Price and Probability: Decomposing the Takeover Effects of Anti-Takeover Provisions

VICENTE CUÑAT, MIREIA GINÉ, and MARIA GUADALUPE[*]

ABSTRACT

We study the effects of anti-takeover provisions (ATPs) on the takeover probability, the takeover premium, and target selection. Voting to remove an ATP increases both the takeover probability and the takeover premium, that is, there is no evidence of a trade-off between premiums and takeover probabilities. We provide causal estimates based on shareholder proposals to remove ATPs and address the endogenous selection of targets through bounding techniques. The positive premium effect in less protected firms is driven by better bidder-target matching and merger synergies.

ANTI-TAKEOVER PROVISIONS (ATPS) SUCH as staggered boards, dual-class shares, poison pills, and similar governance mechanisms have been found to affect firm value (e.g., Gompers, Ishii, and Metrick (2003), Cuñat, Giné, and Guadalupe (2012)).[1] In the context of the takeover debate, proponents argue that ATPs create value by allowing managers to negotiate a higher price in the event of a hostile bid and encouraging more long-term investment (Stein (1988), Harris (1990)). However, they may also reduce or delay the possibility of a takeover (Ryngaert (1988), Pound (1987), Malatesta and Walkling (1988), Comment and Schwert (1995), Karpoff, Schonlau, and Wehrly (2017)). The

[1] Reducing the threat of a takeover can destroy value by weakening managerial discipline (Scharfstein (1988), Bertrand and Mullainathan (2003), Gormley and Matsa (2016)).

trade-off between the price (i.e., premium) and the probability of a takeover has become common wisdom and accepted as fact.

The goal of this paper is to provide causal estimates that allow us to assess the extent of the trade-off between price and probability, and to identify the types of mergers that ATPs deter versus allow to happen. We also aim to identify the channels through which firm-level ATPs create or destroy value for firms and for the economy as a whole. Establishing causal effects is important given the evidence of endogeneity of governance structures (Schoar and Washington (2011), Karpoff, Schonlau, and Wehrly (2017)) and the potential of mergers and acquisitions to create or destroy value (Morck, Schleifer, and Vishny (1990), Maksimovic and Phillips (2001), Burkart, Gromb, and Panunzi (1998), Schoar (2002)).

We start by showing that the expected gains from adopting ATPs can accrue to shareholders in three ways. The first is the effect of such provisions on the *probability* of being acquired (i.e., the deterrent effect). The second is the effect on the *premium* paid conditional on a successful acquisition. This is based on the *price* paid for the target at auction, which is determined in turn by the relative bargaining power of the parties, the degree of competition (i.e., number of bidders), and the potential for synergies. Note that the effect is ex ante ambiguous: for example, if ATPs give managers more bargaining power, removing them should result in a negative effect on the premium while if they attract less competition, removing them implies a positive effect (Bulow and Klemperer (1996)). The third is the *selection* effect, whereby an ATP changes the population of firms that become targets. For example, the additional firms that end up being taken over because they dropped an ATP may be those with the lowest (highest) potential for value creation, implying negative (positive) selection. The first two effects have been the focus of existing research, as they are important determinants of shareholder value. The third, albeit seldom discussed, is also important, as one cannot infer the takeover premium by comparing firms that are taken over with and without ATPs because the population of target firms changes when such provisions are in place.[2]

To provide causal estimates of the components of the expected premium in this setting requires some form of random assignment in the adoption of ATPs. We employ two different specifications. The first is a regression discontinuity design (RDD) for the takeover probability and expected premium. We use data on all shareholder-sponsored proposals (2,882 proposals in 927 different firms) to remove an ATP voted on at annual meetings of S&P 1500 firms between 1994 and 2013. We rely on vote outcomes being random in a narrow interval around the majority threshold, leading to a discrete change in the probability of dropping a provision (see Cuñat, Gine, and Guadalupe (2012, 2013)). The second specification is a matching estimator that is validated using the

---

[2] Note that most existing studies focus on effective (conditional) takeover premiums that are conditional on a takeover offer being made. Since premiums do not exist in the absence of a takeover bid, changes in these "conditional" premiums are subject to selection bias, as we discuss below.

identification strategy proposed by Angrist and Rokkanen (2015).[3] Since the two estimation strategies rely on a different set of assumptions, our study contributes to our understanding by providing consistent results across different techniques that evaluate the effects at different points of the sample.

Even when armed with a source of exogenous variation, we still need to correct for selection problems inherent in the estimation of the conditional premium given the co-determination of premiums with the population of firms taken over. We use the bounding estimation strategy proposed by Lee (2009) to estimate upper and lower bounds for the effect of ATPs on the takeover premium. Because we need a distribution of premiums to apply the bounding technique, which cannot be done at the exact discontinuity, we provide two sets of results for the conditional premium using Lee bounds. The first is on the full sample using validated matching (Angrist and Rokkanen (2015)) and the second is in an interval around the discontinuity.

Across specifications and samples, we find that voting to remove an ATP has a significant positive effect on both the probability of a firm being taken over and the future takeover premium. At the majority threshold (classic RDD), passing a proposal to drop an ATP increases the likelihood of a takeover within five years by 9.2% (1.8% per year) and increases the expected value of future takeover premiums by 4.1%. For firms away from the discontinuity, the effects are smaller but also positive and significant: voting to remove an ATP increases the probability of a takeover within five years by 4.1% (0.8% per year) and increases the expected value of future takeover premiums by 2.6%. These results are intent to treat (ITT) effects that measure the effect of passing a proposal. In Section V, we discuss possible calculations of the effect of the provision itself (treatment on the treated). We also show that the results are similar across different definitions of the share of votes passed (in particular, the different treatment of abstentions across votes).

Total shareholder gains can be expressed as an unconditional premium that includes both firms that experience a takeover (and realize a takeover premium) and those that do not (with a takeover premium of zero). The effect on the expected unconditional premium is not subject to the inherent selection problem of the conditional (i.e., realized) takeover premium because the populations of the treatment and control groups are comparable. However, we would also like to determine whether a given firm is able to obtain a higher or lower premium if it drops the ATP and a merger does happen. To do so, we cannot simply compare the premiums of firms taken over with or without ATPs, as we can only observe takeover premiums for the firms that are taken over and we need to account for different selection patterns in the two groups. We therefore use the bounds methodology developed by Lee (2009) to provide estimates that

---

[3] Angrist and Rokkanen (2015) build on the fact that in the regression discontinuity design, we observe the assignment variable (the vote in our case), which is the only source of heterogeneity. They propose a matching estimator and use the regression discontinuity approach as a tool for validating the conditional independence assumption of the model. We explain the method and intuition further in Section III.

account for selection. Across different specifications and samples (for different intervals around the discontinuity as well as for the full sample), we find that the effect of voting to remove a provision on the conditional premium is never negative and can be as high as a 15% four-week premium, suggesting that more shareholder value is created in less protected firms. This effect is relatively large compared to other effects that the literature finds to be relevant. For example, Eckbo (2009) shows that, in the cross-section, the average difference in premiums between a hostile takeover and a friendly takeover is 5.8%, the average difference in premiums between a public and a private acquirer is 4.9%, and the average difference in premiums between a multiple bidder and a single bidder contest is 7.8%.

The takeover probability effect is in line with the findings of Karpoff, Schonlau, and Wehrly (2017), who, using a different sample and a different local instrument, also find a deterrent effect of ATPs. This consistency between two different causal approaches contrasts with the largely inconclusive prior literature that does not address the endogeneity of adoption.[4]

The positive premium effect is in contrast to the accepted wisdom that ATPs allow managers to obtain higher premiums; we find the opposite to be true in our sample, where there appears to be no trade-off between price and probability. This is all the more important given that studies of the correlation between ATPs and takeover premiums often find that adopting an ATP has a negligible or positive effect on the premium (Comment and Schwert (1995), Bange and Mazzeo (2004), Bebchuk, Coates, and Subramanian (2002), Bates, Becher, and Lemmon (2008), Cotter, Shivdasani, and Zenner (1997)). However, no prior analysis has addressed both the endogeneity of adoption and the selection of targets as we do here.

We next investigate the determinants of the positive premium result (of dropping a provision), which challenges the argument that ATPs give managers bargaining power to extract a higher premium. We find that across specifications, total shareholder value creation (adding up the dollar value of the acquirer and target premiums) is positive. This net value creation seems to come in part from more acquisitions in related industries (with higher potential for synergies) and in part from targets being matched to more valuable acquirers (positive selection).

Across the full sample of firms, the higher premium is also linked to more competition for less protected firms: they have more bidders, more unsolicited bids, more challenged deals, and more deals paid in cash. The competition among bidders increases the overall bargaining power of the target and seems to trump any loss in bilateral bargaining power with each individual bidder. This is consistent with the auction literature that suggests that the surplus of

---

[4] For example, Pound (1987) documents that ATPs reduce the probability of a takeover bid, and Ryngaert (1988) finds that firms with a poison pill are more likely to reject a hostile takeover bid. In contrast, Comment and Schwert (1995) find that poison pills have no effect on takeovers, and Bates, Becher, and Lemmon (2008) find that having a staggered board does not preclude the completion of a takeover once a firm has already received a bid, though it may reduce the likelihood of receiving a bid in the first place.

the seller is determined largely by the number of bidders rather than the individual bargaining power/negotiating skills of each of them (Bulow and Klemperer (1996)). However, for the more contested votes (around the discontinuity) we find a decrease in competition in terms of number of acquirers and number of contested deals and instead find some evidence of a higher acquirer premium, suggesting that for close-call deals the absence of competition allows acquirers to capture some of the surplus generated.

In our last set of tests, we use our empirical framework to determine the fraction of the total increase in value from removing ATPs that comes from its different components. We find that the increase in value operates largely via quantities: over 50% comes from the increased probability of mergers. The premium effect is positive and potentially large. The selection effect is positive and between a quarter and a half of overall value created when we focus on the whole vote support. We cannot assess the direction of selection when we focus on close-call votes, but the bounds suggest that it is potentially large. This result confirms that accounting for selection is important to understand how takeovers create value in the market.

In terms of the generalizability and external validity of our results, our data include over one-third of firms in the S&P 1500 between 1994 and 2013. However, while our results apply to a significant share of firms, we cannot extrapolate the results—without further assumptions—to firms that are never subject to such shareholder votes. In particular, we cannot rule out the possibility that firms in which the premium effect is negative never hold a vote on ATPs. But even in such a case we note that despite the proposals in our sample creating value for all firms on average, some failed to garner strong shareholder support, and the vast majority were opposed by management. This raises questions about governance that are beyond the scope of this paper.

In Section I, we provide a framework for decomposing the unconditional premium. In Section II, we discuss the identification strategies underlying the two specifications that we use in the analyses. In Section III, we describe the data, and in Section IV, we present the results on unconditional premiums and takeover probabilities. Section V reports the results on the conditional premium. In Section VI, we present our decomposition. In Section VII, we offer potential explanations for the positive premiums. Section VIII concludes.

I. Framework: Decomposing the Unconditional Premium

A. Addressing Endogeneity and Selection

We start by developing an analytical framework for examining the effect of ATPs on expected shareholder gains via takeover probabilities and premiums. This framework allows us to establish the elements required for the decomposition of the unconditional premium in Section I.B and to assess all possible sources of bias that we need to address empirically.

We define the treatment dummy variable D, which takes the value of 1 if shareholders vote to drop an ATP and 0 if they vote to keep it. Empirically, we observe the realized premium variable Y, which equals the premium paid

if a takeover takes place and 0 otherwise. The realized premium captures the shareholder gains from the entire population of firms at risk of a takeover. To understand selection issues, we define two latent variables. The first, Y^*, is the potential premium offered for a firm, which is observed only if a takeover takes place. The second, Z^*, is a measure of the latent merger propensity of a firm. A merger happens whenever $Z^* > 0$. We can therefore write the unconditional premium (i.e., not conditional on whether the merger occurred) as $Y = 1[Z^*>0] \cdot Y^*$, where $1[.]$ is the indicator function.

