THE IMPACT OF SHAREHOLDER TAXATION ON MERGERS AND ACQUISITIONS

Eric Ohrn* Nathan Seegert†‡

Abstract

Mergers are an important part of the US economy. Successful mergers efficiently reallocate capital, capture positive synergies, and exploit economies of scale. However, inefficient mergers dampen innovation and decrease efficiency. We demonstrate, within a neoclassical model, that inefficient acquisitions can become value-maximizing to shareholders when dividend tax rates are higher than capital gains tax rates. This tax wedge allows shareholders to free “trapped equity, providing a tax discount to acquisitions. We test this mechanism using variation in the dividend tax rate before and after 2003. Our results suggest that acquiring firms performed 14 percent better after the dividend tax rate was reduced in 2003. These results are especially stark given the sheer magnitude of merger and acquisition activity in the US economy, which in 2014 reached $1.6 trillion.

Keywords: Mergers and Acquisitions, Taxation.
JEL Classification: G34, H20, H25

*Department of Economics, Grinnell College.
†Department of Finance, University of Utah, 1655 East Campus Center Drive, Salt Lake City, UT 84112. Office phone: (801) 585-7131. nathan.seegert@business.utah.edu.
‡We would like to thank Luis Rayo, Joel Slemrod, James Hines Jr., Michael Devereux, Steve Bond, Yona Rubinstein, John Sutton, Shmuel Baruch, Avner Kalay, participants at the Oxford Centre for Business Taxation Doctoral Conference, the International Institute of Public Finance, the National Tax Association, and seminars at the University of Michigan and the London School of Economics.
1 Introduction

For more than 30 years economists have debated whether dividend and capital gains taxes distort corporate investment, (King, 1977; Auerbach, 1979; Bradford, 1981; Poterba and Summers, 1985; Yagan, 2013). In contrast, few studies have considered the impact of shareholder taxation on merger and acquisition behavior—a particularly large component of corporate investment (Auerbach and Reishus, 1987, 1988; King, 1989). This study attempts to fill this void.

To highlight the importance of this topic, consider the role of mergers and acquisitions in the economy. Successful mergers and acquisitions efficiently reallocate capital, capture positive synergies, and generate economies of scale. However, inefficient mergers can dampen innovation, decrease efficiency, and destroy shareholder value. Therefore, any distortion—tax or otherwise—to a firm’s acquisition behavior may impact both the economy at large and the organization of economic activity within and across firms. These distortions may be especially costly given the sheer magnitude of merger and acquisition activity, which in the US totaled $12.78 trillion between 2000 and 2012.

It has long been recognized that frictions in financial markets can cause inefficient acquisitions to be profitable from a firm’s perspective (Jensen and Ruback, 1983). A recent study by Dai, Maydew, Shackelford, and Zhang (2008) claims that, “Taxation is one of the most prevalent market frictions in financial markets, affecting investors decisions and distorting the valuation of assets.” Most of the literature on tax related distortions to acquisitions has focused on the ability of acquires to use a target firm’s net operating losses, increase leverage, or fully utilize tax shields. For example, Hayn (1989) finds that abnormal returns are associated with the target firm having more net operating loss carry-forwards and other tax characteristics, using acquisitions from 1970–1985. Many of these tax advantages were modified or eliminated in the tax reform act of 1986. We extend this literature by empirically testing a different tax motivation based on how shareholder taxation affects the internal cost of capital in a firm, first described by Auerbach and Reishus (1987). Similar to previous tax mechanisms, this tax mechanism is a transfer of benefits from the government to the merging firms.

Fundamentally, shareholder taxes affect the implied cost of capital within a firm, which in
turn, affects the decision of firms to make and accept acquisitions. Shareholder taxes create a tax discount for retained earnings because if the firm distributes the earnings to shareholders some of the earnings go to the government (Harris, Hubbard, and Kemsley, 2001). This tax discount provides motivation for mergers and acquisitions that allow target firm shareholders to “free” their trapped equity (Kraakman, 1988). If acquiring firms are able to extract some of these gains, then a tax discount may be available on the price of the acquisition.

This discount means that some mergers may be profit maximizing from the perspective of both the acquiring and target firms involved, but may be inefficient for the economy as a whole. For example, if dividends are taxed at a rate of 38.6 percent and capital gains are taxed at 20 percent, acquisitions which destroy up to 23 percent of the value of the target firm may be profit maximizing. Importantly, this tax discount is only available for acquiring firms that pay dividends and whose cash flows are, by extension, subject to the dividend tax rate. If, instead, firms distribute retained earnings through share repurchases then the discount disappears. An extended example and a formal model that further elucidate this intuition are presented in Sections 2 and 3.

The tax act of 2003 provides an interesting context to study the effect of shareholder taxes. Prior to the 2003 reform, dividends were taxed as ordinary income with a top rate of 38.6 percent while capital gains were taxed at 20 percent. The reform decreased both rates to 15 percent eliminating the tax discount. This paper uses the tax act to study the impact on mergers and acquisitions, adding to evidence on the tax act’s impact on firm payout polices Chetty and Saez (2005); Brown, Liang, and Weisbenner (2007), ex-dividend returns Cloyd, Li, and Weaver (2006); Chetty, Rosenberg, and Saez (2006), and firm valuation Auerbach and Hassett (2006). Recent studies find that the decrease in the dividend and capital gains tax from the 2003 tax reforms led to a decrease in the cost of equity capital (Lang and Shackelford, 2000; Ayers, Cloyd, and Robinson, 2002; Dhaliwal, Li, and Trezevant, 2003; Dhaliwal, Krull, Li, and Moser, 2005; Blouin, Raedy, and Shackelford, 2003; Dhaliwal, Krull, and Li, 2007). We test whether this decrease in investor taxes also promoted higher quality acquisitions by decreasing the tax discount for retained earnings.

A large literature in economics, finance, and accounting discusses the effect of shareholder taxes (capital gains and dividend taxes) on firm value (for example Modigliani and Miller (1958, 1963); Miller (1977); Miller and Modigliani (1961); Brennan (1970).

Harris, Hubbard, and Kemsley (2001) find evidence that there exists a tax discount for retained earnings in a firm, across tax regimes in the United States and across different countries.
We test the effect of the decrease in dividend tax rate on the value of acquisitions using both long-run and short-run cumulative abnormal returns. Short-run abnormal returns capture the markets initial valuation, put simply, whether the acquiring firm paid too much or too little for the firm. Most studies using short-run returns find small positive cumulative abnormal returns after an acquisition, suggesting the acquirer may have received a slight discount. In contrast, long-run abnormal returns capture the realized synergies from the acquisition. Leading up to an acquisition, acquirers discuss the operating improvements and synergies that justify the acquisition. However, these improvements take time to implement and are thought to take up to two years to be fully realized. Most studies using long-run returns find large negative cumulative abnormal returns after an acquisition, suggesting the transition costs of the acquisition may have been larger than the operating improvements.\(^3\) Several studies have found significant underperformance for subsets of acquisitions, such as those performed by low book-to-market “glamour firms (Rau and Vermaelen, 1998) and firms that use stock as the method of payment Loughran and Vijh (1997).\(^4\)

If shareholder taxes distort acquisition behavior by affecting the implied cost of capital within a firm, then both short-run and long-run abnormal returns will be affected. Specifically, dividend taxes create a tax discount that makes initial valuation of an acquisition to appear favorable and causes some acquisitions to be profitable that otherwise would not be. In this case, short-run abnormal returns after an acquisition will be more positive for firms with a larger tax discount, those that pay a dividend when the dividend tax is large. In contrast, long-run abnormal returns after an acquisition will be more negative for firms that receive a larger tax discount because they are willing to accept mergers with lower synergies in exchange for the benefits of the tax discount.

There is an important literature that discusses the best practices of calculating long-run cumulative abnormal returns in an effort to avoid potential misspecification issues.\(^5\) Generally, tests with long-run cumulative abnormal returns are a joint test of the focal null hypothesis and the

\(^4\)Several studies examine accounting performance productivity data and find the operating improvements are minimal Ghosh (2001); Maksimovic and Phillips (2001).
\(^5\)Long-run cumulative abnormal returns may be biased due to survivor bias, rebalancing bias, and skewness bias if these biases are not accounted for (Lyon, Barber, and Tsai, 1999). Mitchell and Stafford (2000) suggest using calendar portfolio approach suggested by Fama (1998) to avoid biases due to positive cross-correlations for firms that make acquisitions. Harford (2005) suggest controlling for industry because of biases due to clustering of merger activity and mean-reversion in industry-adjusted operating performance noted by Barber and Lyon (1996). Loughran and Ritter (2000) suggest that the monthly returns should be equally weighted rather than value weighted.
asset pricing model used to calculate the long-run cumulative abnormal returns. To limit the bias from misspecification this paper focuses on tests from a difference in difference empirical design.

The difference in difference empirical design allows for differences between groups of firms (in our study firms that pay dividends and firms that repurchase shares) and between time periods (in our study before and after 2003). In terms of model specification, the difference in difference specification allows for different biases between firms that pay a dividend and firms that repurchase shares and between firms making acquisitions before and after 2003. The difference in difference empirical design, therefore, alleviates some of the concern of biases in long-run cumulative abnormal returns.

To further isolate the policy effect we use variation across firms in the percentage of tax-exempt institutional shareholders and method of payments. The tax discount is larger for firms with a low percentage of tax-exempt shareholders and acquisitions made with cash. Therefore, the policy effect should be larger for these two subgroups.

The empirical results support the existence of the tax discount and the hypothesis that shareholder taxation induces inefficient acquisition behavior. Our results suggest that acquiring firms performed 14 percent better after the dividend tax rate was reduced in 2003. Consistent with the theory, the effects are more pronounced for firms with a low percentage of tax-exempt institutional shareholders and acquisitions paid for with cash. Further, this evidence suggests that value maximizing, but socially inefficient, responses to shareholder taxes in years prior to 2003 led newly acquired assets to be operated, on average, at only 86 percent efficiency.

These patterns are consistent with dividend behavior around the 2003 reform. Yagan (2013) documents that within one year of the reform, total dividend payouts rose by 20 percent. Further, Chetty and Saez (2010) show that this response was concentrated among firms that had significant accumulated assets. If mergers and acquisitions are a tax preferred, albeit unconventional, substitute for dividends then when the discount is eliminated in 2003, profit maximizing firms should respond by increasing dividend payments.

