Measuring Private Benefits of Control from the Returns of Voting and Non-Voting Shares[#]

Ingolf Dittmann[§]

January 27, 2003

Abstract:

This paper proposes a new measure for estimating private benefits of control from data on dual-class companies. The proposed measure is the average return difference between non-voting shares and voting shares. We develop a theoretical framework that allows us to compare this new measure with the traditional measure, the relative price difference between voting and non-voting shares. It turns out that the price difference – but not the return difference – suffers from two sources of bias. The price difference will underestimate the value of control for firms which (1) pay higher dividends on non-voting shares than on voting shares or (2) are expected to discontinue their dual-class structure soon. We apply the two measures to data on German and US dual class companies and demonstrate that the price difference severely underestimates the value of control in the US relative to Germany. We conclude that private benefits of control do not differ significantly between these two countries.

JEL Classification Codes: G30, G34

[#] I would like to thank Robert Hauswald, Ernst Maug, Marco Pagano, Michael Rockinger, Kristian Rydqvist, Allan Timmermann, and seminar participants at Dortmund University for helpful discussions and comments. I also thank Christian Bassen for valuable research assistance.

[§] Humboldt-Universität zu Berlin, School of Business and Economics, Spandauer Str. 1, 10178 Berlin, Germany; e-mail: dittmann@wiwi.hu-berlin.de

1. Introduction

This paper examines how private benefits of controlling shareholders can be estimated from stock market data. We reconsider voting and non-voting shares of dual-class companies and derive a new measure for private benefits. This measure is based on the *returns* of the two types of shares, whereas the traditional measure is based on *prices*. We compare the two measures theoretically and empirically. It turns out that the traditional measure suffers from two sources of bias and that these biases lead to false conclusions about the relative size of private benefits in Germany and the US.

It has long been recognized that private benefits of control are reflected in the prices of voting and non-voting shares of dual-class companies (see Lease, McConnell and Mikkelson, 1983, and Levy, 1982). Since then, a sizable empirical literature has emerged that attempts to measure private benefits for different countries. The measure of private benefits used in this literature is the relative price difference between voting and non-voting shares or some variant thereof.[1]

This paper proposes an alternative method for measuring private benefits of control that is based on the average return difference between non-voting and voting shares. The intuition underlying this measure is that the observed returns of voting shares are calculated from prices and dividends but exclude realized benefits from having the right to vote. Hence, the observed return of voting shares underestimates the true return their holders receive. The 'opportunity cost' of holding a voting share is the return on non-voting shares, which is correctly reflected in market data. Hence, the difference between the observed return of non-voting shares and the observed return of voting shares is an estimate of the quantity by which the observed return underestimates the true return on voting shares.

The argument consists of three parts that are closely linked. First, we use a micro-structure model in order to establish a link between the blockholder's private benefits on the one hand and the prices of voting and non-voting shares on the other hand.[2]

[1] The relative price difference has first been proposed by Levy (1982). See Rydqvist (1992) for a review of the literature on dual-class shares.

[2] The model is in the veins of Kyle (1985) and Kyle and Vila (1991) and can be interpreted as an extention of the takeover-premium theory by Zingales (1995) and Rydqvist (1996).

The model shows that the price of voting shares exceeds the price of non-voting shares, because every holder of voting shares expects to make a trading gain from selling his shares in a later period. Although ex-post only some holders of voting shares will realize a trading gain, all of them expect to make such a gain with some probability, so that the expected gain can be treated like a dividend. We argue that this dividend is *not* recorded in stock market data, so that the average return of voting shares calculated from market data underestimates the true return the holders of voting shares expect. This is the key reason why the return difference is a viable measure of private benefits.

The second part of the paper extends the single-period micro-structure model into a dynamic asset pricing model that describes prices and returns of voting and non-voting shares over time. In this setting, we compare the two measures of private benefits with one another and show that the price difference is subject to two sources of bias. The first bias, which is well-known in the literature, is due to the sensitivity of the price difference to differences in dividends between the two types of shares. The second bias is caused by the dependence of the price difference on the stability of the dual-class structure. Ceteris paribus, the price difference yields lower estimates of the value of control if dividends on non-voting shares exceed dividends on voting shares, or if the company is expected to discontinue its dual-class structure soon. The return difference, on the other hand, does not depend on dividends or the stability of the dual-class structure.

In the third part, we apply the two measures to data on German and US dual-class companies. Consistent with the literature, we find a price difference of 4.4% for the US and 12.6% for Germany. In contrast, the return difference yields an estimated annual value of the voting right of about 0.9% for both, the US *and* Germany. Note that the values of the two measures cannot be compared directly, because the price difference estimates the present value of *all future gains* from having the right to vote, whereas the return difference is a measure of the value of the voting right *per period*. We show that the significant gap between the price difference estimate of 4.4% for the US and 12.6% for Germany can be explained by two effects. Variations in the liquidity of the two classes of shares explain 25% of this gap. Most of the remaining 75% is due to variations in the stability of the dual class structure: Dual class structures are discontinued on average every 13 years in the US compared to every 25 years in Germany. We also apply an approximate correction of the price difference for differences in dividends and for the expected duration of the dual-class structure. We find that the adjusted price difference – like the return difference – yields similar results for Germany and the US. We therefore conclude that private benefits of control do not differ significantly between Germany and the US.

A noteworthy by-product of our theory is the finding that non-voting shares have a higher market beta than voting shares. This result is corroborated in our empirical analysis: For both countries, the beta of non-voting shares exceeds the beta of voting shares by 0.06 on average. This result can be easily explained if the voting-share is seen as a portfolio containing one non-voting share (the dividend claim) and one paper confirming the holder's right to vote (the 'voting paper'). If the benefits from the voting paper are less correlated with the market than the dividends, the voting paper – and consequently the portfolio of voting paper and non-voting share– exhibits less systematic risk than the non-voting share alone.

The paper is organized as follows. The next section reviews the related literature. Section 3 presents the micro-structure model and shows how private benefits of a blockholder are reflected in prices and returns of voting and non-voting shares. Section 4 extends this model into a dynamic asset pricing model and derives the properties of the two measures and the link between them. Section 5 presents and describes the two datasets. Section 6 contains our empirical results, and Section 7 concludes. A complete list of all variables used in the paper as well as the proofs of the propositions can be found in the appendix.

2. Review of the literature

Private benefits of control have become an important ingredient of modern corporate governance theory. Grossman and Hart (1988) show that private benefits of incumbents and raiders can explain the existence of takeover contests in spite of the free-rider problem. Since then, many other papers drew heavily on private benefits in order to explain variations in ownership structure across firms and across countries.

The very nature of private benefits is that they are unobservable. Hence, they can be estimated only indirectly. A straight-forward way to do this is to compare the prices paid for blocks in trades among big investors with the prevailing market price at the time of the block trade. Barclay and Holderness (1989) find that the average premium paid in US block trades is 20%. Dyck and Zingales (2001) use data on 412 block trades from 39 countries for an international comparison of private benefits. They find an average block premium of 14% with wide variations across countries. Hanouna, Sarin and Shapiro (2001) measure the value of control as the difference between premia paid on blocks bought to obtain a majority position and blocks bought as a minority stake. They find for the G-7 countries that the value of control is higher in market-oriented economies (US, Canada, UK) than in bank-oriented countries (Japan, France, Italy, Germany).

The second, less obvious way to estimate private benefits employs data on dual-class companies which have issued two types of shares with similar (or equal) dividends but differences in their voting power. The standard theory underlying this approach has been developed by Zingales (1995) and Rydqvist (1996): In a takeover contest between the incumbent and the rival, prices for voting shares are bid up, so that holders of voting shares receive a takeover premium.[3] Section 3 of this paper presents an alternative model that explains the gains to holders of voting shares by small block acquisitions in an illiquid market. The main advantage of our model is that it does not rely on takeover contests which are rare or even non-existent in many countries.

There is a large empirical literature that estimates and analyses the relative price difference between voting and non-voting shares for various countries (see Rydqvist, 1992, for a review). Early papers by, among others, Levy (1982), Lease, McConnell and Mikkelsen (1983) and Megginson (1990) restrict their attention to dual-class companies which pay identical dividends on both classes of shares. Other papers (like Zingales, 1994, or Smith and Amoako-Adu, 1995) also consider dual class companies with systematic dividend differences. Nenova (2000) calculates a variant of the relative price difference for 660 dual class companies from 18 countries in order to compare private benefits of control across countries.

The evidence on the value of control in Germany relative to the US is mixed: Nenova's (2000) results suggest that private benefits are five times as large in Germany than in the US. Dyck and Zingales (2001) arrive at a similar result in their

---

[3] Bergström and Rydqvist (1992) present a related model in which the rival who bids for 100% of the company submits differential bids for voting and non-voting shares in order to pay a premium to the incumbent but not to small shareholders. In their model, differences in the shares' voting power are irrelevant.

analysis of block premia. Hanouna, Sarin and Shapiro (2001) establish, however, that extra premia paid for a *majority* block are smaller in Germany than in the US. Franks and Mayer (2001) point out that premia paid for blocks in Germany are much lower than target bid premia in US takeovers. They attribute the low value of control in Germany to the comparatively strong position of other stakeholders, like minority blockholders, banks, or employees. Our results are midway between the two camps: We find that the value of control is of similar magnitude in Germany and the US. To our knowledge, our paper is the first that attempts to estimate the value of control from returns rather than from prices.

3. The source of the value of a single corporate voting right: A micro-structure model of block acquisitions

This section presents a micro-structure model that describes the trading in voting and non-voting shares of a dual-class firm. There are two types of traders: On the one hand, there is a blockholder who potentially buys voting shares because she can derive private benefits from them. On the other hand, there are a number of potentially informed investors who might observe the quality of the firm and trade on this information. We first show that, in expectation, potentially informed investors gain on their superior information at the cost of the blockholder. We then describe how this expected gain is reflected in stock market data and finally propose two methods to estimate this gain from the data.

Intuitively, our argument goes as follows. We assume that, with some probability, a blockholder is able to obtain additional private benefits if she buys an additional block of voting shares in the market. This block acquisition leads to a temporary price increase of voting shares, because informed investors simultaneously trade small blocks of voting shares in the market in order to gain from their superior information. They can do this because market makers cannot distinguish between orders of informed traders and orders of blockholders which have no information content. Due to short-sell restrictions, potentially informed investors can only obtain the full expected value of their information if they already own voting shares when they receive their signal. Hence, they will buy voting shares in a previous period and, given that there is a large number of potentially informed investors, the price of voting shares will be higher than the price of non-voting shares – even though the terminal payoff is identical for both types of shares.