The structure above gives rise to the classic selection model, which using standard notation and assuming a linear structure can be written as (Heckman (1979), Lee (2009)):[5]

$$Y* = D\beta + X\mu_1 + U \quad \text{(underlying premium)}.$$

$$Z* = D\gamma + X\mu_2 + V \quad \text{(latent merger propensity)}.$$

$$Y = 1[Z* > 0] \cdot Y*. \quad \text{(unconditional premium)}.$$

The first challenge is to find a way to randomly assign the treatment dummy D. If D is randomly assigned, then we can recover the effect of an ATP on the unconditional premium, ΔY, and on the takeover probability, ΔP:

$$\Delta P = \Pr[Z* > 0|D = 1] - \Pr[Z* > 0|D = 0], \tag{1a}$$

$$\Delta Y = E[Y|D = 1] - E[Y|D = 0]. \tag{1b}$$

However, even with a randomly assigned D, one cannot recover β. Nevertheless, β is the parameter of interest to assess the effect of ATPs conditional on a merger taking place. This parameter represents the difference in the price paid for a specific target with or without an ATP in place.

The reason we cannot recover this causal parameter even when we have an instrument for D is the selection of targets: the observed Y is conditional on a merger occurring ($Z^* > 0$), which itself is affected by treatment $E[Y \mid D, X, Z^* > 0] = D\beta + X\mu_1 + E[U \mid D, X, V > -D\gamma - X\mu_2]$.

Typically, existing premium studies compare premiums conditional on a merger happening for firms with and without ATPs, which we can write as

$$\begin{aligned} E[Y|D = 1, X, Z* > 0] &- E[Y|D = 0, X, Z* > 0] \\ &= \beta + E[U|D = 1, X, V > -\gamma - X\mu_2] - E[U|D = 0, X, V > -X\mu_2]. \end{aligned} \tag{2}$$

Equation (2) shows that even with a randomly assigned D (and if U and V are not independent), one cannot recover the causal effect on Y^* because of

---

[5] This model can be generalized to a nonlinear structure.

the sample selection term $E[U \mid D = 1, X, V > -\gamma - X\mu_2] - E[U \mid D = 0, X, V > -X\mu_2]$. We therefore need an identification strategy that not only provides exogenous assignment to treatment but also corrects for selection. Section II describes how we address both of these requirements.

B. Decomposing the Unconditional Premium: Probability, Price, and Selection Effects

Recall from equation (1b) that

$$\Delta Y = E[Y|D = 1] - E[Y|D = 0]$$
$$= \Pr[Z* > 0|D = 1] * E[Y|D = 1, Z* > 0] - \Pr[Z* > 0|D = 0] * E[Y|D = 0, Z* > 0].$$

One can rewrite this equation, after some manipulation, as

$$\Delta Y = \Pr[Z* > 0|D = 1] * \beta \quad \text{(premium)}$$
$$+ E[Y|D = 0, Z* > 0] * \Delta P \quad \text{(probability)}$$
$$+ \Pr[Z* > 0|D = 1] * \{E[Y|D = 1, Z* > 0] - E[Y|D = 1, V > -X\mu_2]\}. \quad \text{(selection)}$$

Each of the terms in the expression represents a different effect of a provision on shareholder value. The first term measures the direct impact on takeover premiums β (times the baseline probability of a merger for the treated group). This is the effect on the premium *conditional* on a takeover taking place, or in other words, how much more/less an acquirer would pay for a given firm. Note that if ATPs give managers more bargaining power, then removing them should lead to negative premiums. However, if removing them attracts more competition for the target or induces better matching between bidder and target, then it may lead to positive premiums. The second term captures the change in merger probabilities (times the premium for the untreated group). This reflects the change in the merger probability from the presence of an ATP and the strength of the provision as an anti-takeover device. The third term is a selection term that captures the change in the population of firms subject to a takeover offer (reflecting the fact that ATPs change the population of firms that become targets). For example, firms that end up being taken over may be the weakest (strongest) firms in the economy as determined by potential value creation, leading to negative (positive) selection.

In the remainder of the paper, we explain how we obtain each of the terms, and we estimate the contribution of each to the overall unconditional premium, as reported in Section V.

II. Identification Strategies

To identify the impact of an additional anti-takeover measure on the two outcomes of interest, we can directly estimate the takeover probability ΔP and the unconditional takeover premium ΔY (as defined in equations (1a) and (1b)).

We can then estimate conditional premiums and their determinants, which requires that we examine how the firms that become a target are selected.

We define y_{ft} as the outcome of interest for firm f at time t, v_{ft} as the votes in favor of a shareholder-sponsored ATP, v_f^* as the majority threshold for a proposal to pass in firm f, and $D_{ft} = 1(v_{ft} \geq v_f^*)$ as an indicator that takes a value of 1 when a proposal passes. The term K is a constant. We can then express the relationship of interest as

$$y_{ft} = K + D_{ft}\vartheta + u_{ft}. \tag{3}$$

The effect of interest is captured by the coefficient θ, while the error term u_{ft} represents all other determinants of the outcome. However, using this expression in a regression is unlikely to give a consistent estimate $\hat{\theta}$ because passing a proposal that induces dropping an ATP is correlated with omitted variables that are themselves correlated with the probability and characteristics of a takeover, so that $E(D_{ft}, u_{ft}) \neq 0$. In the next section, we cover two possible approaches to address this problem.

A. Effect of ATPs on Unconditional Premiums and Merger Probabilities

In our setting, we use two different complementary approaches to calculate the effect of an additional anti-takeover measure on the total unconditional premium and on the probability of a merger. The first is a classic RDD. The second is a matching model that uses the RDD setting to validate the conditional independence assumption (CIA).

A.1. Classic RDD

Identification in the classic RDD setting exploits the facts that assignment into treatment is governed by the running variable (votes) and treatment probability changes discretely around the majority threshold.[6] However, the distribution of other observable and unobservable firm characteristics is continuous around the threshold. In an arbitrarily small interval around the majority threshold, assignment to either side can be considered random. Therefore, a discontinuous increase in the outcome variable around the passing threshold can be interpreted as caused by the treatment.

To estimate this discrete change in the outcome variable, one can use the entire data set, fitting flexible functional forms for the relationship between the vote and the dependent variable in different ways. Lee and Lemieux (2010) suggest the use of different polynomials for observations on either side of the

---

[6] Evidence for the fact that implementation probabilities change discretely at the discontinuity can be found in Cuñat, Gine, and Guadalupe (2012, 2016), Popadak (2014), and Bach and Metzger (2015).

threshold.[7] Alternatively, one can run a local regression on an optimally calculated interval around the discontinuity, as proposed by Imbens and Kalyanaraman (2012, IK) for a local linear regression approach, and extended by Calonico, Cattaneo, and Titiunik (2014, CCT) to a second-order weighted-polynomial regression over an optimal bandwidth that balances efficiency and bias.[8]

A.2. A Matching Model with Validated CIA in an RDD Setting

The downside of the classic RDD is that identification is local and comes from firms with vote outcomes around the discontinuity. To obtain arguably causal estimates for firms with vote outcomes away from the majority threshold, we use a matching estimator. Following the identification strategy in Angrist and Rokkanen (2015), we estimate a matching model in which we validate the CIA using a test that relies only on the standard underlying assumption of RDDs, namely, that all of the heterogeneity in the treatment comes from the running variable.

Angrist and Rokkanen (2015) note that a unique feature of a regression discontinuity setting is that one observes the running variable (v_{ft} in our case), which is the only factor that determines "treatment" ($D = 1$), or as they put it, "RD design takes the mystery out of treatment assignment." In classic matching model applications (without random variation in D), researchers match treated and control firms and assume that the set of controls that they match on is sufficiently rich that any difference between outcomes is driven only by treatment (D). However, unless treatment is randomly assigned, there is no way to know whether one has the right set of matching criteria or whether omitted variables are correlated with assignment and the outcomes.

In regression discontinuity, the running variable is the (only) assignment variable. For example, in our setting we know that $D_{ft} = 1[\, v_{ft} > v_{ft}^*]$. Thus, we know the nature of the possible omitted variables bias anything that is correlated with v_{ft} that also determines outcomes. One can therefore identify the coefficients of interest under a CIA:

$$\mathrm{E}[y_{ft}|v_{ft}, x_{ft}] = \mathrm{E}[y_{ft}|x_{ft}]; D = 0, 1.$$

The CIA states that potential outcomes are mean-independent of the running variable, conditional on a set of controls x_{ft}. If the CIA holds, then the model is identified.

In standard matching models, the CIA is implicit but never tested because the assignment process is not observed. But the RDD setting gives us the

---

[7] If votes are stochastic, the estimator can be interpreted as a weighted average treatment effect that uses all observations, with weights directly proportional to the probability of each firm having a realized vote near the discontinuity (Lee and Lemieux (2010)).

[8] The weights are computed by applying a kernel function on the distance of each observation's score to the cutoff. We then estimate θ as the difference between these nonparametric regression functions on either side of the majority threshold.

running variable, which allows us to test the CIA outside the discontinuity.[9] In practice, testing the CIA amounts to making sure that while there may be a significant relationship between y_{ft} and v_{ft}, the two are mean-independent conditional on x_{ft}. This can be tested by showing that they are mean-independent outside the discontinuity threshold.

In other words, the RDD provides a diagnostic tool to test the validity of the model that is used in a matching estimator. In this sense, the proposed estimator is formally very different from an RDD, but uses the RDD setting to validate the set of variables that participate in the model. This approach stems from the assumption that the only source of heterogeneity in assignment is the vote and relies on an auxiliary regression to test it.

Ideally one wants the CIA to hold in the full support of the running variable.[10] However, this may not be feasible in some cases. In such circumstances, some researchers have proposed limiting the sample to some interval from the threshold over which CIA holds (this is referred to as Bounded CIA in the NBER working paper version of Angrist Rokkanen (2015) and is used, for example, in Hainmueller, Hall, and Snyder. (2015)).[11]

Imposing this additional structure has several advantages. First, this strategy allows us to estimate a matching model using the full sample and test the CIA that underlies identification. This means we can provide estimates for firms with vote outcomes away from the discontinuity while retaining a causal interpretation.[12] Moreover, using our estimates we can build counterfactuals at each vote level that predict what would have happened had that firm voted differently. This implies that we can assess whether there are heterogeneous

---

[9] A further condition required in this identification strategy is the existence of common support, so that the treatment status (removing an ATP) retains meaningful variation after we condition on *X*.

[10] In Angrist Rokkanen (2015), the full-sample CIA is estimated in the ±20 interval for each school because this is the interval over which the samples are "clean" in the sense that the counterfactuals are clear—it avoids a situation whereby a student in the "accepted" sample of the less selective school (O'Bryant) would also be accepted into the more selective school (BLS) where someone not accepted is to BLS would have been below the O'Bryant threshold for acceptance.

[11] The tension between these two strategies is that, on the one hand, it is preferable not to select the interval based on postestimation results (this favors testing for the CIA on the full support only), while on the other hand, testing the CIA on smaller subsamples (because it is less likely to be satisfied as we move away from the discontinuity) provides a gradual sense of the extent to which the CIA is satisfied (this is the Bounded CIA strategy).

[12] Angrist and Rokkanen (2015) also show how to extend this to the fuzzy RDD. Note that throughout the paper, since we do not have information on implementation, we present reduced-form estimates of the intent to treat. One can estimate the impact of implementation by rescaling the intent to treat estimator. The rescaling factor is one over the change in the probability of implementation when a proposal passes. This conversion factor when estimated over the whole vote support ranges between 1.2 (using estimates from Popadak (2014), who finds that the difference in implementation between pass and fail is 84%) and 1.7 (using estimates from Bach and Metzger (2015), who find that average implementation conditional on passing is 59%). Note also that one does not need CIA on implementation to hold to interpret our estimates as causal ITT estimates—the CIA test already takes into account any heterogeneity in implementation on the outcome variables.

effects of ATPs for different levels of vote support. Overall, this approach allows us to test the validity of the matching in a theoretically sound way, subject only to the limitations of applying an asymptotic result to a finite sample. The limit to interpreting results using this identification strategy as causal is the extent to which the CIA is satisfied. We discuss this further below. See section 3 in Angrist and Rokkanen (2015) for a related discussion.

