Stock Price Effects of Permanent and Transitory Shocks

William J. Crowder
Assistant Professor of Economics
College of Business
Department of Economics
University of Texas at Arlington
Arlington, Texas 76019-0479
Office: (817) 273-3147
Fax: (817) 273-3145
E-mail: crowder@uta.edu

and

Mark E. Wohar
Distinguished Lucas Professor
Department of Economics
University of Nebraska at Omaha
Omaha, NE 68182
Office: (402) 554-3712
Fax: (402) 554-3747
E-mail: wohar@unomaha.edu

First draft: October 1995
This revision: May 1997

Abstract
This paper exploits the long-run equilibrium relationship between stock prices and dividends, implied by the present value model, to structurally identify the dynamic time series model that governs the behavior of stock prices. The shocks or innovations to the data are dichotomized into those that leave a permanent imprint on both series and those that have only transitory effects. Unlike previous studies we do not impose arbitrary identification restrictions to decompose the joint process, other than the standard normalization and orthogonality restrictions. We show that the data are not consistent with certain ad hoc restrictions used by other researchers.

Keywords: structural VAR, cointegration, common trends representation.
JEL Classification:C12, C32, E24
1. Introduction

It is now accepted by most financial economists that stock price and dividend series follow a more general nonstationary process than a simple random walk implying that they may be affected by more than one type of disturbance. For instance, if stock price is the sum of a pure random walk and an independent white noise innovation, the process for stock price will follow an ARIMA(0,1,1). If stock prices do not follow a pure random walk (with possible drift) then the series contain a temporary component implying that stock prices (returns) are predictable from past prices (returns). This predictability of stock returns, has been reported by Fama and French (1988) and Poterba and Summers (1988).

The size of this predictable component and its source are important issues that are not yet resolved. The present paper has two main objectives. First, we examine whether a long-run relationship exists between the natural logarithm (log) of real stock prices and the log of real dividends using a data set of quarterly observations. We also perform some analysis on a data set of annual frequency. Such a relationship is implied by a linearized version of the present value pricing relationship for stock prices. Once we establish evidence of a long-run relationship between stock prices and dividends, our second objective is to model the implied dynamic behavior of the two series when subjected to innovations in both the permanent and transitory components. Previous short-run modeling of stock prices has exploited the Vector Autoregressive (VAR) methodology and, in particular, used the VAR to test the set of restrictions implied by the present value model (see Campbell and Shiller, 1987). The approach to the modeling of stock prices taken in this paper is somewhat different to that adopted by other researchers. Drawing from previous literature which finds that stock prices and dividends are nonstationary
processes, this paper employs a structural cointegrated bivariate VAR, to decompose stock prices and dividends into permanent and transitory components based on a single cointegrating vector. The identification procedure employed is similar to that of Blanchard and Quah (1989). But unlike Blanchard and Quah (1989) the identification of the structural model from the reduced form estimates uses no additional restrictions beyond arbitrary normalization and structural error independence. Complete structural identification is achieved by relying on those restrictions implied by cointegration between these two variables. We demonstrate the relative importance of each component in predicting future variations in stock prices by analyzing the one-step-ahead forecast error variance due to each structural innovation. Finally, we show that the inferences regarding the source of the permanent and transitory innovations and their impacts on the data are quite sensitive to the imposition of invalid over-identifying restrictions as is done by previous researchers.

The outline of the paper is as follows. In section 2 we discuss the present value stock price model. In sections 3 and 4 the econometric techniques used are presented. A description of the data is presented in section 5. The empirical results are presented in sections 6 and the paper concludes with section 7.

2. A Cointegrated Model of Stock Price and Dividends

The standard present value stock price model has been used extensively in finance and simply states that the current stock price should be equal to the expected discounted present value of the future stream of dividends to be paid on the asset. This relation is given in equation (1),
where $P_t$ is the stock price in period $t$, $r$ is the constant discount rate, $D_{t+j}$ is the dividend paid in period $t+j$ and $E_t$ is the expectations operator conditioned on information available in period $t$.\(^2\)\(^3\)

Campbell and Shiller (1987) have recently presented an alternative way of testing (1) which appropriately accounts for any possible nonstationarity in the data. Their approach relies on the fact that the dividend and stock price series each have a unit root or are integrated of order one, I(1). The present value model implies that these two series should be cointegrated.

This can be demonstrated by subtracting $D_t/r$ from both sides of (1) to obtain:

$$P_t - \frac{D_t}{r} = \theta \sum_{j=1}^{\infty} \delta^j E_t(D_{t+j} - D_t)$$

where the factor of proportionality $\theta = 1/r$. Since $D_{t+j} - D_t$ is stationary or I(0) stock price and dividends should be cointegrated with cointegration vector $[1 - \theta]^\prime$.

