x'_{it} :

$$\Delta w_{i,t} = d_t + x'_{it}\eta + \Delta e_{i,t}$$

- And the wage volatility at aggregate level can be shown as, under Full-Information Rational Expectation (FIRE):

$$\begin{aligned}\text{Var}_i(\Delta w_{i,t}) &= \text{Var}_i(\Delta e_{i,t}) \\ \text{Var}_{IND}(\Delta w_{i,t}) &= \text{Var}_{IND}(\Delta e_{i,t})\end{aligned}$$

Background Theory

- Campbell and Viceira (2001, p.144), model of one period portfolio decision:

$$\max_{\alpha_t} E_t \left[\beta \frac{C_{t+1}^{1-\gamma}}{1-\gamma} \right]$$

$$\begin{aligned} s.t. \quad C_{t+1} &= W_t(1 + R_{p,t+1}) + Y_{t+1} \\ R_{p,t+1} &= \alpha_t(R_{t+1} - R^f) + R^f \end{aligned}$$

- $R_{p,t+1}$: return on portfolio
 - δ : discount factor
 - α_t : weight on the risky portfolio
- Risk-averse investor with CRRA, cannot trade labor income (uninsurable labor income risk), has access to risk-free and risky asset.

Background Theory

- From F.O.C. and [assuming for joint log normality](#) between consumption and returns gives:

$$\beta E_t[C_{t+1}^{-\gamma} R_{t+1}] = \beta E_t[C_{t+1}^{-\gamma} R^f]$$

$$\Rightarrow E_t r_{t+1} - r^f + \frac{1}{2} Var(r_{t+1}) = \gamma cov(r_{t+1}, c_{t+1})$$

- Solving it for portfolio rule:

$$\alpha = \frac{1}{\rho} \left(\frac{\mu + \frac{\sigma_r^2}{2}}{\gamma \sigma_r^2} \right) + \left(1 - \frac{1}{\rho} \right) \left(\frac{\sigma_{yr}}{\sigma_r^2} \right)$$

Background Theory

- Replace the covariance with correlation:

$$\alpha = \frac{1}{\rho} \left(\frac{\mu + \frac{\sigma_r^2}{2}}{\gamma \sigma_r^2} \right) + \left(1 - \frac{1}{\rho} \right) (\textcolor{red}{corr}_{yr}) \left(\frac{\textcolor{red}{\sigma}_y}{\sigma_r} \right)$$

- α : optimal weight on the risky portfolio
 - $corr_{yr}$: correlation between portfolio return and wage growth
 - ρ : elasticity of consumption w.r.t. financial wealth ($\rho \in (0,1)$)
 - $1 - \rho$: elasticity of consumption w.r.t. labor income
 - μ : expected log excess return
 - γ : coefficient of relative risk aversion
 - σ : standard deviation
- First part related to the sharp ratio and risk aversion (most empirical studies in finance focus on this part, by letting $corr_{yr} = 0$); second part related to the [income hedging motive](#), where this paper is focused on.

Empirical Model 1 (General)

- Probit

$$PR_{i,t} = \alpha + \beta_1 \cdot corr_{yr} + \beta_2 \cdot \sigma_y + \beta_3 \cdot (corr_{yr} \times \sigma_y) + \gamma \cdot X'_{i,t} + \epsilon$$