B. *Estimating the Conditional Premium using Bounds*

The existing literature focuses on the effect of ATPs on the takeover premium conditional on a merger happening. However, as noted in Section I, a remaining challenge is to disentangle which part of this effect is a causal effect, fixing the characteristics of the target firm (e.g., effects that arise from changes in bargaining power, matching with different bidders, changes in competition for target firms, etc.), and which part of the effect is due to selection (i.e., when ATPs are dropped, a different population of firms experience takeovers).

This is a form of selection that is inherent to the problem studied rather than a sampling issue. To correct it, we could have an excluded variable in a Heckman selection model, but these are virtually impossible to find in this setting since any variable that predicts takeovers will also determine the premium. The alternative is to provide bounds for the parameters of interest.

Lee (2009) shows how to use the structure of the underlying model to recover upper and lower bounds for β. If one observes $E[Y \mid D = 1, X, V > -X\mu_2]$, which is the premium from the sample of firms that would have merged even without the ATP but that actually removed it, then one could estimate β from $E[Y \mid D = 1, X, V > -X\mu_2] - E[Y \mid D = 0, X, V > -X\mu_2]$. However, this is never observed. But notice that the sample for which $V > -X\mu_2$ is included in $V > -\gamma -X\mu_2$. This gives us a strategy to provide an upper (lower) bound for β under a monotonicity assumption: if one considers that all counterfactual observations for which we do not see Y are drawn from the lower (upper) end of the Y distribution, we can obtain a lower (upper) bound for β by trimming a proportion p $(1 - p)$ from the observations for Y, where $p = \Pr(-\gamma -X\mu_2 < V < -X\mu_2)/\Pr(-\gamma -X\mu_2 < V)$. In what follows, we call these "sharp Lee bounds" (Lee (2009)).

Note that to apply Lee (2009) bounds we need an empirical distribution of the conditional premium, so it is not possible to estimate bounds at the discontinuity without making additional assumptions. One possibility is to assume that the distribution of premiums on an interval around the discontinuity is a good approximation for the distribution at the discontinuity. One needs to achieve a compromise between a narrow interval (e.g., 10 percentage points around the discontinuity) that brings the results closer to a causal interpretation or a broader interval that produces a more meaningful distribution and increases the power of the test. We use this approach when we decompose the unconditional premium into its components in an interval around the discontinuity. Alternatively, we can extend the results beyond the discontinuity using the CIA-validated matching so we can implement the bounding strategy on the population distribution of premiums using all available observations

adequately weighted. Under the CIA, this approach produces results that can be interpreted as causal and that contain the whole distribution of premiums. Our study thus adds value by combining these various techniques for different subsamples and each with its own maintained assumptions to show similar results.

III. Data Description and Sample Characteristics

We construct a data set that spans 20 years of voting data from ISS-Riskmetrics (ISS-Shareholder Proposals database).[13] These data contain information on all of the proposals voted on in the S&P 1500 universe and an additional 500 widely held firms. We restrict the analysis to the set of ATPs that make up the G-index as defined by Gompers, Ishii, and Metrick (2003). Our main sample consists of 2,881 shareholder-sponsored proposals voted on at annual meetings to change the anti-takeover structure of the firm.

Based on this sample, we construct two different vote metrics. The first is *Simple Vote Rule*, which is computed as Votes For/(Votes For + Votes Against) whenever the pass rule is defined over votes cast. The second metric, *Vote Adjusted for Abstentions*, starts from the simple majority vote rule and computes Votes For/(Votes For + Votes Against + Abstentions) in those cases in which the firm or state rules determine that the cast votes include abstentions. In both cases we use Votes For/(Shares Outstanding) whenever the pass rule is defined over votes outstanding. To gather information about how votes are computed, we fully merge the ISS-Shareholder Proposals database with the ISS-Voting Results database for the period 1997 to 2006. We also merge it with Voting Analytics for the 2007 to 2013 period; we are able to match around half of our observations in that period. Unfortunately, within our universe there is no reliable information about the treatment of abstentions for observations before 1997 and for the unmatched observations after 2006.[14] The advantage of *Simple Vote Rule* is that it is consistently defined across all observations and is comparable with previous studies that use the ISS-Shareholder Proposals data set. The simple majority rule is also very focal and used by ISS, investors, and the SEC to justify their rules and voting recommendations. The advantage of *Vote Adjusted for Abstentions* is that it is closer to what managers publicly report as the vote outcome and hence is more directly determinant for implementation. The disadvantage is that the inclusion of abstentions and broker nonvotes may make this measure easier to manipulate. Bach and Metzger

[13] For the period 1997 to 2013, we use the ISS-Shareholder Proposals data set, formerly known as Riskmetrics, and now part of ISS. For the period 1994 to 1996, we use data from ISS tapes. We would like to thank Ernst Maug and Kristian Rydqvist for providing us with these data (Maug and Rydqvist (2009)).

[14] This translates into not knowing the exact treatment of abstentions for one-third of the sample. On average, two-thirds of the firms use the simple majority rule, so we expect that for one-ninth of the *Adjusted Vote* observations we use the simple rule even though we should have accounted for abstentions. The difference between the two vote criterions is small: on average abstentions represent 1.3% of the votes, and out of 1,851 observations for which we know the exact voting rule using the alternative data set, only 30 observations change from pass to no pass.

([2019](#)) argue that there is manipulation using the *Adjusted Vote* metric in the Voting Analytics data set. As we show below, we find no evidence of manipulation in our sample with either metric. Throughout the paper, we use *Simple Vote Rule* as our main specification, but we report our main results for both rules and show they are very similar across measures.[15]

To obtain our treatment indicator (D), we use information on vote outcomes adjusted by majority rules (simple majority – supermajority) and votes base (votes cast or outstanding). If this information is not available, we use a simple majority rule of 50% of votes cast. We define the distance to the vote as the difference between the vote outcome and the majority threshold ($v_{ft} - v_f^*$).

We match this sample of firms to the SDC platinum database to identify which firms were taken over following a vote. We consider whether a firm is taken over within five years of the vote if at least 50% of its ownership is acquired by a bidder. For firms with multiple votes we treat these as separate events, but we cluster standard errors by firm in our estimates.[16] In most of our analysis, we define the merger premium for firms that are taken over as the cumulative return from four weeks prior to the takeover announcement up to the completion date (as reported by SDC), which, as we show below, gives the most conservative estimates of our effect when in robustness tests we compare it to a range of alternative measures (see Table V). We also obtain information from SDC on the acquirer's premium (available only for listed acquirers), the number of bidders, the number of unsolicited bids, whether the deal was challenged, the percentage that was paid in stock, and whether both firms belong to the same two-digit SIC industry. Financial information comes from Compustat and ownership information from Thomson 13F.

Table IA.I in the Internet Appendix presents information on the evolution of the votes to remove an ATP used in the paper, as well as the takeover probabilities and premium over time. The average probability of a firm experiencing a takeover over the five years following a shareholder vote is 14%. We have a total of 138 (81) targets within 10 (5) percentage points of the majority threshold. The mean conditional premium (the premium paid conditional on a successful merger) is 32.7% and the mean unconditional premium (which assigns zero premium to the unsuccessful mergers) is 4.83%.[17]

Table I presents basic descriptive statistics for the firms in our sample. To assess how firms subject to a shareholder proposal differ from their sampling

---

[15] We also report all of the specification tests and postestimation calculations for *Vote Adjusted for Abstentions* in the Internet Appendix, which is available in the online version of the article on *The Journal of Finance* website.

[16] In an earlier version, we analyzed whether there were differential effects for firms that voted on an issue the first time, second time, third time, etc. We found no significant differences, so we decided to pool the effects.

[17] Note that throughout the paper, we treat all ATPs as if they were identical, although in reality they may not be. Table IA.IV in the Internet Appendix shows that the most frequent provisions, which play a major role in identifying the effect, are: repeal classified board (35%), adopt cumulative voting (16%), and repeal poison pill (14%). Given that there is not enough power at the provision-type level to identify the separate effects of each proposal type within our methodology, we run our analysis for all G-index proposals pooled—a common measure of ATPs.

Table I

Descriptive Statistics

This table describes the sample of 2,882 voted G-index proposals one period before the vote. All accounting variables are obtained from Compustat: Market Value (mkvalt_f), Tobin's Q, defined as the market value of assets (AT+mkvalt_f-CEQ) divided by the book value of assets (AT), balance sheet Deferred Taxes and Investment Tax Credit (TXDITC), Return on Equity (NI/(CEQ+TXDITC)), Return on Assets (NI/AT), Profit Margin (EBITDA/Sale), Liquidity (CHE/Sales), Leverage ((DLTT+DLC)/AT), Capital Expenditures (Capx/AT), and Overhead (XSGA/XOPR). Ownership variables are generated from Thomson 13F database. All monetary values are in 2012 US$. Note that the number of observations may change due to missing values for some of the variables.

	N	Mean	Median	SD	10th Percentile	90th Percentile	Mean SP1500	t-Test
Market Value ($million)	2860	28,928	9,476	58,068	595	73,761	9,561	15.4
Tobin's Q	2755	1.58	1.25	0.98	0.95	2.59	1.96	−14.8
Return on Equity	2863	0.153	0.107	0.220	−0.070	0.292	−0.04	1.25
Return on Assets	2861	0.031	0.031	0.087	−0.023	0.109	0.12	−7.4
Profit Margin (EBITDA/Sales)	2805	0.159	0.169	0.167	0.055	0.385	0.13	0.3
Cash Liquidity (CHE/Sales)	2861	0.091	0.052	0.109	0.006	0.221	0.13	−17.8
Leverage (DLTT+DLC)/AT	2859	0.288	0.277	0.165	0.078	0.505	0.223	18.5
Capital Expenditures (Capx/AT)	2744	0.053	0.043	0.048	0.003	0.109	0.052	0.12
Overheads (SGA/Op.Exp.)	2175	0.282	0.251	0.184	0.077	0.521	0.314	−5.79
Ownership Inst. Shareholders	2696	0.638	0.655	0.193	0.378	0.864	0.680	−4.85
Ownership Herfindahl	2698	0.054	0.041	0.054	0.022	0.089	0.063	−7.61
Vote Simple rule	2882	48.14	47.03	21.83	20	78.1	n.a.	n.a.
Vote Adjusted for Abstentions	2882	47.80	46.7	21.56	20	77.8	n.a.	n.a.

population, we present the characteristics of the average S&P 1500 firms and compare them to firms in our sample. One of the most noticeable differences is that firms in our sample are three times larger than the average S&P 1500 firm. In addition, firms in the voting sample have lower Tobin's *Q*, a slightly higher leverage ratio, and relatively less cash liquidity. However, they are not that different in terms of profitability, return on equity, cash flows, capital expenditures, and overhead. This suggest that while we obtain results for all firms subject to an anti-takeover removal proposal (roughly one-third of the population of S&P 1500 firms), one should exercise caution in extrapolating the results to firms that have never had such a shareholder proposal. In other words, as with any identification strategy, we cannot extrapolate the results outside the sample without making further assumptions.

IV. The Effect of ATPs on Takeover Probability and Unconditional Premiums

A. RDD Estimates of Unconditional Premiums and Takeover Probabilities

A.1. Preliminary Tests to Validate the RDD Identification Strategy

Before presenting results using the classic RDD and the validated matching model, we need to run a series of tests to confirm that the setting of interest is appropriate for using these methods. First, we show there are no pre-existing differences in firm characteristics (or trends in firm characteristics) around the majority threshold, which is an assumption of the RDD (Table IA.II in the Internet Appendix).

