The Theory of Optimal Dividend Policy:
An Intertemporal Approach and Empirical Evidences

Ming-Jang Weng*, Min-Shann Tsai† and Jyh-Lin Wu‡

We propose a theoretical model of optimal dividends based on microfoundation to investigate the relationships between a firm’s expected stream of future net earnings and changes in stockholders’ equity or the smoothing component of the dividend policy. The signaling effects of changes in equity and of dividend policy to the expectations of the future net earning stream are then discussed and tested. Empirical findings using HP and Citi Bank’s quarterly data in the past decade provide robust support for the theoretical predictions of actual changes in stockholders’ equity and to the smoothing component of cash dividends.

JEL classification: C22; G30; G35

Key words: dividend smoothing, net earnings, signaling theory, VAR

* Associate Professor, Department of Applied Economics, National University of Kaohsiung, Taiwan. Correspondence address: 700 Kaohsiung University Rd., Nan-Tzu Dist., Kaohsiung 811, Taiwan. Phone: 886-7-5919186, Fax: 886-7-5919319, E-mail: mjjwn@nuk.edu.tw.
† Assistant Professor, Department of Banking and Finance, National Chi Nan University, Taiwan. E-mail: mstsai@ncnu.edu.tw
‡ Professor, Department of Economics, National Chung Cheng University, Taiwan. E-mail: ecdjlw@ccunix.ccu.edu.tw
Introduction

Dividend policy has been confusing to financial researchers for quite a long time. Conceptually, managers seem to prefer retaining earnings in order to reduce financial costs of new investment opportunities. However, firms always declare dividends, which cause a reduction in useful capital, but they simultaneously pay higher costs to issue new stocks or bonds in the market to fund new investments. Signaling theory is commonly used to explain these confusing aspects of dividend policy. Since managers possess richer information of the firm’s future earnings than outsider investors, then a firm with a generous dividend policy signaling better business prospects will usually present a higher stock price in the market. In addition, public and/or fund managers may refer to the dividend policy to determine the quality of the company due to its information contents (Miller and Rock, 1985). The market price will then promptly react to a firm’s dividend decisions. Therefore, a company with higher dividends will see a better stock price.

Another useful viewpoint of signaling theory indicates that a firm’s announcement of an increase in dividends will cause a positive abnormal return, since it may imply that the manager is optimistic about the firm’s future earnings. March and Merton (1987) and Kao and Wu (1994) proposed a rational signaling model to investigate the information contents of dividends. The model shows that dividends
are smoothed according to the conditions of the manager’s expectations of future earnings. A number of studies did find evidence supportive of signaling theory (Higgins, 1987; Divecha and Morse, 1983; John and Williams, 1985; Miller and Rock, Bar-Yosef and Huffman, 1986). Empirical results support signaling theory, but the econometric specification is perplexing in the existing literature on dividends. In addition, traditional signaling theory only used dividend smoothing to investigate the information contents and explain the phenomenon of dividend declarations, whereas, it is hard to determine if the dividend policy is optimal for shareholders.

The main purpose of this paper is to complete the theory of economic sense of dividend policy. We deduce a model of optimal dividend policy by utilizing an intertemporal approach. Our model not only explains the dividend status as addressed in existing signaling models, but also indicates if the dividend policy is optimal. As in financial theory, regardless of what the financial decisions are, the manager’s purpose is always to maximize the dividend-based utility (or wealth) of shareholders. Since dividends are the most important factor of wealth transferred from a firm to shareholders, then a dividend policy model needs to seriously take shareholders’ utility into account. In our specification, the manager needs to design the dividend policy so as to maximize the representative shareholders’ lifetime utility under the intertemporal restrictions of the firm’s equities and earnings. We then show that the
optimal dividend policy depends on the manager’s expectation of future net earnings, opportunity costs of the company’s endowed equities, interest rates, and shareholders’ utility parameter, i.e., the time preference. Since the manager’s expectation of future net earnings as stated in those signaling models is one of the factors influencing optimal dividend policy, the main implication of traditional signaling models can be adequately derived and interpreted by our model. However there are some differences between our model and traditional signaling theory. In traditional signaling theory, the target dividend is smoothed according to the conditions of the manager’s expectation of future earnings. In our model, the optimal dividend policy depends on the expectation of future “net” earnings instead, and our theory also considers the required returns (opportunity costs) of the initial equity investment. That is, if a firm expects positive future earnings, but they will be less than the required payoff on the initial equity, then it should lower its dividend since it has a negative future “net” earning. This may explain why a firm might possess positive future earnings, but still declare a lower dividend and have a lower stock price in the market. In addition, our model indicates that the optimal dividend policy will be greater (or less) than the smoothed dividend according to if the time preference of shareholders is less (or greater) than the interest rate, which means stockholders are impatient (or patient) about receiving a cash dividend if the outside
investment environment is hot (or cold).

Since our optimal dividend policy is based upon maximization of shareholders' utility, we can further test if a firm’s manager has really made optimal dividend decisions over time by looking at the empirical dividend strategies. We then verify the theoretical findings with an empirical study of HP (Hewlett-Packard) and Citi Bank. The econometric treatments of empirical tests of the present-value model, originally developed by Campbell (1987) and Campbell and Shiller (1987), allow us to map the theory-oriented predictions to the actual changes in equity.