By using a log-linear approximation to the above present value model, Campbell and Shiller (1989) derive what has been termed the dynamic Gordon model (or dividend-ratio model). This model expresses the log dividend-price ratio (or dividend yield), $s_t = \ln(D_t) - \ln(P_t) = d_t - p_t$, (lower case letters here denote natural logarithms) as the expected discounted value of all future one-period growth adjusted discount rates $(r - \Delta d_{t+j})$. Since $r$ and $\Delta d_{t+j}$ are both stationary, $d_t - p_t$ should also be stationary implying a cointegration vector of the form $[1 -1]^\prime$ in the log-linear version of the present value model.

3. Identification of the Structural Dynamic Model

Campbell and Shiller (1987) show that the stock price and dividends series are cointegrated when
they are related by the stock price valuation (present value) model of equation (1). They suggest forming a VAR model with the first difference in one variable (either \( p_t \) or \( d_t \)) and a stationary linear combination of the other two variables, \( s_t = p_t - d_t \). Consider the bivariate model consisting of the first difference in log of stock prices, \( \Delta p_t \), and the log stock price-dividend ratio or spread, \( s_t \), also used by Lee (1995). 4

\[
(3) \quad Z_t = \begin{bmatrix} \Delta p_t \\ s_t \end{bmatrix}
\]

It is assumed there exists a structural moving average representation (MAR) for (3) given by (4):

\[
(4) \quad Z_t = A_0 e_t + A_1 e_{t-1} + \ldots = A(L) e_t, \quad E[e_t e_t'] = I
\]

where \( Z_t \) follows a stationary and invertible process, \( I \) is the \( p \times p \) identity matrix and \( L \) is the lag operator. 5 Estimation consists of estimating the reduced form VAR and inverting it to obtain the reduced form MAR in (5).

\[
(5) \quad Z_t = v_t + G_1 v_{t-1} + \ldots = G(L) v_t, \quad E[v_t v_t'] = \Omega
\]

The structural and reduced form parameters are related via \( v_t = A_0 e_t, \ A_j = G_j A_0 \) and \( A(1) = G(1) A_0 \). Identification of the structural parameters from the reduced form estimates is achieved by imposing suitable restrictions upon \( A_0 \). In general, for a VAR of dimension \( p \), i.e., \( p \) endogenous variables, one requires \( p^2 \) independent restrictions to exactly identify the structural innovations from the reduced form errors. Usually \( p \) restrictions are imposed by normalizing structural error variances to unity and another \( p(p-1)/2 \) independent restrictions are obtained by assuming a diagonal covariance matrix, (i.e., the structural errors are orthogonal). This leaves \( p(p-1)/2 \) additional restrictions that the econometrician must impose to achieve exact
These restrictions may take the form of contemporaneous exclusion restrictions as in Sims (1980) and Bernanke (1986), or long-run neutrality type restrictions as in Blanchard and Quah (1989). They may also be a combination of the two as in King et al. (1991), Gali (1992) and Crowder et al. (1994).

In a recent analysis of the stock price/dividends relationship, Lee (1995) imposes the restriction that the spread innovation has no long run or permanent effect on the level of \( p_t \). Under the assumptions that neither innovation affects the spread in the long run (due to its stationarity in levels) and innovations in the spread have no permanent effect on the price variable, the total impact matrix of (4), \( A(1) \), is of the form

\[
A(1) = \begin{bmatrix}
    a_{11}(1) & 0 \\
    0 & 0
\end{bmatrix}
\]  

where the innovation vector is \( e_t = [e^p_t, e^s_t]' \). This restriction says that the only permanent effect is on the level of the price due to the price innovation. An alternative restriction, which we argue is more consistent with the data, has the \( A(1) \) matrix take the form

\[
A(1) = \begin{bmatrix}
    0 & a_{12}(1) \\
    0 & 0
\end{bmatrix}
\]  

This restriction says that the only permanent effect is on the level of the price but now due to an innovation in the spread.\(^6\)

4. **Identification of a Cointegrated System**

The assumptions made about the time series properties of \( Z_t \) imply that this system can be represented as a cointegrated system. In this section we will demonstrate the general equivalence
between a cointegrated system given by (3), which includes differences, and one given in the
levels of the data and show that restrictions imposed on the representation for \( Z_t \) has important
implications for the representation of \( X_t \).

If a \( p \times 1 \) system \( X_t \) has one or more unit roots in the characteristic equation of its
autoregressive (AR) polynomial and there exists a \( p \times r \) matrix \( \beta \) such that \( \beta'X_t \) has no unit roots,
then it is said that the system \( X_t \) is cointegrated with cointegration vectors given by the columns
of \( \beta \). Engle and Granger (1987) prove that a cointegrated system has a reduced form vector error
correction model (VECM) representation such that:

\[
\Delta X_t = \alpha \beta'X_{t-1} + \sum_{j=1}^{r-1} \Pi_j \Delta X_{t-j} + u_t
\]

where \( \alpha \) is the \( p \times r \) matrix of error correction coefficients. Phillips (1991) demonstrate that a
cointegrated system has an equivalent representation given by defining \( M \) such that \( MX_t = [\Delta X_{1t}, \beta'X_t]' \), where \( \Delta X_{1t} \) are the first \( p-r \) variables in \( \Delta X_t \) and \( \beta'X_t \) are the \( r \) cointegrating relations.

