Recent studies find that consumption is excessively sensitive to income. These studies assume that income is stationary around a deterministic trend. The data, however, do not reject the hypothesis that disposable personal income is a random walk with drift. If income is indeed a random walk, then the standard testing procedure is greatly biased toward finding excess sensitivity. Moreover, if income follows either a more general non-stationary process or a borderline stationary process, this procedure is also seriously biased.

1. Introduction

Since the seminal work of Hall (1978) and Sargent (1978), much research has attempted to test the permanent income hypothesis under the assumption of rational expectations. Some of this research produces evidence contradicting the theory. Most notably, Flavin (1981, 1984) and Bernanke (1985) report that consumption is 'excessively sensitive' to disposable income. To many economists, this excess sensitivity suggests that liquidity constraints are an important determinant of consumer spending.

We argue that the standard test is biased toward finding excess sensitivity. The recent work of Nelson and Plosser (1982) suggests that unit roots are common in economic time series. Dickey and Fuller (1981) and Nelson and Kang (1981, 1984) show that conventional test statistics are inadequate in the presence of unit roots. In particular, a researcher is likely to find a deterministic trend where none exists. Moreover, inappropriate detrending can produce spurious cycles even if the underlying data have no cyclical properties. These
findings motivate our re-examination of tests of the permanent income hypothesis.

The intuition behind our analysis is as follows: Suppose income follows a random walk, so that permanent income equals current income. Suppose also that the permanent income hypothesis is true, and thus consumption equals income. Since the series contain unit roots, standard testing procedures are invalid. Moreover, if both consumption and income are (inappropriately) detrended, then both series will exhibit spurious cycles. Since consumption tracks income perfectly over these seemingly transitory cycles, the econometrician will erroneously conclude that consumption is excessively sensitive to contemporaneous income.

In the second section of this paper, we examine the stochastic properties of disposable personal income per capita over the post-war period. We conclude that income is, or is very close to, a random walk. Thus, shocks to income are essentially permanent.

In the third section, we perform Monte Carlo experiments to demonstrate the bias in the standard test procedure. We begin by assuming that income follows a random walk. We show how use of the conventional test, which assumes income is stationary around a deterministic trend, leads to incorrect inferences. We also generalize the Monte Carlo experiment by allowing income to be subject to both permanent and transitory shocks. Even if the transitory component of income is substantial, the bias toward rejection in the standard test is large. Finally, we consider the case in which income is stationary but strongly autoregressive. We show that the conventional test is again biased toward rejection.

In our final section, we discuss the implications of our results for future research. Since shocks to aggregate income are almost completely permanent, it appears that aggregate data have little power to distinguish among alternative theories of consumption. More generally, our results provide a vivid and concrete example motivated by economic theory of the pitfalls inherent in the conventional and routine use of detrended data.

2. Disposable income is a random walk

In this section we show that quarterly disposable personal income per capita over the past twenty-five years is well approximated as a random walk. In other words, shocks to income are permanent. We use standard Box–Jenkins (1976) tools, which are essentially atheoretical, to identify the stochastic process. To reduce the problem of time aggregation discussed by Working (1960), we examine income in the first month of each quarter.

In table 1, we present the autocorrelations of the level of real disposable personal income per capita \( Y_t \). The autocorrelations begin at one and decline
Table 1


<table>
<thead>
<tr>
<th>Levels</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple</strong></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>1.00</td>
</tr>
<tr>
<td>Second</td>
<td>1.00</td>
</tr>
<tr>
<td>Third</td>
<td>1.00</td>
</tr>
<tr>
<td>Fourth</td>
<td>0.99</td>
</tr>
<tr>
<td>Fifth</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>Partial</strong></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>1.00</td>
</tr>
<tr>
<td>Second</td>
<td>-0.06</td>
</tr>
<tr>
<td>Third</td>
<td>-0.07</td>
</tr>
<tr>
<td>Fourth</td>
<td>-0.21</td>
</tr>
<tr>
<td>Fifth</td>
<td>0.21</td>
</tr>
</tbody>
</table>

only slightly. The partial autocorrelations decline to zero after the first one. A random walk has these properties.

