Journal of Macroeconomics 84 (2025) 103676

Contents lists available at ScienceDirect

Journal of Macroeconomics

journal homepage: www.elsevier.com/locate/jmacro

Research Paper

Comparing the response of different education groups to predictable changes in income

Charles Grant

Department of Economics and Finance, Brunel University, Uxbridge, London UB8 3PH, United Kingdom

ARTICLE INFO

JEL classification:
D12
D91

Keywords:
Consumption
Education
Excess sensitivity

ABSTRACT

The Permanent Income Hypothesis (PIH) can be tested using the excess-sensitivity test, which investigates whether current consumption is affected by predictable changes in income. Several studies have shown that households are overly sensitive to changes in current income, especially if they have low assets. This study shows that university educated households behave according to PIH, but that less educated households are excessively sensitive to predictable changes in income. This result is consistent with the idea that university-educated households are more patient.

1. Introduction

There has been considerable academic debate about the extent to which households respond to predictable changes in income. Several early papers, such as Attanasio and Weber (1995) or Altonji and Siow (1987) argued that households behave according to the Permanent Income Hypothesis (PIH), and do not react to predictable changes in income (this is known as the excess sensitivity test). However, most authors have argued that there is a clear difference in the behaviour of low asset households when compared to high asset households; see Zeldes (1989) or more recently Fisher et al. (2020). In this literature, the predictable change in income is constructed by using past information: current income growth is regressed against past income growth and other information available to the household in earlier time periods. The papers which reject the PIH usually attribute this either to credit constraints (in which households have limited capacity to borrow), or to buffer-stock saving in which households are impatient and risk-averse (and hence households keep a 'buffer-stock' in case they experience a negative income shock). Carroll et al. (2017), in a calibrated model, have shown that the buffer-stock model with quite low levels of heterogeneity in preferences can match the wealth distribution of US households. The empirical studies, such as Zeldes (1989), have mostly divided households by their asset levels, and compared the behaviour of high-asset and low-asset households. They argue that low asset households are likely credit-constrained and thus consume their current income; consequently they are excessively sensitive to predictable changes in their current income. However, current wealth reflects not only risk preferences, and discount rates, but also past income processes, and the past spending behaviour of the household. Those households which habitually consume their current income will end-up with low levels of assets. In this sense, the level of assets is not really truly exogenous and the regressions are reduced-form regressions.

In this paper, households will be divided according to their level of education. The level of education is almost invariably chosen in early adulthood, and does not change over time (and thus is not affected by recent income shocks). Blundell et al. (2016), using six waves of the PSID, provide some preliminary evidence that there are differences in behaviour by education. Calvet et al. (2021), using data from Swedish households, show there are differences between education groups in their wealth-to-income ratios. Moreover, since pursuing a high level of education involves reducing current earnings, such households are likely, on average, to be more patient than other households. However, since they may also have steeper expected earnings profiles, it is not immediately obvious

E-mail address: charles.grant@brunel.ac.uk.

https://doi.org/10.1016/j.jmacro.2025.103676
Received 3 April 2024; Received in revised form 20 March 2025; Accepted 22 March 2025

Available online 4 April 2025
0164-0704/

how education affects consumption. Nevertheless, this paper will show that there are important differences between education groups which may well be a consequence of their different levels of patience.

Crucial to interpreting the results in this paper is the claim, made above, that more highly educated households are more patient; several studies support this claim. Frederick (2005) surveys the early evidence that more patient households invest more years in education. Dohmen et al. (2010), looked at the relationship between patience, risk-aversion and education for a sample of German adults, while Benjamin et al. (2013), in a controlled experiment, provide evidence from a study of Chilean high-school students. The more recent literature on the experimental evidence of the relationship between patience and education is discussed in Sutter et al. (2019). Hanushek et al. (2022) argue that cross country differences in patience can explain a great deal of the differences between countries in student outcomes (based on PISA scores). Angerer et al. (2023) shows that differences in patience among primary school children predicts the schooling which these children later choose.

