A Replication of Robert Hall's Permanent Income Hypothesis Study

Collin Eldridge

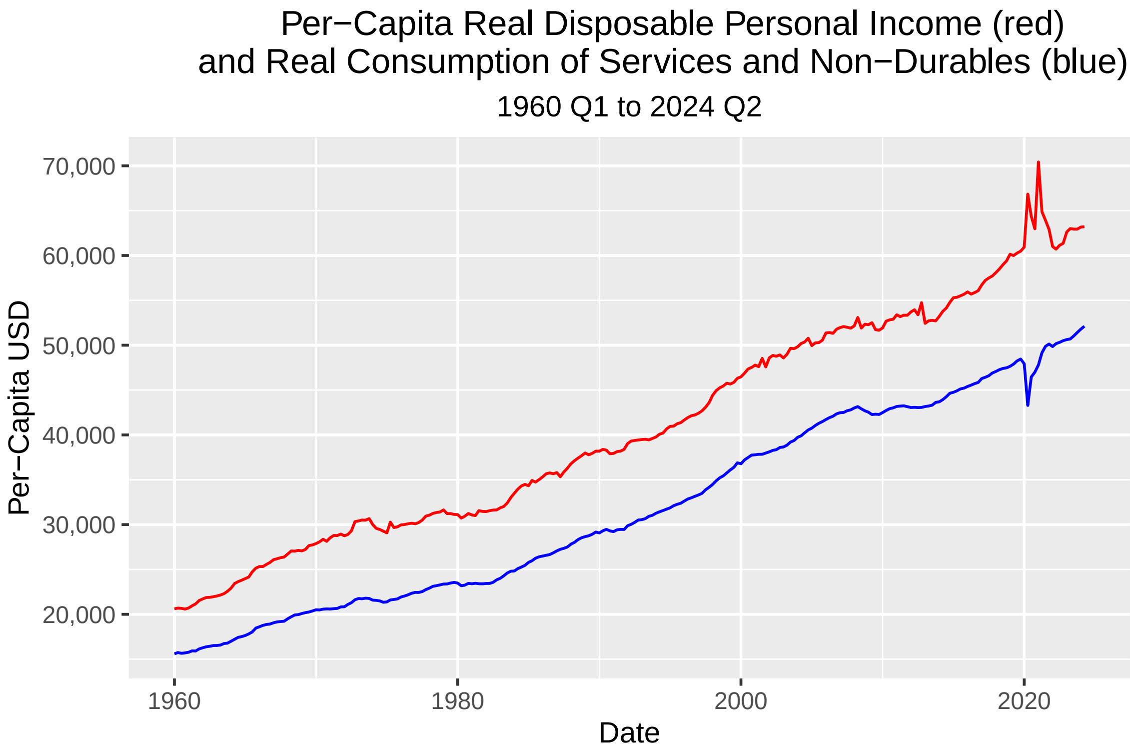
1 Introduction

For over 50 years, economists have worked to understand the consumption and savings patterns of United States citizens (cf. Modigliani & Brumberg, 1954). In a significant contribution to this work, Hall (1978) modeled optimal consumption in the current period as a function solely of consumption from the previous period, plus a mean-zero error term (making the model of consumption from this paper a random walk with a trend). Hall (1978) examines this model using consumption, income, and stock price data, and largely fails to reject it as a plausible model of optimal consumption behavior. I seek to replicate the results of Hall (1978) using updated consumption and income data. Specifically, I aim to answer the question, "After accounting for consumption in the previous period (or equivalently, consumption lagged one period), does consumption lagged two or more periods or income lagged one or more periods have any statistical or economic significance when estimating consumption in the current period?"

Section 2 contains more detailed information on the data I use, Section 3 describes the tests I use, 4 details the results of my analysis, and Section 5 summarizes my findings.

