Discussion Papers In Economics And Business

Takeovers and taxes: Estimates from a two-sided matching model

Kazuki Onji, Roger H. Gordon and Tue Gørgens

Discussion Paper 25-10

September 2025

Graduate School of Economics
The University of Osaka, Toyonaka, Osaka 560-0043, JAPAN

Takeovers and taxes: Estimates from a two-sided matching model

Kazuki Onji*[1], Roger H. Gordon[2] and Tue Gørgens[3]

[1]The University of Osaka
[2]University of California, San Diego
[3]Australian National University

Abstract

In assessing the influence of taxes on corporate takeovers, sorting among firms poses complications. We employ a recently developed matching model that explicitly accounts for sorting behaviors, applying it for the first time to analyze tax implications in corporate takeover markets. Examining Japanese M&A transactions (1999–2018) between publicly-traded corporations, our estimates reveal heterogeneous effects, where the role of acquirer information and control perceptibly moderates the value of loss carryforwards across different transaction types. However, the estimated value generated from loss carryforwards is modest, indicating the stringency of Japan's anti-avoidance rules.

JEL Classification: H25, H32, G34, L20
Keywords: *Corporate tax, corporate reorganization, M&As, matching estimator*

*Onji (corresponding author): kazuki.onji@econ.osaka-u.ac.jp; Gordon: rogordon@ucsd.edu; Gørgens: tue.gorgens@startmail.com. We would like to thank the seminar and conference participants at ANU, Waseda, SOPE2022, AMES2022(Shenzhen), Tokyo U.(CIRJE), IIPF2023, NTA2023. Onji would like to acknowledge support by JSPS KAKENHI Grant Number 19K01695/T23K014200.

1 Introduction

Tax lawyers and policymakers widely regard tax considerations as pivotal in shaping corporate takeover decisions. Tax lawyers often emphasize that deal structure can lead sellers and buyers to realize substantially different after-tax surpluses, making tax attributes a significant determinant or deal pricing.[1] To preempt tax-motivated transactions that erode government revenue without generating social surplus, policymakers incorporate anti-avoidance provisions into tax codes governing corporate takeovers. At the same time, they aim to avoid overly restrictive rules that could suppress economically efficient trades (Vanistendael, 1998).[2]

Despite the practical importance of tax arrangements emphasized by practitioners, early empirical evidence on the role of taxes in domestic takeovers has yielded mixed results (see surveys by Shackelford and Shevlin, 2001; Betton, Eckbo, and Thorburn, 2008; Hanlon and Heitzman, 2010). In contrast, research on foreign direct investment—primarily involving cross-border mergers and acquisitions—consistently finds that higher tax rates reduce transaction volumes (Head and Ries, 2008; Gordon and Hines, 2002). Over time, however, our understanding of domestic contexts has deepened. Recent studies employing more sophisticated methodologies increasingly demonstrate that taxes do influence domestic takeover activity and have begun to elucidate the mechanisms through which these effects operate (Ohrn and Seegert, 2019; Blouin et al., 2021; Onji and Gordon, 2021; Bührle et al., 2023).[3]

This study aims to extend the literature on the roles of taxes in domestic takeover markets by exploring an empirical method that fits the context but remains untested for tax research. Two features of takeover markets are salient. First, outcomes depend not just on a buyer's willingness to purchase but also on a seller's willingness to sell, as evidenced by the existence of a 'poison pill' that target managers can use to resist hostile takeovers. That is, the decision to form a union is two-sided and thus prior relationships can matter. Second, a target's tax attributes, such as loss

[1] For instance, in newsletters targeting business audiences, tax lawyers often emphasize whether a target firm's net operating losses (NOLs) can be transferred to the acquirer—an outcome that hinges on whether the transaction is structured as an asset sale or a stock deal under U.S. tax law (Bottomlee, Bazar, and Walker, 2009). These tax attributes can significantly influence deal pricing.

[2] For example, if target shareholders are subject to taxation upon selling stock, they may be discouraged from participating in transactions, leading to lock-in effects. Vanistendael (1998) notes that most developed countries permit tax-free reorganizations when business continuity and shareholder interests are preserved.

[3] Ohrn and Seegert (2019) examine investor-level taxation and find that when the capital gains tax rate is lower than the dividend tax rate, acquirers are incentivized to reward shareholders through capital gains—via acquisitions—rather than through dividend payments. Blouin et al. (2021) investigate the impact of additional cash generated by the Domestic Production Activities Deduction enacted in 2004, showing that increased liquidity leads to greater acquisition activity.

carryforwards (LCFs), may be valued differently by different sellers. To account for these two features, Hanlon and Heitzman (2010) suggest a 'marriage model' (first considered by Auerbach and Reishus, 1988).

Hanlon and Heitzman (2010) argue that the estimation of tax effects should account for these two features and suggest a 'marriage model' first considered by Auerbach and Reishus (1988) to be a potential way forward.

A marriage model is a special case of matching models, and we take advantage of the recent developments in the econometrics of matching models (see the survey by Chiappori and Salanié, 2016) and consider Fox's (2018) Matching Maximum Score (MMS) estimator, which takes an observed market outcome as an equilibrium and recovers the parameters of a match value function. Several studies have employed this technique to examine corporate takeovers, but none have examined tax issues (Akkus, Cookson, and Hortacsu, 2015; Ozcan, 2015; Linde and Siebert, 2016; Kang, 2019; Chen et al., 2020).

This approach is similar but distinct from multinominal logit (MNL) models that take an observed outcome as a revealed preference over choices with different attributes and recover parameters of a utility/profit function.[4] From a practical perspective, a MMS estimate of a parameter represents the change in match value that arises from a unit increase in a matching characteristic, measured against the impacts of a benchmark variable. This contrasts with the application of limited dependent variable models: variables are assessed on the basis of how much they shift the probability of observing outcomes. Therefore, the MMS estimator provides an additional metric to evaluate the influence of variables.[5] Another practical benefit of this approach is that it provides a clear guideline for constructing counterfactual matches. Although these are important advantages, MMS estimation should be considered complementary rather than replacing existing methods for two main reasons. First, the number of variables in a model may need to be limited to ensure convergence of the estimation algorithm. Second, the method has not yet been integrated into the causal inference framework, rendering interpretation more delicate. Despite these potential shortcomings, MMS estimation has been successful in highlighting joint characteristics that

[4]Arulampalam, Devereux, and Liberini (2019) and Todtenhaupt et al. (2020) use MNL models to examine which locations parent companies choose to purchase targets from in a cross-country setting.

[5]Another complementary approach is to regress premiums on target losses to estimate the extent of capitalization (Henning, Shaw, and Stock, 2000; Chiang, Stammerjohan, and Englebrecht, 2014). The MMS estimator's advantage is that its implementation does not require data on sales prices. Akkus, Cookson, and Hortacsu (2015) consider an extension that incorporate transfers from acquirers to targets.

generate values and providing insights into the efficacy of industrial policies (Fox and Bajari, 2013; Akkus, Cookson, and Hortacsu, 2015).

Our data comprise a sample of takeover transactions concluded from January 1999 trough June 2018 between publicly-traded corporations listed on Japanese stock exchanges. We examine two possible channels of tax effects. First, LCFs are financial assets that can generate future cashflow through reduced tax payments, and as such, the valuation of LCFs depends on their prospects of making taxable income in the future. Where LCF valuations differ between buyers and sellers, potential for gains from reduced taxes arise in takeovers. Second, interest expenses on debts are tax deductible, but firms may underutilize 'interest tax shields' out of concern about potential bankruptcy. Takeovers can increase debt capacity utilization by reducing default risks, allowing the substitution of debt for equity in a firm's capital structure (Lewellen, 1971).[6]

The MMS estimation absorbs firm-level fixed effects, or more accurately, firm characteristics valued equally by all market participants. In our model, we control for pair-specific characteristics such as geographic proximity, market similarity, differences in Tobin's q and age. In a departure from previous studies that typically focus on transactions that fit a particular criterion, we explore different transaction types classified according to the fractions of outstanding shares owned by acquirers before and after the transactions. A wider range of transaction types allows us to explore the roles that a prior relationship can play by comparing cases where potential tax benefits are substantial with cases where potential tax benefits are remote, such as in toehold purchases where acquirers obtain small stakes in the target for the first time.

We detect tax effects of LCFs among some transactions types but not others. Our findings suggest that tax benefits sufficiently outweigh other countervailing forces only when acquirers have sufficient control over their targets before transacting. The degree of control that acquirers exercise over targets matters because this affects acquirers' ability to arrange the targets' structures in a way to enable carryover of losses.

Specifically, the estimated value of LCFs is positive as predicted when acquirers already own the majority of the targets and purchases the remaining stock, although the estimated magnitude is small. In contrast, for transactions in which acquirers initially hold no or minority stakes, the

[6]We have examined the incentive to combine businesses with negatively correlated returns to offset profits and losses contemporaneously, a type of merger incentive that is discussed in textbooks. We omit that analysis because of a complication induced by the reduction in sample size. Preliminary analyses show that the coefficients on correlation measures are generally positive, contrary to predicted tax effects.

estimated values of LCFs are negative and their magnitudes are relatively large. The negative estimates in these low-prior-control samples indicate a countervailing force: an aversion to invest in potentially unprofitable targets (Gallemore and Labro, 2015). Because the measure of tax implications of LCFs tends to be more pronounced when the parties include profitable and unprofitable firms, the variable captures these opposing forces. Overall, these results suggest that the value of tax considerations in the takeover market is dependent on factors such as whether the acquirers are well informed about the targets and have control over their operations.