---

6 However, the difference in difference specification will be biased if the bias from the cumulative abnormal returns changes before and after 2003 differentially for firms that pay dividends and firms that repurchase shares.

7 Chetty and Saez (2010) propose an agency model to explain the dividend payment behavior after 2003. In our context of mergers and acquisitions we find no evidence, tested formally in Appendix C, that agency costs can explain the change in merger and acquisition behavior after 2003.
The findings of this study add to the taxation and merger literature, but also speak directly to two broader literatures. The first literature investigates the effects of shareholder taxation on investment (Harberger, 1962; Feldstein, 1970; King, 1977; Auerbach, 1979; Bradford, 1981). This study finds that even in a (new view) model in which internal investment is unaffected by shareholder taxation, acquisition behavior may be severely distorted. The second literature investigates possible explanations for the empirical puzzle that acquiring firms perform surprisingly poorly following an acquisition (Gregory, 1997; Agrawal and Jaffee, 2000). The results of this study suggest that approximately 10 percent of the puzzle may be attributable to distortions from shareholder taxation.

2 A Simple Example of the Tax Discount

Before the complete model is presented, this section presents a simple example which illustrates the intuition behind the tax discount. Consider an acquiring firm, denoted Firm 1, purchasing a target firm, denoted Firm 2, with the latter having assets worth $100 in present value prior to any shareholder taxation. Assume that dividends are taxed at 38.6 percent ($\tau_d$) and the capital gains tax rate is 20 percent ($\tau_{cg}$), the tax rates in the United States in 2002.

Consider the cost to Firm 1 when purchasing Firm 2. First, if Firm 2’s shareholders keep the firm, the value of the assets is $61.40 because eventually profits must be paid as dividends, in this example taxed at 38.6 percent. Second, if Firm 2’s shareholders sell the firm, the proceeds from the sale will be subject to the capital gains tax ($\tau_{cg} = .2$). Therefore Firm 2’s shareholders are willing to sell the firm for any price greater than $76.75, which after the capital gains tax of 20 percent is $61.40, assuming 100 percent of the value represents capital gains. Thus, Firm 1 pays Firm 2’s shareholders $76.75, which otherwise could have been paid as dividends with an after-tax value of $47.1245 = (1 - .386)76.75.

---

8See also Poterba and Summers (1985, 1983, 1984); Shleifer and Vishny (1986); Poterba, Hall, and Hubbard (1987); Bagwell and Shoven (1989); Goulder and Summers (1989); Gordon and MacKie-Mason (1990); Sinn (1991); Poterba (1991); Cummins, Hassett, Hubbard, and Caballero (1994); Lease, John, Kalay, Loewenstein, and Sarig (1999); La Porta, Lopez-de Silanes, Shleifer, and Vishny (2000); Bell and Jenkinson (2002); Graham (2003); Auerbach and Hassett (2003); Poterba (2004); Bond, Devereux, and Klemm (2007); Desai and Dharmapala (2011) and Dharmapala (2009).

9See also Myers (1984); Agrawal and Madelker (1990); Agrawal, Jaffee, and Madelker (1992).
Now consider the benefits to Firm 1 of buying Firm 2, inclusive of the synergies that may exist. Firm 1 manages Firm 2’s assets such that, before tax, the assets are worth $100\sigma$; where $\sigma = 1$ indicates the assets are managed equally well by either firm, $\sigma > 1$ indicates the assets are better utilized by Firm 1, and $\sigma < 1$ indicates the assets are better utilized by Firm 2. Therefore, after tax, the assets are worth $61.40\sigma = (1 - .386)$100$\sigma$, because the profits are taxed at the dividend tax rate.

The acquisition is value maximizing if the benefits are greater than the costs,$$
$61.40\sigma > 47.1245$
$$

which occurs when,$$
\sigma > \frac{1 - .386}{1 - .2} = \frac{1 - \tau_d}{1 - \tau_{cg}}.
$$

This implies acquisitions where Firm 1 manages Firm 2’s assets above 76.75% efficiency are value maximizing. All acquisitions with a synergy parameter less than one are socially inefficient; however, they are value maximizing as long as the synergy parameter is greater than the ratio $\frac{1-\tau_d}{1-\tau_{cg}}$. Thus, when the dividend tax rate is greater than the capital gains tax rate, firms accept socially inefficient acquisitions.

This inefficiency does not occur if Firm 1 does not pay dividends and instead issues equity to pay for the acquisition. In this case, the benefit is the same, but the cost is the $76.75$ in additional equity Firm 1 issues to buy Firm 2. The acquisition is value maximizing if

$$
$61.4\sigma > 76.75$
$$

or, equivalently, when

$$
\sigma > \frac{1}{1 - .2} = \frac{1}{1 - \tau_{cg}}.
$$

In contrast to the case where Firm 1 pays a dividend and accepts “too many” acquisitions, in this
case Firm 1 makes “too few” acquisitions. For an acquisition to be value maximizing in this case, Firm 1 must manage Firm 2’s assets at 125 percent efficiency or greater.

The tax discount is also absent when Firm 1 finances the acquisition with funds that would otherwise be used to repurchase shares. In this case, Firm 1 must still pay Firm 2 shareholders $76.75, but now the after tax value of these funds are $76.75(1-.2) = $61.4 because share repurchases are taxed at \( \tau_{cg} \). Firm 1 only acquires Firm 2 when

\[
61.4\sigma > 61.4
\]

which occurs only if \( \sigma \geq 1 \), the socially efficient condition.

This simple example demonstrates that the magnitude of the distortion from taxation could be substantial and alludes to the heterogeneity in the tax discount across firms that facilitates empirical tests of the tax discount.

3 Model

This section uses a neoclassical two-period model to show how the tax discount generated by the differential tax treatment of dividends and capital gains may induce value maximizing, but socially inefficient, acquisitions.

3.1 Shareholder Decisions

Consider a single firm, denoted Firm 1, with one shareholder. At the beginning of period 1, Firm 1 is presented with an opportunity to acquire a randomly selected target firm, denoted Firm 2, and owned by one shareholder. The acquisition is characterized by the target firm’s assets, retained earnings \( C \), production technology \( g(\cdot) \), and a synergy parameter \( \sigma \in [0, \infty) \). Firm 1’s value of Firm 2’s assets is \( \sigma(g(C) + C) \), where the synergy parameter captures the ability of Firm 1 to manage Firm 2’s assets. If \( \sigma = 1 \) then Firm 1 manages Firm 2’s assets as well as Firm 2. However, if \( \sigma > 1 \) (respectively \( \sigma < 1 \)) then Firm 1 manages Firm 2’s assets better (worse) than Firm 2, increasing (destroying) social benefit.
Given the acquisition opportunity \((C, g(\cdot), \sigma)\), Firm 1 decides whether to make the acquisition \(Y = 1\) or not \(Y = 0\). Whether taxation creates a distortion in acquisition behavior depends on whether firms have an incentive to make acquisitions that are socially inefficient \((\sigma < 1)\).

Firm 1 begins with \(X\) retained earnings. In period 1, Firm 1 decides the amount of retained earnings to distribute \((D \geq 0)\) and the amount of equity \((E \geq 0)\) to issue. Earnings can be distributed in one of two ways: through dividend payments, which are taxed at the dividend tax rate \(\tau_d\) or share repurchases, which are taxed at the capital gains tax rate \(\tau_{cg}\). We consider these cases separately.\(^{10}\) Firm 1’s distribution and equity issuance choices implicitly define its level of investment for the next period \(I = X + E - D\).

In period 2, investment generates net profits \(f(I)\), where \(f(\cdot)\) is a strictly concave function.\(^{11}\) In addition to equity, shareholders may hold government bonds with an untaxed rate of return of \(r > 0\).\(^{12}\) At the end of period 2, the firm liquidates, returning its principle and profits, which are taxed at the dividend tax rate \(\tau_d\), to shareholders.

### 3.2 Shareholder Payoffs

If Firm 1 does not make an acquisition, \(Y = 0\), its value depends on its choices of distributions and equity \(D_0\) and \(E_0\),

\[
V_0 = (1 - \tau_i)D_0 - E_0 + \frac{(1 - \tau_d)[f(X + E_0 - D_0) + X - D_0] + E_0}{1 + r}
\]  

where \(\tau_i = \tau_d\) if the firm distributes retained earnings through dividends and \(\tau_i = \tau_{cg}\) if the firm distributes retained earnings through share repurchases.

If Firm 1 does make an acquisition, \(Y = 1\), its value depends on the acquisition value \(\sigma(g(C) + C)\) and the amount it pays Firm 2’s shareholder, defined as \(M\).\(^{13}\) The market for acquisitions is assumed

\(^{10}\)In the model, firms would always choose to distribute retained earnings through share repurchases rather than dividends. However, dividends are frequently observed for numerous reasons including tax rules that tax share repurchases as dividend payments if they look like dividend payments. Both cases are considered separately.

\(^{11}\) The net profits function is defined as \(f(I) = F(I) - \delta I\) where \(F(I)\) is the gross production function which includes the depreciation of capital used for production. The functional form of the net profits function is left general in the text of the paper, however, a parametric example where \(f(I) = \frac{1 + e}{1 + e}AI\) is given in ??.

\(^{12}\)The analysis abstracts from general equilibrium effects on the rate of return to government bonds by assuming it is exogenous to the model.

\(^{13}\)The assumption that synergies do not affect the level of investment in the firm does not fundamentally affect the results. A richer model that allows for investment to change with the synergy level is given in ??.
to be competitive such that Firm 2’s shareholder is indifferent between selling the firm, and receiving \((1 - \tau_{cg})M\), and keeping the firm, and receiving the principle and profits net the dividend tax when the firm liquidates in the second period.\(^{14}\) The cost of the acquisition is therefore given by,

\[
M = \frac{1 - \tau_d \frac{g(C) + C}{1 + r}}{1 - \tau_{cg}}.
\]  

The value of the firm when an acquisition is made is

\[
V_1 = (1 - \tau_i)D_1 - E_1 + \frac{(1 - \tau_d)[f(I) + \sigma(g(C) + C) + I - E_1] + E_1}{1 + r},
\]

which includes the assets of the target firm and the updated levels of equity and distributions. Comparing equations (1) and (3) reveals the equilibrium level of investment, \(I\), is the same regardless of the acquisition decision.\(^{15}\) Therefore, Firm 1 can issue more equity (\(E_1 = E_0 + M\)), make fewer distributions (\(D_1 = D_0 - M\)), or do some of both (\(D_0 > D_1 \) and \(E_1 > E_0 \)) to pay for the acquisition.