2 Data

I obtain quarterly data on consumption and income from the Federal Reserve Bank of St. Louis's Federal Reserve Economic Data (FRED) database starting from the first quarter of 1960 and extending to the second quarter of 2024; this produces a total of 258 observations. I measure consumption using the nominal personal consumption expenditures on non-durable goods (available in the "PCND" variable) and the nominal personal consumption expenditures on services (available in the "PCESV" variable). As in Hall (1978), I combine these two sources of consumption into a single variable during my analysis. I measure income using nominal disposable personal income (available in the "DSPI" variable). Though this data is gathered monthly, I use quarterly averages. I use the PCE price index (available in the "PCECTPI" variable) to convert nominal values into real values, and I use population counts (available in the "CNP16OV" variable) to convert aggregate values into per-capita values; as with the nominal disposable personal income data, I use quarterly averages rather than monthly values of the population counts. I abbreviate my transformed data on real per-capita consumption of non-durables and services to "CNS", and I abbreviate my transformed data on real per-capita disposable personal income to "DPI". These two variables are displayed in the time series plot below; I use them to perform my analysis, described in Sections 3 and 4.



3 Methods

The model for consumption that Hall (1978) proposes is

$$c_t = c_{t-1} + \epsilon, \tag{1}$$

where c_t is consumption in the current period, c_{t-1} is consumption lagged one period, and ϵ is a mean-zero error term. To evaluate the statistical and economic significance of this model, I use two models for lagged consumption and income, given in Equations 2 and 3 below.

$$c_t = \beta_0 + \beta_1 c_{t-1} + \beta_2 c_{t-2} + \beta_3 c_{t-3} + \beta_4 c_{t-4} + \epsilon \tag{2}$$

$$c_t = \beta_0 + \beta_1 c_{t-1} + \beta_2 I_{t-1} + \beta_3 I_{t-2} + \beta_4 I_{t-3} + \beta_5 I_{t-4} + \epsilon \tag{3}$$

While Equation 1 models current consumption solely as a function of consumption from the previous period, Equation 2 models current consumption as a function of consumption from the previous four periods and Equation 3 models current consumption as a function of consumption from the previous period *and* income from the previous four periods. If either of these models offer improved predictive power over the model from Hall (1978), then I will have evidence to reject Equation 1 in favor of one (or both) of the other two models. I can evaluate the predictive power of the model in Equation 2 using the following null and alternative hypothesis tests.

$$H_0 : \beta_1 = 1 \text{ and } \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_a : \beta_1 \neq 1 \text{ or at least one of } \beta_2,\ \beta_3, \text{or } \beta_4 \neq 0$$

Rejecting the null hypothesis here is equivalent to rejecting Equation 1, while failing to reject the null hypothesis implies having insufficient evidence to prefer Equation 2 over Equation 1. A similar set of null and alternative hypotheses exist to evaluate the model in Equation 3; these hypotheses are given below.

$$H_0 : \beta_1 = 1 \text{ and } \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_a : \beta_1 \neq 1 \text{ or at least one of } \beta_2,\ \beta_3,\ \beta_4, \text{or } \beta_5 \neq 0$$

As with the case of Equation 2, rejecting this new null hypothesis is equivalent to rejecting Equation 1, while failing to reject the null hypothesis implies having insufficient evidence to prefer Equation 3 over Equation 1.

Though these hypothesis tests are useful in identifying the statistical significance of the coefficients in Equations 2 and 3, I also consider the economic significance of these two models; that is, in the event I reject the null hypothesis for at least one of the two models, I must then examine the coefficient estimates themselves to understand how much more accurate the predictions of current consumption are from the chosen model relative to the model from Hall (1978).

4 Results

The results of estimating the regression models given in Equations 2 and 3 are presented in Table 1, the results of the hypothesis tests for Equation 2 are presented in Table 2, and the results of the hypothesis tests for Equation 3 are presented in Table 3. Note that all of these tables were constructed using the `stargazer` R package developed by Marek Hlavac (Hlavac, 2022).

