The short and long-term interdependencies between stock prices and dividends: 
A panel vector error correction approach

RICKARD PERSSON

Abstract:

This paper examines the short and long-term interdependencies between stock prices and dividends. We utilize firm level data from FTSE ALL SHARE from 1990-2014 and apply panel vector error correction model estimated with Engle & Grangers (1987) two-step procedure. The result show that there is a bi-directional long-term relationship between stock prices and dividends, i.e. an adjustment process is at work when a disequilibrium occurs. We also find a bi-directional short-term relationship. This paper also shows that Lintners model and the present value model are relevant frameworks in stock valuations.

Keywords: Present value model, Lintners model, PVECM, Stock price, Dividends, FTSE, Cointegration

JEL Classifications: C01, C22, C23, G10

Acknowledgements: We appreciate comments from Fredrik Johansson (department of Statistics at Uppsala University), Robert Piterlaff, Joachim Landstrom, William Skoric, Regina Persson and seminar participants at Ekonomistet.

1. Introduction

The present value model (PVM) has been employed frequently to explain and predict stock prices in financial economics literature during the last decades (Macdonald & Poir 1995 p.135; Campbell 2000 p.1530; Blake & Wohar 2006 p.55). Recently, the stock price-dividend relationship has received a fair amount of attention in the literature due to the rapid increase in stock prices in the 1990s and 2000s while fundamentals seemed to remain unchanged (Blake & Wohar 2001 p.22). Some empirical studies also show that the linear PVM fails to explain stock prices in the long run, which has caused some financial economists to question if the PVM still can be used in stock valuations (Esteve & Prats 2010 p.405; McMillan 2010a p.189). However, the empirical evidence that specifically considers the 1990s and 2000s remains scarce and a critique to previous research is that it has mostly examined the stock price-dividend relationship by a stock index and a dividend index (Goddard et al. 2008 p.196), whereas Samuelsson’s dictum suggests that the relationship is more likely to hold on a firm level than on aggregated macro level (Shiller 2000 p.186; Jung & Shiller 2002 p.46; Jung & Shiller 2005 p.223). Moreover, prior research has for the most part examined the long-term relationship between stock prices and dividends, but neglected to examine the short and long-term interdependencies (Sung & Urrutia 1995 p.171). There is, to our knowledge, no empirical study that investigates these interdependencies based on data from the 1990s and 2000s. The aim of this paper is, therefore, to examine the short and long-term interdependencies between stock prices and dividends based on individual firm level data.

This paper applies a panel vector error correction model (PVECM) estimated with Engle & Grangers (1987) two-step procedure. We utilize firm level data from FTSE ALL SHARE in the U.K. for the period 1990-2014. The results show that there are short and long-term bi-directional dependences between stock prices and dividends.

The remainder of this paper is organized as follows: Section 2 provides the literature review, Section 3 gives a review of the theoretical foundation & hypotheses and Section 4 gives a description of the data. Section 5 provides methodology, while Section 6 gives the results and discussion. Finally, Section 7 presents the conclusion.

2. Review of the financial economics literature

Cointegration tests

Campbell & Shiller (1987) show that if the PVM is valid there should exist a long-term relationship between stock prices and dividends i.e. stock price and dividend should be cointegrated. Campbell & Shiller (1987) use the S&P composite stock price index and dividends index from 1871-1986. Some test statistics rejected the null hypothesis of no cointegration (Campbell & Shiller 1987 p.1086). Several studies have, since the paper of Campbell & Shiller (1987), tested stock prices and dividends for cointegration.

---

1 Cointegration is when two variables are non-stationary, but a linear combination of these exist i.e. the residuals are stationary (Pflue & Gentelman 2008 p.76).
Esteve & Prats (2010) use index data from the U.S. market 1871-2004 and test for linear cointegration and threshold cointegration. The results show that there is no linear cointegration but they found evidence of threshold cointegration (Esteve & Prats 2010 p.409). Kanas (2005) finds similar evidence when testing the relationship in four countries based on index data. The original stock price and dividend series are not cointegrated, while the transformed non-linear series are cointegrated (Kanas 2005 p.604). Furthermore, Caporale & Gil-Alana (2004; 2014) use an updated version of the Campbell & Shiller (1987) dataset and test for fractional cointegration and find that the series are fractionally cointegrated (Caporale & Gil-Alana 2004 p.246; Caporale & Gil-Alana 2004 p.1408). Additionally, Han (1996) investigates the present value relationship by applying canonical cointegration regression with the same data as Campbell & Shiller (1987). He finds that the present value relationship does not hold (Han 1996 p.272). Outside the U.S., McMillian (2010b) fails to find cointegration when using sector level data from the U.K. However, Goddard et al. (2008) use data from 104 U.K. firms across the period 1970-2003 and find that stock prices and dividends are, on average, cointegrated at an individual firm level (Goddard et al. 2008 p.201).

**Error correction models**

**Summary**

To sum up, empirical evidence is ambiguous. Studies that apply linear cointegration tests (i.e. the relationship between stock prices and dividends in its simplest form) seem to reject the long-term relationship, while papers that test for threshold cointegration and fractional cointegration are more supportive. Moreover, index data is more often used than individual firm level data. Goddard et al. (2008) use panel data but only test for cointegration. Nasseh & Strauss (2004) do not test the interdependencies in a VECM system instead they estimate two ECMs. This paper uses individual firm level data from the U.K. with a period of time that covers the 1990s and 2000s (1990-2014) and it examines the short and long term interdependencies by applying a PVECM. Table 1 summarizes, to the best of our knowledge, the existing literature regarding studies that apply cointegration tests and ECMs.

### 3. Theoretical foundation & hypotheses

“Economic theory is mathematical analysis. Everything else is just pictures and talk”

- R.E. Lucas, Jr (2001), Nobel laureate

#### 3.1 The Present value model

The PVM is written as:

\[
P_t = \sum_{i=1}^{\infty} E_t (D_{t+i}) / (1+r)^i,
\]

where \( P_t \) is stock price in time period \( t \), \( D_{t+i} \) is dividend in time period \( t+i \), \( r \) is the discount rate and \( E_t \) is the expectations operator at time \( t \). If eq.1 holds empirically, stock prices and dividends should be cointegrated (Campbell & Shiller 1987 pp.1070-1074).

---

1. The assumption of ECM is that it has a tendency to move towards its equilibrium in every period. However, it is possible that the movement does not occur in every period and it is only when deviations from the equilibrium exceed a critical threshold that adjustments occur. 


3. Caporale & Gil-Alana (2004) introduce a testing procedure that allows for a long-memory cointegration relationship (fractional cointegration), because it might be the case that the equilibrium error responds more slowly to shocks (Caporale & Gil-Alana 2004 p.247).