This paper makes several new contributions to the empirical literature. The previous literature can be criticized for placing insufficient emphasis on testing the instrument set. This paper will thoroughly and carefully test the instrument set, the results in this paper suggests weak instruments are a pervasive problem in tests of the Permanent Income Hypothesis. Unlike papers such as Attanasio and Weber (1995) or Fisher et al. (2020), it will report test results which are robust to the presence of weak instruments. This paper will use the Consumer Expenditure Survey (CEX), a survey of US households which has operated on a continuous basis since 1980. The CEX is comfortably the most comprehensive survey of household spending in the US, and hence offers advantages over the PSID (which is discussed in the data section). The previous literature has used a rather short panel of time-periods in their analysis; Zeldes (1989), for example, uses only six time periods. Jappelli and Pistaferri (2010) explains identification requires the number of time periods to be large. By using the CEX, this paper will have a much longer time-series of observations than any previous study. This paper will compare the response of households to predictable changes in income based on their level of education. This contrasts with the previous literature, which has largely looked at how assets affect consumption behaviour. The paper will show that some households do not satisfy the permanent income hypothesis. It will also show, unlike the previous literature, that there are clear differences between education groups. These differences may be due to differences in the level of patience since several of the papers discussed above argue that educated households are more patient.

The rest of the paper is organized as follows. Section 2 describes the methodology which will be applied to testing for the excess sensitivity, that is, testing whether households satisfy the Permanent Income Hypothesis. The study uses the Consumer Expenditure Survey (CEX), and Section 3 describes this dataset. The results are reported in Section 4, and the conclusions are drawn in Section 5.

2. Methodology

In the standard Euler problem, consumer i must choose consumption c in the current period t such that the consumer maximizes expected utility subject to an inter-temporal budget constraint. If the consumer can borrow and lend at the same interest rate r_t then in the consumer's optimal solution, the expectation of marginal utility is held constant. The solution, (see the summary in, Jappelli and Pistaferri, 2010), can be written formally:

$$u'(c_{it-1}) = (1-\delta)^{-1} E_{t-1}[(1+r_t)u'(c_{it})] \tag{1}$$

where $u(\cdot)$ is the utility function, δ is the discount rate and c_{it} is consumption for household i at time t. If the Permanent Income Hypothesis is true then this relationship holds exactly. To test the PIH, Attanasio and Weber (1995) explain the model can be linearized by log-linearizing the first-order condition from an inter-temporally separable optimization problem with isoelastic preferences (where Δ is the first difference).

$$\Delta \ln c_{it} = \sigma \ln(1+r_t) + \beta X_{it} + \varepsilon_{it} \tag{2}$$

Attanasio and Weber (1995) suggest adding some 'taste-shifters' X_{it} (such as family composition) which shift the marginal utility of consumption. The innovation in consumption ε_{it} is the change in consumption between time $t-1$ and time t that is not predictable given tastes X_{it} or the interest rate or any other information which is available at time $t-1$. An implication of this model is that variables known at time $t-1$ should not affect the current change in consumption $\Delta \ln c_{it}$. Thus the Permanent Income Hypothesis can be tested by adding predictable changes in income $\Delta \ln y^p_{it}$ to the regression model, as below.

$$\Delta \ln c_{it} = \alpha \Delta \ln y^p_{it} + \sigma \ln(1+r_t) + \beta X_{it} + \varepsilon_{it} \tag{3}$$

The predictable change in income $\Delta \ln y^p_{it}$ is the change in log-income $\ln y_{it}$ for household i in time t; the superscript p is used to emphasis the predictable change in income is included in this regression. The coefficient α is interpreted as the marginal propensity to consume from changes in predictable income. According to the basic version of the Permanent Income Hypothesis, households should only react to unexpected changes in their income, hence the Permanent Income Hypothesis can be tested by investigating whether α equals zero.

A key issue is how to construct the predictable change in income. Both Altonji and Siow (1987) and Attanasio and Weber (1995) used past income growth and other variables known at time $t-2$ as instruments to predict the current change in income. The validity of these instruments needs to be thoroughly tested using both a Rank-test (e.g. the F-statistic for the first stage regression of the instruments against the instrumented variable) and the Sargan test of the over-identifying instruments (earlier papers have often failed to do this). Both Altonji and Siow (1987) and Attanasio and Weber (1995) argued that when using appropriate instruments, and including a set of household characteristics X_{it} to reflect changes in tastes, the coefficient α was not identically different from zero. Some papers have instead argued that a particular one-off change in income was predictable. Parker (1999) used changes in social security payments while Souleles (1999) and Johnson et al. (2006) looked at federal tax rebates. These three papers have rejected that α equals zero in their regressions, at least for low asset households and hence they have argued that households do not follow the PIH when making their consumption decisions.