3.1 Set-up of the model

We consider one blockholder, N potentially informed investors, and an unspecified number of uninformed investors. All individuals are risk-neutral. The model consists of three stages. At time t = 1, non-voting shares and a part of the voting shares (the free-float) are sold to the investors. At time t = 2, trading takes place in voting and non-voting shares. At time t = 3, all shares are liquidated and receive the liquidation payment $\tilde{L}$ which takes on $\overline{L}$ and $\underline{L}$ ($\underline{L} < \overline{L}$) with equal probabilities. $\tilde{L}$ realizes before time t = 2 but remains unknown to all but one investor until time t = 3. Throughout the paper, symbols with tilde refer to random variables. Realizations of random variables are denoted by the same symbols without tilde.

Before time t = 2, the blockholder learns her private signal $\tilde{B}$, which is equal to $\overline{B} > 0$ with probability α and zero otherwise. $\tilde{B}$ is the additional monetary value of private benefits she can obtain at time t = 3 if she owns an additional block b of voting shares, where b is a constant.[4] We assume that $\tilde{B}$ and $\tilde{L}$ are independent, that $\alpha < ½$, and that $\overline{B} \geq (\overline{L} - \underline{L})b/2$.[5] Also before time t = 2, one of the N potentially informed investors observes $\tilde{L}$, and each of them is equally likely to make this observation.

At time t = 2, each investor (informed and uninformed) and the blockholder can buy or sell voting and/or non-voting shares. We assume that posting an order results in an infinitesimal cost, so that individuals who do not expect a gain from trading do not trade. Thus, only the informed investor and the blockholder will potentially trade. Let y_v and y_{nv} be the informed trader's demand for voting and non-voting shares, respectively. Positive numbers are buy orders and negative numbers sell orders. We assume that short-selling is not allowed, i.e., $-y_v$ and $-y_{nv}$ are bounded by the number of respective shares held by the investor. Let u_v and u_{nv} denote the corresponding

---

[4] We do not explicitly model how the opportunity to obtain private benefits arises. Zwiebel (1995) presents a model in which blockholders form coalitions in order to obtain and share private benefits. If such a coalition becomes unstable (e.g. due to the death of a pivotal blockholder) an outsider might want to buy a block in order to become member of a future coalition.

[5] The last two conditions prevent two tedious case distinctions which we will discuss towards the end of subsection 3.2.

orders of the blockholder. The market is organized as proposed by Kyle (1985): A market maker observes the total orders $y_v + u_v$ and $y_{nv} + u_{nv}$ but not the individual orders. He does know the structure of the game, but not the realizations of $\widetilde{B}$ or $\widetilde{L}$ before time $t = 3$. The market for market making is competitive, so the market maker will set a price equal to the expected liquidation value conditional on the two trading volumes observed: $P_2^v = P_2^{nv} = E(\widetilde{L} \mid y_v + u_v, y_{nv} + u_{nv})$

We assume that the informed trader can only order the quantities $b/2$, 0, or $-b/2$. Further, we assume that the number M of voting shares sold to the public at time $t = 1$ is a multiple of $b/2$: $M = m\, b/2$ with m being an integer, and that $m < N$, i.e. only some of the potentially informed investors can own a block of $b/2$.

3.2 Equilibrium prices and block acquisition premia

As the blockholder is also subject to the infinitesimal cost of posting an order, she will never place an order if $B = 0$. Likewise, she will never place an order for non-voting shares, because private benefits can only be obtained from voting shares. As a consequence, the informed investor will also never trade in non-voting shares, because every trade in non-voting shares would reveal the private signal to the market maker, who would adjust prices immediately.

Proposition 3.1: (Equilibrium at time $t = 2$)

a) The blockholder will
   - buy a block b of voting shares if she observes $B = \overline{B}$.
   - not trade if she observes $B = 0$.

b) The informed investor will
   - buy $b/2$ voting shares if he observes $L = \overline{L}$,
   - sell $b/2$ voting shares if he observes $L = \underline{L}$ and if he owns at least $b/2$ voting shares,
   - not trade if $L = \underline{L}$ and if he does not own at least $b/2$ voting shares.

There are two cases in which the informed investor will realize a trading gain: (1) if he sells $b/2$ voting shares and the blockholder orders b; and (2) if he orders $b/2$ and the blockholder orders nothing. In both cases, the total order is $b/2$, so the market maker cannot infer the informed investor's information from the total order. In case

(1), the blockholder will therefore pay an unfavourably high price and realize a trading loss. Since the market maker makes zero profits, the expected loss of the blockholder equals the expected gain of the informed trader.

Due to the assumed short-sell restriction, the informed investor can sell shares on a negative signal only if he already owns $b/2$ shares. If he does not own $b/2$ shares at time t = 2, he cannot gain from his superior information if $L = \underline{L}$. Therefore, all potentially informed investors will want to buy a block of $b/2$ voting shares at time t = 1. As there are (by assumption) more potentially informed investors than blocks $b/2$ in the free-float, the price of voting shares is bid up to the potentially informed investors' reservation utility that is higher than the expected liquidation value of the shares.

Proposition 3.2: At time t = 1, the price for non-voting shares is equal to the expected liquidation value: $P_1^{nv} = E(\tilde{L}) = \frac{1}{2}(\overline{L} + \underline{L})$. The price for voting shares is bid up to

$$P_1^v = P_1^{nv} + \frac{\alpha(1-\alpha)}{(1-\alpha)N + \alpha m}\frac{(\overline{L} - \underline{L})}{2}. \tag{1}$$

The price of voting shares strictly increases in the probability, α, that the blockholder can obtain additional private benefits. All voting shares are held by potentially informed investors.

Let $x = \frac{\alpha(1-\alpha)}{(1-\alpha)N + \alpha m}\frac{(\overline{L} - \underline{L})}{2}$. We will refer to x as the (expected) *block acquisition premium* in the remainder of this paper.

Proposition 3.2 shows that the price of voting shares is higher if block acquisitions are more likely. This result holds as long as $\alpha < ½$. If α exceeds some threshold above ½, the price of voting shares would decrease with increasing α. Also note that the trading activity of the informed trader translates into a cost to the blockholder if she buys a block b. If her private benefit $\overline{B}$ falls below a threshold below $(\overline{L} - \underline{L})b/2$, these costs are larger than $\overline{B}$, and she will never buy a block b.

Table 1 displays some values of the block acquisition premium, x, for different probabilities α and different amounts of private information $\overline{L} - \underline{L}$. The example assumes that the voting share free float is 40%, which is a typical value for the dual-class firms in the sample we consider in Section 6. In addition, we assume that there are N = 50 potentially informed traders, that the block size b = 2%, and that the value of the firm is normalized to one. The first line shows that a probability α = 1% and private information of ± 5% result in a block acquisition premium of x = 0.001%. If we stipulate that the modelled period is one trading day, the annual block acquisition premium is $250 \cdot x$ = 0.25%, because a year has approximately 250 trading days. An additional expected annual income of 0.25% discounted at a rate of 5.5% results in a relative price difference between voting and non-voting shares of 4.5%. Note that this is about the size of the relative price difference we observe in the US. Hence, the model can explain not only the existence but also the size of the price difference.

3.3 Returns of voting and non-voting shares

Due to the investors' risk-neutrality, the expected return of both types of shares between time t = 1 and t = 3 is zero:

$$E(\tilde{r}_{nv}) = \frac{E(\tilde{L})}{P_1^{nv}} - 1 = 0 \text{ and } E(\tilde{r}_v) = \frac{E(\tilde{L}) + x}{P_1^v} - 1 = 0. \tag{2}$$

Note that by assuming the existence of an alternative investment opportunity with a positive return, we could easily generate a positive return for both types of stock. The important point is that the expected return is equal for both types of shares.

Note, however, that stock market returns are typically calculated from end-of-period prices and beginning-of-period prices only: $r^* = P_3 / P_1 - 1$. Here, prices are adjusted for dividends and stock splits but they are *not* adjusted for block acquisition premia. In expectation, this observed return is therefore not equal for the two types of stock, because $P_1^{nv} = E(\tilde{L})$ but $P_1^v = E(\tilde{L}) + x$:

$$E(\tilde{r}_{nv}^*) = \frac{E(\tilde{P}_3^{nv})}{P_1^{nv}} - 1 = \frac{E(\tilde{L})}{P_1^{nv}} - 1 = 0, \tag{3}$$

but:

$$E(\tilde{r}_v^*) = \frac{E(P_3^v)}{P_1^v} - 1 = \frac{E(\tilde{L})}{P_1^v} - 1 = \frac{P_1^{nv} - P_1^v}{P_1^v} = -\frac{x}{P_1^v} \tag{4}$$

In the remainder of the paper, we will refer to $\tilde{r}$ as the *true* return and to $\tilde{r}^*$ as *observed* return. Equations (2) and (4) show that the expected *observed* return of voting shares is lower than their expected *true* return. The reason for this discrepancy is that the observed return does not incorporate the gain from trading at time t = 2: The observed return is equal to the true return net of the return from the block acquisition premium. Intuitively, each voting share is 'equal to' one non-voting share plus a lottery ticket with expected value x. The observed return ignores the payment which the winning ticket receives at time $t = 2$ and, therefore, underestimates the true return of voting shares.

Consider an outside observer of the game who never observes directly whether a block acquisition takes place or not. In our restricted model world, he can still deduce from the prices at time $t = 2$ and $t = 3$ whether a block acquisition occurred and whether one stockholder holding voting shares realized a block acquisition premium. Hence, this observer could calculate the realized true return r_v. In a richer environment, however, this is not feasible. If, for instance, additional information about the liquidation value becomes public at time $t = 2$ and the outside observer does not observe this information, he cannot deduce the actual block acquisition premium from the prices anymore. Hence, in a realistic environment, block acquisition premia are unobservable and not included in observed returns.

Proposition 3.3: From equations (1), (3) and (4), we obtain: $(P_1^v - P_1^{nv}) / P_1^v = x / P_1^v$ and $E(\tilde{r}_{nv}^*) - E(\tilde{r}_v^*) = x / P_1^v$. Hence, there are two ways to measure the expected block acquisition premium x / P_1^v:

(1) Calculate the relative price difference $(P_1^v - P_1^{nv}) / P_1^v$

(2) Observe a number n of independent outcomes of the game and calculate the return difference $\bar{r}_{nv}^* - \bar{r}_v^* = \frac{1}{n}\sum_{i=1}^{n} r_{nv,i}^* - \frac{1}{n}\sum_{i=1}^{n} r_{v,i}^*$

In our model world, the relative price difference will be the better choice to estimate the block acquisition premium, because only one observation is sufficient to calculate it exactly. In contrast, the return difference is exact only in expectation or asymptotically as n becomes large.

