What Do Tests of Market Efficiency Show In The Presence of The Permanent-Transitory Confusion

Alex Cukierman  
*University of Tel-Aviv*

Allan H. Meltzer  
*Carnegie Mellon University, am05@andrew.cmu.edu*

Follow this and additional works at: [http://repository.cmu.edu/tepper](http://repository.cmu.edu/tepper)  
Part of the [Economic Policy Commons](http://repository.cmu.edu/tepper), and the [Industrial Organization Commons](http://repository.cmu.edu/tepper)
What Do Tests of Market Efficiency Show In The Presence of The Permanent-Transitory Confusion

by: Alex Cukierman and Allan H. Meltzer*

To test for the efficiency of the Treasury bill market, Fama (1975) examined the residuals from a linear regression relating the rate of change in the value of money to a (previous) market forecast of its value. His maintained hypothesis is that if serial correlation of the regression residuals is found, market participants do not use all of the available information in an efficient manner; errors of forecast can be reduced by using the information contained in the persistent deviations of actual values from the forecast values implied by observable, market values.

The simplicity of the test and the intuitive appeal of Fama's interpretation perhaps account for its frequent use and application to other asset markets. Hamburger and Platt (1975) use current values of the forward rates on Treasury bills to forecast future spot rates. They find evidence of positive serial correlation in the residuals from some of their regressions. Figlewski and Wachtel (1981) test the rationality of the individual Livingston price expectations by checking whether forecast errors are serially correlated and find those errors to be serially correlated. They conclude that survey respondents did not use all available information. Mishkin (1981) rejects the rationality of Livingston's inflationary expectations over the period 59-69 for similar reasons.

*University of Tel-Aviv and Carnegie-Mellon University, respectively. Jacob Frenkel, Scott Hein, Jan Kmenta, Rick Mishkin, Francesco Papadia and anonymous referees made useful comments. The first author acknowledges financial assistance from the Foerder Institute for Economic Research, Tel-Aviv University.
(1977), (1979), (1980) and others used very similar procedures to test for the efficiency of the foreign exchange market.

This paper shows that in the presence of some confusion between permanent changes and temporary shocks tests for market efficiency are valid in finite samples only if those samples are not dominated by large permanent changes. When data from a finite sample are taken following a large permanent change, evidence of serial correlation of the residuals does not necessarily imply that markets are inefficient or that there is serial correlation in the underlying population.

To illustrate the problem, we first restate the tests of market efficiency used by Fama and Frenkel. Next we endow the rate of inflation and the rate of exchange with an unobserved components stochastic structure. The unobserved components are a permanent, non-stationary component and a transitory, stationary component, neither of which is ever observed separately. This stochastic structure is a convenient and relatively simple way to model imperfect information about the permanence of shocks that affect the economy. Using this stochastic structure in the context of market efficiency tests we then show that, in finite samples, forecast errors may sometimes appear to be serially correlated even if there is no serial correlation in the population.

I. Tests of Asset Market Efficiency

Fama (1975) proposed the following test of market efficiency. Let \( R_t \) be the nominal return earned in month \( t \) on a Treasury bill with one month to maturity at \( t-1 \). Let \( \Delta_t \) and \( r_t \) be, respectively, the random (at \( t-1 \)) rate of

\footnote{This stochastic structure was proposed by Muth (1960) to justify Friedman's (1957) measure of permanent income. Brunner, Cukierman and Meltzer (1980) use it to explain stagflation. The transitory component is modeled as white noise and the permanent component as a random walk.}
change in the purchasing power of money and the real return on a one month bill during month \( t \). Ignoring second-order terms, we can write:

\[
(1) \quad r_t = R_t + \Delta_t .
\]

\( R_t \) is known at the end of \( t-1 \), so the relationship between the expected real rate and the expected rate of change in the purchasing power of money, given the information available at the end of \( t-1 \), is (from (1)):

\[
(2) \quad E_{r_t} = R_t + E\Delta_t
\]

where \( E \) is the expected value conditioned on the information available at the end of period \( t-1 \). Following Fama we assume that the expected real rate is constant \( (E_{r_t} = E_r \text{ for all } t) \) and rewrite (2) as:

\[
(3) \quad \Delta_t = E_r - R_t + \varepsilon_t
\]

where \( \varepsilon_t = \Delta_t - E\Delta_t \) is the error of forecast in the rate of change of the purchasing power of money. Fama tests the joint hypothesis of market efficiency and constancy of the expected real rate by running the regression,

\[
(3a) \quad \Delta_t = \alpha_0 + \alpha_1 R_t + \varepsilon_t
\]

The null hypothesis is \( \alpha_1 = -1 \) and no serial correlation in \( \varepsilon_t \). Detection of serial correlation in \( \varepsilon_t \) is taken to imply that the market is inefficient.

