Dividend Taxation and Merger Behavior: A New View Explanation For The Post-Merger Performance Puzzle

Eric Ohrn* Nathan Seegert†

September 2013

Abstract

Mergers and acquisitions are a large part of the United States economy. Between 2000 and 2012, the dollar value of merger and acquisition activity was $12.78 trillion. Mergers are critical to economic growth because, done well, they can capitalize on positive synergies and economies of scale, thereby increasing efficiency and creating value for shareholders. However, executed poorly, mergers can dampen innovation, decrease efficiency, and destroy shareholder value. Because of the scale and significance of this topic, large literatures in finance and economics have developed that discuss possible mechanisms which distort merger and acquisition behavior. We follow in this tradition by examining dividend taxation and its affect on merger behavior. We develop a "new view" of corporate taxation model in which dividend taxation causes inefficient mergers. We find strong empirical support for the model, demonstrating long run returns are 8 to 10 percent higher for dividend paying firms following the 2003 US dividend tax reform. These results suggest the tax mechanism we propose may be able to explain, at least in part, the post-merger performance puzzle in the finance literature.

Keywords : Mergers and Acquisitions ; Taxation.

JEL Classification : G34; H20; H25

*ericohrn@umich.edu. Department of Economics, University of Michigan
†nathan.seegert@business.utah.edu. Department of Finance, University of Utah

Mergers and acquisitions are a large part of the United States economy. Between 2000 and 2012, the dollar value of merger and acquisition activity in the US was $12.78 trillion. Mergers are critical to economic growth because, done well, they can capitalize on positive synergies and economies of scale, thereby increasing efficiency and creating value for shareholders. However, executed poorly, mergers can dampen innovation, decrease efficiency, and destroy shareholder value. Because of the scale and significance of this topic, large literatures in finance and economics have developed that discuss possible mechanisms which distort merger and acquisition behavior. In this paper, we examine dividend taxation and its effect on merger behavior. We develop a "new view" of corporate taxation model in which dividend taxation causes inefficient mergers. We find strong empirical support for the model, demonstrating long-run returns are 8 to 10 percent higher for dividend paying firms following the 2003 US dividend tax reform.

There is reason to believe that merger and acquisition behavior is being distorted. The finance literature documents that on average twenty-four months after a merger returns are eighteen percent lower than would be expected in the absence of the merger [Agrawal and Madelker, 1990, Agrawal et al., 1992, Gregory, 1997, Myers, 1984]. This empirical puzzle seems to be robust across countries and through time despite the fact that researchers have been discussing it for twenty years [Agrawal and Jaffee, 2000]. Furthermore, it seems difficult to reconcile this empirical puzzle with a model in which firms seek to maximize shareholder value. This difficulty has inspired many behavioral explanations, such as CEO hubris.

We propose an alternative theory for this puzzle in which value-maximizing behavior in the presence of dividend taxation may result in inefficient mergers and negative post-merger performance. Critically, dividend taxation decreases the value of capital within the firm and by extension the opportunity costs of an acquisition. We formalize this tax mechanism within a "new view" of corporate taxation model characterized by internal investment levels that are undistorted by dividend taxation. In this framework, acquisitions are fundamentally different from internal investment because they produce an implicit transfer from firms to shareholders. Therefore, when the cost of distributing funds through dividend payments is distorted by the dividend tax, firms substitute away from dividends and towards mergers and acquisitions as an alternative channel for transfers.

This tax mechanism has two important implications. First, in reaction to a decrease in the dividend tax rate the mechanism predicts a substitution toward dividend payments. Second, in reaction to a dividend tax rate decrease, long-run post-merger returns should be higher for firms paying a dividend and should be unchanged for firms not paying a dividend. The first implication is consistent with the empirical evidence on dividends following the 2003 dividend tax reform in the United States [Chetty and Saez, 2005]. The second implication

we empirically test using variation created by the 2003 dividend tax reform.

Our empirical examination uses data on nearly 7,000 acquisitions by publicly traded companies between 1998 and 2010. We collect data from Bureau Van Dijks Amadeus Zephyr database, the Compustat North American Fundamentals Quarterly database, and the Center for Research in Security Prices (CRSP) database. In later analyses we extend this data to the years 1992 to 2012, over 20,000 acquisitions by publicly traded companies, using Thomson and Reuters SDC database. We implement a difference-in-difference empirical strategy motivated by the theoretical model's result that post-merger long-run returns should be affected by a dividend tax change for firms paying a dividend but should be unaffected for firms not paying a dividend. In this specification, we use firms that distribute retained earnings through repurchases of shares as a benchmark to control for other time-varying factors. Firms active in repurchasing shares are a good benchmark because they are similar to firms that pay dividends based on observable characteristics, and are unaffected by the dividend tax change. Furthermore, using firms actively repurchasing shares as the control group we empirically test and fail to reject the common trend assumption, the key identifying assumption of the difference-in-difference specification.

This paper contributes to both the economics literature on corporate taxation and the finance literature on mergers and acquisition behavior. We make three significant contributions. First, we develop a model that demonstrates the distortions of the dividend tax on merger and acquisition behavior. Second, we provide empirical evidence supporting the theoretical model and its implications, demonstrating the distortion due to dividend taxation is real and empirically large. Third, we provide a value-maximizing explanation for the post-merger performance puzzle that is theoretically and empirically supported.

1 Documenting and Explaining Merger Behavior

Two literatures have developed which separately explore aspects of merger and acquisition behavior that we seek to study. The first literature is concerned with documenting and explaining why mergers fail relative to benchmark trends. The second literature explores the impact of taxation on mergers.

1.1 The Post-Merger Performance Puzzle

Franks et al. [1991] was the first paper devoted entirely to explaining long-run post-merger performance. The paper introduces techniques for analyzing post-merger performance that the subsequent literature follows. Franks et al. [1991] describe the amount a firm was under- or over-performing over time by comparing a counterfactual performance with the firm's actual performance, measured by its stock price. The cumulative difference between the counterfactual performance and the actual performance of the firm is defined as the Cumulative Abnormal Return or CAR. Subsequent literature improves upon the methods for predicting counterfactual performance, but the Franks et al. [1991] framework of studying CARs for individual firms and the Cumulative Average Abnormal Return (CAAR) for all merging firms is the benchmark in this field of study.

Franks et al. [1991] investigates 399 acquisitions, both mergers and tender offers, that occurred during the period 1975-1984, where both the acquirer and the target were on NYSE/AMEX. The Franks et al. [1991] findings suggest that cumulative average returns for acquiring firms are not statistically different from zero, however the subsequent literature finds different results using Franks et al. [1991]'s methods. For example, Agrawal et al. [1992] repeat the work done by Franks et al. [1991] but extend the time period to cover the years 1955 to 1987 and exclude tender offers from the sample, which would otherwise bias the results towards zero. Agrawal et al. [1992] report five-year CAARs of -0.1026, (interpreted as ten percentage points under-performance relative to counterfactual stock performance). Agrawal and Madelker [1990] find very similar results to Agrawal et al. [1992] when the counterfactual performance is adjusted to control for firm size and the book-to-market ratio.

Gregory [1997] provides the most comprehensive study of post-merger performance to date. The paper computes counterfactual performance using multiple benchmarks. Furthermore, the study, which examines mergers in the United Kingdom (UK) between 1984 and 1992, provides an out-of-sample test of the merger-failure phenomenon observed in the United States. Gregory finds two-year buy-and-hold (or multiplicative) CAARs which vary between -0.1182 and -0.18 using six different counterfactual models. All point estimates are statistically different from zero.

The survey work by Agrawal and Jaffee [2000] covers these results and lists several other similar findings. They conclude "Taken together, we believe that the post Franks, Harris, and Titman (1991) articles suggest strong evidence of an anomaly following mergers." The empirical anomaly that mergers often result in poor performance and that this trend is consistent over time is often called the post-merger performance puzzle.

1.2 Possible Explanations

Several mechanisms have been offered to explain the post-merger failure puzzle. Agrawal and Madelker [1990] propose two mechanisms for abnormal negative post-merger performance: speed of adjustment and size of the merger. The speed of adjustment hypothesis posits that post-merger abnormal returns are the result of the market reacting positively to the merger upon announcement. To test this hypothesis they attempt to determine whether abnormal returns upon announcement are negatively correlated with post-merger cumulative abnormal returns. Looking at post 1960 data they find no evidence of negative correlation, suggesting that the speed of adjustment is not the mechanism driving the post-merger failure puzzle. Agrawal and Madelker [1990] also examines whether the size of the merger, which may impact the speed at which information is capitalized into the price of the acquiring firm stock, impacts firm CARs. They again find no empirical relationship. Franks et al. [1991], Loughran and Vijh [1997] also find no connection between the size of the merger and the post-merger cumulative abnormal return.

Another proposed mechanism, based on work by Myers [1984], is that firms with private information about the value of their own stock are able to buy companies at a discount using their stock when it is overvalued. Loughran and Vijh [1997] find support for this mechanism by demonstrating that mergers paid for with stock fail more often. However, Gregory [1997] finds evidence that abnormal returns are highest for mergers paid for with a combination of stock and cash. In addition, if method of payment were the mechanism driving the post-merger failure the new information should be capitalized into the share price immediately upon merger, not over a period of 24 to 50 months which the literature focuses on.

A theory referred to as "performance extrapolation" has also been considered to explain the post-merger performance puzzle. Rau and Vermaelen [1998] posits that the market overvalues actions of firms that have performed well in the past. Thus, upon announcement of a merger the market overvalues the activities of these "glamour" firms. Similarly, the market undervalues mergers performed by firms with poor past performance, or "value" bidders. As the information of the true value of the merger is revealed, the glamour firm prices drop and the value firm prices increase. Rau and Vermaelen [1998] cite past literature which links higher announcement returns to higher Q-ratios. Rau and Vermaelen [1998] find modest evidence that long-run under-performance is linked to glamour firms. However, their findings are inconsistent with method of payment explanations. Their findings are also inconsistent with the Agrawal and Madelker [1990] results, which do not demonstrate a negative relationship between announcement abnormal performance and subsequent cumulative abnormal returns.

To summarize the post-acquisition stock performance literature, there is ample evidence from a number of studies using varied methodologies and samples to demonstrate that acquiring firms under-perform relative to predicted counterfactual performance. Several explanations have been advanced to account for this apparent anomaly but none have been widely accepted. The post-merger performance puzzle remains an open question two decades after Franks et al. [1991] first described it.

1.3 Mergers and Taxation Literature

We posit that taxation may, at least in part, explain the post-merger performance puzzle. While this is a new mechanism for the literature following Franks et al. [1991], the effect of taxation on merger behavior has been studied in the corporate taxation literature.