Clearly, the world is more complex than depicted in this simple microstructure model. In particular, there are more than only one period, there are risk-averse investors and there is a large variety of stocks among which investors can choose. Incorporating these features into the microstructure model is not a successful strategy, because the model would become too complex to be solved. Instead, we turn to a standard asset pricing model in the next section and incorporate the main insight of the micro model, namely the unobservable block acquisition premium, into this model.

4. Measuring the value of a single corporate voting right: Relative price difference vs. return difference

This section discusses three measures of the value of the voting right:

1. The relative price difference, *RPD* (see Proposition 4.1),
2. the return difference, *RD* (see Proposition 4.2), and
3. the return difference adjusted for differences in systematic risk between the two classes of shares, *ARD* (see Proposition 4.4).

In particular, we explore the relation between *RPD* and *ARD* and show that *RPD* - but not *ARD* - is sensitive to differences in dividends and, more importantly, to the stability of the dual-class structure.

We derive these results by incorporating the unobservable block acquisition premium into a standard single factor (CAPM) time series model.[6] In this formulation, the value of the voting right is equal to the present value of all future expected block acquisition premia. Since block acquisition premia are unobservable (as argued in the previous section), the challenge is to infer these premia indirectly from prices or returns of voting and non-voting shares.

4.1 The relative price difference

Let us consider a dual-class company with voting and non-voting shares. The non-voting share *NV* is entitled to a random dividend stream $\{\tilde{d}_t^{nv}\}_{t=1,2,...}$, where $\tilde{d}_t^{nv}$ denotes the dividend payments to a single non-voting share in period t. Likewise, the voting share *V* receives the dividend stream $\{\tilde{d}_t^{v}\}_{t=1,2,...}$. The two dividend streams need not be identical, although we implicitly assume that they are highly correlated. For the sake of simplicity, we call d_t 'dividend' although it also contains all other direct, pro-rata payments from company to shareholders, including rights issues, liquidation dividends and unification payoffs. Unification payoffs occur if the firm

[6] Our results can be easily generalized to more general factor models. Our choice of a simple single factor model is for expositional convenience only.

terminates the dual-class structure, e.g. in the course of a friendly merger or in order to make the share structure more appealing to investors.

In addition to the dividend stream $\{\tilde{d}_t^v\}_{t=1,2,\ldots}$, voting shares receive an income stream $\{\tilde{x}_t\}_{t=1,2,\ldots}$, where $\tilde{x}_t$ is the expected block acquisition premium in period *t*. The model in the previous section demonstrates how $\{\tilde{x}_t\}$ arises. We further assume that x_t is not recorded in standard stock market data, i.e., adjusted prices are not adjusted for x_t. This will be the driving assumption for our results.

Using the standard dividend discount model, we can readily calculate the price of a non-voting share at the end of period *t*:

$$P_t^{nv} = \sum_{\tau=1}^{\infty} \frac{E(\tilde{d}_{t+\tau}^{nv})}{(1+r^{nv})^{\tau}}, \qquad (5)$$

where r^{nv} is the appropriate discount factor. Accordingly, the market price of a voting share is given by:

$$P_t^v = P_t^{vd} + P_t^{vp}, \text{ where } P_t^{vd} = \sum_{\tau=1}^{\infty} \frac{E(\tilde{d}_{t+\tau}^{v})}{(1+r^{vd})^{\tau}} \text{ and } P_t^{vp} = \sum_{\tau=1}^{\infty} \frac{E(\tilde{x}_{t+\tau})}{(1+r^{vp})^{\tau}} \qquad (6)$$

P_t^{vd} is the present value of the voting share's dividend stream and P_t^{vp} is the present value of future expected block acquisition premia. The three income streams are discounted with individual discount rates, because they may have different systematic risk.

The predominant measure of the value of the voting right in the literature is the relative price difference, often called the voting premium:

$$RPD_t = \frac{P_t^v - P_t^{nv}}{P_t^v}. \qquad (7)$$

Proposition 4.1: The Relative Price Difference (7) is an unbiased measure of the proportion of P_t^v that is due to the value of expected block acquisition premia, P_t^{vp} / P_t^v, if and only if the present value of the dividend stream is identical for voting and non-voting shares.

Note that a sufficient condition for the unbiasedness of *RPD* is that dividends on voting and non-voting shares are identical. Although the statement of Proposition 4.1. is straightforward, it has been neglected in some empirical papers on dual-class companies with systematic dividend differences between the two classes of shares. For such companies *RPD* is a biased measure of the value of control.

4.2 The return difference

We now turn to the returns of the two types of shares. Clearly, the true returns of voting and non-voting shares over period t are $r_t^{nv} = (P_t^{nv} + d_t^{nv}) / P_{t-1}^{nv}$ and $r_t^{v} = (P_t^{v} + d_t^{v} + x_t) / P_{t-1}^{v}$, respectively. Under the assumption that the Capital Asset Pricing Model holds, we can write:

$$\begin{aligned} \tilde{r}_t^{nv} &= r^f + \beta^{nv}(\tilde{r}_t^m - r^f) + \tilde{\varepsilon}_t^{nv} \\ \tilde{r}_t^{v} &= r^f + \beta^{v}(\tilde{r}_t^m - r^f) + \tilde{\varepsilon}_t^{v}, \end{aligned} \tag{8}$$

where r^f is the risk-free rate and $\tilde{r}_t^m$ is the return on the market portfolio in period t. $\tilde{\varepsilon}_t^{nv}$ and $\tilde{\varepsilon}_t^{v}$ are zero-mean innovations which are uncorrelated with the market premium $\tilde{r}_t^m - r^f$. The return difference between non-voting shares and voting shares is:

$$\tilde{r}_t^{nv} - \tilde{r}_t^{v} = (\beta^{nv} - \beta^{v})(\tilde{r}_t^m - r^f) + \tilde{\varepsilon}_t, \tag{9}$$

with $\tilde{\varepsilon}_t = \tilde{\varepsilon}_t^{nv} - \tilde{\varepsilon}_t^{v}$.

The return difference in (9) is not operational, because, due to the unobservability of x_t, the return r_t^v cannot be calculated. What can be calculated is the observed return $r_t^{v*} = (P_t^v + d_t^v) / P_{t-1}^v = r_t^v - x_t / P_{t-1}^v$. Note that the observed return underestimates the true return for voting shares. For non-voting shares, the two are identical, i.e., $r_t^{nv*} = r_t^{nv}$. An operational form of the return difference then reads:

$$RD_t \equiv \tilde{r}_t^{nv*} - \tilde{r}_t^{v*} = \frac{\tilde{x}_t}{\tilde{P}_{t-1}^v} + (\beta^{nv} - \beta^{v})(\tilde{r}_t^m - r^f) + \tilde{\varepsilon}_t, \tag{10}$$

Proposition 4.2: If the betas of voting and non-voting shares are identical and $\{\tilde{x}_t / \tilde{P}_{t-1}^v\}_{t=1,2,...}$ is a stationary time series, the average return difference, $RD = \bar{r}_t^{nv*} - \bar{r}_t^{v*} = \frac{1}{T}\sum_{t=1}^{T}(r_t^{nv*} - r_t^{v*})$, is an unbiased estimator of the expected value of block acquisition premia per period, $E(\tilde{x}_t / \tilde{P}_{t-1}^v)$.

First note that the return difference measures the expected value of block acquisition premia *per period*, whereas the relative price difference (7) measures the *present value of all future* expected block acquisition premia. Only in a one-period world – like the microstructure model in the previous section – the two are identical. Please also note that Proposition 4.2 still holds if the present value of dividends differs between the two types of shares. Hence, the return difference is a meaningful measure of the value of the voting right if dividends are not identical.

At first glance, the assumption that the betas of voting and non-voting shares of one and the same company are identical seems intuitive. There are good reasons, however, why this might not be the case. As the notation in equation (6) suggests, we can interpret each voting share P_t^v as a portfolio of one dividend paper P_t^{vd} and one 'voting' paper P_t^{vp}. Then at any time *t*, the portfolio beta is a weighted sum of the betas of its constituents:

$$\beta^v = \frac{P_{t-1}^{vd}}{P_{t-1}^{v}}\beta^{vd} + \frac{P_{t-1}^{vp}}{P_{t-1}^{v}}\beta^{vp}, \tag{11}$$

Proposition 4.3: If dividends on voting and non-voting shares have the same systematic risk (i.e., $\beta^{vd} = \beta^{nv}$) and the income stream due to the voting right has less systematic risk than the dividend stream (i.e., $\beta^{vp} < \beta^{nv}$), then voting shares have less systematic risk than non-voting shares: $\beta^v < \beta^{nv}$.

The two additional assumptions made in Proposition 4.3 appear sensible: Usually dividends of voting and non-voting shares are highly correlated, so their systematic risk will be similar. Block acquisition premia, however, are infrequent high payments which are therefore unlikely to be strongly correlated with the market return. Therefore, under sensible assumption, we expect voting shares to be less exposed to market risk than non-voting shares.

If Proposition 4.3 holds, a major assumption of Proposition 4.2 is violated and the return difference is biased upwards. In order to remove this bias, we consider the *adjusted return difference*

$$ARD_t \equiv \tilde{r}_t^{nv*} - \tilde{r}_t^{v*} - (\beta^{nv} - \beta^v)(\tilde{r}_t^{m} - r^f) = \frac{\tilde{x}_t}{\tilde{P}_{t-1}^{v}} + \tilde{\varepsilon}_t. \tag{12}$$

Compared to the simple return difference RD_t in (10), ARD_t in (12) is adjusted for differences in risk between the two types of shares.

Proposition 4.4: If $\{\tilde{x}_t / \tilde{P}^v_{t-1}\}_{t=1,2,...}$ is a stationary time series that is uncorrelated with the market premium $\{\tilde{r}^m_t - r^f\}$, then $\hat{a}$ from the regression $r^{nv*}_t - r^{v*}_t = \hat{a} + \hat{b}(r^m_t - r^f) + \hat{\varepsilon}_t$ is an unbiased estimator of the expected value of block acquisition premia per period, $E(\tilde{x}_t / \tilde{P}^v_{t-1})$.

We will refer to $\hat{a}$ as the Adjusted Return Difference (*ARD*). The assumption that the block acquisition premium is uncorrelated with the market return is crucial for the result. If block acquisition premia are positively correlated with the market return, for instance because block acquisitions are more likely in bull markets than in bear markets, *ARD* underestimates $E(\tilde{x}_t / \tilde{P}^v_{t-1})$ whereas *RD* overestimates $E(\tilde{x}_t / \tilde{P}^v_{t-1})$. In this case, the two estimates can be interpreted as lower and upper bound of the true value.