This paper is organized as follows. In section 2, we present the theory of the intertemporal approach to dividend smoothing, followed by a brief introduction to econometric procedures in testing the present-value model in section 3. In section 4, the VAR (vector autoregression) estimation of the model is presented using HP and Citi Bank’s quarterly data over 1992Q1-2002Q2, due to the availability of the data resources. In section 5, conclusions are summarized.

The Model

We extend the intertemporal model to decisions of optimal dividend policy in corporate finance. The intertemporal approach was first proposed in the 1980s by Buiter (1981), Sachs (1981), Svensson and Razin (1983), and Obstfeld (1986), and
has been well applied to several areas of international financial research, such as
capital mobility, consumption smoothing, and the validity and sustainability of
theoretical forecasts of actual current accounts (e.g., Sheffrin and Woo, 1990; Otto,
1992; Ghosh, 1995a; Ghosh and Ostry, 1995; Cashin and McDermott, 1998; Bergin
and Sheffrin, 2000). The method has also been applied to studies of bond prices
(Campbell and Shiller) and in dealing with tax smoothing/government deficits
(Huang and Lin, 1993; Ghosh, 1995b).

The scenario of our model is presented below. Consider a representative firm in
which a change in the firm’s equity can be interpreted as earnings minus dividends,
i.e., as retained earnings in period $t$. The retained earnings thus are defined as
earnings after taxes and debt minus dividends. The manager is facing an
intertemporal equity constraint, in per capita terms, as follows,

$$\Delta S_t = S_{t+1} - S_t = rS_t + Y_t - D_t = RE_t$$ (1)

where $S_t$ denotes the equity at the beginning of time $t$; $r$ is the market risk-free
interest rate; $Y_t$ is the firm’s net earnings ($i.e.$, earnings after taxes, debt, and equity
interest during period $t$); $D_t$ is the cash dividend to be determined and distributed
per share at the end of time $t$; and $RE_t$ is the firm’s retained earnings. Therefore,$rS_t$ can be viewed as the required earnings, $i.e.$, the so-called opportunity cost, of a
firm with an initial collection of investment \( S_t \). \( Y_t \) can then also be interpreted as the excess earnings from \( S_t \).

Now, the manager has to determine how much of a cash dividend per share should be proposed to the board of directors so as to fulfill stockholders’ expectations. Therefore, we assume there exists an infinitely-lived representative stockholder whose lifetime utility, \( U \), known to the manager, at time 0 is given by,

\[
U = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t u(D_t) \right\}, \tag{2}
\]

where \( u(\cdot) \), the instantaneous preference, is assumed to represent increasing and concave in cash dividends, \( D \), i.e., \( u'(\cdot) > 0, \ u''(\cdot) < 0 \). \( \beta \), the subjective discount factor which ranges from 0 to 1, means that utilities are valued lower, the later they are received. \( E \) indicates the expectation of a rational stockholder based on all information available to date. That is, the manager considers not only shareholders’ myopic satisfaction about their investment returns, but also their long-run comfort of dividend flow; in other words, the firm’s prospects of future development are considered as well.

Herein, we further assume that information on net earnings, \( Y_t \), are exogenously given to the manager when making a decision about the cash dividend (see Nakamura and Nakamura, 1985). Forward iteration of equation (1) with the transversality
condition held gives us,

\[ D_i^p = \frac{r}{1+r} \mathbb{E} \left( \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} D_{t+1} \right) = rS_i + \frac{r}{1+r} \mathbb{E} \left( \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} Y_{t+1} \right) = rS_i + Y_i^p, \quad (3) \]

where \( D_i^p \) and \( Y_i^p \) are defined as the “permanent dividend” and “permanent net earnings” (analogous to Nakamura and Nakamura), respectively. Equation (3) thus states that the expected discounted stream of the dividend, or the permanent dividend, must equal the expected sum of the discounted future net earnings (i.e., permanent net earnings) plus required earnings (i.e., initial interest claims on equities). The optimization problem faced by a firm’s manager now is to maximize \( U \) in equation (2), subject to the intertemporal equity constraint of equation (1). With no “Ponzi game”\(^1\) assumption and a CRRA (constant relative risk aversion) form of the utility function, i.e.,

\[ u(D_i) = \frac{D_i^{1-\sigma} - 1}{1 - \frac{1}{\sigma}}, \quad \sigma > 0, \]

where \( \sigma \) is the elasticity of intertemporal substitution, it turns out that the optimal dividend policy necessarily follows\(^2\):

\[ D_i^* = \frac{1}{\theta} \left[ rS_i + \frac{r}{1+r} \mathbb{E} \left( \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} Y_{t+1} \right) \right] = \frac{1}{\theta} \left[ rS_i + Y_i^p \right] = \frac{1}{\theta} D_i^p, \quad (4) \]

\[ \theta = \frac{r}{(1+r) - (1+r)^{\sigma} \beta^{\sigma}}. \]
where $D_t^*$ denotes the optimal path for the cash dividend, and $\theta$ reflects the dividend-tilting effect. Equation (4) indicates that the optimal cash dividend to be distributed is proportional to the permanent net earnings, i.e., the permanent stream of expected net earnings, plus its required returns. This differs from traditional signaling theory in two ways. First, even though equation (4) is analogous to Kao and Wu’s relation between their target dividend and permanent earnings, their target dividend is smoothed according to the conditions of the manager’s expectation of future earnings; however, in our theory, the optimal dividend policy depends on the expectation of future “net” earnings instead. Second, equation (4) also considers the opportunity cost of initial equity holdings. That is, even conventional signaling models rely on a similar argument of permanent earnings, but lack consideration of the required returns from the preexisting stock of equity. This implies that a firm should lower its dividends even though it expects positive future earnings that are less than the required payoff on the initial equity. Moreover, from $\theta$, we know that if