A further criticism of estimates of the marginal propensity to consume is discussed in Jappelli and Pistaferri (2010). They note that the conditional expectation of the forecast errors in Eq. (3) must be zero, that is $E_{t-1}(\varepsilon_{it}) = 0$. The empirical analog of this expectation is that the average holds over a large number of time periods rather than over a large number of households. If there are shocks to the aggregate economy (such as over the business cycle), then the cross-sectional average of the forecast errors in each time period will not converge to zero as the number of households becomes large. Much of the previous literature has used rather short time periods. For example, the sample used by Altonji and Siow (1987) includes 14 years of data while the sample used by Zeldes (1989) uses only 6 time periods. This problem has sometimes been addressed by adding time dummies in the Euler equation, but as Jappelli and Pistaferri (2010) explain, time dummies do not solve the problem if the effect of the aggregate shock differs between households.

The early studies using household data were limited in the number of time periods they could include in the analysis, but the CEX, used in this paper, now has around forty years (80 quarters) of data publicly available. Hence we argue that there are sufficient time-periods to address the criticism of Jappelli and Pistaferri (2010). This paper will make a much more thorough analysis of the choice of instruments. The longer time-series also improves the behaviour of the Sargan test, which is known to have weak power in small samples. We will find that the instruments are often weak (the instruments poorly predict the change in current income). When this happens, the CLR-test can be used to test the coefficient of the marginal propensity to consume, rather than conventional t-tests.

3. Data

Household level US data has been to investigate the PIH from the earliest studies (see Attanasio and Weber, 1995 for an explanation of why national account data is inappropriate). For the US, studies have used either the Consumer Expenditure Survey (CEX) or the Panel Survey of Income Dynamics (PSID). The measure of spending in the CEX is more comprehensive of the two surveys, and it is the survey which is used in this paper. The CEX is a representative survey of US households conducted each quarter by the Bureau of Labor Statistics. The survey was originally designed to survey US household spending in order to construct an accurate measure of household inflation. Since 1980, the survey has operated on a continuous basis, in which households provide very detailed information on all aspects of their expenditure. Each quarter, around 5000 (or 7500 after 1998) households report, in detail, their income, consumption and other particulars of their household circumstances (such as family composition, and the age and education of family members). Each household provides information on the household's spending over four successive quarters, with one-quarter of households dropping out of the survey and replaced each quarter.

The CEX dataset contains very detailed information on almost all categories of spending undertaken by the household and remains the most comprehensive survey of household spending in the US. Nevertheless, most recent studies have used the PSID rather than the CEX, despite its disadvantages. The PSID has operated continuously since 1968; households in the PSID are continuously reinterviewed, originally, on an annual basis. Up until 1999 the PSID only measured food expenditure (from which some researchers have attempted impute an overall measure of household consumption based on the household's characteristics). Since 1999 the PSID has used a broader measure of consumption than just food spending, and switched to surveying the households biannually. If this fuller measure of household spending in the PSID had been utilized, then it necessarily entails a relatively short number of time periods. Blundell et al. (2016) could only use 6 waves in their study. The new, more complete measure of consumption covers only 72 percent of the expenditure reported in the CEX. Andreski et al. (2014) argue "*the CEX remains the most comprehensive household expenditure survey*" for US households. Given the importance of having a long time-series of observations for a broad measure of consumption, this paper will use the CEX.

The sample of households included in the study excludes student households, households with multiple non-related adults, and households in which the principle earner is neither the head of household nor their spouse. The survey includes details on family size, the number of children in the household, the marital status of the household head, their gender, age and education, and information about whether the wife works. To ensure the results can be easily compared to earlier studies, the paper will include a measure of non-durable consumption, and follow the construction of non-durable consumption in Attanasio and Weber (1995). Using non-durable consumption assumes durable consumption is additively separable in utility. This measure includes food, alcohol, tobacco, fuel and utilities, transport services, personal care items, entertainment services, and housing services and maintenance. It excludes housing costs, consumer durables, cars, and health spending. The nominal values of household non-durable consumption are deflated by the consumer price index, published by the BLS (where the index equals 100 in 1983). The income measure for each household is the total after tax income, and hence it includes transfers and social payments. The analysis will also include the real interest rate in the regressions. The measure used is the three month treasury bill, adjusted for the rate of inflation.

In the CEX, each household is interviewed four times before dropping out of the survey. In order to estimate a dynamic model equation (3), we can construct a synthetic cohort based on household characteristics that do not change over time. Households in the survey report their year-of-birth and the level of education of the household head. The highest level of education of the household head is almost invariably completed before the household enters full time into the workforce, and is completed in the household head's late teens or early twenties. We define cohorts based on year-of-birth and four education groups (less than full high school, completed high school, two-year college degree, four year university degree). The year-of-birth cohorts and the education groups

Table 1
Year-of-birth cohorts with education group.