4.3 The relation between RPD and ARD

The next proposition provides a direct link between *RPD* and *ARD* under two simplifying assumptions. The derivation of this result can be found in the Appendix.

Proposition 4.5: Assume in addition to the assumptions of Proposition 4.4 that expected dividends and expected block acquisition payoffs grow at a constant rate g: $E(d^{nv}_{t+\tau}) = E(d^{nv}_t)(1+g)^\tau$, $E(d^v_{t+\tau}) = E(d^v_t)(1+g)^\tau$, $x_{t+\tau} = x_t(1+g)^\tau$ and that the expected return of the two dividend streams is identical: $r = r^{vd} = r^{nv}$. Then we obtain:

$$RPD_t = \frac{E(ARD_t)}{r^f - g} - \frac{1+g}{r-g}\frac{(d^{nv}_t - d^v_t)}{P^v_t}. \quad (13)$$

Equation (13) contains the *expectation* of ARD_t, because the block acquisition premium is unobservable and ARD_t is equal to this premium only in expectation. In empirical applications, E(ARD_t) should be substituted by the estimate *ARD* described in Proposition 4.4. Note that in case of equal dividends, (13) reduces to

$RPD_t = E(ARD_t)/(r^f - g)$. If dividends on non-voting shares are higher than dividends on voting shares, RPD_t will become smaller, whereas $E(ARD_t)$ remains unaffected. This effect is demonstrated in Table 2A which displays values of *RPD* depending on the growth rate g and the dividend difference $(d_t^{nv} - d_t^v)/P_t^v$. Throughout the table, E(*ARD*), i.e., the value of the corporate voting right per period, is kept constant at 1%. The upper row shows, for instance, that the *RPD* is 18.2% if there are no dividend differences and if there is no growth. If the dividend difference increases to 1%, the *RPD* drops to 9.8%.

Finally note that up to now the whole analysis in this section assumes that the dual class structure lasts forever. In reality, however, firms can decide to convert their non-voting shares into voting shares and thereby discontinue the dual class structure. Proposition 4.6 takes this possibility into account:

Proposition 4.6: Assume in addition to the assumptions of Proposition 4.5 that the dual class structure will be discontinued after period *n*. Then we obtain:

$$RPD_t = c(r^f, g, n-t)\frac{E(ARD_t)}{1+g} - c(r, g, n-t)\frac{(d_t^{nv} - d_t^v)}{P_t^v} \qquad (14)$$

with the factor $c(r,g,n-t) = \left[1 - \left(\frac{1+g}{1+r}\right)^{n-t+1}\right] \Big/ \left[1 - \left(\frac{1+g}{1+r}\right)\right] - 1$.

Admittedly, the assumption that the time of the termination of the dual class structure is fixed and known is simplistic. Clearly, alternative (and more complicated!) expressions could be derived under more realistic assumptions. For our purposes, however, equation (14) is sufficient. Note that the factor $c(r,\ g,\ n-t)$ increases monotonically in $(n-t)$, the number of remaining years of the dual class structure. Hence, firms with an unstable dual class structure (i.e. small *n*) will have a lower *RPD* than firms with a stable dual class structure – even if the *ARD* does not differ between firms. As a consequence, the *RPD* yields spurious results in the comparison of two sets of firms, if the two sets differ in the stability of the dual class structure. The *ARD*, on the other hand, does not depend on the stability of the dual class structure, because the *ARD* measures the value of control (i.e. the block acquisition premium) *per period* rather than the present value of all future premia.

Table 2B illustrates the dependence of *RPD* on the stability of the dual-class structure. The table displays values of *RPD* from equation (14) depending on the growth rate *g* and the number of years, *n*, before the dual-class structure is discontinued. E(*ARD*) is kept constant at 1%, and we assume that there are no dividend differences. In the zero-growth scenario, the relative price difference decreases from 18.2%, if the dual-class structure persists forever, to 7.0%, if the structure will be terminated after 10 years. Also note that the sensitivity of *RPD* to changes in the stability of the dual-class structure increases as growth increases.

Proposition 4.6 assumes that the two classes of stock are treated equally in a stock unification. In practice, this is true for most stock unifications in the US and Germany. In some cases, however, voting shares receive more favourable terms than non-voting shares. Instead of including this effect explicitly into formula (14), we will treat this differential payoff as another 'unobservable' income component of voting shares: We estimate the last return difference over the period during which the stock unification was announced. If the terms are more favourable for voting shares, this return difference will be lower than otherwise. Thereby, benefits (or losses) from stock unifications are included in our estimate of the *ARD*. The *RPD* – if calculated before the announcement date – includes the expected benefits from stock unifications by construction, because it takes into account all future differential payoffs.

5. The Datasets

5.1 The US dataset

We searched Standard & Poor's "The Ticker Symbol Book 1997 Edition" for companies with at least two classes of shares outstanding. From the identified 101 companies, we obtained price, adjusted return, market value and volume series from Datastream. Information on the ownership structure, on dividends, and on charter provisions concerning the two classes' voting and conversion rights was compiled from SEC filings for all years available. 19 of the initially 101 companies were discarded: 6 of them because the series available from Datastream are shorter than one year; 8 are incorporated in a foreign country; for two companies the two classes do not differ in their voting right; in two cases the second class is a tracking stock; and for one firm, the voting share can receive higher dividends than non-voting shares. We also searched the SEC filings for announcements of events that effectively terminate the dual class structure. For 18 firms, the dual class structure was terminated in the course of a merger. Twelve companies converted non-voting shares into voting shares. In two cases, an outright tender offer is recorded, but in both cases the bid was identical for the two classes of shares. There was no takeover contest in our sample. Annual risk-free rates and equity premia have been obtained from Kenneth R. French's website.

5.2 The German dataset

Most of the German dataset has been constructed from Karlsruher Kapitalmarkt Datenbank (KKMDB), a scientific database that contains German stock market data from 1960 onwards. Dual class companies have been identified by their stock identification number: The first five digits identify the firm and the last digit the class of shares. For all these dual-class firms we compiled ownership information, the number of outstanding shares and charter provisions regarding voting power and dividend differences from *Handbuch der deutschen Aktiengesellschaften*, the German equivalent of Moody's Manual. We excluded firms for which (1) there are no differences in voting rights, (2) one or both classes are subject to trading restrictions or (3) we could not find any information on the voting arrangement. In addition, we excluded one company that unified its dual class structure twice within eight years only to issue new non-voting shares a few weeks later both times. No other company introduced new non-voting shares after unifying its dual-class structure. Price, return, dividend and volume data stem from KKMDB. For 50 of the 96 companies in our sample, one or both classes of shares ceased trading during our observation period. For these, we searched the leading German business papers *Handelsblatt*, *Börsenzeitung*, and *Frankfurter Allgemeine Zeitung* for the reason of the discontinuation and its first announcement. Two companies went bankrupt, 28 companies converted their non-voting shares into voting shares, one company bought back its non-voting shares, in three firms the majority shareholder issued a tender offer to minority shareholders of both classes, and in 9 cases the company was involved in a merger. In four of the nine merger cases, old non-voting shares were converted into new non-voting shares and old voting shares into new voting shares, i.e. the dual-class structure was not abandoned. For seven companies we could not find any information in the newspapers. The returns on the full Frankfurt market portfolio (DAFOX) come from KKMDB. For the risk-free rate we use the average one-month money market rate in Frankfurt which has been provided by the German central bank.

5.3 Description of the two datasets

For both datasets, we construct annual firm observations. Adjusted returns are calculated from data of the last day of December each year with non-zero trading volume for both types of shares. If there was no trading during the whole month, the two adjacent returns are set to a missing value. In years in which a stock unification was announced, the return is calculated from the beginning of the year until one month after the announcement. Analogously, unadjusted prices and (for the US dataset) the market value are taken from the last day of June with non-zero trading volume. The US dataset contains 980 firm year observations and spans the period from 1973 to 2001. The German dataset contains 1065 observations from 1960 to 2001. Note that the US dataset suffers from a survivorship bias by construction: All firms that introduced a second class of shares after about 1997 or unified their dual class structure (or went bankrupt) before about 1996 are not included in our dataset. The German dataset, on the other hand, does not suffer from such a bias: It contains all dual-class companies between 1960 and 2001.

The firms in our US dataset are very heterogeneous concerning their voting and dividend arrangements. 26 companies have a class voting arrangement, i.e., the non-voting shares elect a minority of directors (usually 25%, never more than 33%) whereas the majority of directors is determined by voting shareholders.[7] In these class voting arrangements non-voting shares usually have a 10% voting right on 'all other matters'. For 39 firms, non-voting shares have indeed no voting right and in 17 firms non-voting shares have a voting right which never exceeds 10%. The charter of 35 firms states that dividends on non-voting and voting shares must be identical. For 23 firms, non-voting shares *can* receive higher dividends than voting shares, and for 17 firms, non-voting shares *must* receive higher dividends than voting shares. For the latter group, the charter typically specifies a percentage (between 5% and 20%) by which the non-voting dividend must exceed the voting dividend. In practice, only 16

---

[7] As both classes of shares have some voting power, it would be more precise to call them 'superior-voting' and 'inferior-voting' shares. We nevertheless continue to use the terms 'voting' and 'non-voting shares'.

firms *did* pay higher dividends in at least one year. Note that nine companies that *must* pay higher dividends on non-voting shares never paid any dividends on any of their shares. For these companies, the 'must' provision seems to be a commitment device not to pay any dividends. For seven firms, we could not find any information on the dividend arrangement.

Due to legal requirements, German non-voting shares have much more homogeneous characteristics. All of them are entitled to a small minimum dividend (between 2% and 10% of the face value), and voting dividends must not be higher than non-voting dividends. Moreover, the charter of 71 firms specifies that non-voting dividends must exceed voting dividends by a fixed amount that varies between 1% and 4% of the face value. There are only two situations in which German non-voting shares have the right to vote: (1) if the minimum dividend has not been paid for two years or (2) in case of charter amendments that affect the dividend rights of non-voting shares.

Table 3 displays means and medians of a number of variables that describe the two datasets. The table shows that both samples are remarkably similar in terms of market capitalization and ownership structure: The average market capitalization is 2.1 billion dollars in the US sample and 1.9 billion euro in the German sample. Mean and median of the largest block of voting shares is around 50% in both samples and the sum of the largest three blocks of voting shares is about 60% on average in both samples. The median block of voting shares owned by US directors and officers is 52.5%, indicating that most US dual class firms are controlled by the management.[8] For the German sample, a comparable figure could not be obtained. In the US sample, non-voting shares have a much higher turnover than voting shares: On average, dollar turnover in non-voting shares is six times higher than dollar turnover in voting shares. In contrast, voting and non-voting shares are traded with equal frequency for the average German dual class company. The proportion of voting shares varies between 2.3% and 99.99% in the US with an average of 36.2%. In Germany, the law prohibits a capital structure with more than 50% non-voting shares. Here, the proportion of voting shares is on average 68.7% and varies between 50% and 97.5%. For 35 US

---

[8] For a more detailed description of managerial ownership in US dual-class companies see DeAngelo and DeAngelo (1985). In their sample, the median voting block controlled by directors and officers is 56.4%, i.e. somewhat higher than in our sample.

companies, the holders of voting shares have the right to convert their shares into non-voting shares. In Germany such a conversion right is implicitly prohibited because it might lead to more than 50% non-voting shares.