In evaluating estimated magnitudes of LCF effects, we take geographical distance, which is known to affect takeover activities, as the baseline variable against which we measure the effects of other explanatory variables. Taking into account the aforementioned aversion to invest in unprofitable targets, our best estimate is that a one-standard-deviation increase in our measure of LCF incentives increases the match value by up to 0.54 times that of a 100 km increase in geographical distance for the sample of full acquisition of subsidiaries. This value is much smaller than that of another variable that is known to matter, market distance, which measures differences in the industry classifications of acquirers and targets (e.g. Andrade, Mitchell, and Stafford, 2001): the corresponding magnitude of that variable is 3.2.

Our results regarding debt capacity utilization are inconclusive since the effects of debt capacity utilization are imprecisely estimated. Although the estimate is positive and large for a sample in which acquirers already own the majority of targets and purchase the remaining stocks, we cannot rule out other explanations. In sum, we do not find evidence of debt utilization as a strong motive for takeovers.[7]

Our study primarily contributes to a body of studies that investigate the role of taxes in corporate takeover activities, particularly on factors that moderate the values of LCFs. This study is the first to highlight the role of information and control in affecting the value of LCFs to our knowledge, providing a specific case that speaks to the importance of internal coordination in tax avoidance emphasized by Gallemore and Labro (2015). The prevailing view in the U.S. appears to be that the tax rule that limits the utilization of tax losses in case of large ownership changes, defined in the Internal Revenue Code section 382, prevents takeovers motivated by loss trafficking, as emphasized earlier by Moore and Pruitt (1987). However, the precise efficacy of such specific

[7]Our study also adds to a burgeoning accounting literature on tax loss carryforwards (Langenmayr and Lester, 2018; Bethmann, Jacob, and Müller, 2018; Dobridge, 2021).

provisions has been questioned; for instance, Choi, Curtis, and Hayashi (2019) question its efficacy citing the lack of strong response over the four-month period during the Global Financial Crisis when the Internal Revenue Service relaxed the rule for the banking sector. Nonetheless, broader evidence suggests that anti-avoidance rules generally do deter tax-motivated takeovers. Utilizing EU-wide variations in 'anti-avoidance' rules, Bührle et al. (2023) and Hoehl (2021) show that tighter rules reduce the frequency of acquisition, corroborating the notion that the anti-avoidance rule deters loss trafficking. The trafficking restrictions in Japan are not as stringent as in the U.S. (Onji and Gordon, 2021). Nevertheless, our estimate of the value of LCFs is small. In conjunction with previous studies, our results suggest that tax rules can largely prevent windfall gains from LCFs.

Our analysis suggests that the anti-avoidance rules are largely successful in mitigating tax-motivated takeovers in the sample, but this result raises the question of whether the rules were too stringent for an economy that was stagnating. Bührle et al. (2023) emphasize a trade-off involved in anti-avoidance rules: stringent rules mitigate abuse and protect tax revenues but can impair economic growth by reducing risk taking and innovation. Whether we ought to use the tax system to encourage corporate takeovers in the aftermath of recessions needs to be carefully weighed because such measures exacerbate the complexity of the tax system and possibly disrupt markets. For example, the unexpected lifting of the loss limitation rule in the Wachovia Takeover Turmoil of 2008 seemingly increased the price offered to take over a troubled financial institution by more than sevenfold (Hanlon and Heitzman, 2010). If, however, a rapid reorganization of the corporate sector is a priority, temporarily relaxing rules on loss carryover may constitute a valid counter-cyclical policy.

This paper is organized as follows. Section 2 reviews the literature. Section 3 outlines the MMS estimator with an emphasis on its application to the corporate takeover market. Section 4 introduces the data. Section 5 discusses the empirical implementation. Section 6 presents the results. Section 7 offers discussions. Section 8 concludes.

2 Literature

2.1 Loss carryforwards

Evidence from early U.S. studies on targets' tax attributes is largely mixed (Shackelford and Shevlin, 2001; Betton, Eckbo, and Thorburn, 2008). In a seminal study, Auerbach and Reishus (1987b) estimate tax benefits in 318 mergers conducted from 1968 through 1983 and find that, in a significant minority of cases (20%), the tax benefits of Net Operating Losses (NOLs) appear large enough; their subsequent paper based on the marriage model, however, finds little effect of NOLs (Auerbach and Reishus, 1988). Plummer and Robinson (1990) consider transactions in which acquirers could benefit from NOLs before a tightening of the loss-carry rule in 1982 and find no difference in cumulative abnormal return (CAR) between targets with and without NOLs for 29 taxable acquisitions conducted from 1970 to 1982. Henning, Shaw, and Stock (2000) find that acquirers pay shareholder-level taxes in the form of higher acquisition premiums, but they detect no effect of NOLs in a sample of 1,071 transactions conducted from 1990 through 1994. Erickson (1998) examines the transaction structures (i.e. taxable or tax-free) of 340 transactions from 1985 through 1988 and find that, while acquirers' NOLs and tax rates increase the probability of cash transactions, targets' NOLs do not have any influence.

In contrast, Hayn (1989) examines 640 taxable and tax-free mergers from 1970 through 1985 and finds that targets' NOLs and tax credits are important in explaining CAR. Scholes and Wolfson (1990) consider the composite effects of tax changes under the Tax Reform Act of 1986 and argue that the reduction in takeover activities after the reform is attributable to the policy change.

Subsequent studies on tax loss carryforwards emphasize the role of the ability of acquiring companies to carryover losses upon transactions, as highlighted by Moore and Pruitt (1987). Devos, Kadapakkam, and Krishnamurthy (2009) observe that the loss limitation restriction would have limited the value of NOLs to acquirers, and they decide not to consider NOLs in their study. Erickson, Ton, and Wang (2019) note that the loss limitation rule limits the carryover of NOLs in targets but not in acquirers. They examine a sample of 1,986 acquisitions from 1987 through 2015 and find higher CARs for pairs in which acquirers have NOLs and targets are profitable. Onji and Gordon (2021) document an increase in a specific tax avoidance strategy to utilize LCFs following a Japanese tax reform in 2001 and report a detectable fall in book tax payments upon undertaking the strategy. Bührle (2021) predicts that the leniency of loss carryforward rules affects the incentives to found venture capital firms and uses variations across 28 E.U. nations in a difference-in-differences analysis to identify the tax effects using a sample of venture capital firms from 1999 through 2017. Bührle et al. (2023) and Hoehl (2021) utilize the same variation and find that the number of acquisitions involving targets with losses is systematically related to the loss carry rules across the E.U.

Overall, the lack of evidence reported in studies on the U.S. may in part be due to the success of the loss limitation rule, in conjunction with the legal doctrines restricting aggressive tax avoidance, in preventing tax-motivated mergers.

2.2 Debt capacity

In the trade-off theory of capital structure, firms weigh the tax deductibility of interest payments against expected bankruptcy costs (Kraus and Litzenberger, 1973). The theory implies that although some firms are too financially constrained to take full advantage of interest deductibility, mature firms with steady cash flow have excess capacity to take on additional debt. It has long been recognized that mergers can increase debt capacity by reducing default risks, allowing the substitution of debt for equity in a firm's capital structure (Lewellen, 1971). The aforementioned study by Auerbach and Reishus (1987a) finds that although the level of debt on average increases after takeovers, the debt-to-equity ratio remains roughly the same. Auerbach and Reishus (1988) use the marriage model and show that potential targets with higher debt-to-equity ratios are more likely to be taken over, contrary to the prediction that firms with unused debt capacity are targeted. In sum, these early studies do not find large effects of tax incentives on increased debt capacity.

Later studies offer suggestive evidence on the influence of debt capacity. The aforementioned paper by Erickson (1998) on the choice of acquisition structure finds that the probability of debt-financed taxable cash transactions increases with tax rates faced by acquirers. While this finding is not direct evidence for debt capacity being a motive for acquisition, the finding indicates debt tax shields are more valuable to high tax rate firms.

Devos, Kadapakkam, and Krishnamurthy (2009) consider a sample of 264 mergers from 1980 to 2004 in the U.S., and define the net present value (NPV) of the merger as the difference between the combined value of the merged firm and the sum of the standalone values of the acquirer and the target. To construct counterfactual, they employ the forecast of firm-specific cash flows published by Value Line to estimate the expected cash flows and investment expenditures of hypothetical non-merged firms. They decompose the NPV into operating and financial synergies. Operating synergies are defined as an increase in operating cash flows due to revenue enhancements or cost reductions, whereas financial synergies are defined as a decrease in financing costs due to interest tax shields or lower capital costs. They find that financial synergy is less than 17% of total synergy, concluding that tax benefits are relatively small.

Overall, while inconclusive, the empirical literature provides indirect evidence in support of the theoretical prediction regarding the debt capacity.

3 Methodology

A corporate takeover market is characterized by the jointness of the decision making and competition for matching partners. Target firms may accept or resist takeover offers from acquirers. A deal between two firms reduces the options available to other firms in the market. Given that the market has many features of matching games, there is a concern that standard regression techniques may be inadequate.

