Estimation of Unexpected Income and Consumption Growth Innovation, and Consumption Insurance

Yuqi Qian

December 2025

Abstract

This study first regresses income and three types of constructed consumption variables on demographic variables with PSID data from 1999 to 2023. The above estimated econometric model is used to construct the unexpected income growth innovation and consumption growth innovation. It shows that the strictly defined nondurable consumption has the smallest variance among all three measurements in the majority of the time period, while the loosely defined nondurable consumption has the largest variance across the whole time period. The strictly defined consumption variance-income variance ratio obtained is consistent with the existing literature and shows excess

smoothness of consumption. We also introduce a theoretical model for the evolution of income consumption covariances. We argue that with appropriate identification methods can the exact insurance level be extracted.

1 Introduction

The idea of consumption risk sharing is related to the nature of consumer characteristics. For example, assuming a CRRA(Constant Relative Risk Averse) utility function, the agent will be risk-averse and tend to smooth consumption inter-temporally. The presence of financial markets allows the agent to save or borrow to some extent. Ideally, with the perfect risk-sharing, there will not be an idiosyncratic risk for any individual, which means the agent's consumption will not be affected by own individual income shocks at all.

However, in the real world, a representative consumer's current non-durable goods consumption is usually affected by the individual income shocks. It could result from the incompleteness of the financial market, the liquidity constraint, and some other reasons. This conflicted realistic and theoretical phenomenon makes the direct measurement of consumption insurance degree interesting.

To identify what is the proportion that individual consumption is secured through the financial market, we need to first identify the income shocks and consumption shocks. Previous studies such as Blundell, Pistaferri, and Preston (2008) uses two steps estimations by regressing non-durable goods on food consumption, then regressing imputed consumption on demographic characteristics. The shortcoming of this method is obvious that the measurement errors may play a big role because of too many estimations.

To overcome this possible problem, this study uses one step estimation by directly regressing consumption on demographic characteristics with family level PSID data from 1999 to 2023. We argue that it should have minimized the measurement errors.

Ghosh and Theloudis (2023) also use one step estimation, but the relative ratio of $\frac{Var(c_t)}{Var(y_t)}$ in their paper is quite large, around 0.8235. [1]. It is not consistent with the previous literature that consumption shows excess smoothness, for example Campbell and Deaton (1989). It is probably because they use loosely defined non-durable consumption variable. As will show later in this study, the loosely defined non-durable consumption has a largest variance among our different types of measurement.

In order to overcome this possible bias, this study constructs three different types of measurement for non-durable goods consumption: the strictly defined nondurable consumption, the non strictly defined nondurable consumption, and the total consumption.

This study can be briefly divided into three parts. In the first part, a panel random effect regression model is constructed to estimate the unexpected income and consumption growth innovation. And in the second part, we calculate the implied variance and covariance of the unexpected income and consumption growth innovation. In the last part, this study introduces a theoretical model to study the evolution of consumption and income variances and covariances.

[1] $Var(c_t) = 0.126$ and $Var(y_t) = 0.153$

2 Literature Review

Some previous literature has used at least two types of definition to define non-durable goods, for example Johnson, Parker, and Souleles (2006) and Lusardi (1996). Where Johnson, Parker, and Souleles (2006) use two different definitions to observe how households react in non-durable consumption to unexpected income changes. And Lusardi (1996) excludes the semi-nondurable goods and services to study how consumption reacts to predictible income growth.

With respect to the consumption income variance ratio, except for Campbell and Deaton (1989), which has a ratio of around 0.25, Backus, Kehoe, and Kydland (1992) concludes that from 1954 to 1989, the consumption income variance ratio for the U.S. is also around 0.24.

In addition, there is also some literature discusses the consumption inequality and income inequality. Aguiar and Bils (2015) first use direct data to construct the measure of consumption and then also use demand system corrected CE expenditure data to study on the income inequality and consumption inequality evolution across time. Both methods show significantly greater increase in consumption inequality than the direct measurement of CE household expenditure data. Krueger and Perri (2006) use CE data and document that recent increase in U.S. income inequality has not been accomplished by a corresponding increase in consumption inequality. A theoretical model is developed to allow within-group income equality affects consumption inequality in a world in which a full set of consumption claims can be traded, so that this model can align with data.