6. Empirical Results

6.1 The relative price difference RPD

Table 4 displays the relative price difference *RPD* as defined in (7) for the US and Germany. *RPD* is calculated for each company and each year on the last day of June on which non-zero trades are recorded for both types of stock. Observations after the announcement of a stock unification were deleted. If series ended within the sample and an announcement could not be found, the last year of observations was discarded. Table 4 shows that the average *RPD* in the US sample is 4.38% compared to 12.62% in the German sample. The difference is highly significant. These numbers are consistent with earlier results: Previous *RPD* estimates are 5.44% for the US (Lease, McConnell and Mikkelsen, 1983) and 26% for Germany (Hoffmann-Burchardi, 1999).[9] Nenova (2000) finds an average 'total vote value as share of firm value' of 2.0% for the US and 9.5% for Germany. Nenova's (2000) numbers are not directly comparable with our *RPD*, because her measure also depends on the number of shares outstanding for each of the two classes.

Table 4 also shows the average *RPD* for several subgroups depending on the dividend arrangements. If the dividend decision is exogenous, i.e., independent of other variables that determine the value of the voting right, the *RPD* must be smaller if dividends on non-voting shares are higher than on voting shares. This is the case in the US sample: Firms that always paid equal dividends according to their SEC filings display an average *RPD* of 5.0% compared to 1.9% for companies that did pay higher dividends on non-voting shares at least once. Likewise, there is a significantly lower *RPD* for companies with charters specifying that non-voting shares must receive higher dividends compared to companies where voting and non-voting shares receive equal dividends. Companies that *can* pay higher dividends on non-voting shares, however, have a higher average *RPD* than companies with equal dividends, although

---

[9] Both papers calculate the relative price difference with the price of the non-voting share in the denominator. Consequently, their estimates are somewhat larger than ours by construction.

this difference is not significant. This result might be interpreted as a first indication that dividends are not exogenous. An outright rejection of the exogeneity assumption comes from the German dataset. German companies that systematically pay higher dividends on non-voting shares have a significantly higher average *RPD* (14%) than firms that do not pay an additional dividend on non-voting shares (10%).

Table 5 contains more evidence on the endogeneity of the dividend arrangements. In the German dataset, the dividend dummy is strongly correlated with the combined size of the three largest blocks of voting shares.[10] For the US, we find that the MUST dummy[11] is positively correlated with the size of the largest voting block and negatively correlated with the size of the second largest voting block. The CAN dummy is positively correlated with the size of the second largest voting block. In addition, firms with a charter stating that non-voting dividends must be higher than voting dividends are significantly more likely to grant a conversion right to the holders of voting shares. It is not clear from Table 5, whether the dividend structure causes the ownership structure or vice versa, but it becomes apparent that the dividend dummies are correlated with variables that clearly have an influence on the value of the voting right. Therefore, controlling for differences in dividends is not a valid method to alleviate the sensitivity of *RPD* to differences in dividends.

6.2 The simple return difference (RD)

Table 6 presents the mean and median annual return difference for the full samples and for the same subsamples as Table 4. The return on non-voting shares exceeds the return on voting shares on average by 1.5% in the US and 1.2% in Germany. The difference between the two countries is not significant. The results vary across dividend subgroups but none of the differences is significant. A comparison of the significance levels in Tables 4 and 6 indicates that the signal-to-noise ratio is much lower for the *RD* estimates than for the *RPD* estimates.

---

[10] If we regress *RPD* on the additional dividend dummy and the three block size variables in Table 5B (not shown in the tables), the coefficient on the dummy still has the 'wrong' sign as in Table 4. Hence, the dividend dummy seems to proxy also for other or unobservable variables that determine the value of the voting right.

[11] See Table 3 for a definition of the MUST and the CAN dummy variables.

6.3 Liquidity adjustments

Amihud and Mendelson (1986, 1989) show theoretically and empirically that more liquid stocks have *ceteris paribus* higher prices and lower returns. Hence, the positive values for *RPD* and *RD* might both be due to differences in liquidity between the two types of shares. We measure liquidity based on the dollar trading volume, because we do not have data on bid-ask spreads. In particular, we use the following measure:

$$LR_t = \log(VO_t^v) - \log(VO_t^{nv}), \qquad (15)$$

where VO_t^v is the dollar volume of voting shares traded in year *t*. If we drop the logs and just use the ratio of the volumes, we obtain similar results but lower levels of significance. Alternative measures of liquidity could be based on the number of shares traded, the number of shares traded relative to the number of shares outstanding or the number of shares traded relative to the number of shares in the free float. The problem with liquidity ratios based on these measures is that they also depend on the relative number of voting shares or the ownership structure of the firm, so that we might run into another endogeneity problem. In order to find the most informative measure of liquidity, we ran regressions of simple returns (of voting shares and non-voting shares) on the different liquidity measures. The results are not shown in the tables. The log of the dollar volume turned out to be the most significant measure in these regressions in both datasets. Intuitively, this is sensible because investors are concerned about liquidity in *dollar* values rather than in number of shares or proportion of the free-float. Hence, we use the log of the dollar volume in (15).

The results of a regression of the *RPD* on the liquidity ratio *LR* are shown in Panel A of Table 7. For the US, the liquidity ratio has indeed a significant positive effect on the *RPD*. Comparing the estimate of the intercept (6.5%) with the average *RPD* from Table 4 (4.4%) reveals that the liquidity adjustment leads to a *higher RPD* estimate. The reason is that the liquidity ratio *LR* is negative for most companies as Table 3A displays. For most US companies, non-voting shares are more liquid than voting shares, so that the *RPD* in Table 4 is biased downwards.

In the German sample, *LR* is significantly *negatively* related to the *RPD*. If we include the voting ratio, i.e. the number of outstanding voting shares divided by the number of outstanding non-voting shares, into this regression, the liquidity ratio becomes insignificant and the voting ratio is highly significant. Hence, the liquidity ratio (15) proxies for the relative number of outstanding shares in the German dataset.

If we control for this effect, liquidity itself has no significant impact on the *RPD* in the German sample. On the other hand, the voting ratio is insignificant in the US sample and does not affect the estimate of the liquidity ratio. We conclude that liquidity has a significant effect on the *RPD* in the US but not in Germany. Adjusting for liquidity reduces the difference in the *RPD* between the US and Germany by 2.1 percentage points or by 25%.

Panel B of Table 7 shows regression results of the return difference on the liquidity ratio and the voting ratio. For the US data, neither ratio is significant which we attribute to the low signal-to-noise ratio in the return difference. In the German sample, the liquidity ratio is significantly negatively related to the return difference if the voting ratio is not included in the regression. As in the case of the *RPD* regressions, the liquidity ratio becomes insignificant if the voting ratio is included. Altogether, liquidity seems to have no effect on the return difference.

6.4 The adjusted return difference (ARD)

Now we turn to the adjusted return difference *ARD* by regressing the simple return difference *RD* on the market premium as Proposition 4.4 suggests. Thereby, we obtain an estimate of the expected annual block acquisition premium that is unaffected by differences in risk between the two types of shares. We perform two types of regressions. In the pooled regression, we do not allow for different specifications across companies. In the 'firm-specific beta difference' regression we allow the market premium coefficient to vary across firms but still restrict the intercept, i.e. the *ARD*, to be equal across all firms. The results are displayed in Table 8. Our first observation from this table is that the estimates of the beta difference is significant and around 0.06 for both countries. Thus Proposition 4.3 is verified which states that non-voting shares have higher systematic risk than voting shares.

The *ARD* estimates are much lower than the average *RD* from Table 6. For the US, the *ARD* is 0.95% and for Germany 0.81%. The difference is not significant. In fact, the two values are even not significantly different from zero. The comparatively high overall significance of the regression with 'firm-specific beta difference' for the German sample indicates that the beta difference varies a lot across firms. For the US sample, however, overall significance is poor. Here, the beta difference is rather homogeneous across companies.

6.5 The connection between *RPD* and *ARD*

Even though our *ARD* estimates are not significantly different from zero, they are reasonably robust across countries and regressions and the question arises, why the *RPD* suggests that the value of the voting right is higher in German than in the US, whereas the *RD* suggests that the value of the voting right is equal for the two countries. In this subsection, we will use Propositions 4.5 and 4.6 in order to explain the differing results. Both formulae (13) and (14) depend on four variables which we can estimate from our data: the risk-free rate, the return on non-voting shares, the voting dividend yield d_t^v / P_t^v and the non-voting dividend yield d_t^{nv} / P_t^v. The estimates are shown in the upper panel of Table 9. It is interesting to note that the dividend is larger in Germany (2.8%) than in the US (1.1%), although the dividend difference $d_t^{nv} / P_t^v - d_t^v / P_t^v$ is virtually identical (0.15%) for the two samples. In order to use equation (13) we need an assumption on the growth rate (of the dividends and the value of the voting right). We start our analysis by assuming zero growth for the time being. If we plug all these estimates, the zero-growth assumption and the estimated *ARD* from the pooled regressions in Table 8 into formula (13), we obtain an implied *RPD* of 16.2% for the US and 12.9% in Germany (columns 1 and 5 of Table 9). The values for the implied *ARD* are calculated in a similar way from the estimated *RPD* after solving equation (13) for E(ARD_t). For Germany, the implied *RPD* (12.9%) is consistent with the estimated *RPD* (12.6%). For the US sample, however, there is a large discrepancy. The *ARD* implies an *RPD* of 16.2% whereas the estimated *RPD* is only 6.5%. The discrepancy between the two estimates increases even further if positive growth rates are assumed (not shown in the table).

We now turn to formula (14) that also takes into account the duration of the dual-class structure. In the US sample, there are 379 observations from 1996 on, and 30 stock unifications. Therefore, US firms unify their shares (on average) every 12.6 years. In the German sample, there are 1,065 observations and 37 verified stock unifications. For seven companies, at least one of the two stocks disappeared and no information on the reason could be found. So the total number of stock unifications lies between 37 and 44, meaning that German dual-class companies discontinue their dual-class structure after 24.2 to 28.8 years. Columns (2), (3), (6) and (7) of Table 9 show the implied values of *RPD* and *ARD* from formula (14) using a duration of 12.6 years for US and 25 years for German dual-class structures. For Germany, we find a perfect fit of the model if we assume a growth rate of 3%. For the US, the implied *RPD* (under 3% growth: 8.9%) is still higher than the estimated *RPD* (6.5%) but the difference is much smaller than under the assumption of a permanent dual-class structure (column 1).