In this study, we consider Fox's (2018) MMS estimator, which is based on a theoretical matching market model. With the MMS estimator, analysts specify a functional form for the value generated from matching and estimate parameters of that function, taking observed outcomes as an equilibrium. The MMS estimator is semiparametric because the distribution of the error terms is left unspecified. The following exposition draws on the concise description provided by Chen and Song (2013).

3.1 Theoretical model

Following previous applications in the context of takeover markets (Akkus, Cookson, and Hortacsu, 2015; Ozcan, 2015; Linde and Siebert, 2016; Kang, 2019; Chen et al., 2020), we assume a market is two-sided, consisting of acquirers indexed by $i = 1, 2, \ldots, I_d$ and targets indexed by $j = 1, 2, \ldots, J_d$ where d indicates a particular market. A two-sided market presumes sellers/targets and buyers/acquirers as distinct, like men and women in a heterosexual marriage market. In practice, acquirers can become targets so that the market is better characterized as one-sided in the long run, but a Monte Carlo study by Akkus, Cookson, and Hortacsu (2015) shows that the estimator is useful and informative even when data are generated from a one-sided matching market. We also assume one-to-one matching for simplicity, but this is not strictly necessary.

We denote the matching between i and j as (i, j). In a given match, acquirer i realizes $V_a(i, j)$, which is a valuation function of firm i combining with firm j, in exchange for a payment u_{ij}. In addition to the synergy generated from combining tangible assets, $V_a(i, j)$ includes tax benefits to buyers, such as use of LCFs and any additional tax costs. Target j receives the payment u_{ij} and also realizes $V_b(i, j)$. In theory, target firms can remain unmatched, in which case they realize a value normalized to 0. However, unmatched firms are not used in our estimation. If a target incurs capital gains tax, $V_b(i, j)$ can be negative, and u_{ij} would need to be high enough to compensate for the capital gains tax liability. The total value generated from a match (i, j) is $V(i, j) = V_a(i, j) + V_b(i, j)$.

In equilibrium, matches (i, j) and (i', j') are pairwise stable. Two matches are pairwise stable if they satisfy the condition

$$V_a(i, j) - u_{ij} \geq V_a(i, j') - \tilde{u}_{ij'} \tag{1}$$

where $\tilde{u}_{ij'} = \Big(V_b(i', j') + u_{i'j'}\Big) - V_b(i, j')$. In inequality (1), $\tilde{u}_{ij'}$ is the minimum transfer that target j' is willing to accept if target j' were to switch its partner from i' to i. In other words, acquirer i prefers to buy j and pay u_{ij} rather than buying j' even if j' were to accept the minimum transfer $\tilde{u}_{ij'}$. In equilibrium, (1) holds for all $i, i' \in I$ and all $j, j' \in J$. In other words, acquirer i prefers to buy j rather than buying j' even when j' offers the best price it would be willing to offer.

The key step in connecting the theoretical model to the empirical estimator is to recognize that the equilibrium condition in inequality (1) implies

$$V(i, j) + V(i', j') \geq V(i, j') + V(i', j) \text{ for all } i, i' \in I \text{ and all } j, j' \in J. \tag{2}$$

This 'sum of match values inequality' states that swapping partners yields no greater sum of values in equilibrium. We relegate its derivation to the appendix.

3.2 Empirical model

In the implementation, we specify the value $V(i,j) = V_a(i,j) + V_b(i,j)$ generated from a match as

$$V(i,j) = X'_{ij}\beta + \eta_i + \delta_j + \epsilon_{ij}. \tag{3}$$

In equation (3), X_{ij} is a vector of joint characteristics that determines pair-specific match values. We consider a linear specification with a parameter vector β. The parameters η_i and δ_j are firm-specific characteristics that are valued equally by market participants. These parameters are not identified because they are eliminated when constructing the inequalities for estimation. The term ϵ_{ij} represents other match-specific characteristics. All terms are observable to market participants but only X_{ij} is observable to the econometrician.

A key assumption for interpreting the estimates in this structural framework is the rank order property condition

$$X'_{ij}\beta > X'_{ij'}\beta \text{ if and only if } \Pr[(i,j) \in \Omega] > \Pr[(i,j') \in \Omega] \tag{4}$$

where Ω is a set of all possible matches (Fox, 2010). The rank order property condition states that if the component involving X_{ij} of the value generated from a match (i,j) is greater than that of (i,j') then the probability of observing the formation of the match (i,j) is greater than that of the match (i,j'). Intuitively, for the variation in X_{ij} to inform us about the value of β, the stochastic component ϵ_{ij} should not be a large determinant of the outcomes.

Additional notation is necessary when describing the estimator. Assume the economy consists of D takeover markets indexed by $d = 1, 2, \ldots, D$ and that takeovers occur between firms within a given market and not across, so that markets are independent.

The MMS estimator maximizes the objective function

$$Q(\beta) = \sum_{d=1}^{D} \sum_{[(i,j),(i',j')] \in G^d} \mathbb{1}[X'_{ij}\beta + X'_{i'j'}\beta \geq X'_{ij'}\beta + X'_{i'j}\beta] \tag{5}$$

where G^d is a set of all combinations of pairs of observed matches $[(i,j),(i',j')]$ from market d. If market d contains 10 matches, for example, then G^d has 45(=10(10-1)/2) elements and the inner summation in (5) sums the 45 indicator functions. The inequality within the indicator function is called a matching maximum score inequality and is the empirical analogue of the sum of match values inequalities (2). In words, $\hat{\beta}$ is a parameter vector that maximizes the number of correctly predicted matching maximum score inequalities.

3.3 Estimator

While a conditional logit model is a convenient tool to analyze value functions such as (3), the logit approach assumes the error terms to be iid and follow the Type I Extreme Value distribution, leading to the Independence of Irrelevant Alternatives property that is known to impose restrictive substitution patterns. Advanced logit models, such as a mixed logit model, relax some aspects of this property, but not fully. Since the MMS estimator does not require the error term to follow any specific distribution, the approach is not susceptible to such misspecification errors (Akkus, Cookson, and Hortacsu, 2015; Ozcan, 2015). However, a trade-off for this flexibility is that the MMS estimator achieves set identification, so that the estimator narrows down the parameter values to a specific range, rather than determining a single point estimate (Fox, 2018). Another aspect of the MMS estimator is that it identifies the *relative* importance of covariates, requiring parameters to be interpreted in relation to a normalized variable, which can make evaluating the nominal magnitude of effects less straightforward. That said, the MMS estimator allows analysts to assess each variable's contributions to the value function directly, while with logit models, one assesses contributions indirectly through the probability of a match occurring. Thus, while care is needed in interpretation, the MMS estimator offers a robust semiparametric framework capable of modeling complex, two-sided matching processes without imposing restrictive assumptions on unobservable error distributions.

3.4 Optimization

The objective function (5) is not continuous, let alone differentiable, so standard gradient-based optimisation methods are not applicable. Instead, we use the iterative Differential Evolution (DE) algorithm (Price, Storn, and Lampinen, 2006), which is well-suited for this kind of problem. We start by defining a search space within the parameter space. Then, DE is initialised by randomly generating a set of candidate solutions (a 'population' of 'agents' exploring the space). In each iteration, the algorithm evaluates the performance of each agent using the objective function. Then, for each agent, a new 'trial' agent is created by combining the positions of current agents in the population, using a formula that includes some randomness. If this trial agent performs better, that is, if it leads to more correct predictions of MMS inequalities, it replaces the current one. This process repeats until the algorithm converges. Because the objective function has flat regions, the model is not point-identified. That is, there will be a range of parameter values that achieve the maximum.

The DE algorithm depends on several tuning parameters: the number of agents, the number of iterations generating new agents, the probability of a new trial agent replacing an existing agent, and the search space. To explore the sensitivity to the tuning parameters, we carry out a simulation analysis where the design is calibrated to our empirical setting in terms of the number of estimation parameters and inequalities.[8] In our simulation, we find that the optimization is robust to tuning parameters except for the size of the search space. The search space size matters little when the number of estimation parameters is two. For a model with five parameters to be estimated, choosing a relatively narrow search space produces confidence intervals with approximately correct coverage, indicating an advantage in limiting search space. A large search space has an advantage in guarding against missing the global optimum, but it may lead to estimates converging to a local optimum. Hence, there is an advantage in combining both small and large search spaces.

Based on the insight from the simulation, to estimate the parameters in the full sample, we apply several search spaces and see which parameter estimates perform well in terms of the number of correctly predicted MMS inequalities. To ensure accuracy, we then run the MMS estimator with a search space containing all these estimates. We elaborate on the procedure in the appendix. Briefly, we run the DE algorithm 500 times with different starting agents for each search space ranging from $[-50, 50]$ to $[-5, 5]$, common across all parameters. The range of $[-50, 50]$ is excessively wide by design. Since the impact of each variable is measured relative to the effects of the geographical distance between the headquarters, other variables are unlikely to exert 50-fold greater effects. We try wide search areas to ensure that the optimum does not lie outside. Based on the inspection of the distribution of the fit statistic, we select a final search area as reported in Appendix Table A1. Further details of the implementation can be found in the appendix.