3 Unexpected Income and Consumption Estimation

3.1 Data and statistics

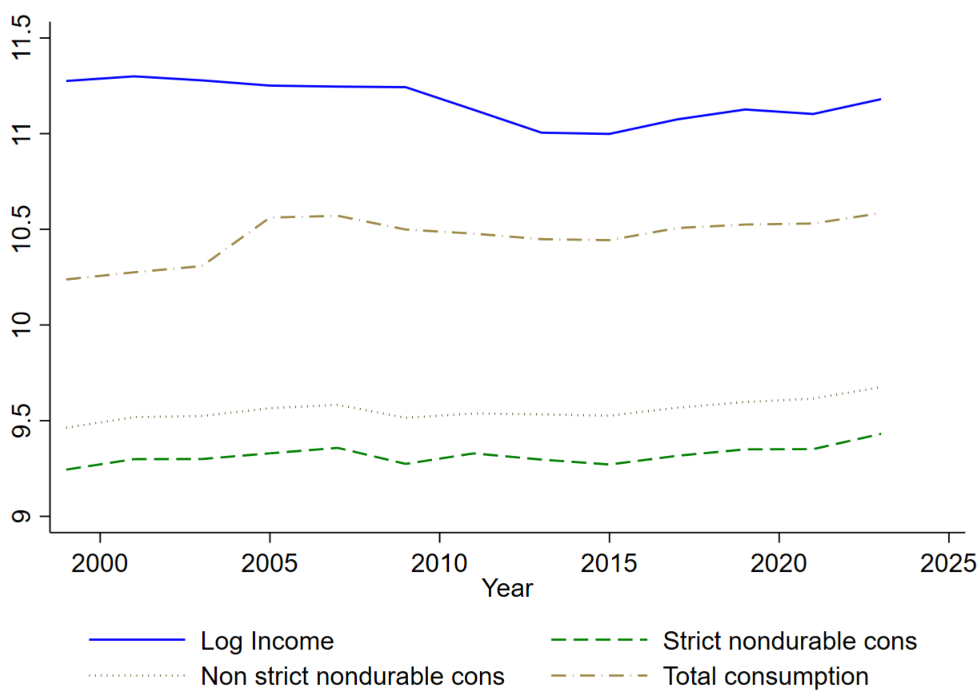
For the part 1 Panel regression, the income data and expenditure data are from Panel Study of Income Dynamics (PSID) 1999 to 2023 waves, 13 waves in total, with 2 years gap. All demographic variables are also from PSID. There are 18,890 households across 24 years. The PSID are a series of longitudinal surveys conducted by University of Michigan. This first survey was conducted in 1968. Original surveys did not ask questions about household expenditures. However, starting from 1999, dozens of expenditure questions have been asked in each survey, for example, the amount of food expenditure, the amount of different types of utility expenditure(water, electricity, and gas), housing expenditure, health care expenditure, and total expenditure.

The PSID surveys also ask different income questions, for example, taxable income, labor income, and transfer income.

Here, this study defines household income as taxable income plus transfer income. The strictly defined nondurable goods include food, utility, gasoline, parking, taking bus transport, taking train transport, taking taxi, and taking cab. In addition, this study also define a intermediate measurement between strictly defined nondurable goods consumption and total consumption. The non strictly defined durable goods expenditure include all items listed above, vehicle lease, health care, and child care. In addition, all income and consumption variables are deflated following the BEA (Bureau of Economic Analysis) data.

Here shows the trend of average household income, strictly defined nondurable goods consumption, non strictly defined nondurable goods consumption, and total consumptions:

Figure 1: Time trend of income and consumption



3.2 Estimation Method

This study can be briefly divided into two parts. In the first part, this study is going to utilize the PSID (Panel Study of Income Dynamics) data from 1999 to 2023 to reveal the unexpected part in household income, strictly defined nondurable consumption, non strictly defined nondurable consumption, and the total consumption. To do so, this study utilizes panel random effects models:

$$\begin{aligned} Depvari_{i,t} = \alpha + \beta_1 * Race_{i,t} + \beta_2 * Educ_{i,t} + \beta_3 * Famsiz_{i,t} + \beta_4 * Age_{i,t} \\ + \beta_5 * Age^2_{i,t} + \beta_6 * Unemp + + \sum_j \gamma_j * RegDum^j_{i,t} + \epsilon_{i,t}. \end{aligned} \tag{1}$$

Here, the dependent variables take 4 forms: log taxable income plus transfer income $lnY_{i,t}$, log strictly defined nondurable consumption $lnC^{st}_{i,t}$, log non strictly defined nondurable consumption $lnC^{unst}_{i,t}$, and log total consumption $lnC^{total}_{i,t}$. Control variables are mainly demographic variables. They include race $Race_{i,t}$, which is a dummy variable with value of 1 for the non-white households. $Educ_{i,t}$ is the educational year of the head of a household. $Famsiz_{i,t}$ is the family size. $Age_{i,t}$ and $Age^2_{i,t}$ capture the quadratic life cycle income curves. $Unemp$ is a continuous variable constructed from 0 to 1, which indicates the unemployed weeks in 52 weeks a year. $\sum_j \gamma_j * RegDum^j_{i,t}$

denote regional dummies. And the standard errors are clustered from the individual level.