Let \( X_t \) be given by,

\[
X_t = \begin{bmatrix} p_t \\ d_t \end{bmatrix}
\]

so that \( Z_t \) in equation (3) is related to \( X_t \) in (9) by,

\[
Z_t = MX_t = \begin{pmatrix} (1-L) & 0 \\ 1 & -1 \end{pmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} \Delta X_{1t} \\ \beta'X_t \end{bmatrix}
\]

where \( \beta = [1 \ -1]' \) as implied by the present value model described in section 2. Johansen
(1991) has proved that if the elements of \( X_t \) are integrated of order one or less, so that I(2)
processes are ruled out, the product matrix \( \xi = [\alpha_\perp \Pi^*(1) \beta_\perp]^{-1} \) is non-singular, where \( \Pi^*(1) \) is the total impact matrix from the AR representation of (8)\(^8\) and the \( pxK \) matrices \( \alpha_\perp \) and \( \beta_\perp \) are the orthogonal complements to the matrix of error correction coefficients, \( \alpha \), and the matrix of cointegration vectors, \( \beta \), respectively (i.e., \( \alpha'\alpha_\perp = \beta'\beta_\perp = 0 \)). This result gives a necessary and sufficient condition for the MAR of (8) to exist. Assuming that log price and log dividends are both I(1) and cointegrated, \( \xi \) exists by definition and the MAR of (8) is given by:

\[
(11) \quad \Delta X_t = u_t + C_1 u_{t-1} + C_2 u_{t-2} + \ldots = C(L)u_t
\]

where \( C(L) \) is the inverse of the reduced form AR polynomial representation of \( \Delta X_t \). The common trends representation (CTR) of Stock and Watson (1988) is achieved by decomposing the \( C(L) \) polynomial as \( C(1) + (1-L)C^*(L) \) where \( C^*(L) = (C(L) - C(1))/(1-L) \). This yields the reduced form CTR for \( X_t \):

\[
(12) \quad X_t = C(1)(1-L)^{-1}u_t + C^*(L)u_t
\]

where \( (1-L)^{-1}u_t \) is a pure random walk and \( C(1) \), the total impact matrix of the reduced form, must have reduced rank of \( k \) under the assumption of cointegration rank \( r \). Equation (12) in \( X_t \) can be directly compared with equation (5) in \( Z_t \). The reduced form CTR demonstrates that the common trends are formed by the accumulation of the reduced form errors, \( u_t \). Johansen (1991) demonstrates that the total impact matrix of the MAR is given by \( C(1) = \beta_\perp \xi \alpha_\perp' \), where \( \alpha_\perp' \) represents the linear combination of the reduced form errors that form the common trends and \( \beta_\perp \xi \) represents the factor loadings that transmit the effects of the permanent innovations to each of the variables in \( X_t \). This interpretation is consistent with King, Plosser, Stock and Watson (1991) (KPSW) and highlights the role played by variables that are weakly exogenous with
respect to the long-run parameter matrices $\alpha$ and $\beta$. Only those variables that are weakly exogenous can be the sources of the common trend.

The identification of the structural innovations from the reduced form CTR is achieved by noting that $\beta'A(1) = \beta'C(1) = 0$, i.e., the common trends must be eliminated by (lie in the null space of) the cointegration vectors, implying that one suitable choice of $A(1)$ is $\beta^0_{\perp}$ such that the $k$ structural innovations that accumulate to form the common trends are independent. From (4), (10), and (12) it is seen that $C(1)u_t = A(1)e_t$ and thus $C(1)\Omega C(1)' = A(1)A(1)'$. These conditions along with assumption that the structural innovations are orthogonal yield $k(k+1)/2$ restrictions on the reduced form parameters, leaving $k(k-1)/2$ additional restrictions needed to just identify the structural innovations that accumulate to form the common trends. The $r$ structural innovations associated with transitory movements can be identified by noting that the transitory and permanent innovations must be orthogonal. Given that $A(1)\alpha = C(1)\alpha = 0$ from the definition of cointegration, this yields $r(r+1)/2$ restrictions on the transitory innovations. Thus, there are but $r(r-1)/2$ extra restrictions required to achieve exact identification of the transitory innovations.9

The above discussion demonstrates that knowledge of the cointegration space, $\beta$, can yield restrictions that aid in the identification of the structural innovations.10 Specifically, only $k(k-1)/2 + r(r-1)/2$ additional restrictions are needed to just identify the structural model in a cointegrated system. In the bivariate model of stock price and dividends, the existence of cointegration is all that is necessary to achieve exact identification.11 This is most easily seen by exploiting the relationship that $C(1) = \beta_{\perp}\xi\alpha_{\perp}'$, where $\beta = [1 -1]'$ and thus $\beta_{\perp} = [1 1]'$. This yields a $C(1)$ matrix of the form:
where \( \alpha_\perp = [\alpha_{1\perp} \alpha_{2\perp}]' \). Thus the reduced form CTR is given by:

\[
C(1) = \begin{bmatrix}
\xi \alpha_{1\perp} & \xi \alpha_{2\perp} \\
\xi \alpha_{1\perp} & \xi \alpha_{2\perp}
\end{bmatrix}
\]

