Working Paper No. 2016006

北京大學
汇丰商学院
Peking University HSBC Business School
商界军校

PHBS

Sticky Dividends: A New Explanation

Hyun Joong Im

Ya Kang

Chang Yong Ha

Sticky dividends: A new explanation*

Hyun Joong Im[†]
HSBC Business School
Peking University

Ya Kang[‡]
NUS Business School
National University of Singapore

Chang Yong Ha[§]
HSBC Business School
Peking University

December 15, 2016

Abstract

Past studies show that firms adjust dividends very slowly to their dividend targets. This paper reinvestigates the dynamics of corporate dividend policy using a generalized partial adjustment model. A novel feature of our model is that it allows managers to consider the earnings *history* via adaptive expectation formation of future earnings for their payout decisions. Thus, our model captures the spirit of the finding in recent literature that managers target to stay consistent with the historical dividend policy while considering future earnings prospects. We show that firms adjust dividends to their target payouts much faster than previously documented. This study also shows that their target dividends are predominantly driven by firm-specific effects, and tend to become significantly more stable when managers form future earnings prospects adaptively. Thus, dividend-smoothing behavior could arise from their attempts to conform to the target payouts, thereby leading to higher dividend adjustment speeds.

JEL classification: D9, G30, G35

Keywords: Payout Policy, Speed of Adjustment, Dividend Dynamics, Adaptive Expectations

*Corresponding Author: Chang Yong Ha

†Address: HSBC Business School, Peking University, University Town, Nanshan District, Shenzhen, 518055, China; Tel: +86 (0)755 2603 3627; Fax: +86 (0)755 2603 5344; Email: hyun.im@phbs.pku.edu.cn

‡Address: Department of Finance, NUS Business School, National University of Singapore, BIZ 2 Building #B1-3, 1 Business Link, 117592, Singapore; Tel: +65 8647 3952; Email: kangya@u.nus.edu

§Address: HSBC Business School, Peking University, University Town, Nanshan District, Shenzhen, 518055, China; Tel: +86 (0)755 2603 3651; Fax: +86 (0)755 2603 5344; Email: cyha@phbs.pku.edu.cn

1 Introduction

Since Lintner (1956) studied corporate dividend policy and practice using a partial adjustment model, extensive prior research has documented a series of empirical findings and their plausible explanations.[1] Yet dividends remain one of the most contested and thorniest puzzles in corporate finance (Allen et al., 2000). Research in more recent years, in particular, provides evidence that many of those empirical findings and underlying theories are to be revised or refuted. Among others, Brav et al. (2005), using survey and field interviews with financial executives, provide a new perspective on various aspects of corporate payout policy such as managers' beliefs and stances concerning dividend policy and its determinants. Of particular interest for this paper is their finding that more than four-fifths of executives target to remain consistent with historical dividend policy and take lagged dividends as a benchmark when choosing the current dividend policy. Also, the majority of firms are known to tie their dividends to the sustainable future earnings. While these managerial tendencies are in line with dividend conservatism, they also offer some clues on how firms and managers are likely to set the dividend targets.

Building on the documented managerial attention to past dividend history and future earnings prospects in setting today's dividend policy, this study aims to offer a novel insight into the mechanism through which firms' actual dividends remain sticky.[2] To that end, we propose a generalized partial adjustment model with adaptive expectations for future earnings.[3] In our proposed model, the managerial attention to past dividends is reflected in the way managers form the future earnings prospects which has also been documented to be an important consideration for dividend payout decision. Hence, our model does capture the spirit of managers' tendency to consider both past dividends and future earnings in setting the dividend targets. Our model is also consistent with managers' motive to maintain smooth dividends because of asymmetric response of the market to dividend increases and cuts. By allowing managers to set target dividends based on expected future earnings,[4] our model can generate a smoother path of target dividends provided that

[1] See Allen and Michaely (1995) and DeAngelo et al. (2008) for excellent reviews of the related literature.

[2] In Internet Appendix A.1, we present the analysis of the time-series evolution of dividends for our cross-section of the firms following Lemmon et al. (2008). A preliminary examination reveals that the presence of a permanent or long-run component that leads to highly persistent cross-sectional differences in dividend ratios, as well as a transitory or short-run component that leads to a gradual convergence in dividend ratios. In addition, both nonparametric and parametric (ANCOVA) analyses of variance decomposition show that the time-invariant firm-specific components are the major source of total variation in dividends. That is, the majority of the total variation in dividends comes from cross-sectional differences as opposed to time-series variation. See Internet Appendix A.1 for further details.

[3] Chow (2011) provides a statistical reason and strong econometric evidence for supporting the adaptive expectations hypothesis in economics.

[4] Setting dividend targets in this manner is in line with the signaling hypothesis of dividends (Bhattacharya, 1979; Miller and

managers form expectations adaptively when assessing future earnings prospects. Note that with adaptive expectation formation, future earnings prospects are formed as a weighted average of past earnings with geometrically declining weights.

Among the reported empirical results lie the slow adjustments of dividends toward target payouts. For example, Fama and Babiak (1968) and Fama and French (2002) report quite low adjustment speeds that are 0.37 and 0.33, respectively. Given the volatility in firms' earnings, it has been a puzzle that actual dividends paid out do not reflect that volatility. Our model allows us to reexamine the adjustment speed of dividends to payout targets by explicitly modeling the dividend target formation process. Existing research often attributes smooth dividends to firms' reluctance to change dividends due to asymmetric information (i.e., signaling effect (Bhattacharya, 1979)) or agency conflicts (e.g., irrelevance of short-term profits to dividend decision (Easterbrook, 1984)).[5] One important implication of those theories is that the manager's information set for dividend decision is likely to contain a longer series of past dividends as well as future earnings prospects. Incorporating this aspect of firms' dividend decisions, this study provides an alternative and richer explanation for this long-lived puzzle by showing that firms' target dividend payouts themselves are much "smoother" than previously documented. While volatile target payouts in conventional models result in fairly low speeds of adjustment, our estimation results suggest that firms tend to adjust their dividend payouts to the targets much faster.