Second, we show that the distribution of the frequency of votes is continuous around the discontinuity. A discrete and significant jump in density to either side of the discontinuity would be indicative of strategic behavior around the majority threshold, such that the continuity assumption would be violated. This does not appear to be true in our data, and we believe the main differences with respect to Bach and Metzger (2019) arise from the limited overlap between the two samples—only 16% of observations are common, due to different proposal coverage, years, and sampling (See the Internet Appendix for more details on differences between the samples).[18] In fact, Figure IA.1, Panel A, in the Internet Appendix shows a smooth overall distribution of votes for the simple vote. Figure IA.1, Panel B, in the Internet Appendix shows that the formal continuity test proposed by McCrary (2008) rejects the discontinuity

---

[18] There are important sample differences with respect to Bach and Metzger (2019): we focus on G-index proposals for S&P 1500 firms, their main sample focuses on a broader set of proposals of the Russell 3000 index that only includes 16% of takeover-related proposals; they also focus on a different time period. When Bach and Metzger (2019) use ISS data for 2003 to 2011, they focus on the 10 proposals with most favorable votes. Only six of them belong to the G-index. They also have incomplete sampling due to missing information. The fraction of observations in our sample present in their sample ranges from 9% to 25% depending on subsamples. Section III in the Internet Appendix describes the main differences between these two databases for our time period.

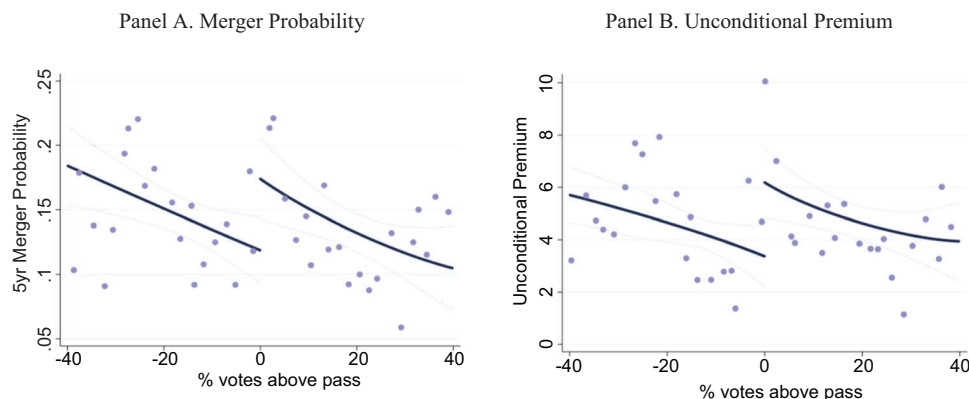Panel A. Merger Probability   Panel B. Unconditional Premium

Figure 1. Unconditional merger probability and premium. Linear regression using the Imbens and Kalyanaraman (2012) approach to select the bandwidth. Dots represent the simple means by bins of 2% vote intervals. (Color figure can be viewed at wileyonlinelibrary.com)

of the density function at the majority threshold. Figure IA.1, Panel C, in the Internet Appendix plots results of the discontinuity test for the adjusted vote and shows no statistical difference at the threshold (note that, visually, there is a small discrete change, but this does not resemble in any way the large discontinuity documented by Bach and Metzger (2019; Figure 1). We also test for discontinuity in the votes for subperiods and by proposal using both vote definitions, and find no evidence of manipulation in any subsample (see Table IA.V, Panel A and B, in the Internet Appendix). These tests confirm that our data set represents a good setting in which to apply the classic RDD using both vote definitions.

A.2. Results using the RDD

We now present estimates of the effect of passing a proposal to remove an ATP on the takeover probability and the expected premium using the RDD.

We begin by presenting graphical evidence using all of our data. Figure 1, Panel A, shows the relationship between the merger probability and the distance from the majority threshold (percentage of votes above pass in the horizontal axis) for *Simple Vote Rule*. The dots represent simple means in bins of 2% vote intervals, and the solid line is a running linear regression using the Imbens and Kalyanaraman (2012) approach to select the bandwidth. Overall, the downward-sloping line suggests that higher shareholder support for dropping ATPs is associated with a lower likelihood of a takeover. On the basis of this evidence alone we would wrongly conclude from the correlation that the more likely firms are to drop ATPs, the less likely they are to be taken over. However, this is driven by unobserved characteristics. In fact, at the majority threshold we see a discrete change upward in the function, suggesting a positive causal effect of voting to drop the provision on the takeover probability.

The size of this discrete change is the regression discontinuity estimate, that is, the local causal effect of the vote outcome.

Figure 1, Panel B, shows the same graph with the unconditional premium in the vertical axis. We again observe a negative overall relationship between the two variables but a clear positive change at the discontinuity, suggesting that voting to drop an ATP increases the unconditional premium that firms expect to receive.[19]

Table II presents regression estimates of the effect at the discontinuity seen in Figure 1, Panels A and B, using four different estimation methods. Columns (1) to (4) report results for the nonparametric test, which is a means test of the outcome variable, calculated on an increasingly narrow interval of votes around the majority threshold. Columns (5) and (6) report the regression discontinuity estimates using polynomial controls of order two and three (respectively) to each side of the discontinuity. Columns (7) and (8) report results of running local regressions on an optimal bandwidth around the discontinuity. In particular, column (7) reports the Imbens and Kalyanaraman (2012) local regression estimate, and column (8) reports the Calonico, Cattaneo, and Titiunik (2014) estimate. Columns (9) to (12) replicate the specifications of columns (5) to (8) using *Vote Adjusted for Abstentions*.

Panel A reports results for the probability of a takeover within five years of a shareholder vote. The results show no effect on average of passing a proposal when all observations are included (column (1)). The differential probability of experiencing a takeover within five years of the vote is 4.76% in the 10% interval and between 7.7% and 9.6% in the narrower intervals.[20] Using the specifications in columns (5) to (8), this effect ranges from 8.7% to 11.7%. The results when using *Vote Adjusted for Abstentions* in columns (9) to (12) are very similar with estimates ranging from 7% to 9.2%. These effects are sizeable when compared with the sample-wide average five-year probability of a takeover of 14%.

In Panel B of Table II, we explore the effects of ATPs on the unconditional expected premium received by shareholders in subsequent takeover transactions over five years. We focus on unconditional premiums (we assign zero premium to firms that do not undergo a merger within five years).

The results in columns (1) to (4) of Table II, Panel B, report results for the fully nonparametric means comparison approach. The effect of dropping an ATP is an increase in the expected premium of between 2.6% (in the 10% interval) and 4.3% (closer to the threshold). Columns (6) and (7), which use the flexible polynomial approach, show expected premiums of about 5%. The local regression approach produces slightly smaller estimates of 4.1% and 3.6% (IK and CCT). The results with *Vote Adjusted for Abstentions* in columns (9) to (12)

---

[19] The graphs are very similar when using *Vote Adjusted for Abstentions*. See Figure IA.3, Panels A and B, in the Internet Appendix.

[20] A possible explanation for the difference in the size of the effects is that the estimation of θ in a broad interval is biased due to the endogenous adoption of proposals. For example, if firms with a lower ex ante likelihood of receiving an offer are more likely to drop ATPs, a sample-wide estimate like that in column (1) would be biased downward.

Table II
Takeover Probability and Premiums around the Majority Threshold

This table presents the effect of passing an anti-takeover proposal on the probability of becoming a target and on premiums. Panel A displays the probability of becoming a target, which is estimated over the next five years after the vote using SDC data. Panel B displays the unconditional premium of becoming a potential target. Premiums are computed as the price offer to target four weeks prior to announcement until completion. Column (1) estimates are based on the full sample. Column (2) restricts the sample to observations with a vote share within 10 points of the threshold, column (3) to five points, and so forth. Column (5) and (6) introduce a polynomial in the vote share of order two and three (Lee and Lemieux (2010)), one on each side of the threshold, and use the full sample. Column (7) uses the local linear regression approach by Imbens and Kalyanaraman (2012) with a triangular kernel function. Column (8) uses the nonparametric approach proposed by Calonico, Cattaneo, and Titiunik (2014). Columns (9) to (12) are estimated using an alternative vote measure that adjusts for abstentions. All columns control for year fixed effects; standard errors are clustered by firm. The reported bandwidth is expressed in percentage vote. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

	Simple Vote Rule								Vote Adjusted for Abstentions			
					Polynomial	Polynomial			Polynomial	Polynomial		
	Full	±10	±5	±2.5	2	3	IK	CCT	2	3	IK	CCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Probability of Becoming a Takeover Target over the Next Five Years												
Yes simple	−1.24%	4.76%	7.71%**	9.60%*	8.80%**	11.7%**	9.18%***	8.69%**	9.23%**	9.01%*	8.08%**	7.03%*
	(2.11)	(3.26)	(3.81)	(5.34)	(4.25)	(5.11)	(3.11)	(4.25)	(4.34)	(5.18)	(3.39)	(4.72)
Bandwidth							27.7	13.2			23.3	13.5
Observations	2,881	883	457	249	2,881	2,881	2,882	2,882	2,881	2,881	2882	2882
R^2/Z	0.000	0.004	0.010	0.015	0.006	0.007	2.95	2.05	0.005	0.006	2.38	1.65
Panel B: Unconditional Premium, Four Weeks												
Yes simple	0.194	2.616**	3.012*	4.321	5.327***	5.243**	4.12***	3.59*	5.02***	3.48	3.27**	2.66
	(0.794)	(1.212)	(1.707)	(2.662)	(1.711)	(2.251)	(1.52)	(2.02)	(1.746)	(2.294)	(1.60)	(2.09)
Bandwidth							24.6	14.3			22.6	13.8
Observations	2,881	883	457	249	2,881	2,881	2,882	2,882	2,881	2,881	2,882	2,882
R^2/Z	0.000	0.007	0.008	0.013	0.035	0.036	2.69	1.76	0.035	0.036	2.04	1.27

are qualitatively and quantitatively similar, although in column (12) the result is not statistically distinguishable from zero. Again, these effects are substantial against an average unconditional premium of 4.8% in the sample as shown in Table IA.I, Panel B, in the Internet Appendix.

The estimates are by construction local and since they are quite large it is natural to wonder whether they extrapolate to the rest of the sample. It is possible, for instance, that the very large estimates apply only to firms with close-call votes. To address this question, in the next section, we employ a validated matching model as a complementary estimation approach.

B. Estimating Unconditional Premiums and Takeover Probabilities using a Validated Matching Model

B.1. Testing the CIA and Preliminary Results

As described in Section II, the RDD setting is an ideal setting to test the CIA that underlies any matching model. This is what we do in Table III.

More specifically, in Table III we test whether conditioning on an explicit model for the determinants of takeover allows us to eliminate the relationship between the running variable (the vote) and the outcome variables (takeover probability and unconditional premium) at each side of the discontinuity. To satisfy the CIA, in the remainder of the paper we use a model that includes as regressors natural variables capturing the takeover probability and premium. These are firm size and performance the year before the vote (in terms of sales, market value, profit margin, cash liquidity), firm governance the year before the vote (percentage of equity controlled by institutional owners and E-index), measures related to market performance the year before the vote (average Tobin's Q of the industry and average market value of the industry), and year dummies.