From (10) we get \( C(1) = M^{-1} G(1)(1-L) \), where \( G(1) \) is the total impact matrix from (5), thus

\[
X_t = \begin{bmatrix}
p_t \\
d_t
\end{bmatrix} = \begin{bmatrix}
\xi \alpha_{1\perp} & \xi \alpha_{2\perp}
\end{bmatrix} \begin{bmatrix}
\zeta_{1t} \\
\zeta_{2t}
\end{bmatrix} + \begin{bmatrix}
G_{11}(L) & G_{12}(L) \\
G_{21}(L) & G_{22}(L)
\end{bmatrix} \begin{bmatrix}
u_{1t} \\
u_{2t}
\end{bmatrix}
\]

where \( \zeta_{it} = (1-L)^{-1}u_{it} \). From (10) we get \( C(1) = M^{-1} G(1)(1-L) \), where \( G(1) \) is the total impact matrix from (5), thus

\[
\begin{bmatrix}
(1-L)^{-1} & 0 \\
(1-L)^{-1} & 1
\end{bmatrix} \begin{bmatrix}
g(1)_{11} & g(1)_{12} \\
g(1)_{12} & g(1)_{12}
\end{bmatrix} (1-L) = \begin{bmatrix}
c(1)_{11} & c(1)_{12} \\
c(1)_{11} & c(1)_{12}
\end{bmatrix} = \begin{bmatrix}
\xi \alpha_{1\perp} & \xi \alpha_{2\perp} \\
\xi \alpha_{1\perp} & \xi \alpha_{2\perp}
\end{bmatrix}
\]

implying that \( g(1)_{11} = c(1)_{11} \) and \( g(1)_{12} = c(1)_{12} \). It is easy to see in this form that the long-run restriction imposed on (5) by Lee (1995) is \( g(1)_{12} = 0 \) which implies that \( c(1)_{12} = 0 \). Since \( \xi \) is of full rank, \( c(1)_{12} = 0 \) iff \( \alpha_{2\perp} = 0 \). This in itself implies (due to the definition of the orthogonal complement) that \( \alpha_p = 0 \) (\( \alpha_p \) is the error correction coefficient in the stock price equation). This restriction implies that shocks to the stock price have permanent effects on stock prices and dividends, while shocks to dividends have no long-run impact on stock prices or dividends. The alternative restriction we proposed in (7), that \( \alpha_{1\perp} = 0 \), can also be tested and implies that shocks to dividends have permanent effects on stock prices and dividends, while shocks to stock prices have no permanent impact on stock prices or dividends.

Gonzalo and Granger (1995) discuss how to obtain an estimate of \( \alpha_\perp \) using the Johansen procedure. In the case where \( k=1 \), as in the analysis below, the permanent innovations can be identified from restrictions on the \( \alpha_\perp \) matrix. Such restrictions are testable within the Johansen
framework using likelihood ratio or Wald tests with asymptotic $\chi^2$ distributions.

It is important to reiterate that imposing either of these restrictions, that $\alpha_1 = 0$ or $\alpha_2 = 0$, is over-identifying in the bivariate cointegrated model. Strictly speaking, neither of these restrictions is necessary to just identify the structural model and as such, these restrictions are testable.

5. Data

Two data sets are employed in this paper--quarterly and annual. The results for the quarterly analysis are reported throughout while those of the annual analysis are left for in footnote. The quarterly data set is the same data set used by Lee (1995) and consists of real value-weighted stock prices and dividends obtained from the Center for Research in Securities Prices (CRSP) covering the sample 1926Q1 to 1991Q4. The annual observations on real stock prices and dividends are for the Standard and Poor's Composite Price Index and cover the period 1871-1986. A detailed description of the annual data set is contained in Shiller (1989, Ch. 26) and Campbell and Shiller (1989).

6. Testing the Long-Run Restrictions

There are two long-run restrictions imposed by Lee (1995) to achieve identification of the structural innovations from the reduced form errors. The first restriction is that the spread between log stock price and log dividend is stationary in levels. In section II this was referred to as the cointegration restriction since it implies that (log) $P_t$ and $D_t$ [or alternatively $p_t$ and $d_t$] are cointegrated with cointegration vector $\beta = [1 \ -1]'$. To test this hypothesis we use the maximum likelihood estimation procedure advocated by Johansen (1988, 1991).