### Table 1

Summary literature review: Cointegration & ECM:s

<table>
<thead>
<tr>
<th>Author</th>
<th>Data &amp; market</th>
<th>Testing long-term relations</th>
<th>Testing short relations</th>
<th>Testing inter-relations</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esteve &amp; Prats (2010)</td>
<td>Index data: U.S. 1871-2004</td>
<td>Yes, by cointegration &amp; threshold cointegration</td>
<td>No</td>
<td>No</td>
<td>No cointegration but threshold cointegration</td>
</tr>
<tr>
<td>Caporale &amp; Gil-Alana (2004; 2014)</td>
<td>Index data: U.S. 1871-1995 1871-2010</td>
<td>Yes, by fractional cointegration</td>
<td>No</td>
<td>No</td>
<td>Fractional cointegration</td>
</tr>
<tr>
<td>Han (1996)</td>
<td>Index data: U.S. 1871-1986</td>
<td>Yes, by canonical cointegration</td>
<td>No</td>
<td>No</td>
<td>No cointegration</td>
</tr>
<tr>
<td>McMellan (2010b)</td>
<td>Sector index data: U.K. 1986-2009</td>
<td>Yes, by cointegration</td>
<td>No</td>
<td>No</td>
<td>No cointegration</td>
</tr>
<tr>
<td>Saltoglu (1998)</td>
<td>Index data: U.S. 1871-1987</td>
<td>Yes, by the cointegrated VAR</td>
<td>Yes, implied by the VAR</td>
<td>No</td>
<td>Find both short and long-term relationship</td>
</tr>
<tr>
<td>Sung &amp; Urrutia (1995)</td>
<td>Index data: U.S. 1930-1990</td>
<td>Yes, implied by ECM in both ways</td>
<td>Yes, implied by the ECM</td>
<td>Yes, only short-term</td>
<td>Short and long-term relations in both ways</td>
</tr>
</tbody>
</table>

#### 3.2. Causality from stock prices to dividends

Sung & Urrutia (1995) derive models of causality from stock prices to dividends and vice versa based on the PVM and Lintner’s (1956) dividend model (LDM). First, we derive the relationship from stock prices to dividends based on Sung & Urrutia (1995 pp.172-173). Consider LDM written as:

\[
D_t = \gamma D_{t-1}^T + (1 - \gamma) D_{t-1} + u_t
\]

where \( D_t \) is the dividend paid in time period \( t \), \( \gamma \) (0 \leq \gamma \leq 1) is the speed of adjustment, \( D_{t-1}^T \) is the target dividend and \( D_{t-1} \) is dividend in time period \( t-1 \) and \( u_t \) is the error term. The present value of future expected target dividends (\( P_T^{T} \)) is defined as:

\[
P_T^T = \sum_{t=1}^{\infty} E_t (D_{t+1}^T)/(1+r)^t ,
\]

where \( E_t \) is the expectations operator at time \( t \) and \( r \) is the discount rate. Sung & Urrutia (1995 pp.172-173) argue since the firm’s earnings, which is principal determinant of target dividends \(^5\), seems to follow a geometric random walk it is reasonable to assume that target dividend also follows a geometric random walk\(^6\) of the form \( \ln D_{t+1}^T = u + \ln D_{t+1}^T + v_t \) where \( u \) is a positive parameter and \( v_t \) i.i.d. \( \mathcal{N}(0, \sigma_v^2) \). Using the aforementioned assumption, thus \( P_T^T = (1+g) D_{t+1}^T/r \Rightarrow D_{t+1}^T = (r-g) P_T^T/(1+g) \), where \( g \) is the constant percent increase. Let \( D_{t+1}^T = \alpha P_T^T \), where \( \alpha = (r-g)/(1+g) \) and plug this expression into eq.2, thus:

\[
D_t = \gamma \alpha P_T^T + (1 - \gamma) D_{t-1} + u_t
\]

\(^5\) \( D_{t}^T = e_t + d_t \) where \( e_t \) is earnings in time period \( t \) and \( d_t \) is the desired payout ratio (Renneboog et al. 2009 p.188).

\(^6\) An implementation of the random walk, where the forecast for next quarterly value will equal to the previous quarterly value plus a constant percentage increase.

The short & long-term interdependencies between stock prices and dividends:

A panel vector error correction approach
By recursively substituting values of $D_{t-1}$, ..., $D_{t-(n+1)}$ into the right-hand side:

\[(5) \quad D_t = y\alpha P_t^{TG} + u_t + (1 - \gamma)^2 y\alpha P_{t-1}^{TG} + (1 - \gamma)^2 y\alpha P_{t-2}^{TG} + \ldots + (1 - \gamma)^2 y\alpha P_{t-n}^{TG} + (1 - \gamma)^2 y\alpha u_{t-n} + (1 - \gamma)^{n+1} D_{t-n}\]

With $\lim_{n \to 0} (1 - \gamma)^n D_{t-n} = 0$, for $0 < \gamma \leq 1$ this sums up to:

\[(6) \quad D_t = y\alpha \sum_{i=0}^{\infty} (1 - \gamma)^i P_t^{TG} + u_t\]

Assuming that stock price ($P_t$) is equal to $P_t^{TG}$ plus a disturbance term ($b_t$): $P_t = P_t^{TG} + b_t$, thus:

\[(7) \quad D_t = y\alpha \sum_{i=0}^{\infty} (1 - \gamma)^i P_{t-1} - y\alpha \sum_{i=0}^{\infty} (1 - \gamma)^i b_{t-1} + \sum_{i=0}^{\infty} (1 - \gamma)^i u_{t-1}\]

Eq. 7 reads that current dividends are affected by current and past stock prices.

### 3.3 Causality from dividends to stock prices

From eq. 8-13 we refer to Sung & Urrutia (1995 pp.173-175, 185-188). Dividend smoothing means that current dividend depends on the firm target dividend ($\overline{D}_{t-1}$) and on some average of the firm’s past dividends ($\overline{D}_{t-1}$). Under these conditions LDM can be written as:

\[(8) \quad D_t = y\alpha \gamma \overline{D}_t^{TG} + (1 - \gamma) \overline{D}_{t-1} + u_t\]

For expositional clarity $\overline{D}_{t-1}$ can be expressed as:

\[(9) \quad \overline{D}_{t-1} = (1 - \omega) \sum_{i=0}^{\infty} \omega^i D_{t-1-i}\]

where $\omega$ (0 $\leq$ $\omega$ <1) is the weights of past dividends. Plug eq. 8 and 9 into eq. 1 thus:

\[(10) \quad P_{t} = y\alpha \sum_{i=1}^{\infty} E_i(D_t^{TG}) \overline{D}_{t-1} + (1 - \gamma) (1 - \omega) \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \omega^j E_i(D_{t-1-j}) (1 + r)^i\]

The second expression of eq. 10 includes the investor’s anticipation of the firm’s dividend smoothing into the PVM, which can be rewritten as:

\[P_t = \frac{P_t + \sum_{i=0}^{\infty} \omega^i D_{t-1-j}}{1 + r - \omega}\]

By replacing the above expression into eq. 10 and solve for $P_t$:

\[(11) \quad P_t = \frac{(1 + r - \omega) y\alpha \gamma \sum_{i=1}^{\infty} E_i(D_t^{TG}) (1 + r)^i + (1 - \gamma) (1 - \omega) \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \omega^j E_i(D_{t-1-j}) (1 + r)^i}{(\gamma + r - \gamma \omega) (1 + r)^i}\]

Eq. 3 states that $P_t^{TG} = \sum_{i=1}^{\infty} E_i(D_t^{TG})(1 + r)^i$ and by using the assumption of that the $\overline{D}_t^{TG}$ follows a geometric random walk, thus $P_t^{TG} = (1+g)(\overline{D}_t^{TG}/(r-g))$. Substituting this expression into eq. 11 gives:

\[(12) \quad P_t = \frac{(1 + r - \omega) y\alpha (1+g)}{(\gamma + r - \gamma \omega) (r-g)} \overline{D}_t^{TG} + (1 - \gamma) (1 - \omega) \sum_{i=0}^{\infty} \omega^i D_{t-1-i}\]

Combining eq. 8 and 9 provides the following expression:

\[\gamma P_t^{TG} = \overline{D}_t - (1 - \gamma) (1 - \omega) \sum_{i=0}^{\infty} \omega^i D_{t-1-i}\]

Substituting the above expression into eq. 12 and rearranging the terms, gives:

\[(13) \quad P_t = \frac{1}{(\gamma + r - \gamma \omega) (r-g)} \left( \frac{1 + r - \omega) (1+g)}{(r-g)} + (1 - \gamma) (1 - \omega) \right) \overline{D}_t - \frac{(1 - \gamma) (1 + r - \omega) (1+g)}{(\gamma + r - \gamma \omega) (r-g)} \sum_{i=0}^{\infty} \omega^i D_{t-1-i}\]
or

\[ R_t = \beta_0 D_t + \theta_1 (1 - \omega) \sum_{j=0}^{\omega} \omega^j D_{t-1-j} \]

Eq. 13 states that the current stock price is affected by current and past dividend. Note that current dividends, i.e. \( \beta_0 \), affect the stock prices positively, while the sum of the coefficients of past dividends, i.e. \( \theta_1 \), is negative.