Note that the return *r* in equations (13) and (14) is the *expected* return on non-voting shares. For the implementations in columns (1) – (3) and (5) – (7) in Table 9, we use the *historical* return as an estimate for the expected return. There is growing evidence, however, that historical returns have been higher than expected returns in the second half of the 20th century – especially in the US. Claus and Thomas (2001) and Fama and French (2002) estimate equity premia between 3.5% and 4.5%. If we use the risk-free rate plus 4% equity premium instead of historical returns (columns (4) and (8)), the implied *RPD* falls to 8.65% for the US, i.e. the gap between implied and estimated *RPD* becomes even smaller.

As standard errors are small for the *RPD* estimates and high for the *ARD* estimates, the *implied ARD* values will be more reliable than the *estimated ARD* values – if the assumptions underlying equation (14) are correct. Note that the implied ARD for Germany (0.81%) is only slightly larger than the implied ARD for the US (0.74%). Hence, not only our *ARD* estimates but also the *RPD* estimates (once adjusted for liquidity, dividend differences and duration of the dual class structure) suggest that the value of a single corporate voting right does not differ between Germany and the US.

7. Conclusions and Further Notes

This paper compares two methods for the estimation of private benefits of control from stock market data on dual-class companies: the relative price difference and the return difference. The price difference measures the *present value* of the right to vote, whereas the return difference estimates the value of the right to vote *per period*. We show theoretically and empirically that the price difference is sensitive to differential dividends between voting and non-voting shares and to the stability (i.e. the expected duration) of the dual-class structure. The return difference, on the other hand, is by construction insensitive to dividend differences or to the stability of the dual-class structure.

We also present evidence that the differences in dividends between voting and non-voting shares are correlated with the ownership structure of the firm. Hence, the dividend decision is an endogenous variable. As a consequence, the common procedure to remove the sensitivity of the price difference to differential dividends by controlling for differential dividends is problematic and might yield misleading results.

In the empirical analysis, it turns out that the signal-to-noise ratio is much lower for price difference estimates than for return difference estimates. The theoretical reason is that the return difference is equal to the annual value of the voting right only in expectation, whereas the price difference is equal to the present value of the voting right with certainty. Hence, there is a trade-off between bias and variance when choosing between the two measures. The decision, which method to use, should be made depending on the dataset and the objective of the analysis. If the aim is to find determinants of the value of control for a single country, the price difference seems to be the more suitable measure. However, if two groups of companies are to be compared and the groups differ in terms of dividends or the stability of the dual-class structure, the return difference should be used. A third possibility is to use the price difference but to adjust it for the two types of biases along the line of Proposition 4.6. Note, however, that this proposition holds only under a number of simplifying assumptions that are unlikely to hold in general. In principal, adjustments under more realistic assumptions could be derived and used to estimate private benefits of control from the price difference.

The key idea of this paper is that holders of voting shares receive an expected gain from having the right to vote – in addition to any dividends and other pro-rata payments from the company to the shareholders. The micro-model in Section 3 demonstrates how such a gain can arise from block trades in illiquid markets. This model can be seen as a generalization of the existing takeover-premium theory by Zingales (1995) and Rydqvist (1996), because a takeover bid is an extreme form of a block acquisition that renders the market completely illiquid. Even though our model is more complicated than existing theory, it has at least two advantages. First, in case of a takeover, our model can explain a takeover premium without rivals. The role of the rival in the standard theory is to bid up the price. In our model, this role is performed indirectly by informed traders who might launch a takeover because they know that the company is undervalued. The second advantage of our model is that it does not depend on takeovers. Hence, it offers a possible explanation for price differences between voting and non-voting shares in countries where takeover contests are uncommon (like, for instance, in Germany) or even impossible.

Appendix

List of variables used

ARD_t	adjusted return difference over period *t*
b	number of shares the blockholder might potentially want to acquire (constant)
$\widetilde{B}$	blockholder's private benefit of owning an additional block *b* of voting shares with realizations 0 and $\overline{B} > 0$.
$\widetilde{d}_t^{\,v}$	dividend paid for each voting share at the end of period *t*
$\widetilde{d}_t^{\,nv}$	dividend paid for each non-voting share at the end of period *t*
g	terminal growth rate
$\widetilde{L}$	random liquidation payment with realizations $\overline{L}$ and $\underline{L}$ with $\underline{L} < \overline{L}$.
M	number of voting shares in the free-float
m	number of blocks of size *b*/2 in the free-float (i.e., $m = 2\,M\,/\,b$)
N	number of potentially informed investors
n	number of years until the dual-class structure is terminated
p	probability that blockholder orders *b* shares after observing $B = \overline{B}$
$\widetilde{P}_t$	price at the end of period *t* of voting share ($\widetilde{P}_t^{\,v}$) or non-voting share ($\widetilde{P}_t^{\,nv}$)
$\widetilde{r}_t^{\,v}$	*true* return of voting shares over period *t*
$\widetilde{r}_t^{\,nv}$	*true* return of non-voting shares over period *t*
$\widetilde{r}_t^{\,v*}$	*observed* return of voting shares over period *t* (net of block acquisition premia)
$\widetilde{r}_t^{\,m}$	return on the market portfolio over period *t*
r^f	risk-free rate
RD_t	return difference ($\widetilde{r}_t^{\,nv*} - \widetilde{r}_t^{\,v*}$)
RPD_t	relative price difference ($(P_t^v - P_t^{nv})/P_t^v$)
u_v, u_{nv}	blockholder's order of voting shares and non-voting shares, respectively
$\widetilde{x}_t$	block acquisition premium in period *t*
y_v, y_{nv}	informed trader's order of voting shares and non-voting shares, respectively
α	probability that $\widetilde{B} = \overline{B}$
β^v, β^{nv}	equity beta of voting shares and non-voting shares, respectively
β^{vd}	equity beta of the dividend claim of voting shares
β^{vp}	equity beta of the hypothetical voting paper (= voting share less dividend claim of voting share)
δ	probability that informed investor owns *b*/2 voting shares
$\widetilde{\varepsilon}_t$	zero-mean innovations

Proof of Proposition 3.1

We will first calculate the equilibrium price for voting shares at time $t = 2$ under the assumption that the two players stick to the strategies described in the proposition. After that we will show that deviating from these strategies is not optimal.

Let δ be the probability that the informed trader owns $b/2$ voting shares at time $t = 2$. Note that there are m blocks of size $b/2$ in the free-float. As voting shares trade above their liquidation value at time $t = 1$ (as we will shortly see), owning more than $b/2$ voting shares has no benefits but only costs. Therefore, $\delta = m/N$.

Depending on the realizations of $\tilde{B}$ and $\tilde{L}$ and on whether the informed trader owns voting shares, there are six different pairs of orders (u_v, y_v). The table below displays these six outcomes and the total order flow $u_v + y_v$ that is observed by the market maker together with the probabilities of these outcomes. The right column shows the realization of $\tilde{L}$ which the market maker will try to deduce from the total order flow.

		blockholder's order		**Realization of $\tilde{L}$**
		$u_v = 0$ (prob = $1 - \alpha$)	$u_v = b$ (prob = α)	
informed investor's order	$y_v = -b/2$ prob = $\delta/2$	$u_v + y_v = -b/2$ prob = ½ $(1 - \alpha)\,\delta$	$u_v + y_v = b/2$ prob = ½ $\alpha\delta$	$\underline{L}$
	$y_v = 0$ prob = $(1 - \delta)/2$	$u_v + y_v = 0$ prob = ½$(1 - \alpha)(1 - \delta)$	$u_v + y_v = b$ prob = ½ $\alpha(1 - \delta)$	$\underline{L}$
	$y_v = b/2$ prob = ½	$u_v + y_v = b/2$ prob = ½$(1 - \alpha)$	$u_v + y_v = 3b/2$ prob = ½ α	$\overline{L}$

In four of these six potential outcomes, the market maker immediately knows the signal $\tilde{L}$ and therefore chooses the respective prices: $P_2^v(-b/2) = P_2^v(0) = P_2^v(b) = \underline{L}$ and $P_2^v(3b/2) = \overline{L}$. Only if $u_v + y_v = b/2$, the market maker does not know the value of $\tilde{L}$. According to the probabilities in the table, he will set:

$$P_2^v(b/2) = E(\tilde{L} \mid u_v + y_v = b/2) = \frac{1-\alpha}{1-\alpha+\delta\alpha}\overline{L} + \frac{\delta\alpha}{1-\alpha+\delta\alpha}\underline{L} \qquad (16)$$

If the blockholder observes $B = \overline{B}$, she buys a block if the benefit $\overline{B}$ is larger than the expected cost of buying the block b, i.e., if:

$$\overline{B} \geq b\frac{\delta}{2}(P_2^v(b/2) - \underline{L}) = \frac{(\overline{L} - \underline{L})}{2} b \frac{1-\alpha}{1-\alpha+\delta\alpha}\delta \tag{17}$$

Under the maintained assumption that $\overline{B} \geq (\overline{L} - \underline{L})b/2$, condition (17) holds and the blockholder will buy a block whenever she observes $B = \overline{B}$.

Given the blockholder's trading strategy, it is straightforward to verify the optimality of the informed trader's strategy.

Proof of Proposition 3.2

A potentially informed investor who owns *b*/2 voting shares just after time *t* = 1 expects the following revenues from selling these shares at time *t* = 2 or liquidating them at time *t* = 3:

$$\left[\frac{N-1}{N}E(\widetilde{L}) + \frac{1}{N}\left(\frac{1}{2}\overline{L} + \frac{(1-\alpha)}{2}\underline{L} + \frac{\alpha}{2}P_2^v(b/2)\right)\right]\frac{b}{2} = \left[E(\widetilde{L}) + \frac{\alpha}{2N}\left(P_2^v(b/2) - \underline{L}\right)\right]\frac{b}{2}$$

$$= \left[E(\widetilde{L}) + \frac{\alpha(1-\alpha)}{(1-\alpha)N + \alpha m}\frac{\overline{L} - \underline{L}}{2}\right]\frac{b}{2}$$

This proves equation (1). It remains to show that P_1^v increases in α. Taking the first derivative of P_1^v with respect to α and equating this to zero yields the solution $\alpha = z - \sqrt{z^2 - z}$ with $z = N/(N\text{–}m)$. For $\alpha < z - \sqrt{z^2 - z}$, P_1^v strictly increases in α. As $z - \sqrt{z^2 - z} > 0.5$, P_1^v strictly increases in α if $\alpha < 0.5$.