After the panel regression, the unexpected income series $y_{i,t}$ and the unexpected consumption series $c_{i,t}$ are the residuals of the previous regression models.

3.3 Estimation Results

Before working on the unexpected income and consumption series, here are some brief results of part 1 panel regression, i.e., the demographic regression:

TABLE I.
DEMOGRAPHIC REGRESSION OF INCOME AND CONSUMPTION

Parameters	**Log income**	**Log nondurable (strict)**	**Log nondurable (non strict)**	**Total consumption**
Race (β_1)	-0.3370^{***}	-0.1727^{***}	-0.2317^{***}	-0.2291^{***}
	(0.0188)	(0.0072)	(0.0079)	(0.0086)
Education (β_2)	0.0016^{***}	0.0004^{***}	0.0005^{***}	0.0004^{**}
	(0.0004)	(0.0001)	(0.0002)	(0.0002)
Family size (β_3)	0.1648^{***}	0.1500^{***}	0.1642^{***}	0.1403^{***}
	(0.0045)	(0.0019)	(0.0022)	(0.0021)
Age (β_4)	-0.0276^{***}	0.0031^{***}	0.0075^{***}	0.0066^{***}
	(0.0009)	(0.0002)	(0.0003)	(0.0003)
Age^2 (β_5)	0.0000^{***}	-0.0000^{***}	-0.0000^{**}	-0.0000^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Unemployment (β_6)	-1.6187^{***}	-0.2311^{***}	-0.2733^{***}	-0.2577^{***}
	(0.0404)	(0.0131)	(0.0141)	(0.0130)
North East (γ_1)	0.1671^{*}	0.0726^{**}	0.1158^{***}	0.1935^{***}
	(0.0969)	(0.0311)	(0.0313)	(0.0417)
North Central (γ_2)	-0.0351	-0.0385	0.0129	0.0421
	(0.0962)	(0.0302)	(0.0305)	(0.0411)
South (γ_3)	-0.0069	0.0345	0.0834^{***}	0.1138^{***}
	(0.0934)	(0.0297)	(0.0299)	(0.0404)
West (γ_4)	0.0905	0.0293	0.0628^{**}	0.1998^{***}
	(0.0951)	(0.0303)	(0.0305)	(0.0409)

[1] Standard errors in parentheses.
[2] $* : p < 0.1, ** : p < 0.05, *** : p < 0.01.$

Non white families on average have 33.7% less income than white families after controlling other demographic factors. The figure for strictly defined nondurable goods consumption is around half of that, which is around 17.3%. With one more family member, the household is expected to have 16.5% more income, and to consume 15% more on strictly defined non durable goods. The relationship between monetary variables (income and consumption) and age is expected to be close to linear. Being unemployed significantly decrease earnings and consumption. And as expected, the figure for the effect of being unemployed on consumption is much smaller than that for the effect on income. Most of the regional dummies coefficients are not statistically significant. Except that a household living in the North East region on average spends 19.4% more overall than living in Alaska, Hawaii or abroad, and a household living in the south region overall spends 11.4% higher than living in Alaska, Hawaii or abroad. Western families spend most overall, around 20.0% higher than the baseline.

4 Unexpected Income and Consumption Autocovariance

Now, let's focus on the residuals of the panel regression. After taking the first order difference, the available data start from 2001.

As shown in table II, The income growth rate innovation variance is increasing over time from 0.8113 in 2001 to 1.5033 in 2023, which almost doubled in the last 20 years. The first order autocovariance of income is

TABLE II.
INCOME GROWTH RATE INNOVATION AUTOCOVARIANCE

	$Var(y_t)$	$Cov(\Delta y_t, \Delta y_{t+1})$	$Cov(\Delta y_t, \Delta y_{t+2})$
2001	0.8113	-0.2111	0.0211
2003	0.7660	-0.2668	-0.0435
2005	0.8541	-0.2918	-0.0092
2007	0.8854	-0.2526	-0.0079
2009	0.9105	-0.2702	0.0077
2011	1.1255	-0.3593	-0.0365
2013	1.2531	-0.3149	-0.0251
2015	1.3806	-0.3798	-0.0013
2017	1.4746	-0.3040	-0.0820
2019	1.2152	-0.2715	-0.0097
2021	1.3236	-0.3357	
2023	1.5033		

[1] Results here show the autocovariance of income growth rate innovation.

negative and significant, which shows the oscillator property of income innovation. The second order autocovariance is not significant and is quite close to zero in most years.