We also present the autocorrelations of the change in disposable income per capita. These autocorrelations are all small. The Q-statistic for the null hypothesis that the first twenty-four autocorrelations are zero is 29.5, which is not significant at even the ten percent level. Thus, the change in disposable income appears to be approximately white noise.

Finally, we present a simple univariate forecasting equation. When we regress $Y_t - Y_{t-1}$ on $Y_{t-1}$, $Y_{t-1} - Y_{t-2}$, and a time trend, we obtain, with standard errors in parentheses,

$$Y_t - Y_{t-1} = 169.8 - 0.11 Y_{t-1} + 0.10 (Y_{t-1} - Y_{t-2}) + 2.4 \text{ Time},$$

$$R^2 = 0.06, \quad \bar{R}^2 = 0.03, \quad D.W. = 1.99, \quad s.e. = 34.5.$$  

Although some of the coefficients appear significant using conventional significance levels, one should not draw this conclusion. As Dickey and Fuller (1981) show, if $Y_t$ in fact follows a random walk, the conventional tests are inadequate. In particular, a $t$-statistic of approximately 3.1 is required to reject, using a two-tailed test at the five percent level, the null hypothesis that the coefficient on the trend is zero. The adjusted $R$-squared clearly suggests that the forecastability of income changes from lagged values and from trend is negligible.
Thus, from this examination of the data, we conclude that income is approximately a random walk with drift:

\[ Y_t = 20.8 + Y_{t-1}, \quad s.e. = 35.0. \]  

(3.5)

The conventional F-statistic for the null hypothesis that income is a random walk with drift against the alternative in eq. (1) is 1.99. This statistic is not significant with the conventional critical value. As Dickey and Fuller make clear, however, an even higher critical value is in fact required. Using the correct small sample critical value would thus also fail to reject the random walk specification.

3. Testing the permanent income hypothesis

In this section we examine tests of the permanent income hypothesis. For expositional purposes, we first consider an economy in which income follows exactly a random walk. We then consider the case in which income follows a more general non-stationary process. Finally, we show that the problems in the non-stationary case also arise when income follows a borderline stationary process.

3.1. Random walk income

We begin by supposing that income follows exactly a random walk with drift. We then examine the standard tests of the permanent income hypothesis performed on detrended data. Even when the permanent income hypothesis is true, these tests often yield rejections of the theory that appear both statistically and economically significant.

If income follows a random walk, then permanent income equals current income. If, in addition, the permanent income hypothesis is true, then consumption (C,) also equals current income. Consider the now standard test of the theory, derived first by Hall (1978) and examined more fully by Flavin (1981), that changes in consumption are not forecastable. We might regress the change in consumption on lagged income:

\[ C_t - C_{t-1} = \mu + \pi Y_{t-1} + \nu_t. \]  

(3)

Since \( C_t = Y_t \) in our fictitious economy, the test that \( \pi = 0 \) is just the test that consumption follows a random walk.

If income follows a random walk with drift, then permanent income equals current income plus a constant. This constant is inessential and thus we do not discuss it.
Suppose that, in an economy in which consumption and income are generated by eq. (2), an econometrician first detrended the data on both consumption and income and then performed the regression test (3). What would he find? We answer this question by a Monte Carlo experiment assuming 25 years of quarterly data and normal errors. Based on 1000 replications, we find the median value of his t-statistic would be -2.21. Using the conventional ‘five percent’ critical value of 1.96, the econometrician would reject the null hypothesis 61 percent of the time. Using the conventional ‘one percent’ critical value of 2.58, he would reject 30 percent of the time.