There is very little reason to assume that the rate of change of the purchasing power of money (or the rate of price change) is drawn from a distribution with constant mean value. \(^5\) Suppose the mean of the distribution of \( \Delta \) is not stationary

---

\(^2\) See Fama (1975) section I and equation (5). We have simplified the notation.  
\(^3\) Fama later allowed for a fluctuating real rate. Our point applies to that case as well.  
\(^4\) In Fama's words (1975, p. 273): "Nonzero autocorrelations imply that the market is inefficient; one can improve on the market's assessment of the expected value of \( \Delta \) ..."  
\(^5\) A study by Beveridge and Nelson (1981) suggests that most aggregate economic time series are non-stationary in this sense.
but changes at discrete intervals. A stochastic process in which there is confusion
about the permanence of shocks has this property. Let

\[ \Delta_t = \Delta_t^p + \Delta_t^q \]  

(4)

where

\[ \Delta_t^p - \Delta_{t-1}^p \sim N(0, \sigma_{\Delta}^2) \]
\[ \Delta_t^q \sim N(0, \sigma_{\Delta}^2) \]
\[ \Delta_t^p - \Delta_{t-1}^p \quad \text{and} \quad \Delta_t^q \quad \text{are statistically independent.} \]

Frenkel (1977), (1979) and (1980) studied the efficiency of the foreign ex-
exchange market. In his model, \( s_t \) is the logarithm of the spot rate of exchange in
period \( t \) and \( f_{t-1} \) is the logarithm of the one period forward exchange rate prevailing
in the previous period. Assuming that \( f_{t-1} \) is an unbiased forecast of \( s_t \) that
reflects all the information available in \( t-1 \), the forecast error is serially uncorrelated. Formally,

\[ s_t = f_{t-1} + \eta_t \]

(5)

where \( \eta_t = s_t - f_{t-1} \). Frenkel tests the efficiency of the foreign exchange market
by running the regression

\[ s_t = a + bf_{t-1} + \eta_t \]

(5a)

and by testing the hypothesis that \( a = 0 \), \( b = 1 \) and \( \eta_t \) is serially uncorrelated.

Suppose now that (as was the case with the rate of change in the purchasing
power of money) the logarithm of the rate of exchange is composed of two stochastic
components, one permanent \( (s_t^p) \) and one transitory \( (s_t^q) \)

\[ s_t = s_t^p + s_t^q, \quad s_t^p - s_{t-1}^p \sim N(0, \sigma_{s_t^p}^2), \quad s_t^q \sim N(0, \sigma_{s_t^q}^2) \]

(6)

\[ s_t^p - s_{t-1}^p \quad \text{and} \quad s_t^q \quad \text{are statistically independent.} \]
II. A More General Model and the Appearance of Serial Correlation.

In this section, we reformulate the two models presented above in a more general model that collapses to either of the two models in particular cases. Let \( y_t \) be a stochastic variable. The optimal forecast of \( y_t \) made in period \( t-1 \), using the information available during \( t-1 \), is \( \text{Ey}_t \). Let

\[
X_{t-1} = c_0 + c\text{Ey}_t
\]

be a market variable observed in period \( t \) that reflects the market forecast, as of \( t-1 \), of \( y_t \). By definition

\[
y_t = \text{Ey}_t + (y_t - \text{Ey}_t)
\]

solving for \( \text{Ey}_t \) from (7) and substituting for the first \( \text{Ey}_t \) in (8)

\[
y_t = -\frac{c_0}{c} + \frac{1}{c} X_{t-1} + (y_t - \text{Ey}_t)
\]

Efficiency of this general model can be tested by running the regression

\[
y_t = \beta_0 + \beta X_{t-1} + u_t
\]

When \( y_t = \Delta_t \), \( X_{t-1} = R_t \) (9) reduces to equation (3a) used by Fama to test the efficiency of the Treasury Bill market. For \( c = 1 \), \( \beta_0 = c_0 \) becomes an estimate of the (assumed) constant expected value of the real rate of interest. When \( y_t = s_t \), \( X_{t-1} = f_{t-1} \), equation (9) reduces to equation (5a) used by Frenkel to test the efficiency of foreign exchange markets. The hypothesis of market efficiency takes the form \( c_0 = 1 \), \( c = 1 \) and no serial correlation in \( u_t \).