$$\beta = \frac{1}{(1 + \rho)} \leq \frac{1}{(1 + r)}$$

where $\rho$ is the stockholder’s rate of time preference, this implies $\theta \leq 1$, which in turn states that the discrepancy between the objective market rate of interest, $r$, and subjective rate of time preference, $\rho$, determines whether the firm’s manager chooses a pattern of per capita cash dividends which tilts towards the present ($\theta < 1 \Rightarrow D_t^* > D_t^p$) or the future ($\theta > 1 \Rightarrow D_t^* < D_t^p$). If $\theta = 1$, i.e.,
\( D_i^* = D_i^p \), there is no tilting component to the dividend, meaning that the dividend is smoothed. In other words, \( D_i^* \) will be greater (or less) than the smoothed dividend, \( D_i^* \biggr|_{0<1} \), as \( \theta < (>)I \), \( i.e., \theta > (_)r \), which means that the stockholders are impatient (or patient) about acquiring the cash dividend. Combining equations (1) and (3) gives us

\[
\Delta S_t = S_{t+1} - S_t = (Y_t - Y_t^p) - (D_t - D_t^p) \quad \Box
\]  

Equation (5) shows that changes in equity result from two components. The first is the dividend-smoothing motive, which stabilizes dividends in the face of shocks to net earnings, that is, the representative stockholder regards changes in equity as a buffer to smooth dividends in the face of shocks to the net earning stream. The second component is the dividend-tilting motive, whereby the manager tilts the dividend towards the present or future, depending on the relative magnitudes of shareholders’ subjective discount rate and the market objective interest rate. In order to remove the dividend-tilting effect, we then define the associated optimal change in equity under dividend smoothing as:

\[
\Delta S_t^* = rS_t + Y_t - \theta D_t^* = Y_t - Y_t^p. \tag{6}
\]

Then, with some manipulation\(^3\) we end up with

\[
\Delta S_t^* = -\mathbb{E} \left( \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} \Delta Y_{t+i} \right), \tag{7}
\]
where $\Delta Y_{t+i} = Y_{t+r} - Y_{t+i-1}$. Equation (7) expresses changes in equity as the negative sum of the discounted present value of expected future changes in net earnings. This implies that a firm will gain in equity only if it expects its flow of net earnings to fall in the future, which means that a rational manager will “save for a rainy day” (i.e., lower the cash dividend) when expecting the firm’s prospective profits to decline, and vice versa. The intuition behind equation (7) is quite straightforward. A firm’s manager will propose a stingy dividend policy (i.e., retain more earnings now) if he/she foresees that the firm’s net earnings will worsen in the future. This results in an increase in the equity, which gives us a useful application to signaling theory. Once a firm’s equity increases (i.e., it is saving more earnings in its own pocket), the market will then take it as a “signal”, or “warning”, to predict that the firm may profit less in the future, and vice versa. Therefore, the optimal smoothed dividend, i.e., without the tilting effect, can be easily shown to be

$$0D^*_r = rS_t + Y_t + E \left( \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} \Delta Y_{t+i} \right).$$

(8)

This means that the current optimal smoothed dividend depends not only on both the required and excess earnings in this period but also on expectations of future performance of excess earnings. More precisely, a rational firm should not make its dividend policy only by looking at its current earnings. Therefore, the optimal smoothed dividend should even be negative when expecting a declining net earning
stream, although some positive net earnings and required earnings exist.

**Econometric Procedures in Estimating the Optimal Dividend Flow**

From equation (7), we know that the current change in equity serves as a good predictor of the future income stream of net earnings. Therefore, in discussing whether or not the change in equity is optimal, we can also justify a firm’s dividend policy over time. In order to estimate this optimal change in equity, we then utilize techniques developed by Campbell to first estimate an unrestricted $p$-th order of VAR in $\Delta Y_t$ and $\Delta S_t^*$ as follows,

$$
\begin{bmatrix}
\Delta Y_t \\
\Delta S_t^*
\end{bmatrix} = 
\begin{bmatrix}
a(L) & b(L) \\
c(L) & d(L)
\end{bmatrix}
\begin{bmatrix}
\Delta Y_{t-1} \\
\Delta S_{t-1}^*
\end{bmatrix} + 
\begin{bmatrix}
v_{1t} \\
v_{2t}
\end{bmatrix},
$$

or, (9)

$$
Z_t = \Phi Z_{t-1} + V_t,
$$

where $\Delta S_t^*$ is the actual dividend-smoothing component of the change in equity as in equation (6)$^4$. For simplicity, we assume that $p = 1$ for later derivation of econometric procedures for the present-value model, but will proceed to choose an optimal $p$-th order of VAR with some criteria when performing the empirical study. $v_1$ and $v_2$ are disturbances with conditional means of zero; and $Z_t = [\Delta Y_t \Delta S_t^*]'$ and $\Phi$ are the transition matrices of the $p$-th order VAR. We can now discuss the implications of equation (9) for the VAR system. One implication is that $\Delta S_{t-1}^*$ must
negatively Granger-causes $\Delta Y_i$, i.e., $b < 0$, according to the discussion and equation (7). The signaling effect can thus be tested. Then the $i$-period-ahead expectation on $Z_i$ is simply

$$E(Z_{t+i}) = \Phi'Z_t.$$ 

Therefore, $E_t(\Delta Y_{t+i}) = [1 \ 0] E_t(Z_{t+i}) = [1 \ 0] \Phi' Z_t$. The infinite sum in equation (7) is thus

$$\Delta \hat{S}_i = -[1 \ 0] \frac{1}{1 + r} \hat{\Phi} \left[I_2 - \frac{1}{1 + r} \hat{\Phi} \right]^{-1} \left[ \Delta Y_i \ \Delta S_i^* \right] = \begin{bmatrix} \hat{\Gamma}_Y & \hat{\Gamma}_S \end{bmatrix} \begin{bmatrix} \Delta Y_i \\ \Delta S_i^* \end{bmatrix}$$