Would the econometrician conclude that the rejection was economically significant? Consider Flavin’s (1981) measure of excess sensitivity. In the first-order case, the just-identified system of equations is

\[ Y_t = \delta + \rho Y_{t-1} + \epsilon_t, \]

\[ C_t - C_{t-1} = \alpha + \beta (E_{t-1} Y_t - Y_{t-1}) + \nu_t, \]

where \( \beta \) is the measure of the excess sensitivity. According to the theory, predictable changes in income should not be related to changes in consumption; that is, \( \beta = 0 \). Flavin estimates \( \beta \) by noting that

\[ E_{t-1} Y_t = \delta + \rho Y_{t-1}. \]

Thus, eqs. (3) and (4) are the reduced form system and

\[ \pi = \beta (\rho - 1). \]

The test for excess sensitivity is thus just the test that \( \pi = 0 \) in eq. (3). The excess sensitivity parameter, \( \beta \), is recovered from the estimates of \( \pi \) and \( \rho \).

As already noted, the econometrician would reject the hypothesis that \( \pi = 0 \) much too frequently. That is, the hypothesis of no excess sensitivity would be rejected 61 percent of the time using a test with a nominal size of five percent. The estimate of \( \beta \) would also be large. Since \( C_r = Y_r \), eqs. (3) and (4) are in fact the same, except for normalization. In particular, the estimate of \( \rho \) is always one plus the estimate of \( \pi \). Therefore, the econometrician would always infer that \( \beta = 1 \), that is, complete excess sensitivity. Hence, he would always conclude that the extent of excess sensitivity is economically significant.

### 3.2. More general non-stationarity

While our example above concerns an income series that follows exactly a random walk, the same issues arise for any non-stationary series. Here we

---

4By the Frisch–Waugh (1933) theorem, first detrending the data is numerically equivalent to estimating the trend simultaneously with the other parameters.
consider the case in which income is the sum of a permanent (random walk) component and a transitory (white noise) component. We show how the extent of the bias depends on the relative importance of the two components.

The sum of a random walk and white noise is an IMA(1, 1) process. That is, income is described as

\[ Y_t = Y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \]  

If \( \theta = 0 \), then income follows a random walk. If \( \theta = 1 \), then income is white noise. For any \( \theta \) less than one, income is non-stationary.

Suppose consumption is set according to the permanent income hypothesis:

\[ c_t = \left( r + w + \alpha \right) Y_t. \]  

Non-human wealth evolves as

\[ H_{t+1} = (1 + r) \left[ W_t + Y_t - C_t \right]. \]  

Finally, we set initial non-human wealth and initial income equal to zero and the quarterly real interest rate to 1.25 percent.

Suppose our econometrician obtained data generated by this economy, detrended it, and then performed regression (3). Table 2 presents the results of this Monte Carlo experiment for various values of \( \theta \). The bias is largest for values of \( \theta \) close to zero. Even at \( \theta = 0.50 \), the median value of the test statistic is \(-1.88\). In 46 percent of the cases he would obtain a test statistic exceeding 1.96 in absolute value. In other words, while he would think his significance level is 5 percent, it would actually be 46 percent. Moreover, the median estimate of the excess sensitivity parameter, \( \beta \), is 0.21. Thus, the bias in the standard procedures is only somewhat smaller under the assumption that income is subject to both permanent and transitory shocks.

3.3. Borderline stationarity

Our examples above concern income series that are not stationary around a deterministic trend. The problem of bias, however, does not disappear if

\[ \text{While there is no drift term in income, its inclusion would not alter the Monte Carlo results. In particular, a drift term of } \omega \text{ would increase each value of income by } \omega t, \text{ which would be eliminated by the detrending. Note also that the results are invariant with respect to the variance in the innovation.} \]
Table 2
Results of Monte Carlo experiment; income is IMA(1,1). a