2 Data and methodology

This study uses annual accounting data from the CRSP/Compustat Merged Database (CCM) for the years 1988–2014. Firms with standard industrial classification (SIC) codes between 6000 and 6999, between 4900 and 4999, or between 9000 and 9999 are excluded as these firms focus on financial services, are regulated utilities, or are government entities. We require that each firm have at least 12 years of observations and there be no gaps in the middle of the sample period. We drop observations if the dividend to total assets ratio (denoted $D_{i,t}$), earnings to total assets ratio (denoted $E_{i,t}$), or a proxy for Tobin's Q as measured by the sum of the book value of debt and market value of equity divided by the book value of total assets (denoted $Q_{i,t}$) are missing. All variables are winsorized at the 1st and 99th percentiles to minimize the effects of outliers.

Rock, 1985; John and Williams, 1985).

[5]See Leary and Michaely (2011) for a comprehensive survey of the theoretical models.

There are a total of 28,063 firm-year observations, corresponding to 1,383 firms. Industry dummies are constructed according to Fama and French's (1997) 48 industry classification.

Waud (1966) shows that a conventional partial adjustment model and an adaptive expectations model yield indistinguishable empirical specifications as far as estimation is concerned. Hence, one cannot tell whether the estimated coefficient of the lagged dividend ratio is driven by the speed of dividend adjustment (γ) or the speed of expectations revision (ρ). See Internet Appendix A.2 for a detailed discussion of the identification problem. Our model presented in this section has a novel feature in that it includes the ingredients of both the partial adjustment model and the adaptive expectations model. This feature allows us to sort out the respective effects of dividend adjustment speed (γ) and expectations revision speed (ρ) in the dynamics of corporate dividend policy. In addition, our model takes into account the unobserved firm heterogeneity in setting dividend targets.

A generalized partial adjustment model of dividends with an adaptive expectations formation process in the panel data setting consists of the following three equations:

$$D_{i,t} - D_{i,t-1} = \gamma(D^{\star}_{i,t} - D_{i,t-1}) + \pi_j + \kappa_t + \nu_{i,t}; \quad (1)$$

$$D^{\star}_{i,t} = \alpha E^{e}_{i,t} + \beta Q_{i,t-1} + \mu_i; \quad (2)$$

$$E^{e}_{i,t} - E^{e}_{i,t-1} = \rho(E_{i,t-1} - E^{e}_{i,t-1}), \quad (3)$$

where $D_{i,t}$ and $D^{\star}_{i,t}$ denote the actual and target dividend ratios of firm i in year t. Equations (1) and (2) describe a partial adjustment model similar to the conventional partial adjustment model. Equation (2) describes how the target dividend is determined. We modify the target dividend equation in the conventional partial adjustment model so that the target dividend ratio is determined by unobservable, but estimable, expected earnings ($E^{e}_{i,t}$) rather than previous-period earnings ($E_{i,t-1}$). To make γ and ρ separately identifiable, we also include Tobin's Q measured at the beginning of the year ($Q_{i,t-1}$) as an additional observable determinant of the target payout ratio.[6] In the first two equations, assumptions on error components are the same as those in the conventional partial adjustment model, as described in Internet Appendix A.2. Equation (3) describes the adaptive expectations formation process, as described in Internet Appendix A.2.

[6]If the second equation is specified as $D^{\star}_{i,t} = \alpha E^{e}_{i,t} + \mu_i$, one can obtain the following reduced-form regression model:

$$D_{i,t} = [(1-\gamma) + (1-\rho)]D_{i,t-1} - (1-\gamma)(1-\rho)D_{i,t-2} + \gamma\rho\alpha E_{i,t-1} + \eta_i + \xi_{i,t},$$

where $\eta_i = \gamma\rho\mu_i$ and $\xi_{i,t} = [(\kappa_t + \nu_{i,t}) - (1-\rho)(\kappa_{t-1} + \nu_{i,t-1})]$. However, one cannot identify γ and ρ separately by estimating this reduced-form model.

Substituting Equation (2) into Equation (1), substituting $\rho E_{i,t-1} + (1-\rho)E^{e}_{i,t-1}$ for $E^{e}_{i,t}$, substituting $\frac{1}{\gamma\alpha}[D_{i,t-1} - (1-\gamma)D_{i,t-2} - \gamma\beta Q_{i,t-1} - \gamma\mu_i - \pi_j - \kappa_t - \nu_{i,t-1}]$ for $E^{e}_{i,t-1}$, and rearranging the equation gives the following reduced form:

$$\begin{aligned} D_{i,t} &= [(1-\gamma)+(1-\rho)]D_{i,t-1} + \gamma\rho\alpha E_{i,t-1} - (1-\gamma)(1-\rho)D_{i,t-2} + \gamma\beta Q_{i,t-1} - \gamma(1-\rho)\beta Q_{i,t-2} \\ &\quad + \gamma\rho\mu_i + \rho\pi_j + [(\kappa_t + \nu_{i,t}) - (1-\rho)(\kappa_{t-1} + \nu_{i,t-1})]. \end{aligned} \tag{4}$$

This can be rewritten as the following standard dynamic panel regression model:

$$D_{i,t} = \delta_1 D_{i,t-1} + \delta_2 E_{i,t-1} + \delta_3 D_{i,t-2} + \delta_4 Q_{i,t-1} + \delta_5 Q_{i,t-2} + \eta_i + \xi_{i,t}, \tag{5}$$

where $\delta_1 = (1-\gamma)+(1-\rho)$, $\delta_2 = \gamma\rho\alpha$, $\delta_3 = -(1-\gamma)(1-\rho)$, $\delta_4 = \gamma\beta$, $\delta_5 = -\gamma(1-\rho)\beta$, $\eta_i = \gamma\rho\mu_i$, and $\xi_{i,t} = \rho\pi_j + [(\kappa_t + \nu_{i,t}) - (1-\rho)(\kappa_{t-1} + \nu_{i,t-1})]$. The error term $\xi_{i,t}$ is an MA(1) process if each of κ_t and $\nu_{i,t}$ is assumed to be white noise.[7] A consistent estimator can be obtained using System GMM suggested by Arellano and Bover (1995) and Blundell and Bond (1998). The delta method is employed in order to estimate structural parameters (γ, ρ, α, β) as nonlinear combinations of regression coefficients.[8]

3 Results

Before we present our main results, we first estimate the conventional partial adjustment models that can be viewed as a special case of our generalized model in the sense that the speed of expectations revision (ρ) is set to 1. Thus, managers in this model form future earnings prospects based only on current earnings. Table 1 reports estimation results based on three different estimation methods, i.e., Pooled OLS, Within

[7]This does not imply that the actual residuals always follow the process implied by the specification. However, in both Difference GMM and System GMM, a different error structure would result in a different set of valid instruments as suggested by the Sargan-Hansen test of overidentifying restrictions. A less restrictive assumption such as MA(1), compared with the case of MA(0), allows for a smaller number of valid instruments.