(10)

where $I_2$ is a $(2 \times 2)$ identity matrix, and variables with the hat-like circumflex indicate estimates. $\Delta \hat{S}_i$ is our model’s prediction of the change in equity that will be compared to the actual dividend-smoothing component of the change in equity, $\Delta S_i^*$, in equation (6). Therefore, to evaluate the performance of the VAR model of equation (9), we need to obtain the estimated coefficient matrix $\hat{\Phi}$, then plug it into equation (10) to calculate $[\hat{\Gamma}_Y \hat{\Gamma}_S]$ and test the hypothesis $[\Gamma_Y \ \Gamma_S] = [0 \ 1]$. If we accept the hypothesis that our model’s prediction of the change in equity is equal to the actual optimal change in equity, $\Delta S_i^*$, then we can conclude that the firm has an optimal dividend policy over time. Moreover, both the theory-based model
prediction of smoothed dividend and the actual smoothed dividend can then be computed. Finally, we may compare the standard deviations of predicted and actual (dividend-smoothing) changes in stockholders’ equity and smoothed dividend as well to quantify the flexibility of equity movements and dividend policy, respectively.

**Empirical Evidence**

Quarterly data from 1992Q1 to 2002Q2 for both HP and Citi Bank were used to justify the theoretical model’s prediction of the changes in equity to the actual series. All data are from COMPUTSTAT and are expressed in per capita terms, *i.e.*, per share. The sample periods chosen are based on the availability of information in COMPUTSTAT as interpreted in section 1. An averaged annual market interest rate of 5%, *i.e.*, 1.25% quarterly, was calculated using US discount rates over the investigation period from the International Financial Statistics (IFS). Further details on the construction of all series are specified whenever needed in the following discussion.

The first step in the analysis is to verify if \( rS_t + Y_t \) and \( D_t \) are nonstationary, *i.e.*, \( I(1) \), and cointegrated. Table I shows unit-root test statistics for several series which are used in the estimation procedure. Next, we have to obtain an estimate of \( \theta \) in equation (6) in order to construct the (stationary) dividend-smoothing component
of the change in stockholders’ equity by removing the non-stationary component of the actual series associated with dividend tilting.

From Table I, \( rS_t + Y_t \) and the \( D_t \) series both exhibit stationarity at the 5% and 1% significance levels for HP, respectively; nevertheless, these two series are nonstationary for Citi Bank at the 5% significance level. The OLS (ordinary least squares) regression can then give an efficient estimate of the tilting effect for HP; on the other hand, the cointegrated relationship between \( rS_t + Y_t \) and \( D_t \) for Citi Bank, therefore, was estimated using Phillips and Hansen’s (1990) FM (fully modified) correction method; the results are tabulated in Table II. The advantage of the FM procedure (over use of OLS) is that hypothesis tests based on the FM regression are asymptotically normal when variables encountered nonstationarity. Thus we can formally test if \( \theta = 1 \) to justify the existence of dividend-tilting effects.

In Table II, the OLS estimate of the dividend-tilting effect for HP, \( \hat{\theta}_{OLS} = 0.8048 \), is not significantly different from unity; however, the simple \( t \) test of \( \theta = 1 \) is rejected at the 1% significance level for Citi Bank (\( \hat{\theta}_{FM} = 4.1146 \)), which means that there exists no dividend tilting effect for HP, but Citi Bank’s dividend distribution may tilt towards the future. The LC test statistic of Hansen (1992) for the null hypothesis of cointegration between \( rS_t + Y_t \) and \( D_t \) for Citi Bank is 0.0849, which confirms that cointegration does exist in Citi Bank. Moreover, Hansen’s stability tests of the constant parameter in Citi Bank’s FM cointegrating regression,
show the consistent inference with LC test; the mean-F and sup-F tests for the null of stable $\theta$ enforce that the cointegrated relationship between $rS_t + Y_t$ and $D_t$ is fairly stable for Citi Bank. In other words, HP has distributed adequate cash dividends of its permanent stream of net earnings to its shareholders. On the contrary, Citi Bank’s dividend policy has been too conservative over time, in that it has distributed smaller cash dividends of its permanent flow of net earnings to its shareholders than it should have due to its perception of bleak prospects. These events may have resulted from the historical outlook of HP which has been confident about the burst of e-business and hi-tech industries since 1990s, while, at the same time, Citi Bank was on the brink of collapse in 1991 due to the fall of the real estate market and to loans to troubled Latin American nations which soured.