Cohort	Year-of-birth	Years included
1	1915–1925	1981–1984
2	1926–1936	1981–1995
3	1937–1947	1981–2006
4	1948–1958	1981–2017
5	1959–1969	1992–2019
6	1970–1980	2003–2019
7	1981–1991	2014–2019

Table 2
Four education groups: definition.

Group	
Ed1	Did not complete high school
Ed2	Obtained a high school certificate
Ed3	Two year college degree
Ed4	Four year university degree (or higher)

are shown in Tables 1 and 2. The number of households in each education group is roughly equal. For each cohort and each quarter, we can construct the average for each of the variables included in the regression (where we take the average of log-consumption and log-income). This means i represents the cohort rather than the individual household in Eq. (3). Table 1 reports the choice of year-of-birth cohort used in the analysis, and in which years the cohort was included in the analysis. Cohorts are included in the analysis when their age is between 29 and 65; this means, for instance, those household heads born between 1915 and 1925 are only included when observed between 1981 and 1984. The average cell-size (number of households in each cohort-quarter) is roughly 200. The regressions investigate the effect of predictable income growth on consumption growth. The size of the income changes over time was, on average, more-or-less the same for all education groups. The variance of the income changes was also similar across three of education groups, the exception was that the variance of log-income growth was much higher for the lowest education group (around 0.11) compared to the other education groups (around 0.06).

4. Results

The basic regression results are reported in Table 3. In each regression the change in log-consumption is regressed against the change in log-income. The real interest rate and a set of quarterly dummies are also included in each regression. Some regressions also include household characteristics as additional explanatory variables. These additional variables are interpreted as taste-shifters, which change the desired level of consumption independently of any income change. In the regression, both the change in income, and the interest rate are instrumented using age, age-squared, and the second, third and fourth lag of both the interest rate and income growth as instrumental variables (see, Grant, 2024, for a detailed discussion of how to choose the instruments when testing the Permanent Income Hypothesis). Throughout the regressions are estimated using a simple IV regression. In this framework, if predictable changes in income cause changes in consumption then we can reject the permanent income hypothesis (this formulation of the test is known as the excess sensitivity test).

The first set of results are reported in Table 3. In columns (1)–(3) the effect of the income growth (instrumented) on consumption growth is imposed to be the same for all four education groups. The results in column (1) do not include household characteristics in the regression, and the predictable change in income is significant in this regression. This result is in line with Attanasio and Weber (1995). The results in column (2) include the change in family-size as an explanatory factor, and this variable is significant in the regression. When this variable is included in the regression, income growth remains significant. In column (3), more household characteristics are included in the regression. Again, the predictable change in income remains significant (but only at the 10 percent significance level). The fact that income growth remains significant when household characteristics are included in the regression is a rejection of the permanent income hypothesis. However, the instruments need to be tested using both the Sargan test (on the over-identifying instruments) and the Rank-test.

The problems associated with weak instruments are discussed in Staiger and Stock (1997). Weak instruments occur when the instruments predict the instrumented variable, but with a low R-squared value. If there is a problem with weak instruments, the regression output is likely to mis-report the standard-errors in the regression thus inference is likely to be unreliable. If only one variable in the regression is instrumented then the F-statistic for the first stage regression of the instruments against the instrumented variable can be reported. Critical values for a variety of scenarios are reported in Stock and Yogo (2005), but it is also very common to use the rule-of-thumb critical value of 10 (which is known to work reasonably well). When there are several variables being instrumented Sanderson and Windmeijer (2016) showed that the F-statistic can be replaced by the Cragg-Donald statistic (the minimum eigenvalue rank test statistic), see Cragg and Donald (1993). Stock and Yogo (2005) have also published critical values for this scenario.

In column (1) of Table 3, the Rank-test (the Cragg and Donald minimum eigenvalue) gives a value of 10.63. Although this suggests the results are not affected by a weak instrument problem, screening based on the result of the Rank test is known to

Table 3
The pooled Euler Regression with education groups.