### 3.3 Equilibrium Behavior

This section derives the equilibrium behavior of the firm. An equilibrium is characterized by a set \((E^*, D^*, Y^*)\); Firm 1’s choices of equity issuances, distribution payments, and acquisition decision. The analysis focuses on a subset of illustrative cases, leaving the full set to be discussed in Appendix A.

The analysis focuses on the cases in which Firm 1 either has an abundance of cash (\(f'(X - M) < r\)) and chooses to distribute earnings, whether it makes an acquisition or not, or has scarce cash (\(((1 - \tau_i)f'(X) > r)\) and issues equity. Optimally a firm will never simultaneously issue equity and distribute earnings because, in this case, the firm’s value could be increased by reducing both equity and distributions by $1, lowering its tax bill by \(t_d r/(1 + r)\). Appendix A considers the cases in which Firm 1 has an intermediate level of cash (\(f'(X) \in [r, r/(1 - \tau_i)]\)) and neither distributes

\(^{14}\)The assumption that the market for acquisitions is competitive does not affect the results. For details see Appendix A.3 that considers the case where the target firm has an exogenous level of bargaining power.

\(^{15}\)Investment does not change with the acquisition by construction. \(??\) shows the results are robust to the case where synergies augment Firm 1’s technology, causing equilibrium investment levels to differ with acquisition behavior.
earnings or issues equity and the cases where Firm 1 only distributes earnings without an acquisition \( f'(X) < r \) and \( f'(X - M) > r \).

The equilibrium levels of equity issuances and distribution payments are derived from the first order conditions of equations (1) and (3). When Firm 1 has an abundance of cash it distributes earnings such that \( D_1^* = D_0^* - M > 0 \). When Firm 1 has scarce cash it issues equity such that \( E_1^* = E_0^* + M > 0 \). The equilibrium level of investment is greater when Firm 1 has an abundance of cash because the internal cost of capital is lower. However, if the firm has an abundance of cash the equilibrium level of investment is undistorted by the distribution tax rate.

In equilibrium, the firm accepts an acquisition opportunity if the value of the firm with the acquisition, given in equation (3), is greater than the value of the firm without the acquisition, given in equation (1). The difference in the value of the firm, with and without an acquisition, can be written as\(^\text{16}\),

\[
V_1 - V_0 = \begin{cases} 
(1 - \tau_{cg})\sigma M - (1 - \tau_i)M, & \text{if } D_1 > 0 \text{ and } E_1 = 0 \\
(1 - \tau_{cg})\sigma M - (1 - \tau_d/(1 + r))M, & \text{if } E_1 > 0 \text{ and } D_1 = 0.
\end{cases}
\]

The benefit of an acquisition is the value of the assets and synergy net of taxes, given by the first term \((1 - \tau_{cg})\sigma M\). The cost of the acquisition differs based on whether the firm distributes retained earnings through dividend payments \((D_1 > 0 \text{ and } \tau_i = \tau_d)\), distributes retained earnings through repurchasing shares \((D_1 > 0 \text{ and } \tau_i = \tau_{cg})\), or issues equity \(E_1 > 0\). The cost of the acquisition, if the firm distributes earnings, is \(M\) forgone distributions in the first stage, with an after tax value of \((1 - \tau_i)M\). The cost of the acquisition, if the firm issues equity, is the opportunity cost of the equity, \((1 - \tau_d/(1 + r))M\). These differences in costs lead to differences in the set of acquisitions Firm 1 accepts, characterized by a threshold synergy level \(\sigma^*\).

The threshold synergy value is determined by the values of \(\sigma\) such that the difference in firm

\(^\text{16}\)The full derivation of the difference between the value of the firm with and without an acquisition is given in Appendix A.2.
values, given in equation (4), are greater than zero,

\[
\sigma^* = \begin{cases} 
1 - \tau_i & \text{if } D_1 > 0 \& E_1 = 0 \\
1 & \text{if } E_1 > 0 \& D_1 = 0.
\end{cases}
\] (5)

Firm 1 is said to over acquire (respectively under acquires) if \( \sigma^* < 1 \) (respectively \( \sigma^* > 1 \)). Firm 1 is said to acquire efficiently if \( \sigma^* = 1 \).

**PROPOSITION 1** When Firm 1 has limited cash (e.g. \( r < f'(X) \)) it under acquires. In contrast, when Firm 1 has abundant cash (e.g. \( r > (1 - \tau)f'(X - M) \)) it over acquires if the tax rate on distributions is greater than the capital gains tax rate (e.g. dividend payments where \( \tau_i = \tau_d > \tau_{cg} \)) and acquires efficiently if the tax rate on distributions is equal to the capital gains tax rate (e.g. share repurchases \( \tau_i = \tau_{cg} \)).

**PROOF:**

\[
\tau_i > \tau_{cg} \Rightarrow 1 - \tau_{cg} > 1 - \tau_i \Rightarrow 1 > \frac{1 - \tau_i}{1 - \tau_{cg}} = \sigma^*
\]

\[
\tau_i = \tau_{cg} \Rightarrow 1 - \tau_{cg} = 1 - \tau_i \Rightarrow 1 = \frac{1 - \tau_i}{1 - \tau_{cg}} = \sigma^*
\]

\[
0 < \tau_{cg} \Rightarrow 1 - \tau_{cg} < 1 \Rightarrow 1 < \frac{1}{1 - \tau_{cg}} = \sigma^* \]

Proposition 1 predicts that before 2003, when the dividend tax rate was greater than the capital gains tax rate, dividend paying firms accepted socially inefficient acquisitions. However, these socially inefficient acquisitions became unprofitable after the 2003 tax change when the dividend tax rate was harmonized with the capital gains tax rate. In contrast, firms that repurchased shares acquired efficiently before and after 2003.

Proposition 1 provides two testable implications. First, the model predicts in response to the dividend tax decrease in 2003 that firms should make fewer acquisitions and increase their dividend payments. Second, the model predicts firms that pay a dividend accepted lower quality acquisitions before 2003. The first implication is consistent with previous studies on the dividend tax rate change of 2003 (Yagan, 2013; Chetty and Saez, 2005). The second implication is tested in Section 6 using firms that repurchase shares as a control group because their incentives are undistorted before and after 2003.
The empirical analysis uses variation in distribution strategies and tax policy to test the predictions from the model. In addition, the empirical analysis uses variation in institutional shareholders, who are not subject to the distribution tax, to further isolate the distortions described in the model. Intuitively, the distortions from the distribution tax are muted as the percentage of tax-exempt shareholders increases. Therefore, the distortions to acquisition behavior should be stronger for firms with low percentages of institutional shareholders. Appendix A.4 formally shows this intuition by explicitly modeling the objectives of the firm when some exogenous percentage of shareholders are tax-exempt.

4 Data Collection

Data on acquisitions and firm characteristics are collected to test the predictions of the model. This section describes the sources, collection, and cleaning of these data, as well as descriptive statistics.

4.1 Firm Characteristics

Acquiring firm characteristics such as dividend payments, share repurchases, assets, sales, cash flows, marginal Q, the Hadlock-Pierce Financial Distress measure (Hadlock and Pierce (2010)), and retained earnings are taken from the Compustat North American Fundamentals Quarterly database.\textsuperscript{17} These variables control for correlation between acquiring behavior, payout behavior, firm size, financial distress, and productivity. For each acquisition, the characteristics of the acquiring firms are averaged over the two years prior to the acquisition. Each acquisition is also matched to the average level of institutional holding in the years prior to the acquisition (constructed from 13f files contained in the Thomson Reuters Institutional Holdings database).

The effect of an acquisition on firm value is measured as the cumulative abnormal return (CAR) 24 months after an acquisition. This measure provides a market valuation of the firm relative to its predicted value in the absence of the acquisition.\textsuperscript{18} In addition, this measure allows for a

\textsuperscript{17} presents definitions of Cash Flows, Marginal-Q, and the HP Index. Compustat provides data on shares repurchased starting with the first quarter of 2004. The procedure outlined in Stephens and Weisbach (1998), which approximates shares repurchased as the dollar value of decreases in shares outstanding, is used to construct repurchases prior to 2004. This method is internally valid as it closely approximates repurchase behavior in years when Compustat records actual repurchase data.

\textsuperscript{18} The propositions of the model may also be used to generate testable hypotheses regarding the number of acquisi-
direct comparison of the results with previous studies documenting the post-merger performance puzzle. The difference in difference specification (discussed in section 5) controls for any possible misspecification from the cumulative abnormal returns if they are not systematically related to the firm’s method of distribution (Lyon, Barber, and Tsai, 1999; Kothari and Warner, 2006)). The cumulative abnormal returns are calculated with data from the Center for Research in Security Prices (CRSP) database.\textsuperscript{19}

4.2 Acquisition Data

Acquisition characteristics are taken from the Bureau Van Dijk Amadeus Zephyr database, with permission of Zephyr. The sample window is restricted to acquisitions that occur between the dates of January 1, 1998 and December 31, 2008, five years before and after the 2003 tax reforms. The sample is further restricted in several ways. First, acquisitions made by private companies not listed on the NYSE or NASDAQ are excluded because they do not have stock price data by which to measure performance. Second, less than 100 percent purchases of the target firm are excluded to avoid acquisition type effects. Finally, acquisitions made by firms with more than twenty 100 percent acquisitions in the sample are excluded.\textsuperscript{20}

4.3 Descriptive Statistics

Table 1 provides descriptive statistics on all variables used in the empirical analysis by distribution firm groups. The descriptive statistics highlight 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 almost all variables, firms that pay dividends and firms that repurchase shares are more similar to each other than to firms that do neither or both. The similarity of these groups supports the use of firms that repurchase shares as a control for firms

\textsuperscript{19}\textsuperscript{?? provides details on the method used to calculate cumulative abnormal returns.

\textsuperscript{20}The results are robust to setting other limits on the maximum number of mergers per firm. This exclusion limits the effects from overlapping mergers and excludes mostly large banking firms from the sample.
that pay a dividend.