Columns (1) and (3) ((5) and (7)) of Table III show that the correlation between the vote and the takeover probability (unconditional premium) on either side of the threshold ($D = 0$ and $D = 1$) is negative and in most instances highly significant. The effect is also rather large in most cases. For example, the coefficient in column (1) implies that a 10% increase in the vote outcome reduces the takeover probability by 2.51 percentage points (from a mean takeover probability of 14%). This result reflects the fact that the vote outcome and our dependent variables are highly correlated. However, once we condition on our model (even-numbered columns of Table III), the point estimates drop by a full order of magnitude, getting closer to zero, and the correlation becomes statistically insignificant. For example, the 2.51 percentage point effect in column (1) drops to 0.12 percentage points in column (2) (and is highly insignificant, with standard errors of a similar magnitude as in column (1)). This result shows that the outcomes and the vote are mean-independent conditional on a number of variables, which supports the assumption that vote and takeover probability are

Table III
Conditional Independence Tests

This table reports results on the conditional independence assumption for our two outcome variables: Takeover Probability and Unconditional Premium. Columns (1), (3), (5), and 7 present evidence on the initial relationship between the running variable (i.e., the vote) and the two outcome variables for observations to the left or right of the cutoff. Columns (2), (4), (6), and (8) present evidence for the model that controls for firm characteristics one year prior to the vote, where controls include ln Sales, Profit Margin, ln Market Value, Cash Liquidity, Percentage of Institutional Ownership, Average Industry Tobin's Q, Average Industry Market Value, and the Entrenchment Index. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

	Simple Vote Rule							
	Takeover Probability				Unconditional Premium			
	$D=0$		$D=1$		$D=0$		$D=1$	
	[−50,0)		[0,50]		[−50,0)		[0,50]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote	−0.00251***	−0.000127	−0.00160*	−0.000470	−0.0825***	−0.0417	−0.0249	−0.0400
	(0.000848)	(0.000878)	(0.000838)	(0.000951)	(0.0313)	(0.0300)	(0.0348)	(0.0439)
ln Sales		−0.0135		−0.0261*		−1.446***		−0.654
		(0.0146)		(0.0147)		(0.500)		(0.677)
Profit Margin		0.235***		0.00846		−0.536		0.210
		(0.0748)		(0.0807)		(2.560)		(3.723)
ln Market Value		−0.0173		−0.0149		0.190		−1.102*
		(0.0137)		(0.0138)		(0.468)		(0.638)
Cash Liquidity		0.110		0.192*		0.721		7.674*
		(0.104)		(0.0984)		(3.543)		(4.540)
Inst. Own.		0.0241		−0.183**		4.103*		−11.89***
		(0.0634)		(0.0796)		(2.168)		(3.674)
Av. Ind. Tobin's Q		0.00581		0.0163		0.274		1.116**
		(0.0109)		(0.0105)		(0.373)		(0.486)
Av. Ind. Mkt. Value		0.0699***		0.0249*		1.337***		0.321
		(0.0113)		(0.0128)		(0.387)		(0.590)
Entrench. Index		0.0156**		0.0135		0.446		1.187**
		(0.00793)		(0.0100)		(0.271)		(0.463)
Year dummies		Y		Y		Y		Y
Observations	1,284	1,284	1,095	1,095	1,284	1,283	1,095	1,095
R^2	0.007	0.139	0.003	0.089	0.005	0.094	0.000	0.085

Panel A. Conditional Independence Test – Merger Probability

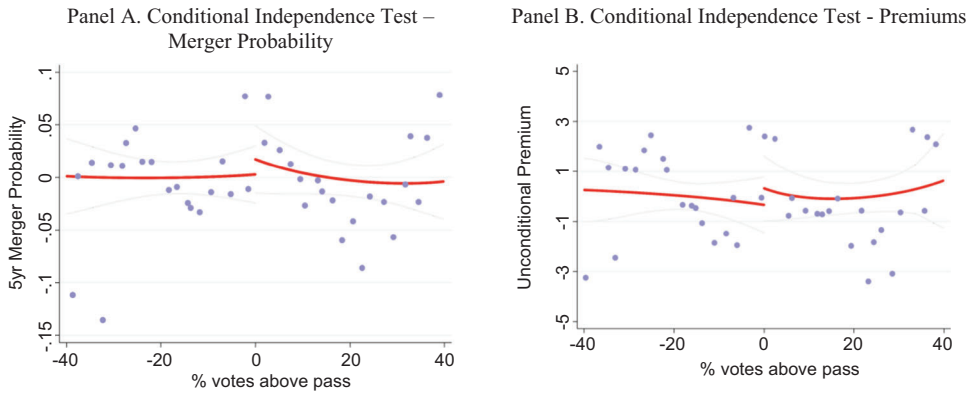Panel B. Conditional Independence Test - Premiums

Figure 2. Conditional independence (CIA) tests. Residuals of two independent linear models (one to each side of the discontinuity) using the same covariates as in the matching model. Dots represent the simple means by bins of 2% vote intervals. (Color figure can be viewed at wileyonlinelibrary.com)

conditionally independent in the $D = 0$ (votes did not pass) region.[21] Column (4) shows that vote and takeover probability are also conditionally independent in the $D = 1$ (votes passed) region. A similar pattern emerges for the unconditional premium in columns (5) (large, significant effect) and (6) (smaller, insignificant effect). In columns (7) and (8), for $D = 1$ there is no significant relationship between the unconditional premium and vote (column (7)), and the coefficient continues to be insignificant in column (8). Table IA.IX in the Internet Appendix reports similar results for the adjusted votes measure.

Complementing the formal CIA analysis, in Figure 2 we plot the residuals of regressions that include the covariates in Table III excluding shareholder votes. In particular, Figure 2 plots firm outcomes (takeover probability in Panel A and unconditional premium in Panel B) against the residuals obtained from regressing the outcomes on our model, on each side of the threshold. The figure plots the residual means in 2% bins and uses local linear regression estimation of the outcome variables as a function of the vote. If the CIA holds after we condition on our model, the remaining relationship between firm outcomes (takeover probability or premium) and the vote outcome should be relatively flat. We see that the estimated relationship is indeed statistically flat on both sides of the threshold for both variables (and within the confidence bands), indicating that the model does a good job making the running variable uncorrelated with potential outcomes along the vote support. The same is true for the adjusted vote measure as can be seen in Figure IA.4, Panels A and B, in the Internet Appendix.

---

[21] Note that the R^2 in odd-numbered columns is low, meaning that factors other than the vote outcome help explain whether a firm is taken over, but what is important for this test is that those other factors are not omitted variables in our regressions that would determine assignment (i.e., correlated with outcomes and the vote).

Having made the running variable, which determines assignment to treatment, conditionally independent of outcomes, we move on to using matching methods to compare treated and control groups. We first test whether the calculated propensity scores for treatment and control groups pass the common support test. The logit model for the propensity score is calculated using the same model as in the CIA tests. (See the common support test in Figure IA.2 in the Internet Appendix. The formal balancing test (Dehejia and Wahba (1999)) also shows that the covariates are balanced).

B.2. Results using the CIA-Validated Matching Model

After testing for the CIA and establishing that we have common support, we can match firms on either side of the discontinuity based on our model. First, as in Angrist and Rokkanen (2015), we use the estimated propensity score (see Table IA.VI in the Internet Appendix) to provide a propensity-score-weighted matching estimator of the effect of passing a shareholder-sponsored proposal to remove an ATP. This amounts to weighting treated observations ($D = 1$) by $1/p$ and control observations ($D = 0$) by $1/(1-p)$, where p is the propensity score estimated using our model. We also add to the reweighted regression the variables included in the CIA model as controls to reinforce the matching procedure. The results are reported in Table IV, Panel A. For *Simple Vote Rule* we find that passing an ATP leads to a 4.1% increase in the probability of takeover (column (1)) and a 2.6% increase in the unconditional premium (column (2)). The analogous results for *Vote Adjusted for Abstentions* are 3.6% and 2.3%, respectively.

We also obtain similar results if we use a different matching estimator, such as the nearest-neighbor matching estimator with replacement (Table IV, Panel B), which delivers a 3.4% increase in takeover probability and a 2.5% increase in the unconditional premium. The results for *Vote Adjusted for Abstentions* are virtually unchanged, with estimates of 3.25% and 2.4%, respectively.

Note that, in general, while the adjustment for the treatment of abstentions is small, it can still affect the RDD estimates, which are based largely on the observations near the majority threshold.[22] However, this adjustment has almost no effect on the estimators based on matching that rely on the full set of available observations.

Three results are noteworthy here. First, confirming what we find when comparing results at the 10% interval relative to those at the discontinuity, the estimates away from the discontinuity are smaller than the discontinuity estimates, suggesting that firms around the discontinuity (with contentious votes) stand to benefit more from removing ATPs than firms away from the discontinuity, on average.[23] Second, the results away from the discontinuity are still

---

[22] Within our sample, abstentions account for an average of 4.8% of the votes (median 2.5%) for those observations for which we have full information about them.

[23] Unfortunately, one cannot apply this estimation strategy to cumulative abnormal returns (CARs) on the day of the vote itself. This is because, while the CAR for firms at the

Table IV
Matching Estimates of the Unconditional Premium in the Full Sample

This table reports CIA estimates of the effect of passing a G-index proposal on the takeover probability and the unconditional premium (four weeks before announcement to completion). Panel A reports results from a linear reweighting estimator and Panel B reports results from a nearest-neighbor matching procedure with replacement and two matches per observation. Controls are the same as in Table III: ln Sales, Profit Margin, ln Market Value, Cash Liquidity, Percentage of Institutional Ownership, Average Industry Tobin's Q, Average Industry Market Value, and Entrenchment Index. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

	Simple Vote Rule		Vote Adjusted for Abstentions	
	Takeover Probability (1)	Unconditional Premium (2)	Takeover Probability (3)	Unconditional Premium (4)
Panel A: Propensity Score Weighting				
Yes	4.1%**	2.59***	3.62%*	2.34**
	(1.97)	(0.97)	(1.9)	(1.00)
t-Statistic	2.05	2.66	1.83	2.33
Model	Y	Y	Y	Y
Observations	2,123	2,123	2,129	2,129
Panel B: Nearest-Neighbor Matching				
Yes	3.36%*	2.53***	3.25%*	2.39***
	(1.92)	(0.81)	(1.92)	(0.80)
t-Statistic	1.74	3.10	1.69	2.98
Observations	2,379	2,379	2,379	2,379

positive, significant, and economically large. The mean takeover probability (over five years) in this sample is 14%, and voting to remove an ATP increases that probability by 4.1 percentage points. Correspondingly, the mean unconditional takeover premium is 4.8%, and voting to remove an ATP increases the premium by 2.6 percentage points. Third, these matching estimates can be interpreted as *causal* for a broad set of firms under stronger identification assumptions than the RDD assumptions. Although our sample does not comprise the full set of listed firms in the United States, it represents a substantial share of the S&P 1500 index (931 distinct firms). See Section II in the Internet Appendix for further discussion and tests of heterogeneous effects along the voting support.[24]

discontinuity is the surprise outcome that reveals information (thus reflecting the full value of the vote, which the paper estimates), returns away from the discontinuity are likely expected by the market and therefore contain no information on the vote (see section 1B in Cuñat, Gine, and Guadalupe (2012)).

[24] As additional robustness checks, we evaluate whether these effects are the result of voting on any proposal—rather than of voting to remove an ATP. In Table IA.VII, Panels A and B, in the Internet Appendix, we replicate the analysis in Tables II and III using other (non-antitakeover)

Before moving on to analyze conditional premiums, it is worth noting that our estimates capture the effect of voting to remove the provision, but since votes are not binding (the decision as to whether to follow shareholders' recommendation is left to managers), they are estimated on the basis of those firms that implemented the proposal because it passed and that would not have implemented it otherwise. In other words, these are ITT effects. To obtain the effect of treatment on the treated (the effect on outcomes that operates through the actual removal of the provision), they need to be re-scaled by the inverse of the change in the probability of removing the provision following a positive vote. Although we cannot estimate these conversion factors directly from our data, we can approximate them using the reported implementation differentials from the pre-existing literature. This would imply multiplying the regression discontinuity results by a factor of three and the matching results by a factor between 1.2 and 1.7 to obtain a Wald estimate of the treatment on the treated.[25]

Finally, we find that these effects emerge only when voting to remove ATP. Voting to drop other types of provisions has no effect on takeover probabilities or premiums, which suggests the results are not related to "voting" per se but more specifically to the takeover channel (see Table IA.VII in the Internet Appendix).

V. Effect on the Conditional Premium

In Section IV, we obtain causal estimates for the effect of treatment on the unconditional premium ΔY and the takeover probability ΔP. However, we would also like to recover the effect on the premium itself, β, that is, the expected premium after accounting for selection that a given firm would observe if it removed the ATP. Given the potentially quite strong selection in the data (our estimated ΔP is rather large), it is not possible to infer the value of β from either ΔY or the difference in realized premiums.