	(1)	(2)	(3)
Δy_{it}	0.393**	0.290**	0.259*
	(0.081)	(0.091)	(0.105)
interest rate	−0.017	−0.015	−0.013
	(0.062)	(0.060)	(0.059)
Δ family-size		0.329**	0.194**
		(0.060)	(0.054)
Δ wife-work			0.016
			(0.055)
Δ couple			0.146
			(0.085)
Δ children			0.107*
			(0.046)
$Wald$	23.51	10.19	6.11
(p-val)	(0.000)	(0.006)	(0.047)
CLR	16.59	7.46	4.46
(p-val)	(0.000)	(0.006)	(0.033)
Rank Test	10.63	8.56	7.42
Sargan	12.85	11.10	10.96
(p-val)	(0.045)	(0.085)	(0.089)
R^2	0.30	0.36	0.37
N	2,100	2,100	2,100
Instruments: age, age-sq, r_{t-2}, r_{t-3}, r_{t-4}, Δy_{it-2}, Δy_{it-3}, Δy_{it-4}.			

Notes: Standard errors in parenthesis. Income and consumption are measured in logs. All regressions included quarter dummies; * denotes significant at the 5 percent significance level while ** denotes significant at the 1 percent significance level. The Wald test, CLR test and AR test are all tests of the variable Δy_{it}. Although both the CLR test and the AR test allow for weak instruments, the CLR test is preferred. The Rank test is the Cragg and Donald minimum eigenvalue, for which Stock and Yogo (2005) have published critical values.

lead to poor behaviour of the test of the variable of interest; the CLR test should still reported (this test is efficient under strong instruments, and there is no reason to prefer the results of the t-test). The Sargan test rejects the over-identifying instruments, as a result no inference can be drawn from the results in column (1). (Note that the CLR test has weak power to reject the null hypothesis when the Sargan test rejects the over-identifying instruments). The results in column (2) include the change in family-size as an explanatory variable. This time the Sargan test is passed (at the 5 percent level). The Rank-test reports a value of 8.56, which is below the Stock and Yogo critical value (and below the rule-of-thumb value of 10). This means that there is a weak instrument problem in this regression. When the instruments are weak (as occurs in this case) Andrews et al. (2006) argue that the Wald test is biased, and will typically over-reject the null hypothesis when instruments are weak. Any inference based in this test will be unsafe (results for the Wald test are nevertheless reported in the table).

Although two alternative tests have been proposed in the literature when the instruments are suspected to be weak: the AR-test and the CLR-test, Andrews, Andrews et al. (2006) argue that the CLR-test proposed in Moreira (2003) should be used in the over-identified case. Consequently, we will only discuss results for the CLR-test. In the second column of Table 3 reports results for the CLR-test, in this case a test of whether the growth in log-income is zero. The null hypothesis is clearly rejected at the 5 percent level, and we can conclude that predictable income changes predict the change in consumption. That is, the results reject the Permanent Income Hypothesis. Column (3) adds some additional explanatory variables to the regression, interpreted as taste-shifters. The Sargan test is passed (at the 5 percent level) but the Rank-test reports a value of 7.42, which is below the Stock and Yogo critical value (and below the rule-of-thumb value of 10). Again there is a problem of weak instruments. The effect of the change in log-income is again investigated using the CLR-test, which rejects the null hypothesis that the effect is zero. The third column of Table 3 adds a full set of household characteristics to the regression. The table shows that the Sargan test is passed (at the 5 percent level) but that the Rank-test is below the critical value. To test the null hypothesis that the effect of the predictable change in income is zero, the table reports the CLR-test. The null hypothesis is rejected (using the 5 percent significance level). Hence we can conclude that predictable changes in income predicts changes in consumption, and the results mean we can reject the Permanent Income Hypothesis.

4.1. Differences between education groups

In the results reported in Table 3, the effect of income growth on consumption growth was the same for all education groups. However, it is possible that there are differences between low education groups and high education groups. Carroll et al. (2017) argued that households differ by their level of patience, and they believe the Permanent Income Hypothesis may well describe the behaviour of more patient households. We believe that more patient households are likely to invest more in their education. Frederick (2005), Dohmen et al. (2010), and Angerer et al. (2023) have all provided evidence that more highly educated households are more patient. As a result, this paper runs some regressions in which the effect of predictable income growth on consumption growth is

Table 4
Euler Regression for each education group.