4.4 Synergy Threshold Impacts of JGTRA 2003

For top income tax payers, the Jobs and Growth Tax Relief Reconciliation Act (JGRTRA) of 2003 decreased the tax rate on qualified dividends from 38.6 percent to 15 percent and lowered the tax rate of long-term capital gains from 20 percent to 15 percent.\textsuperscript{21}

Prior to 2003 the model suggests the distortion due to the tax discount for dividend paying firms was substantial, causing the threshold synergy level to be as low as 76.75 percent efficiency (Proposition 1).\textsuperscript{22} However, after 2003 when the dividend and capital gains tax rates are harmonized, the threshold level is 1 and only efficient acquisitions are undertaken. Therefore, the difference in quality of acquisitions undertaken by dividend paying firms before and after 2003 may be substantial. In contrast, the acquisition behavior of firms that repurchase shares is undistorted before and after the JGTRA (Proposition 1).\textsuperscript{23}

5 Empirical Design

The remainder of the paper seeks to determine the empirical importance of the tax discount on acquisition behavior. Section 5.1 defines the baseline difference in difference empirical specification used to test the hypothesis that the tax discount is empirically large enough to cause distortions in acquisition behavior. Section 5.2 provides a triple difference empirical specification using variation in institutional holdings. Finally, section 5.3 tests the key identifying assumption of the difference

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Income Level 2002 & Tax Rate 2002 & Income Level 2003 & Tax Rate 2003 & Dividend Tax Rate Change \\
\hline
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 & \\
\hline
\end{tabular}
\end{center}

\textsuperscript{21}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 had been held 60 days during the 121-day period that begins 60 days before the ex-dividend date.

\textsuperscript{22}The tax discount is given by the ratio \((1 - \tau_d)/(1 - \tau_{cg})\) that before 2003 equaled 0.7675.

\textsuperscript{23}The threshold synergy level for firms that issue equity was affected by JGTRA such that the threshold, given by \(1/(1 - \tau_{cg})\), equaled 1.25 prior to 2003 and 1.176 after (Proposition 1).
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dividend Only</td>
<td>479.7</td>
<td>336.4</td>
<td>-695.2</td>
<td>2,631</td>
</tr>
<tr>
<td></td>
<td>(1,548)</td>
<td>(3,114)</td>
<td>(6,542)</td>
<td>(7,522)</td>
</tr>
<tr>
<td>Repurchase Only</td>
<td>3,604</td>
<td>4,699</td>
<td>1,266</td>
<td>19,641</td>
</tr>
<tr>
<td></td>
<td>(8,264)</td>
<td>(36,023)</td>
<td>(3,433)</td>
<td>(108,733)</td>
</tr>
<tr>
<td>Neither</td>
<td>0.162</td>
<td>0.230</td>
<td>0.148</td>
<td>0.244</td>
</tr>
<tr>
<td></td>
<td>(0.345)</td>
<td>(0.500)</td>
<td>(0.728)</td>
<td>(0.374)</td>
</tr>
<tr>
<td>DIV and REP</td>
<td>2.233</td>
<td>2.174</td>
<td>3.187</td>
<td>1.988</td>
</tr>
<tr>
<td></td>
<td>(1.426)</td>
<td>(1.422)</td>
<td>(2.355)</td>
<td>(1.150)</td>
</tr>
<tr>
<td>Retained Earnings</td>
<td>-5.076</td>
<td>-4.946</td>
<td>-4.287</td>
<td>-6.013</td>
</tr>
<tr>
<td></td>
<td>(1.176)</td>
<td>(1.207)</td>
<td>(1.196)</td>
<td>(1.180)</td>
</tr>
<tr>
<td>CAR 12 mo.</td>
<td>-8.002</td>
<td>-6.117</td>
<td>-11.282</td>
<td>-5.472</td>
</tr>
<tr>
<td></td>
<td>(29.45)</td>
<td>(28.28)</td>
<td>(35.08)</td>
<td>(22.25)</td>
</tr>
<tr>
<td></td>
<td>(18.86)</td>
<td>(18.16)</td>
<td>(22.51)</td>
<td>(13.46)</td>
</tr>
<tr>
<td>Baseline Obs.</td>
<td>421</td>
<td>2,213</td>
<td>1,044</td>
<td>3,062</td>
</tr>
<tr>
<td>Instr. Hold. %</td>
<td>0.603</td>
<td>0.664</td>
<td>0.568</td>
<td>0.686</td>
</tr>
<tr>
<td></td>
<td>(0.252)</td>
<td>(0.282)</td>
<td>(0.286)</td>
<td>(0.686)</td>
</tr>
<tr>
<td>Instr. Hold. Obs.</td>
<td>336</td>
<td>1,990</td>
<td>855</td>
<td>2,778</td>
</tr>
</tbody>
</table>

Notes: Dividend Only: firms that paid dividend, but did not repurchase shares prior to acquisition. Repurchase Only: firms that regularly repurchase shares, but have not paid a dividend. Neither: firms that do not distribute retained earnings regularly. DIV and REP: firms that both regularly pay dividends and repurchase shares.
in difference empirical specification.

5.1 Difference in Difference (DD) Empirical Specification

The differential impact of the dividend tax reform of 2003 (JGTRA) across firms with different distribution strategies naturally suggests a difference in difference (DD) empirical strategy. The policy change affects the acquisition incentives for dividend paying firms after 2003 but not firms that repurchase shares (Proposition 1). In addition, firms that repurchase shares are similar to dividend paying firms in terms of firm characteristics (as discussed in section 4). For these two reasons, firms that repurchase shares act as a natural control group.

The baseline difference in difference empirical specification uses as the dependent variable post-acquisition performance, $P_{i,t}$, measured by the cumulative abnormal return twenty-four months after an acquisition, where $i$ denotes the acquiring firm and $t$ denotes the year. The specification also controls for differences between time periods and differences between firms with different distribution strategies by including indicator variables. The indicator variable $d^D$ equals one if the acquisition is made by a dividend paying firm and zero otherwise. Similarly, the indicator variable $d^A$, equals one if the acquisition is made after the policy change and zero otherwise; giving the baseline specification,

$$P_{i,t} = \beta_0 + \beta_1 d^D_i + \beta_2 d^D_id^A_t + z_{i,t}\gamma + \lambda_t + \epsilon_{i,t}.$$  \hspace{1cm} (6)

The baseline specification includes year fixed effects $\lambda_t$ and other controls for firm characteristics, $z_{i,t}$, described in section 4.1. The policy effect is captured by the DD term, $\beta_2$, that estimates the effect of the policy change on the post-acquisition performance of dividend paying firms. The DD coefficient is predicted to be positive because dividend paying firms accept higher quality acquisitions as a result of the dividend tax rate decrease in 2003, while the quality of acquisitions performed by share repurchasing firms is unchanged.

24 A firm that regularly paid a dividend between 1998 and 2002, but did not repurchase shares during this period, is considered to be a dividend paying firm.

25 Firms that repurchase shares are identified in the data as firms that repurchased shares but do not pay dividends. Firms that pay dividends and repurchase shares are excluded because their value of $\tau$ is unclear and, on average, they are much larger than firms that solely pay dividends or repurchase shares. Firms that do not repurchase shares or pay dividend are excluded because they are generally younger and growing faster than firms that actively pay out retained earnings to shareholders.
Concerns of bias in the estimation of the standard errors are addressed in two ways: multi-way clustering and a triple difference (Bertrand, Duflo, and Mullainathan, 2004). The standard errors in Table 2 are clustered at the firm and year level to address concerns of serial correlation and allow for an arbitrary covariance structure. The multi-way clustered robust standard errors are similar to those in Acemoglu and Pischke (2003) and follow methods presented in Cameron, Gelbach, and Miller (2011), Petersen (2009) and Thompson (2011). The triple difference reported in Table 3 uses variation in institutional shareholders as an additional set of control groups.

5.2 Institutional Shareholder Heterogeneity Specifications

Intuitively, the effect of the tax discount on acquisition behavior should be larger for firms with a higher percentage of taxable shareholders. Therefore, using variation across firms in the percentage of the firm owned by institutions, which are tax-exempt, provides additional tests of the empirical importance of the tax discount. The DD specification can be updated in two ways to use this additional variation. First, the DD specification may be estimated in two subsamples: firms with a low percentage of institutional shareholders and firms with a high percentage of institutional shareholders. The DD coefficient from the subsample of firms with a low percentage of institutional shareholders is predicted to be positive and larger than the coefficient from the subsample of firms with a high percentage of institutional shareholders.

Second, a triple difference (DDD) empirical specification can be used to capture the heterogeneity in the effect of the tax discount on acquisition behavior across firms with varying levels of institutional shareholders. The DDD empirical specification interacts the indicator variables from the DD specification with the percentage of the firm held by institutions, \( I_{i,t} \), which is a continuous variable,

\[
P_{i,t} = \beta_0 + \beta_1 d^{D}_i + \beta_2 I_{i,t} + \beta_3 d^{D}_i d^{A}_i + \beta_4 d^{D}_i I_{i,t} + \beta_5 d^{A}_i I_{i,t} + \beta_6 d^{D}_i d^{A}_i I_{i,t} + z_{i,t} \gamma + \lambda t + \epsilon_{i,t}. \tag{7}
\]

The theory predicts the coefficient on the triple interaction, \( \beta_6 \), is negative because the DD effect is smaller when the percentage of institutional shareholders is larger. The policy effect for firms with zero percent institutional shareholders is given by the DD coefficient, \( \beta_3 \), and the difference
in policy effects between firms with one-hundred and zero percent institutional ownership is given by the DDD coefficient, $\beta_6$. The theory predicts the DD coefficient to be positive and the sum of the DD and DDD coefficients to be zero.

The triple difference specification provides a particularly strong test of the implications of the model because it uses three layers of variation to isolate the policy effect. Section Appendix B.1 reports additional tests of the tax discount using variation in the value of the acquisition and whether the acquisition is paid for with cash or not.