[8]We use the following nonlinear combinations of coefficients to obtain the structural parameters. First, dividing $-\delta_5$ by δ_4 gives an estimate of $1-\rho$:

$$\frac{-\delta_5}{\delta_4} = \frac{\gamma(1-\rho)\beta}{\gamma\beta} = (1-\rho),$$

and therefore $\rho = 1 + \frac{\delta_5}{\delta_4}$. Second, we can get $(1-\gamma)$ using the equation for δ_1:

$$(1-\gamma) = \delta_1 - (1-\rho) = \delta_1 + \frac{\delta_5}{\delta_4},$$

and therefore $\gamma = 1 - \delta_1 - \frac{\delta_5}{\delta_4}$. Finally, $\alpha = \frac{\delta_2}{\gamma\rho}$ and $\beta = \frac{\delta_4}{\gamma}$.

Groups, and System GMM estimators. Regardless of estimation methods, the parameter estimates, $\widehat{\alpha}$ and $\widehat{\beta}$, for target payout determinants are significantly positive at the 1% significance level. The estimated speed of adjustment ($\widehat{\gamma}$) is comparable to the estimates reported in previous studies.[9] The Sargan-Hansen test of overidentifying restrictions does not reject the specification in Column (3).[10] Note, however, that because the partial adjustment model and adaptive expectations model are observationally equivalent in their estimable forms, the parameter estimate ($\widehat{\gamma}$) which we just interpreted as the speed of dividend adjustment may, in fact, represent the speed of expectations revision ($\widehat{\rho}$).[11]

[Insert Table 1 Here]

In Table 2, we report the main regression results for our generalized partial adjustment model. Although the estimation results are qualitatively similar across the estimation methods, our System GMM estimates reported in Columns (3) and (4) are considered better as they are known to be consistent and efficient. Moreover, our models as reported in those two columns are strongly supported by the Sargan-Hansen tests and Arellano-Bond second-order serial correlation tests. Several aspects of the estimates are of particular interest. First, the estimated speed of adjustment ($\widehat{\gamma}$) is much higher than those reported in the existing literature. Note that, regardless of estimation methods, $\widehat{\gamma}$ is also much higher than the adjustment speed estimated in Table 1. Although the results are qualitatively similar across estimation methods, $\widehat{\gamma}$ is somewhat higher with System GMM estimates ($\widehat{\gamma}^{OLS} = 0.76$; $\widehat{\gamma}^{WG} = 0.71$; $\widehat{\gamma}^{SGMM} = 0.82 \sim 0.88$). This finding corroborates our intuition that "sticky" dividends may not be evidence that firms do not actively reassess how much they should pay in dividends, but that they actively align their dividends with the "smooth" target payouts. Consequently, the actual dividends tend to be smooth as well and the adjustment speeds are, in fact, higher than previously documented. Second, the speed of expectations revision ($\widehat{\rho}$) is much lower than 1. Note that the speed is implicitly assumed to be 1 in the conventional partial adjustment models.[12] This result indicates that managers consider a longer history of past performances rather than previous-year earnings in setting

[9] Fama and French (2002) report an estimate of about 0.30. Dewenter and Warther (1998), on the other hand, obtain a much lower average estimate of 0.055 for 313 US firms studied. A somewhat higher speed from the Within Groups estimation in Table 1 may be driven by the short-panel bias (Nickell, 1981).

[10] In Columns (3), we report the set of instruments used in first-differenced equations and level equations. Arellano and Bond's (1991) second-order serial correlation test suggests that the error term ξ_{it} is an MA(1) process. This reduces the number of lags available as instruments.

[11] The parameter estimate based on System GMM does not lie between OLS and Within Groups estimators, but the goodness-of-fit for System GMM model is higher than that for Within Groups model. In any case, our main regression results are reported in Table 3.

[12] Thus, conventional partial adjustment models impose a strong restriction on the way managers form future earnings prospects and set the target dividends.

the target payouts, offering a plausible explanation for dividends' tendency to lag behind earnings (Fama and Babiak, 1968).

[Insert Table 2 Here]

Coefficients for all but one of the variables incorporated in Equation (5) are positive and significant.[13] As evidenced by the significantly positive $\widehat{\alpha}$ and $\widehat{\beta}$ in Equation (2), future earnings prospects and growth opportunities have positive influences on target dividends. We implement the analysis of covariance (ANCOVA) to further examine the relative importance of various determinants in capturing the variation in target dividends. Table 3 shows, as predicted, that the total sum of squares in the generalized model (0.593) is only a small fraction (5.8%) of the conventional model counterpart (10.145), which confirms that target dividends remain far more stable over time in the generalized model.

[Insert Table 3 Here]

Similarly, Figure 1 shows that the volatility of target dividends in the generalized model is far below that in the conventional model. The ANCOVA results also show that time-invariant firm-specific effects are the major source of the total variation. It is interesting to note that while the total variation explained by time-varying determinants is less than 25%, their incremental contribution is much smaller. Intuitively, this suggests that much of the explanatory power of existing (target) dividend determinants comes from the cross-sectional, as opposed to time-series, variation. Overall, our results provide some new evidence that firms' target payout polices may not be as puzzling as previously thought. Rather, it may be the case that managers set target payouts cautiously by conditioning them on a longer stretch of available earnings data. The smooth dividend paths observed in the market, therefore, may be rational responses to target payouts determined in such a way, resulting in higher speeds of adjustment to the targets.

[Insert Figure 1 Here]

[13] The coefficients for $Q_{i,t-2}$ are only marginally insignificant in the last two columns. For example, in Column (4), its p-value is 0.111.