Auerbach and Reishus [1987] examine mergers that occurred before 1986. During this time period, acquiring firms could gain windfall tax benefits generated by mergers (especially through the acquisition of tax loss carry-forwards or the basis step-up of assets). Auerbach and Reishus [1987] offer preliminary results that suggest: 1) many acquisitions provide an opportunity for acquiring firms to receive some sort of tax benefit, 2) in a small minority of cases, these benefits are large relative to the value of the acquiring firm, and 3) even in cases where relatively large tax benefits exist, there is no strong evidence that this is the driving force in the merger. In addition to the benefits of attaining tax loss carry-forwards, Auerbach and Reishus [1987] also identify a possible tax wedge; when transactions are made via nontaxable stock transactions, the shareholders of the acquiring company may pay a premium because they are able to diversify their portfolio without realizing capital gains. Given this tax wedge theory, the acquirer should realize larger gains when paying with stock.[1] Although Auerbach and Reishus [1987] do not find strong evidence that merger behaviors are driven by tax attributes of mergers, the authors do find that tax loss carry-forwards, at least prior to 1987, are the most beneficial tax attributes of acquired firms.

A subsequent paper by the same authors, Auerbach and Reishus [1988], present empirical findings that support the theory that some types of tax benefits can influence merger behavior. Auerbach and Reishus [1988] find that tax loss carry-forwards (TLCFs) on the part of the acquiring firm do influence the takeover of firms that earn positive income. In these cases, the TLCFs can be used to offset the income of the acquired firm. Auerbach and Reishus [1988] also find that firms that issue share repurchases in the two yeas prior are less likely to merge which is "hard to reconcile with the theory that firms seek acquisitions to free trapped equity." This pattern does not exist in our data.

[1] This theory runs counter to the method of payment empirical findings.

In reviewing the results from the 1987 and 1988 Auerbach and Reishaus papers, the reader should keep in mind that many of the tax benefits of mergers were eliminated in provisions contained in Tax Reform Act of 1986.

The corporate finance literature presented here predicts that taxation should be a driving force behind mergers if 1) the tax attributes of acquiring or acquired firms may be better used after the merge or 2) if there exists some "trapped equity" incentives that make mergers more attractive options than some opportunity cost of outside investment. While strong empirical results are generally lacking in this literature, some of these theories may be applied to help us understand the post-merger performance puzzle. Our theory seeks to bridge the gap between these two literatures. We explain that mergers, and in particular inefficient mergers, are driven by tax-motivated incentives.

2 The Dividend Tax And Merger Behavior

2.1 Model Introduction

We consider a firm's decisions to acquire other firms within a standard model, following Auerbach [1979], where competitive firms finance their projects through retention of earnings, sales of equity, and issuance of one term debt. The model abstracts from risk and other complexities in order to present the theoretical results in a straightforward and conventional setting. The model formalizes the intuition that a firm's decision to acquire another firm depends on the marginal benefit of additional capital in the firm, which, in turn, depends on whether a firm is distributing dividends and the dividend tax rate.

We model the benefits of mergers as an enhancement to the production capabilities of the firm. Modeling mergers in this way captures production synergies in a straightforward manner and is flexible enough to capture various other benefits of mergers with minimal reinterpretation. We assume acquiring firms gain the production capacity of the target firm at the present value costs of its production stream. Without taxation a firm will be indifferent about acquiring another firm if there are no synergies and no additional costs, or equivalently the synergies exactly offset the additional costs. In this case merger behavior is efficient and the firm only accepts mergers when the productive synergies outweigh the costs of merging.

In this "new view" of corporate taxation model, traditional investment choices are unaffected by dividend taxation. However, acquisition investments are fundamentally different

as they produce an implicit transfer from firms to shareholders.[2] When dividend payments are taxed it increases the cost of using dividend payments to transfer earnings to shareholders causing firms to substitute toward mergers and acquisitions as an alternative method. Therefore, an unintended consequence of the dividend tax is that it creates an incentive for firms to make otherwise inefficient acquisitions as an alternative method of making transfers to shareholders, as long as the inefficiency is less costly than the tax liability.

2.2 Model

The firm maximizes the present value of the net cash flows to shareholders,

$$\int_0^\infty e^{-\rho t}[(1-\theta)D_t - E_t]dt, \tag{1}$$

consisting of after-tax dividends $(1-\theta)D_t$ minus equity issuances E_t.[3]

There are two state variables; the capital stock, K_t and the production factor, A_t, which is augmented through time by the accumulated synergies obtained through acquisitions. Each instant the firm draws an acquisition opportunity from the synergy distribution $f(z)$ which is positive everywhere on the support $[0, \bar{z}]$.[4] Draws from the distribution are assumed to be identically and independently distributed. In equilibrium a firm chooses a range of acceptable mergers, $z^A = \bar{z} - z^T$, such that all merger opportunities above the threshold, z^T, are accepted and all below are rejected. As the range of acceptable acquisitions increases, the probability increases that an acceptable acquisition is drawn,

$$P(\bar{z} - z^A) = 1 - F(\bar{z} - z^A), \tag{2}$$

but the average synergy of an acquisition accepted decreases. Therefore, in choosing the threshold value, z^T, firms balance the benefits from increasing the average quality of the merger opportunity with increasing the frequency with which the firm makes an acquisition.

When a firm makes an acquisition, it obtains the production capabilities of the target firm (at cost) and any production synergies that exist between the two firms. The synergies that exist between the two firms augment the acquiring firm's production factor. The expected

[2]We restrict equity issuances to be positive in the model abstracting from firms that repurchase stock as a method of transferring earnings from the firm to shareholders.

[3]The model is presented in continuous time for analytical ease, all of the results are robust to modeling time discretely.

[4]Restricting merger opportunities to be on a positive support models the fact that the two companies could continue to be run separately. All transaction and integration costs are captured in the cost function.

change in the firm's production factor,

$$\dot{A} = P(\bar{z} - z^A)E[z|z \geq \bar{z} - z^A] \equiv \tilde{z}, \qquad (\lambda_A) \tag{3}$$

increases with the probability of a merger and the average quality of the merger conditional on the merger being accepted.

The cost of the acquisition is the present value of the production capabilities of the target plus any additional costs of combining the two firms. The additional costs of combining the two firms may include fixed costs to acquiring a firm, a premium paid to the shareholders of the target firm, or any adjustment costs. These additional costs are modeled as a function of the range of acceptable mergers, z^A, and is assumed to be increasing and sufficiently convex with respect to z^A.[5,6] The additional cost of an acquisition comes at the expense of having less capital in the firm, fewer current and future dividends, more debt, or more equity issuance according to the capital stock law of motion,

$$\dot{K} = [AQ(K_t + B_t) - rB_t - \delta(K_t + B_t)](1 - \tau) + E_t - D_t - c(z^A) \qquad (\lambda_k). \tag{4}$$

Therefore, the opportunity cost of an acquisition depends on the acquiring firms current financing decisions; whether the acquirer is issuing equity, investing out of retained earnings, or is paying out dividends. This intuition is formalized in the first order conditions where equity issuances (Lagrange multiplier μ_E) and dividend payments (Lagrange multiplier μ_D) are constrained to be nonnegative and instantaneous borrowing, B_t, is constrained to be less than some borrowing limit (Lagrange multiplier μ_B).

In equilibrium, for firms that are not credit constrained, $\mu_B = 0$ and so from the first order condition with respect to debt,

$$\frac{\partial H}{\partial B_t} = \lambda_k[A_t Q'(K_t + B_t) - r - \delta](1 - \tau) - \mu_B = 0, \tag{5}$$

it is clear the firm will borrow until the marginal product of capital in the firm is equal to the interest rate plus the depreciation rate. For firms that issue new equity to finance investment, $\mu_t^E = 0$ implying from the first order condition with respect to equity issuances,

[5]The assumption that the cost function is increasing in the probability a merger is accepted assumes the fixed costs of merging is sufficiently large. For example, suppose the additional costs are equal to a fixed cost and a fraction of the synergies, $c(z^A) = P(z > \bar{z} - z^A)\Phi + \nu E[z|z > \bar{z} - z^A]$ where Φ is the fixed cost and ν is the fraction of the synergies paid to the target firm's shareholders. Then this assumption holds as long as $\Phi > \nu(\partial E[z|z > \bar{z} - z^A]/\partial z^A)/(\partial P(z > \bar{z} - z^A)/\partial z^A)$.

[6]Formally the additional cost function is sufficiently convex when $c''(z^A) > c'(z^A)\frac{\partial^2 \tilde{z}}{\partial (z^A)^2}/\frac{\partial \tilde{z}}{\partial z^A}$.

$$\frac{\partial H}{\partial E_t} = -1 + \lambda_K + \mu_E = 0 \tag{6}$$

that the shadow price of capital within the firm must equal 1.

Lemma 1: For firms that issue new equity, the shadow value of capital within the firm does not depend on the dividend tax rate.

Alternatively, for firms that pay dividends, $\mu_t^D = 0$ implying from the first order condition with respect to dividend payments,

$$\frac{\partial H}{\partial D_t} = (1 - \theta) - \lambda_K + \mu_D = 0, \tag{7}$$

that the shadow price of capital within the firm equals $1 - \theta$. Therefore, the shadow price of capital within the firm is highest for firms issuing equity to build capital, is lower for firms not engaging in equity issuances or dividend payouts, and is the lowest for firms that payout a dividend.

Lemma 2: For dividend paying firms, the shadow value of capital within the firm decreases with the dividend tax rate.

From the first order condition with respect to dividend payments, in equation (7), for firms that pay a dividend

$$\lambda_K = 1 - \theta \tag{8}$$

therefore, $\partial \lambda_K / \partial \theta < 0$.

In equilibrium, the firm sets its acquisition threshold such that the expected marginal benefit of accepting a wider range of acquisition opportunities equals the increase in expected costs of accepting a wider range. The first order condition with respect to the set of acceptable acquisitions,

$$\frac{\partial H}{\partial z^A} = -\lambda_K c'(z^A) + \lambda_A \frac{\partial \tilde{z}}{\partial z^A} = 0, \tag{9}$$

perfectly demonstrates this tradeoff.

Lemma 3: Firms with a lower opportunity cost of capital within the firm accept a wider range of mergers.