The estimation results shown above suggest a spurious long run relationship between $rS_t + Y_t$ and $D_t$ for HP, therefore, its actual dividend-smoothing component of the changes in equity is just the difference between them. However, Citi Bank’s results enable us to take the residuals from the FM cointegrating regression of $rS_t + Y_t$ on $D_t$ as the actual dividend-smoothing component of the changes in equity. Having the actual dividend-smoothing component of the changes in equity and the net earning stream series for both firms in hand, we are able to run the VAR estimation described in the previous section, except that the VAR requires that stationary variables be used. Fortunately, as reported in Table I, both the differences in intertemporal net earnings, $\Delta Y$, and actual changes in equity, $\Delta S^*$, are stationary at least at the 5% significance levels for HP and Citi Bank, respectively; therefore, they are qualified for use in general VAR estimations.
Another problem in performing VAR is the choice of an optimal lag order; we referred to Akaike’s Information Criterion (AIC), and the Schwarz Bayesian Criterion (SBC) to select the optimal lag length for the VAR estimation. Table III indicates quite consistent selection of the VAR(1) model using AIC and SBC for both HP and Citi Bank. We therefore report the first-order VAR statistical results and draw conclusions based on them.

A summary of VAR results is given in Table IV. The respective standard error coefficients of the VAR estimations are reported in parentheses. First, the signaling effect discussed in the previous section requires a negative $b$ in the $\Phi$ matrix (refer to equation (9)). The VAR estimates of $b$ are -0.2443 and -0.8882 for HP and Citi Bank, respectively, which are negatively significant at the 1% significance level, and both are consistent with our model’s prediction of the signaling effect. That is, the null hypothesis that $\Delta S^+$ negatively Granger-causes $\Delta Y$ cannot be rejected, which implies that if a firm makes a conservative cash dividend per share this time due to pessimistic expectation about its future, it will retain more earnings and result in a higher level of equity now, i.e., $\Delta S^+ > 0$. The market, then, will take it as a warning “signal” that this firm may profit less from its business in the future, and vice versa. This in turn generates the perfect fit of expected changes in equity to the actual series (Fig. 1) and close predictions of expected smoothed dividends to actual ones (Fig. 2) for both HP and Citi Bank. Moreover, the formal tests by the Wald statistics$^8$ reported in Table V for the restriction of $[\Gamma_y \Gamma_s]=[0 \ 1]$ can not be rejected at the 5% significance level for both HP and Citi Bank which conclude the outstanding performance of our theoretical modeling and empirical studies.

Finally, in Table V, the estimated coefficient matrix, $[\hat{\Gamma}_y \hat{\Gamma}_s]$, for both HP and
Citi Bank are quite close to \([0 \ 1]\). This has been confirmed by the Wald statistics for the formal test of the coefficient restrictions implied by the present-value relationship, \textit{i.e.}, equation (10), which says that both firms’ VAR estimation tracks the dynamic behavior of the changes in equity precisely. The correlation between the actual and predicted series of changes in equity reaches 0.9999 and 0.9991 for HP and Citi Bank, respectively. This is quite obvious from Fig. 1. We then computed the ratio of the variance of \(\Delta S^*\) to that of \(\Delta \hat{S}\). The actual changes in equity for HP were about 112\% of volatility relative to the series generated from the theoretical model estimation. In other words, this says that the actual changes in stockholders’ equity might have fluctuated more than they were supposed to during the time period being investigated. Nonetheless, the volatility almost perfectly fitted the case of Citi Bank. Finally, correlations between the actual and predicted smoothed dividends read 0.9748 and 0.9996 for HP and Citi Bank, respectively. This can also be confirmed by the perfect fitness shown in Fig. 2. The ratios of the variance of actual smoothed dividends to that of model’s prediction are 0.9918 for HP and 1.0033 for Citi Bank, which implies that both firms have implemented optimally flexible policies regarding smoothed dividends over time.

**Conclusions**

The main purpose of this paper was to develop a theory of dividend smoothing by utilizing an intertemporal approach and then to justify the theoretical prediction using HP and Citi Bank’s quarterly data from 1992Q1 to 2002Q2.

In section 2, we provided a theory of dividend smoothing based on the microfoundation of a representative agent model. The model then produced quite
plausible results specifying the relation between the optimal dividends and permanent net earnings, and is an important application of signaling theory. In order to judge if the theoretical implications of the theory of dividend smoothing are robust, sections 3 gave a full discussion of econometric procedures for testing the theoretical findings. In section 4, the empirical estimation was implemented using HP and Citi Bank’s quarterly data in a first-order VAR.

From the empirical findings for both HP and Citi Bank, we showed that the theory is fully practical. Our theory-based VAR estimation of dividend smoothing provided a statistically confident prediction of the dynamic behavior of the actual changes in stockholders’ equity and the smoothing component of cash dividends.
Table I: Phillips-Perron Tests for Unit Roots

<table>
<thead>
<tr>
<th></th>
<th>HP</th>
<th>Citi Bank</th>
</tr>
</thead>
<tbody>
<tr>
<td>$rS_t + Y_t$</td>
<td>-2.1288*</td>
<td>-0.9232</td>
</tr>
<tr>
<td>$D_t$</td>
<td>-3.1976**</td>
<td>0.9087</td>
</tr>
<tr>
<td>$Y_t$</td>
<td>-1.9824*</td>
<td>-1.6109</td>
</tr>
<tr>
<td>$\Delta Y_t \equiv Y_t - Y_{t-1}$</td>
<td>-5.8512**</td>
<td>-9.9058**</td>
</tr>
<tr>
<td>$\Delta S_t^*$</td>
<td>-2.0421*</td>
<td>-4.9100**</td>
</tr>
</tbody>
</table>