<table>
<thead>
<tr>
<th>Value of θ</th>
<th>ρ</th>
<th>π</th>
<th>β</th>
<th>t-test</th>
<th>Percent rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.91</td>
<td>-0.09</td>
<td>1.00</td>
<td>-2.21</td>
<td>61</td>
</tr>
<tr>
<td>0.1</td>
<td>0.89</td>
<td>-0.09</td>
<td>0.80</td>
<td>-2.18</td>
<td>60</td>
</tr>
<tr>
<td>0.2</td>
<td>0.87</td>
<td>-0.09</td>
<td>0.62</td>
<td>-2.14</td>
<td>59</td>
</tr>
<tr>
<td>0.3</td>
<td>0.82</td>
<td>-0.08</td>
<td>0.46</td>
<td>-2.10</td>
<td>56</td>
</tr>
<tr>
<td>0.4</td>
<td>0.76</td>
<td>-0.08</td>
<td>0.32</td>
<td>-2.01</td>
<td>53</td>
</tr>
<tr>
<td>0.5</td>
<td>0.67</td>
<td>-0.07</td>
<td>0.21</td>
<td>-1.88</td>
<td>46</td>
</tr>
<tr>
<td>0.6</td>
<td>0.55</td>
<td>-0.06</td>
<td>0.13</td>
<td>-1.68</td>
<td>38</td>
</tr>
<tr>
<td>0.7</td>
<td>0.39</td>
<td>-0.04</td>
<td>0.07</td>
<td>-1.44</td>
<td>27</td>
</tr>
<tr>
<td>0.8</td>
<td>0.20</td>
<td>-0.02</td>
<td>0.03</td>
<td>-1.06</td>
<td>17</td>
</tr>
<tr>
<td>0.9</td>
<td>0.05</td>
<td>-0.01</td>
<td>0.01</td>
<td>-0.62</td>
<td>9</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.16</td>
<td>5</td>
</tr>
</tbody>
</table>

a Y_t is generated by eq. (6) and C_t by eq. (7). Then eqs. (3) and (4) are estimated on detrended data. Each sample is 100 periods long. Distribution is based on 1000 replications. Detrended data are generated by taking residuals from regression on linear trend. Percent rejections for the null hypothesis that ρ = 0 is based on conventional 'five percent' critical value of t-test of 1.96.

Income is in fact stationary. In particular, if the income series is barely stationary (that is, has a root close to the unit circle), the asymptotic theory justifying standard test procedures is misleading for samples of typical size. To illustrate the problems of conventional inference with barely stationary series, we consider an example in which income is in fact stationary.

Suppose income follows a first-order autoregressive process with autoregressive parameter ρ:

\[ Y_t = \phi + \rho Y_{t-1} + \epsilon_t. \]  

(10)

As before, income is set according to the permanent income hypothesis in eqs. (7), (8), and (9). We set initial non-human wealth equal to zero, initial income equal to its unconditional mean of \( \phi/(1 - \rho) \), and the quarterly real interest rate to 1.25 percent.

Again suppose our econometrician obtained data generated by this economy, detrended it, and then performed regression (3). Table 3 presents the results of this Monte Carlo experiment for various values of ρ. The bias is largest for values of ρ close to one. At ρ = 0.95, the median value of the test statistic is -1.37. In 27 percent of the cases he would obtain a test statistic exceeding 1.96 in absolute value. Moreover, the median estimate of the excess

6While there is no true trend in income in this example, it is straightforward to verify that the results of the Monte Carlo experiment are invariant with respect to the true trend. Hence, we set the true trend equal to zero for simplicity. Note that the results are also invariant with respect to \( \phi \) and to the variance of the innovation in income.
Table 3
Results of Monte Carlo experiment; income is AR(1).8