### 3.4 Hypotheses

Section 3.1 states that there should be a long-term relationship implied by the cointegration, which suggest that stock prices will error-correct (adjust) to the long-term relationship when a disequilibrium occurs. Vivian & Wohar (2010) and Saltoglu (1995) (recall table 1) find that there are short-term dynamics and implied by Sung & Urrutia (1995) framework, which suggests that there should exist interdependencies, we state the following research hypotheses:

- **H_1**: Dividends have short-term effects on stock prices
- **H_2**: Stock prices adjust to the long-term relationship
- **H_3**: Stock prices have short-term effects on dividends
- **H_4**: Dividends adjust to the long-term relationship
- **H_5**: There are short and long-term interdependencies

### 4. Data & sample selection procedure

#### 4.1 Data & variables

Three historical lists (1996, 2002 and 2008) with firm level data from FTSE ALL SHARE in the U.K. from 1990-2014 have been extracted from the database THOMSON REUTERS DATASTREAM and aggregated into one panel. We compare the aggregated dataset to historical name lists for every year to validate how many years the individual firm has been listed at FTSE. The motivation for choosing the U.K. market is that it is one of the biggest markets (Daya et al. 2012 p.1056; Yeandle et al. 2014 p.4). The variables extracted from DATASTREAM are stock price (PS) and net dividend per share (DPS) expressed in pence. In the U.K. dividends are generally paid semi-annually (Vivian & Wohar 2010 p.654). These two variables are taken from the end of the 1st and 3rd quarter of the year. The variables are adjusted for inflation and logged.

#### 4.2 Sample criteria & sample selection procedure

The criteria by which firms are included in the sample are as follows: the firm must have at least 30 observations. The motivation for this criterion is to get sufficient power in the panel unit root test (PURT) (see section 5.1). The firms must pay dividends during the aforementioned period of time, however a maximum gap of 4 observations without paying any dividends has been allowed due to the financial crisis in 2007-2009, which implies the variables are calculated as follows: LnDPS=Ln(DPS+1) and LnPS=Ln(PS+1). Financial institutions and investment firms cannot be included in the sample. The sample selection procedure is summarized in table 2.

#### Table 2

<table>
<thead>
<tr>
<th>Procedure</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of firms after aggregation of the FTSE lists</td>
<td>2,227</td>
</tr>
<tr>
<td>Less duplicates</td>
<td>873</td>
</tr>
<tr>
<td>Less financial and investment firms</td>
<td>406</td>
</tr>
<tr>
<td>Less firms that did not pay dividends or have a gap of &gt;4 obs without paying dividends</td>
<td>50</td>
</tr>
<tr>
<td>Less firms with &lt;30 observations on FTSE ALL SHARE 1990-2014</td>
<td>670</td>
</tr>
<tr>
<td><strong>Total firms included in the final sample</strong></td>
<td><strong>228</strong></td>
</tr>
</tbody>
</table>

*Note*: The average firm has 41.98 observations. In total there are 9,573 observations. Descriptive statistics over the variables LnDPS and LnPS can be found in appendix B.

The short & long-term interdependencies between stock prices and dividends: A panel vector error correction approach
5. Methodology

We apply a PVECM to examine the short and long-term interdependencies in the stock-dividend relationship. The vector ECM is appropriate when economics relationships are interrelated and share long and short-term relations (Asteriou & Hall 2011 p.358; Wooldridge 2008 p.643; Hill et al. 2011 p.499). To estimate a PVECM certain conditions must be fulfilled, which include that the series must be integrated of the same order and cointegrated, therefore we first describe the PURT and the panel cointegration tests before we present the model.

5.1 Panel unit root test

The first generation of PURT (introduced by Levin et al. (2002) and Im et al. (2003)) focuses on panels where there errors are assumed to be independent, while the second generation of PURT focuses on cases where the errors are allowed to be cross-sectional dependent (Pesaran et al. 2013 p.94). The assumption of cross-sectional independence is however rather restrictive in empirical applications (Werkman 2012 pp. 271-272), especially in the context of financial data where, for example, international shocks are present. Therefore, the second generation of PURT suggested by Pesaran (2007) is applied when the series of LnPS and LnDPS are integrated of the same order.

Pesaran (2007) PURT is based on the t-ratio of the OLS estimate of $\hat{b}$ in the following cross-sectionally augmented Dickey-Fuller autoregression (CADF):

$$\Delta Y_{it} = \alpha_i + \beta Y_{it-1} + \sum_{j=0}^{p} \delta_j \Delta Y_{it-1} + \sum_{j=0}^{p} \delta_j \Delta Y_{it-j} + e_{it}$$

where $\Delta$ is the first difference operator, $\bar{Y}_i = \sum_{t=1}^{T} Y_{it}$, $\Delta \bar{Y}_i = \sum_{t=1}^{T} \Delta Y_{it}$, $p$ is the lag length and $e_{it}$ is the regression error. The cross-sectional averages, $\bar{Y}_i$ and $\Delta \bar{Y}_i$, are included in eq.14 as a substitute for the unobserved common factor (Palm et al. 2010 p.116). Pesaran (2007) constructs the cross-sectional augmented IPS (CIPS) test, which is an average of the individual CADF tests. A linear time trend can also be included in eq.14, but this will be at cost of a lower power of the test (Pesaran 2007 p.291). If a time trend is included the test needs the time dimension ($T$) to be at least 30 before the power starts to increase with the number of firms ($N$) (ibid.).

5.2 Panel cointegration test

The most widely known and cited panel cointegration tests are the residual based proposed by Pedroni (1999; 2004) and Kao (1999) (Allowd & Basher 2008 p.128; Westerlund et al. 2014 p.2). Moreover, Westerlund (2007) proposes an alternative panel cointegration test that is based on structural rather than residual dynamics (Westerlund 2007 p.710). All of the aforementioned tests are applied, because they are used in the literature of economics. The following sections briefly describe the test procedures.


Pedroni (1999-2004) introduced several test statistics that test the null hypothesis of no cointegration in non-stationary panels. The seven test statistics allow for heterogeneity in the panel (Neal 2014 p.685). The test statistics are, as previously mentioned, residual based and constructed from the following regressions:

$$Y_{it} = \alpha_i + \delta_i t + \beta_1 x_{1it} + \cdots + \beta_M x_{mit} + \epsilon_{it}$$

(15)

$$\Delta Y_{it} = \beta_1 \Delta x_{1it} + \cdots + \beta_M \Delta x_{mit} + n_{it}$$

(16)

$$\hat{e}_{it} = \bar{Y}_i \hat{e}_{it-1} + \mu_{it}$$

(17)

$$\epsilon_{it} = \bar{Y}_i \epsilon_{it-1} + \sum_{k=1}^{K} \bar{Y}_i \epsilon_{it-k} + \bar{u}_{it}$$

(18)

where $\Delta$ is the first difference operator, $i=1,2,\ldots,N$ is the number of individuals in the panel, $t=1,2,\ldots,T$ is the number of time periods, $m=1,2,\ldots,M$ is the number of number of regressors, $k=1,2,\ldots,K$ is the number of lags in the ADF regression, and $\delta$ is the linear time trend. The statistics are computed in the following steps: Estimate the regression in eq.15 and collect the residuals. Estimate eq.16 and compute the residuals. Eq. 16 is used to calculate the long–run variance of $n_i$. Use the residuals in eq.15 to calculate the appropriate autoregression choosing either eq.17 (for the non-parametric statistics) or eq.18 (for the parametric statistics) (Pedroni 1999 pp.659-663). The aforementioned steps are used to construct all of the seven statistics in Pedroni (1999 p.661).
The Kao (1999) test is based on Dickey-Fuller (DF) and the Augmented Dickey-Fuller (ADF) (Dickey & Fuller 1981) test and tests the null hypothesis of no cointegration (Kao 1999 p.6). Kao (1999) is based on the following regression:

\[
Y_{it} = \alpha_i + \beta x_{it} + \epsilon_{it}
\]

where \(i=1,2,N\) is the number of individuals in the panel, \(t=1,2,T\) is the number of time periods. The DF test can then be applied to the following regression:

\[
\hat{\alpha}_i = \rho \hat{\epsilon}_{it-1} + \nu_{it}
\]

where \(\hat{\epsilon}_{it}\) is the estimate of \(\epsilon_{it}\) in eq.19. The null hypothesis of no cointegration is defined as \(H_0: \rho = 1\) against the alternative \(H_1: \rho < 1\).