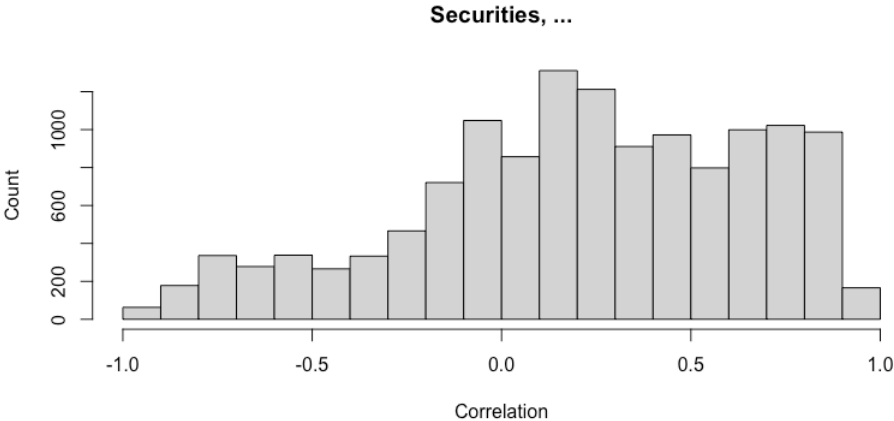
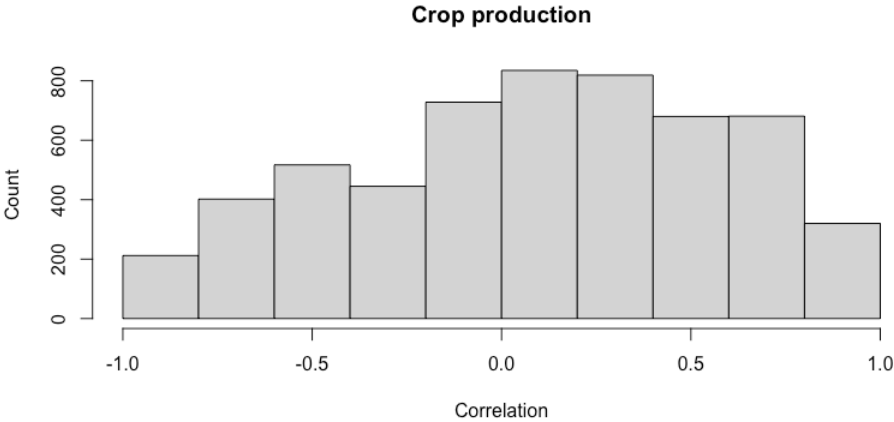
- $PR_{i,t}$: 0/1 dummy for participation in risky asset market
- $X'_{i,t}$: Demographic variables

- Assumption: $corr_i(M, Y)$ is fixed by individual within a time span:

$$corr_{yr} = corr_i(M, Y) = \sum_{t=1}^T \frac{(SP500_t - \overline{SP500})(Y_t - \bar{Y})}{(T-1)(\sigma_Y \cdot \sigma_M)}$$

Table 2: Correlation by Occupation

IND	Description	Estimate
Financial		
6870	Banking and related activities	0.2123
6880	Savings institutions, including credit unions	0.3165
6970	Securities, commodities, funds, trusts, and other financial investments	0.2116
6992	Agencies, brokerages, and other insurance related activities	0.2034
Necessity		
0170	Crop production	0.0921
0380	Coal mining	-0.0488
0770	Construction	0.1369
2390	Rubber products, except tires, manufacturing	0.0490
Luxury		
1691	Cut and sew, and apparel accessories and other apparel manufacturing	0.3003
1770	Footwear manufacturing	0.3182
2890	Coating, engraving, heat treating, and allied activities	0.2472
3690	Other transportation equipment manufacturing	0.3035
25th	50th	75th
0.1182	0.1625	0.2049



Results

Table 3

	<i>Dependent variable:</i>		
	Own STMF		
	(1)	(2)	(3)
$\sigma_y \times corr_{yr}$	−0.198*** (0.007)	−0.142*** (0.007)	−0.090*** (0.007)
Wealth FE	N	Y	Y
Individual FE	N	N	Y
N	1,679,287	1,679,287	1,679,287

Results

Table 3

	<i>Dependent variable:</i>			
	Entering Market			
	(1)	(2)	(3)	(4)
dy	0.041*** (0.006)	0.042*** (0.006)	0.048*** (0.006)	0.026*** (0.006)
Market Index FE	N	Y	Y	Y
Individual FE	N	N	Y	Y
Time FE	N	N	N	Y
N	1,845,180	1,845,180	1,845,180	1,845,180