In table III, we compare three types of the consumption growth innovation variance, which are strictly defined nondurable consumption, non strictly defined nondurable consumption, and total consumption. The consumption variances also exhibit an increasing trend, but not as obvious as income variance.

Comparing three measurements. The strictly defined nondurable goods and non strictly defined nondurable goods both have a slightly increasing

TABLE III.
THE CONSUMPTION GROWTH RATE AUTOCOVARIANCE

	Strictly Defined		Non strictly Defined		Total consumption	
	$Var(\Delta c_t)$	$Cov(\Delta c_t, \Delta c_{t+1})$	$Var(\Delta c_t)$	$Cov(\Delta c_t, \Delta c_{t+1})$	$Var(\Delta c_t)$	$Cov(\Delta c_t, \Delta c_{t+1})$
2001	0.1698	-0.0831	0.2365	-0.0968	0.2563	-0.1076
2003	0.1983	-0.0946	0.2666	-0.0914	0.2772	-0.1020
2005	0.2188	-0.1163	0.2942	-0.1228	0.2866	-0.1159
2007	0.2265	-0.1094	0.3103	-0.1126	0.2900	-0.1139
2009	0.2517	-0.1357	0.3623	-0.1480	0.3363	-0.1387
2011	0.2694	-0.1487	0.3940	-0.1557	0.3419	-0.1294
2013	0.2716	-0.1352	0.3973	-0.1487	0.3356	-0.1153
2015	0.2560	-0.1261	0.3756	-0.1401	0.2947	-0.1027
2017	0.2485	-0.1180	0.3570	-0.1326	0.2511	-0.0691
2019	0.2289	-0.1137	0.3299	-0.1284	0.1965	-0.0777
2021	0.2601	-0.1410	0.3619	-0.1487	0.2225	-0.0866
2023	0.2538		0.3382		0.2164	

[1] Results here show the autocovariance of different types of consumption growth rate innovation.

trend in the last 20 years, while the overall consumption variances remained stable. And the non strictly defined non-durable consumption growth innovation seems to have the largest variance, followed by the total consumption, while the strictly defined nondurable consumption has the smallest variance. And although the strictly defined nondurable consumption has the smallest variance, the $cov(\Delta c_t, \Delta c_{t-1})$ remains large, which implies a strong correlation between two periods' innovation.

TABLE IV.
INTERTEMPORAL COVARIANCE BETWEEN INCOME GROWTH RATE INNOVATION AND CONSUMPTION GROWTH RATE INNOVATION

	$Cov(\Delta c_t, \Delta y_t)$	$Cov(\Delta c_t, \Delta y_{t+1})$	$Cov(\Delta c_{t+1}, \Delta y_t)$
2001	0.0589	-0.0093	-0.0116
2003	0.0613	-0.0147	-0.0171
2005	0.0728	-0.0131	-0.0176
2007	0.0768	-0.0034	-0.0224
2009	0.0771	-0.0050	-0.0249
2011	0.1294	-0.0329	-0.0475
2013	0.1385	-0.0039	-0.0427
2015	0.1157	-0.0300	-0.0248
2017	0.1090	-0.0011	-0.0137
2019	0.0779	0.0006	-0.0274
2021	0.1028	-0.0223	-0.0388
2023	0.0664	-	-

[1] Results here show the intertemporal covariance between the income growth rate innovation and strictly defined nondurable consumption growth rate innovation.

Table IV show the covariance between income growth innovation and strictly defined nondurable consumption growth innovation. The contemporaneous covariance first increased and then decreased in the last 20 years. And the intertemporal covariances are negative and non-negligible.

5 Consumption Insurance

5.1 Theory

Follow Blundell, Pistaferri, and Preston (2008), assume unexpected income follows

$$y_{i,t} = P_{i,t} + u_{i,t} \tag{2}$$

here $y_{i,t}$ denotes unexpected income (the natural log), $P_{i,t}$ denotes the permanent part, and $u_{i,t}$ denotes the transitory part.

Assume that $P_{i,t}$ follows random walk process, that is:

$$P_{i,t} = P_{i,t-1} + \eta_{i,t} \tag{3}$$

where $\eta_{i,t}$ is serially uncorrelated.