5.2.2 Westerlund (2007)

Westerlund (2007) proposes four-panel tests, which are based on the following equation:

\[
\Delta Y_{it} = \delta_i d_i + \alpha_i (y_{it-1} - \beta x_{it-1}) + \sum_{t=1}^{p_1} \alpha_{ij} \Delta Y_{it-j} + \sum_{t=1}^{p_2} \alpha_{ij} \Delta x_{it-j} + \epsilon_{it}
\]

where \(\Delta\) is the first difference operator, \(d_i\) holds the deterministic components, \(p_i\) is the lag order that is allowed to vary across individuals. Eq. 21 is actually just a single equation version of the ECM, where \(\alpha_i\) is the error correction (EC) parameter. The problem with eq.21 is how to estimate the \(\alpha_i\). This problem can be solved, by estimating the following regression:

\[
\Delta Y_{it} = \delta_i d_i + \alpha_i y_{it-1} + \lambda_i x_{it-1} + \sum_{t=1}^{p_1} \alpha_{ij} \Delta Y_{it-j} + \sum_{t=1}^{p_2} \alpha_{ij} \Delta x_{it-j} + \epsilon_{it}
\]

If \(\alpha_i < 0\) there is EC, which indicates that \(y_{it}\) and \(x_{it}\) are cointegrated. If \(\alpha_i = 0\) the EC is absent, i.e. there is no cointegration (Westerlund 2007 p.712). The null hypothesis of no cointegration for cross sectional unit \(i\) is implemented as a test of \(H_0: \alpha_i = 0\) vs. \(H_1: \alpha_i < 0\). The panel tests have the alternative hypothesis that the panel is cointegrated as a whole (Westerlund 2007 p.710). The group mean tests have the alternative that at least one individual is cointegrated (Ibid.).

5.3 Panel vector error correction model

In the parlance of economics, deviations from a long-run equilibrium are possible, but these errors are characterized by a mean revision back to its long-run equilibrium (Pfaff & Gentleman 2008 p.76). The question is how to model this dynamic behavior. Engle & Granger (1987) that proposed a two-step estimation technique to model dynamic behavior of \(I(1)\) variables that are cointegrated, which is implemented in this paper. In the first step, the following model is estimated:

\[
\ln \text{PS}_{it} = \alpha_i + \beta_i \ln \text{DPS}_{it} + \epsilon_{it}
\]

where \(i=1,2,\ldots,N\) is the number of firms in the panel, \(t=1,2,\ldots,T\) is the number of time periods, \(\ln \text{PS}\) is the log of stock prices, \(\ln \text{DPS}\) is the log of dividends and \(\epsilon\) is the residuals. The residuals are obtained:

\[
\hat{\epsilon}_{it} = \ln \text{PS}_{it} - \bar{\epsilon}_i - \bar{\beta}_i \ln \text{DPS}_{it}
\]

The lagged residual (\(\hat{\epsilon}_{it-1}\)) now contains information about the long-term relationship and the adjustment process to its long run equilibrium (Asteriou & Hall 2011 p.365). I define the lagged residuals as the error correction term (ECT\(_{it-1}\)). The next step in Engle & Granger (1987) two-step procedure is to estimate a system of equations where the error correction term is incorporated with the short dynamics (Hill et al. 2011 pp.499-502). The system is written as:

\[
\Delta \ln \text{PS}_{it} = \alpha_{11} + \beta_{11} \Delta \ln \text{PS}_{it} + \sum_{k=1}^{3} \theta_{11} \Delta \ln \text{DPS}_{it-k} + \lambda_{11} \text{ECT}_{it-1} + \mu_{11t}
\]

\[
\Delta \ln \text{DPS}_{it} = \alpha_{21} + \sum_{k=1}^{3} \theta_{21} \Delta \ln \text{PS}_{it-k} + \lambda_{12} \text{ECT}_{it-1} + \mu_{21t}
\]

where \(\Delta\) is the first difference operator, \(k\) is the lag length, \(\alpha, \beta, \theta, \gamma\) and \(\lambda\) are slope coefficient in equation \(a\) and \(\mu\) is the residuals. The lag length is set to three in the model, however the chosen lag
length could be debated, therefore we test with different number of lag length to examine if it affects the inference regarding the research hypotheses. Since, all variables in model 25 are I(0) the significance of the parameters can be tested by applying the OLS technique (Hill et al. 2011 p.501). In model 25, it is intuitive that dividend announcements in time period t will affect stock price in time period t. According to eq. 13 the coefficient of $\Delta \ln D_{it}$ ($\beta_{11}$) is expected to be positive while the sum of the lagged coefficients of $\Delta \ln D_{it}$ ($\theta_{11}$) are expected to be negative. In model 25, we expect the signs of the coefficients of LNPS to be positive. Further, in model 25 it is reasonable to assume that the decided dividend amount has been decided earlier than time period t, therefore dividends in time period t cannot be affected stock prices in time period t, hence we only include lagged periods.

The coefficient of $\lambda_{11}$ is expected to $-1 < \lambda_{11} \leq 0$ and the coefficient of $\lambda_{12}$ is expected to $0 \leq \lambda_{12} < 1$ (Hill et al. 2011 p.500). Recall section 3.4, to accept $H_1$ at least one of the coefficients of LNPS in model 25 must be significantly different from zero and to accept $H_2$ the ECT term in model 25 needs to be $-1 < \lambda_{11} < 0$. To accept $H_3$ at least one of the coefficients of LNPS in model 25 must be significantly different from zero and to accept $H_4$ the ECT term in model 25 needs to be $0 < \lambda_{12} < 1$. The aforementioned hypotheses can be tested with a Wald test. To accept $H_5$ hypothesis 1-4 needs to be accepted.

### 5.3.1 Model estimation methodology: FE-OLS

The fixed-effect (FE) OLS model assumes that each firm is different and that individual heterogeneity can be captured by an individual intercept (Mátyás & Sevestre 2008 p.23). If the error terms are correlated the random effects model should be used (Mátyás & Sevestre 2008 p.186). We test if FE-OLS is preferred over the random effects by using the Hausman test. We reject the null hypothesis of the Hausman test (see appendix A) hence the FE-OLS is preferred. The FE-model is:

\[
Y_{it} = \alpha_i + \beta_1 X_{1it} + \cdots + \beta_k X_{kit} + e_{it}
\]

where $\alpha_i$ (i=1,…,n) is the fixed effect for firm i, $Y_{it}$ is the dependent variable, $X_{it}$ is the independent variable and $e_{it}$ is the error term. Note that the slope coefficient $\beta_k$ is the same for each firm. Eq. 23-25 is estimated with FE-OLS estimation methodology. Estimating a PVECM with the FE-OLS estimation methodology is not unprecedented for example Coers & Sanders (2013) and Jiang & Liu (2014) use the same approach.