To complete the generalization of the models of section I, we endow \( y_t \) with the same qualitative stochastic structure that \( \Delta_t \) and \( s_t \) have. We therefore assume that

\[
y_t = y_t^p + y_t^q, \quad \Delta y_t^p = y_t^p - y_{t-1}^p \sim N(0, \sigma_p^2), \quad y_t^q \sim N(0, \sigma_q^2),
\]

\[
\Delta y_t^p \text{ and } y_t^q \text{ are statistically independent.}
\]
\(y^p_t\) and \(y^q_t\) are the permanent and transitory components of \(y_t\) respectively.\(^6\) If individuals in the market can observe the permanent and the transitory components of \(y_t\) separately, the best forecast of \(y_t\) as of \(t-1\) is \(y^p_{t-1}\), and the forecast error \(\Delta y^p_t + y^q_t\) is serially uncorrelated. In general, individuals observe \(y_t\) but can never observe the permanent and the transitory components of \(y_t\) separately. Given the information that is available in \(t-1\) (including all past values of \(y\) up to and including period \(t-1\)) the best, in the mean square sense, forecast of \(y_t\) is given by

\[
E y_t = \theta \sum_{i=0}^{\infty} (1-\theta)^i y_{t-1-i} \quad 0 \leq \theta \leq 1
\]

where \(\theta\) is a known function of the ratio of variances \(\sigma^2_p / \sigma^2_q\). Using (11) we can write the forecast error which appears in equation (8a) as

\[
y_t - E y_t = y^q_t - \theta \sum_{i=0}^{\infty} (1-\theta)^i y^q_{t-1-i} + \sum_{i=0}^{\infty} (1-\theta)^i \Delta y^p_{t-1}
\]

To illustrate that the realization of a large permanent shock may cause forecast errors to appear to be serially correlated for a time, even if there is no serial correlation in the population, we consider the effect of the realization of a large permanent change, \(\Delta y^p_t\), which occurs between period \(t-1\) and period \(t\). It is shown in the appendix that the product of two adjacent forecast errors for any period following period \(t\) can be written:

\[
u_{t+j+1} u_{t+j} = (1-\theta)^{2j+1} ((\Delta y^p_t)^2 - \sigma^2_p) + [C_{t+j} + (1-\theta)^{2j+1} \sigma^2_p]
\]

\[+ (1-\theta)^j \Delta y^p_t A_{t+j} + B_{t+j} \quad j \geq 0\]

where \(A_{t+j}\) and \(B_{t+j}\) are linear combinations of present and past values of \(y^q\) and

\(^6\) \(y_t = \Delta_t, y^p_t = \Delta^p_t, y^q_t = \Delta^q_t\) imply \(\sigma^2_p = \sigma^2_p\Delta\) and \(\sigma^2_q = \sigma^2_q\Delta\). An analogous set of definitions holds for the case \(y_t = s_t\).

\(^7\) The proof of this statement appears in Muth (1960). For this stochastic structure \(E y_t\) is also the best forecast, as of \(t-1\), of the permanent component of \(y\) in that period. For further details see Brunner, Cukierman & Meltzer (1980).
\( \Delta y^p \), do not involve \( \Delta y^p_t \), and have zero unconditional expected values. \( C_{t+j} \) also
does not involve \( \Delta y^p_t \) and the unconditional expected values of \( C_{t+j} + (1-\theta)^2 j + 1 \sigma^2_p \)
is zero. Past values of \( \Delta y^p_t \) and \( y^q_t \) which appear in the expressions for \( A_{t+j} \),
\( B_{t+j} \) and \( C_{t+j} \) are weighted more lightly the further away they are from period \( t+j \).

By a 'large' value of \( \Delta y^p_t \) we mean that it is large in comparison to the
following magnitudes:
1) the variance \( \sigma^2_p \) of \( \Delta y^p_t \);
2) the normal values of \( A_{t+j} \);
3) the normal values of \( B_{t+j} + C_{t+j} + (1-\theta)^2 j + 1 \sigma^2_p \).