We first begin by testing for the existence of cointegration between the natural log of
stock price, \( p_t \), and natural log of dividend, \( d_t \). We allow for a trend in the data generating process which necessitates using a specific set of critical values. These can be found in table 1 of Osterwald-Lenum (1992).\(^{13}\) Table 1 presents results for quarterly data for \( \kappa = 5 \) in the VECM of (8). The estimation yields trace test statistics of 10.175 for \( r = 0 \) and 0.215 for \( r \leq 1 \), where \( r \) is the number of cointegration vectors. Comparing these to the 95% critical values of 15.41 and 3.76 the evidence is not favorable to the present value model. Allowing for \( \kappa = 2 \) in the VECM yields trace test statistics of 16.639 for \( r = 0 \) and 0.284 for \( r \leq 1 \), where \( r \) is the number of cointegration vectors. Comparing these to the 95% critical values, there is evidence that for this shorter lag length the series are indeed cointegrated. This lag length of \( k=2 \) does not eliminate all significant serial correlation from the residuals of the dividend equation of the VAR and thus all subsequent analysis is performed with a lag of \( k=5 \) in the VAR which does ensure white noise residuals in both the stock price as well as the dividend equations, as evidenced by the Q-statistics in Table 1.\(^{14}\) We have strong grounds for imposing cointegration even at the higher lag length in light of augmented Dickey and Fuller (1979) (ADF) test results which rejected the null of a unit root in the log dividend-price ratio series using lag lengths up to six. Similarly, when we extract the common stochastic trend from the log price and log dividends series, see figure 1, and test the transitory components for a unit root (which would be a violation of the permanent/transitory decomposition implied by cointegration) we reject the unit root null hypothesis in both transitory components. The ADF t-statistic for the dividends transitory component is -5.40 which is significant at very high levels of significance. The ADF t-statistic for the transitory component of stock price is -3.15 which is significant at conventional levels as well.\(^{15}\)
Maintaining the null of one cointegration vector, the restriction on that vector implied by the present value model that $\beta = [1 -1]'$ can be tested using a likelihood ratio (LR) test proposed by Johansen (1988). The test yields a likelihood ratio statistic of 1.789 which is asymptotically distributed as a $\chi^2$ variate with one degree of freedom. This null cannot be rejected at conventional levels of significance.\textsuperscript{16} Using a recursive procedure suggested by Hansen and Johansen (1993), we tested the stability of the cointegrating relationship over the sample period. The results, which are not presented but are available upon request, suggest that only during the period around the Great Depression is there any evidence against the restriction that $\beta = [1 -1]'$.

The over-identifying restriction imposed by Lee (1995) is on the orthogonal complement to the error correction coefficients. This is equivalent to restricting the error correction coefficients themselves. Estimation of $\alpha_\perp$ indicates that the element of $\alpha_\perp$ associated with dividends is clearly the largest. Lee's restriction is $\alpha = [0 1]'$ implying that stock prices are weakly exogenous with respect to the long-run parameters $\alpha$ and $\beta$.\textsuperscript{17} This hypothesis can also be tested using a LR test developed by Johansen (1988). Testing the null hypothesis of $\alpha = [0 1]'$ yields a test statistic of 4.681 which is distributed as a $\chi^2$ variate with one degree of freedom. This null can be rejected at the 5% level. The corresponding t-statistic from table 1 is -2.13 which is also significant at the 5% level. This casts doubt on Lee's over-identification restriction.

We next test our over-identifying restriction, $\alpha = [1 0]'$ or that dividends are weakly exogenous with respect to the long-run parameters. The null hypothesis of $\alpha = [1 0]'$ is tested using the LR statistic which yields a test statistic of 2.019, again distributed as a $\chi^2(1)$ and so is not rejected. The t-statistic of the null hypothesis $\alpha_d = 0$ is 1.41 which is not significant at conventional levels of significance. These results provide support for our alternative over-
identifying restriction and suggest that dividends are weakly exogenous, so that the dividend innovation is accumulated to form the common trend. Intuitively, dividends do not adjust to any of the disequilibria given by \((p_t - d_t)\).

The test for weak exogeneity of dividends cannot be rejected at the 15% or higher level of significance, implying that the dividends series is causally prior to the stock price. It does not imply strict (or strong) exogeneity of dividends, which would require no short-run feedback. To test the joint hypothesis that all of the lagged stock prices in the dividend equation of the VECM are insignificantly different from zero, a standard F-test is calculated. The results of this test (reported in Table 1) yield an F-statistic of 3.945 which is significant at the 0.004 level of significance. This indicates that, in fact, there is substantial short-run feedback from stock prices to dividends. Furthermore, the joint restriction on the lagged dividends in the price equation yields an F-statistic of 0.4293, which is clearly insignificant and suggests very little short-run feedback from dividends to stock prices. Based on this evidence, we conclude that long-run causality runs one way, from dividends to stock prices, but there is short-run feedback from prices to dividends.\(^{18}\)

If dividends are weakly exogenous we can identify innovations to dividends as the permanent component. This means that innovations to stock prices must be the transitory component. Figure 1 plots the actual series as well as the permanent and transitory components of each series. The left panels plot the actual series as a solid line with the permanent component reflected in the dotted line. The right panels plot the stationary deviations from the common stochastic trend that represent the transitory portion of each series.