### 5.3 Common Trend Assumption

The key identifying assumption is that the unobserved change in performance dividend paying firms would have experienced in the absence of the policy change is captured by the change in performance by firms that repurchase shares. This “common trend” assumption is not directly testable across time periods with the policy change, but can be tested in the years before the policy change. Intuitively, this test determines whether firms’ performance is significantly affected by year and firm group (e.g. firms that pay a dividend) specific idiosyncratic shocks that could be correlated with the dividend tax policy change. The empirical test regresses performance (CAR twenty-four months after an acquisition) on year fixed effects $\lambda_t$ and an indicator variable $d_i^D$ of whether the firm pays dividends interacted with year fixed effects,

$$P_{i,t} = \beta_0 + \lambda_t + \lambda_t d_i^D + \epsilon_{i,t}. \quad (8)$$

The coefficients on the year fixed effects estimate the common trend between firms that pay a dividend and repurchase shares. The coefficients on the year fixed effects interacted with the dividend firm indicator estimate the difference in trend between firms that pay a dividend and firms that repurchase shares. A failure to reject the joint F-test that all of the coefficients on the interaction, $d_i^D \lambda_t$, are zero is a failure to reject the common trend assumption during the years before the policy change. Table 4 in Appendix B reports the empirical test fails to reject the common trend assumption for firms that pay dividends, firms with a low percentage of institutional shareholders, and firms that pay dividends and also have a low percentage of institutional shareholders.
6 Empirical Analysis

As a result of a dividend tax decrease, the model demonstrates firms will redistribute more retained earnings through dividends substituting away from mergers and acquisitions. This substitution will show up in the data in two ways; first firms will undertake fewer mergers and second the mergers that are undertaken will, on average, have greater synergies. Section 6.1 provides evidence that dividend paying firms did perform fewer mergers after the dividend tax decrease and Section 6.2 provides evidence the acquisitions dividend paying firms made had greater synergies. Section Appendix B.1 provides additional evidence looking at subsamples of large, small, and cash financed acquisitions and considering short-run returns. Finally, Section 6.3 relates the estimates to the post-acquisition performance puzzle.

6.1 Merger Activity Analysis

This section reports two measures of changes in the quantity of mergers a firm undertakes after the dividend tax decreases in 2003. The first measure predicts whether a firm will make an acquisition in a given year. The second measure predicts the average number of acquisitions a firm will undertake. Both measures are conditioned on the dividend tax rate, whether the firm redistributes retained earnings through share repurchases or dividends, and other control variables. Figure 1 presents a graphical implementation of the difference-in-difference research design, the coefficients from the analysis are reported in Table 5 and are statistically significant at the 1 percent level.

Panel A of Figure 1 reports the probability a firm makes an acquisition in a given year conditional on the year being before or after the dividend tax cut in 2003 and the method by which the firm distributes retained earnings. After the dividend tax decrease in 2003, the probability a dividend firm made an acquisition decreased by roughly 17.5 percent, from 32.5 percent to 26.8 percent. In contrast, there is no change in the probability share repurchase firms made an acquisition after 2003. Panel B reports the number of acquisitions per year a firm undertakes. Prior to 2003, dividend paying firms performed 0.51 mergers per year. After 2003, this number dropped nearly 29 percent to 0.36 mergers per year. In contrast, mergers per year for share repurchasing firms stayed nearly constant at 0.41. This evidence is consistent with firms that paid dividends before 2003 shifting
6.2 Merger Performance Analysis

Table 2 reports estimates from the difference-in-difference (DD) specification described in equation (6), where the dependent variable is the cumulative abnormal return twenty four months after an acquisition. The DD estimates are conducted on the full sample and on four subsamples to test the implications of the model. If the tax discount is important empirically, then after an acquisition, the performance of firms that pay a dividend should increase when the dividend tax rate is reduced in 2003. To control for other time varying aspects, the change in performance of dividend paying firms is compared with firms that repurchase shares. This effect should be larger for firms with a higher percentage of taxable shareholders, and therefore, a low percentage of institutional shareholders. The estimates from Table 2 support these hypotheses from the model and suggest the distortion from dividend taxation is empirically important.
The DD estimate in the full sample, specification (1), suggests the dividend tax cut in 2003 led dividend paying firms to perform 3.360 points better after an acquisition. This estimate of the policy effect is statistically significant at the ten percent level. The policy effect is larger for the subsets of firms that pay dividends and have a low percentage of institutional shareholders, specifications 2 and 4. In contrast, firms that have a large share of tax-exempt shareholders do not experience an increase in performance after the policy change, specifications 3 and 5. Statistically, the model rejects the null hypothesis that the DD estimates in specifications 2 and 3, and similarly in specifications 4 and 5, are the same at the five and one percent levels.

The policy effect demonstrated in Table 2 can be seen visually in Figure 2. Panels A and B graph the average performance after an acquisition across time for firms that pay dividends and firms that repurchase shares. Panel A demonstrates that the performance of dividend and share repurchase firms before 2003 look similar, but the performance of firms that pay a dividend start to outperform share repurchase firms after 2003. This pattern is more pronounced for firms with a low percentage of institutional shareholder, graphed in Panel B; consistent with the hypothesis that firms with a higher percentage of taxable shareholders will experience a larger effect. Panels A and B also demonstrate that the policy effect reported in Table 2 is consistent across years.

Panels C and D of Figure 2 extend the analysis of Table 2 to each month, between one and twenty-four months after an acquisition. Panel C demonstrates, for firms that repurchase shares, the post acquisition performance after 2003 is similar to, or slightly worse than, the performance before 2003. In contrast, Panel D demonstrates the post acquisition performance of firms that pay dividends is better after 2003, when the dividend tax rate was reduced. The patterns in Panels C and D are consistent for each month between one and twenty-four months after an acquisition.

---

26 The standard errors are two-way clustered as suggested by Petersen (2009). The results are robust to other clustering methods including not clustering and the a triple difference analysis reported in Table 3.

27 Panel C of Figure ?? in ?? finds firms with a high percentage of institutional holdings that were less affected by the policy change do not experience the same pattern of performance, acting as a placebo test.
Table 2: Difference-in-Difference Analysis

<table>
<thead>
<tr>
<th>Specification</th>
<th>Full Sample</th>
<th>&lt; 50th</th>
<th>&gt; 50th</th>
<th>&lt; 25th</th>
<th>&gt; 25th</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIV Firm</td>
<td></td>
<td>1.498</td>
<td>-2.400</td>
<td>0.121</td>
<td>-1.402</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.633)</td>
<td>(1.748)</td>
<td>(4.081)</td>
<td>(3.944)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.963)</td>
<td>(2.973)</td>
<td>(4.296)</td>
<td>(5.134)</td>
</tr>
<tr>
<td>Controls</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Year FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Equality Test</td>
<td></td>
<td>P = 0.032**</td>
<td>P = 0.004***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADJ. R-Square</td>
<td>0.172</td>
<td>0.183</td>
<td>0.070</td>
<td>0.218</td>
<td>0.109</td>
</tr>
<tr>
<td>Observations</td>
<td>2,634</td>
<td>1,139</td>
<td>1,187</td>
<td>547</td>
<td>1,779</td>
</tr>
</tbody>
</table>

Notes: Specifications (1) through (5) present coefficients from regressions from equation (6) with different samples. The dependent variable in all of the regressions is the cumulative abnormal return twenty-four months after an acquisition. The equality test measures whether the DD coefficient on the interaction (Low Tax x DIV Firm) is different across the institutional shareholder samples. All specifications include controls for marginal Q, financial distress, total assets, retained earnings, cash flow, and year fixed effects. Standard errors are reported in parentheses and are clustered at the acquiring firm and year level and are robust to heteroskedasticity. Statistical significance at the 1 percent level is denoted by ***, the 5 percent by **, and the 10 percent by *.
Notes: Panels A and B graph the twenty-four month mean cumulative abnormal return for acquisitions performed in years 2000-2008 for groups sorted by their distribution strategies, where group means are equalized in year 2001 to ease the comparison of trends then smoothed away from 2003 using a 3-year moving average. Panels C and D graph the evolution of cumulative abnormal returns for the twenty-four months after an acquisition is completed. Group averages are derived through the following procedure: cross-sectional regression of twenty-four month CAR on controls for firm size, retained earnings, and marginal Q are run in each year. Residual group means for the treatment and control group are then calculated and added to the mean investment percent for each year. All means are count weighted.
Table 3 presents coefficient estimates for the difference-in-difference-in-difference (DDD) model presented in equation (7). The coefficients of interest in all specifications are the DD coefficient (Low Tax x DIV Firm) and the DDD coefficient (Low Tax x DIV Firm x In Hold %). The DD coefficient estimate is the effect of lowering the dividend tax rate on the post-acquisition performance of firms that pay dividends and are not held by institutional shareholders. The sum of the DD and DDD coefficients is the effect of lowering the dividend tax rate on post acquisition performance of firms that pay dividends and are held 100 percent by institutional shareholders. The theory suggests that the DD coefficient should be positive and the sum of the DD and DDD coefficients should be zero because firms held by tax-exempt shareholders should not have been affected by the dividend tax decrease.

The DDD estimates suggests that high dividend taxation induces inefficient acquisition behavior, consistent with the model. Specification (1) reports that the decrease in dividend tax rate in 2003 led the post-acquisition performance to increase by 14.15 percentage points for firms that paid a dividend and were held exclusively by taxable shareholders. In contrast, there was no increase in post-acquisition performance for firms that paid a dividend but were exclusively held by tax-exempt shareholders (the DD plus DDD coefficients), as the theory predicts. These estimates are robust to excluding acquisitions in 2003 (Specification 2).

To put the DDD estimates in context, consider the magnitude of the change in performance predicted by the model. The change in the synergy threshold is given by the difference between the synergy level prior to 2003,

\[
\frac{1 - \tau_d}{1 - \tau_{cg}} = \frac{1 - 0.386}{1 - 0.2} = 0.7675
\]

and 1, the synergy level after 2003. The model predicts that this change in synergy levels in a simple model, where acquisition opportunities are uniformly distributed, would lead to an increase of almost 12 percent in post-acquisition performance. This estimate is remarkably similar to 14.15 DD estimate in the triple difference specification reported in Table 3.29

---

28 The sum of the DD and DDD coefficients is negative and is not statistically different than zero.

29 In the simple model with uniformly distributed acquisition opportunities the average acquisition quality is \((x + 0.7675)/2\) before 2003 and \((x + 1)/2\) after 2003. The increase in average performance is then given by \((x + 1)/2 - (x + 0.7625)/2 = 0.11875\).
Table 3: Difference in Difference in Difference Analysis

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Cumulative Abnormal Returns:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24 Month</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>DIV Firm</td>
<td>-2.674</td>
</tr>
<tr>
<td></td>
<td>(3.032)</td>
</tr>
<tr>
<td>In. Hold</td>
<td>-3.514</td>
</tr>
<tr>
<td></td>
<td>(2.145)</td>
</tr>
<tr>
<td>Low Tax × DIV Firm</td>
<td>14.154***</td>
</tr>
<tr>
<td></td>
<td>(5.382)</td>
</tr>
<tr>
<td>Low Tax × In. Hold</td>
<td>9.445***</td>
</tr>
<tr>
<td></td>
<td>(3.262)</td>
</tr>
<tr>
<td>DIV Firm × In. Hold</td>
<td>2.886</td>
</tr>
<tr>
<td></td>
<td>(2.436)</td>
</tr>
<tr>
<td>Low Tax × DIV Firm × In. Hold (DDD)</td>
<td>-19.768***</td>
</tr>
<tr>
<td></td>
<td>(6.488)</td>
</tr>
</tbody>
</table>