An alternative approach is to estimate eq.23 with a fully modified OLS (FMOLS)\(^7\) obtain the residuals and apply the FE-OLS estimating methodology to model 25 (Herreras et al. 2013 p. 1488; Liddle & Lung 2013 p.525). The FMOLS and FE-OLS procedure is applied regarding eq.23, because both approaches are used and recognized in the literature of economics.

In panel data sets, the residuals could be correlated (Petersen 2009 p.434). If the residuals are correlated across firms the OLS standard errors (SE) can be biased, which might lead to incorrect inference of the estimated coefficients (Ibid.). Even White (1980) SE would be biased when residuals are not independent (Hoechle 2007 p.282; Petersen 2009 p.475). We test for heteroscedasticity, serial correlation and cross sectional dependence using a modified Wald statistic for groupwise heteroscedasticity (Greene 2000 p.598), Wooldridge’s test for autocorrelation (Wooldridge 2002 p.282) and Pesaran’s (2004) test for cross sectional dependence. If the null hypothesis of the aforementioned tests are rejected we use Driscoll & Kraay (1998) (D&K) SE which produces heteroscedastic and autocorrelated consistent (HAC) SE that are well calibrated even when cross-sectional dependence is present.

### 5.3.2 Model estimation methodology: Mean-group estimator

The FE-OLS technique assumes homogenous slopes, however this assumption is quite restrictive (Pesaran & Smith 1995 p.102; Smith et al. 1999 p.621; Phillips & Moon 2000 pp. 270-274). The econometric literature offers two important techniques that relax the assumption of homogenous slope coefficients i.e. the mean-group (MG) estimator suggested by Pesaran & Smith (1995) and the pooled mean group (PMG) estimator suggested by Smith et al. (1999). The MG estimator estimates N regression and averaging the coefficients and allows all coefficients to be heterogeneous (Blackbourne III & Frank 2007 p.198). The PMG estimator relies on a combination of both pooling and averaging the coefficients i.e. the intercept, the short coefficients are allowed to be heterogeneous but the long-term coefficients would be equal across all N (Ibid.). We cannot think of a theoretical argument where the long-term coefficient would be equal across all sample firms, hence the MG-estimator suggested by Pesaran & Smith (1995) is applied to relax the assumption of homogenous slopes.

\(^7\)Cointegrated links between non-stationary series lead to endogeneities in the regressors (Phillips 1995 p.1024). The OLS technique is not designed to consider long-term endogeinities (Ibid.), but the FMOLS is designed to estimate cointegrated relations by modifying the OLS with corrections that consider endogeneity and serial correlation (Lyhagen et al. 2007 p.8)
5.3.3 Model estimation methodology: Generalized method of moments

The generalized method of moments (GMM) is an estimation method that deals with endogeneity issues and it also helps to avoid parametric assumptions in the data analysis (Arellano 2003; Park et al. 2011 p.2605; Bistri 2015 pp.517-518). The construction of the PVECM (model 25) suggests that the independent variables are not strictly exogenous, hence the GMM estimator is appropriate to use (Roodman 2006 p.86).

To estimate a GMM with Engle & Granger (1987) two-step procedure is not unprecedented, for example Costantini & Martini (2010) and Polemis & Fotis (2013) use the same approach. The GMM model is estimated with a technique suggested by Arellano & Bover/Blundell &Bond (1995; 1998). The number of instruments should not be N< when estimating the GMM (Renneboog et al. 2009 p.63). For the GMM to yield unbiased estimators the instruments must be valid, which can be tested with the Sargan test. Further, the GMM estimator is very sensitive to the choice of instruments and the lag specification of the instruments. We specify our model by using the dependent variable as instruments and restrict the lags so that it begins with 2 lags and ends with p lags until the null hypothesis of the Sargan test is not rejected and the number of instruments is <N. Robust SE with Windmeijer (2005) correction is used otherwise the SE would be biased.

5.4 Panel Granger causality test

If two variables are cointegrated, granger causality must exist in at least one direction (Pfaff & Gentleman 2008 p.76). However, it is not established if granger causality exists in both directions hence it is appropriate to test for panel granger causality (PGC), i.e. if DPS can be used to predict PS and vice versa. Dumirescu & Hurlin (2012) suggest a simple Granger (1969) test based on heterogeneous panel data sets. Consider the following linear model:

\[ Y_{it} = \alpha_i + \sum_{k=1}^{K} y_{ik} Y_{it-1} + \sum_{k=1}^{K} \beta_{ik} X_{it-1} + \epsilon_{it} \]

where x and y are two stationary variables observed for N individual for T periods, K is the lag order and \( \alpha_i \) is the fixed effect. Dumirescu & Hurlin (2012) propose to ”test the homogenous non-causality test (HNC) hypothesis, by considering both the heterogeneity of the regression model and the causal relations” Dumirescu & Hurlin 2012 p.1453. The null hypothesis of the HNC test is defined as:

\[ H_0: \beta_i = 0 \quad \forall i = 1, ... , N \]

with \( \beta_{1i},... , \beta_{Ki}, \beta_i \) may differ across firms under the alternative hypothesis (i.e. model heterogeneity). Further, the PGC test assumes that under \( H_0 \) there are \( N_i < N \) individual processes with no causality from x to y. The alternative hypothesis is defined as follows:

\[ H_1: \beta_i = 0 \quad \forall i = 1, ... , N \quad \beta_i \neq 0 \quad \forall i = N_i + 1, N_i + 2, ... , N \]

where \( N_i \) is unknown but satisfies the condition 0 ≤ \( N_i / N \) < 1. If \( N_i = N \) there is no causality for any individuals in the panel, hence the null hypothesis is not rejected and if \( N_i = 0 \) there is causality for all of the individuals in the panel (Dumirescu & Hurlin 2012 p.1453). The panel statistics is the cross-sectional average of the individual Wald statistics (Ibid.).

5.5 Assumptions & limitations

We assume that every firm pays dividends twice a year. For some firms this may not be the case as they may pay once or four times per year. However, since the firms in the U.K. have a tradition of paying dividends twice a year it is reasonable to make this assumption in the data collection procedure. 

**Limitation 1.** Eq.1 states stock prices are driven by discount rates and dividends. Discount rates are not incorporated in the models, which may impact the results. However, since discount rates are i(0) imply that in the long-term, stock prices must be driven by dividends (Chen et al. 2013 p.844). The ECT coefficients in model 25 will not be affected by this assumption, but the short coefficients are most likely to be affected with weaker significance level. We leave it to future research to incorporate firm specific discount rates.

**Limitation 2.** There is a trade-off between pair in the PURT and to introduce survival bias in the sample, where we have chosen to prioritize the pair in the PURT. 

**Limitation 3.** Stock repurchases (SR) have increased rapidly during the last decade (Haw et al. 2011 p. 152), which would affect the use of the PVM (Stoi et al. 2009 p.173). This paper does not consider SR, which most likely should affect the significance level of the short coefficients.
6. Results & Discussion

6.1 Panel unit root and panel cointegration test

Table 3 reports the Pesaran (2007) CIPS test for series of LnDPS and LnPS. The results are reported for lag augmentations from P=0 to P=3 (inclusive). For both the LnDPS and LnPS series the CIPS test fails to reject the null hypothesis of that the series are I(1) for P=1,2,3 for both specifications, which indicates that the series are I(1).