[8]Briefly, the simulation generates joint characteristics of pairs for all combinations of firms in each market, forms pairs based on the Kuhn-Munkres (KM) algorithm, also known as the Hungarian algorithm. In our application, the KM algorithm assigns acquiring firms to target firms by identifying an optimal matching that maximizes the total value in the market. This approach provides a practical means of approximating the implication of the sum of match values inequality suggested by the theoretical model. The matched pairs are then used in the MMS estimation procedure to obtain parameters of the match value function.

3.5 Confidence intervals

The asymptotic sampling theory for the MMS estimator is based on there being a large number of markets. (We define markets in the next section.) To estimate confidence intervals (CI), we can therefore use a subsampling method. Specifically, this method works by drawing random samples of markets from the data and recomputing the estimates multiple times.

The procedure to estimate the main sample takes many hours on a desktop because we run estimations for many search areas to ensure we find the global optimum and not a local optimum. To cut down on the computation time, we apply the following procedure to estimate CI.

Our prior is that an optimum in a subsample is probably near the point estimates obtained from the full sample. Therefore, we initially set the search space around the point estimates. However, we are concerned about the possibility that the true optimum in fact lies outside the initial search space. To ensure that the space is not too restrictive, we recenter the search space when 'too many' estimates cluster near the boundary (the rule is 5% of subsamples). In addition, we try a narrower search space around the point estimates as we learn from the simulation that a narrower area helps locate the global maximum.

4 Data and empirical specification

4.1 Data sources and sample selection

The data source for M&A transactions is the MARR Database maintained by a consulting firm, RECOF Co. The MARR Database records M&A transactions involving prominent companies in Japan based on publicly available information such as news reports and corporate press releases. This database's primary users are businesses, serving as a ledger chronicling transactions ranging from initial equity participation, obtaining 100 percent of outstanding shares, and mergers. We use transactions from January 1999 through June 2018.

The first step in our sample construction is to select transactions in which both acquirers and targets are traded on the Japanese stock exchanges. We limit the sample to publicly-traded corporations because we match financial statement data from another database using stock exchange codes. We remove dissolved deals using entries on terminations of negotiations, which are listed separately from original deal announcements. We also remove transfers of divisions (e.g.

sales of a factory), reorganizations into a holding company structure, and additional equity purchases that do not result in an acquirer becoming a majority owner.

The database reflects transaction-level activity, resulting in multiple entries for firms involved in several deals. Consequently, the dataset is inherently restricted to firms participating in deals, omitting those that remained unmatched. While it is conceptually possible to augment the data involving unmatched firms, this is beyond the scope and requirements of the current analysis.

4.2 Tax variables

Loss carryforwards

To construct a measure of incentives to utilize LCFs, we draw from Auerbach and Reishus (1988) who estimate potential gains by comparing the present discounted values of LCFs from separate entities with those from a combined entity as follows:[9]

$$LCFGAIN_{ij} = \frac{\widehat{LCF}_{combined} - (\widehat{LCF}_i + \widehat{LCF}_j)}{(TA_i + TA_j)/2},$$

where subscripts i, j, and *combined* refer to the acquirer, target, and the combined firm, respectively, and where $\widehat{LCF}$ denotes the present discounted value of LCFs. To compute $\widehat{LCF}$, we assume the projected income in each future period is equal to the current net income. The projected income for the combined firm is the sum of incomes from both firms. We assume an annual discount rate of 6% and tax losses to expire after 5 years. The potential gains differential is weighted by the average total assets (TA) of two firms. We expect the associated coefficient in (5) to be positive.

Debt capacity utilization

In the trade-off theory of capital structure, firms weigh the tax deductibility of interest payments against expected bankruptcy costs (Kraus and Litzenberger, 1973). The theory implies that although some firms are too financially constrained to take full advantage of interest deductibility, mature firms with steady cash flow have excess capacity to take on additional debt. A potential benefit of a merger is better utilization of debt finance by reducing default risk (Lewellen, 1971). Auerbach and Reishus (1988) measure the tax benefits of debt finance using targets'

[9] We draw specifics from the working paper version of the paper (Auerbach and Reishus, 1987a, p. 30).

debt-to-equity ratios, presuming that targets with lower debt-to-equity ratios underutilize debt capacity. Their measure is target-specific and hence the effect is not identified within this framework. Therefore, we depart from the previous approach by considering whether or not the targets are financially constrained firms and acquirers are 'cash cows.'

Fazzari, Hubbard, and Petersen (1988) argue that financially constrained firms finance investments with internal cashflow since they face higher borrowing and equity finance costs. They and other subsequent studies show higher investment-cashflow sensitivities for groups of firms that are more likely to be financially constrained, such as young and non-dividend paying firms (Carreira and Silva, 2010). However Kaplan and Zingales (1997) and others debate the efficacy of using the investment-cash sensitivities to measure financial constraint. Almeida, Campello, and Weisbach (2004) argue that the sensitivities of cash to cashflow provide a better measure of financial constraints, reasoning that financially constrained firms are more likely to hold cash to smooth their cash flows and avoid future financing problems. Under this hypothesis, financially constrained firms should have a higher sensitivity of cash holding to cashflow than unconstrained firms.

We use the sensitivity of cash to cashflow as a measure of financial constraints. Following the standard approach in the literature (e.g., Almeida, Campello, and Weisbach, 2004), we regress changes in cash holdings on cashflow, controlling for Tobin's q and changes in sales. To obtain firm-specific estimates of cashflow sensitivity, we extend this basic model to a multilevel mixed-effects linear regression (Searle, Casella, and McCulloch, 1992) which allows for random firm-level heterogeneity in the cashflow coefficient.

Specifically, we estimate the sensitivity of cash to cashflow for firm j in period T using the auxiliary model

$$\Delta Cash_{js}/K_{js} = \alpha_{0t} + \alpha_{1t}Q_{js} + \alpha_{2t}\Delta S_{js} + (\alpha_{3t} + \gamma_{jt})CF_{js}/K_{js-1} + \zeta_{jt} + \psi_s + u_{js},$$
$$s = t-6, \ldots, t-1, \quad t = 1997, \ldots, 2018, \quad (6)$$

where Δ denotes a change from period $s-1$ to s. The dependent variable is the change in $Cash_{js}$, defined as cash holdings plus bank deposits excluding those maturing after one year, weighted by the book value of fixed assets, K_{js}. The control variable Q_{js} is Tobin's q, S_{js} is sales, and CF_{js} is cashflow. The term ψ_s represent a year dummy, and u_{js} is an idiosyncratic error term. The parameters α_{1t}, α_{2t} and α_{3t} represent fixed effects, while γ_{jt} is a random firm-specific coefficient and ζ_{jt} is a random firm-specific constant. The parameters depend on t since we estimate separately for each t. The model is estimated by the maximum likelihood method, assuming γ_{jt}, ζ_{jt} and u_{js} are independent of everything else and normally distributed. We drop firms with two or fewer observations. After estimation, best linear unbiased predictions (BLUPs) of the firm-specific coefficients, $\hat{\gamma}_{jt}$, are calculated from the residuals and the variance estimates. To ensure non-negativity, our measure of cash sensitivity CS_{jt} is defined as $\hat{\gamma}_{jt}$ with the minimum normalized to 0 (by subtracting the minimum value): $CS_{jt} = \hat{\gamma}_{jt} - \min_j[\hat{\gamma}_{js}]$. The model is estimated for all available firms.

Cash-cow firms typically operate in stable output markets and generate a steady stream of cash flows. As an indicator for cash-cow firms, we consider the ratio of cash flow to the total assets. Cash cows should be producing higher cash flows than comparable firms, so we use the ratio of cash flow weighted by total assets to that of the industry median as a marker, $(CF_{it}/TA_{jt-1})/(\overline{CF_{it}/TA_{jt-1}})$, where $\overline{CF_{it}/TA_{jt-1}}$ is the median asset-weighted cash flow of firms in the same industry as firm i.[10] To reduce the influence of extreme values and to ensure non-negativity, our cash cow measure CC_{it} is defined as the weighted cash flow winsorized at the 5th and 95th percentiles and the minimum normalized to 0 (by subtracting the minimum value).

As an indicator of the potential interest tax shield arising from higher debt capacity utilization, we construct the following measure:

$$fincon_{ijt} = CC_{it} CS_{jt} \Big(\mathbb{1}[age_{jt} \leq 10] + 1 \Big),$$

where subscript i refers to the acquirer and subscript j refers to the target. In the last term on the right-hand side, $\mathbb{1}[age_j \leq 10]$ is an indicator for target j being 10 years old or less. Intuitively, young firms tend to be more financially constrained than old established firms, so we incorporate corporate age as supplementary information.[11] [12]

[10] An indicator for any dividend payment is an alternative marker for cash cow. We are not incorporating dividend policy in the analysis because the measure would not provide useful information in the context of Japan. Generally, one-third of U.S. public companies pay dividends, but more than 80% of Japanese public companies do. An alternative control is the fraction of profits paid out as dividends. However, given a steady payout behavior, a higher payout rate does not necessarily indicate cash cows: firms with low current profits tend to maintain steady dividend payments. Additionally, unlike in the U.S. where aggregate repurchase often exceeds dividend payments in good years, the practice of repurchase is relatively new and dividend remains the main way to remunerate shareholders in Japan.