	(4)	(5)	(6)
Δy_{it} × Ed1	0.501**	0.424**	0.451*
	(0.162)	(0.161)	(0.179)
Δy_{it} × Ed2	0.000	−0.013	0.010
	(0.252)	(0.239)	(0.254)
Δy_{it} × Ed3	0.553**	0.458**	0.470**
	(0.178)	(0.172)	(0.172)
Δy_{it} × Ed4	0.320	0.130	0.109
	(0.236)	(0.231)	(0.234)
interest rate	−0.011	−0.009	−0.004
	(0.065)	(0.062)	(0.062)
Δ family-size		0.322**	0.193**
		(0.066)	(0.057)
Δ wife-work			−0.028
			(0.069)
Δ couple			0.090
			(0.096)
Δ children			0.126*
			(0.051)
$Wald$	27.29	14.69	13.39
(p-val)	(0.000)	(0.011)	(0.020)
CLR	28.46	15.48	15.70
(p-val)	(0.000)	(0.005)	(0.004)
Rank Test	2.48	2.38	2.32
Sargan	15.14	13.95	13.40
(p-val)	(0.233)	(0.303)	(0.340)
R^2	0.25	0.32	0.32
N	2,100	2,100	2,100

Instruments: age, age-sq, r_{t-2}, r_{t-3}, r_{t-4}, Δy_{it-2}, Δy_{it-3}, Δy_{it-4}.

Notes: Standard errors in parenthesis. Income and consumption are measured in logs. All regressions included quarter dummies; * denotes significant at the 5 percent significance level while ** denotes significant at the 1 percent significance level. The Wald test, CLR test and AR test are all tests of the variable Δy_{it}. Although both the CLR test and the AR test allow for weak instruments, the CLR test is preferred. The Rank test is the Cragg and Donald minimum eigenvalue, for which Stock and Yogo (2005) have published critical values.

allowed to differ between the four education groups reported in Table 2 (less than full high school, completed high school, two-year college degree, four year university degree). These results are reported in Table 4.

In column (4), which does not include household controls, the t-test shows that income growth is significant for the first education group (those who did not complete high school) and the third education group (those with some college but not a full 4-year college degree), but is not significant for the other education groups. The results are almost exactly the same when the change in family-size is included in the regression in column (5) or when all the household controls are included, in column (6). The Sargan test never rejects the over-identifying instruments in these three regressions. The Rank test (the Cragg and Donald minimum eigenvalue) reports quite low statistics: for example, the rank statistic is 2.48 in column (4), suggesting the instruments are weak. The CLR test (a joint test that income growth is zero for all four education groups), which is robust to weak instruments, reports are value of 28.46 in column (4), which is significant at the 5 percent level (and also at the 1 percent level). The CLR test for the four education groups is also significant in column (5), which includes the change in family-size; and in column (6) in the table, where all the household characteristics are included in the regression. These results suggest that we can reject that the Permanent Income Hypothesis holds for all households, but also that there might some education groups for which the Permanent Income Hypothesis does hold.

4.2. Separate regressions for each education group

The analysis in Table 4 suggests that more educated households may behave differently from less well educated households. To investigate this further, Table 5 reports separate regressions for each education level. Attanasio and Weber (1995) argue that tests for the Permanent Income Hypothesis need to account for taste-shifters. Hence all four regressions include the change in family-size in the regression (results are very similar if all the household controls are included). The table shows, using the t-test, that income growth is significant for the lowest education group, in the column labelled column (7). And it is also significant for the ‘some college’ education group, in the column labelled column (9), but income growth is not significant in the column (8) and column (10).

In each of the four regressions the Sargan test for the over-identifying instruments is passed. However, the Rank test (the Cragg and Donald minimum eigenvalue) is always well below the rule-of-thumb value of 10 (and also the Stock and Yogo critical value), meaning that in each of these regressions there is a weak instrument problem. The table reports the CLR test for income growth, which is robust to weak instruments. The CLR test is significant at 5 percent in both column (7) and column (9). Column (8) reports the regression for those who have completed high school. While the CLR test is not statistically significant in column (8), the very

Table 5
Separate Euler Regression for each education group.

	(7)	(8)	(9)	(10)
Δy_{it} × Ed1	0.474*			
	(0.190)			
Δy_{it} × Ed2		0.198		
		(0.149)		
Δy_{it} × Ed3			0.362**	
			(0.137)	
Δy_{it} × Ed4				0.064
				(0.157)
interest rate	−0.063	0.021	−0.032	0.024
	(0.156)	(0.102)	(0.110)	(0.105)
Δ family-size	0.141	0.448**	0.315**	0.455**
	(0.142)	(0.102)	(0.085)	(0.102)
$Wald$	6.22	1.77	6.96	0.16
(p-val)	(0.044)	(0.413)	(0.030)	(0.921)
CLR	5.34	0.99	4.85	0.04
(p-val)	(0.020)	(0.325)	(0.027)	(0.844)
Rank Test	1.65	3.25	4.38	4.28
Sargan	2.86	5.31	4.60	8.55
(p-val)	(0.826)	(0.504)	(0.595)	(0.200)
R^2	0.13	0.40	0.41	0.53
N	525	525	525	525
Instruments: age, age-sq, r_{t-2}, r_{t-3}, r_{t-4}, Δy_{it-2}, Δy_{it-3}, Δy_{it-4}.				