Table 4 reports Pedroni (1999, 2004) and Kao (1999) panel cointegration test. Pedroni’s (1999; 2004) panel cointegration test is based on two specifications. The model specification with intercept rejects the null hypothesis of no cointegration for all eleven test statistics at a 1% significance level. The model specification with deterministic trend and intercept rejects the null hypothesis of no cointegration in nine of eleven test statistics at the 1 percent significance level. Kao (1999) test rejects the null hypothesis of no cointegration at the 1 percent significance level. Table 5 reports Westerlund (2007) panel cointegration test. Westerlund (2007) test is based on two model specifications i.e. with intercept only and with deterministic trend and intercept. All test statistics of the two model specifications reject the null hypothesis of no cointegration at the 1 percent significance level. To sum up, tiny-nine out of thirty-one test statistics rejected the null hypothesis of no cointegration when applying both the residual and the structural based panel cointegration tests. This indicates that the series of LnDPS and LnPS are cointegrated.

6.2 Panel vector error correction model

Table 6 reports the results of the specified PVECM. The PVECM is estimated with the three estimation methodologies. The FE-OLSa model estimates the first and second stage in Engle & Granger (1987) two-step procedure with FE-OLS, while the FE-OLSB model estimates the first stage with FM-OLSP and the second stage with FE-OLS.

---

**Table 3**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lag</th>
<th>Intercept</th>
<th>Zt-bar</th>
<th>Intercept &amp; trend</th>
<th>Zt-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>LnDPS</td>
<td>P=0</td>
<td>0.758</td>
<td>-0.325</td>
<td>1.902</td>
<td>2.923</td>
</tr>
<tr>
<td></td>
<td>P=1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=2</td>
<td>5.765</td>
<td>7.995</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=3</td>
<td>3.391</td>
<td>9.293</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LnPS</td>
<td>P=0</td>
<td>0.127</td>
<td>2.723</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=1</td>
<td>-0.513</td>
<td>1.786</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=2</td>
<td>0.485</td>
<td>2.426</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P=3</td>
<td>1.285</td>
<td>4.952</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: H0: series is I(1).

---

**Table 4**

<table>
<thead>
<tr>
<th>Pedroni (1999:2004) 1</th>
<th>Intercept 2</th>
<th>Intercept &amp; trend 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test statistic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel v-Statistic</td>
<td>14.30*</td>
<td>5.17*</td>
</tr>
<tr>
<td>Panel rho-Statistic</td>
<td>-11.48*</td>
<td>-6.50*</td>
</tr>
<tr>
<td>Panel PP-Statistic</td>
<td>-10.32*</td>
<td>-8.91*</td>
</tr>
<tr>
<td>Panel ADF-Statistic</td>
<td>-12.08*</td>
<td>-12.38*</td>
</tr>
<tr>
<td>Group rho-Statistic</td>
<td>-6.29*</td>
<td>-6.63</td>
</tr>
<tr>
<td>Group PP-Statistic</td>
<td>-10.17*</td>
<td>7.65*</td>
</tr>
<tr>
<td>Group ADF-Statistic</td>
<td>-13.91*</td>
<td>-13.28*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kao (1999) 4</th>
<th>Test statistic</th>
<th>T-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF</td>
<td>-0.17*</td>
<td></td>
</tr>
</tbody>
</table>

Note: 1 H0: No cointegration between the series LnDPS & LnPS. Automatic lag selection is based on AIC. 2 Trend assumption: No deterministic trend. 3 Trend assumption: Deterministic intercept & trend. 4 H0: No cointegration between the series LnDPS & LnPS. Automatic lag selection is based on AIC.

* significant at the 1 % level
Table 5
Westerlund (2007) panel cointegration test

<table>
<thead>
<tr>
<th>Test statistic</th>
<th>Intercept 1</th>
<th>Intercept &amp; trend 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Z-value</td>
<td>Z-value</td>
</tr>
<tr>
<td>Gi</td>
<td>-15.65*</td>
<td>-14.00*</td>
</tr>
<tr>
<td>Gz</td>
<td>-6.92*</td>
<td>-12.52*</td>
</tr>
<tr>
<td>P1</td>
<td>-12.36*</td>
<td>-14.40*</td>
</tr>
<tr>
<td>Ps</td>
<td>-21.43*</td>
<td>-29.31*</td>
</tr>
</tbody>
</table>

Note: H0: No cointegration between the series LnDPS & LnPS. Automatic lag selection is based on AIC. * adds a constant to the cointegration relationship. ** allows for a deterministic trend in the cointegration relationship.

* significant at the 1 % level

Model MGb estimates the first and second stage with Pesaran & Smith (1995) MG estimator, while model MGa estimates the first stage with the FM-OLS technique and the second stage with Pesaran & Smith (1995) MG estimator. The GMM model estimates the first stage with FM-OLS and applies

Table 6
The panel vector error correction model

(25a) \( \Delta \text{LnPS}_{1} = \alpha_{11} + \beta_{11} \Delta \text{LnDPS}_{1} + \theta_{11} \Delta \text{LnPS}_{1-1} + \theta_{12} \Delta \text{LnDPS}_{1-2} + \theta_{13} \Delta \text{LnPS}_{1-3} + \lambda_{11} \text{ECT}_{1} + \mu_{11} \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Exp. sign</th>
<th>FE-OLS 1</th>
<th>FE-OLS 2</th>
<th>MGa</th>
<th>MGb</th>
<th>GMM 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>+/-</td>
<td>0.0223**</td>
<td>0.0220**</td>
<td>0.0111*</td>
<td>0.0112*</td>
<td>0.0256*</td>
</tr>
<tr>
<td>( \Delta \text{LnDPS}_{1} )</td>
<td>+</td>
<td>0.0634</td>
<td>0.0634</td>
<td>0.5108*</td>
<td>0.4771*</td>
<td>-0.0412</td>
</tr>
<tr>
<td>( \Delta \text{LnDPS}_{1-1} )</td>
<td>-</td>
<td>-0.0439**</td>
<td>-0.0451**</td>
<td>-0.0070</td>
<td>-0.0338</td>
<td>-0.0636*</td>
</tr>
<tr>
<td>( \Delta \text{LnDPS}_{1-2} )</td>
<td>-</td>
<td>-0.0175**</td>
<td>-0.0184**</td>
<td>-0.0994**</td>
<td>-0.0980**</td>
<td>-0.0042</td>
</tr>
<tr>
<td>( \Delta \text{LnDPS}_{1-3} )</td>
<td>-</td>
<td>-0.0379</td>
<td>-0.0389</td>
<td>-0.0038</td>
<td>0.0143</td>
<td>-0.1242*</td>
</tr>
<tr>
<td>( \text{ECT}_{1} )</td>
<td>-</td>
<td>-0.1104*</td>
<td>-0.1089*</td>
<td>-0.2146*</td>
<td>-0.1712*</td>
<td>-0.1035*</td>
</tr>
<tr>
<td>Obs</td>
<td>8,661</td>
<td>8,661</td>
<td>8,661</td>
<td>8,661</td>
<td>8,661</td>
<td></td>
</tr>
<tr>
<td>(i) Wald-test statistic</td>
<td>14.11**</td>
<td>14.48*</td>
<td>75.85*</td>
<td>64.86*</td>
<td>26.34*</td>
<td></td>
</tr>
<tr>
<td>(ii) Wald-test statistic</td>
<td>28.77**</td>
<td>27.73*</td>
<td>361.22*</td>
<td>239.63*</td>
<td>28.75*</td>
<td></td>
</tr>
<tr>
<td>Sargan statistic</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>225.07</td>
</tr>
</tbody>
</table>