Total differentiating the first order condition with respect to the set of acceptable acquisitions, in equation (9), and rearranging demonstrates,

$$\frac{dz^A}{d\lambda_K} = \frac{-c'(z^A)}{\lambda_K c''(z^A) - \lambda_A \tilde{z}''(z^A)} < 0 \tag{10}$$

which is negative by the assumption that the additional cost of mergers function is sufficiently convex.[7]

These first order conditions characterize the optimal finance, investment, and payout strategies of a firm, from which three lemmas are derived. The equilibrium behavior of the firm in this model is the same as the equilibrium behavior of firms under the "new view" of corporate finance (e.g. firms do not issue equity and pay dividends at the same time and the dividend tax does not distort internal investment behavior). This model extends "new view" models with the addition of the characterization of a firm's acquisition behavior, summarized in the first result

Result 1: The set of acceptable acquisitions is increasing with respect to the dividend tax rate for firms that pay dividends and is independent of the dividend tax rate for firms not paying a dividend.

Result 1 follows directly from lemmas 1-3, details are given in appendix 7.1.

This result demonstrates the dividend tax distorts firm acquisition behavior, causing wealth maximizing firms to accept suboptimal acquisitions, from a pareto efficient criteria. This result has important policy ramifications suggesting contrary to most "new view" of corporate taxation models that economic activity is not neutral with respect to dividend taxation.[8] Furthermore, the deadweight loss from this distortion is likely to be large given the level of the dividend tax rate and the value of merger and acquisition behavior in the economy.

In the following analysis we use the first order conditions above to solve for the shadow price of capital and debt to demonstrate the affects of dividend taxation in terms of the market value of equity. The following analysis has two objectives. First, the following analysis reiterates Result 1 in terms of the market value of equity connecting firm's acquisition policy to its long-run abnormal returns; used in the empirical section to test the model and Result

[7] See footnote 6.

[8] This model explicitly considers the opportunity cost of an acquisition which intuitively affects the marginal benefit and cost calculation. This model is consistent with work by Hans Sinn Werner who claims in his 1991 article, "only when it is clear where an additional dollar used for investment comes from and where its returns are going, is it possible to calculate the tax burden on marginal investment and to find out which minimum pretax return is required to make this investment profitable," Sinn [1991].

1. Second, the following analysis demonstrates the model provides a possible explanation for the post-merger performance puzzle in the finance literature.

2.3 Dividend Taxation and The Post-Merger Performance Puzzle

This section considers the theoretical implications of the model pertaining to the post-merger performance puzzle. Previous research has documented that on average, firms' abnormal post-merger returns are negative, a phenomenon which seems unexplainable in a competitive equilibrium. This section demonstrates that our model provides a possible explanation consistent with firms maximizing shareholder value.

The market value of equity in a firm is given by,

$$V = \lambda_k K + \mu_B B \tag{11}$$

for each point in time by Euler's theorem assuming the production function of the firm is linearly homogenous. Replacing the shadow price values of capital and debt using the optimal behavior derived from equations (5), (6), (7), and (9) this equation reduces to the intuitive relationship,

$$V = (1-\theta)(KP_K - B) \tag{12}$$

which states the market value of equity is given by the value of capital minus debt net of dividend taxation. Consider two firms, A and B, with the market value as given in equation (12). Suppose firm A decides to acquire firm B, either by decreasing investment or equivalently by taking out a loan of size

$$B^{A+} = V^B, \tag{13}$$

to obtain the net synergies between firms given in reduced form by, ζ.[9] The new market value for firm A is given by

$$\begin{aligned} V^{A,new} &= (1-\theta)(K^A P_K + K^B P_K - B^A - B^B - B^{A+} + \zeta) \\ &= V^A + \theta V^B + (1-\theta)\zeta \end{aligned} \tag{14}$$

which is the old market value of firm A plus the market value of firm B multiplied by the dividend tax rate plus the value of the synergies net the dividend tax. This merger

[9]Net synergies are defined as the benefit received from the augmentation of A_t minus the costs of making the acquisition.

does not affect the market value for firm B shareholders, unless they are able to be paid a premium through negotiations; however, firm A shareholders receive a gain. Specifically, firm A shareholders receive a transfer equal to θV^B and the benefits from the synergies, net the dividend tax. In equilibrium, firms make acquisitions such that the benefit is positive, $\theta V^B + (1-\theta)\zeta > 0$. When the dividend tax rate is zero this condition implies only acquisitions with positive net synergies, ζ, are accepted. However, when the dividend tax rate is not zero this condition implies some acquisitions with negative net synergies are accepted, and even worse acquisitions are accepted the larger the dividend tax. This condition reiterates Result 1 in terms of the market value of the firm.

Because this condition is derived from optimal behavior assuming equilibrium prices, this is a general equilibrium condition incorporating not only the direct changes in the tax burden but also indirect effects incurred from moving to a new market equilibrium. In addition, while this result was derived under the assumption of debt-financing the result is more general. In fact, in a Miller equilibrium the attractiveness of this acquisition is independent of whether it is financed with debt or equity.

The intuition for this result is that mergers and acquisitions are transfers to shareholders and a substitute for dividend payments. Consider two scenarios, one in which the firm acquires one dollar of stock from another firm and one in which a firm distributes a dollar of retained earnings through a dividend payment which the shareholder then invests in the capital market. In both scenarios the original firm's growth path remains the same and the household owns, either directly or indirectly, some additional stock. The difference in the two scenarios arises from the fact that a tax liability is triggered in the second scenario leaving the household with $(1 - \theta)$ in stock value of the second firm instead of one dollar in the first scenario. The net advantage for the shareholder is directly related to the size of the dividend tax just as equation (14) reveals.

Optimal merger behavior in the model seems to correspond with the empirical findings in the finance literature that on average short-run abnormal returns are positive but long-run abnormal returns are negative. All information known at the time of the announcement is immediately capitalized into the market value of the firm and its short-run abnormal returns. At the time of the announcement shareholders of the acquiring firm know, at least in expectation, that the combined benefits from the tax incentive transfer and the net synergies will be positive, causing positive short-run abnormal returns. Over the next roughly twenty-four months the value of ζ is realized and is capitalized into the market value of the firm and its long-run abnormal returns. In equilibrium, the expectation of ζ is decreasing with respect to the dividend tax rate, causing the long-run abnormal returns to decrease as well.

This section provides a mechanism with firms that maximize shareholder value that can explain, at least in part, the post-merger performance puzzle and that is empirically testable. The question remains whether, even theoretically, this mechanism is large enough to explain this puzzle. To answer this question, consider the tax rates that prevailed in the United States before and after the 2003 tax rate changes. Dividend payments went from being taxed at a rate as high as 39.6 in the 1990s to having a top marginal tax rate of 15 percent after 2003.[10] Prior to 2003 this implies the tax incentive would be almost forty percent of the market value of the purchased firm, an amount sufficient to produce a real economic impact. In contrast, after 2003 the tax incentive drops to fifteen percent of the market value of a purchased firm, a difference of almost twenty-five percent. Given the tax rates that prevailed in the United States before and after 2003 we expect there to be a large tax incentive for acquisitions and a large change in those incentives after 2003.

3 Data Collection and Construction

To test the theoretical predictions of the model, data describing merger and firm characteristics. The abnormal performance of the acquiring firm post-merger are compiled using this data and details are provided in appendix 7.2. The sources, collection, and cleaning of these data are described in the following three subsections. The last subsection uses these data to report descriptive statistics and test assumptions of the empirical design.

3.1 Merger Data

Merger characteristics are taken from the Bureau Van Dijk Amadeus Zephyr database, with permission of Zephyr. We examine mergers that occurred between the dates of January 1, 2000 and December 31, 2010. We restrict our sample in several ways. First, acquisitions made by private companies not listed on the NYSE or NASDAQ are excluded because they lack a stock price by which to measure performance. We further limit our sample to 100 percent acquisitions to avoid merger type effects. Finally, we exclude firms with more than twenty 100 percent acquisitions during our sample period.[11] This exclusion limits the effects from overlapping mergers and excludes mostly large banking firms from our sample.

The Zephyr database has extensive data on mergers and acquisitions. For our analysis

[10] The top marginal income tax rate in the United States decreased from 39.6 in 2000 to 39.1 in 2001 and 38.6 in 2002. From 2003-2007 the marginal tax rate on dividend payments was only 5 percent for low income households and dropped to 0 percent from 2008-2012.

[11] The results are robust to setting other limits on the maximum number of mergers per firm.

the most pertinent features capture the timing of the merger, including the announcement and official date, and merger type. We match the merger data from Zephyr with data from CRSP and Compustat databases.

3.2 Stock Prices and Firm Characteristics

Stock prices are taken from the CRSP database. These data allows us to calculate the abnormal returns for acquiring firms (using the month end closing stock price) and the dollar value of repurchased shares (using the quarter end closing stock price). Acquiring firm characteristics are taken from the Compustat North American Fundamentals Quarterly database. From the Compustat database we extract acquiring firm data on dividends, share repurchases, assets, sales, and retained earnings.

Cash dividends paid per year are referred to simply as dividends throughout this paper. Compustat provides the dividends paid as a number reported quarterly that accumulates throughout the year. To transform the variable into cash dividends paid per quarter, we use the accumulating variable value for the first quarter of the year and then for each subsequent quarter take the difference in the accumulating variable to be the cash dividends paid in that quarter. These values are used to determine which firms regularly pay a dividend to identify them as the treatment group in the empirical analysis.

Compustat provides data on shares repurchased starting with the first quarter of 2004. We follow Stephens and Weisbach [1998] in imputing shares repurchased as the dollar value of decreases in shares outstanding from Compustat. This method is internally valed as it closely approximates repurchase behavior in years when Compustat records actual repurchase data. As with the dividends, these values are used to determine which firms regularly repurchase shares to identify them as the control group.

Throughout the proceeding analysis, we also use firm controls in an effort to eliminate any correlation between merging behavior, payout behavior, firm size, retained earnings, and productivity. To control for these characteristics, we include in our analysis assets, retained earnings, percent cash merger, sales, percent stock merger, marginal q, and the book-to-market ratio. We average these by firm for the two years prior to the merger to attain average firm characteristics.

3.3 Descriptive Statistics

The empirical strategy compares long-run stock returns of dividend paying firms making acquisitions before and after the dividend tax rate change in 2003. The empirical strategy uses firms that regularly repurchase shares to control for all other time varying variables that may affect long-run stock returns after an acquisition. Table 1 demonstrates the similarity of firms that regularly pay a dividend and firms that repurchase shares. On average, dividend firms have more retained earnings but fewer assets than firms that repurchase shares, however they are more similar in these categories than firms that do neither or both. In fact, across most of these statistics firms that pay dividends and firms that repurchase shares are more similar than firms that do neither or both, supporting the use of firms that repurchase shares as a control for firms that pay a dividend.