[11] The basic specification takes young firms as those that are at or less than 10 years old. Sensitivity analysis attempts different cut-off values. Note also that this specification does not utilize dividend payout policy which provides information on cash constraint in certain cases.

[12] In a preliminary examination, we tested dummy variable specifications; however, the parameter estimate for the

4.3 Control variables

In the q-theory of mergers, firms with high market-to-book ratios invest by buying under-deployed assets of another firm (Jovanovic and Rousseau, 2002). While this theory predicts 'high buys low', Rhodes-Kropf, Robinson, and Viswanathan (2005) argue that 'high buys less high' better describes data. They show that although targets' assets are, on average, valued below acquirers', targets' valuations are much higher than average firms. To allow for both possibilities parsimoniously, we consider the absolute value of the difference between the q of an acquirer and a target, as in Rhodes-Kropf and Robinson (2008) where they compare actual and counterfactual mergers using probit regression. If the q-theory describes our data well, then this variable should have a positive coefficient. If firms that match are alike, the sign should be negative.[13]

The earlier literature on mergers finds that the geographical proximity between potential pairs of firms is an important control (Akkus, Cookson, and Hortacsu, 2015; Bena and Li, 2014). We consider a measure of the distance between headquarters in 100 km.

Another common control is a measure of business similarity. To measure differences in product markets compactly, we convert pairs of industry classifications into a scale of 0 to 4 following Ozcan (2015). Based on 4-digits Japanese SIC, the product market distance is zero if both firms share a 4-digit SIC. If two firms share a 3-digit industry but belong to separate 4-digit SICs, the distance is 1, and so on. Finally, we consider the corporate age gap as the absolute value of the difference in log age, where age is defined as the number of years since a business was founded.

4.4 Transaction types

The tax implications of takeover transactions vary depending on the type of transaction. If arranged carefully, acquirers can benefit from target LCFs as soon as they complete tax-free mergers satisfying the conditions for loss transfer under Japanese institutions. In contrast, taxes are likely to have little impact on acquirers' decision to obtain small stakes in the target for the

measure failed to converge, likely due to an excessive number of zero observations. Accordingly, we operationalized the measure as a continuous variable, rather than representing it through a dummy variable.

[13]We have tried alternative specifications in our preliminary analysis. One is an interaction of q which should capture assortative matching. We subtract the industry-time median of q to account for differences across industries in application. Another is an indicator for both pairs having lower than the median q to capture parings between two weak firms. We report the specification using just the absolute spread since a parsimonious model allows us to limit the number of parameters to be estimated and the model with absolute spread has a better fit. We compute approximate q as the sum of market capitalization and total liabilities, minus total current assets, divided by total assets.

first time (toehold purchase). For toehold purchases, any tax effects depend on anticipated future gains.

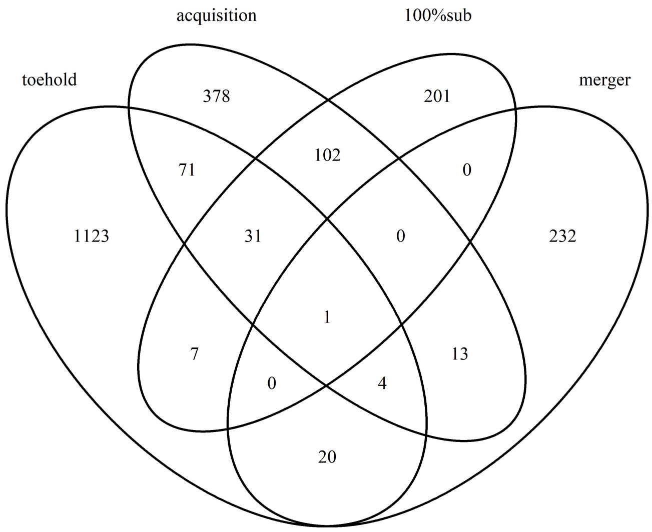
Our analysis uses the primary transaction categories in the MARR Database and implements the estimator for each transaction category separately with some modifications. The database's 'mergers' include transactions that eliminate corporate borders, in which two legally separate entities become one, as well as the establishment of a common holding company that owns two transacting parties. Other transactions are categorized based on prior and post-transaction ownership, and in some cases, the effective influence of acquirers. 'Toehold purchases' are stock acquisitions that result in minority ownership of targets in which acquirers have not held any stakes before. 'Acquisitions' are stock purchases by a minority corporate shareholder that result in majority ownership of a target, including cases in which the acquirer gains effective control even with less than 50% ownership. In implementation, we also examine acquisitions that result in 100% ownership and those that result in less than 100% ownership (but still majority). 'Full purchase of subsidiaries' is a stock purchase that results in 100% ownership of a subsidiary for which an acquirer either holds a majority stake or effectively controls the subsidiary. Appendix Table A2 provides a summary of transaction types.[14] Since the MARR Database chronicles transactions as they materialize, we observe multi-stage takeovers that progress in multiple stages, such as a toehold purchase followed up by a 100% acquisition.

Figure 1 provides a visual illustration of the overlap of firm pairs across transaction type samples. The leftmost oval, for example, represents a set of toehold purchases. The numbers shown in each intersected segment are the number of pairs that overlap with other transaction types. Of all the toehold purchases (N=1,257), 1,123 cases (89%) are not pursued further (as of June 2018, when we obtained the data). The intersection between the set of toehold purchases and acquisitions (the oval to the right) contains 107 cases (8.5%), indicating the fraction of targets in toehold purchases that are acquired subsequently. Similarly, intersections with other ovals show that 3.1% become wholly-owned subsidiaries and 2.0% are merged into acquirers. Of targets where acquirers purchased over 50% of the shares (N=600), the majority are not pursued further (74.8%); a large fraction (22.3%) become wholly-owned subsidiaries; and the fraction of follow-up mergers (3.0%) is larger than in the toehold sample. After becoming a wholly-owned subsidiary (N=342), follow-up mergers are rare and the sample contains only one case (0.3%). Finally, 85.9% of mergers (N=270)

[14] We are not able to utilize additional transaction information on stock transfer and exchange out of concern for the completeness of coverage.

involve parties with no prior stock involvement. For the overlapping transactions, Table A3 in the appendix shows the months elapsed between the transactions. The average time-to-event ranges from 28 to 76 months across categories. The longest average period is from toehold purchases to the full purchase of the subsidiary. Some transactions are followed rather quickly; one acquisition following a toehold purchase took only a third of a month.

Figure 1: Overlaps of firm pairs across the transaction samples



Notes: Each oval in this Venn diagram represents a type of M&A transaction. Numbers are counts of firm pairs, including observations with missing information on joint characteristics.

4.5 Markets

Mechanically, a 'market' plays two roles in implementing MMS estimation. First, the market determines the set of MMS inequalities in the objective function (5). In (5), the right-hand side of the inequality inside the indicator function is the sum of values from two counterfactual matches constructed by swapping partners in actual matches of the same transaction type. These partner swaps occur in a market, forming a set G^d of potential matches. For example, consider a market where electric equipment manufacturers are bought and sold. Counterfactual matches from that market consist of an acquirer in the market paired with an electric equipment target bought by another acquirer. The counterfactual matches do not involve a buyer of a construction firm (unless that firm also bought an electric equipment manufacturer in the same year).

Second, markets are the basis for estimating confidence intervals. The MMS estimator's limiting distribution is complex, and a formula for the standard errors has not been derived. To obtain a confidence interval, a subsampling procedure repeats the estimation for many subsets of markets drawn from all markets in the sample as described in section 3.5.

To form a market, we adopt a procedure set out in prior studies on mergers (Bena and Li, 2014; Ozcan, 2015). We first divide corporate pairs into 'buckets,' defined over the targets' industry at 4-digits and April–March financial-year. If a bucket contains five (actual) matches or more, that bucket is considered a single market. We combine buckets with fewer than five matches with other buckets in the same 3-digit category and see if the combined bucket contains five or more matches. We repeat this procedure and drop matches if the final 1-digit bucket contains fewer than five pairs.

5 Results

5.1 Control variables

Table 1 presents MMS estimates multiplied by a one-standard-deviation change in the respective variables. We relegate coefficient estimates, which are harder to interpret, to Appendix Table A5. The five columns represent different transaction types: 1 is acquisition, 2 and 3 are a subset of acquisition that results in less than 100% ownership and 100% ownership, respectively, 4 is full purchase of subsidiaries, and 5 is toehold purchase. The coefficients are normalized relative to the effects of a base variable, namely, the distance between headquarters in 100 km. Similar to previous studies, our data reveal the negative effects of geographical distance, and we normalize its coefficient to -1.

To illustrate the interpretation of the MMS estimates, we consider the coefficient for market distance from column 1, Table 1. This coefficient, pertaining to a variable commonly found to be highly influential in the literature, is -6.6 with a confidence interval of $[-13.1, -2.4]$. This indicates a substantial negative effect, corroborating previous studies (e.g. Andrade, Mitchell, and Stafford, 2001).