Controls ✓ ✓
Year FE ✓ ✓
2003 Mergers Excluded ✓
Adj. R-Square 0.148 0.145
Observations 2,326 2,102

Notes: Specifications (1) and (2) present coefficients from regressions from equation (7). The 24 month cumulative abnormal return is the dependent variable. All specifications include controls for marginal Q, financial distress, total assets, retained earnings, and cash flow and year fixed-effects. Specification (2) excludes mergers performed in the tax change year. Standard errors are robust to heteroskedasticity. *** indicates statistical significance at the 1% level, ** at 5%, and * at 10%.
Figure 3 graphs the performance of firms that repurchase shares (Panels A and C) and firms that pay a dividend (Panels B and D), each with high and low percentages of institutional shareholders (Panels A and B or Panels C and D respectively). Comparing these panels demonstrates that firms that pay a dividend and have a low percentage of institutional shareholders experience the largest increase in performance after the policy change. The other groups, depicted in Panels A-C, act as placebo tests and control for other possible factors that may affect performance differentially before and after 2003. These panels demonstrate that the average performance before and after 2003 is similar, or slightly better before 2003, for all months after an acquisition. In contrast, Panel D demonstrates an increase in post-acquisition performance for firms that pay a dividend and have a low percentage of institutional shareholders. This analysis reinforces the evidence in support of the tax discount model reported in Table 3.

### 6.3 Explaining the Post-Merger Performance Puzzle

This paper finds evidence consistent with existing studies of the post-merger performance puzzle (Franks, Harris, and Titman (1991), Agrawal and Madelker (1990), Agrawal, Jaffee, and Madelker (1992), Agrawal and Jaffee (2000), Loughran and Vish (1997), Myers (1984), Gregory (1997)) that the average firm under-performs by nearly 19 percent after an acquisition. The results demonstrate that the lowering of the dividend tax in 2003 improved the post-acquisition performance of dividend paying firms by roughly 17 percent of the post-merger performance puzzle (Table 2, Specification (1)). Given that roughly 52 percent of acquisitions are made by firms that pay a dividend (Table 1) the dividend tax explanation accounts for roughly 9 percent of the puzzle (17 percent for roughly 52 percent of the population). This is consistent with the evidence from the full sample that firms performance increased roughly 10 percent from -19 percent to -17 percent after 2003.\[^{30}\]

\[^{30}\]In extrapolating these numbers to the entire population, both the average institutional holding percentage and the percentage of firms that pay dividends must be considered. From Table 1, the average percentage of firms held by institutions varies only slightly by distribution type. Thus, the intensity of the treatment is similar across all firms and the baseline estimates may be extrapolated by considering only the breadth of the tax distortion.
Figure 3: Triple Difference Performance Analysis

(A) Repurchasing and High Institutional

(B) Dividend and High Institutional

(C) Repurchasing and Low Institutional

(D) Dividend and Low Institutional

Notes: Figure 3 graphs the evolution of cumulative abnormal returns for the twenty-four months after an acquisition is completed. Each panel presents CARs averaged over mergers performed prior to and after the 2003 tax rate decreases for groups of firms. Other determinants of abnormal returns are controlled for in the following manner: for each month post merger, the CAR is regressed on controls for marginal Q, cash flow, assets, retained earnings, and financial constraint. The residuals are then averaged across groups and by time period. Mean residuals are added to population averages in each post merger month. All means are count weighted.
7 Conclusion

This paper suggests that shareholder taxation distorts acquisition behavior of firms. Even in a model in which internal investment is undistorted by shareholder taxes, a tax discount on acquisitions may exist and induce inefficient acquisitions. The presence of this distortion is empirically tested by exploiting the heterogeneity of the discount across firms that pay dividends, repurchase shares, or issue equity prior to 2003 and the elimination of the discount by way of the 2003 reform.

The results of the empirical analysis support the predictions of the model. The magnitude of the empirical estimates suggest that, after the elimination of the tax differential in 2003, the post-acquisition performance increased by more than 14 percentage points for dividend paying firms. This result implies that prior to 2003, the tax wedge led these firms to engage in acquisitions in which newly acquired assets operated, on average, at only 86 percent efficiency. This suggests a large distortion due to shareholder taxation.

Since 2003, dividends and capital gains have been taxed at the same rate. However, as pressure has mounted to increase taxes on investment income, many economists and legislators have argued that given fixed revenue goals, increased dividend taxes, and therefore preferential treatment of capital gains, is the optimal solution. A dividend tax increase would address concerns of distortions due to the bias against equity financed investments by C corporations. However, our research suggests that any possible gains earned by increasing the dividend tax must be weighed against the distortion to acquisition behavior caused by the reintroduction of a tax discount on acquisitions. Given the magnitude of acquisition activity in the US economy - the first half of 2013 saw $437 billion in acquisition transactions - the cost of this distortion is large and must be considered in any discussion of changes to tax rates on distributions to shareholders.
References


Appendix A  Model Derivations and Extensions

Appendix A.1  Optimal Equity and Distribution Policies

The marginal value of issuing equity or paying a dividend is found by taking the first order conditions of the shareholder optimization. Optimally a firm will never simultaneously issue equity and distribute earnings because in this case the firm’s value could be increased by reducing both equity and distributions by $1, lowering its tax bill by $t_d r/(1 + r)$. The marginal change in firm value with respect to equity issuance,

\[ \frac{\partial V}{\partial E} = \frac{1}{1 + r} \left( (1 - \tau_i) f'(X + E^*) - r \right), \]

increases with increased profits, which are subject to taxation on distributions, and decreases by the rate of return \( r \) by which shareholders discount future repayments. In contrast, the marginal change in firm value with respect to distributions,

\[ \frac{\partial V}{\partial D} = \frac{1}{1 + r} \left( (1 - \tau_i) r - (1 - \tau_i) f'(X - D^*) \right), \]

increases with the rate of return \( r \) because shareholders prefer distributions today relative to distributions tomorrow. The marginal change in firm value with respect to distributions decreases due to lower investment. Distributions in both periods are subject to the tax on distributions.

Three cases emerge from these first order conditions. In the first case, the firm does not distribute retained earnings \( (D = 0) \), and equity issuances increase investment \( (E > 0) \) until the marginal benefit equals the marginal cost,

\[ (1 - \tau_i) f'(X + E) = r. \]

In the second case, the firm does not issue equity \( (E = 0) \), and distributions decrease investment \( (D > 0) \) until the marginal benefit equals the marginal cost

\[ f'(X - D) = r, \]

where the distribution tax does not distort investment decisions for firms making distributions but does for firms issuing equity. In the third case, the firm does not distribute earnings or issue equity \( (D = 0, E = 0) \), and the firm is at a kink point such that

\[ f'(X) \in \left( r, \frac{r}{1 - t_d} \right). \]

These three cases demonstrate that the determinant of the relevant case depends on the production technology and the amount of initial cash holding \( X \). For example, when the firm has a relatively large amount of initial cash, such that the slope of the net profit function is small \( f'(X) < r \), the firm will distribute earnings until the rate of return increases and exactly equals \( r \). In contrast, when the firm has a relatively small amount of initial cash, such that the slope of the net profit function is large \( (1 - t_d) f'(X) > r \), the firm will issue equity until the rate of return
Appendix A.2  The Value of the Firm with and without an Acquisition

The value of the firm making an acquisition is given by,

\[ V_M = (1 - \tau_i)D - E + \frac{(1 - \tau_d)[f(I) + \sigma(g(C) + C) + I - E] + E}{1 + r}, \]

where investment \( I = X + E - D - M \). Investment can be rewritten as a function of the equilibrium distributions and equity issuances without a merger \((D_0, E_0)\) where distributions with an acquisition are given by \( D = D_0 - M \), and equity issuances with an acquisition are given by \( E = E_0 + M \). Using this expression for investment, the value of the firm making an acquisition can be written as,

\[ V_M = (1 - \tau_i)D - E + \frac{(1 - \tau_d)[f(I) + \sigma(g(C) + C) + I - E] + E}{1 + r} = V + M(\sigma - \theta) - \sigma \tau_{cg} M + \theta \phi \tau_i M + \lambda \left( \frac{\theta - 1}{1 + r} + \tau_d \left( 1 - \theta \phi \right) \right) M \]

where \( \theta = \phi = 1 \) when the firm distributes earnings, \( \theta = 1 \) and \( \phi = 0 \) when the firm issues equity, \( \theta = 0 \) when the firm neither distributes earnings nor issues equity, and \( \lambda = 1 \) if equity used to finance the acquisition is returned to shareholders and \( \lambda = 0 \) otherwise. The text makes the assumption \( \lambda = 0 \), but the propositions 1 and 2 are robust to assuming \( \lambda = 1 \).

Firms with a Change in Policy

This section considers the distortions when the firms’ optimal distribution and equity issuance policy changes with an acquisition. There are three cases to consider: distribution paying firm without an acquisition issues equity with an acquisition, distribution paying firm without an acquisition does not distribute earnings or issues equity with an acquisition, and a firm that neither distributes earnings nor issues equity without an acquisition issues equity with an acquisition. In the first case, after making an acquisition the firm no longer pays a dividend and issues equity equal to \( E = M - D_0 \). In the third case, after making an acquisition the firm issues equity equal to \( E = M \).

\[
V_M - V = \begin{cases} 
M(\sigma(1 - \tau_{cg}) - (1 - \tau_i)), & \text{if } D > 0 \\
M(\sigma(1 - \tau_{cg}) - (1 - \lambda \tau_i/(1 + r))), & \text{if } E > 0 \\
M(\sigma(1 - \tau_{cg}) - (1 - \lambda(\tau_i + r)/(1 + r))), & \text{if } D = E = 0 
\end{cases}
\]

When the difference between the value of the firm with the acquisition and the firm without
an acquisition is positive the firm makes the acquisition. For each of these three cases this occurs when the synergy level of the acquisition opportunity is greater than,

$$\sigma > \begin{cases} 
\frac{M - \tau_i D_0 - \tau_i/(1 + r)(M - D_0)}{M(1 - \tau_{cg})} & \text{if } D > 0 \land M > D_0 \\
\frac{(1 - \tau_i)(M + rD_0)}{(1 + r)M(1 - \tau_{cg})} & \text{if } D > 0 \land M = D_0 \\
\frac{1 - \tau_i/(1 + r)}{1 - \tau_{cg}} & \text{if } D_0 = E_0 = 0 \land E = M.
\end{cases}$$

**Appendix A.3 Noncompetitive Market for Acquisitions**

This section loosens the assumption that the market for acquisitions is competitive such that the target firm is paid more than their value of the firm. Proposition 1 has to be updated such that firms that repurchase shares accept inefficient acquisitions but less inefficient acquisitions than firms that pay dividends. Further, the distortion for firms that repurchase shares does not depend on the dividend tax, and thus is unaffected by the dividend tax rate change in 2003. Now harmonizing the capital gains and dividend tax rates is not sufficient to eliminate the distortion to acquisition behavior for firms issuing equity, but setting both tax rates to zero is.