4 Conclusion

This study proposes a generalized partial adjustment model of dividends in which managers form future earnings prospects adaptively and set the target dividends based on the earnings prospects. The main contribution of this study is to present new evidence with respect to the dynamic behavior of firms' dividend policies. We show that the slow adjustments of dividends to target payouts reported using conventional models largely stem from a strong restriction imposed on the way firms decide their dividend targets. Given that firms' earnings are quite volatile, the target payouts themselves will be more volatile when managers set the targets solely based on the previous earnings compared to when they use adaptive expectations. This will, in turn, lead to larger deviations of actual dividend payouts from the targets and hence slower speeds of dividend adjustments, *ceteris paribus*, making it more challenging to account for firms' dividend payout policies. If target dividends set by managers are smoother, on the other hand, actual dividends observed in the market will become more in line with the targets, driving up the speed of adjustment. Our model offers an insight that smooth dividend paths could be a consequence of managers' attempts to match dividend payouts with the targets. A variance decomposition analysis shows that firm-specific effects are predominant sources of variations in target payouts, suggesting that the majority of the total variation in target dividends is due to time-invariant factors.

Internet Appendix

The internet appendix contains empirical evidence for sticky dividends and a detailed discussion of the identification problem related to the conventional partial adjustment model and the adaptive expectations model.

References

[1] Allen, F. and Michaely, R., 1995, "Dividend policy," in R. A. Jarrow, V. Maksimovic, and W. T. Ziemba (eds.), *Handbooks in Operations Research & Management Science: Finance*, Elsevier Science, New York.

[2] Allen, F., Bernado, A. E., and Welch, I, 2000, "A Theory of dividend based on tax clienteles," *Journal of Finance*, 55: 2499–2536.

[3] Arellano, M. and Bond, S. R., 1991, "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58: 277–297.

[4] Arellano, M. and Bover, O., 1995, “Another look at the instrumental variable estimation of error-components models,” *Journal of Econometrics*, 68: 29–52.

[5] Bhattacharya, S., 1979, “Imperfect information, dividend policy, and “the bird in the hand” fallacy,” *Bell Journal of Economics*, 10: 259–270.

[6] Blundell, R. W. and Bond, S. R., 1998, “Initial conditions and moment restrictions in dynamic panel data models,” *Journal of Econometrics*, 87: 115–143.

[7] Brav, A., Graham, J. R., Harvey, C. R., and Michaely, R., 2005, “Payout policy in the 21st century,” *Journal of Financial Economics*, 77: 483–527.

[8] Chow, G. C., 2011, “Usefulness of adaptive and rational expectations in economics,” CES Working Paper No. 221.

[9] DeAngelo, H., DeAngelo, L., and Skinner, D. J., 2008, “Corporate payout policy,” *Foundations and Trends in Finance*, 3: 2–3.

[10] Dewenter, K. L., and Warther, V. A., 1998, “Dividends, asymmetric information, and agency conflicts: Evidence from a comparison of the dividend policies of Japanese and U.S. firms,” *Journal of Finance*, 53: 879–904.

[11] Easterbrook, F. H., 1984, “Two agency-cost explanations of dividends,” *American Economic Review*, 74: 650–659.

[12] Fama, E. and Babiak, H., 1968, “Dividend policy: an empirical analysis,” *Journal of the American Statistical Association*, 63: 1132–1161.

[13] Fama, E. and French, K., 1997, “Industry cost of equity,” *Journal of Financial Economics*, 43: 153–193.

[14] Fama, E. and French, K., 2002, “Testing trade-off and pecking order predictions about dividends and debt,” *Review of Financial Studies*, 15: 1–34.

[15] John, K. and Williams, J., 1985, “Dividends, dilution, and taxes: A signaling equilibrium,” *Journal of Finance*, 40: 1053–1070.

[16] Leary, M. T. and Michaely, R., 2011, “Determinants of dividend smoothing: Empirical evidence,” *Review of Financial Studies*, 24: 3197–3249.

[17] Lemmon, M. L., Roberts, M. R., and Zender, J. F., 2008, “Back to the beginning: Persistence and the cross-section of corporate capital structure,” *Journal of Finance*, 63, 1575–1608.

[18] Lintner, J., 1956, “Distribution of incomes of corporations among dividends, retained earnings, and taxes,” *American Economic Review*, 46: 97–113.

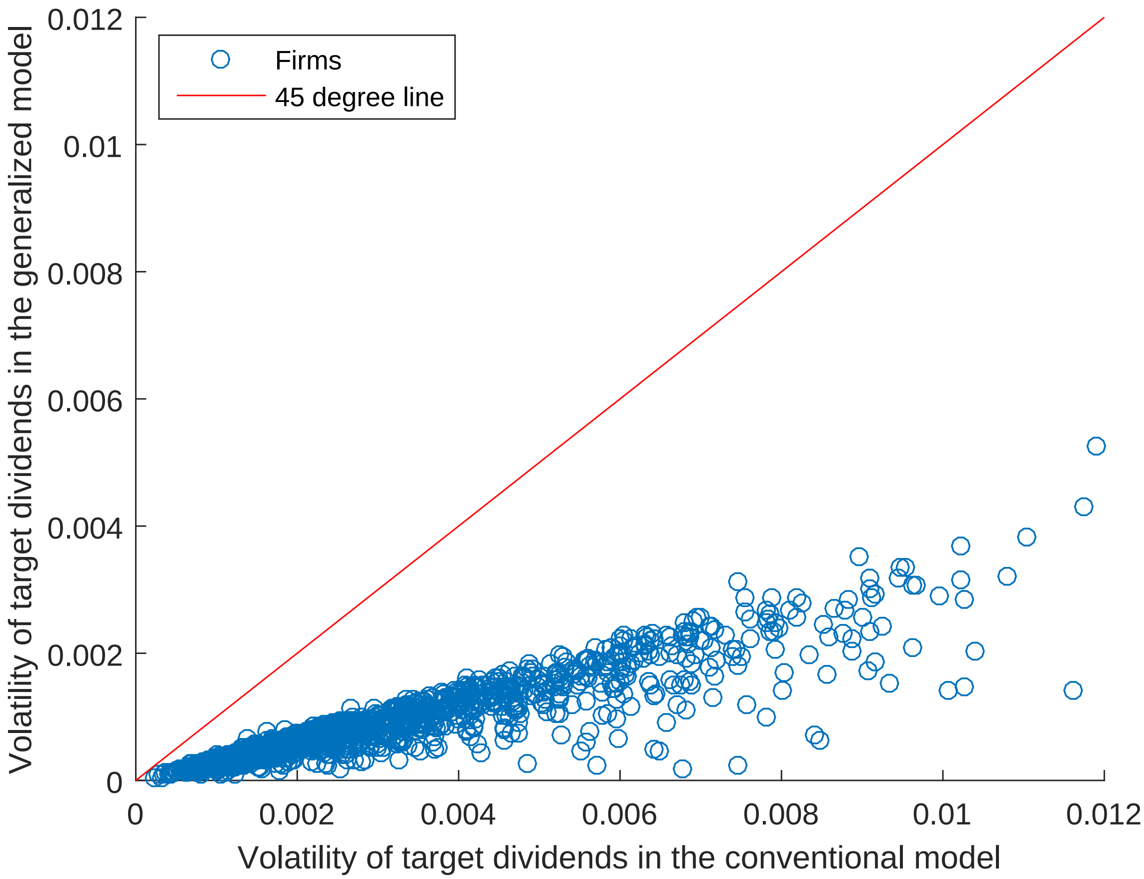
[19] Miller, M. and Rock, K., 1985. “Dividend policy under asymmetric information,” *Journal of Finance*, 40: 1031–1051.