Note: The second and third columns report the unit root tests for HP and Citi Bank, respectively. $S_t$ is the equity; $r$ is the market risk-free interest rate; $Y_t$ is the firm’s net earnings; $D_t$ is the cash dividend. The Newey-West automatic truncation lag selection for Phillips-Perron (PP) equals \(\text{Integer}\{4(T/100)^{2/9}\}=3\), where \(\text{Integer}\{\}\) is the Gauss function, and \(T\) is the number of observation. Neither constant nor time trend is included in the PP testing regression. ‘*’ and ‘**’ means test statistic significant at the 5%, and 1% significance levels, respectively. MacKinnon critical values for rejection of hypothesis of a unit root at the 5%, and 1% significance levels are -1.95, and -2.63, respectively. This table indicates that $rS_t + Y_t$, $D_t$ and $Y_t$ exhibit stationary for HP, then OLS regression is used in generating Table II. On the other hand, $rS_t + Y_t$, $D_t$ and $Y_t$ exhibit nonstationary for Citi Bank, therefore, the FM (fully modified) correction method is used in producing Table II.
Table II: Estimation for Cash-Dividend tilting Effects (θ)

<table>
<thead>
<tr>
<th></th>
<th>HP (OLS)</th>
<th>Citi Bank (FM OLS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ (t value for H0: θ=1)</td>
<td>0.8048(-0.3225)</td>
<td>4.1146**(7.8951)</td>
</tr>
<tr>
<td>LC test</td>
<td>N/A</td>
<td>0.0849</td>
</tr>
<tr>
<td>Mean-F</td>
<td>N/A</td>
<td>0.3845</td>
</tr>
<tr>
<td>Sup-F</td>
<td>N/A</td>
<td>1.9486</td>
</tr>
</tbody>
</table>

Note: ‘*’ and ‘**’ means test statistic significant at the 5%, and 1% significance levels, respectively. The second row is the test result for the null of θ = 1. θ is the dividend tilting effect and is estimated from following equation:

\[ rS_t + Y_t = \theta D_t + \varepsilon_t \]

The OLS estimate of θ can’t significantly reject θ = 1 for HP (\( \hat{\theta}_{OLS} = 0.8048 \)). However, the null of θ = 1 is rejected by the FM estimation at the 1% significance level for Citi Bank (\( \hat{\theta}_{FM} = 4.1146 \)). Therefore, there exists no dividend tilting effect for HP, but Citi Bank’s dividend distribution may tilt towards the future. The third to fifth rows are the cointegration test results between \( rS_t + Y_t \) and \( D_t \) for Citi Bank. LC test is based on Hansen (1992) for H0: a cointegrated relationship (5% and 1% critical values are 0.575 and 0.898, respectively). Mean-F and Sup-F are based on Hansen (1992) as well for H0: cointegrating vector is stable (5% and 1% critical values for Mean-F are 4.57, 6.78, and Sup-F are 12.4, 16.2, respectively). Neither constant nor time trend is included in FM estimation. The LC test statistic is 0.0849, which means the cointegration does exist between \( rS_t + Y_t \) and \( D_t \) for Citi Bank. The mean-F and sup-F tests confirm that this cointegrating relationship is stable.
Table III: Determination of Optimal Lag Length in VAR Estimation

<table>
<thead>
<tr>
<th>Lag</th>
<th>HP AIC</th>
<th>HP SBC</th>
<th>Citi Bank AIC</th>
<th>Citi Bank SBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14.7606†</td>
<td>15.0139†</td>
<td>11.8031†</td>
<td>12.0564†</td>
</tr>
<tr>
<td>2</td>
<td>15.2194</td>
<td>15.6459</td>
<td>12.2637</td>
<td>12.6903</td>
</tr>
<tr>
<td>3</td>
<td>15.4879</td>
<td>16.0912</td>
<td>12.5759</td>
<td>13.1792</td>
</tr>
<tr>
<td>4</td>
<td>15.7830</td>
<td>16.5667</td>
<td>13.3530</td>
<td>14.1367</td>
</tr>
<tr>
<td>5</td>
<td>16.6697</td>
<td>17.6374</td>
<td>14.0064</td>
<td>14.9741</td>
</tr>
<tr>
<td>6</td>
<td>16.9932</td>
<td>18.1486</td>
<td>14.5789</td>
<td>15.7343</td>
</tr>
<tr>
<td>7</td>
<td>18.3294</td>
<td>19.6762</td>
<td>15.0634</td>
<td>16.4102</td>
</tr>
<tr>
<td>8</td>
<td>19.0113</td>
<td>20.5532</td>
<td>15.7781</td>
<td>17.3200</td>
</tr>
</tbody>
</table>

Note: †: indicates the minimum value of AIC or SBC amongst all lag lengths. We referred to Akaike’s Information Criterion (AIC), and the Schwarz Bayesian Criterion (SBC) to select the optimal lag length for the VAR estimation. This table indicates quite consistent selection of the VAR(1) model using AIC and SBC for both HP and Citi Bank.
Table IV: VAR Results for HP and Citi Bank

<table>
<thead>
<tr>
<th></th>
<th>HP</th>
<th>Citi Bank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta Y_{t-1}$</td>
<td>$\Delta S^*_{t-1}$</td>
</tr>
<tr>
<td>$\Delta Y_t$</td>
<td>0.2153</td>
<td>-0.2443***</td>
</tr>
<tr>
<td></td>
<td>(0.1580)</td>
<td>(0.1037)</td>
</tr>
<tr>
<td>$\Delta S^*_t$</td>
<td>0.2430</td>
<td>0.7569***</td>
</tr>
<tr>
<td></td>
<td>(0.1580)</td>
<td>(0.1036)</td>
</tr>
</tbody>
</table>