First, the acquiring firm pays the target firm’s shareholders a convex combination of the value of the firm to the target firm’s shareholders and the acquiring firm’s shareholders,

$$M = \psi V_A + (1 - \psi)V_T,$$

where $\theta$ is the exogenous bargaining power of the target firm such that $\psi \in (0, 1).$ The price at which the target firm’s shareholders are indifferent between selling and holding on the firm is given by

$$V_T = \frac{(1 - \tau_d)(g(C) + C)}{(1 - \tau_{cg})(1 + r)}.$$

Similarly, the value of the target firm to the acquiring firm is the present discounted value of the firm with the synergies $\sigma$ and discounted by the tax on distributions, $V_A = (1 - \tau_d)\sigma(g(c) + C)/(1 + r)$.

Second, given the new price the difference in value of the acquiring firm with and without the acquisition is given by,

$$V_M - V = \begin{cases} 
\frac{(1 - \tau_{cg})\sigma}{1 + \psi(\sigma(1 - \tau_{cg}) - 1)}M - M + \tau_iM, & \text{if } D > 0 \\
\frac{(1 - \tau_{cg})\sigma}{1 + \psi(\sigma(1 - \tau_{cg}) - 1)}M - M, & \text{if } E > 0.
\end{cases}$$

The threshold synergy value is determined by the values of $\sigma$ such that the difference in firm values is greater than zero,

$$\sigma \geq \begin{cases} 
\frac{(1 - \psi)(1 - \tau_i)}{(1 - \tau_{cg})(1 - \psi(1 - \tau_i)), & \text{if } D > 0 \\
\frac{(1 - \psi)}{(1 - \tau_{cg})(1 - \psi(1 - \tau_i))}, & \text{if } E > 0.
\end{cases}$$

**PROPOSITION 1b** When firm 1 has limited case (e.g. $r < f'(X)$) and $\psi \neq \tau_{cg}$ it does not
accept the socially efficient set of acquisitions. In contrast, when firm 1 has abundant cash (e.g. \( r > (1 - \tau)f'(X - M) \)) it over acquires if the tax rate on distributions is greater than the capital gains tax rate (e.g. \( \tau_i > \tau_{cg} \)) but acquires less socially inefficient acquisitions if the tax rate on distributions is greater than zero and equal to the capital gains tax rate (e.g. \( \tau_i = \tau_{cg} > 0 \)).

**PROOF:**

\[
\tau_d > \tau_{cg} \\
1 > \frac{(1 - \psi)(1 - \tau_d)}{(1 - \psi(1 - \tau_d))(1 - \tau_{cg})} = \sigma^T
\]

where \( \sigma^T \) is the threshold synergy value.

\[
\frac{d}{d\tau_i} \frac{(1 - \psi)(1 - \tau_i)}{(1 - \tau_{cg})(1 - \psi(1 - \tau_i))} < 0
\]

\( \tau_{cg} \neq \psi \Rightarrow 1 - \tau_{cg} \neq 1 - \psi \Rightarrow 1 \neq \frac{1 - \psi}{1 - \tau_{cg}} = \sigma \]

**Appendix A.4 Institutional Shareholders**

This section considers the incentives of firms to make acquisitions when the fraction \( \rho \) of the firm’s shareholders are institutional shareholders that are tax-exempt.

The value of the firm with shareholders that are tax-exempt is given by,

\[
V_\rho = \rho V_I + (1 - \rho)V \\
= (1 - (1 - \rho)\tau_i)D - E + \frac{(1 - (1 - \rho)\tau_i)(f(I) + I - E) + E}{1 + r}
\]

The difference in value for the firm with and without an acquisition is given by,

\[
V_M - V = \begin{cases} 
(1 - \tau_{cg})(1 - (1 - \rho)\tau_d)\sigma M - M + \tau_i M, & \text{if } D > 0 \\
1 - \tau_d \\
1 - \tau_d
\end{cases}
\]

The threshold synergy value is determined by the values of \( \sigma \) such that the difference in firm values is greater than zero,

\[
\sigma \geq \begin{cases} 
(1 - (1 - \rho)\tau_i)(1 - \tau_d) \\
(1 - \tau_{cg})(1 - (1 - \rho)\tau_d), & \text{if } D > 0 \\
1 - \tau_d \\
(1 - \tau_{cg})(1 - (1 - \rho)\tau_d), & \text{if } E > 0.
\end{cases}
\]

**PROPOSITION 2** When all shareholders of the acquiring firm are tax-exempt the distortion is the same for firms that pay dividend and firms that repurchase shares.

**PROOF:**
Set $\rho = 1$,

$$
\sigma \geq \begin{cases} 
1 - \tau_d, & \text{if } D > 0 \text{ and } \tau_i = \tau_d \\
1 - \tau_{cg}, & \text{if } D > 0 \text{ and } \tau_i = \tau_{cg}.
\end{cases}
$$

Proposition 2 demonstrates the difference in acquisition behavior for firms that pay dividends and repurchase shares goes to zero as the percent of shareholders that are tax-exempt goes to 100 percent. Proposition 2 provides an additional empirical test, using variation in institutional shareholders, to test the model.

**Appendix B  Common Trend Assumption**

Table 4 reports the empirical test of the common trend assumption for three samples: firms that pay dividends (Column 1), firms with a low percentage of institutional shareholders (Column 2), and firms that pay dividends and also have a low percentage of institutional shareholders (Column 3). Column (1) reports that none of the coefficients of the year fixed-effects interacted with the dividend firm indicator are statistically significant. In addition, the joint F-test fails to reject the null hypothesis that all of the coefficients are zero; and therefore, the common trend assumption, at the forty-five percent level. Similarly, columns (2) and (3) fail to reject the common trend assumption for firms with a low percentage of institutional shareholders and firms that pay dividends and also have a low percentage of institutional shareholders.

**Appendix B.1 Additional Empirical Evidence**

**Appendix B.2 Additional Merger Activity Analysis**

This appendix expands on the brief evidence reported in section 6.1. The empirical DD results, presented in Table 5 echo the graphical findings in Figure 1. Specifications (1) and (2) perform the regression counterpart to Panel (A) of Figure 1 without and then with controls variables. Specifications (3) and (4) perform the counterpart to Panel (B). Focusing on Specification (2), the estimated coefficient on the Low Tax x Div Only interaction suggests that after 2003, the merger probability of dividend paying firms decreased by 4.4 percentage points or by 13.5 percent relative to the change in merger activity by share repurchasing firms. The Specification (4) results suggest that mergers per year for dividend paying firms decreased by 0.115 or by 22.5 percent relative to share repurchasing firms. The Specification (4) results may also reinterpreted to say that after 2003, the time in between mergers for dividend paying firms increased by about 5 months relative to share repurchasing firms. Overall, the empirical analysis lines up very nicely with the graphical results and suggests that the 2003 tax rate equalization decreased the merger activity of dividend paying firms. From these results, we conclude that (1) prior to 2003 the dividend capital gains tax wedge induced tax motivated mergers and acquisitions for dividend paying firms and (2) approximately 20 percent of mergers and acquisitions by dividend paying firms prior to 2003 were tax motivated.

**Appendix B.2.1 Subsample Large, Small, and Cash Acquisitions**

This section reports the difference in difference estimates in different samples. The first comparison splits the sample into large and small acquisitions, calculated as the deal value (the amount the
Table 4: Test Common Trend Assumption

<table>
<thead>
<tr>
<th>Interaction Coefficient</th>
<th>Dividend</th>
<th>Low Inst. Hold %</th>
<th>Div x Low In. Hold %</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1(Year1999) \times 1(\text{Group}))</td>
<td>-4.148</td>
<td>-23.406</td>
<td>1.008</td>
</tr>
<tr>
<td>(5.079)</td>
<td>(15.604)</td>
<td>(19.135)</td>
<td></td>
</tr>
<tr>
<td>(1(Year2000) \times 1(\text{Group}))</td>
<td>1.636</td>
<td>-2.296</td>
<td>-1.057</td>
</tr>
<tr>
<td>(2.287)</td>
<td>(2.37)</td>
<td>(2.668)</td>
<td></td>
</tr>
<tr>
<td>(1(Year2001) \times 1(\text{Group}))</td>
<td>1.289</td>
<td>0.768</td>
<td>3.214</td>
</tr>
<tr>
<td>(2.421)</td>
<td>(2.195)</td>
<td>(2.871)</td>
<td></td>
</tr>
<tr>
<td>(1(Year2002) \times 1(\text{Group}))</td>
<td>4.653</td>
<td>1.053</td>
<td>1.614</td>
</tr>
<tr>
<td>(3.115)</td>
<td>(2.332)</td>
<td>(3.625)</td>
<td></td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.046</td>
<td>0.052</td>
<td>0.05</td>
</tr>
<tr>
<td>(\text{Observations})</td>
<td>976</td>
<td>758</td>
<td>758</td>
</tr>
<tr>
<td>(F\text{-stat})</td>
<td>0.92</td>
<td>0.88</td>
<td>0.40</td>
</tr>
<tr>
<td>(\text{d.f.})</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>(p\text{-value})</td>
<td>0.450</td>
<td>0.476</td>
<td>0.807</td>
</tr>
</tbody>
</table>

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: \(CAR_{i,24} = \beta_0 + \sum_k \delta_k \gamma_k + \sum_k \lambda_k \gamma_k d_G\) where \(d_G\) is an indicator variable for the group.
Table 5: DD Merger Activity Analysis

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Merger Probability</th>
<th>Mergers per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification</td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Div Firm</td>
<td>0.021</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Low Tax x Div Firm</td>
<td>-0.048***</td>
<td>-0.044***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Controls</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Adj. R-Square</td>
<td>0.014</td>
<td>0.016</td>
</tr>
<tr>
<td>Observations</td>
<td>16,251</td>
<td>16,251</td>
</tr>
</tbody>
</table>

Notes: Specifications (1) through (4) present coefficients from regressions from equation (6). In Specifications (1) and (2), the dependent variable is a firm-year indicator that is equal to 1 for years in which the firms performed a merger and 0 in years when no merger was performed. In Specifications (3) and (4), the dependent variable is the number of mergers performed by each firm in each year. Specifications (2) and (4) include firm level controls for assets, cash, sales, retained earnings, and book-to-market. These controls are averages over the 3 years preceding the firm-year dependent variable observation. All specifications include year fixed effects. Standard errors are reported in parentheses and are two-way clustered at the acquiring firm and year level and are robust to heteroskedasticity. Statistical significance at the 1 percent level is denoted by ***, the 5 percent by **, and the 10 percent by *.
The theory predicts that the average post-acquisition performance of a large acquisition will increase more after the policy change than the performance of a small acquisition because the tax discount is larger for larger acquisitions. The difference in difference estimate in the large sample is 7.126 and in the small sample is 1.115, consistent with the model's prediction. These point estimates are not statistically different from zero, possibly due to the decrease in sample size, but are statistically different at the 10.4 percent level.