<table>
<thead>
<tr>
<th>True value of ρ</th>
<th>Median estimate of</th>
<th>Percent</th>
<th>Median estimate of</th>
<th>β</th>
<th>Median estimate of</th>
<th>Percent</th>
<th>Median estimate of</th>
<th>Percent</th>
<th>Median estimate of</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.91</td>
<td>-0.09</td>
<td>1.00</td>
<td>-2.21</td>
<td>61</td>
<td>0.995</td>
<td>0.91</td>
<td>-0.06</td>
<td>0.68</td>
<td>-2.10</td>
</tr>
<tr>
<td>0.99</td>
<td>0.91</td>
<td>-0.05</td>
<td>0.49</td>
<td>-1.96</td>
<td>50</td>
<td>0.98</td>
<td>0.91</td>
<td>-0.03</td>
<td>0.30</td>
<td>-1.77</td>
</tr>
<tr>
<td>0.95</td>
<td>0.88</td>
<td>-0.01</td>
<td>0.11</td>
<td>-1.37</td>
<td>28</td>
<td>0.90</td>
<td>0.85</td>
<td>-0.01</td>
<td>0.04</td>
<td>-0.99</td>
</tr>
<tr>
<td>0.80</td>
<td>0.75</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.72</td>
<td>12</td>
<td>0.50</td>
<td>0.47</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.36</td>
</tr>
<tr>
<td>0.00</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.16</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8 $Y_t$ is generated by eq. (10) and $C_t$ by eq. (7). Then eqs. (3) and (4) are estimated on detrended data. Each sample is 100 periods long. Distribution is based on 1000 replications. Detrended data are generated by taking residuals from regression on linear trend. Percent rejections for the null hypothesis that $\pi = 0$ is based on conventional ‘five percent’ critical value of $t$-test of 1.96.

sensitivity parameter, $\beta$, is 0.11. Thus, while the bias in the standard procedures is smaller under the assumption that income is stationary, the bias is nonetheless large.

4. Conclusions

Tests of the permanent income hypothesis that assume income is stationary around a deterministic trend are biased toward rejection if income follows a non-stationary, or barely stationary, process. We believe this finding casts doubt on the conclusion that consumption is excessively sensitive to income.

Researchers testing the permanent income hypothesis should entertain the possibility that income is not stationary around a deterministic trend. Flavin (1981, pp. 1005–1006) reports that her rejection of the theory is decisive only with detrended data. Bernanke (1985, p. 54) reports that using non-detrended data typically leads to non-convergence of his estimation procedure or to nonsensical results. As Plosser and Schwert (1978) discuss, using first-differenced data is often preferable to using detrended data. It is not generally valid simply to use the raw non-stationary series, because the asymptotic distribution theory justifying the hypothesis testing assumes stationarity of the regressors.

Our finding that disposable personal income is approximately a random walk suggests that measuring the excess sensitivity of consumption is not possible using aggregate post-war data.7 Such a test entails measuring the

7It is still possible, however, that modeling income as a multivariate process might permit such measurement.
response of consumption to transitory income, but shocks to aggregate income appear completely permanent.

We do not mean to imply that measured consumption should equal measured income. Even given that the univariate process of income is a random walk, there are a variety of explanations for the failure of consumption exactly to equal income. The real interest rate may not be constant [Mankiw (1981), Michener (1984)]; the utility function may not be quadratic [Zeldes (1983)]: consumption goods may be durable [Hayashi (1983)]; the utility function may not be additively separable over time and between the measure of consumption, leisure and other goods [Mankiw, Rotemberg and Summers (1985)]; or taste shifts may be important [Garber and King (1983), Hall (1984)]. Our point in this paper is that even if these factors are irrelevant, as is often assumed, the standard testing procedure is invalid.

Although transitory income is not important in post-war aggregate data, the responsiveness of consumption to transitory income is nonetheless important for evaluating the impact of certain policies. Our findings suggest that tests for the importance of liquidity constraints may require data on individuals. It is likely that individuals experience large variation in transitory income, even though such variation is not found in aggregate data. Recent studies of micro-data, however, appear to reach mixed conclusions.

References


Frisch, Ragnar and Frederick V. Waugh, 1933, Partial time regressions as compared with individual trends, Econometrica 1, 387–401.


Hall, Robert E., 1984, The role of consumption in economic fluctuations, NBER working paper no. 1391.

*See, for example, Bernanke (1984), Hall and Mishkin (1982), Hayashi (1984), Runkle (1984), Shapiro (1984), and Zeldes (1984).


Runkle, David E., 1984, Liquidity constraints and the permanent income hypothesis: Evidence from panel data, Unpublished paper (Brown University, Providence, RI).