[20] Nickell, S. J., 1981, “Biases in dynamic models with fixed effects,” *Econometrica*, 49: 1417–1426.

[21] Waud, R. N., 1966, “Small Sample Bias Due to Misspecification in the “Partial Adjustment” and “Adaptive Expectations” Models,” *Journal of the American Statistical Association*, 61: 1130–1152.

[22] Windmeijer, F, 2005, “A finite sample correction for the variance of linear efficient two-step GMM estimators,” *Journal of Econometrics*, 126: 25–51.

Figure 1: Comparison of volatilities of target dividends: conventional vs. generalized models



Note: This figure plots within-firm volatilities of target dividends in the generalized model against those in the conventional model. Each circle represents a firm among 1,376 firms.

Table 1: Estimation results for conventional partial adjustment models of dividends

ESTIMATION METHOD VARIABLES	(1) Pooled OLS $D_{i,t}$	(2) Within Groups $D_{i,t}$	(3) System GMM $D_{i,t}$
First-lagged dividends ($D_{i,t-1}$)	0.821***	0.616***	0.877***
	(0.012)	(0.023)	(0.021)
First-lagged earnings ($E_{i,t-1}$)	0.007***	0.005***	0.003**
	(0.001)	(0.001)	(0.001)
First-lagged Tobin's Q ($Q_{i,t-1}$)	0.001***	0.001***	0.000**
	(0.000)	(0.000)	(0.000)
Constant	0.002	0.007***	-0.000
	(0.001)	(0.001)	(0.001)
Firm fixed effects	No	Yes	Yes
Industry fixed effects	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes
Number of observations	28,063	28,063	28,063
Number of firms	1,383	1,383	1,383
Goodness of fit—$(Corr(D_{i,t}, \widehat{D}_{i,t}))^2$	0.747	0.743	0.745
First-order serial correlation (p-value)			0.000
Second-order serial correlation (p-value)			0.000
Third-order serial correlation (p-value)			0.872
Sargan-Hansen test (p-value)			0.154
Dividend adjustment speed ($\widehat{\gamma}$) or	0.179***	0.384***	0.123***
expectations revision speed ($\widehat{\rho}$)	(0.012)	(0.023)	(0.021)
Target-expected earnings sensitivity ($\widehat{\alpha}$)	0.041***	0.013***	0.022***
	(0.003)	(0.002)	(0.008)
Target-Tobin's Q sensitivity ($\widehat{\beta}$)	0.005***	0.002***	0.002**
	(0.001)	(0.000)	(0.001)
Instruments for first-differenced equations			
			$D_{i,t-4}, \cdots, D_{i,t-8}$
			$E_{i,t-4}, \cdots, E_{i,t-8}$
			$Q_{i,t-4}, \cdots, Q_{i,t-8}$
Instruments for level equations			
			$\Delta D_{i,t-3}$
			$\Delta E_{i,t-3}$
			$\Delta Q_{i,t-3}$
			Ind. dummies
			Year dummies

Note: In all three columns, we report standard errors that are asymptotically robust to both heteroskedasticity and serial correlation. In the last column, we report two-step GMM coefficients and standard errors which use the finite-sample correction proposed by Windmeijer (2005). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 2: Estimation results for generalized partial adjustment models of dividends

ESTIMATION METHOD VARIABLES	(1) Pooled OLS $D_{i,t}$	(2) Within Groups $D_{i,t}$	(3) System GMM $D_{i,t}$	(4) System GMM $D_{i,t}$
First-lagged dividends ($D_{i,t-1}$)	0.565***	0.474***	0.616***	0.599***
	(0.020)	(0.023)	(0.077)	(0.070)
First-lagged earnings ($E_{i,t-1}$)	0.006***	0.005***	0.003***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Second-lagged dividends ($D_{i,t-2}$)	0.306***	0.226***	0.276***	0.287***
	(0.018)	(0.018)	(0.072)	(0.066)
First-lagged Tobin's Q ($Q_{i,t-1}$)	0.001***	0.001***	0.000**	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Second-lagged Tobin's Q ($Q_{i,t-2}$)	-0.000***	-0.000**	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.002*	0.004***	-0.001*	-0.001**
	(0.001)	(0.001)	(0.000)	(0.000)
Firm fixed effects	No	Yes	Yes	Yes
Industry fixed effects	Yes	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	26,507	26,507	26,507	26,507
Number of firms	1,383	1,383	1,383	1,383
Goodness of fit—$(Corr(D_{i,t}, \widehat{D}_{i,t}))^2$	0.769	0.767	0.767	0.768
First-order serial correlation (p-value)			0.000	0.000
Second-order serial correlation (p-value)			0.327	0.225
Sargan-Hansen test (p-value)			0.467	0.101
Dividend adjustment speed ($\widehat{\gamma}$)	0.762***	0.712***	0.881***	0.823***
	(0.063)	(0.089)	(0.230)	(0.189)
Expectations revision speed ($\widehat{\rho}$)	0.673***	0.815***	0.502**	0.578***
	(0.061)	(0.090)	(0.215)	(0.175)
Target-expected earnings sensitivity ($\widehat{\alpha}$)	0.012***	0.008***	0.007**	0.008***
	(0.001)	(0.001)	(0.003)	(0.003)
Target-Tobin's Q sensitivity ($\widehat{\beta}$)	0.001***	0.001***	0.001**	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Instruments for first-differenced equations				
			$D_{i,t-4}, \cdots, D_{i,t-8}$	$D_{i,t-4}, \cdots, D_{i,t-10}$
			$E_{i,t-4}, \cdots, E_{i,t-8}$	$E_{i,t-4}, \cdots, E_{i,t-10}$
			$Q_{i,t-4}, \cdots, Q_{i,t-8}$	$Q_{i,t-4}, \cdots, Q_{i,t-10}$
Instruments for level equations				
			$\Delta D_{i,t-3}$	$\Delta D_{i,t-3}$
			$\Delta E_{i,t-3}$	$\Delta E_{i,t-3}$
			$\Delta Q_{i,t-3}$	$\Delta Q_{i,t-3}$
			Ind. dummies	Ind. dummies
			Year dummies	Year dummies