In order to analyze the consistency of the dynamic behavior of each of the variables with
the implications of the theoretical model, impulse response functions (IRFs) are calculated from
the structural common trends representation. Figure 2 displays the graphs of the accumulated
impulse response functions (IRFs) from the common trends representation implied by the
cointegrated system. These IRFs were obtained without making any restrictions to identify the
model beyond the arbitrary normalization of each equation and the independence of the structural
innovations.19 Examining the IRFs due to a (permanent) dividend shock (Figure 2 left panel)
leads one to conclude that the basic predictions of the theoretical model are satisfied. An
innovation to the permanent component causes stock prices and dividends to rise permanently.
The initial disequilibrium, shown by the IRF of the price-dividend ratio \((p_t - d_t)\) variable, is large
but quickly dissipates as the stock price adjusts to its new higher long run equilibrium. The new
equilibrium is achieved in, roughly, fourteen quarters. A shock to stock price (the temporary
component) (Figure 2 right panel) has no significant long-run effect on either stock prices or
dividends. The transitory shock does imply a negative short-run relationship between prices and
dividends, causing a large increase in the price-dividend ratio, which is relatively long-lasting.

The contribution of each of the structural innovations in our common trends model to the
dynamic behavior of each series can be analyzed through decompositions of the forecast error
variance (VDCs). The exogeneity of dividends and the long run relation that the growth rate of
dividends equals the growth rate of stock prices can be readily observed through analyzing these
variance decompositions. Table 2 presents the percentage of the forecast error variance that is
explainable by each of the structural innovations of stock price and dividend series along with
a 95 percent confidence interval. The permanent innovation accounts for a large proportion of
the forecast error variance in the dividend equation [see Table 2, panel B], especially at longer
forecast horizons. The stock price forecast errors are to a large degree explained by shocks to stock prices in the short run [see Table 2, panel A, left column], while innovations to dividends explain a larger amount of stock price forecast errors at long forecast horizons [see Table 2, panel B, left column]. Approximately 50% of the short-run, up to four quarters, variation in stock price is due to the transitory innovation.

We wish to examine the effects of improperly imposing weak exogeneity on stock price series as is done in Lee (1995). While we know that incorrectly imposing weak exogeneity on stock prices will alter the true long-run properties of the system, it is not clear if the short-run dynamics are as adversely affected. Figure 3 displays the IRFs generated by imposing the (incorrect) weak exogeneity restriction on the stock price variable. In this model, dividend innovations are associated with the transitory shock and stock price innovations correspond to the permanent shock. Note that the IRF paths are essentially the reverse of those in figure 2. By imposing an invalid over-identifying restriction, the dynamic responses yield results that are the opposite of those obtained from a valid specification. Furthermore, the magnitudes are also quite different. For example, in figure 2 the first period response of stock price to a transitory (stock price) innovation is approximately 40% greater than the analogous response given in figure 3. Imposing an invalid restriction not only leads to the wrong conclusion about the source of the transitory component but also leads to a distortion in the magnitudes of the effects of the innovations.

7. Conclusions
Using quarterly stock price and dividend data spanning the period 1926-1919, this paper exploits the long-run equilibrium relationship between stock prices and dividends, found by several
previous studies, to identify the structural model that governs the behavior of stock prices from reduced form estimates. Shocks or innovations to stock prices/dividends are classified as permanent and transitory. Unlike previous studies, we do not impose arbitrary identification restrictions to decompose the joint process, other than the standard normalization and orthogonality restrictions. We show that the cointegration relations implied by the equilibrium between stock prices and dividends imply restrictions on these variables that aid in the identification of the cointegrated structural VAR.

We also show that the data are not consistent with certain ad hoc restrictions used by other researchers. We then examine the response of stock prices and dividends to permanent and temporary innovations. Finally, variance decomposition analysis is performed. The procedure in this paper is more explicit than previous work about the implications of the model, and the results reveal more information about the linkages between stock prices and dividends.
References


Table 1
Cointegration Analysis
Quarterly Data: 1926Q1-1991Q4

<table>
<thead>
<tr>
<th>H₀:</th>
<th>Trace Test (k=5)</th>
<th>Trace Test (k=2)</th>
<th>10% Critical(^a)</th>
<th>5% Critical(^a)</th>
<th>Eigenvalues (k=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>r = 0</td>
<td>10.175</td>
<td>16.639</td>
<td>13.325</td>
<td>15.410</td>
<td>0.0377</td>
</tr>
<tr>
<td>r ≤ 1</td>
<td>0.215</td>
<td>0.284</td>
<td>2.687</td>
<td>3.762</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

(k=5)

<table>
<thead>
<tr>
<th>Equation</th>
<th>( \hat{\beta}_c )</th>
<th>( \hat{\alpha} )</th>
<th>LQ-Stat(64)(^b)</th>
<th>F-(\Sigma\Delta p_{ei})</th>
<th>F-(\Sigma\Delta d_{ei})</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_t )</td>
<td>1.000</td>
<td>-0.066 (0.031)</td>
<td>44.3670 {0.97}</td>
<td>3.3414 {0.01}</td>
<td>0.4293 {0.787}</td>
</tr>
<tr>
<td>( d_t )</td>
<td>-1.102 (0.072)</td>
<td>0.038 (0.027)</td>
<td>78.518 {0.104}</td>
<td>3.945 {0.004}</td>
<td>66.688 {0.000}</td>
</tr>
</tbody>
</table>