Assume that $u_{i,t}$ follows MA(q) process:

$$u_{i,t} = \epsilon_{i,t} + \sum_{k=1}^{q} \lambda_k \epsilon_{i,t-k} \tag{4}$$

where $\epsilon_{i,t}$ is also serially uncorrelated.

By subtraction, then

$$\Delta y_{i,t} = \eta_{i,t} + \Delta u_{i,t}. \tag{5}$$

To represent the correlation between income and consumption, assume that

$$\Delta c_{i,t} = \phi_{i,t}\eta_{i,t} + \psi_{i,t}\epsilon_{i,t} + \pi_{i,t}. \tag{6}$$

here $\pi_{i,t}$ is independent of $\eta_{i,t}$ and $\epsilon_{i,t}$, which represents specific consumption innovations.

The implied covariance is

$$cov(\Delta y_{i,t}, \Delta y_{i,t+s}) = \begin{cases} Var(\Delta u_{i,t}) + Var(\eta_{i,t}), s = 0 \\ Cov(\Delta u_{i,t}, \Delta u_{i,t+s}), s \neq 0 \end{cases}$$

(7)

Then

$$cov(\Delta c_{i,t}, \Delta c_{i,t+s}) = cov(\eta_{i,t}, \eta_{i,t+s}) \cdot \phi_t \cdot \phi_{t+s} + cov(\epsilon_{i,t}, \epsilon_{i,t+s}) \cdot \psi_t \cdot \psi_{t+s} + cov(\pi_{i,t}, \pi_{i,t+s}) \tag{8}$$

The covriance betwenn consumption and income is:

$$cov(\Delta c_{i,t}, \Delta y_{i,t+s}) = \begin{cases} \phi_t Var(\eta_{i,t}) + \psi_t Cov(\epsilon_{i,t}, \sum_{j=0}^{q} \theta_j \cdot \epsilon_{i,t-1}, s = 0 \\ \psi_t Cov(\epsilon_{i,t}, \Delta u_{i,t+s}), s \neq 0 \end{cases}$$

(9)

The next step will be to use some identification method to identify the parameters ϕ_t and ψ_t.

Things I am going to do

The next thing to do should be to find the appropriate identification method to calculate ϕ_t and ψ_t for different types of consumption. Previous literature suggests GMM is available. I am going to verify it next.

6 Conclusion

This paper first utilizes the PSID data from 1999 to 2023 (13 waves) to construct three types of nondurable consumption variables. By regressing income and consumption variables on demographic variables, the unexpected income and consumption growth rate innovation is constructed. The panel regression results are economically consistent.

By calculating the variance covariance matrix of income and different measures of consumption, this study finds that there is an increase in income variance in the last 20 years. The consumption variance also show similar trend, but the pattern is not as obvious as income variance. Comparing three measurement of consumption, both strictly defined nondurable goods and non strictly defined nondurable goods have an increasing trend in variance in the last 20 years, while the overall consumption(durable + nondurable) variances remained stable.

Strictly defined nondurable goods have the smallest variance before 2017, but after 2017 the overall consumption tends to have the smallest variance. Regardless of which year, the non strictly defined nondurable goods consumption tends to have the largest variance in its growth innovation.

References

Aguiar, Mark and Mark Bils (2015). "Has consumption inequality mirrored income inequality?" In: *American Economic Review* 105.9, pp. 2725–2756.

Backus, David K, Patrick J Kehoe, and Finn E Kydland (1992). "International real business cycles". In: *Journal of political Economy* 100.4, pp. 745–775.

Blundell, Richard, Luigi Pistaferri, and Ian Preston (2008). "Consumption inequality and partial insurance". In: *American Economic Review* 98.5, pp. 1887–1921.

Campbell, John and Angus Deaton (1989). "Why is consumption so smooth?" In: *The Review of Economic Studies* 56.3, pp. 357–373.

Ghosh, Anisha and Alexandros Theloudis (2023). "Consumption partial insurance in the presence of tail income risk". In: *arXiv preprint arXiv:2306.13208*.

Johnson, David S, Jonathan A Parker, and Nicholas S Souleles (2006). "Household expenditure and the income tax rebates of 2001". In: *American Economic Review* 96.5, pp. 1589–1610.

Krueger, Dirk and Fabrizio Perri (2006). "Does income inequality lead to consumption inequality? Evidence and theory". In: *The Review of Economic Studies* 73.1, pp. 163–193.

Lusardi, Annamaria (1996). “Permanent income, current income, and consumption: Evidence from two panel data sets”. In: *Journal of Business & Economic Statistics* 14.1, pp. 81–90.