Under these three restrictions, the first term on the right hand side of (13) is
positive and dominates all the other terms on the right hand side of (13) for small
values of \( j \). As a result the product of two adjacent forecast errors in (13) is
positive for small non-negative values of \( j \). If the difference \( (\Delta y^p_t)^2 - \sigma^2_p \) is
sufficiently large, it makes the product \( u_{t+j, t+j} \) positive not only for \( j = 0 \)
but also for several successive values of \( j \). However, at some point the term
\( (1-\theta)^2 j + 1 \) becomes small enough to obliterate the dominance of \( (\Delta y^p_t)^2 - \sigma^2_p \) on the
sign and magnitude of \( u_{t+j+1, t+j} \). From this point on, unless there are additional
large permanent changes, the sign of \( u_{t+j+1, t+j} \) is randomly distributed around
zero.

The product of adjacent forecast errors tends to be positive immediately after
the occurrence of a large permanent change and for several following periods. An econo-
metrician who uses a finite sample that is dominated by observations taken from such a
period is likely to detect serially correlated forecast errors, even if the errors are not

---

8 Explicit expressions for \( A_{t+j} \), \( B_{t+j} \) and \( C_{t+j} \) appear in the Appendix.
9 One could define those normal values as being within a two or three standard
error band of expected values of the expression shown in (2) and (3) of the text (which
are zero). Since the main point can be made without introducing this additional
machinery, we do not specify the precise meaning of 'normal'.
10 Since \( \Delta y^p_t \) is large in comparison to \( \sigma^2_p \), \( (\Delta y^p_t - \sigma^2_t)^2 \) is positive and large.
Since \( \Delta y^p_t \) is large in comparison to normal values of \( A_{t+j} \), \( (\Delta y^p_t)^2 \) is also large in
comparison to \( \Delta y^p_t A_{t+j} \).
serially correlated in the population. The problem has two sources: (1) the permanent-transitory confusion and (2) the length of the sample. For very large samples, the problem does not arise; the sample estimate of serial correlation does not deviate from the expected value (zero) of the true serial correlation between adjacent forecast errors. However, when a small finite sample is taken immediately after the occurrence of a large permanent change, the sample estimate of the serial correlation between forecast errors is likely to be positive. The positive estimate is caused by the occurrence of a permanent change that is large in comparison to the change that is expected on the basis of the distribution of $\Delta y^p$ in equation (10). The probability that a large permanent change will occur is low in comparison to the probability that smaller permanent shocks will occur.

The probability that a large permanent shock occurs, although small, is not zero. When such a change occurs, individuals learn about the change gradually. The reason is that the structure of the optimal forecast depends mainly on the relatively frequent small values in the distribution of $\Delta y^p$. As a result, the size of the permanent change is underestimated for several periods. Gradual learning implies that forecast errors appear to be serially correlated during this interval. There is no way that this seeming serial correlation can be used by individuals to improve their forecast, however. Everyone knows that there is no serial correlation in the population. The serial correlation appears only after the fact to an econometrician who draws a finite sample that includes the period in which the large permanent change occurred. His detection of serial correlation in that sample does not imply that markets are inefficient.  

\textsuperscript{11}Strictly speaking the scenario described in the text will induce the appearance of serial correlation only if a large permanent change today is not followed during the next few periods by a large change in the opposite direction. Note also that although large permanent shocks that have occurred in the distant past (previously to period $t$) enter into the right hand side of eq. (13), their effect is negligible because they are multiplied by very small weights. See also eq. (A3)-(A5) of the appendix.
When there is no permanent-transitory confusion, the problem disappears. To illustrate this point consider the case in which $\sigma_q^2 = 0$ so that all shocks are permanent and individuals get a direct observation on the permanent component. In this case $\theta = 1$ and all the terms involving $\Delta y_t^p$ on the right hand side of (13) drop out for $j \geq 1$. As a result, for $j \geq 1$, the product $u_{t+j} u_{t+j+1}$ is unaffected by the large permanent shock, $\Delta y_t^p$. In other words, the large permanent shock affects only the product of the forecast error for the period in which it occurs and the forecast error of the immediately following period. The products of adjacent forecast errors in all subsequent periods are unaffected by $\Delta y_t^p$.