To put this magnitude into perspective, first, consider that a pair of firms sharing the same 3-digit but not 4-digit industry classification exhibits a unit difference in market distance. One standard deviation of market distance in this sample is approximately 1.5 (Appendix Table A4). Interpreted structurally, a one-standard-deviation increase in the market distance reduces the match value by as much as 6.6 times the change in value arising from a 100 km increase in geographical distance. Second, for a more tangible comparison, Tokyo and Osaka are approximately 400 kilometers apart in straight-line distance. A one-standard-deviation change in market distance is roughly equivalent to $1.65 (= 6.6/4)$ times the Tokyo-Osaka displacement effects. Overall, this estimated magnitude suggests a substantial negative effect of market distance. This effect is consistent across the transaction samples, with a somewhat larger effect for column 3, which represents a combined sample of mergers pooled with acquisitions that result in 100 percent post-transaction ownership.

Table 1: Matching maximum score estimates from various takeover transactions: Standardized

	1	2	3	4	5
variables	all acquisitions	acquisitions < 100%	acquisitions = 100%	100% subsidiaries	toehold purchase
market distance	−6.66	−4.50	−13.80	−3.02	−4.90
	[−13.187, −2.486]	[−10.695, −1.650]	[−31.647, −4.950]	[−6.776, −1.065]	[−12.460, −4.074]
q distance	−1.67	−4.07	−8.74	−3.36	−1.62
	[−4.152, −0.472]	[−9.072, −1.230]	[−23.084, −2.901]	[−6.840, −0.322]	[−4.307, −0.215]
lcfgain	0.11	0.09	−0.07	0.04	−0.57
	[−0.006, 0.289]	[0.006, 0.223]	[−0.158, −0.005]	[−0.003, 0.103]	[−1.341, −0.018]
fincon	5.65	−3.83	−0.34	2.59	−2.22
	[−0.714, 11.822]	[−9.726, 0.255]	[−0.636, −0.030]	[0.309, 5.952]	[−5.700, 0.060]
agegap	−0.56	−2.58	−6.57	−0.23	−2.85
	[−1.563, −0.113]	[−7.344, −0.544]	[−18.300, −2.192]	[−0.650, 0.040]	[−6.935, −0.898]
no. ineq	626	275	307	263	2100
no. markets.	31	17	19	16	105
subsample size	9	5	6	5	32
pct. correct	0.84	0.85	0.89	0.89	0.72

Notes: This table shows estimates of values generated from one-standard-deviation increase in each variable relative to the value change arising from 100 km increase in geographical distance. We multiply coefficients and confidence intervals reported in Appendix Table A5 by the standard deviations of each variable from the matched sample reported in Table A4. The scaling of standard deviations of $LCFGAIN_{it}$ in the summary statistics table are adjusted back in this computation.

The coefficient on q distance is expected to be negative under the hypothesis of assortative matching. The coefficients are consistently negative across the samples, indicating ‘like buys like’ in the merger emphasized by Rhodes-Kropf and Robinson (2008). The estimated magnitudes of these effects are heterogeneous: small in the toehold sample (−1.6) and large in 100% acquisition (−8.7) in terms of a one-standard-deviation increase in the q distance. The confidence intervals lie in the negative region.[15]

The coefficients on differences in corporate age are also consistently negative across the samples. Confidence intervals are relatively tightly estimated, with all but one lying in the negative region.

[15] We included the absolute difference in q and a dummy indicating whether both firms had below-median industry q; however, their inclusion did not qualitatively affect the results.

This result could reflect the tendency of younger corporations not taking over older counterparts with large assets. Additionally, if age gaps capture differences in corporate cultures, the estimate reflects the large costs associated with combining two workforces with different ways of running a business.

5.2 Tax variables

For the acquisition sample, the coefficient on $LCFGAIN_{it}$ is positive, and its confidence interval is mostly in the positive range. To assess this coefficient's magnitude, consider the standard deviation of $LCFGAIN_{it}$ (0.023, shown in Appendix Table A4). A one-standard-deviation increase in $LCFGAIN_{it}$ increases the match value by 11% of the 100 kilometer displacement effects. Therefore, compared to the influence of market distance, this magnitude is modest.

We find that the LCF effects in the acquisition sample are heterogeneous, varying with post-transaction ownership level. Column 2 reports results for acquisitions that result in partial ownership, defined as post-transaction ownership between 50% and less than 100%.[16] To illustrate the magnitude of the effect, Figure 2 compares the impacts of a one-standard-deviation increase in $LCFGAIN$ across samples. The estimate from column 2 shows a narrow confidence interval around 9% of the geographical displacement effect, entirely within the positive region.

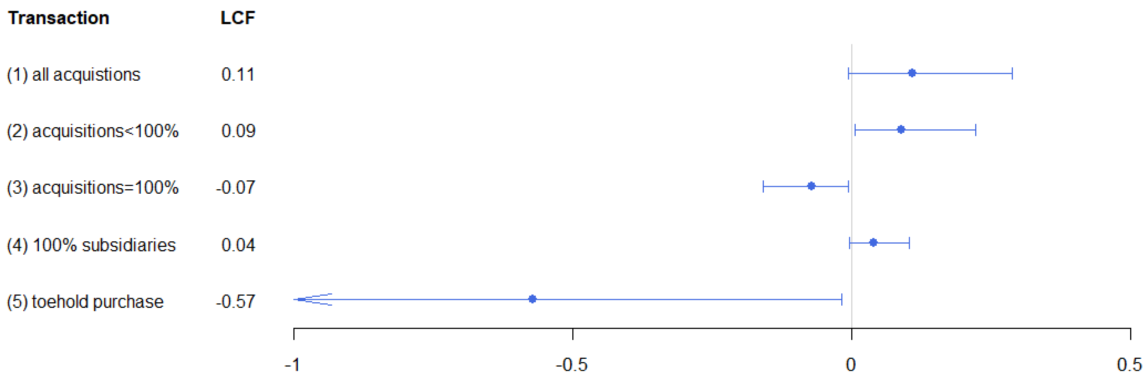
Column 3 presents results for acquisitions that result in full ownership, combined with mergers. Ideally, mergers and acquisitions would be analyzed separately; however, due to limited sample size, we combined them to ensure estimable confidence intervals. In contrast to the previous two columns, the coefficient on $LCFGAIN$ is negative, with the confidence interval entirely in the negative region. The effect size is modest, amounting to approximately -7% of the geographical displacement effect.

The aforementioned result contrasts with the estimate from a sample of transactions that bring majority-owned subsidiaries to full ownership, as presented in column 4. The estimated effect is 4% of the displacement effect with the confidence interval mostly in the positive region. The last column presents estimates from a sample of toehold purchases for which there are no immediate tax benefits accruing from LCFs. The point estimate is negative and its magnitude is large in

[16] The number of inequalities in this sample is less than half that of the previous one–not because most acquisitions result in full ownership, as the majority do not reach 100% ownership. Rather, the sample size is smaller because we exclude markets with fewer than five transactions from the estimation, as discussed in section 4.5. Market definitions vary depending on the estimation sample.

comparison: −57% of the geographical displacement effect. This large estimate is partly due to the large estimated coefficient, but the large standard deviation of *LCFGAIN* (5.7) is the main contributing factor.

Figure 2: Value increase from a one-standard-deviation increase in *LCFGAIN*



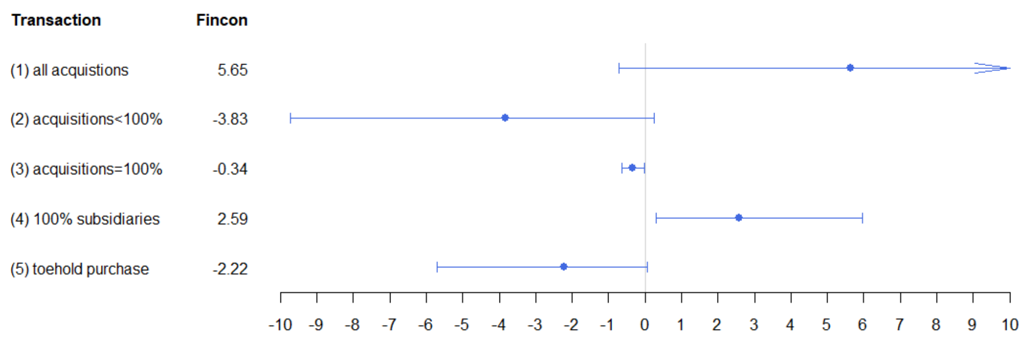
Notes: This figure is shows estimated values generated from a one-standard-deviation increase in *LCFGAIN* relative to the change in value arising from a 100 km increase in geographical distance from Table 1.

Generally, the coefficients on debt capacity are imprecisely estimated, as illustrated in Figure 3, which has a wider range. In particular, the confidence intervals for acquisition samples 1 and 2 almost span the base search range, which indicate irrelevance of this variable in these samples. However, the coefficient for acquisitions that bring ownership to 100% is the most precise, and its confidence interval is in the negative range. The coefficient for the 100% purchase of subsidiaries is relatively precise, with its confidence interval being in the positive range. The confidence interval for the toehold sample lies almost entirely in the negative region. Overall, since variable construction requires multiple steps, the noise in the measurement likely contributes to the wide estimate, making it difficult to discern signals from noise. This said, the debt capacity variable appears to add no value to samples 1 and 2, has a negative influence in 3 and 5, and adds a positive value in 4.

6 Discussion

In our data, the MMS estimator can detect key economic forces underlying takeover activities: the estimates on non-tax variables are generally precise, and their signs agree with previous studies; the estimated model predicts the observed outcome over 80% of the time. Therefore, these estimates provide a basis for considering taxes' contribution in the value function.