Notes: Standard errors in parenthesis. Income and consumption are measured in logs. All regressions included quarter dummies; * denotes significant at the 5 percent significance level while ** denotes significant at the 1 percent significance level. The Wald test, CLR test and AR test are all tests of the variable Δy_{it}. Although both the CLR test and the AR test allow for weak instruments, the CLR test is preferred. The Rank test is the Cragg and Donald minimum eigenvalue, for which Stock and Yogo (2005) have published critical values.

weak instruments (the Rank-test reports a value of 3.25) results in very wide confidence bounds, such that the AR test reports a confidence bound which includes both ‘zero’ and ‘one’. That is, the fact that the instruments poorly predict the instrumented variables means the test has extremely low power, and is not able to reject the null hypothesis. As a result we do not believe the results can provide any inference about whether these households (those households who have obtained a high-school certificate) satisfy the Permanent Income Hypothesis. In the last column, those with a full four-year college degree, the Sargan test is passed, the instruments predict the instrumented variables (although there is a weak instrument problem); the results show that income growth is not significant in the CLR test. The confidence bounds are much narrower than those for column (8), such that high estimates of the coefficient (as in column 7) would have been significant. The evidence suggests that these educated households behave as predicted by the Permanent Income Hypothesis.

5. Conclusion

Papers which have estimated the marginal propensity to consume of US households have often implemented the excess sensitivity test to examine whether households behave according to the Permanent Income Hypothesis. While Attanasio and Weber (1995) argued that when using household level data and allowing for taste-shifters the Permanent Income Hypothesis held, other studies, such as Zeldes (1989), Souleles (1999), and more recently Fisher et al. (2020), have clearly rejected the hypothesis. These papers have found that low asset households respond to predictable changes in income, which clearly violates the Permanent Income Hypothesis.

A problem with much of the previous literature is that the instruments have been inadequately tested. In this paper we report both the Sargan test of the over-identifying instruments and the Rank-test for weak instruments (the Cragg and Donald minimum eigenvalue). In almost all regressions, the instruments were found to be weak. Andrews et al. (2006) argue that when instruments are weak, the CLR-test should be used to test the variables of interest (in this case to test the null hypothesis that predictable changes in income does not predict changes in consumption). This paper finds that we can reject this null hypothesis and thus find the Permanent Income Hypothesis does not hold.

Although the results show that US households do not, as a whole, follow the PIH, there may be some households for which the PIH is a good description of their behaviour. Studies by Zeldes (1989) and, Souleles (1999) argued that low asset households did not follow the Permanent Income Hypothesis. However, in a multi-period framework, the level of assets is chosen by the household. For example, households which habitually spend their current income will end up with few assets. Note that in this example, it is the underlying impatience of these households which means they do not behave according to the Permanent Income Hypothesis and have few assets. Carroll et al. (2017) argued that households differ by their level of patience: and believe the Permanent Income Hypothesis may well describe the behaviour of more patient households. Although patience itself is not directly observed, higher levels of education are associated with more patience (since education entails deferred consumption). Frederick (2005), Dohmen et al. (2010), and Angerer et al. (2023) have all provided evidence for this relationship between education and patience. Moreover, unlike the level of assets, education does not change over the adult life-cycle; households overwhelmingly have completed their education in their early twenties (or earlier). This paper has divided the US population into four education groups (less than high school, high school certificate, two year college degree, four year university degree).

By running a separate regression for each education group, the paper has found that lowest education groups do not follow the Permanent Income Hypothesis. Similarly the results showed that household with less than 4 years of university also did not satisfy the Permanent Income Hypothesis. For the group who completed high school, the 95 percent confidence interval was extremely large, and as a direct consequence, the null hypothesis (that predictable income was zero) could not be rejected. Nevertheless, the point estimates for the three lowest educated groups (their marginal propensity to consume) were found to be between 0.2 and 0.5. The results for university graduates (those who have competed a four year university degree) were very different to the estimates for the less educated households. For university graduates, the predictable change in income does not predict current consumption growth, and these educated households have a marginal propensity to consume remarkably similar to the value of about 0.05 implied by the certainty-equivalent permanent income hypothesis.