By contrast, when there is some degree of confusion between permanent changes and transitory shocks, the large permanent shock of period $t$ contaminates the products of forecast errors for many periods to come. This causes forecast errors to be on the same side of zero for a number of periods after the occurrence of the large permanent change.

The failure to reject serial correlation, in tests of the efficiency of the Treasury bill market, is misleading if applied to samples taken shortly after large permanent changes in the rate of change in the purchasing power of money. Similarly, the serial correlation test may support incorrect inferences about the efficiency of the foreign exchange market if it is applied during or shortly after large permanent changes in the exchange rate.

**Conclusion**

A basic inference problem that confronts individuals in a changing world is the detection of the permanence of shocks. Following relatively rare, but large, permanent shocks, slow, but optimal, learning may invalidate serial correlation as

---

12 See eq. (A7) of the appendix.
13 For a recent application of this methodology see Hafer and Hein (1982).
a test of market efficiency in finite samples. Forecast errors will show evidence of serial correlation following large, unanticipated, permanent changes even if the market is efficient.

Some examples illustrate the problem. Mishkin (1981) tests the rationality of the Livingston data on inflationary expectations during the 1959-69 period and finds that rationality is strongly rejected. However, when he performs the same test over a longer (1954-76) period, rationality is not rejected. These results are consistent with the framework proposed here. The period 1959-69 is a relatively short period that includes a large permanent change in the rate of inflation. Standard tests of market efficiency may, as we saw, reject market efficiency in such periods even if markets are efficient. The fact that rationality is not rejected when this period is included as part of a longer time period supports this view. A similar explanation applies to rejection of rationality for expected exchange rate changes during the German hyperinflation by Frenkel (1977) and also by Papadia (1982) who used survey based inflationary expectations for Italy and Denmark. The common element in each of these cases is that there is a strong likelihood that the samples examined were in all those cases dominated by large permanent changes. Similarly, the serial correlation detected by Figlewski and Wachtel (1981) in forecast errors of individual respondents to the Livingston survey may reflect a confusion between permanent and transitory changes in the rate of inflation following a large permanent change rather than inefficient use of information.

We illustrate this point using a particular stochastic process in which permanent and transitory shocks occur but cannot be distinguished for some time after their occurrence. Similar qualitative results obtain using other stochastic processes to model the permanent-transitory confusion; for example, the permanent component can be a random walk and the transitory component a first or higher order Markov process.
Since evidence of serial correlation in finite samples may be caused by the confusion between permanent and transitory shocks following rare but violent permanent changes as well as by inefficient use of information, such evidence is ambiguous and does not necessarily imply that markets are inefficient. As a matter of fact such evidence could be used, under the assumption that markets are efficient, by econometricians or economic historians looking backwards to detect periods in which large permanent changes have occurred.
Leading (12) by \( j \) periods

\[
(Al) \quad u_{t+j} = y_{t+j} - Ey_{t+j} = y^q_{t+j} - \theta \sum_{i=0}^{\infty} (1-\theta)^i y^q_{t+j-i} + \sum_{i=0}^{\infty} (1-\theta)^i \Delta y^p_{t+j-i}
\]

Leading (Al) by one period and multiplying the resulting expression for \( u_{t+j} \) by the expression for \( u_{t+j} \) in (Al)

\[
(A2) \quad u_{t+j+1} u_{t+j} = \left( y^q_{t+j+1} - \theta \sum_{i=0}^{\infty} (1-\theta)^i y^q_{t+j-i} + \sum_{i=0}^{\infty} (1-\theta)^i \Delta y^p_{t+j-i} \right) \cdot \left( y^q_{t+j} - \theta \sum_{i=0}^{\infty} (1-\theta)^i y^q_{t+j-i} + \sum_{i=0}^{\infty} (1-\theta)^i \Delta y^p_{t+j-i} \right)
\]