The value of β can be bounded using the method in Lee (2009). The proposed bounds rely on the assumption of monotonicity of the effect of ATPs on the

shareholder votes. We find no effect of those proposals on either the takeover probability or the premium. The results are particularly different for Table IA.VII, Panel B, in the Internet Appendix, which estimates the effects over the entire vote support, while they are heterogeneous and imprecise in Table IA.VII, Panel A, in the Internet Appendix, near the discontinuity, where we cannot always rule out the possibility that they are different from those in Table II. Taken as a whole, the results suggest that what drives our main results is not just some signal around shareholder activism (as proposed by Bach and Metzger (2015)), which should arise after any type of vote, but rather an effect that appears only after the removal of an ATP.

[25] The conversion factors take into account in which part of the distribution of votes they are estimated, to match the estimation sample in our paper. The conversion factor at the discontinuity comes from Cuñat, Gine, and Guadalupe (2012). For results outside the discontinuity, we use estimates from Popadak (2014) (using data from Shark Repellent) and Bach and Metzger (2015) (using data from Voting Analytics). Note that the data used to calculate the conversion factors differ from ours in terms of period, provision type, and vote measurement, so they have to be taken as approximations, given that the marginal firms that implement proposals may differ across them.

selection criterion. That is, the assumption that the causal effect of ATPs on the probability of a takeover can be heterogeneous, but must run in the same direction across firms, (i.e., always positive or always negative).[26] The bounds are calculated by trimming the distribution of premiums of the treated group. The trimming procedure can be seen as implementing the best- and worst-case selection scenario, given the estimated change in the probability of a takeover.[27]

A. Bounding β in the Whole Vote Support

In Table V, Panel A, column (1) estimates the bounds proposed by Lee (2009) calculated for the entire vote support. The procedure requires that the estimates of the coefficients that determine selection in the first stage be interpreted as causal. We achieve this by using the same linear reweighting as in Table IV, that is, we use the weights obtained using the propensity score and we assume that the same CIA holds. This method yields estimates of β for the four-week premium that are bounded between a nonstatistically significant –2.3% and a significant 5.8%. This means that the direct premium effect of dropping an ATP on a given targeted firm is positive.

The remaining columns in Panel A of Table V use additional premium measures in different windows for the purposes of robustness. Column (2) reports the effect on the target premium computed as the change in price one week before announcement until completion (i.e., a shorter run-up) and column (3) reports abnormal returns using the Fama-French factors plus momentum (FFM) factors in a short window (±five days) around the announcement. These shorter premium windows show significant positive bounds both for the lower and the upper bounds (between 5% and 11.5%). Columns (4) to (6) employ longer time windows. Column (4) reports CARs for a very long window—from the day of the vote to the takeover announcement day plus one. This specification accounts for the fact that some of the expected effect of the merger could have been incorporated as early as the day of the vote. Columns (5) and (6) report the CARs using the FFM factors from (–42, 5) trading days around the announcement and days (–42, until completion), respectively. These latter tests allow us to assess the difference between using announcement and completion dates. All estimates unambiguously show a positive premium effect. What is noteworthy here is that while, by construction, a bounding strategy (such as Lee (2009)) is likely to give broad bounds, the results that we obtain allow us to reject the hypothesis that the effect on the conditional premium is

---

[26] Note that this is a weaker assumption than the standard monotonicity assumption necessary for instrumental variable regressions.

[27] In our application, the calculation of the bounds involves first calculating the increase in the probability of a takeover induced by the treatment, relative to the probability of the treated firms $q = [(\Pr(Z^* > 0 \,|\, D = 1)) - \Pr(Z^* > 0 \,|\, D = 0)]/\Pr(Z^* > 0 \,|\, D = 1)$. Then, from the observed population of mergers in the treated group (those for which the ATP proposal passes), we compute the upper (q) and lower ($1 - q$) quantile of observed premiums. The upper (lower) bound of β is then calculated as the average of observed takeover premiums above (below) the lower (upper) quantile minus the average premium of the control group (firms that did not pass the ATP proposal).

Table V
Target Conditional Premiums, Lee Bounds Estimates

This table reports the results of the effect of passing a G-index proposal on different premium measures for the target company. Panel A estimates are obtained using the Lee (2009) methodology and propensity score matching to account for selection in the universe of targeted companies. Panel B estimates restrict the sample to votes within the (−10,10) interval. Column (1) reports the effect on the target premium computed as the change in price four weeks before announcement until completion. Column (2) reports the effect on the target premium computed as the change in price one week before announcement until completion. Columns (3) and (4) report premiums based on cumulative abnormal returns using the FFM factors for different windows (±5 days) and (Vote/+1 day), both relative to the announcement date. The run-ups for the target company are computed as the abnormal returns from (−42, 5) trading days around the announcement in column (5) and as the abnormal return (−42, until Completion) in column (6), always using the FFM factors. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

	Premium Four Eeeks before Announce. to Completion (1)	Premium One week before Announce. to Completion (2)	CAR (−5,5) FFM (3)	CAR (Vote,Ann+1) (4)	Run-up (−42,5) FFM (5)	Run-up (−42, Completion) FFM (6)
			Simple Vote Rule			
		Panel A: Upper and Lower Sharp Lee Bounds (with Propensity Score Weights, Full Sample)				
Lower Bound Estimation						
Yes	−2.3	4.95*	5.27***	2.6	2.72	3.29
	(3.23)	(2.9)	(2.11)	(19.3)	(2.93)	(3.39)
Z	−0.71	1.71	2.49	0.13	0.93	0.84
Upper Bound Estimation						
Yes	5.75**	11.49***	10.12***	52.24**	9.17***	16.52***
	(2.89)	(2.48)	(1.92)	(16.16)	(2.61)	(3.93)
R^2/Z	1.99	4.62	5.25	3.23	3.51	4.80
# Sel. Obs.	418	416	418	406	418	408
Observations	2,379	2,379	2,379	2,379	2,379	2,379
		Panel B: Upper and Lower Sharp Lee Bounds in (−10, 10) Interval				
Lower Bound Estimation						
Yes	−2.02	−1.18	−1.79	−52.24**	2.53	−0.27
	(4.57)	(4.56)	(3.23)	(23.1)	(5.27)	(7.45)
Z	−0.44	−0.26	−0.55	−2.25	0.48	−0.04
Upper Bound Estimation						
Yes	14.63**	22.93***	11.06***	47.73*	20.96***	27.39***
	(5.99)	(6.80)	(4.11)	(26.96)	(5.71)	(8.279)
R^2/Z	2.44	3.37	2.69	1.77	3.67	3.30
# Sel. Obs.	135	134	133	124	133	130
Observations	883	883	883	883	883	883

negative.[28] The bound estimates solve the selection problem and rely on very weak assumptions at the cost of not determining the coefficient with precision. However, in our case the bounds rule out the possibility that the conditional premium is negative. In Section VII, we explore a number of hypotheses that may explain this nonnegative effect.

The results of this section use the full distribution of votes as supported by the CIA assumption tests of Section IV.B.1. These are necessary inputs that allow us to calculate the decomposition of the unconditional premium in Section VI.

B. Bounding β Near the Discontinuity Threshold

An alternative approach is to perform the decomposition of the unconditional premium using results at the majority threshold. Unfortunately, the bounding technique relies on having an empirical distribution of votes and mergers when vote outcomes tend to the majority threshold, which does not exist in a finite sample. Alternatively, we can rely on a bandwidth around the discontinuity that is sufficiently narrow to approach the RDD intuition but sufficiently broad to get meaningful results.

In Table V, Panel B, we report results on a 10% vote interval around the discontinuity. The results are generally consistent with those in Panel A but sometimes less precise at the lower bound. There are two mechanical reasons for this loss of precision. First, as we approach the discontinuity, we have fewer observations and mergers and hence we lose power. Second, the effect on the probability of a merger is higher in the RDD results than for the matching estimator, resulting in more severe selection problems that take the bounds further apart. In general, it is not possible to maintain a reasonable sample size and take the limit of the bounds to the discontinuity without imposing additional functional assumptions.[29,30]

The results show that the upper bound is consistently positive and significant. The lower bound is generally statistically indistinguishable from zero, except for the most volatile premium measure (that measuring returns from the vote until merger completion in column (4)), which yields broad and uninformative bounds.[31]

---

[28] All of the results in Table V, Panels A and B, are based on the simple vote rule. Results are similar using the adjusted vote rule. See Table IA.X, Panels A and B, in the Internet Appendix.

[29] In Table V, Panel B, we obtain bounds for the conditional merger premium (four-week and one-week premium) for different bandwidths around the discontinuity. These include the intervals used by the local regression estimators in Table III with their relevant weights. At very close intervals, the combination of a smaller sample and more severe selection problems generates broad bounds. These are mostly distributed on the positive side, especially for the one-week premium, which is less volatile, but a negative lower bound cannot be ruled out.

[30] See also Table IA.VIII in the Internet Appendix for an RDD calculation of the unconditional premiums in Table VI.

[31] Throughout the paper, we use raw returns after checking that our results are not driven by outliers. We replicated all of our main tests, using winsorized variables, and obtain similar results. These results can be found in earlier versions of this paper.

VI. Decomposing the Unconditional Premium: Takeover Probability, Takeover Premium, and Selection Effects

We now have all the elements necessary to evaluate the contribution of price, probability, and selection effects to the overall estimated unconditional takeover premium ΔY using the decomposition in equation (2).[32] We run the decomposition for the full sample (based on the CIA-validated matching model) and for the $\pm 10\%$ interval. Table VI reports the results.

For the full sample, we find that 52% of the premium is driven by the takeover probability effect (note that the treatment effect on the takeover probability is estimated without selection bias, so this number does not change with the bounding exercise). Using our lower bound estimate for β (–2.3), we find that selection accounts for 61% and that the premium contributes negatively (–13%). However, from Table V, the lower bound premium is not statistically distinguishable from zero and other windows for the premium yield unambiguously positive estimates. With our upper bound estimate for β (5.7), 32% of the unconditional premium is explained by the effect on premiums, holding the population constant, and 17% by selection.

At the discontinuity (Panel B), the effect of takeover probabilities the largest at 57% and the premium effect ranges from 84% to –12% (although the negative estimate is based on a negative premium effect that is statistically indistinguishable from zero). The selection effect at the discontinuity cannot be signed, although it is potentially large, ranging between 55% and –41%.

The results above imply that while half of the value implications of dropping an ATP can be attributed to an increased probability of experiencing a takeover, nonnegligible amounts are driven by the positive premium. Selection effects are positive and large for the full sample and imprecisely estimated for the RDD sample, although still potentially large. This paints a very different picture than existing literature (which does not address selection and endogeneity) and shows that failing to account for the endogenous selection of targets induces substantial bias.

VII. Understanding Positive Target Premiums

We next explore the possible drivers for the positive effect on the conditional target premium by looking at what else changes when firms pass a proposal. Given that the population of merged companies changes with the removal of an ATP, we analyze these effects using Lee bounds (full descriptive statistics of all the variables in this section can be found in Table IA.I in the Internet Appendix). We start by analyzing the effects for the *Simple Vote* in whole matched sample in Table VII and then focus on the effects closes to the discontinuity (in

---

[32] We take the estimates for ΔP and ΔY from column (1) of Table II (RDD) and columns (1) and (3) of Table IV (matching). We compute $\Pr[Z^* > 0 \mid D=1] = 13.5$ using the probabilities of each observation being treated and $E[Y \mid D=0, Z^* > 0] = 29.6$ using the probabilities of each observation not being treated from the matching model. The bounds on β and the selection term come from column (1) in Table V.