Note: $\Delta S^*_t$ is the change of equity in theory; $\Delta Y_t$ is the change of firm’s net earnings. VAR(1) model is used for both HP and Citi Bank,

$$\begin{bmatrix} \Delta Y_t \\ \Delta S^*_t \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \Delta Y_{t-1} \\ \Delta S^*_{t-1} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}. $$

Standard errors are in parentheses, (), under estimates. ‘*’, ‘**’ and ‘***’ mean that test statistics are significant at the 10%, 5%, and 1% significance levels, respectively. Neither constant nor time trend is included in VAR estimation. The VAR estimates of $b$ are -0.2443 and -0.8882 for HP and Citi Bank, respectively, which are both negatively significant at the 1% significance level. Then the null hypothesis that $\Delta S^*$ negatively Granger-causes $\Delta Y$ cannot be rejected for both firms, which is consistent with our model prediction of the signaling effect. The market, then, will take today’s $\Delta S^*$ as a warning “signal” of the firm’s future profit prospect ion, i.e., if $\Delta S^*_t > (<)0$, firms may profit less (more) in the future.
Table V: Other Statistics for the Present-Value Models

<table>
<thead>
<tr>
<th></th>
<th>HP</th>
<th>Citi Bank</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{F}_y ) \begin{bmatrix} \hat{F}_s \end{bmatrix} )</td>
<td>\begin{bmatrix} 0.0164 &amp; 0.9401 \end{bmatrix}</td>
<td>\begin{bmatrix} 0.0408 &amp; 0.9348 \end{bmatrix}</td>
</tr>
<tr>
<td>Wald [d.f.]</td>
<td>2.3667[2]</td>
<td>0.6732[2]</td>
</tr>
<tr>
<td>Correlation(( \Delta S^\prime_y, \Delta \hat{S}_y ))</td>
<td>0.9999</td>
<td>0.9991</td>
</tr>
<tr>
<td>Correlation(( \theta D^\prime_y, \theta \hat{D}_y ))</td>
<td>0.9748</td>
<td>0.9996</td>
</tr>
<tr>
<td>( \frac{Var(\Delta S^\prime)}{\frac{Var(\Delta \hat{S})}} )</td>
<td>1.1228</td>
<td>1.0400</td>
</tr>
<tr>
<td>( \frac{Var(\theta D^\prime)}{\frac{Var(\theta \hat{D})}} )</td>
<td>0.9918</td>
<td>1.0033</td>
</tr>
</tbody>
</table>

Note: \( \begin{bmatrix} \hat{F}_y \ \hat{F}_s \end{bmatrix} \) is the coefficient matrix in equation (10),

\[
\Delta \hat{S}_y = \hat{F}_y \Delta Y_y + \hat{F}_s \Delta S^\prime_y.
\]

The null hypothesis for the formal test of the theory is \( \begin{bmatrix} F_y & F_s \end{bmatrix} = \begin{bmatrix} 0 & 1 \end{bmatrix} \). Wald test statistics for the null are reported in the third row, which show the null can not be rejected at the 5% significance level for both HP and Citi Bank, where \( \chi^2 \) degrees of freedom are in [ ]. Therefore, we conclude that our theoretical modeling and empirical studies perform quite successfully. Correlations between actual and predicted series of changes in equity are shown in the fourth row, and reach 0.9999 and 0.9991 for HP and Citi Bank, respectively. The ratios of the variance of \( \Delta S^\prime \) to that of \( \Delta \hat{S} \) are reported in the sixth row and indicate the
actual changes in equity for HP were about 112% of volatility relative to the series generated from the theoretical model estimation. This says that the actual changes in stockholders’ equity might have fluctuated more than they were supposed to during the time period being investigated. Nonetheless, the volatility almost perfectly fitted the case of Citi Bank. Finally, correlations between the actual and predicted smoothed dividends are shown in the fifth row and read 0.9748 and 0.9996 for HP and Citi Bank, respectively. The ratios of the variance of actual and predicted smoothed dividends are shown in the seventh row, which are 0.9918 for HP and 1.0033 for Citi Bank. It says that both firms have implemented optimally flexible policies regarding smoothed dividends over time.
Figure 1: HP and Citi Bank’s Actual and Predicted Changes of Equity (per share)

Note: This in turn follows the empirical studies interpreted above and generates the perfect fit of expected changes in equity to the actual series for both HP and Citi Bank. We showed that the theory is fully practical.
Figure 2: HP and Citi Bank’s Actual and Predicted Smoothed-Dividend (per share)

Note: This also follows from the empirical performance stated above, and therefore generates the close predictions of expected smoothed dividends to actual ones for both HP and Citi Bank. We concluded that the theory is fully practical.
Mathematical Appendix

Derivation of equation (7):