Table 1: Descriptive Statistics

	(1) Dividend Only	(2) SBB Only	(3) Neither	(4) Dividend and SBB
Retained Earnings	374.3	182.7	-268.9	2,009
	(1,040)	(2,264)	(3,104)	(5,636)
Total Assets	3,141	3,796	1,242	17,967
	(10,919)	(27,569)	(4,268)	(81,356)
Cash Flow	0.144	0.185	0.114	0.201
	(0.331)	(0.444)	(0.696)	(0.318)
Marginal q	2.058	2.377	3.227	2.019
	(1.315)	(1.505)	(2.258)	(1.212)
CAR 12 mo.	-5.322	-5.419	-9.223	-5.040
	(29.37)	(30.80)	(40.27)	(23.00)
CAR 24 mo.	-17.73	-18.38	-27.52	-13.34
	(18.53)	(19.32)	(23.79)	(13.77)
Observations	382	1,841	861	2,555

Notes: These descriptive statistics support the use of firms that repurchase shares as the benchmark, or control, for firms that pay a dividend based on the similarity across these observable characteristics.

Dividend Only: firms that regularly pay a dividend but have not repurchased shares.

SBB only: firms that regularly repurchase shares but have not paid a dividend.

Neither: firms that do not distribute retained earnings regularly.

Dividend and SBB: firms that regularly pay dividends and repurchase shares.

The advantage of using a difference-in-difference (DD) strategy is that it is able to control for differences in unobservable characteristics between dividend and share repurchasing firms that may affect long-run stock returns after an acquisition. Despite the advantages of the (DD) strategy, a smaller concern remains that unobservable characteristics will change affecting dividend paying firms and share repurchasing firms differentially. This concern is a violation of the common trend assumption, the key identifying assumption in difference-in-difference empirical strategies. Formally, the common trend assumption states that the effect of changes in unobservable variables will have the same effect on the two subgroups. This assumption is empirically testable by regressing,

$$CAR_{i,t} = \beta_0 + \sum_k \delta_k \lambda_k + \sum_k \gamma_k \lambda_k \mathrm{x} d_{D,i} \tag{15}$$

where long-run stock returns are given by our twenty-four month CAR estimates, λ_k are year fixed effects, and $d_{D,i}$ is an indicator variable of whether the firm pays out dividends or not.[12] The coefficients on the year fixed effects, δ_k, estimate the common trend. The coefficients on the year fixed effects interacted with the dividend firm indicator, γ_k, estimate the difference in trend between firms that pay a dividend and firms that repurchase shares. Therefore, a failure to reject the joint test that all γ_k are zero is a failure to reject the common trend assumption.

Table 2 reports this empirical test of the common trend assumption for three subgroups: firms that pay dividends, firms with a low percentage of institutional ownership, and firms that pay dividends and have a low percentage of institutional ownership. Column (1) reports that none of the coefficients of the year fixed effects interacted with the dividend firm indicator are statistically significant. The joint F-test fails to reject the null, and therefore, the common trend assumption, at the fifty percent level. Similarly, Columns (2) and (3) report that none of the coefficients on the other subgroups are statistically significant and the F-test fails to reject at the thirty-seven and forty-six percent levels respectively.

[12]This regression is run only over firms that pay dividends or share repurchases. Firms that do not distribute retained earnings or distribute using both dividends and share repurchases are excluded.

Table 2: Common Trend Assumption

	Dividend	Low Inst.	Div x Low Inst.
1998	-4.289	-3.046	-2.508
	(3.036)	(1.868)	(3.220)
1999	-7.901	-6.658	-6.120
	(5.032)	(4.424)	(5.147)
2000	2.790	3.807	6.095
	(5.911)	(5.623)	(6.110)
2001	-1.720	-0.650	1.731
	(4.625)	(4.095)	(5.018)
2002	-0.193	0.962	3.325
	(5.314)	(4.862)	(5.880)
R^2	0.029	0.030	0.033
Observations	2223	2223	2223
F-stat	0.87 d.f. 5	1.09 d.f. 5	0.94 d.f. 5
p-value	0.500	0.365	0.457

Notes: The estimation equation is run over the set of firms paying a dividend paying or repurchasing shares, excluding firms that do neither or both: $CAR_{i,t} = \beta_0 + \sum_k \delta_k \lambda_k + \sum_k \gamma_k \lambda_k \text{x} d_{D,i}$.

4 Empirical Design and Results

The goal of the empirical strategy is to separate the effect of the dividend tax on the quality of acquisitions undertaken from other unobserved factors that may have changed between 1998 and 2010.[13] We use variation in the dividend tax rate, across firm payout strategies, and across firm percentage institutional ownership to identify the effect. The Jobs and Growth Tax Relief Reconciliation Act of 2003 cut the tax rate on qualified dividends from the ordinary income tax rates to the long-term capital gains tax rate.[14] For top income tax

[13]The capital gains tax also decreased in 2003, but by only 5 percent. Using firms that repurchase shares, that may be affected by the capital gains tax reform, will bias our estimates toward zero.

[14]A dividend is a qualified dividend if 1) it was paid after December 31, 2002, 2) paid by a U.S. corporation or other entity that qualifies for benefits under U.S. tax laws and treaties and 3) the stock most have been held 60 days during the 121-day period that begins 60 days before the ex-dividend date.

payers this effectively lowered the tax rate on dividends from 38.6 percent to 15 percent.[15] This tax rate change gives us variation in the dividend tax rate across the years in our sample. The dividend tax rate change may not be the only relevant factor to acquisition policy that changed in our sample years. For example, differences in market conditions before and after 2003 could cause differences in performance. To control for these other time-varying factors we use variation across firms in payout strategy. Specifically, we use firms that distribute profits through share repurchases as a benchmark to control for all other time-varying factors. Firms that repurchase shares are a good benchmark because they are similar in terms of observable characteristics, as Table 1 demonstrates, and firms that repurchase shares are unaffected by the dividend tax rate change.

A difference-in-difference (DD) specification is used to exploit the variation in the dividend tax rate and across firm payout strategies. The goal is to isolate the effect of taxation on the performance of firms after an acquisition by comparing differences in performance before and after the tax rate change for firms affected by the tax rate change relative to firms unaffected by the tax rate change. To isolate this effect, we limit our sample in two ways. First, we use only firms that regularly pay a dividend and firms that regularly repurchase shares, excluding firms that do both or neither. Second, we limit the sample to firms with a low percentage of institutional ownership because many institutional ownership are exempt from the dividend tax and therefore, firms with a high percentage of institutional ownership will be unaffected by the change: Lakonishok and Vermaelen [1986], Michaely [1991], Robin [1991]. The dependent variable of interest is the cumulative abnormal returns twenty-four months after an acquisition, though other months are examined in appendix 7.3. The specification includes dummy variables for being a dividend paying firm, d_D, and being in the high dividend tax years 1998-2002, d_0, and the interaction term, which is the (DD) estimate. Finally, the specification includes time fixed effects λ_t and merger specific covariates $\mathbf{z}$ as given by

$$R = \beta_0 + \beta_1 d_D + \beta_2 d_0 + \delta_1 d_D d_0 + \lambda_t + \boldsymbol{z}\boldsymbol{\gamma} + u \tag{16}$$

where the subscripts for time, t, and merger, i, have been suppressed. The coefficient of interest is δ_1 which captures the effect of high dividend tax rates on the stock return

Dividend Tax Rate Change

Income Level 2002	Tax Rate 2002	Income Level 2003	Tax Rate 2003
0 - $6,000	10	0 - $7,000	5
$6,000 - $27,950	15	$7,000 - $28,400	5
$27,950 - $67,700	27	$28,400 - $68,800	15
$67,700 - $141,250	30	$68,800 - $143,500	15
$141,250 - $307,050	35	$143,500 - $311,950	15
over $307,050	38.6	over $311,950	15

[15]

performance of dividend paying firms after an acquisition.

Alternatively, we can estimate a difference-in-difference-in-difference (DDD) specification using the variation across firms with high and low institutional ownership. This specification allows us to control for time-varying factors that affect all dividend paying firms and all firms with a low percentage of institutional ownership. However, we are unable to control for unobserved time varying factors which differentially affect dividend paying firms with a low percentage of institutional ownership. Though, we fail to reject the common trend assumption, reported in Table 2, which dampens some of this concern.

As in the (DD) specification, the dependent variable of interest in the (DDD) specification is the cumulative abnormal returns twenty-four months after an acquisition. In addition to the indicator variables for being a dividend paying firm, d_D, and being in the high dividend tax years 1998-2002, d_0, from the (DD) specification we now include an indicator variable for being a firm with lower than 50 percent institutional ownership, d_L, and the additional interaction terms, $d_0 d_L$, $d_D d_L$, and the (DDD) estimate $d_0 d_D d_L$.[16] Once again, the specification includes time fixed effects λ_t and merger specific covariates $\mathbf{z}$ as given by

$$R = \beta_0 + \beta_1 d_D + \beta_2 d_0 + \beta_3 d_L + \delta_1 d_D d_0 + \delta_2 d_0 d_L + \delta_3 d_D d_L + \delta_4 d_0 d_D d_L + \lambda_t + \boldsymbol{z\gamma} + u \quad (17)$$

where the subscripts for time, t, and merger, i, have been suppressed. The coefficient of interest δ_4 captures the policy effect of higher dividend taxation on a firm's long-run stock returns after an acquisition.

4.1 Non-Parametric Difference-in-Difference Analysis

Table 3 calculates the unconditional means of the twenty-four month CARs for each of the eight subgroups. The third row calculates the differences between the high-tax and low-tax period for firms paying dividends and repurchasing shares with high and low institutional ownership. Row (3) Column (1) reports that dividend paying firms with low institutional ownership performed 5.47 percent worse in the high tax period than in the low tax period. This simple difference conflates the tax rate change and all other unobserved changes, creating the need for a benchmark to compare it to. Column (2) in Row (3) provides this benchmark, reporting that firms that repurchase shares with low institutional ownership performed 3.8 percent better in the high tax period than in the low tax period. This difference implies market conditions for acquisitions were better in the high-tax period, 1998-2002,

[16] The results are robust to specifying the percentage of institutional ownership as a continuous variable instead of a dummy variable.

than in the low-tax period, 2003-2010. Thus in the counterfactual case without a tax rate change we would have expected dividend paying firms to perform better in the high-tax period, contrary to what we observe. Taken together these two differences imply that the higher taxes in the high-tax period led dividend paying firms with low institutional ownership to perform 9.27 percent worse than they would have with the lower tax rates. This result is the non-parametric difference-in-difference (DD) estimate that corresponds to the parametric estimates in section 4.2 (Table 4 Column (1)).