Note: In all four columns, we report standard errors that are asymptotically robust to both heteroskedasticity and serial correlation. In the last two columns, we report two-step GMM coefficients and standard errors which use the finite-sample correction proposed by Windmeijer (2005). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Variance decompositions of target dividends

Panel A. Target dividends from the conventional partial adjustment model

VARIABLES		(1) $D^{\star}_{i,t}$	(2) $D^{\star}_{i,t}$	(3) $D^{\star}_{i,t}$	(4) $D^{\star}_{i,t}$	(5) $D^{\star}_{i,t}$
Past earnings ($E_{i,t-1}$)		0.123		0.134		0.017
Tobin's Q ($Q_{i,t-1}$)			0.093	0.104		0.012
Firm-specific effects (μ_i)					0.967	0.773
Number of Observations		28,063	28,063	28,063	28,063	28,063
Root MSE		0.018	0.018	0.017	0.004	0.000
Adjusted R-Squared		0.123	0.093	0.227	0.965	1.000
Total Sum of Squares	10.145					

Panel B. Target dividends from the generalized partial adjustment model

VARIABLES		(1) $D^{\star}_{i,t}$	(2) $D^{\star}_{i,t}$	(3) $D^{\star}_{i,t}$	(4) $D^{\star}_{i,t}$	(5) $D^{\star}_{i,t}$
Expected earnings ($E^{e}_{i,t}$)		0.128		0.138		0.015
Tobin's Q ($Q_{i,t-1}$)			0.116	0.126		0.017
Firm-specific effects (μ_i)					0.963	0.745
Number of observations		22,019	22,019	22,019	22,019	22,019
Root MSE		0.005	0.005	0.004	0.001	0.000
Adjusted R-Squared		0.128	0.117	0.227	0.959	1.000
Total Sum of Squares	0.593					

Note: We compute the partial sum of squares for each effect in the model and then normalize each estimate by the total sum of squares. For example, in Column (5) of Panel B, 74.5% of the total sum of squares can be attributed to unobserved firm-specific effects (μ_i). Expected earnings are computed as follows: $E^{e}_{i,t} = \sum_{k=1}^{5}(1-\widehat{\rho})^{k-1}\widehat{\rho}E_{i,t-k}$ where $\widehat{\rho}$ is the estimated speed of expectation revision reported in Column (4), Table 3. To compute fixed effects in target dividends, we go through the following procedures. First, we compute within-firm average residuals in the dynamic regression model. Second, we add the mean of time effects to the within-firm average residuals to get firm-specific effects in dividends (η_i). Finally, we divide the firm-specific effects in dividends by $\widehat{\gamma}$ ($\widehat{\gamma\rho}$) to estimate firm-specific effects in target dividends (μ_i) in the conventional (generalized) model.

Internet Appendix for
"Sticky dividends: A new explanation"

Hyun Joong Im[*] — HSBC Business School, Peking University

Ya Kang[†] — NUS Business School, National University of Singapore

Chang Yong Ha[‡] — HSBC Business School, Peking University

December 15, 2016

Abstract

This internet appendix contains empirical evidence for sticky dividends and a detailed discussion of the identification problem related to the conventional partial adjustment model and the adaptive expectations model.

JEL classification: D9, G30, G35

Keywords: Payout Policy, Speed of Adjustment, Dividend Dynamics, Adaptive Expectations

[*]Address: HSBC Business School, Peking University, University Town, Nanshan District, Shenzhen, 518055, China; Tel: +86 (0)755 2603 3627; Fax: +86 (0)755 2603 5344; Email: hyun.im@phbs.pku.edu.cn

[†]Address: Department of Finance, NUS Business School, National University of Singapore, BIZ 2 Building #B1-3, 1 Business Link, 117592, Singapore; Tel: +65 8647 3952; Email: kangya@u.nus.edu

[‡]Address: HSBC Business School, Peking University, University Town, Nanshan District, Shenzhen, 518055, China; Tel: +86 (0)755 2603 3651; Fax: +86 (0)755 2603 5344; Email: cyha@phbs.pku.edu.cn

A Appendix

A.1 Empirical evidence for sticky dividends

We begin our analysis by studying the evolution of dividend ratios for our cross-section of firms in the spirit of Lemmon et al. (2008). Figure A.1 presents the average dividend-to-total assets ratios of three actual portfolios in "event time." The figure is constructed in the following manner. Each calendar year, we sort firms into terciles (i.e., three portfolios) based on their dividend ratios, which we denote: High, Medium, and Low. The portfolio formation year is denoted event year 0. We then compute the average dividend for each portfolio in each of the subsequent 20 years, holding the portfolio composition constant (but for firms that exit the sample). We repeat these two steps of sorting and averaging for every year in the sample period. This process generates 27 sets of event-time averages. We then take the average of the average dividend-to-total assets ratios in each "event time," which are shown in bold lines. Surrounding dotted lines represent 95% confidence intervals.