\(^a\) Critical values are taken from table 1. of Osterwald-Lenum (1992). \(^b\) LR test of 64\(^{th}\) order serial correlation from the level VAR to correspond to a sample size of 264. \( \hat{\beta}_c \) is the normalized estimate of the cointegration vector. The column of \( \hat{\alpha} \) corresponds to the estimate of the error correction coefficient associated with each equation in the VECM. Numbers in parentheses are asymptotic standard errors and numbers in brackets are marginal significance levels. \( p_t \) and \( d_t \) are natural log of real stock price and dividend, respectively.
Table 2  
Panel A  

One-Step Ahead Forecast Error Variance Decompositions  
(Due to Innovations in Stock Price)

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Stock Price</th>
<th>Dividend</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42.57</td>
<td>25.25</td>
</tr>
<tr>
<td></td>
<td>[0.32, 96.20]</td>
<td>[0.13, 92.54]</td>
</tr>
<tr>
<td>2</td>
<td>43.45</td>
<td>22.98</td>
</tr>
<tr>
<td></td>
<td>[0.27, 96.55]</td>
<td>[0.48, 90.77]</td>
</tr>
<tr>
<td>3</td>
<td>43.85</td>
<td>18.67</td>
</tr>
<tr>
<td></td>
<td>[0.30, 96.71]</td>
<td>[2.70, 83.68]</td>
</tr>
<tr>
<td>4</td>
<td>42.52</td>
<td>15.99</td>
</tr>
<tr>
<td></td>
<td>[0.36, 96.13]</td>
<td>[5.57, 75.45]</td>
</tr>
<tr>
<td>8</td>
<td>38.15</td>
<td>11.11</td>
</tr>
<tr>
<td></td>
<td>[0.54, 93.62]</td>
<td>[4.15, 66.43]</td>
</tr>
<tr>
<td>12</td>
<td>33.82</td>
<td>9.19</td>
</tr>
<tr>
<td></td>
<td>[1.07, 92.21]</td>
<td>[3.23, 64.44]</td>
</tr>
<tr>
<td>16</td>
<td>30.01</td>
<td>8.23</td>
</tr>
<tr>
<td></td>
<td>[1.66, 89.36]</td>
<td>[2.65, 62.01]</td>
</tr>
<tr>
<td>20</td>
<td>26.93</td>
<td>7.64</td>
</tr>
<tr>
<td></td>
<td>[2.37, 86.56]</td>
<td>[2.29, 61.34]</td>
</tr>
<tr>
<td>24</td>
<td>24.51</td>
<td>7.44</td>
</tr>
<tr>
<td></td>
<td>[3.05, 83.97]</td>
<td>[1.97, 60.25]</td>
</tr>
<tr>
<td>36</td>
<td>19.35</td>
<td>6.70</td>
</tr>
<tr>
<td></td>
<td>[4.36, 76.97]</td>
<td>[1.48, 59.00]</td>
</tr>
<tr>
<td>48</td>
<td>16.65</td>
<td>6.40</td>
</tr>
<tr>
<td></td>
<td>[4.96, 72.28]</td>
<td>[1.22, 58.57]</td>
</tr>
<tr>
<td>60</td>
<td>14.81</td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>[5.06, 68.11]</td>
<td>[1.04, 59.22]</td>
</tr>
</tbody>
</table>

Numbers in brackets represent the 95% confidence interval. The procedure used to simulate these confidence bounds takes random draws from the posterior distribution of the VAR. The simulation is conducted for 1000 iterations. The VDCs are then ordered and the 2.5th and the 97.5th percentiles are extracted.
Panel B

Variance Decompositions of Stock Prices and Dividends
(Due to Innovations in Dividend)