Table 3: Returns Means

	Low Institution Ownership		High Institution Ownership		
	Dividends	Share Purchases	Dividends	Share Purchases	
High-Tax	-21.36	-21.46	-19.84	-16.91	
Low-Tax	-15.89	-25.26	-16.84	-16.35	
Differences	-5.47	3.8	-3	-.56	-6.83

Notes: This table reports the mean levels of the cumulative abnormal returns 24 months after an acquisition. The top row reports $\bar{R}_{0,D,L}, \bar{R}_{0,S,L}, \bar{R}_{0,D,H}, \bar{R}_{0,S,H}$ and the second row reports $\bar{R}_{1,D,L}, \bar{R}_{1,S,L}, \bar{R}_{1,D,H}, \bar{R}_{1,S,H}$.

Columns (3) and (4) in Row (3) repeat this exercise for firms with high levels of institutional ownership. These differences report dividend paying firms with high institutional ownership performed 2.44 percent worse in the high-tax period than in the low-tax period, relative to firm's that repurchase shares. This 2.44 percent difference may be due to time-varying differential effects on firms paying a dividend and firms repurchasing shares. If this is the case then we want to control for this difference and subtract it off of the policy effect found for firms with a low percentage of institutional ownership. However, this 2.44 percent difference may be due to the fact that firms with a high percentage of institutional share holders are still affected by the dividend tax rate, but to a lesser extent. In this case we would not want to subtract this effect off of the policy effect found for firms with a low percentage of institutional ownership because it will bias the policy effect toward zero. The difference between the tax effects for low institutional ownership firms and high institutional ownership firms gives the result that dividend paying firms with low institutional ownership performed 6.83 percent worse because of the high tax rates. This result is the non-parametric triple difference estimate that corresponds to the parametric triple difference estimates reported in section 4.2 (Table 5 row 7 Column (2)).

4.2 Parametric Difference-in-Difference Analysis

This section reports the parametric difference-in-difference (DD) and difference-in-difference-in-difference (DDD) estimates from the empirical models in equations (16) and (17). These estimates demonstrate higher dividend tax rates in the late 1990s and early 2000s led firms to perform 8 to 10 percent worse two years after an acquisition relative to what their returns would have been with the lower dividend tax rates after 2003. This evidence is consistent with the the model in section 2 which demonstrates dividend paying firms have an incentive, proportional to the dividend tax rate, to make acquisitions with negative net synergies. Section 5 runs several alternative specifications to rule out other possible models, including payment method and firm monitoring, that could explain this result. Together these sections provide strong evidence in support of the theoretical model and the negative impact of dividend taxation on merger and acquisition behavior.

Table 4 reports the parametric difference-in-difference (DD) estimates for the sub-sample of firms with a low percentage of institutional ownership. Limiting the sample to firms with a low percentage of institutional ownership concentrates the analysis on firms most likely to be affected by the dividend tax rate change, since many institutional shareholders are tax-exempt.[17] In all three specifications the difference-in-difference estimator is statistically and economically significant. Specification (2) controls for firm characteristics such as sales, marginal q, retained earnings, and total assets and specification (3) includes year fixed effects. The point estimates are similar across these specifications, estimating post-merger long-run returns were 8 to 10 percent lower because of higher dividend taxation.

[17]The (DDD) specification uses the full sample of firms that pay dividends or repurchase shares and includes a third difference for firms with low and high percentages of institutional ownership. The (DD) estimates are robust to using the full sample and controlling for the level of institutional ownership.

Table 4: Difference-in-Difference Estimates: Low Institutional Firms Sub-Sample

	(1)	(2)	(3)
Dividend Firms β_1	8.387***	5.559*	8.826**
	(2.143)	(3.167)	(3.097)
High-Tax β_2	3.286	0.895	0.736
	(2.146)	(1.850)	(4.350)
Dividend x High-Tax δ_1	-9.271**	-8.004*	-9.911**
	(4.922)	(4.133)	(4.291)
Firm Controls		✓	✓
Year Fixed Effects			✓
R^2	0.010	0.076	0.108
Observations	927	798	798

Notes: This table reports difference-in-difference estimates. The dependent variable is the cumulative abnormal return 24 months after an acquisition. Robust standard errors are in parentheses. The results remain statistically significant clustering at the group level (e.g. firms that pay dividends and are in the high-tax years 1998-2002) and clustering by firm.

Table 5 reports the parametric difference-in-difference-in-difference (DDD) estimates using the full sample of mergers by firms paying dividends or repurchasing shares. This specification is more flexible than the (DD) estimates and is able to further control for differences between firms that repurchase shares and firms that pay a dividend by using the difference between them in the sub-samples of firms with low and high percentages of institutional ownership. However, to the extent that firms with high percentages of institutional share holders are affected by the dividend tax rate change then the (DDD) estimates will be an underestimate of the true policy effect of higher dividend taxation. Specification (1) runs the (DD) specification using the full sample of mergers by firms paying dividends or repurchasing shares. The (DD) estimate is statistically significant and has a smaller point estimate, -3.773, than in Table 4 because it conflates firms with tax-exempt shareholders with firms with taxable shareholders. The (DDD) estimate, given by δ_4 is statistically and economically significant across all specifications. Specification (2) does not control for firm controls or year fixed effects and therefore, exactly replicates the non-parametric estimates in Table 3. Controlling for year fixed effects and firm controls increases the magnitude of the point estimate from -6.836 to -10.704. The point estimates are similar across these specifications and similar to the (DD) results, estimating post-merger long-run returns were

8 to 10 percent lower because of higher dividend taxation.

Table 5: Difference-in-Difference-in-Difference Estimates

	(1)	(2)	(3)	(4)	(5)
Dividend Firms β_1	2.319**	-0.495	-0.565**	-1.354***	-2.178***
	(1.093)	(1.172)	(0.172)	(0.010)	(0.233)
High-Tax β_2	-0.542	-0.558	-13.398***	-0.922***	-8.663***
	(0.948)	(1.118)	(1.202)	(0.109)	(1.943)
Low Inst. Ownership β_3		-8.915***	-10.566***	-6.453***	-8.101***
		(1.470)	(0.481)	(0.926)	(0.533)
Dividend x High-Tax δ_1	-3.773**	-2.437	-0.548	-1.162*	1.501**
	(1.922)	(2.717)	(0.491)	(0.529)	(0.609)
High-Tax x Low Inst. δ_2		4.366**	7.328***	3.591***	6.359***
		(2.154)	(0.583)	(0.564)	(0.511)
Dividend x Low Inst. δ_3		9.868***	11.520***	8.143***	11.281***
		(2.432)	(0.449)	(0.390)	(0.391)
Div x H-T x Low Inst. δ_4		-6.836*	-7.954***	-8.282***	-10.704***
		(4.033)	(0.698)	(0.058)	(0.827)
Year Fixed Effects			✓		✓
Firm Controls				✓	✓
R^2	0.089	0.082	0.125	0.061	0.129
Observations	2223	2223	2223	2223	2223

Notes: This table reports difference-in-difference-in-difference (DDD) estimates. The dependent variable is the cumulative abnormal return 24 months after an acquisition. Robust-clustered standard errors are in parentheses. The standard errors are clustered at the group level (e.g. firms that pay dividends, have low institutional ownership, and are in the high-tax years 1998-2002). The results remain statistically significant without clustered standard errors.

4.3 Graphical Difference-in-Difference Analysis

Figure 1 demonstrates the (DDD) estimation graphically, expanding the analysis from twenty-four months after an acquisition to any month between one and thirty. Each sub-figure graphs the average long-run returns before and after the tax rate change. Firms that repurchase shares and have a high percentage of institutional ownership are shown in Figure 2(a), and with a low percentage of institutional ownership in Figure 2(b). Below that firms that pay dividends and have a high percentage of institutional ownership are shown in Figure 2(c), and with a low percentage of institutional ownership in Figure 2(d).

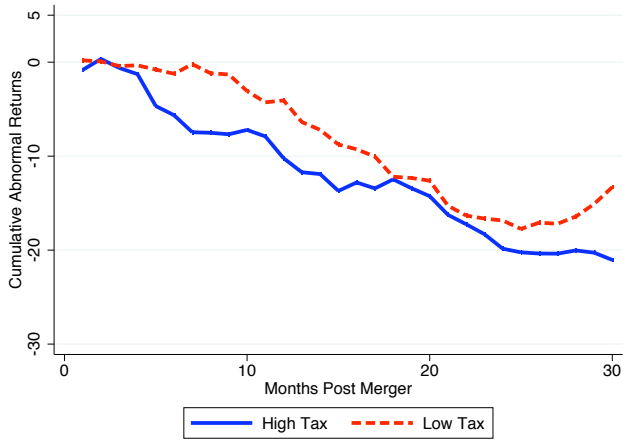
The theory suggests firms that distribute retained earnings through share repurchases should not be affected by the dividend tax rate change. To the extent that there are no other unobserved changes between the high-tax and low-tax periods, this suggests the long-run returns for these firms should look similar before and after the dividend tax rate change. For almost all months between one and thirty months after an acquisition the long-run returns look very similar before and after the dividend tax reform, denoted by the similarities in the blue solid (high-tax period) and red dashed (low-tax period) lines in Figures 2(a) and 2(b). This suggests unobservable changes between the high-tax and low-tax periods are minimal. Similarly, the theory suggest firms that have shareholders that are tax-exempt from the dividend tax, measured here by the percentage institutional ownership, should also not be affected by the dividend tax rate change. Therefore, to the extent that firms with a high percentage of institutional ownership are tax-exempt, the blue solid line and red dashed lines in Figure 2(c) should also be similar, absent any other unobservable changes. In contrast, to the previous three figures, the theory suggests firms paying a dividend with taxable shareholders should be affected by the dividend tax reform and cause a difference in the high-tax period (blue solid) line and low-tax period (red dashed) line in Figure 2(d).

These figures support the empirical analysis demonstrating the returns in the high-tax and low-tax periods are similar for firms repurchasing shares and firms paying a dividend but with tax-exempt shareholders while the low-tax period led to higher returns for firms paying a dividend with non-tax-exempt shareholders. Formally, the policy effect is captured by subtracting the difference in lines in Figures 2(a), 2(b), and 2(c) from the difference in lines in Figure 2(d). These graphs suggest the empirical results in Table 5 are robust to using any month between one and thirty months after an acquisition as the dependent variable.