Opening the parentheses on the right hand side of (A2), rearranging so that terms involving \( \Delta y^p_t \) appear separately and adding and subtracting \( (1-\theta)^2 \frac{1}{p} \) we obtain equation (13) in the text where

\[
(A3) \quad A_{t+j} = y^q_{t+j+1} + (1-2\theta) y^q_{t+j} - 2\theta \sum_{i=1}^{\infty} (1-\theta)^i y^q_{t+j-i} + \Delta y^p_{t+j+1} + 2\sum_{i=0}^{\infty} (1-\theta)^{i+1} \Delta y^p_{t+j-i}
\]
\[ (A4) \quad B_{t+j} = y_t^{q} \left( \sum_{i=0}^{\infty} (1-\theta)^i \Delta y_{t+j-1}^p - \theta \sum_{i=0}^{\infty} (1-\theta)^i y_{t+j-1} \right) + y_{t+j}^q \left( \sum_{i=0}^{\infty} (1-\theta)^i y_{t+j+1}^p - \theta \sum_{i=1}^{\infty} (1-\theta)^i y_{t+j-1}^p \right) + \theta^2 \sum_{i=0}^{\infty} \sum_{s=0}^{\infty} (1-\theta)^{i+s} y_{t+j-i}^p y_{t+j-1-s} + \sum_{i=0}^{\infty} \sum_{s=0}^{\infty} (1-\theta)^{i+s} \Delta y_{t+j-1}^p \Delta y_{t+j+1-s} \]

\[ \quad - \theta \left( \sum_{i=0}^{\infty} \sum_{s=0}^{\infty} (1-\theta)^{i+s} y_{t+j-i}^p \Delta y_{t+j-1}^p + \sum_{i=0}^{\infty} \sum_{s=0}^{\infty} (1-\theta)^{i+s} \Delta y_{t+j}^p \right) \]

\[ (A5) \quad C_{t+j} = -\theta (y_{t+j}^q)^2 + \theta^2 (1-\theta) \sum_{i=0}^{\infty} (1-\theta)^{2i} (y_{t+j-i-1}^q)^2 + (1-\theta) \sum_{i=0}^{\infty} (1-\theta)^{2i} (\Delta y_{t+j-i}^p)^2 \]

\( A_{t+j} \) involves all the terms which multiply \( \Delta y_{t}^p \). \( B_{t+j} \) includes all the cross products except those which involve \( \Delta y_{t}^p \). \( C_{t+j} \) includes all second powers of the unobserved stochastic components except \( (\Delta y_{t}^p)^2 \). Using equation (10) in (A3) and (A4) it follows that:

\[ E \ A_{t+j} = 0, \quad E \ B_{t+j} = 0 \]

where the \( u \) under the expected value operator means that the expected value is an unconditional expected value. The second quality follows immediately by noting that \( B_{t+j} \) involves only cross products terms. Using (10) in (A5)
\[ EC_{t+j} = (\theta^2(1-\theta)[1 + (1-\theta)^2 + (1-\theta)^4 + ...] - \theta)^2 \]
\[ + (1-\theta)[1 + (1-\theta)^2 + ...(1-\theta)^2(j-1) + (1-\theta)^2(j+1) + ...]\]
\[ = \frac{1}{\theta(2-\theta)} \left[ \frac{(1-\theta)^2}{p} - \theta^2 \frac{2}{q} \right] \]

But \((1-\theta)^2 \frac{2}{p} - \theta^2 \frac{2}{q} = 0\) since as shown in Cukierman (1982) eq. (11) and Part B of the Appendix \(\frac{2}{p} / \frac{2}{q} = \frac{\theta^2}{(1-\theta)}\) (See also Muth (1960) pp. 302-4). Hence

\[(A6) \quad EC_{t+j} = -(1-\theta)^2 j+1 \frac{2}{p} \]

from which it follows that the expected value of \(C_{t+j} + (1-\theta)^2 j+1 \frac{2}{p}\) is also zero.

Note that past changes in the permanent component are weighted less heavily the further away they are in the past.

Note that when \(\frac{2}{q} = 0\) and therefore \(\theta = 1\),

\[ A_{t+j} = \Delta y_{t+j+1}^p, \quad C_{t+j} = 0 \quad \text{for} \quad j \geq 0 \]

and

\[ B_{t+j} = \begin{cases} 0 & \text{for} \ j = 0 \\ \Delta y_{t+j+1}^p & \Delta y_{t+j}^p \quad \text{for} \ j > 0 \end{cases} \]

As a result equation (13) reduces in this case to

\[(A7) \quad u_{t+j+1} + u_{t+j} = \Delta y_{t+j+1}^p \Delta y_{t+j}^p \quad j \geq 0 \]

so that \(\Delta y_{t+j}^p\) appears only in the expression for \(u_{t+j+1} u_{t+j}\).