Proof of Proposition 4.5

Due to the constant growth assumption and due to $r = r^{vd} = r^{nv}$, we obtain:

$$P_t^{nv} = \sum_{\tau=1}^{\infty}\frac{d_t^{nv}(1+g)^\tau}{(1+r)^\tau} = \frac{1+g}{r-g}d_t^{nv} \tag{18}$$

and
$$P_t^v = \sum_{\tau=1}^{\infty}\frac{d_t^v(1+g)^\tau}{(1+r)^\tau} + \sum_{\tau=1}^{\infty}\frac{x_t(1+g)^\tau}{(1+r^{vp})^\tau} = \frac{1+g}{r-g}d_t^v + \frac{1+g}{r^{vp}-g}x_t \tag{19}$$

As block acquisition premia are assumed to be uncorrelated with the market return, we obtain $r^{vp} = r^f$. Hence:

$$RPD_t = \frac{P_t^v - P_t^{nv}}{P_t^v} = \frac{1+g}{r^f - g}\frac{x_t}{P_t^v} + \frac{1+g}{r-g}\left(\frac{d_t^v}{P_t^v} - \frac{d_t^{nv}}{P_t^v}\right) \tag{20}$$

From equation (12), we get $E(ARD_t) = x_t / P_{t-1}^v = (1+g)x_t / P_t^v$, which we substitute into (19) in order to obtain equation (13).

Proof of Proposition 4.6

If the dual class structure is terminated after period n, we get:

$$P_t^{nv} = \sum_{\tau=1}^{n-t} \frac{d_t^{nv}(1+g)^\tau}{(1+r)^\tau} + \frac{P_n}{(1+r)^{n-t}} = c(r,g,n-t)d_t^{nv} + \frac{P_n}{(1+r)^{n-t}} \tag{21}$$

and

$$P_t^v = c(r,g,n-t)d_t^v + \frac{P_n}{(1+r)^{n-t}} + c(r^f,g,n-t)x_t \tag{22}$$

Here, P_n denotes the value of the unified stock at the end of period n. It follows that

$$RPD_t = \frac{P_t^v - P_t^{nv}}{P_t^v} = c(r,g,n-t)\frac{x_t}{P_t^v} + c(r^f,g,n-t)\left(\frac{d_t^v}{P_t^v} - \frac{d_t^{nv}}{P_t^v}\right) \tag{23}$$

Now, equation (14) follows immediately after substituting in $E(ARD_t) = (1+g)x_t / P_t^v$.

References

Amihud, Yakov and Haim Mendelson (1986): “Asset pricing and the bid-ask spread”, *Journal of Financial Economics* 17, 223-249.

Amihud, Yakov and Haim Mendelson (1989): “The effects of beta, bid-ask spread, residual risk, and size on stock returns”, *Journal of Finance* 44, 479-486.

Barclay, Michael J. and Clifford G. Holderness (1989): “Private benefits from control of public corporations”, *Journal of Financial Economics* 25, 371-395.

Bergström, Clas and Kristian Rydqvist (1992): “Differentiated bids for voting and restricted voting shares in public tender offers”, *Journal of Banking and Finance* 16, 97-114.

Claus, James and Jacob Thomas (2001): “Equity premia as low as three percent? Evidence from analysts’ earnings forecasts for domestic and international stock markets”, *Journal of Finance* 56, 1629-1666.

DeAngelo, Harry and Linda DeAngelo (1985): “Managerial ownership of voting rights: A study of public corporations with dual classes of common stock”, *Journal of Financial Economics* 14, 33-69.

Dyck, Alexander and Luigi Zingales (2001): “Private benefits of control: An international comparison”, NBER working paper 8711.

Fama, Eugene F. and Kenneth R. French (2002): “The equity premium”, *Journal of Finance* 57, 637-659.

Franks, Julian and Colin Mayer (2001): “Ownership and control of German corporations”, *Review of Financial Studies* 14, 943-977.

Grossman, Sanford J. and Oliver D. Hart (1988): “One share – one vote and the market for corporate control”, *Journal of Financial Economics* 20, 175-202.

Hanouna, Paul, Atulya Sarin and Alan C. Shapiro (2001): “Value of corporate control: Some international evidence”, working paper 01-4, University of Southern California.

Hoffmann-Burchardi, Ulrike (1999): “Corporate governance rules and the value of control – A study of German dual-class shares”, London School of Economics FMG discussion paper No. 315.

Kyle, Albert S. (1985): “Continuous auctions and insider trading”, *Econometrica* 53, 1315-1335.

Kyle, Albert S. and Jean-Luc Vila (1991): “Noise trading and takeovers”, *RAND Journal of Economics* 22, 54-71.

Lease, Ronald C., John J. McConnell and Wayne H. Mikkelson (1983): “The market value of control in publicly-traded corporations”, *Journal of Financial Economics* 11, 439-471.

Levy, Haim (1982): “Economic evaluation of voting power of common stock”, *Journal of Finance* 38, 79-93.

Megginson, William L. (1990): “Restricted voting stock, acquisition premiums, and the market value of corporate control”, *Financial Review* 25, 175-198.

Nenova, Tatiana (2000): “The value of corporate votes and control benefits: A cross-country analysis”, discussion paper, Harvard University.

Rydqvist, Kristian (1992): “Dual-class shares: A review”, *Oxford Review of Economic Policy* 8, 45-57.

Rydqvist, Kristian (1996): “Takeover bids and the relative prices of shares that differ in their voting rights”, *Journal of Banking and Finance* 20, 1407-1425.

Smith, Brian F. and Ben Amoako-Adu (1995): “Relative prices of dual class shares”, *Journal of Financial and Quantitative Analysis* 30, 223-238.

Zingales, Luigi (1994): “The value of the voting right: A study of the Milan stock exchange experience”, *Review of Financial Studies* 7, 125-148.

Zingales, Luigi (1995): “What determines the value of corporate votes?”, *Quarterly Journal of Economics* 110, 1047-1073.

Zwiebel, Jeffrey (1995): “Block investment and partial benefits of corporate control”, *Review of Economic Studies* 62, 161-185.

Table 1:

Block acquisition premium and relative price difference

This table displays values for the (daily) block acquisition premium x from equation (1) depending on the probability α of a block acquisition and the size of insider information $\overline{L} - \underline{L}$. The value of the firm is normalized to one: $(\overline{L} + \underline{L})/2 \equiv 1$. The number of potentially informed traders, N, is 50, the voting share free float is 40%, and the block size, b, is 2%. The annual block acquisition premium is equal to the daily premium multiplied by 250. The relative price difference is equal to the annual block acquisition premium divided by the risk-free rate of 5.5%.

α	$\overline{L} - \underline{L}$	block acquisition premium		relative price difference
		daily	annual	
1%	10%	0.0010%	0.248%	4.51%
2%	10%	0.0020%	0.492%	8.94%
3%	10%	0.0029%	0.732%	13.31%
1%	20%	0.0020%	0.496%	9.02%
2%	20%	0.0039%	0.984%	17.89%
3%	20%	0.0059%	1.464%	26.61%

Table 2:

Potential biases of *RPD*: A numerical example

Panel A displays values for *RPD* depending on the growth rate g and the dividend difference $(d_t^{nv} - d_t^{v})/P_t^{v}$, assuming that the dual-class structure is stable (from equation 13). Panel B shows values for *RPD* depending on the growth rate g and the duration of the dual-class structure, n, assuming that the dividend difference is zero (from equation 14). In both panels, the expected *ARD* is 1%, the risk-free rate, r^f, is 5.5% and the return on non-voting shares, r, is 12%.

Panel A: Sensitivity of *RPD* to dividend differences

growth g	dividend difference		
	0%	1%	2%
0%	18.2%	9.8%	1.5%
2%	28.6%	18.4%	8.2%
4%	66.7%	53.7%	40.7%

Panel B: Sensitivity of *RPD* to the stability of the dual-class structure

growth g	duration of dual-class structure n			
	∞	30	20	10
0%	18.2%	14.3%	11.6%	7.0%
2%	28.6%	17.8%	13.5%	7.5%
4%	66.7%	22.7%	15.9%	8.1%

Table 3:

Description of selected firm-specific variables

Panel A: US continuous variables

name	description	number of observations	mean	median	minimum	maximum
market value	yearly market values were averaged over years for each firm to generate one observation – in million dollar	82 firms	2,072.9	267.9	6.8	80,640.4
block1	size of largest voting block	469 firm-years	48.90%	50.00%	0.00%	93.90%
block2	size of second-largest voting block	469 firm-years	8.65%	7.78%	0.00%	38.00%
block1-3	size of the largest three blocks taken together	469 firm-years	61.24%	63.20%	0.00%	97.50%
officers	share of superior voting shares held by directors and officers	459 firm-years	49.82%	52.46%	0.00%	93.90%
liquidity ratio	log of the dollar turnover of voting shares minus the log of the dollar turnover of nonvoting shares	964 firm-years	-1.83	-1.64	-8.45	2.18
voting proportion	number of outstanding voting shares divided by the total number of outstanding (voting and non-voting) shares	964 firm-years	36.23%	37.11%	2.33%	99.99%
voting ratio	number of outstanding voting shares divided by the number of outstanding non-voting shares	964 firm-years	42.31	0.59	0.02	10,61
premium	difference between realized market premium and risk-free rate	29 years (1973-2001)	6.25%	6.25%	-35.94%	31.57%

Panel B: US dummy variables

name	description	number of observations	frequency of ones
DID	dummy variable that indicates whether the company paid higher dividends at least once according to SEC filings	82 firms	16
CAN	dummy variable that indicates whether the company's charter specifies that dividends can be higher on inferior-voting shares	82 firms	23
MUST	dummy variable that indicates whether the company's charter specifies that dividends must be higher on inferior-voting shares	82 firms	17
convertible	superior-voting shares are convertible into inferior-voting shares	82 firms	35

Panel C: German continuous variables

name	description	number of observations	mean	median	minimum	maximum
market value	yearly market values were averaged over years for each firm to generate one observation – in million euro	94 firms	1,932.5	168.5	3.4	48,736.4
block1	size of largest voting block	1047 firm-years	51.42%	51.00%	0.00%	100.00%
block2	size of second-largest voting block	1047 firm-years	5.99%	0.00%	0.00%	50.00%
block1-3	size of the largest three blocks taken together	1047 firm-years	58.47%	60.00%	0.00%	100.00%
liquidity ratio	log of the euro turnover of voting shares minus the log of the euro turnover of nonvoting shares	863 firm-years	0.10	-0.05	-3.41	5.32
voting proportion	number of outstanding voting shares divided by the total number of outstanding (voting and non-voting) shares	1025 firm-years	68.70%	66.67%	50.00%	97.50%
voting ratio	number of outstanding voting shares divided by the number of outstanding non-voting shares	1025 firm-years	3.93	2.00	1.00	39.00
premium	difference between realized market premium and risk-free rate	41 years (1960-2001)	4.40%	3.23%	-38.54%	70.00%

Panel D: German dummy variable

name	description	number of observations	frequency of ones
additional dividend	dummy variable that is one if the company's charter provides that inferior voting shares receive an additional dividend	96 firms	71

Table 4:

The relative price difference (RPD)

Panel A displays location measures for the relative price difference $RPD_t = (P_t^v - P_t^{nv}) / P_t^v$ for different subsamples. RPD_t has been calculated on June 30th (or the closest earlier day in June for which both types of stock were traded) for each company year in the sample. For each group, the table displays the mean and median *RPD*, the p-value of the two-sided t-test for zero mean, and the p-value of the Wilcoxon signed rank test for zero median. Panel B displays p-values of the two-sample t-test for equal mean and the two-sample Wilcoxon signed rank test for equal median.