Figure 3: Value increase from a one-standard-deviation increase in debt capacity



Notes: This figure shows estimated values generated from a one-standard-deviation increase in the debt capacity measure (*FINCON*) relative to the value change arising from 100 km increase in geographical distance from Table 1.

6.1 Loss carry forward

Interestingly, estimated effects of LCF values are heterogeneous across different types of transactions. Although the estimated values are positive for 100% acquisition of subsidiaries and acquisitions that result in less than 100% ownership, LCFs reduce match values in other transactions. To rationalize this finding, the conflicting effects of LCFs, as pointed out by Amir and Sougiannis (1999), seem relevant: LCFs are financial assets that can generate cashflow, but their presence may signal a firm's likelihood of future losses. Firms with LCFs are not necessarily in financial distress, as in the cases of prospective young firms with startup costs, but some are. The estimated model allows for influences from associating with underperforming firms: the value function includes the q distance, which captures a match between prospective and underperforming firms. However, the coefficients on $LCFGAIN$ appear to capture residual negative values from associations with underperformers. The negative estimate from the sample of toehold purchases that does not derive immediate tax benefits is a case in point.

The signs on $LCFGAIN$ are opposite for the two transaction types with the same 100% post-transaction ownership but with different pre-transaction ownership cf. columns 3 and 4, Table 1. These contrasting results suggest the importance of information and control in deriving benefits from LCFs. In acquisitions that involve 100% ownership (column 3), acquirers are minority shareholders with limited rights. Minority shareholders have the right to inspect books but they lack the power to nominate and elect corporate directors on their own. Consequently, while acquirers with minority stakes possess more insights into targets than non-shareholders, they likely lack direct access to executive information. Limited information can cause concern about hidden information that reduces the expected value of merging firms with LCFs. In contrast, acquirers in the column 4 sample are majority shareholders who have the power to nominate and elect directors on their own. A minority shareholder can also elect directors, but must form a coalition with other shareholders. Having two-thirds ownership further allows shareholders to decide unanimously on restructuring matters, such as mergers, acquisitions, and sales of divisions. These powers are important because they allow acquirers to make necessary arrangements to qualify for tax-favorable treatment. For instance, under Japanese law, even when a transaction qualifies for a tax-free merger, additional requirements apply to carryover target losses in a new entity.[17] Therefore, the positive value generated by LCFs in the sample of 100% acquisitions of subsidiaries is likely due to the acquirers' greater information and control over their targets, which enables them to be better informed and pre-arrange target structures in a way to meet the letters of the law.

If the coefficient on *LCFGAIN* also captures aversion to loss-making targets, what are the net effects of tax incentives? The model does not allow us to quantify the relative contributions of these three forces exactly, but we can estimate a bound with additional assumptions. Suppose that the contribution of geographical distance is identical across different transaction types. Suppose likewise for the negative signal conveyed by LCFs regarding the future profitability of targets. Since the coefficient on *LCFGAIN* in the toehold sample would have neither tax nor rescue motives, the effects of a negative signal generated from one unit increase in *LCFGAIN* are equivalent to a loss of 9.9 times the 100 km displacement effects (Appendix Table A5, column 5). Applying this estimate to the result for all acquisitions (Appendix Table A5, column 1), the combined effects of tax and rescue motives are 0.34 of a 100 km displacement effect from a one-standard-deviation increase in *LCFGAIN*.[18] Given the gross estimate of 0.11, the net effect of taxes is in the range $[0.11, 0.34]$ under the assumptions. In other words, the value of tax effects from the sample of all acquisitions is at most 0.34 of the displacement effect. Because this upper bound represents 5.1% of the effects of market distance, the value of LCFs is modest.[19]

[17] In a prominent tax lawsuit, an acquirer appointed its president as a vice president of a target firm to fulfill one of these requirements (Hironaka, Kitamura, and Noda, 2015).

[18] $(4.87 + 9.9) \times 2.34/100$. 2.34 is the standard deviation of *LCFGAIN* (rescaled) from Appendix Table A4, column 1.

[19] The observation of modest LCF effect is robust to relaxing the assumption on the equivalent effects of geographical distance across transactions. If anything, the upper bound would be smaller since geographical distance plays a weaker role among the sample of toehold purchases, as suggested by the smaller gap in means of geographical distance between matched and counterfactural samples presented in table A4. We also estimate a linear probability model pooling a toehold sample and full purchases of subsidiaries and confirm the smaller geographical distance's ability to predict matches among the toehold sample.

6.2 Debt capacity utilization

The results indicate positive values generated from debt capacity utilization among transactions that fully buy up the remaining shares of subsidiaries (Table 1, column 4). Surprisingly, the effect magnitude for debt capacity is much larger than that for LCFs: a one-standard-deviation increase in $LCFGAIN$ and financial constraints is equivalent to 4% and 259% of the 100 km displacement effects, respectively. This is surprising given the economic environment in Japan during the sample period. Since 2000, average contract interest rates on loans and discounts from domestic banks have been below 2%. Low interest rates reduce the value of the debt shield; thus, the tax incentive for utilizing loans is lower than in a high interest rate environment.

One way to rationalize this result is to consider the following: the benefits of debt capacity arise under specific conditions regarding the relationship between financially constrained firms and cash cow firms. Possibly, becoming a 100% subsidiary of a cash-rich firm boosts the target's borrowing capacity because Japanese parent companies often provide guarantees on loans formally or informally. However, this explanation requires a discontinuous or at least a rapid increase in debt capacity as ownership in the target approaches 100%. Moreover, this explanation does not conform with the negative coefficient obtained in the sample which includes acquisitions of previously minority-owned targets that result in 100% ownership (Table 1, column 3). Instead, the evidence points to an alternative mechanism at work, distinct from the debt capacity hypothesis. In a common practice to reconstruct subsidiaries in financial distress, parent companies buy up remaining shares of subsidiaries to make them wholly owned.[20] By gaining full control, parent companies can restructure their targets without interference from other shareholders and reap full benefits in the event of successful recovery. In addition, assisting a wholly-owned subsidiary has a tax advantage. Generally, intra-group loans are tax inefficient because a subsidized portion of the interest payment is included in the taxable income of the recipient but is deducted only up to a limit from the recipient's income. However, since 2010, subsidized loans do not trigger tax liabilities if a corporate lender has held 100% of the borrower's stock. In addition, since 2002, intra-group loans are tax-free when a group files consolidated tax returns available to a parent and its wholly-owned subsidiaries. In sum, the debt capacity hypothesis does not explain our data as well. Instead, the variable likely captures the combined effects of business and tax benefits of

[20] For example, the retail giant Aeon increased its ownership of a supermarket chain in distress, Daiei, from 44% to 100% in January 2015.

restructuring wholly-owned subsidiaries.

6.3 Alternative interpretation of the estimate

Under the structural interpretation, the estimates are our best guess of parameters in a match value function. Recall that estimates are obtained by maximizing the number of correctly predicted MMS inequalities with the sum of values generated in two actual matches on the left side and the sum of values generated from counterfactual matches constructed by swapping partners on the right side. An alternative interpretation is that the estimates are weights indicating the importance of each variable in predicting match outcomes. Under this reduced-form interpretation, the estimates on LCFs indicate a small but perceptible influence of LCFs in predicting match outcomes in certain types of transactions. Therefore, the results provide a summary of data that help assess the usefulness of variables in prediction, and not just estimates of structural parameters.

The structural model implies that match characteristics affect match values, presuming a causal relationship. However, if some characteristics are correlated with other factors not included in the model, the estimates capture correlational relationships that are not necessarily causal. As discussed earlier, $LCFGAIN$ capture firms' aversion to association with partners in potential distress, in addition to tax incentives. Additionally, some parent firms state that the objective of taking over their partially-owned subsidiaries in financial distress is to facilitate restructuring by gaining full rein over target operations. The estimated effects on $LCFGAIN$ in the sample of full purchase of subsidiaries would capture such motives, which are intertwined with tax motives. For instance, while the forgiveness of loans is treated in principle as a donation that is an expense for tax purposes only up to a certain limit, loan forgiveness for the purpose of rescuing subsidiaries is fully tax deductible under certain conditions. Consequently, our $LCFGAIN$ estimates, particularly in the context of full subsidiary purchases, likely reflect a complex interplay of tax and non-tax motives.

7 Conclusion

This study examines tax motives in corporate takeovers and estimates their influences. Our sample comprises takeovers among publicly-traded corporations in Japan from January 1999 through June 2018. Our estimates suggest that the gross values of LCFs in takeovers are heterogeneous by type of transaction, depending on how well the acquirers are informed about the targets and how much control they have over their operations. The estimated magnitude of the net value of LCFs is relatively small even for a favorable transaction type, indicating the efficacy of the anti-avoidance rule in deterring tax-motivated transactions among large corporations. The debt capacity hypothesis does not explain our data as well. Instead, our measure of debt capacity utilization likely captures the combined effects of business and tax benefits of restructuring wholly-owned subsidiaries.