Table VI
Decomposing the Shareholder Value Effect

This table provides a decomposition of the change in shareholder value induced by the passing of a proposal to eliminate an ATP. We provide an estimate of the three components that affect shareholder value—changes in the premium, changes in the probability of a takeover, and changes in the population of firms that are put into play. We provide both the lower and the upper bound values since we use Lee (2009) to estimate the change in takeover premium β. In column (1), the Change in Shareholder Value is the unconditional takeover premium under the CIA model in Panel A, and the RDD IK estimate in Panel B. In column (2), Premium Effect is the change in takeover premium β times the baseline probability of merger ($\Pr[Z^*>0 \mid D = 1]$). In column (3), Takeover Probability Effect is the result of the change in the probability of merger ($\{\Pr[Z^*>0 \mid D = 1] - \Pr[Z^*>0 \mid D = 0]\}$ times the baseline premium ($E[Y \mid D = 0 , Z^*>0]$). In column (4), we provide an estimate of the selection effect. Using the probabilities of the matching model, we calculate that the baseline probability is 14.4 and the baseline premium is 32.7.

Simple Vote Rule			
Change in Shareholder Value	Premium Effect	Takeover Probability Effect	Selection Effect
ΔY (1)	$\beta * \Pr[Z^*>0 \mid D = 1]$ (2)	$\{\Pr[Z^*>0 \mid D = 1] - \Pr[Z^*>0 \mid D = 0]\} * E[Y \mid D = 0 , Z^*>0]$ (3)	$\Pr[Z^*>0 \mid D = 1] * \{ E[Y \mid D = 1 , Z^*>0] - E[Y \mid D = 1 , V > -\mu_2]\}$ (4)
Panel A: Using Validated Matching for Uncond. Premium and Probability from Table IV			
Lower Bound Estimation of $\beta = -2.3$			
2.59%	−0.33	1.34	1.59
	−13%	**52%**	**61%**
Upper Bound Estimation of $\beta = 5.7$			
2.59%	0.82	1.34	0.43
	32%	**52%**	**17%**
Panel B: Using RDD-IK for Unconditional Premium and Probability from Table II (−10,10)			
Lower Bound Estimation of $\beta = -2.02$			
2.61%	−0.30	1.48	1.42
	−12%	**57%**	**55%**
Upper Bound Estimation of $\beta = 14.6$			
2.61%	2.19	1.48	−1.06
	84%	**57%**	**−41%**

the ±10% interval) in Table VIII. Note that these are not only different estimation strategies but also that the effects are evaluated at different points of the distribution of votes (weighted average of the full sample versus close-call votes) that can yield different point estimates.[33] Table IA.XII, Panels A and B, in the Internet Appendix replicate the results using the *Adjusted* vote. Since the results are extremely similar for both vote definitions, we discuss only the

---

[33] See Section II of the Internet Appendix for a discussion of the heterogeneous effects of ATPs at different levels of vote support and for an analysis of the difference between the two samples.

Simple vote here and refer the reader to the Internet Appendix for the exact point estimates using the *Adjusted* vote.

We first evaluate the total value/synergies created by deals in less protected versus more protected firms. We find that voting to remove an ATP leads to more value-creating deals using a number of different measures, both for the full sample (Panel A of Table VII) and for close-call votes in the in the ±10% interval (Panel A of Table VIII). For example, for the full sample (for close-call votes), bidder and target firms are between 16% and 28% (14% and 38%) more likely to belong to the same two-digit SIC industry relative to a sample mean of 63%. These results indicate that the deals are more likely to be related mergers with greater synergies than financial or diversifying mergers. Column (2) shows that in less protected deals, targets are matched with relatively larger acquirers, as measured by their market capitalization. The ratio of target to acquirer market capitalization four weeks before the announcement is 0.7 to 1.3 lower in the full sample (relative to a mean of 1.11 and standard deviation of 4.35; the interval is not significant for close-call votes). This finding suggests that on average the positively selected targets are matched to relatively more valuable and potentially productive acquirers, but this effect is not present closer to the discontinuity. If we cumulate the dollar value of the premium of the target and the acquirer, which measures the total value/synergies created by the deal, the upper bound estimate of the effect is quite large and positive both in dollar value terms (up to US$7.3 billion higher for the full sample and US$9.2 billion higher for close-call votes, column (3)) and as a share of total market cap (up to 14% of the target and acquirer's value for the full sample and 22% for close-call votes, column (4)). The lower bound of the synergy estimates is not significant in either sample. Nonetheless, the results suggest that, if anything, there is net additional value creation in the market when ATPs are removed, with this true both for close-call votes and for the full sample. This finding is important as it suggests that the presence of ATPs suppresses the realization of deals that have more value-creating potential and hence potentially represents a net loss to the economy.

The results above indicate that less protected firms receive a higher premium and the deals of those firms create more market value. But what happens to the acquirer premium? Here we find differences between the two samples, as can be seen in columns (5) to (8) in Tables VII and VIII. In the full sample we cannot clearly sign the acquirer premium—it does not seem that acquirers are systematically able to extract a higher share of the synergies created given that the results depend heavily on the measure of premium used. In contrast, for close-call votes (Table VIII) we find a systematic pattern: the upper bound of the effect on the acquirer premium is always significant, large, and positive, where the lower bound is insignificant but also positive in several cases. We interpret these results as indicating that acquirers can extract/appropriate some of the value created in close-call votes, but less so over the full sample of deals.

The share of the surplus extracted by acquirers can change with changes in relative bargaining power when targets are less protected or with changes in

Table VII
Merger Effects—Lee Bounds with Propensity Score Weights (Full Sample)

This table reports results on the effect of passing a G-index proposal on different merger outcomes. All estimates are obtained using the Lee (2009) methodology to account for selection in the universe of targeted companies. Panel A presents results for different measures of matching and acquirer premium. Column (1) reports the effect on the likelihood of the target and acquirer being in the same two-digit SIC code, column (2) on the relative size of target versus acquirer, column (3) on a measure of total synergies, and column (4) on total synergies as a percentage of total market capitalization. Column (5) reports the effect on the acquirer premium (change in price four weeks before announcement until one day after). Column (6) reports a premium based on cumulative abnormal returns (FFM) on a (−5/+5) window around announcement. The run-up of the acquirer is measured as the abnormal return on a (−42/+5) window around announcement in column (7) and as (−42/Completion) in column (8). Panel B, columns (1) to (3) report the effect on the number of bidders, the deal being unsolicited and the deal being challenged, column (4) reports the effect on the percentage of stock paid for the target, and columns (5) and (6) present the effects on activism events. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Panel A: Value Creation and Value Split

| | Matching | | | | Acquirer Premium | | | |
| | Simple Vote Rule | | | | | | | |
	Same Two-Digit SIC (1)	Size Target Rel. to Acquiror (2)	Total Synergies FFM (3)	Total Synergy/Total Mkt Cap (4)	Acquirer Premium (5)	Acquirer CAR(−5,5) (6)	Run-Up CAR(−42,5) (7)	Run-Up CAR (−42,Comp) (8)
Lower Bound Estimation								
Yes	0.161**	−1.31**	−321,515	−0.03	−8.81***	−4.46***	−7.14**	−3.12
	(0.06)	(0.51)	(2,446,489)	(0.03)	(3.18)	(1.37)	(2.90)	(4.45)
Z	2.30	−2.57	−0.13	−1.01	−2.77	−3.25	−2.46	−0.70
Upper Bound Estimation								
Yes	0.28***	−0.73	7,317,733***	0.14***	−0.25	1.06	5.78**	17.6***
	(0.07)	(0.48)	(1,854,619)	(0.03)	(2.35)	(1.38)	(3.27)	(3.74)
Z	3.95	−1.53	3.95	4.63	0.11	0.77	1.76	4.7
#Sel	418	356	262	262	281	278	278	271
Observations	2,379	2,379	2,379	2,379	2,379	2,379	2,379	2,379

(Continued)

Table VII—*Continued*

Panel B: Competition and Activism

	Competition				Activism	
	Number of Bidders (1)	Unsolicited Deal (2)	Challenged Deal (3)	Stock Percentage (4)	Num 13D Events Two Years Prior (5)	Dummy 13D Event Two Years Prior Annoucement (6)
Lower Bound Estimation						
Yes	0.14**	0.039	0.10***	−26.83***	0.051	0.072
	(0.06)	(0.03)	(0.03)	(7.22)	(0.077)	(0.044)
Z	2.20	1.27	2.66	−3.71	0.66	1.62
Upper Bound Estimation						
Yes	0.315***	0.12***	0.22***	−3.46	0.445***	0.289***
	(0.04)	(0.02)	(0.03)	(5.39)	(0.083)	(0.074)
I	6.40	5.01	7.26	−0.64	5.33	3.87
#sel	418	418	418	217	639	639
Observations	2,379	2,379	2,379	2,379	2,379	2,379

Table VIII

Merger Effects—Lee Bounds in (−10,+10) Interval

This table reports results on the effect of passing a G-index proposal on different merger outcomes. All estimates are obtained using Lee (2009) bounds on the restricted (−10,10) vote interval. Panel A presents results for different measures of matching and acquirer premium. Column (1) reports the effect on the likelihood of the target and acquirer being in the same two-digit SIC code, column (2) on the relative size of target versus acquirer, column (3) on a measure of total synergies, and column (4) total synergies as a percentage of total market capitalization. Column (5) reports the effect on the acquirer premium (change in price four weeks before announcement until one day after). Column (6) reports a premium based on cumulative abnormal returns (FFM) on a (−5/+5) window around announcement. The run-up of the acquirer is measured as the abnormal return on a (−42/+5) window around announcement in column (7) and as (−42/Completion) in column (8). Panel B, columns (1) to (3) report the effect on the number of bidders, the deal being unsolicited, and the deal being challenged, column (4) reports the effect on the percentage of stock paid for the target, and columns (5) and (6) show the effects on activism events. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Simple Vote Rule

Panel A: Value Creation and Value Split

| | Matching | | | | Acquirer Premium | | | |
	Same Two-Digit SIC (1)	Size Target Rel. to Acquiror (2)	Total Synergies FFM (3)	Total Synergy/Total Mkt Cap (4)	Acquirer Premium (5)	Acquirer CAR(−5,5) (6)	Runup CAR (−42,5) (7)	Runup CAR (−42,Comp) (8)
Lower Bound Estimation								
Yes	0.14	−0.24	5,461,211*	0.07	−1.81	−0.42	3.20	8.44
	(0.094)	(1.03)	(3,003,063)	(0.074)	(3.35)	(2.28)	(4.67)	(7.33)
Z	1.46	−0.23	1.83	0.94	−0.54	−0.19	0.69	1.15
Upper Bound Estimation								
Yes	0.38***	−0.05	9,208,777***	0.22***	7.20**	4.68**	13.4***	22.94***
	(0.061)	(0.34)	(2,876,863)	(0.049)	(3.52)	(2.65)	(4.82)	(6.36)
Z	6.18	−0.15	3.20	4.42	2.04	2.04	2.78	3.61
#sel	135	140	101	101	103	103	104	103
Observations	883	883	883	883	883	883	883	883

(Continued)

Table VIII—*Continued*

Panel B: Competition and Activism

| | Competition | | | | Activism | |
	Number of Bidders (1)	Unsolicited Deal (2)	Challenged Deal (3)	Stock Percentage (4)	Num 13D Events Two Years Prior Annouce. (5)	Dummy13D Event Two Years Prior Ann. (6)
Lower Bound Estimation						
Yes	−0.30**	−0.12**	−0.22***	−28.08**	−0.142	−0.054
	(0.08)	(0.042)	(0.051)	(11.04)	(0.200)	(0.096)
Z	−3.75	−3.0	−4.21	−2.54	−0.71	−0.56
Upper Bound Estimation						
Yes	0.16	0.04	0.076	−10.43	−0.049	−0.025
	(0.15)	(0.072)	(0.09)	(11.03)	(0.132)	(0.076)
Z	1.05	1.05	0.79	−0.94	−0.38	−0.34
#sel	135	135	135	88	240	240
Observations	883	883	883	883	883	883

the extent of competition for targets. We cannot measure bargaining power directly, but we can capture competition using several proxies. In Panel B, columns (1) to (4) of Tables VII and VIII we find that competition increases for the full sample, whereas it does not (and actually seems to decrease) for the sample at close-call votes, which would explain the difference in results on the acquirer premium. In the full sample, less protected deals observe between 0.14 and 0.32 more bidders (the mean number of bidders is 1.24 and over 90% of firms have only one or two bidders). In addition, the probability of deal being a challenged is between 10% and 22% higher (sample mean is 16%). These effects are statistically and economically significant. Targets also appear to be subject to more unsolicited deals (with 4% to 12% higher probability, although the lower bound result is not significant), and observe a higher fraction of the deal settled in cash (indicating more competition as in Offenberg and Pirinsky (2015)).[34] For close-call votes, the lower bounds are negative and significant but the upper bound is of the opposite sign and equally large, which may be due to the lack of power to estimate effects in this subsample. However, the overall picture is not of an increase in competition but rather if anything, a decline.