From the definition of \( Y_t^p \) that

\[
Y_t^p = \frac{r}{1+r} E \left( \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} Y_{t+i} \right) = \frac{r}{1+r} E \left( \sum_{i=0}^{\infty} \frac{F}{1+r} \right) Y_t
\]

\[
= \frac{r}{1+r} \left( \sum_{i=0}^{\infty} \frac{F}{1+r} \right) Y_t = \frac{r}{1+r} \cdot \frac{1}{1-F} Y_t = \frac{r}{1+r-F} Y_t,
\]

where \( F \) denotes the forward operator, \( F Y_t = Y_{t+1} \), \( i = 0, 1, 2, \ldots \). Equation (6) can then be rewritten as

\[
\Delta S_t^* = r S_t + Y_t - \theta D_t^* = Y_t - Y_t^p = Y_t - \frac{r}{1+r-F} Y_t
\]

\[
= \frac{1-F}{1+r-F} \Delta Y_t = -\frac{1}{r} \left( \frac{r}{1+r-F} \Delta Y_{t+1} \right) = \frac{1}{r} \left( \frac{r}{1+r} - \frac{1}{1+r-F} \right) E \left( \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} \Delta Y_{t+i} \right)
\]

\[
= \frac{1}{1+r} E \left( \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} \Delta Y_{t+i} \right) = -E \left( \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} \Delta Y_{t+i} \right). \tag{7}
\]

Derivation of equation (10):

From equation (7) and the fact that \( E_i \left( \Delta Y_{t+i} \right) = \begin{bmatrix} 1 & 0 \end{bmatrix} \Phi^i Z_i \), we then have,

\[
\Delta \hat{S}_t = -\left( \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} E_i(\Delta \hat{Y}_{t+i}) \right) = -\left( \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} \begin{bmatrix} 1 & 0 \end{bmatrix} \hat{\Phi}^i Z_i \right)
\]

\[
= \begin{bmatrix} 1 & 0 \end{bmatrix} \left( \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} \hat{\Phi}^i \right) Z_i = \begin{bmatrix} 1 & 0 \end{bmatrix} \left( \begin{bmatrix} 1 & 0 \end{bmatrix} \hat{\Phi} \right) \begin{bmatrix} I_2 - \frac{1}{1+r} \hat{\Phi} \end{bmatrix}^{-1} \begin{bmatrix} \Delta Y_t \end{bmatrix}
\]

\[
= \hat{F}_Y \begin{bmatrix} \Delta Y_t \end{bmatrix}, \text{ where } \begin{bmatrix} \hat{F}_Y & \hat{F}_S \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} I_2 - \frac{1}{1+r} \hat{\Phi} \end{bmatrix}^{-1}. \tag{10}
\]
References


Footnotes:

* Associate Professor, Department of Applied Economics, National University of Kaohsiung, Taiwan. Correspondence address: 700 Kaohsiung University Rd., Nan-Tzu Dist., Kaohsiung 811, Taiwan. Phone: 886-7-5919186, Fax: 886-7-5919319, E-mail: mjweng@nuk.edu.tw.
† Assistant Professor, Department of Banking and Finance, National Chi Nan University, Taiwan. E-mail: mstsai@ncnu.edu.tw
‡ Professor, Department of Economics, National Chung Cheng University, Taiwan. E-mail: ecdjlw@ccunix.ccu.edu.tw

1 The ability of any business to roll over its debt perpetually brings to mind the notorious Boston financier Charles Ponzi who used to pay exorbitant interest to lenders out of an ever-expanding pool of deposits, without ever “investing” one penny. Mr. Ponzi was indicted in Federal Court in November 1920, and his bank eventually collapsed. The stockholder’s refusal to finance a Ponzi game means that he will demand that the net present value of his equity position to be zero, i.e., full repayment of his investment. O’Connell and Zeldes (1988) offer a theoretical study of “Ponzi games.”

2 As we maximize equation (2) subject to equation (1), Bellman’s formula gives us Euler’s equation. Then substitution of the CRRA form of the utility function results in equation (4).

3 Please refer to the mathematical appendix.

4 We define $\Delta S^*$ as the optimal change of equity under dividend smoothing in equation (6). In association with the empirical studies, we view the actual change in equity (dividend-smoothing component) as this optimal change.

5 Please refer to the mathematical appendix.

6 If the constant term and time trend are added to our empirical studies of VAR estimation, the above formula has to be modified as:
\[ \Delta \hat{S}_t = -[I \ 0] \left[ \begin{bmatrix} \frac{1}{1+r} \Phi & 0 \\ \frac{1}{1+r} \Phi \end{bmatrix} \right]^{-1} \left( I_2 - \left[ I_2 - \Phi \right] \left( \hat{A} + \hat{B} \cdot t \right) \right) + \frac{1}{r} \left[ I_2 - \Phi \right]^{-1} \left( \hat{A} + \hat{B} \cdot t \right) + \left( \frac{1+2r}{r^2} I_2 - \frac{1}{r} I_2 - \Phi \right) + \left[ \frac{1}{1+r} \Phi \right] \left( I_2 - \left[ I_2 - \Phi \right] \right) \left( I_2 - \Phi \right)^{-1} \left( I_2 - \Phi \right)^{-1} \right] \cdot \hat{B} \]

, where \( \hat{A} \) and \( \hat{B} \) are estimated VAR coefficient matrices of constant term and time trend, respectively.

\(^7\) As stated in Otto, it follows from equation (5) that in the steady state, \( \Delta S = -\Delta Y \hat{B} \).

In our analysis, if we substitute \( \Delta S \) and \( \Delta Y \) with the means of \( \Delta S \) and \( \Delta Y \) series, respectively, the market interest rate implied by the model’s steady-state restriction is implausibly large and even negative for HP. Therefore, we follow Otto’s methodology to replace \( \Delta S \) and \( \Delta Y \) by their deviations from respective means in empirical studies.

\(^8\) Please refer to Campbell and Shiller (1987) for more details in implementing this test.