The second comparison splits the sample based on whether the method of payment is cash or not. The theory predicts that cash deals by dividend paying firms would experience a larger increase in performance after the policy change than non-cash deals because the tax discount is larger for firms paying for the acquisition in cash, as opposed to equity. The difference in difference estimate in the sample of cash deals is 5.332 and in the non-cash sample is -2.227, consistent with the model's prediction. The difference in difference point estimate is statistically different from zero in the cash deals sample, but not in the non-cash deals, and the estimates are statistically different from each other at the 11.5 percent level.

These additional tests provide evidence that dividend taxation distorts acquisition behavior. Appendix C provides further evidence by testing an alternative explanation. The alternative explanation is that changes in shareholder monitoring could cause an increase in post-acquisition performance for firms paying a dividend after the policy change. The test uses variation in the percentage of the firm owned by the largest institutional shareholders, who have the most incentive to monitor the firm, and tests whether the performance of these firms increase after the policy change. The empirical tests find no evidence that changes in monitoring increased the performance of dividend paying firms after the policy change.

**Appendix B.2.2 Short-Run Cumulative Abnormal Returns**

Table 7 reports the DD analysis using short-run returns. Panel A reports the results for the acquiring firms, Panel B for target firms, and Panel C for joint target and acquiring returns. In contrast to long-run returns which measure the realized synergies of an acquisition, the short-run returns measure the initial evaluation of an acquisition based on the available information, mostly evaluations of assets. When there is a tax discount, acquiring firms are able to pay less for the assets trapped within a firm. Therefore, when the tax discount decreases, the acquisition appears to be less favorable based on assets, suggesting short-run returns for dividend paying firms will decrease after 2003.

The DD estimates for the acquiring firm, reported in Panel A, are negative and more negative in the subsamples with more taxable shareholders, consistent with the tax discount. The DD estimates for the target firms' returns, reported in Panel B, are not statistically significant. The DD estimates for the joint returns follow a similar, though sightly weaker, pattern as the returns for the acquiring firm.

**Appendix C Monitoring and Acquisition Behavior**

The model and the empirical evidence suggest the heterogeneity in returns before and after 2003 is due to changes in the dividend tax rate. This section tests an alternative hypothesis that the heterogeneity is due to changes in monitoring. Intuitively, firms that are more heavily monitored should perform better acquisitions. If the decrease in the dividend tax rate caused dividend paying...
Table 6: DD Analysis Different Samples

<table>
<thead>
<tr>
<th>Specification</th>
<th>Full Sample</th>
<th>Large</th>
<th>Small</th>
<th>Cash</th>
<th>Not Cash</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIV Firm</td>
<td>-1.498</td>
<td>-4.783</td>
<td>1.439</td>
<td>-4.075</td>
<td>2.294</td>
</tr>
<tr>
<td></td>
<td>(1.633)</td>
<td>(3.864)</td>
<td>(2.562)</td>
<td>(2.559)</td>
<td>(3.145)</td>
</tr>
<tr>
<td>Low Tax X DIV Firm</td>
<td>3.360*</td>
<td>7.126</td>
<td>1.115</td>
<td>5.332*</td>
<td>-2.227</td>
</tr>
<tr>
<td></td>
<td>(1.963)</td>
<td>(4.831)</td>
<td>(2.704)</td>
<td>(3.06 )</td>
<td>(3.732)</td>
</tr>
<tr>
<td>Controls</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Year FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Equality Test</td>
<td></td>
<td></td>
<td></td>
<td>P = 0.104</td>
<td>P = 0.115</td>
</tr>
<tr>
<td>Adj. R-Square</td>
<td>0.172</td>
<td>0.151</td>
<td>0.136</td>
<td>0.163</td>
<td>0.215</td>
</tr>
<tr>
<td>Observations</td>
<td>2,634</td>
<td>523</td>
<td>439</td>
<td>1,146</td>
<td>563</td>
</tr>
</tbody>
</table>

Notes: Specifications (1) through (5) present coefficients from regressions from equation (6) with different samples. The dependent variable in all of the regressions is the cumulative abnormal return twenty-four months after an acquisition. “Large” (“Small”) acquisitions are those with above (below) median values of deal value divided by aquirer’s assets. To construct Cash, the probability of a Cash acquisition is predicted by running a probit regression of cash on the value of the acquisition divided by the total assets of the acquiring firm. Cash is then predicted only for acquisitions in which deal method of payment is not available. All specifications include controls for marginal Q, financial distress, total assets, retained earnings, cash flow, and year fixed effects. Standard errors are reported in parenthesis and are clustered at the acquiring firm and year level and are robust to heteroskedasticity. Statistical significance at the 1 percent level is denoted by ***, the 5 percent by **, and the 10 percent by *.
Table 7: Short-Run DD Analysis: Acquirer CARs

<table>
<thead>
<tr>
<th>Specification</th>
<th>Full Sample</th>
<th>3 Day Acquirer Returns</th>
<th>Institutional Holding Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>(4)</td>
<td>(5)</td>
<td></td>
</tr>
<tr>
<td><strong>Panel A:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Div Firm</td>
<td>-0.001</td>
<td>-0.000</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td>-0.004</td>
<td>-0.009**</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Low Tax x Div Firm</td>
<td>-0.004</td>
<td>-0.009**</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Equality Test</td>
<td>P = 0.514</td>
<td>P = 0.048**</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5,274</td>
<td>3,348</td>
<td>1,926</td>
</tr>
<tr>
<td></td>
<td>1,557</td>
<td>3,717</td>
<td></td>
</tr>
<tr>
<td><strong>Panel B:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Div Firm</td>
<td>-0.028***</td>
<td>-0.078***</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.018)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Low Tax x Div Firm</td>
<td>0.032</td>
<td>0.055</td>
<td>-0.051</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.054)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>Equality Test</td>
<td>P = 0.289</td>
<td>P = 0.095*</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>328</td>
<td>228</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>143</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td><strong>Panel C:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Div Firm</td>
<td>-0.006</td>
<td>-0.008</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.017)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>Low Tax x Div Firm</td>
<td>-0.006</td>
<td>-0.001</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.024)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Equality Test</td>
<td>P = 0.715</td>
<td>P = 0.083*</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>395</td>
<td>267</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>107</td>
<td>288</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Specifications (1) through (5) present coefficients from regressions from equation (6) with different samples. The dependent variable in all of the regressions is the cumulative abnormal return 3 days after an acquisition. The equality test measures whether the DD coefficient on the interaction (Low Tax x DIV Firm) is different across the institutional shareholder samples. All specifications include controls for acquiring firm assets, cash flows, sales, and book to market, and year fixed effects. Standard errors are reported in parentheses and are clustered at the payout method and year level and are robust to heteroskedasticity. Statistical significance at the 1 percent level is denoted by ***, the 5 percent by **, and the 10 percent by *. 

46
Table 8: Largest Institutional Shareholder Heterogeneity Analysis

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>24 Month CARs</th>
<th>Large Shareholder Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification</td>
<td>&lt; 50th</td>
<td>&gt; 50th</td>
</tr>
<tr>
<td>DIV Firm</td>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>-4.469</td>
<td>0.958</td>
</tr>
<tr>
<td></td>
<td>(2.799)</td>
<td>(1.949)</td>
</tr>
<tr>
<td>LOW TAX X DIV Firm</td>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>1.430</td>
<td>3.191</td>
</tr>
<tr>
<td></td>
<td>(3.247)</td>
<td>(2.596)</td>
</tr>
<tr>
<td>Controls</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Year FE</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Equality Test</td>
<td>P = 0.230</td>
<td>P = 0.141</td>
</tr>
<tr>
<td>Adj. R-Square</td>
<td>0.096</td>
<td>0.187</td>
</tr>
<tr>
<td>Observations</td>
<td>1,134</td>
<td>1,192</td>
</tr>
</tbody>
</table>

Notes: Specifications (1) through (4) present coefficients from regressions from equation (6) with different samples. The dependent variable in all of the regressions is the cumulative abnormal return twenty-four months after an acquisition. The equality test measures whether the DD coefficient on the interaction (Low Tax x DIV Firm) is different across the largest institutional holdings subsamples. All specifications include controls for marginal Q, financial distress, total assets, retained earnings, cash flow, and year fixed effects. Standard errors are reported in parenthesis and are robust to heteroskedasticity. Statistical significance at the 1 percent level is denoted by ***, the 5 percent by **, and the 10 percent by *.

firms to be more heavily monitored, this could explain the results presented in the paper. To test this hypothesis, this section uses variation across firms in the percentage of the firm owned by the largest institutional shareholders who have the most incentive to monitor the firm.

The empirical specifications reported in Table 8 are similar to those in Table 2, except Table 8 splits the data based on firms that have a higher percentage owned by their largest investors. If monitoring is able to explain the empirical results, we would expect the specifications in Table 8 to be similar to the results in Table 2. In contrast, the estimates in Columns (2), (3), and (4) have the opposite sign as the benchmark model and the estimate in Column (1) is not statistically significant. These empirical results are inconsistent with the shareholder monitoring model but are consistent with the tax mechanism.