The paper argues that educated households are more patient, which is why they are more likely to invest in education. Their greater patience means they are better able to smooth consumption over time (assuming that these differences are not due to systematic differences in the income process across education groups). This is likely to have implications for public policy, both for short-term policy choices (such as a temporary tax-cut), and over longer time-frames (such as saving for pension provision).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmacro.2025.103676.

Data availability

Data will be made available on request.

References

Altonji, J.G., Siow, A., 1987. Testing the response of consumption to income changes with (noisy) panel data. Q. J. Econ. 102, 293–328.
Andreski, P., Li, G., Samancioglu, M.Z., Schoeni, R., 2014. Estimates of annual consumption expenditures and its major components in the PSID in comparison to the CE. Am. Econ. Rev. 104 (5), 132–135.
Andrews, D.W.K., Moreira, M.J., Stock, J.H., 2006. Optimal two-sided invariant similar tests for instrumental variables regression. Econometrica 74 (3), 715–752.
Angerer, S., Bolvashenkova, J., Glätzle-Rützler, D., Lergetporer, P., Sutter, M., 2023. Children's patience and school-track choices several years later: Linking experimental and field data. J. Public Econ. 220, 104837.
Attanasio, O., Weber, G., 1995. Is consumption growth consistent with intertemporal optimization? Evidence from the consumer expenditure survey. J. Polit. Econ. 106, 1121–1157.
Benjamin, D.J., Brown, S.A., Shapiro, J.M., 2013. Who is 'behavioral'? Cognitive ability and anomalous preferences. J. Eur. Econ. Assoc. 11 (6), 1231–1255.
Blundell, R., Pistaferri, L., Saporta-Eksten, I., 2016. Consumption inequality and family labor supply. Am. Econ. Rev. 106 (2), 387–435.
Calvet, L.E., Campbell, J.Y., Gomes, F., Sodini, P., 2021. The Cross-Section of Household Preferences. Technical Report, National Bureau of Economic Research.
Carroll, C.D., Slacalek, J., Tokuoka, K., White, M.N., 2017. The distribution of wealth and the marginal propensity to consume. Quantative Econ. 8 (3), 977–1020.
Cragg, J.G., Donald, S.G., 1993. Testing identifiability and specification in instrumental variable models. Econometric Theory 9 (2), 222–240.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? Am. Econ. Rev. 100 (3), 1238–1260.
Fisher, J.D., Johnson, D.S., Smeeding, T.M., Thompson, J.P., 2020. Estimating the marginal propensity to consume using the distributions of income, consumption, and wealth. J. Macroecon. 65, 1–14.
Frederick, S., 2005. Cognitive reflection and decision making. J. Econ. Perspect. 19 (4), 25–42.
Grant, C., 2024. A new investigation into whether households are excessively sensitive to predictable changes in income. Empir. Econ. 1–16.
Hanushek, E.A., Kinne, L., Lergetporer, P., Woessmann, L., 2022. Patience, risk-taking, and human capital investment across countries. Econ. J. 132 (646), 2290–2307.
Jappelli, T., Pistaferri, L., 2010. The consumption response to income changes. Annu. Rev. Econ. 2, 479–506.
Johnson, D.S., Parker, J., Souleles, N.S., 2006. Household expenditure and the income tax rebates of 2001. Am. Econ. Rev. 96, 1589–1610.
Moreira, M.J., 2003. A conditional likelihood ratio test for structural models. Econometrica 71, 1027–1048.
Parker, J., 1999. The reaction of household consumption to predictable changes in social security taxes. Am. Econ. Rev. 89, 959–973.
Sanderson, E., Windmeijer, F., 2016. A weak instrument F-test in linear IV models with multiple endogenous variables. J. Econometrics 190, 212–221.
Souleles, N.S., 1999. The response of household consumption to income tax refunds. Am. Econ. Rev. 89, 947–958.
Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.
Stock, J., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Andrews, D.W. (Ed.), Identification and Inference for Econometric Models. Cambridge University Press, New York, pp. 80–108.
Sutter, M., Zoller, C., Glätzle-Rützler, D., 2019. Economic behavior of children and adolescents–A first survey of experimental economics results. Eur. Econ. Rev. 111, 98–121.
Zeldes, S.P., 1989. Consumption and liquidity constraints: An empirical investigation. J. Polit. Econ. 97, 305–346.