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4

	<i>Dependent variable:</i>			
	Exiting Market			
	(1)	(2)	(3)	(4)
dy	−0.048*** (0.006)	−0.046*** (0.006)	−0.040*** (0.006)	−0.051*** (0.006)
Market Index FE	N	Y	Y	Y
Individual FE	N	N	Y	Y
Time FE	N	N	N	Y
N	1,845,180	1,845,180	1,845,180	1,845,180

Note: *p<0.1; **p<0.05; ***p<0.01

Empirical Model 2 (Volatility by Occupation)

- Probit

$$PR_{i,t} = \alpha + \beta_1 \cdot corr_{yr} + \beta_2 \cdot dy + \beta_3 \cdot \ln(y) + \gamma \cdot X'_{i,t} + \epsilon$$

- $PR_{i,t}$: 0/1 dummy for participation in risky asset market
- dy : log change in labor income
- $X'_{i,t}$: Demographic variables

- The wage volatility at individual level:

$$volatility_i = var_i(dy_{i,t})$$

- Occupation level:

$$volatility_{IND} = var_{IND}(dy_{IND,t})$$

Table 13: Ranking of occupations by wage volatility

IND	Description	Estimate	N
Bottom 10			
6675	Internet publishing and broadcasting	0.0506	8
3095	Commercial and service industry machinery manufacturing	0.0632	25
1670	Knitting fabric mills, and apparel knitting mills	0.0966	24
1470	Fiber, yarn, and thread mills	0.0974	25
0480	Not specified type of mining	0.0978	7
2680	Aluminum production and processing	0.1044	54
1690	Apparel accessories and other apparel manufacturing	0.1073	17
7181	Other consumer goods rental	0.1112	10
0580	Natural gas distribution	0.1179	107
3770	Sawmills and wood preservation	0.1276	139
Top 10			
8891	Personal and household goods repair and maintenance	2.7035	14
0280	Fishing, hunting and trapping	2.3729	28
5295	Musical instrument and supplies stores	1.3572	12
7071	Lessors of real estate, and offices of real estate agents and brokers	0.8310	201
8990	Nail salons and other personal care services	0.7580	157
5690	Other direct selling establishments	0.7450	102
6190	Taxi and limousine service	0.7127	167
6590	Sound recording industries	0.7000	23
8563	Promoters of events, agents and managers for public figures	0.6813	19
8970	Barber shops	0.6482	25
Financial			
6870	Banking and related activities	0.2727	1814
6970	Securities, commodities, funds, trusts, and other financial investments	0.3792	919
25th	50th	75th	
0.2292	0.2911	0.3730	

Results

Table 12: Probit Regression (Individual Level with $sd(dy) > 0$)

	<i>Bottom 10</i>		<i>Top 10</i>	
	Own STMF		Own STMF	
	(1)	(2)	(1)	(2)
$corr_{yr}$	0.338*** (0.011)	0.277*** (0.011)	-0.079*** (0.007)	-0.057*** (0.007)
dy	-0.342*** (0.011)	-0.232*** (0.011)	-0.126*** (0.003)	-0.061*** (0.003)
ln(y)	0.654*** (0.006)	0.408*** (0.008)	0.370*** (0.004)	0.182*** (0.005)
Individual FE	N	Y	N	Y
N	146,498	146,498	181,166	181,166

Table 13: Probit Regression (Industrial Level)

	<i>Bottom 10</i>		<i>Top 10</i>	
	Own STMF		Own STMF	
	(1)	(2)	(1)	(2)
$corr_{yr}$	0.022 (0.018)	0.036* (0.019)	-0.014 (0.013)	-0.025* (0.014)
dy	-0.346*** (0.026)	-0.236*** (0.026)	-0.153*** (0.010)	-0.080*** (0.010)
ln(y)	0.872*** (0.017)	0.585*** (0.020)	0.486*** (0.009)	0.269*** (0.010)
Individual FE	N	Y	N	Y
N	29,511	29,511	61,607	61,607