[Insert Figure A.1 Here]

Several features of the figure are noteworthy. First, there exist a great deal of cross-sectional differences in the dividend-to-total assets ratios in the initial portfolio formation period, ranging from 0.0% to 13.1%. Second, there is significant convergence among all three portfolio averages over the event time. After 20 years, the High dividend portfolio has declined from 3.8% to 2.9%, whereas the Low dividend portfolio has increased from 0.0% to 1.0%. Finally, despite the convergence, the average dividend across the portfolios 20 years later remains significantly different, both statistically and economically. The average dividend-to-total assets ratios in the High, Medium, and Low portfolios after 20 years are 2.9%, 1.7%, and 1.0%, respectively. When compared to the average within-firm standard deviation of dividend-to-total assets ratios (1.1%), this differential is economically large.

Therefore, a preliminary examination of dividend ratios suggests the presence of a permanent or long-run component that leads to highly persistent cross-sectional differences in dividend ratios, as well as a transitory or short-run component that leads to a gradual convergence in dividend ratios.

We then move on to a variance decomposition of dividend-to-total assets ratios. We begin with a non-parametric framework. More precisely, we compute the within- and between-firm variation of dividend ratios, finding that these estimates are 1.39% and 1.74%, respectively. Thus, the between-firm variation is approximately 25% larger than the within-firm variation. Intuitively, this suggests that dividend varies significantly more across firms, as opposed to within firms over time, consistently with the patterns observed in Figure A.1. We now turn to a parametric framework, analysis of covariance (ANCOVA), which enables us to decompose the variation in actual dividends attributable to different factors. Table A.1 shows that firm-specific effects (η_i) account for 51.2% of the total sum of squares (14.505) in the specification reported in Column (5).

[Insert Table A.1 Here]

A.2 An identification problem

In this subsection we show that a conventional partial adjustment model and an adaptive expectations model yield indistinguishable empirical specifications for the dividend adjustment process.[1] Nesting both models as special cases, we have shown that our proposed model allows partial adjustment behavior and expectation updating to work together to characterize firms' dynamic dividend adjustment behavior. The conventional partial adjustment models of dividends found in the literature can be specified as the following two equations:

$$\begin{aligned} D_{i,t} - D_{i,t-1} &= \gamma(D^{\star}_{i,t} - D_{i,t-1}) + \pi_j + \kappa_t + \nu_{i,t}; \\ D^{\star}_{i,t} &= \alpha E_{i,t-1} + \mu_i, \end{aligned}$$

where $D_{i,t}$ and $D^{\star}_{i,t}$ denote the actual and target dividends (scaled by assets) of firm i in year t.[2] The target dividend ratio, $D^{\star}_{i,t}$, is determined by observed earnings (scaled by assets) ($E_{i,t-1}$) and unobserved firm-

[1] Although their proof is not done in the panel data setting, Waud (1966) first shows that a conventional partial adjustment model and an adaptive expectations model yield indistinguishable empirical specifications as far as estimation is concerned.

[2] The error term in the partial adjustment equation consists of three parts. π_j and κ_t represent unobserved industry-specific and year-specific effects, and $\nu_{i,t}$ represents the idiosyncratic error with zero mean and no serial correlation. Note that the error components π_j and κ_t can be replaced by industry dummies and year dummies, respectively.

specific effects (μ_i).[3] γ denotes the speed of adjustment, which measures how fast firms adjust to their target or optimal dividends.

With some substitutions and re-parameterizations, we finally obtain the following standard dynamic panel regression model:

$$D_{i,t} = b_1 D_{i,t-1} + b_2 E_{i,t-1} + \eta_i + \xi_{i,t},$$

for $i = 1, \cdots, N$ and $t = 2, \cdots, T$ where $b_1 = (1-\gamma)$ and $b_2 = \gamma\alpha$. Therefore, the speed of adjustment can be estimated as $\widehat{\gamma} = 1 - \widehat{b}_1$. Similarly, the sensitivity of target dividends to earnings can be estimated as $\widehat{\alpha} = \widehat{b}_2/(1-\widehat{b}_1)$.

We now consider an adaptive expectations model of dividends to highlight a major potential cause of the reported slow dividend adjustment speeds. It may arise from the fact that the dynamic panel regression models used to estimate the adjustment speed can also be derived by assuming that firms adaptively form expectations of their earnings to determine their actual dividend policies. The expectation formation process is specified as follows:

$$E^e_{i,t} - E^e_{i,t-1} = \rho(E_{i,t-1} - E^e_{i,t-1}),$$

where $0 < \rho \leq 1$. $E_{i,t-1}$ is the earnings ratio observed in period $t-1$ and $E^e_{i,t-1}$ and $E^e_{i,t}$ are the earnings ratios expected to prevail in periods $t-1$ and t, respectively. ρ represents the proportion of the expectation error taken to be permanent rather than transitory. For example, if $\rho = 1$, then all of the error is taken to be permanent and $E^e_{i,t} = E_{i,t-1}$.[4]

Assume now that the expected earnings ratio ($E^e_{i,t}$) determines the actual dividend ratio $D_{i,t}$:

$$D_{i,t} = \alpha E^e_{i,t} + \mu_i + \pi_j + \kappa_t + \nu_{i,t}.$$

Again, standard substitutions and re-parameterizations will give the following standard dynamic panel re-

[3]Fama and French (2002) also model a target dividend ratio as a function of observed earnings, but do not include unobserved firm-specific effects.

[4]A firm's expected earnings can be expressed as a weighted average of its past observed earnings with geometrically declining weights if $0 < \rho < 1$. However, a firm's earnings series is a Martingale if $\rho = 1$.

gression model:

$$D_{i,t} = b_1 D_{i,t-1} + b_2 E_{i,t-1} + \eta_i + \xi_{i,t}, \tag{1}$$

for $i = 1, \cdots, N$ and $t = 1, \cdots, T$ where $b_1 = (1 - \rho)$ and $b_2 = \rho\alpha$. Therefore, the speed of expectations revision can be estimated as $\widehat{\rho} = 1 - \widehat{b}_1$.