Figure 1: Difference High-Tax and Low-Tax Time Periods

(a) Share Repurchases and High Institutional Ownership

(b) Share Repurchases and Low Institutional Ownership



(c) Dividend and High Institutional Ownership

(d) Dividend and Low Institutional Ownership

5 Durability Of Empirical Results

This section presents several robustness tests which reinforce the baseline estimates in Tables 4 and 5 and strengthen the connection between the empirical results and the theoretical model. For these robustness tests, we rely on additional data from Thomson and Reuters' Securities Data Company (SDC) to expand the time series to the years 1992-2012. The SDC data allows us to examine smaller subsets, such as non-cash mergers and mergers done by firms not paying a dividend, because the data set more than doubles the initial set of mergers considered from just under 7,000 to just over 20,000. While the SDC data increases the number of observations, it does so by using data up to 10 years on both sides of the tax rate change in 2003. Using data further from the tax rate change in 2003 increases the possibility of conflating effects, for example other tax rate changes in the 1990s. On whole, these robustness checks provide further evidence that dividend taxes distort merger and acquisition behavior and that this tax mechanism may be responsible for the post-merger performance puzzle.

First, we present baseline specifications (as in Table 5) using the expanded data set. The results are reported in Table 6. Using this data two additional tests are run and reported in the appendix. Appendix 7.3 tests the robustness of our baseline results to considering different time horizons. Appendix 7.4 tests whether the results are biased due to mergers that occur in quick succession. These tests suggest that the results are robust to different time horizons and the bias from mergers occurring in succession is small. Furthermore, these tests suggest the baseline estimates may be underestimates of the true policy effect of higher dividend taxation.

Second, we examine two alternative models which could generate our baseline results. Regression results are presented in Tables 7 and 8. We find no evidence that the method of payment nor shareholder monitoring models are able to explain the change in post-merger long-run returns reported in this paper.

5.1 Parametric Difference-in-Difference Analysis Using SDC Data

Table 6 presents baseline specifications as in Tables 4 and 5 using the expanded data set. Column (1) reproduces the (DD) results reported in Column (3) of Table 4. The magnitudes of the (DD) estimates, -9.911 and -9.289, are close in both data sets, though the estimate in Table 6 is not statistically significant. The (DDD) estimates in Column (5) of both Table 5 and 6 are also similar, -10.704 and -14.872, and statistically significant. The similarities between the estimates using the Zephyr and SDC data sets provides some additional confidence in these estimates and lessens the concern of possible biases in the SDC due to the longer time series.

Table 6: Robustness: New Data Set 1992-2012

	(1)	(2)	(3)	(4)	(5)
Dividend Firms β_1	10.844	1.312	1.320***	-0.003	-0.083
	(6.900)	(1.071)	(0.088)	(0.060)	(0.062)
High-Tax β_2	3.954	-11.942***	-13.261***	-11.371***	-14.441***
	(11.548)	(1.751)	(1.462)	(0.071)	(1.088)
Low Inst. Ownership β_3		-4.623	-5.594***	-9.998***	-11.075***
		(4.650)	(0.148)	(0.091)	(0.550)
Dividend x High-Tax δ_1	-9.289	6.461***	6.461***	6.752***	7.117***
	(7.689)	(1.816)	(0.172)	(0.038)	(0.220)
High-Tax x Low Inst. δ_2		0.711	1.775***	11.358***	12.295***
		(5.425)	(0.213)	(0.006)	(0.595)
Dividend x Low Inst.δ_3		4.328	5.589***	8.879***	10.263***
		(4.714)	(0.134)	(0.015)	(0.552)
Div x H-T x Low Inst. δ_4		-2.692	-3.943***	-13.657***	-14.872***
		(5.510)	(0.154)	(0.073)	(0.645)
Year Fixed Effects	✓		✓		✓
Firm Controls	✓			✓	✓
Low Inst. Subset	✓				
R^2	0.094	0.067	0.128	0.056	0.131
Observations	665	4207	4207	4207	4207

Notes: This table repeats the estimates in Table 5 with mergers data from Thomson and Reuters Securities Data Company (SDC) years 1992-2012. The dependent variable is the cumulative abnormal return 24 months after an acquisition. Robust standard errors are in parenthesis. Standard errors are clustered at the group level for the (DDD) specification. Statistical significance denoted, 10 percent *, 5 percent **, and 1 percent ***. The results are robust without clustering the standard errors and clustering at the firm level.

5.2 Method of Payment Alternative Explanation

This subsection tests whether the merger method of payment could explain the (DD) and (DDD) results. The finance literature, notably Travlos [1987], Wansley et al. [1983] and Martin [1996], have found a link between the method of payment and abnormal returns: cash payments are correlated with larger negative abnormal returns. The following analysis tests whether the negative abnormal returns observed in this paper are due to tax effects or effects associated with cash financing.

To separate these effects we run two different specifications, the first reported in Columns (2) and (5) and the second in Columns (3) and (6) of Table 7. The first specification runs the benchmark analysis on the subset of mergers that were financed with less than 50 percent cash payments. Here we find the magnitude of the (DD) and (DDD) estimates increase in this subset, strengthening the empirical results and contrary to what we would expect if the empirical results were due to method of payment effects.

The second specification replaces the indicator variable for firms that pay a dividend with an indicator variable for 100 percent cash financed mergers, and then runs the analysis on the subset of mergers by non-dividend paying firms. If method of payment effects could explain the empirical results we would expect to see the (DD) and (DDD) results to look similar in this specification. In contrast, estimates of $\tilde{\delta}_1$ and $\tilde{\delta}_4$ are the opposite sign and statistically significant at the 1 percent level. These results corroborate the tax mechanism in explaining the empirical results and furthermore, suggest the tax effect may be the mechanism driving the method of payment correlation noted in the finance literature.

Table 7: Alternative Explanation: Method of Payment

	(1)	(2)	(3)	(4)	(5)	(6)
Dividend Firms β_1	10.844	15.779		-0.083	1.704	
	(6.900)	(10.687)		(0.062)	(2.645)	
Cash Payments $\tilde{\beta}_1$			-10.980***			3.150***
			(0.000)			(0.469)
High-Tax β_2	3.954	4.789	-6.896***	-14.441***	-15.618***	-4.867
	(11.548)	(14.397)	(0.000)	(1.088)	(5.669)	(2.748)
Low Inst. Ownership β_3				-11.075***	-14.944	-3.194***
				(0.550)	(10.061)	(0.407)
Dividend x High-Tax δ_1	-9.289	-20.614*		7.117***	3.287	
	(7.689)	(12.064)		(0.220)	(4.259)	
Cash x High-Tax $\tilde{\delta}_1$			14.154***			-4.168***
			(0.001)			(0.536)
High-Tax x Low Inst. δ_2				12.295***	20.309*	-1.639**
				(0.595)	(12.005)	(0.457)
Dividend x Low Inst.δ_3				10.263***	11.976	
				(0.552)	(10.205)	
Cash x Low Inst. $\tilde{\delta}_3$						-2.323*
						(0.932)
Div x H-T x Low Inst. δ_4				-14.872***	-21.408*	
				(0.645)	(12.225)	
Cash x H-T x Low Inst. $\tilde{\delta}_4$						7.913***
						(1.215)
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
$< 50\%$ Cash Mergers		✓			✓	
Non Dividend Firm Mergers			✓			✓
R^2	0.094	0.040	0.018	0.131	0.126	0.185
Observations	665	233	90	4207	2876	3419

Notes: This table reports the estimates limiting the data to mergers paid with less than 50 percent cash, (2) and (5), mergers made by firms not paying a dividend, (3) and (6), and the benchmark specifications with all of the data, (1) and (4). Specifications in (1), (2), and (3) limit the data to firms with low percentages of institutional ownership. The dependent variable is the cumulative abnormal return 24 months after an acquisition. Robust standard errors are in parenthesis. Standard errors are clustered at the group level for the (DDD) specification. Statistical significance denoted, 10 percent *, 5 percent **, and 1 percent ***. The results are robust without clustering the standard errors and clustering at the firm level. Data comes from Thomson and Reuters Securities Data Company (SDC) and Compustat, years 1992-2012.

5.3 Shareholder Monitoring Alternative Explanation

Chetty and Saez [2010] provide a second alternative hypothesis that may be able to explain the empirical results presented in this paper based on firm monitoring by shareholders. The shareholder monitoring model predicts that when the dividend tax rate decreases, shareholders will increase their monitoring of corporate managers. This increase in monitoring could generate the increased acquisition performance that we observe after the 2003 dividend tax reform if our measure of tax status, low percentage of institutional ownership, is also measuring monitoring within the firm. According to this shareholder monitoring model, the increase in monitoring and by extension increase in acquisition performance should be larger when the firm has large block shareholders with strong incentives to monitor.

To test the ability of the monitoring model to explain the DD and DDD empirical results we observe, we replace the indicator variable for low institutional ownership with an indicator variable for firms with a low percentage of stock owned by the top five shareholders and the top shareholders generally. The percentage of the firm owned by a few shareholders is a measure of the monitoring within a firm because a shareholder with a larger percent of the firm's stock will have a greater incentive to monitor the firm. In contrast, this is not a good measure for tax status as the top shareholders of a firm may be taxable.

If monitoring is able to explain the empirical results we would expect the specifications with the measures of percentage owned by top shareholders to be similar to the results in the benchmark specifications. In contrast, the (DD) estimates in Columns (2) and (3) and the (DDD) estimates in Columns (5) and (6) have the opposite sign as the benchmark model and are not statistically significant. These empirical results are inconsistent with the shareholder monitoring model but are consistent with the tax mechanism.

Table 8: Alternative Method: Firm Monitoring

	(1)	(2)	(3)	(4)	(5)	(6)
Dividend Firms β_1	10.844	3.028	2.502	-0.083	1.276***	1.732***
	(6.900)	(2.786)	(3.397)	(0.062)	(0.126)	(0.082)
High-Tax β_2	3.954	-14.272***	-12.117**	-14.441***	-12.742***	-12.692***
	(11.548)	(5.104)	(6.007)	(1.088)	(1.458)	(1.484)
Low Inst. Ownership β_3				-11.075***		
				(0.550)		
Top Inst. $\tilde{\beta}_3$					-0.680	
					(2.418)	
High Inst. $\tilde{\beta}_3$						-0.322
						(3.169)
Dividend x High-Tax δ_1	-9.289	2.581	2.842	7.117***	5.958***	5.597***
	(7.689)	(4.057)	(4.443)	(0.220)	(0.206)	(0.150)
High-Tax x Low Inst. δ_2				12.295***		
				(0.595)		
Dividend x Top $\tilde{\delta}_3$					1.802	
					(2.503)	
Dividend x High $\tilde{\delta}_3$						0.840
						(3.188)
Div x H-T x Low Inst. δ_4				-14.872***		
				(0.645)		
Div x H-T x Top $\tilde{\delta}_4$					0.313	
					(2.917)	
Div x H-T x High $\tilde{\delta}_4$						1.157
						(3.582)
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
Low % Top 5 Sub-Sample		✓				
Low % Top Holders Sub-Sample			✓			
R^2	0.094	0.124	0.118	0.131	0.124	0.124
Observations	665	652	638	4207	4207	4207

Notes: This table reports the benchmark specifications in (1) and (4). Specification (2) limits the data to firms with a low percentage of stock held by the top five shareholders. Specification (3) limits the data to firms with a low percentage of stock held by top shareholders generally. Specifications (5) and (6) replace the indicator variable on low institutional ownership (used to control for the tax status of the shareholders) with an indicator variable for firms with a low percentage held by the top five shareholders, (5) , and an indicator for firms with a low percentage held by top shareholders generally, (6), (used to control for the level of monitoring in the firm). The dependent variable is the cumulative abnormal return 24 months after an acquisition. Robust standard errors are in parenthesis. Standard errors are clustered at the group level for the (DDD) specification. Statistical significance denoted, 10 percent *, 5 percent **, and 1 percent ***. The results are robust without clustering the standard errors and clustering at the firm level. Data comes from Thomson and Reuters Securities Data Company (SDC) and Compustat, years 1992- 2012.