Empirical Model 3 (Hedging Motive)

- Probit

$$PR_{i,t} = \alpha + \beta_1 \cdot LOC + \beta_2 \cdot HIR + \beta_3 \cdot LOC \cdot HIR + \gamma \cdot X'_{i,t} + \epsilon$$

- $PR_{i,t}$: 0/1 dummy for participation in risky asset market
- LOC : 0/1 dummy for individuals with low income growth-market returns correlation
- HIR : 0/1 dummy for individuals with high income growth volatility

Results

Table 7: Probit Regression with Personal Level Risk

	<i>Dependent variable:</i>		
	Own STMF		
	(1)	(2)	(3)
<i>Low Corr</i> \times <i>High Risk</i>	0.110*** (0.006)	0.060*** (0.006)	0.026*** (0.006)
<i>LowCorr</i> ₂₅	−0.098*** (0.003)	−0.014*** (0.003)	−0.013*** (0.003)
<i>HighRisk</i> ₇₅	−0.061*** (0.003)	0.027*** (0.003)	0.030*** (0.003)
Wealth FE	N	Y	Y
Individual FE	N	N	Y
N	1,679,287	1,679,287	1,679,287

Table 8: Probit Regression with Occupation Level Risk

	<i>Dependent variable:</i>		
	Own STMF		
	(1)	(2)	(3)
<i>Low Corr</i> \times <i>High Risk</i>	0.008 (0.006)	0.025*** (0.006)	0.022*** (0.006)
<i>LowCorr</i> ₂₅	−0.065*** (0.003)	0.004 (0.003)	−0.006* (0.003)
<i>HighRisk</i> ₇₅	−0.201*** (0.003)	−0.117*** (0.003)	−0.027*** (0.003)
Wealth FE	N	Y	Y
Individual FE	N	N	Y
N	1,679,287	1,679,287	1,679,287

Results

Table 9: Probit Regression with Time Variant Occupation Level Risk Table 10: Probit Regression with Time Variant Correlation and Occupation Level Risk

	<i>Dependent variable:</i>		
	Own STMF		
	(1)	(2)	(3)
<i>Low Corr</i> \times <i>High Risk</i>	0.067*** (0.006)	0.052*** (0.006)	0.060*** (0.007)
<i>LowCorr</i> ₂₅	−0.083*** (0.003)	−0.002 (0.003)	−0.014*** (0.003)
<i>HighRisk</i> ₇₅	−0.158*** (0.003)	−0.158*** (0.003)	−0.126*** (0.003)
Wealth FE	N	Y	Y
Individual FE	N	N	Y
N	1,679,283	1,679,283	1,679,283

	<i>Dependent variable:</i>		
	Own STMF		
	(1)	(2)	(3)
<i>Low Corr</i> \times <i>High Risk</i>	0.024*** (0.008)	0.039*** (0.008)	0.024*** (0.008)
<i>LowCorr</i> ₂₅	−0.017*** (0.004)	0.068*** (0.004)	0.074*** (0.004)
<i>HighRisk</i> ₇₅	−0.152*** (0.004)	−0.165*** (0.004)	−0.127*** (0.004)
Wealth FE	N	Y	Y
Individual FE	N	N	Y
N	1,157,427	1,157,427	1,157,427

Empirical Model EX (Risky Share)

- Since the risky share information is only available after 2013, so this part of the empirical research will have less observations. The risky share is measured as:

$$Risky\ Share_{i,t} = \frac{Value\ of\ stock\ and\ mutual\ fund_{i,t}}{Value\ of\ total\ financial\ assets_{i,t}}$$