It can be clearly seen that the reduced-form equations for the partial adjustment model and the adaptive expectations model are indistinguishable. Hence, one cannot tell whether the estimated coefficient of the lagged dividend ratio ($\widehat{b}_1$) is driven by the speed of dividend adjustment (γ) or the speed of expectations revision (ρ). That is, one cannot separately identify γ and ρ using the dynamic panel data regression model described above. Therefore, we use a generalized partial adjustment model with an adaptive expectations formation process which renders both parameters identifiable. The model is described in Section 2.

References

[1] Fama, E. and French, K., 2002, "Testing trade-off and pecking order predictions about dividends and debt," *Review of Financial Studies*, 15: 1–34.

[2] Lemmon, M. L., Roberts, M. R., and Zender, J. F., 2008, "Back to the beginning: Persistence and the cross-section of corporate capital structure," *Journal of Finance*, 63, 1575–1608.

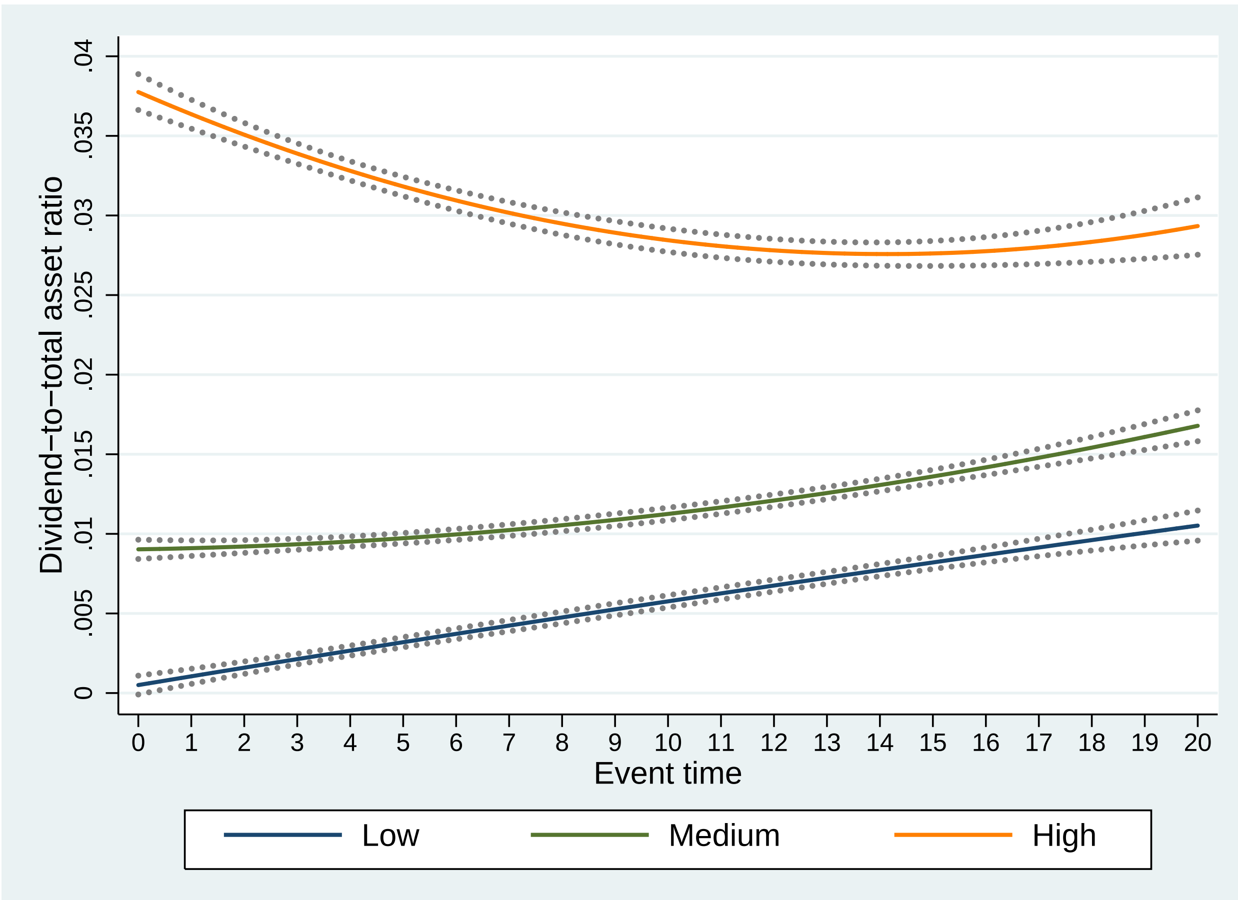
[3] Waud, R. N., 1966, "Small Sample Bias Due to Misspecification in the "Partial Adjustment" and "Adaptive Expectations" Models," *Journal of the American Statistical Association*, 61: 1130–1152.

Table A.1: Variance decompositions of actual dividends

VARIABLES		(1) $D_{i,t}$	(2) $D_{i,t}$	(3) $D_{i,t}$	(4) $D_{i,t}$	(5) $D_{i,t}$
Past earnings ($E_{i,t-1}$)		0.084		0.083		0.004
Tobin's Q ($Q_{i,t-1}$)			0.041	0.047		0.005
Firm-specific effects (η_i)					0.625	0.512
Number of Observations		28,267	28,267	28,267	28,267	28,267
Root MSE		0.022	0.022	0.021	0.014	0.014
Adjusted R-Squared		0.078	0.041	0.124	0.605	0.617
Total Sum of Squares	14.505					

Note: We compute the partial sum of squares for each effect in the model and then normalize each estimate by the total sum of squares. For example, in Column (5), 51.2% of the total sum of squares can be attributed to unobserved firm-specific effects (η_i).

Figure A.1: Average dividend-to-total assets ratios of actual dividend portfolios in event time



Note: This figure plots average dividend-to-total assets ratios of actual dividend portfolios. To obtain this figure, first, for each calendar year from 1988 to 2014, we sort firms into terciles based on dividend-to-total assets ratios (denoted as High, Medium, and Low) and calculate the average ratios for each of the three portfolios in each of the subsequent 20 years, holding the portfolio composition constant. We repeat this process for all the years from 1988 to 2014. Second, we take the average of the average dividend-to-total assets ratios in each "event time", which are shown in bold lines. Surrounding dotted lines represent 95% confidence intervals.