(25b) \( \Delta \text{LnDPS}_{1} = \alpha_{41} + \gamma_{11} \Delta \text{LnPS}_{1-1} + \gamma_{12} \Delta \text{LnPS}_{1-2} + \gamma_{13} \Delta \text{LnPS}_{1-3} + \lambda_{11} \text{ECT}_{1} + \mu_{12} \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Exp. sign</th>
<th>FE-OLS 1</th>
<th>FE-OLS 2</th>
<th>MGa</th>
<th>MGb</th>
<th>GMM 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>+/-</td>
<td>0.0092</td>
<td>0.0974</td>
<td>0.0140*</td>
<td>0.0140*</td>
<td>0.0057***</td>
</tr>
<tr>
<td>( \Delta \text{LnPS}_{1-1} )</td>
<td>+</td>
<td>0.0960*</td>
<td>0.0934*</td>
<td>0.0202</td>
<td>0.0136</td>
<td>0.0354</td>
</tr>
<tr>
<td>( \Delta \text{LnPS}_{1-2} )</td>
<td>+</td>
<td>0.1022*</td>
<td>0.1003*</td>
<td>0.0516*</td>
<td>0.0514*</td>
<td>0.1989*</td>
</tr>
<tr>
<td>( \Delta \text{LnPS}_{1-3} )</td>
<td>+</td>
<td>0.0540</td>
<td>0.0529</td>
<td>0.0272**</td>
<td>0.0291**</td>
<td>0.1379**</td>
</tr>
<tr>
<td>( \text{ECT}_{1} )</td>
<td>+</td>
<td>0.1309*</td>
<td>0.1335*</td>
<td>0.1393*</td>
<td>0.1350*</td>
<td>0.3155*</td>
</tr>
<tr>
<td>Obs</td>
<td>8,661</td>
<td>8,661</td>
<td>8,661</td>
<td>8,661</td>
<td>8,661</td>
<td></td>
</tr>
<tr>
<td>(i) Wald-test statistic</td>
<td>4.03**</td>
<td>3.93**</td>
<td>17.24*</td>
<td>17.52*</td>
<td>22.72*</td>
<td></td>
</tr>
<tr>
<td>(ii) Wald-test statistic</td>
<td>12.27**</td>
<td>12.50*</td>
<td>152.87*</td>
<td>166.68*</td>
<td>40.69*</td>
<td></td>
</tr>
<tr>
<td>Sargan statistic</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>223.86</td>
</tr>
</tbody>
</table>

Note: *Estimates eq.23 and eq.25,1 with FE-OLS D & K (1998) SE are used, which produces HAC consistent SE even when cross sectional dependence is present. ** Estimates eq.23 with FM-OLS and eq.25,1 with FE-OLS D & K (1998) SE are used. \( \hat{ \alpha } \) estimates eq.23 with the MG estimator. \( \hat{ \lambda } \) estimates eq.23 with FE-OLS and eq.25,1 with the MG estimator. * Estimates eq.23 with FM-OLS and eq.25,1 with GMM. Robust SE with Windmeijer (2005) correction is used. \( \hat{ \alpha } \) = \( \Delta \text{LnPS}_{1} = \Delta \text{LnDPS}_{1} = \Delta \text{LnPS}_{1-1} = \Delta \text{LnPS}_{1-2} = \Delta \text{LnPS}_{1-3} = \lambda_{12} \text{ECT}_{1} + \mu_{12} \). The instruments are valid. T-statistics in the parentheses.

*, **, *** significant at the 1, 5 and 10 % level

The short & long-term interdependencies between stock prices and dividends:
A panel vector error correction approach
the GMM to the second stage of Engle & Granger’s (1987) two-step procedure. The null hypothesis of the Sargan test is not rejected hence the instruments are valid.

In model 25, the coefficient of $\Delta \text{LnDPS}$ is positive for the FE-OLS and the MG models and negative in the GMM model, but it is only significant in the MG models. The sum of the coefficients of $\Delta \text{LnDPS}$, $\Delta \text{LnDPS}_{-1}$, and $\Delta \text{LnDPS}_{-3}$ are, as expected, negative across all estimated models. The coefficient of ECT$_{-1}$ in model 25 is $< 0$ and significant across all estimated models. The (i) Wald-test statistics in model 25 show that the null-hypothesis of no short-term effects is rejected for all estimated models, which implies dividends have, on average, short-term effects on stock prices. Examining the (ii) Wald-test statistics in model 25 reveals that the null hypothesis of no adjustment to the long-term relationship is rejected across all estimated models, which implies that stock prices, on average, adjust to the long-term relationship.

In model 25, the coefficients of the lagged $\Delta \text{LnPS}$ is as expected positive, however it is only significant in two out of three lagged coefficients across all estimated models. The ECT$_{-1}$ in model 25 is $< 0$ and significant across all estimated models. The null hypothesis of no short-term effects is rejected for all estimated models, which implies that stock prices have, on average, short-term effects on dividends. The (ii) Wald-tests statistics in model 25 reveal that the null-hypothesis of no adjustment to the long-term relationship is rejected across all estimated models, which implies that dividends, on average, adjust to the long-term relationship.

Since, $H_1$ is accepted $H_2$ is also accepted, which indicates that there, on average, are long and short-term interdependencies. Despite the different assumptions about the data in the employed estimation methodologies the inference regarding the short and long-term dependencies remains the same, which indicates that the results are robust. We also estimate the models with different number of lag lengths ($k=1,k=2,k=4$) (recall discussion in section 5.3 regarding the choice of the lag length in the original model), which does not affect inference of the research hypotheses. The results of these estimates can be provided by request.

6.3 Panel Granger causality test

Table 7 reports the results of Dumirescu & Hurlin (2012) PGC test. The tests are performed with first differenced variables. The results are reported for lag augmentations from $P=1$ to $P=3$ (inclusive). The results showing that the null hypothesis of that $\text{LnDPS}$ does not homogeneously cause $\text{LnPS}$ is rejected regardless of the number of lags included. The null hypothesis of that $\text{LnPS}$ does not homogeneously cause $\text{LnDPS}$ is rejected regardless of the number of lags included. The outcome of the PGC tests reveals that $\text{LnDPS}$, on average, can be used to predict $\text{LnPS}$ and vice versa, which supports the $H_{1,1}$ and supports the notion that there, on average, are short-term (inter) dependencies.

6.4 Discussion of the results

The statistical perspective

As expected, the significance level of the short coefficients in model 25 is weaker significance levels than the ECT, which may be due to the limitations (see section 5.5). There are some differences amongst the estimation methodologies in the magnitudes of some coefficients. The most striking difference is that a 10 percent increase in $\text{DPS}$, will lead to on average approximately a 5 percent increase in stock price in the MG-model, while the FE-OLS and the GMM model suggest that a 10 percent increase in $\text{DPS}$ would, on average, have no impact on current stock price. However, when we apply HAC SE on the FE-OLS, b model the coefficient of $\Delta \text{LnDPS}$, becomes significant at the 5 percent level i.e. a 10 percent increase in $\text{DPS}$, will lead to, on average, approximately a 0.64 percent increase in stock price (see appendix C).

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Lag</th>
<th>Zbar-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{LnDPS}$ does not homogeneously cause $\text{LnPS}$</td>
<td>$P=1$</td>
<td>1.80*</td>
</tr>
<tr>
<td>$\text{LnPS}$ does not homogeneously cause $\text{LnDPS}$</td>
<td>$P=1$</td>
<td>25.41*</td>
</tr>
<tr>
<td>$\text{LnDPS}$ does not homogeneously cause $\text{LnPS}$</td>
<td>$P=2$</td>
<td>2.60*</td>
</tr>
<tr>
<td>$\text{LnPS}$ does not homogeneously cause $\text{LnDPS}$</td>
<td>$P=2$</td>
<td>8.40*</td>
</tr>
<tr>
<td>$\text{LnDPS}$ does not homogeneously cause $\text{LnPS}$</td>
<td>$P=3$</td>
<td>2.87*</td>
</tr>
<tr>
<td>$\text{LnPS}$ does not homogeneously cause $\text{LnDPS}$</td>
<td>$P=3$</td>
<td>39.10*</td>
</tr>
</tbody>
</table>

* significant at the 1% level
The relation to existing literature

The ECT in table 6 reveals that the speed of adjustment is not very fast (about 10-30 percent per time unit), which is in line with Saltugu (1998), Sung & Urrutia (1995) and Nasseh & Strauss (2004). The estimated ECT in this study is clearly a violation against the efficient market hypothesis (Saltugu 1998 p. 373). However, because we do not incorporate discount rates and SRs one should be very careful to make any statements or conclusions in this matter.