6 Conclusion

Mergers and acquisitions continue to be an important way for economic activity to be reorganized; as the first half of 2013 saw $437 billion in merger and acquisition transactions. Our model demonstrates a potentially large distortion in merger and acquisition behavior caused by the dividend tax. Empirically we find long-run returns are 8 to 10 percent higher as a result of lower dividend taxation that resulted from the 2003 dividend tax reform in the US. These findings suggest the deadweight loss caused from dividend taxation could be a significant drag on the economy.

The theoretical model provides explanations for two empirical puzzles; the increase in dividend payments following the 2003 dividend tax reform in the US [Chetty and Saez, 2005] and the post-merger performance puzzle [Agrawal and Jaffee, 2000]. While there have been other proposed mechanisms to explain these empirical puzzles ours is the first to do so with firms that maximize shareholder value in a competitive equilibrium.

The intuition for the mechanism in the model follows from the insight that mergers and acquisitions act as a transfer from firms to shareholders. Without dividend taxation the model demonstrates firms only make an acquisition when the expected net synergies from the acquisition are positive. In contrast, with dividend taxation the model demonstrates firms have an incentive to use acquisitions as an alternative channel to transfer retained earnings to shareholders, avoiding the dividend tax. This leads firms that maximize shareholder value to sometimes make acquisitions with negative net synergies to gain the tax benefits from making transfers through acquisitions.

Through a series of empirical tests we find strong evidence that dividend taxation is distorting merger and acquisition behavior. The empirical results collaborate the theoretical claim that the distortion due to the dividend tax is economically large. This suggests that the tax mechanism we propose is large enough to explain both the post-merger performance puzzle and the increase in dividend payments after the dividend tax reform of 2003. Furthermore, these results on the distortion of the dividend tax is important for policy makers to consider when determining if and how to tax dividends.

References

Agrawal, Anup and G.N. Madelker, "Large Sheholders and the Monitoring of Managers: The Case of Anti-takeover Charter Amendments," *Journal of Financial and Quantitative Analysis*, 1990, *25(2)*, 143–161.

_ **and J. Jaffee**, "The Post-Merger Performance Puzzle," in Cary Cooper and Alan Gregory, eds., *Advances in Mergers and Acquistions*, Elsevier Science, 2000, pp. 7–41.

_ **,** _ **, and G.N. Madelker**, "The Post-Merger Performance of Acquiring Firms: A Re-examination of an Anomoly," *Journal of Finance*, 1992, *47*, 1605–1621.

Auerbach, Alan J., "Wealth Maximiation and the Cost of Capital," *Quarterly Journal of Economics*, 1979, *93(3)*, 433–436.

_ **and David Reishus**, "The Impact of Taxation on Mergers and Acquisitions," in Alan J. Auebach, ed., *Mergers and Acquistions*, UMI, 1987, pp. 69–82.

_ **and** _ , "The Effects of Taxation on the Merger Decision," in J. Auerbach, Alan, ed., *Corporate Takovers: Causes and Consequences*, UMI, 1988, pp. 157–190.

Bessembinder, Hendrik and Feng Zhang, "Firm characteristics and long-run stock returns after corporate events," *Journal of Financial Economics*, 2013.

Chetty, Raj and Emmanuel Saez, "Dividend Taxes and Corporate Behavior: Evidence from the 2003 Dividend Tax Cut," *Quarterly Journal of Economics*, 2005, *120(3)*, 791–833.

_ **and** _ , "An Agency Theory of Dividend Taxation," *American Economic Journal: Economic Policy*, 2010, *2(3)*, 1–21.

Fama, Eugene F. and Kenneth French, "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 1992, *47*, 427–465.

_ **and** _ , "Common Risk Factos in the Returns on Stock and Bonds," *Journal of Financial Economics*, 1993, *33*, 3–56.

Franks, Julian, Robert Harris, and Sheridan Titman, "The Postmerger Share-Price Performance of Acquring Firms," *Journal of Financial Economics*, 1991, *29*, 81–96.

Gregory, Alan, "An Examination of the Long Run Performance of UK Acquiring Firms," *Journal of Business Finance and Accounting*, 1997, *24*, 971–1002.

Lakonishok, Josef and Theo Vermaelen, "Tax-induced trading around ex-dividend days," *Journal of Financial Economics*, 1986, *16* (3), 287–319.

Loughran, Tim and Anand M. Vijh, "Do Long-Term Shareholders Benefit from Corporate Acquisitions," *Journal of Finance*, 1997, *52*, 1765–1790.

Majluf, Nicholas S. Myers Stewart C.and, "Corporate Financing and Investment Decisions when Firms have Information that Investors Do Not Have," *Journal of Financial Economics*, 1984, *13*, 187–221.

Martin, Kenneth J, "The method of payment in corporate acquisitions, investment opportunities, and management ownership," *The Journal of Finance*, 1996, *51* (4), 1227–1246.

Michaely, Roni, "Ex-Dividend Day Stock Price Behavior: The Case of the 1986 Tax Reform Act*," *The Journal of Finance*, 1991, *46* (3), 845–859.

Rau, Raghavendra P. and Theo Vermaelen, "Glamour, value and the post-acquistions performance of acquiring firms," *Journal of Financial Economics*, 1998, *49(2)*, 223–253.

Robin, Ashok J, "The impact of the 1986 tax reform act on ex-dividend day returns," *Financial Management*, 1991, pp. 60–70.

Sinn, Hans-Werner, "Taxation and the cost of capital: The" old" view, the" new" view, and another view," in "Tax Policy and the Economy, Volume 5," The MIT Press, 1991, pp. 25–54.

Stephens, Clifford P and Michael S Weisbach, "Actual share reacquisitions in open-market repurchase programs," *The Journal of Finance*, 1998, *53* (1), 313–333.

Travlos, Nickolaos G, "Corporate takeover bids, methods of payment, and bidding firms' stock returns," *The Journal of Finance*, 1987, *42* (4), 943–963.

Wansley, James W, William R Lane, and Ho C Yang, "Abnormal returns to acquired firms by type of acquisition and method of payment," *Financial management*, 1983, pp. 16–22.

7 APPENDIX

7.1 Dividend Taxation and The Post-Merger Performance Puzzle

This section considers the theoretical implications of the model pertaining to the post-merger performance puzzle. The previous research has documented that on average firms' abnormal returns post-merger are negative a phenomenon which seems unexplainable in equilibrium. Several behavioral explanations have been suggested as explanations for this empirical puzzle. This section demonstrates the model presented above provides a possible explanation for this empirical puzzle consistent with a firm maximizing shareholder value. In addition, the model produces some empirically testable implications for the post-merger performance puzzle separate from the behavioral explanations previously suggested.

Intuitively, the minimum pretax return required to make an acquisition profitable depends on where the investment comes. For example, if the merger is financed by the firm retaining additional earnings, rather than distributing them as dividends, then the marginal dollar for investment through mergers comes from dividend payments. From the shareholder's point of view the firm could distribute earnings to them, triggering a tax liability, then with the post-personal-tax returns the shareholder could invest in and receive future dividends of another firm. Alternatively, the firm could acquire another firm allowing the shareholder to receive future dividends from the newly acquired firm without triggering the dividend tax. Therefore, dividend taxation creates a wedge encouraging the tax preferred acquisitions by decreasing the minimum pretax return required to make an acquisition profitable.

This intuition is formalized using the costate equation for the production technology and the first order condition with respect to the acceptable merger opportunities. The costate equation for the production technology,

$$-\dot{\lambda}_A = \lambda_k(1-\tau)Q(K_t + B_t) - \rho\lambda_A \tag{18}$$

relates the shadow price for production technology with the shadow price for capital within the firm and the time varying shadow price for production technology. This relationship is substituted into the first order condition with respect to the acceptable range of mergers to create a condition,

$$\lambda_K \left[\frac{(1-\tau)Q(K_t + B_t)}{\rho}\frac{\partial \tilde{z}}{\partial z^A} - c'(z^A)\right] + \frac{\dot{\lambda}_A}{\rho}\frac{\partial \tilde{z}}{\partial z^A} = 0, \tag{19}$$

that demonstrates firms should accept mergers until the marginal benefit to capital within

the firm and to the production technology is exhausted. This condition implies, with the assumption that in equilibrium $\dot{\lambda}_A > 0$, that in equilibrium firms sacrifice some capital for the benefits of incorporating the synergies into the production technology. Furthermore, firms with a low shadow price of capital within the firm, such as those paying a dividend, accept a wider range of mergers in equilibrium and therefore sacrifice more capital within the firm reiterating proposition 1 and its corollaries.

In equilibrium, firms sacrifice capital within the firm, depressing cumulative average returns, in return for increased production technology. This tradeoff maximizes shareholder value despite depressed cumulative average returns because it avoids the additional tax liability that would be triggered by the firm distributing profits through dividends. This behavior predicted by the model is empirically testable. Specifically, the model predicts after a merger the average cumulative abnormal returns of dividend paying firms will be lower than firms not paying a dividend. This follows from the differences in opportunity costs of capital discussed in corollary 2. In addition, the model predicts the average post-merger cumulative abnormal returns of dividend paying firms will increase in response to a dividend tax rate cut while the average post-merger cumulative abnormal returns of firms not paying a dividend should be unaffected, as demonstrated in corollary 3.

7.2 Appendix: Result 1

This appendix explicitly demonstrates how result 1 follows from lemmas 1-3.

Result 1 The set of acceptable mergers is increasing with respect to the dividend tax rate, for firms that pay dividends, and is independent of the dividend tax rate for firms not paying a dividend.

Result 1 follows directly from lemmas 1-3.