Conclusion & Next Step

- This paper has examined the relation between income volatility and portfolio decision. Theory model have shown that people can hedge their labor income risk though stock market. The evidence from the first test offers support for this theory.
 - As a second test, I divided the sample into different groups, corresponding to the volatility level. The results provide extra support to the first exercise. On top of that, I tested the hypothesis that “the higher a worker’s wage volatility, the lower his exposure to the stock market”, by showing that people facing the lowest level of income risk (both at individual and occupation level) are more “exposed” to the stock market.
-
- Use other data to support my results, i.e., NLSY97, PSID, DNB (Netherland)...
 - Robustness test with more dimensions
 - Structural LC model

Other projects (if time allows)

- I am interested in the **consumption and income dynamic**, **risk and uncertainty**, and **inequality** of households.
 1. The Disaggregated Household Inequality (Brown bag November 2022)
 - Japanese data 2000-2018
 2. Consequences of Speculation in Housing Market (10%)
 3. The Price of Uncertainty in Macroeconomic and Financial Market (1%)
- In short, what I want to do is to use **micro data** to answer **macro questions**, and to develop **computational** solution for HA model.

Plots



Table 3: Engel Curves from Family Income and Expenditure Survey

Good Category	Share (L/H)	Elasticity	
		Low Income	High Income
Food	19.3/15.7	0.69(0.05)	0.45(0.05)
Food (Outside)	4.4/4.6	0.85(0.08)	0.36(0.08)
Entertainment	2.1/2.1	2.01(0.14)	1.26(0.11)
Entertainment (Durable)	0.9/1.2	2.18(0.42)	2.23(0.31)
Book	1.2/1.2	0.45(0.08)	0.59(0.08)
Service	4.8/6.3	0.82(0.13)	0.34(0.13)
Housing (Rent/Land)	6.2/2.5	0.13(0.05)	−0.03(0.08)
Housing (Maintenance)	2.0/2.1	3.22(0.53)	1.78(0.39)
Transportation	1.7/3.0	0.67(0.20)	1.37(0.14)
Vehicle	8.8/7.5	1.82(0.22)	1.27(0.19)
Network	5.1/3.8	0.09(0.09)	−0.03(0.08)
Clothing	4.0/5.0	1.41(0.15)	1.10(0.13)
Furniture	3.3/3.2	1.50(0.16)	0.78(0.15)
Health	3.7/3.3	0.54(0.12)	0.26(0.10)
Utility	7.6/5.5	0.48(0.14)	0.44(0.13)
Education	3.9/8.1	0.47(0.26)	1.95(0.28)
Extra	7.4/6.7	0.66(0.12)	0.52(0.09)
Social	6.5/6.8	1.82(0.18)	1.41(0.16)

Appendix

- If X_t is conditionally lognormal distributed:

$$\ln E_t[X_t] = E_t[\ln(X_t)] + \frac{1}{2} \text{Var}_t[\ln(X_t)]$$

- Recall:

$$\ln E_t[C_{t+1}^{-\gamma} R_{t+1}] = \ln E_t[C_{t+1}^{-\gamma} R^f]$$

$$\begin{aligned} \Rightarrow E_t[\ln((C_{t+1})^{-\gamma}(R_{t+1}))] + \frac{1}{2} \text{Var}_t[\ln((C_{t+1})^{-\gamma}(R_{t+1}))] \\ = E_t[\ln((C_{t+1})^{-\gamma}(R^f))] + \frac{1}{2} \text{Var}_t[\ln((C_{t+1})^{-\gamma}(R^f))] \end{aligned}$$

$$\begin{aligned} \Rightarrow -\gamma E_t C_{t+1} + E_t r_{t+1} + \frac{1}{2} (\sigma_r^2 + \gamma^2 \sigma_c^2 - 2\gamma \sigma_{cy}) \\ = -\gamma E_t C_{t+1} + r^f + \frac{1}{2} (\gamma^2 \sigma_c^2) \end{aligned}$$