This paper supports the notion that stock prices and dividends are cointegrated, which is in line with Campbell & Shiller (1987). On the contrary, the results of this study are not in line with index data studies like McMillan (2010b), who also uses data from U.K. However, it is in line with Goddard et al. (2008) that use firm level data. This indicates that the PVM is more likely to hold on a firm level, as suggested by Samuelsson’s dictum, than on aggregated macro level. However, Vivian & Wohar (2010) use index data and find cointegration. The difference lies in data. McMillan (2010b) almost covers the same period as this paper, whereas Vivian & Wohar (2010) covers 1900-2004. Recall, that the stock-dividend relationship has been questioned during the 1990s and 2000s, this paper shows that the relationship holds, on average, at an individual firm level, which implies that the PVM can be used in stock-valuations. Comparing the U.K. studies with the U.S. studies there seems to be similar pattern, for example the panel data study of Nasseh & Strauss (2004) find cointegration, while the index data study by Esteve & Prats (2010) fails to find linear cointegration.

Contrary to Sung & Urrutia (1995) who find that the short relationship from dividends to stock price is iak, this paper finds a stronger short-term relationship. A possible explanation could be that Sung & Urrutia (1995) use index data. On the other hand, Nasseh & Strauss (2004 p.2014) also find that the relationship from dividends to stock prices is iak when using panel data. This indicates that there might be some differences in the stock-dividend relationship between countries. The difference might be due that SR represents a larger share of total payouts in the U.S. than in the U.K. and that Nasseh & Strauss (2004) & Sung & Urrutia (1995) do not consider SR.

Practical implications

For investors, this study shows that, on average, stock prices can be used to predict dividends and vice versa. Due to cointegration and the interrelationship, investors know that stock prices can just rise to a certain point without an increase in dividend payments and vice versa. The evidence for cointegration supports validity of PVM. The empirical evidence for a cointegration relationship at an individual firm level also suggests that investors should be able to distinguish whether a rise in the stock price is sustainable. The fact that this paper finds, on average, short dependencies indicates that the PVM is usable even in the short term.

7. Conclusion

This paper examines the short and long-term interdependencies between stock prices and dividends. We utilize firm level data from FTSE ALL SHARE in the U.K. for the period 1990-2014 and apply a PVECVM estimated with Engle & Granger’s (1987) two-step procedure.

The results show that there is a bi-directional long-term relationship between stock prices and dividends i.e. there is an adjustment process at work when a disequilibrium occurs. We also find that there is bi-directional short relationship and that stock prices can be used to predict dividends and vice versa. Despite the different assumptions about the data in the employed estimation methodologies the inference regarding the short and long-term dependencies remains the same, which indicates that the results are robust. Further, the results show that the PVM and LDM still is relevant theoretical framework when analysing the stock price-dividend relationship from a stock valuation perspective.

There are suggestions for future research. For example, it could be of interest to divide the firms into industries in order to examine the differences in the magnitude of the error mechanism or the short-term effects, due to the impact of the influence of industry.
References

Drukker, D.J. & Kratz, A.C. 1999, "Consistent covariance matrix estimation with spatially dependent panel data", The review of economics and statistics, vol. 81, no. 4, pp. 590-591
APPENDIX A

### Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
</table>

### Pearson’s correlation

#### Statistical outputs from STATA 13 and E-views 8.1

#### PURT table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>lags</th>
<th>zt-bar</th>
<th>p-value</th>
<th>t-bar</th>
</tr>
</thead>
</table>

#### Null CIPS test: series is t(1).

The short & long-term interdependencies between stock prices and dividends: A panel vector error correction approach.
Cointegration table 4

Pedroni Residual Cointegration Test
Series: LNINPS, LNINDPS
Sample: 1995Q1 2014Q2
Included observations: 975
Cross-sections included: 228
Null Hypothesis: No cointegration
Trend assumption: Deterministic intercept and trend

Automatic lag length selection based on AIC with lags from 6 to 10
Average absolute value of the off-diagonal elements = 0.28

Wooldridge test for autocorrelation in panel data
Ho: No first order autocorrelation
Pr > F = 0.0000

Regression with Driscoll-Kraay standard errors
Method: Fixed-effects regression
Number of obs = 9661
Group variable: firm
Number of groups = 228
Group variable selected length and Bulletin heading

PE-OLS

FE-OLSa

Fixed-effects (within) regression
Number of obs = 8661
Group variable: firm
Number of groups = 228

Cointegration table 5

Calculating Almon ECM panel cointegration tests
Results for H0: no cointegration
With K 228 series and 1 covariate
Average AIC selected lag length: 1.21
Average AIC selected lead length: 1.37

Cointegration table 6

PVECM table 6

FE-OLSb

Dependent Variable: LNINPS
Method: Panel Fully Modified Least Squares (FMOLS)
Panel corrected (adj): 1995Q2 2014Q2

EKENOMIKUM: Centre for economic sciences, May 2015

R. Persson
Wooldridge test for autocorrelation in panel data

Average absolute value of the off-diagonal elements = 0.288

---

Test:

(1) \text{dlninPS} = 0
(2) \text{dlllninPS} = 0
(3) \text{dllllninPS} = 0

Prob > F = 0.0000

---

MG a

Peasara & Smith (1999) Mean Group estimator

Mean Group type estimation

Number of obs = 8661
Number of groups = 228

---

MG b

Peasara & Smith (1999) Mean Group estimator

All coefficients represent averages across groups (group variable: firm)
Coefficient averages computed as unweighted means

Mean Group type estimation

Number of obs = 8661
Number of groups = 228

---

The short & long-term interdependencies between stock prices and dividends: A panel vector error correction approach
### Panel Granger causality table 7

<table>
<thead>
<tr>
<th>R. Persson</th>
<th>ch11(217)</th>
<th>232.8405</th>
<th>ch11(219)</th>
<th>232.7001</th>
<th>0.3459</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob &gt; ch12</td>
<td>0.3601</td>
<td>0.3459</td>
<td>Prob &gt; ch12</td>
<td>0.3459</td>
<td>0.3459</td>
</tr>
</tbody>
</table>

#### Pairwise Dumitrescu Hurlin Panel Causality Tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DLNPS does not homogeneously cause DLNPS</td>
<td>W-Statistic for DLNPS</td>
<td>7.10817</td>
<td>1.9E-12</td>
<td>0.0000</td>
</tr>
<tr>
<td>DLNPS does not homogeneously cause DLNPS</td>
<td>N-Statistic for DLNPS</td>
<td>3.70573</td>
<td>25.4217</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

#### Panel Granger causality table 7

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DLNPS does not homogeneously cause DLNPS</td>
<td>W-Statistic for DLNPS</td>
<td>7.10817</td>
<td>1.9E-12</td>
<td>0.0000</td>
</tr>
<tr>
<td>DLNPS does not homogeneously cause DLNPS</td>
<td>N-Statistic for DLNPS</td>
<td>3.70573</td>
<td>25.4217</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

#### Pairwise Dumitrescu Hurlin Panel Causality Tests

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DLNPS does not homogeneously cause DLNPS</td>
<td>W-Statistic for DLNPS</td>
<td>7.10817</td>
<td>1.9E-12</td>
<td>0.0000</td>
</tr>
<tr>
<td>DLNPS does not homogeneously cause DLNPS</td>
<td>N-Statistic for DLNPS</td>
<td>3.70573</td>
<td>25.4217</td>
<td>0.0000</td>
</tr>
</tbody>
</table>