Panel A: Location measures

group description	number of observations	mean	p-value of t-test	median	Wilcoxon p-value
All US observations	980	4.38%	< 0.0001	2.15%	< 0.0001
dividends on both types of shares must be equal (A)	391	4.57%	< 0.0001	2.19%	< 0.0001
dividends on inferior-voting shares can be higher (B)	291	5.05%	< 0.0001	1.92%	< 0.0001
dividends on inferior-voting shares must be higher (C)	228	1.63%	0.0096	1.23%	< 0.0001
firms which always paid equal dividends (D)	779	5.03%	< 0.0001	2.26%	< 0.0001
firms which paid higher dividends on inferior-voting shares (E)	201	1.87%	0.0124	1.69%	< 0.0001
All German observations	1065	12.62%	< 0.0001	11.77%	< 0.0001
without additional dividend	342	10.05%	< 0.0001	8.63%	< 0.0001
with additional dividend	723	13.84%	< 0.0001	13.36%	< 0.0001

Panel B: Pairwise comparisons

	p-value of t-test	Wilcoxon p-value
all observations: US versus Germany	< 0.0001	< 0.0001
US: equal vs. higher actual dividends (D vs. E)	0.0003	0.0218
must be equal vs. must be higher (A vs. C)	0.0007	0.0281
must be equal vs. can be higher (A vs. B)	0.5766	0.9180
Germany: with versus without additional dividend	0.0002	0.0002

Table 5:

Correlations between dividend dummies and other firm-specific variables

The table shows correlations between the dividend dummies and a number of variables that are described in Table 3. First, for each company in the sample, the mean of the annual observations of the individual variables was calculated. The table displays the Pearson correlations between these means and the dummy variables across firms. P-values of the t-test for zero correlation are shown in parentheses.

Panel A: Correlations for the US dataset

variables	dividend dummy variables		
	DID	**CAN**	**MUST**
block1	15.50% (0.1644)	-2.78% (0.8043)	19.71% (0.0760)
block2	2.21% (0.8439)	23.99% (0.0299)	-22.41% (0.0430)
block1-3	16.44% (0.1399)	9.77% (0.3825)	-13.86% (0.2142)
officers	19.27% (0.0829)	11.35% (0.3098)	7.51% (0.5025)
convertible	16.89% (0.1341)	17.86% (0.1130)	26.45% (0.0177)

Panel B: Correlations for the German dataset

variables	additional dividend
block1	16.21% (0.1147)
block2	19.62% (0.0554)
block1-3	26.28% (0.0097)

Table 6:

The return difference (RD)

Panel A displays location measures for the return difference $RD_t \equiv r_t^{nv*} - r_t^{v*}$ for different subsamples. RD_t has been calculated for each company and each calendar year in the sample. For each group, the table displays the mean and median *RD*, the p-value of the two-sided t-test for zero mean, and the p-value of the Wilcoxon signed rank test for zero median. Panel B displays p-values of the two-sample t-test for equal mean and the two-sample Wilcoxon signed rank test for equal median.

Panel A: Location measures

group description	number of observations	mean	p-value of t-test	median	Wilcoxon p-value
All US observations	947	1.54%	0.0046	0.55%	0.0140
dividends on both types of shares must be equal (A)	379	1.64%	0.0795	0.39%	0.1883
dividends on inferior-voting shares can be higher (B)	279	0.92%	0.2886	0.64%	0.0896
dividends on inferior-voting shares must be higher (C)	219	2.05%	0.0482	0.71%	0.2186
firms which always paid equal dividends (D)	752	1.12%	0.0684	0.50%	0.1193
firms which paid higher dividends on inferior-voting shares (E)	195	3.16%	0.0065	0.87%	0.0172
All German observations	1008	1.20%	0.0495	0.79%	0.0073
without additional dividend	321	0.43%	0.7496	0.01%	0.7496
with additional dividend	687	1.55%	0.0144	1.18%	0.0074

Panel B: Pairwise comparisons

	p-value of t-test	Wilcoxon p-value
all observations: US versus Germany	0.6739	0.6393
US: equal vs. higher actual dividends (D vs. E)	0.1288	0.1662
must be equal vs. must be higher (A vs. C)	0.7788	0.7920
must be equal vs. can be higher (A vs. B)	0.5843	0.7145
Germany: with versus without additional dividend	0.3894	0.2740

Table 7:

Liquidity adjustments for *RPD* and *RD*

Panel A shows regression results of simple OLS regressions of RPD on the liquidity ratio and the voting ratio. The liquidity ratio is the log of the dollar turnover of voting shares minus the log of the dollar turnover of nonvoting shares. The dollar turnover is calculated as the total number of shares traded during the calendar year times the unadjusted stock price at the end of June. The voting ratio is the number of outstanding voting shares divided by the number of outstanding non-voting shares. Panel B shows similar results for *RD* as the dependent variable. Standard errors are shown in parentheses. ***, **, *, ° indicate results that are significant at the 1%, 5%, 10% and 15% level, respectively.

Panel A: Regression results of *RPD* on liquidity and voting ratios

dataset	number of observations	intercept	liquidity ratio	voting ratio	adjusted R^2
US data	980	6.464%*** (0.487%)	1.143%*** (0.190%)		0.0349
US data	980	6.490%*** (0.489%)	1.150%*** (0.190%)	-0.0003% (0.0006%)	0.0342
German data	863	12.792%*** (0.520%)	-0.909%** (0.364%)		0.0060
German data	826	14.405%*** (0.751%)	-0.136% (0.487%)	-0.372%*** (0.132%)	0.0163

Panel B: Regression results of *RD* on liquidity and voting ratios

dataset	number of observations	intercept	liquidity ratio	voting ratio	adjusted R^2
US data	947	1.229%° (0.773%)	-0.168% (0.298%)		-0.0007
US data	947	1.226%° (0.778%)	-0.169% (0.299%)	0.00003% (0.00086%)	-0.0018
German data	822	1.528%** (0.709%)	-1.003%** (0.505%)		0.0036
German data	790	1.418% (1.014%)	-1.041%° (0.663%)	-0.004% (0.175%)	0.0026

Table 8:

The adjusted return difference (*ARD*)

The table displays regression results of *RD* on the market premium, which is the difference between market return and risk-free rate. The intercept is labelled 'ARD' and the slope coefficient is labelled 'beta difference'. 'Pooled estimation' refers to a simple OLS regression in which intercept and slope coefficients are equal for all firms. In the 'firm-specific beta difference' estimation method, only the intercept is equal for all firms; the slope is estimated for each firm separately. Here, we only report the average slope coefficient. Standard errors are shown in parentheses. ***, **, *, ° indicate results that are significant at the 1%, 5%, 10% and 15% level, respectively.

dataset	**estimation method**	**ARD (intercept)**	**beta difference**	**average beta difference**	**adjusted R^2**	**p-value of model F-test**
US (all observations)	pooled	0.945%° (0.641%)	0.0610* (0.0353)		0.0021	0.0838
US (all 82 firms)	firm-specific beta difference	1.034%° (0.643%)		0.0844** (0.0392)	0.0124	0.1879
US (73 firms with at least 5 obs.)	firm-specific beta difference	1.147%* (0.646%)		0.0589° (0.0374)	0.0131	0.1680
Germany (all observations)	pooled	0.805 (0.638)	0.0533** (0.0267)		0.0030	0.0461
Germany (all 92 firms)	firm-specific beta difference	1.116%* (0.638%)		0.1076** (0.0500)	0.0482	0.0011
Germany (70 firms with at least 5 obs.)	firm-specific beta difference	0.928% (0.653%)		0.0672* (0.0355)	0.0316	0.0128

Table 9:

Connection between *RPD* and *ARD*

The upper panel of the table contains the sample averages of six variables. The voting dividend yield is d_t^v / P_t^v, the non-voting dividend is d_t^{nv} / P_t^v. The dividends d_t include payoffs from rights issues. The value for *RPD* is taken from Table 7 for the US and from Table 4 for Germany. The *ARD* has been taken from the pooled regression in Table 8. The middle panel displays the maintained assumptions on the growth rate *g*, on the duration of the dual class structure *n* (in years) and on the expected return on non-voting shares. 'Historical return' refers to the sample mean of returns of non-voting shares shown in the upper panel of the table. The lower panel shows the results for *RPD* and *ARD* if the values of the upper two panels are plugged into formula (13) in case of infinite duration and formula (14) in case of finite duration. For the *ARD* results, the two formulae are first solved for E(ARD_t) before plugging in the estimated and assumed values.

	USA				Germany			
Estimated inputs (mean in dataset)								
risk-free rate	5.49%				5.59%			
return on non-voting shares	13.94%				10.05%			
voting dividend yield	1.03%				2.74%			
non-voting dividend yield	1.18%				2.89%			
RPD	6.46%				12.62%			
ARD	0.95%				0.81%			
Assumptions	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
growth rate	0.00%	0.00%	3.00%	3.00%	0.00%	0.00%	3.00%	3.00%
duration of dual class structure	∞	12.6	12.6	12.6	∞	25.0	25.0	25.0
choice of expected return *r*	historical return	historical return	historical return	4% + risk-free rate	historical return	historical return	historical return	4% + risk-free rate
Implied values								
RPD	16.18%	7.61%	8.90%	8.65%	12.88%	9.32%	12.57%	12.50%
ARD	0.41%	0.82%	0.71%	0.74%	0.79%	1.05%	0.81%	0.81%