Table 1: Regression Results for Models of Real Consumption (CNS) and Real Income (DPI)

	Dependent variable:	
	CNS in Current Period	
	‘(2)‘	‘(3)‘
CNS Lagged 1 Period	0.819***	0.914***
	(0.063)	(0.022)
CNS Lagged 2 Periods	0.146*	
	(0.082)	
CNS Lagged 3 Periods	0.042	
	(0.082)	
CNS Lagged 4 Periods	−0.005	
	(0.064)	
DPI Lagged 1 Period		0.178***
		(0.031)
DPI Lagged 2 Periods		−0.040
		(0.035)
DPI Lagged 3 Periods		−0.013
		(0.035)
DPI Lagged 4 Periods		−0.049
		(0.032)
Intercept	95.541	−254.098**
	(78.249)	(103.579)
Observations	254	254
R^2	0.999	0.999
Adjusted R^2	0.999	0.999
Residual Std. Error	391.163 (df = 249)	363.054 (df = 248)
F Statistic	46,548.680*** (df = 4; 249)	43,236.870*** (df = 5; 248)
Note:		*p<0.1; **p<0.05; ***p<0.01
		Consumption is per-capita and is of non-durables and services
		Income is per-capita disposable personal income

Table 2: F Test Results for Lagged Consumption Regression

Statistic	N	Mean	St. Dev.	Min	Max
Res.Df	2	251.000	2.828	249	253
RSS	2	38,783,433.000	967,704.700	38,099,162.000	39,467,703.000
Df	1	4.000		4	4
Sum of Sq	1	1,368,541.000		1,368,541.000	1,368,541.000
F	1	2.236		2.236	2.236
Pr(>F)	1	0.066		0.066	0.066

The results in Table 2 apply to Equation 2
The results in Table 3 apply to Equation 3

Table 3: F Test Results for Lagged Consumption and Income Regression

Statistic	N	Mean	St. Dev.	Min	Max
Res.Df	2	250.500	3.536	248	253
RSS	2	36,078,042.000	4,793,705.000	32,688,380.000	39,467,703.000
Df	1	5.000		5	5
Sum of Sq	1	6,779,323.000		6,779,323.000	6,779,323.000
F	1	10.287		10.287	10.287
Pr(>F)	1	0.000		0	0

The results in Table 2 apply to Equation 2
The results in Table 3 apply to Equation 3

Perhaps the two most relevant values in the above three tables are the p-values associated with the F test statistics presented at the bottoms of Tables 5 and 6. The p-value of about 0.066 given in Table 5 means that, at a 5% level of significance, I lack evidence to prefer the model in Equation 2 over that from Equation 1. The statistical insignificance of this model is coupled with economic insignificance visible in the relatively small estimated coefficients presented in Table 4. However, the p-value of less than 0.001 given in Table 6 means that, at a 5% level of significance, I have evidence to prefer the model in Equation 3 over that from Equation 1. Further inspection of Table 4 suggests that income lagged one period has a statistically discernible effect on current consumption, as evidence by the t-test performed on the coefficient for the "DPI Lagged 1 Period" variable. Despite this statistically significant result, however, Table 4 also reveals relatively small coefficient estimates for β_2 through β_5 in Equation 3. This means that, while my hypothesis tests do suggest that income lagged one period has a statistically discernible effect on current consumption, the model from Hall (1978) still captures the relationship between current consumption and lagged variables with reasonable accuracy.

5 Conclusion

Despite the nearly 50-year gap between my analysis and the analysis of Hall (1978), our results are nearly identical. Specifically, we both find that, while the model of current consumption given in Equation 1 can be improved by adding additional variables, those additional variables offer little economic significance even if they offer statistical significance. However, the path I take to arrive at this conclusion differs slightly from the route of Hall (1978). Perhaps the greatest difference between our respective analyses is that I do not use stock price data in my analysis and subsequently do not find evidence of a lag between changes in wealth and consumption behavior. Meanwhile, Hall (1978) finds that lagged income has neither a statistically significant nor an economically significant effect on current consumption; I find a statistically significant relationship between current consumption and lagged income, but not an economically significant one. These differences in our analyses are likely attributable to the data we use; I do not use stock price data, and Hall (1978) does not have access to the date range of data that I use. Despite these differences, the similarities shared between my analysis and that of Hall (1978) reinforce the significance of the contribution that paper has made toward the field of economics.

6 References

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, *86*(6), 971–987.

Hlavac, M. (2022). *Stargazer: Well-formatted regression and summary statistics tables*. Social Policy Institute. https://CRAN.R-project.org/package=stargazer

Modigliani, F., & Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross section data, post keynesian economics. *Franco Modigliani*, *1*, 388–436.