<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>Stock Price</th>
<th>Dividend</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>57.32</td>
<td>74.65</td>
</tr>
<tr>
<td></td>
<td>[3.20, 99.64]</td>
<td>[7.39, 99.86]</td>
</tr>
<tr>
<td>2</td>
<td>56.41</td>
<td>77.00</td>
</tr>
<tr>
<td></td>
<td>[2.92, 99.59]</td>
<td>[9.16, 99.52]</td>
</tr>
<tr>
<td>3</td>
<td>56.02</td>
<td>81.30</td>
</tr>
<tr>
<td></td>
<td>[2.81, 99.63]</td>
<td>[16.00, 97.29]</td>
</tr>
<tr>
<td>4</td>
<td>57.37</td>
<td>83.95</td>
</tr>
<tr>
<td></td>
<td>[3.25, 99.60]</td>
<td>[24.47, 94.42]</td>
</tr>
<tr>
<td>8</td>
<td>61.82</td>
<td>88.79</td>
</tr>
<tr>
<td></td>
<td>[5.19, 99.44]</td>
<td>[32.73, 95.84]</td>
</tr>
<tr>
<td>12</td>
<td>66.18</td>
<td>90.80</td>
</tr>
<tr>
<td></td>
<td>[7.51, 98.92]</td>
<td>[35.40, 96.77]</td>
</tr>
<tr>
<td>16</td>
<td>69.87</td>
<td>91.72</td>
</tr>
<tr>
<td></td>
<td>[10.58, 98.33]</td>
<td>[37.77, 97.32]</td>
</tr>
<tr>
<td>20</td>
<td>73.02</td>
<td>92.32</td>
</tr>
<tr>
<td></td>
<td>[13.22, 97.55]</td>
<td>[38.42, 97.70]</td>
</tr>
<tr>
<td>24</td>
<td>75.47</td>
<td>92.56</td>
</tr>
<tr>
<td></td>
<td>[15.91, 96.94]</td>
<td>[39.46, 98.02]</td>
</tr>
<tr>
<td>36</td>
<td>80.58</td>
<td>93.26</td>
</tr>
<tr>
<td></td>
<td>[22.41, 95.61]</td>
<td>[40.93, 98.51]</td>
</tr>
<tr>
<td>48</td>
<td>83.32</td>
<td>93.58</td>
</tr>
<tr>
<td></td>
<td>[27.62, 94.99]</td>
<td>[41.33, 98.78]</td>
</tr>
<tr>
<td>60</td>
<td>85.18</td>
<td>93.98</td>
</tr>
<tr>
<td></td>
<td>[31.64, 94.93]</td>
<td>[40.34, 98.95]</td>
</tr>
</tbody>
</table>

Numbers in brackets represent the 95% confidence interval. The procedure used to simulate these confidence bounds takes random draws from the posterior distribution of the VAR. The simulation is conducted for 1000 iterations. The VDCs are then ordered and the 2.5th and the 97.5th percentiles are extracted.
Figure 1
Figure 2
Figure 3
Endnotes


2. The assumption of a constant discount rate is made for convenience. A necessary condition for stock price and dividends to be cointegrated is that the discount rate is itself stationary. Crowder and Hoffman (1996) present evidence that the real risk-free rate in the U.S. is indeed stationary.

3. The transversality condition \((1+r)^n \mathbb{E}_t P_{t+n} \to 0\) as \(n \to \infty\) must hold to ensure against price bubbles.

4. This is the triangular representation of a cointegrated system discussed by Phillips (1991).

5. As Lippi and Reichlin (1993) demonstrate, this specification makes an additional implicit assumption concerning the existence of non-fundamental MA representations of \((5)\).

6. What we call the spread innovation here is later shown to actually be a dividend innovation.

7. The first \(k\) rows of \(M\), where \(k = p - r\), are given by the \(k \times p\) selection matrix with \((1-L)\) on the main diagonal and the final \(r\) rows are given by the \(r\) cointegration vectors suitably normalized.

8. The matrix \(\Pi' (1)\) is the sum of the short-run dynamics coefficient matrices in \((9)\); i.e., \(\Pi' (1) = \Pi'_1 + \Pi'_2 + \ldots + \Pi'_{k-1}\).

9. Crowder et al. (1994) provide a perspicuous discussion of the identification of the structural model from a reduced form cointegrated system that builds upon the initial work of King et al. (1991).

10. Since estimates of \(\beta\) are superconsistent (see Stock (1987)), they can be treated as if known in subsequent statistical analysis.

11. The bivariate case is a simple but important one since the existence of cointegration imposes enough (testable) restrictions on the system that no further ad hoc restrictions are needed to identify the model. This important proof is made by Crowder (1995) who demonstrates the testability of the restrictions imposed to identify structural VARs from the reduced form estimates.

12. A more detailed description of the quarterly data is contained in Lee (1995). We would like to thank Bong-Soo Lee for kindly providing us with his data.

13. We tested the restriction that the data have linear trends that are eliminated by the cointegration vector, corresponding to case 1 in Osterwald-Lenum (1992), against the more general model where the linear trends are not eliminated by the cointegrating relationship, corresponding to case 2* in Osterwald-Lenum. We could not reject the restriction. We then
tested an even more restrictive specification that did not allow linear trends in the data only a non-zero mean in the cointegrating relationship, case 1* in Osterwald-Lenum. This restriction was rejected by the data.

14. The lag length k corresponds to the lag length in the VAR. The lag length in the VECM is equal to k-1. In a recent Monte Carlo study examining the small sample properties of various cointegration estimators, Inder (1993) demonstrates that the problems associated with under-parameterization far-outweigh the problems of over-parameterization.

15. The lag lengths in the regressions were 4 and 3 for the dividends and stock price transitory components, respectively. These were the minimum lags needed to reduce the residuals to statistical white noise.

16. The LR test of the alternative hypothesis that $\beta = [1 \ 0]'$ for the vector $X_t = \{p_t, d_t\}$ yields a test statistic of 9.669 and is easily rejected at conventional levels of significance.


18. The results for annual data (not shown) are qualitatively similar, however, the results are not as strong statistically.

19. We simulated the 90% confidence intervals for each of the IRFs, but the bounds were so tight that for many of the plots it was difficult to distinguish the mean responses from the confidence intervals. The plot of the IRFs with the confidence intervals are available from the authors upon request.