Result 1 claims $\frac{\partial z^A}{\partial \theta}$ is zero for firms not paying a dividend and the sign is positive for firms paying a dividend.

$$\frac{\partial z^A}{\partial \theta} = \frac{\partial z^A}{\partial \lambda_K} \frac{\partial \lambda_K}{\partial \theta}$$

For firms not paying a dividend we know,

$$= \underbrace{\frac{\partial z^A}{\partial \lambda_K}}_{\substack{<0 \\ \text{Lemma 3}}} \underbrace{\frac{\partial \lambda_K}{\partial \theta}}_{\substack{=0 \\ \text{Lemma 1}}} = 0$$

by lemma 3 and lemma 1. For dividend paying firms we know,

$$= \underbrace{\frac{\partial z^A}{\partial \lambda_K}}_{\substack{<0 \\ \text{Lemma 3}}} \underbrace{\frac{\partial \lambda_K}{\partial \theta}}_{\substack{<0 \\ \text{Lemma 2}}} > 0$$

by lemma 3 and lemma 2.

7.3 Generating Cumulative Abnormal Returns

We estimate the cumulative abnormal returns using the same method as Gregory [1997] and Agrawal et al. [1992] to match the post-merger performance puzzle literature. Our results are robust to alternative estimates of long-run stock returns, see Bessembinder and Zhang [2013]'s recent advance, but we report estimates following Gregory [1997]'s method for better comparability. The difference-in-difference empirical strategy we use controls for any bias from the cumulative abnormal return estimation because the procedure is the same for all subgroups in the sample. Therefore, each of the subgroups are evaluated on a level playing field.[18] The cumulative average return at time t is the multiplicative sum of abnormal returns from period 1 to t where the abnormal return for firm i at time t is calculated as the actual performance of the firm at time t minus the performance of the firm specific benchmark. The firm specific benchmark is the predicted counterfactual performance of firm i at time t in the hypothetical world where the firm did not merged.

In generating CARs for acquiring firms, the choice of benchmark generating model and the time period over which to estimate the coefficients of the model are important decisions that must be considered carefully. Following the comparative analysis by Gregory [1997] we use the Fama French Value-Weighted Three Factor Model described in Fama and French [1992, 1993]. The model is described in depth below. The monthly return of the firm, R_{it} minus the risk free rate of return, R_{ft}, is regressed on 1) the performance of the market R_{mt} minus the risk free rate, 2) the value weighted return on small firms minus the value weighted return on large firms, SMB, and 3) the value weighted return on high book-to-market value

[18] The difference-in-difference empirical strategy is not able to control for heterogeneous biases across subgroups but this concern seems small given the subgroups in this analysis.

firms minus the value weighted return of low book to market value firms.

Abnormal Returns Model: Fama and French [1992, 1993] Value Weighted Three Factor Model:

$$a_{it} = R_{it} - [R_{ft} + \hat{\beta}_i(R_{mt} - R_{ft}) + \hat{\gamma}_i(\text{SMB}) + \hat{\delta}_i(\text{HML})] \tag{20}$$

$$\text{CAR}_{it} = \prod_{j=1}^{t}(1 - (a_{ij}/100)) \tag{21}$$

$$\text{CAAR}_t = \frac{1}{n}\sum_i \text{CAR}_{it} \tag{22}$$

where

a_{it} =the abnormal return on company i in month t

R_{it} =return on company i in month t;

R_{ft} =risk-free (treasury bill) return in month t;

R_{mt} =return on market in month t;

SMB =the value-weighted return on small firms minus the value-weighted return on large firms;

HML =the value-weighted return on high BMV firms minus the value-weighted return on low BMV firms.

CAR_{it} =the cumulative abnormal return for firm i in month t

CAAR_t =the cumulative average abnormal return in month t

Gregory [1997] highlights the importance of the choice of time period over which to estimate the coefficients of the model, demonstrating a significant bias when using pre-event data. Thus, we use post-event data to estimate the model and generate abnormal returns, cumulative abnormal returns, and cumulative average abnormal returns. We generate CARs using up to 24 months of post return data (subject to availability). We limit our sample to firms that have at least 12 months of post-merger data available, excluding firms that do not continue to operate or were themselves purchased. This data limitation likely biases our CAR estimates towards insignificance. We are able to use acquisitions in the last year of our sample, 2010, by using stock price data through 2012 to estimate CARS.

Figure 1 depicts the CAAR estimates, from one month to thirty months after an acquisition, for the eight separate groups used in the difference-in-difference-in-difference empirical method we employ. Twenty-four months after an acquisition the CAAR estimates are approximately -18 percentage points below benchmark returns and are statistically significant

from zero at the 1 percent confidence level. These estimates are very similar to Gregory [1997] Fama French CAAR estimates despite a different set of mergers and time period. These estimates reinforce that the post-merger performance puzzle is still a prominent empirical phenomenon. As several authors have pointed out, the fact that this phenomenon continues to exist more than 20 years after the puzzle was first identified and after several explanations have been suggested means that economists still do not understand why firms continue to merge despite these failing market outcomes.

7.4 Appendix: Varying Performance Time Horizons

This subsection tests the robustness of our baseline results to considering different time horizons. The results in Table 9 reports the (DD) and (DDD) estimates using long-run returns 24, 18, and 12 months after the acquisition. The benchmark analysis, repeated in Columns (1) and (4), uses 24 months after an acquisition for comparability with much of the post-merger performance puzzle literature. For both the (DD) and (DDD) estimates the magnitudes are largest after 18 months, though the (DD) results are not statistically significant. This pattern suggests the negative synergies are fully realized somewhere between 12 and 18 months after the acquisition and implies that the baseline (DD) and (DDD) estimates, reported in Tables 4 and 5, may underestimate the effect of dividend taxation on acquisition performance.

Table 9: Robustness: Varying Performance Outcome Time Horizons

	24 Month	18 Month	12 Month	24 Month	18 Month	12 Month
	(1)	(2)	(3)	(4)	(5)	(6)
Dividend Firms β_1	7.223	14.849**	10.844	0.974***	-0.394*	-0.083
	(8.016)	(7.515)	(6.900)	(0.111)	(0.177)	(0.062)
High-Tax β_2	11.617	-2.691	3.954	-6.376**	-11.966***	-14.441***
	(14.584)	(14.125)	(11.548)	(2.397)	(1.431)	(1.088)
Low Inst. Ownership β_3				-8.400***	-14.073***	-11.075***
				(0.761)	(0.433)	(0.550)
Dividend x High-Tax δ_1	-6.845	-12.608	-9.289	5.574***	6.777***	7.117***
	(9.670)	(8.626)	(7.689)	(0.318)	(0.366)	(0.220)
High-Tax x Low Inst. δ_2				9.770***	12.450***	12.295***
				(0.780)	(0.533)	(0.595)
Dividend x Low Inst.δ_3				8.825***	14.907***	10.263***
				(0.829)	(0.456)	(0.552)
Div x H-T x Low Inst. δ_4				-14.114***	-18.229***	-14.872***
				(0.837)	(0.560)	(0.645)
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
R^2	0.026	0.070	0.094	0.023	0.049	0.131
Observations	665	665	665	4207	4207	4207

Notes: This table reports estimates for 24, 18, and 12 months after an acquisition for both the difference-in-difference (DD) and the difference-in-difference-in-difference (DDD) specification. Robust standard errors are in parenthesis. Standard errors are clustered at the group level for the (DDD) specification. Statistical significance denoted, 10 percent *, 5 percent **, and 1 percent ***. The results are robust without clustering the standard errors and clustering at the firm level. Data comes from Thomson and Reuters Securities Data Company (SDC) and Compustat, years 1992 - 2012.

7.5 Eliminating Mergers in Quick Succession

One concern with the data set constructed for the baseline analyses is that a single acquiring firm may perform several mergers in quick succession. If one merger closely follows another then the long-run returns of the second may conflate the effects from the previous merger. To manage this concern, we perform (DD) and (DDD) analyses excluding mergers that occur in quick succession. Results are presented in Table 10.

Table 10: Robustness: Eliminating Mergers in Quick Succession

	(1)	(2)	(3)	(4)	(5)	(6)
Dividend Firms β_1	10.844	10.857	13.100*	-0.083	-0.002	-0.144
	(6.900)	(7.571)	(7.854)	(0.062)	(0.065)	(0.138)
High-Tax β_2	3.954	0.707	3.112	-14.441***	-14.101***	-14.760***
	(11.548)	(13.111)	(13.133)	(1.088)	(1.541)	(1.578)
Low Inst. Ownership β_3				-11.075***	-12.320***	-15.433***
				(0.550)	(0.747)	(0.637)
Dividend x High-Tax δ_1	-9.289	-9.363	-12.423	7.117***	4.559***	4.890***
	(7.689)	(8.903)	(9.289)	(0.220)	(0.485)	(0.545)
High-Tax x Low Inst. δ_2				12.295***	10.680***	14.083***
				(0.595)	(0.821)	(0.658)
Dividend x Low Inst.δ_3				10.263***	10.040***	12.221***
				(0.552)	(0.765)	(0.646)
Div x H-T x Low Inst. δ_4				-14.872***	-11.925***	-15.101***
				(0.645)	(0.878)	(0.728)
Year Fixed Effects	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
Mergers Within 2 Years Dropped		✓			✓	
Mergers Within 3 Years Dropped			✓			✓
R^2	0.094	0.071	0.065	0.131	0.112	0.119
Observations	665	376	315	4207	3011	2234

Notes: This table reports the robustness specifications limiting the data by dropping mergers that occurred within 2 years of another merger in the same firm, (2) and (5), 3 years of another merger in the same firm (3) and (6) and the benchmark specifications with all of the data, (1) and (4). The dependent variable is the cumulative abnormal return 24 months after an acquisition. Robust standard errors are in parenthesis. Standard errors are clustered at the group level for the (DDD) specification. Statistical significance denoted, 10 percent *, 5 percent **, and 1 percent ***. The results are robust without clustering the standard errors and clustering at the firm level. Data comes from Thomson and Reuters Securities Data Company (SDC) and Compustat, years 1992- 2012.

Columns (1) and (4), in Table 10, are the benchmark estimates using the full data set from Table 6. Columns (2) and (5) exclude mergers that occurred fewer than 2 years after another merger within the acquiring firm. Similarly, Columns (3) and (6) exclude mergers that occurred fewer than 3 years after another merger. The (DD) and (DDD) estimates remain similar to the benchmark estimates in both of these sub-samples. The estimates in Columns (3) and (6) are actually larger in magnitude than the benchmark model and the (DDD) estimate is statistically significant. These results suggest that the bias from conflating mergers is small, which is consistent with the fact that most acquiring firms in the sample make only one acquisition during the sample period and that the distribution of the number of mergers per acquiring firm is roughly exponential.