We interpret the evidence above as suggesting that voting to remove an ATP makes entering the bidding contest less attractive in close-call votes, reduces competition, and allows the acquirer to appropriate some of the surplus created. In contrast, for nonclose-call votes, competition between bidders erodes any premium for the acquirer.

Finally, we explore the potential role of activist investors in delivering higher premiums.[35] Greenwood and Schor (2009) provide evidence that activists can put firms into play and therefore collect high takeover premiums. We find that, in the full sample, among firms that are taken over, the likelihood of a 13D activist campaign prior to a merger announcement is higher for firms that passed a proposal to get rid of an ATP.[36] This results does not hold, however, in the close-call sample. Activist investors may therefore be an additional channel through which higher premiums are observed for the full sample, but not for the close-call votes sample.

[34] The use of cash in takeovers has also been linked to overvalued targets (see Shleifer and Vishny (2003), Malmendier, Opp, and Saidi (2016)). Given that we are isolating the effect of ATPs on a given target, our results suggest that dropping an ATP may lead to the overvaluation of the target. However, over- and undervaluation of targets is more likely to operate through the selection of targets, and such effects are captured by the selection part of our decomposition (see Section VI).

[35] An activist investor that acquires more than 5% beneficial ownership is required to disclose crossing the 5% threshold in Schedule 13D within 10 days of crossing the threshold if it intends to influence control.

[36] We found that the unconditional probability of being the target of activists after passing a vote is unchanged (unreported).

VIII. Conclusion

In spite of the attention devoted to the consequences of ATPs, there is limited empirical evidence of their effect on takeovers in terms of probabilities and premiums or of their effect on value creation/destruction economy wide. To investigate the effect of ATPs, one must first jointly assess their effect on takeover probabilities, merger premiums, and target selection.

This paper provides estimates—addressing the endogenous adoption of provisions and target selection—of the effects of ATPs and identifies several channels through which they destroy value. First, having an ATP in place reduces the likelihood of a takeover happening. Second, the deals that take place when a firm is protected by an ATP are "worse" along several dimensions: they involve worse targets, they involve smaller acquirers, they are more likely to be between firms in unrelated businesses and hence are less likely to create value, and they create fewer synergies.

The more protected the firm, the lower the premium paid for the target. We find evidence that for the full sample this is due at least in part to the fact that more protected firms attract less competitive bidding. This is likely because ATPs deter bidders, and also because the worse the firm's governance, the more difficult it will be to realize synergies. This mechanism does not apply, however, to close-call votes, where we see no increase in competitive bidding when less protected and where bidders can realize a positive premium.

In sum , we find no apparent trade-off between takeover price (the premium) and probability—a trade-off typically presented by managers as the rationale for the adoption of ATPs. In our results, both price and probability are significantly lower when an ATP is in place. In fact, the gains from dropping an ATP accrue almost exclusively to the target shareholders.

We find similar results using different identification strategies that employ different vote definitions and are run on different subsamples, each with its own strengths and weaknesses. Specifically, we obtain causal effects at the vote discontinuity for takeover probabilities and unconditional premiums. We are also able to bound the effects on the conditional premium for an interval around the discontinuity as well as away from the discontinuity using CIA-validated matching. When using the latter, we find that our results apply to most of the support of vote outcomes. Thus, while our results cannot necessarily be extrapolated to firms that never hold such shareholder votes, the evidence suggests that they do apply to the majority of firms in this population, which is about-one third of the S&P 1500.

Existing literature fails to account for selection when computing the takeover premium. We show that this selection effect can be quite large, and we provide a framework for assessing how much of the overall expected premium associated with removing an ATP is driven by probability, price (premium), and selection effects.

Although we present new results and answer a number of previously unanswered questions, our analysis leaves a number of questions open. For example, we take ATPs to be identical and do not consider heterogeneity across

types of proposal or firms. Furthermore, if takeover deals are good for shareholders of the target firms and for the economy as a whole, why do so many firms keep ATPs in place? Given that the firms in our sample tend to be large listed companies, our results would be consistent with the view that ATPs have a positive role in young entrepreneurial firms, but become value-destroying for more mature firms (see Johnson, Karpoff, and Yi (2015, 2017)). There is ample evidence of inefficiencies in internal governance and the political economy of decision-making within firms that make such provisions sticky, and mature firms may find themselves off equilibrium with an above-optimal level of anti-takeover protection. We leave these important questions to future research.

REFERENCES

Angrist, Joshua D., and Miikka Rokkanen, 2015, Wanna get away? RD identification away from the cutoff, *Journal of the American Statistical Association* 110, 1331–1344.

Bach, Laurent, and Daniel Metzger, 2015, Why do shareholder votes matter?*Swedish House of Finance Research Paper*13-01.

Bach, Laurent, and Daniel Metzger, 2019, How close are close shareholder votes?*Review of Financial Studies* 32, 3183–3214.

Bange, Mary M., and Michael A. Mazzeo, 2004, Board composition, board effectiveness, and the observed form of takeover bids, *Review of Financial Studies* 17, 1185–1215.

Bates, Thomas, David Becher, and Michael Lemmon, 2008, Board classification and managerial entrenchment: Evidence from the market for corporate control, *Journal of Financial Economics* 87, 656–677.

Bebchuk, Lucian A., John C. Coates, and Guhan Subramanian, 2002, The powerful antitakeover force of staggered boards: Further findings and a reply to symposium participants, *Stanford Law Review* 55, 885–917.

Bertrand, Marianne, and Sendhil Mullainathan, 2003, Enjoying the quiet life? Corporate governance and managerial preferences, *Journal of Political Economy* 111, 1043–1075.

Bulow, Jeremy, and Paul Klemperer, 1996, Auctions versus negotiations, *American Economic Review* 86, 180–194.

Burkart, Mike, Denis Gromb, and Fausto Panunzi, 1998, Why higher takeover premia protect minority shareholders, *Journal of Political Economy* 106, 172–204.

Calonico, Sebastian, Matias Cattaneo, and Rocio Titinuik, 2014, Robust nonparametric confidence intervals for regression-discontinuity designs, *Econometrica* 82, 2295–2326.

Comment, Robert G., and William Schwert, 1995, Poison or placebo? Evidence on the deterrence and wealth effects of modern antitakeover measures, *Journal of Financial Economics* 39, 3–43.

Cotter, James F., Anil Shivdasani, and Marc Zenner, 1997, Do independent directors enhance target shareholder wealth during tender offers?*Journal of Financial Economics* 43, 195–218.

Cuñat, Vicente, Mireia Giné, and Maria Guadalupe, 2012, The vote is cast: The effect of corporate governance on shareholder value, *Journal of Finance* 67, 1943–1977.

Cuñat, Vicente, Mireia Giné, and Maria Guadalupe, 2013, Corporate governance and value: Evidence from "close calls" on shareholder governance proposals, *Journal of Applied Corporate Finance* 25, 44–54.

Cuñat, Vicente, Mireia Giné, and Maria Guadalupe, 2016, Say pays! Shareholder voice and firm performance, *Review of Finance* 20, 1799–1834.

Dehejia, Rajeev H., and Sadek Wahba, 1999, Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs, *Journal of American Statistical Association* 94, 1053–1062.

Eckbo, B. Espen, 2009, Bidding strategies and takeover premiums: A review, *Journal of Corporate Finance* 15, 149–178.

Gompers, Paul, Joy Ishii, and Andrew Metrick, 2003, Corporate governance and equity prices, *The Quarterly Journal of Economics* 118, 107–156.

Gormley, Todd A., and David A. Matsa, 2016, Playing it safe? Managerial preferences, risk, and agency conflicts, *Journal of Financial Economics* 122, 431–455.

Greenwood, Robin, and Michael Schor, 2009, Investor activism and takeovers, *Journal of Financial Economics* 92, 362–375.

Hainmueller, Jens, Andrew B. Hall, and James M. Snyder, Jr., 2015, Assessing the external validity of election RD estimates: An investigation of the incumbency advantage, *Journal of Politics* 77, 707–720.

Harris, Ellie G., 1990, Antitakeover measures, golden parachutes, and target firm shareholder welfare, *RAND Journal of Economics* 21, 614–625.

Heckman, James, 1979, Sample selection bias as a specification error, *Econometrica* 47, 153–161.

Imbens, Guido, and Karthik Kalyanaraman, 2012, Optimal bandwidth choice for the regression discontinuity estimator, *Review of Economic Studies* 79, 933–959.

Johnson, William C., Jonathan M. Karpoff, and Sangho Yi, 2015, The bonding hypothesis of takeover defenses: Evidence from IPO firms, *Journal of Financial Economics* 117, 307–332.

Johnson, William C., Jonathan M. Karpoff, and Sangho Yi, 2017, The lifecycle effects of firm takeover defenses, SSRN Working paper.

Karpoff, Jonathan M., Robert Schonlau, and Eric Wehrly, 2017, Do takeover defense indices measure takeover deterrence?*Review of Financial Studies* 30, 2359–2412.

Lee, David, 2009, Training, wages, and sample selection: Estimating sharp bounds on treatment effects, *Review of Economic Studies* 76, 1071–1102.

Lee, David, and Thomas Lemieux, 2010, Regression discontinuity designs in economics, *Journal of Economic Literature* 48, 281–355.

Maksimovic, Vojislav, and Gordon Phillips, 2001, The market for corporate assets: Who engages in mergers and asset sales and are there efficiency gains?*Journal of Finance* 56, 2019–2065.

Malatesta, Paul H., and Ralph A. Walkling, 1988, Poison pill securities: Stockholder wealth, profitability, and ownership structure, *Journal of Financial Economics* 20, 347–376.

Malmendier, Ulrike, Marcus M. Opp, and Farzad Saidi, 2016, Target revaluation after failed takeover attempts: Cash versus stock, *Journal of Financial Economics* 119, 92–106.

Maug, Ernst, and Kristian Rydqvist, 2009, Do shareholders vote strategically? Voting behavior, proposal screening, and majority rules, *Review of Finance* 13, 47–79.

McCrary, Justin, 2008, Manipulation of the running variable in the regression discontinuity design: A density test, *Journal of Econometrics* 142, 698–714.

Morck, Randall, Andrei Shleifer, and Robert Vishny, 1990, Do managerial objectives drive bad acquisitions?*Journal of Finance* 45, 31–48.

Offenberg, David, and Christo Pirinsky, 2015, How do acquirers choose between mergers and tender offers?*Journal of Financial Economics* 116, 331–348.

Popadak, Jillian, 2014, A corporate culture channel: How increased shareholder governance reduces firm value, SSRN Working paper.

Pound, John, 1987, The effects of antitakeover amendments on takeover activity: Some direct evidence, *Journal of Law and Economics* 8, 3–35.

Ryngaert, Michael, 1988, The effect of poison pill securities on shareholder wealth, *Journal of Financial Economics* 20, 377–417.

Scharfstein, David, 1988, The disciplinary role of takeovers, *The Review of Economic Studies* 55, 185–199.

Schoar, Antoinette, 2002, Effects of corporate diversification on productivity, *The Journal of Finance* 57, 2379–2403.

Schoar, Antoinette, and Ebonya L. Washington, 2011, Are the seeds of bad governance sown in good times? National Bureau of Economic Research paper 17061.

Shleifer, Andrei, and Robert Vishny, 2003, Stock market driven acquisitions, *Journal of Financial Economics* 70, 295–311.

Stein, Jeremy C., 1988, Takeover threats and managerial myopia, *Journal of Political Economy* 96, 61–80.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication code.