Beyond our substantive findings, our application of the MMS estimator illuminates several important considerations for future research. One of the key conceptual appeals of the MMS estimator is that it allows analysts to account for the two-sided nature of corporate takeover decision-making. In application, we find that the determination of search spaces, a key tuning parameter in estimation, must be determined systematically, albeit at the cost of lengthening computation time. Despite the relative complexity in implementation, the method offers an advantage over MNL models: the framework allows us to directly assess the monetary value of certain characteristics about a matching pair. We find that another variable, besides a scaling variable, that is known to be influential is highly helpful in providing a benchmark for assessment. This feature has an advantage over MNL models, where we must make assessments indirectly through the influence of variables on the probability of observing certain outcomes. Thus, while careful implementation is required, the MMS estimator provides a powerful and intuitive framework for directly valuing match characteristics in two-sided settings, offering a valuable complement to existing methods.

Our study has some limitations. First, the sample is based on transactions between publicly-traded companies in Japan. Tax rules on takeovers are diverse across the world, potentially generating stronger tax effects elsewhere. Takeovers involving smaller companies in Japan are outside the scope of our analysis, and tax effects may be more pronounced among smaller companies that face less scrutiny from the tax authority. We leave the assessment of the generalizability of the findings to future studies. The second limitation stems from the fundamental choice between matching and search models in analyzing markets for takeovers. Our sample includes pairings of publicly-traded firms that face disclosure requirements so that a matching model that assumes perfect information has some traction, but not all information relevant to takeovers is public. Explicitly accounting for search frictions is beyond the scope of this study, and future studies that consider estimators based on search models are of interest.

References

Akkus, Oktay, J Anthony Cookson, and Ali Hortacsu (2015). "The determinants of bank mergers: A revealed preference analysis". *Management Science* 62.8, pp. 2241–2258.

Almeida, Heitor, Murillo Campello, and Michael S Weisbach (2004). "The cash flow sensitivity of cash". *The Journal of Finance* 59.4, pp. 1777–1804.

Amir, Eli and Theodore Sougiannis (1999). "Analysts' interpretation and investors' valuation of tax carryforwards". *Contemporary Accounting Research* 16.1, pp. 1–33.

Andrade, Gregor, Mark Mitchell, and Erik Stafford (2001). "New evidence and perspectives on mergers". *Journal of Economic Perspectives* 15.2, p. 103.

Arulampalam, Wiji, Michael P Devereux, and Federica Liberini (2019). "Taxes and the location of targets". *Journal of Public Economics* 176, pp. 161–178.

Auerbach, Alan J and David Reishus (1987a). "The effects of taxation on the merger decision". NBER Working Paper 2192.

— (1987b). "The impact of taxation on mergers and acquisitions". *Mergers and Acquisitions*. Ed. by Alan J Auerbach. Chicago: University of Chicago Press, pp. 69–86.

— (1988). "The effects of taxation on the merger decision". *Corporate Takeovers: Causes and Consequences*. Ed. by Alan J Auerbach. Chicago: University of Chicago Press, pp. 157–190.

Bena, Jan and Kai Li (2014). "Corporate innovations and mergers and acquisitions". *The Journal of Finance* 69.5, pp. 1923–1960.

Bethmann, Inga, Martin Jacob, and Maximilian A Müller (2018). "Tax loss carrybacks: Investment stimulus versus misallocation". *The Accounting Review* 93.4, pp. 101–125.

Betton, Sandra, B. Espen Eckbo, and Karin S. Thorburn (2008). "Corporate Takeovers". *Handbook of Empirical Corporate Finance*. Ed. by B. Espen Eckbo. Handbooks in Finance. San Diego: Elsevier, pp. 291–429.

Blouin, Jennifer L, Eliezer M Fich, Edward M Rice, and Anh L Tran (2021). "Corporate tax cuts, merger activity, and shareholder wealth". *Journal of Accounting and Economics* 71.1, p. 101315.

Bottomlee, Thomas W., Jason S. Bazar, and Arthur C. Walker (2009). "Don't ignore a target's NOLs: The price and structure of your deal can depend on them". *M&A Journal* 9.7.

Bührle, Anna Theresa (2021). “Do tax loss restrictions distort venture capital funding of start-ups?”

Bührle, Anna Theresa, Elisa Casi-Eberhard, Barbara Stage, and Johannes Voget (2023). “The value of a loss: the impact of restricting tax loss transfers”. TRR 266 Accounting for Transparency Working Paper Series 128.

Carreira, Carlos and Filipe Silva (2010). “No deep pockets: Some stylized empirical results on firms’ financial constraints”. *Journal of Economic Surveys* 24.4, pp. 731–753.

Chen, Guoli, Sterling Huang, Philipp Meyer-Doyle, and Denisa Mindruta (2020). “Generalist vs. Specialist CEOs and Acquisitions: Two-sided Matching and the Impact of CEO Characteristics on Firm Outcomes”. *Strategic Management Journal*, pp. 1–32.

Chen, Jiawei and Kejun Song (2013). “Two-sided matching in the loan market”. *International Journal of Industrial Organization* 31.2, pp. 145–152.

Chiang, Wei-Chih, William Stammerjohan, and Ted D Englebrecht (2014). “Pricing target NOLs in mergers and acquisitions from the participating firms’ perspective”. *Advances in Accounting* 30.1, pp. 32–42.

Chiappori, Pierre-André and Bernard Salanié (2016). “The econometrics of matching models”. *Journal of Economic Literature* 54.3, pp. 832–61.

Choi, Albert H, Quinn Curtis, and Andrew T Hayashi (2019). “Crisis-driven tax law: The case of section 382”. *Florida Tax Review* 23, pp. 1–72.

Devos, Erik, Palani-Rajan Kadapakkam, and Srinivasan Krishnamurthy (2009). “How do mergers create value? A comparison of taxes, market power, and efficiency improvements as explanations for synergies”. *Review of Financial Studies* 22.3, pp. 1179–1211.

Dobridge, Christine L (2021). “Tax loss carrybacks as firm fiscal stimulus: A tale of two recessions”. *The Accounting Review* 96.6, pp. 153–181.

Erickson, Merle (1998). “The effect of taxes on the structure of corporate acquisitions”. *Journal of Accounting Research* 36.2, pp. 279–298.

Erickson, Merle, Karen Ton, and Shiing-wu Wang (2019). “The effect of acquirer net operating losses on acquisition premiums and acquirer abnormal returns”. *The Journal of the American Taxation Association* 41.2, pp. 103–124.

Fazzari, Steven M, R Glenn Hubbard, and Bruce C Petersen (1988). “Financing constraints and corporate investment”. *Brookings Papers on Economic Activity* 19.1, pp. 141–206.

Fox, Jeremy T (2010). “Identification in matching games”. *Quantitative Economics* 1.2, pp. 203–254.

— (2018). “Estimating matching games with transfers”. *Quantitative Economics* 9.1, pp. 1–38.

Fox, Jeremy T and Patrick Bajari (2013). “Measuring the efficiency of an FCC spectrum auction”. *American Economic Journal: Microeconomics* 5.1, pp. 100–146.

Gallemore, John and Eva Labro (2015). “The importance of the internal information environment for tax avoidance”. *Journal of Accounting and Economics* 60.1, pp. 149–167.

Gordon, Roger H and James R Hines (2002). “International taxation”. *Handbook of Public Economics*. Ed. by Alan J Auerbach and Martin Feldstein. Vol. 4. Elsevier, pp. 1935–1995.

Hanlon, Michelle and Shane Heitzman (2010). “A review of tax research”. *Journal of Accounting and Economics* 50.2, pp. 127–178.

Hayn, Carla (1989). “Tax attributes as determinants of shareholder gains in corporate acquisitions”. *Journal of Financial Economics* 23.1, pp. 121–153.

Head, Keith and John Ries (2008). “FDI as an outcome of the market for corporate control: Theory and evidence”. *Journal of International Economics* 74.1, pp. 2–20.

Henning, Steven L, Wayne H Shaw, and Toby Stock (2000). “The effect of taxes on acquisition price and transaction structure”. *Journal of the American Taxation Association* 22.s-1, pp. 1–17.

Hironaka, Akihiro, Michito Kitamura, and Masaki Noda (2015). “Japan”. *Tax Disputes and Litigation Review—The Third Edition*. Ed. by Simon Whitehead. London: Law Business Research Ltd., pp. 227–239.

Hoehl, Rebecca (2021). “Anti-tax loss trafficking rules and the acquisition of loss-carrying firms”. Unpublished manuscript available at SSRN 3971019.

Jovanovic, Boyan and Peter L Rousseau (2002). “The Q-theory of mergers”. *American Economic Review* 92.2, pp. 198–204.

Kang, Gu Sang (2019). “The effects of technological similarity and diversity on merger and innovation”. KIEP Working Paper 19-08.

Kaplan, Steven N and Luigi Zingales (1997). “Do investment-cash flow sensitivities provide useful measures of financing constraints?” *The Quarterly Journal of Economics* 112.1, pp. 169–215.

Kraus, Alan and Robert H Litzenberger (1973). “A state-preference model of optimal financial leverage”. *The Journal of Finance* 28.4, pp. 911–922.

Langenmayr, Dominika and Rebecca Lester (2018). “Taxation and corporate risk-taking”. *The Accounting Review* 93.3, pp. 237–266.

Lewellen, Wilbur G (1971). “A pure financial rationale for the conglomerate merger”. *The Journal of Finance* 26.2, pp. 521–537.

Linde, Sebastian and Ralph Siebert (2016). "Do Mergers Among Multimarket Firms Create Value?" CESifo Working Paper 6139.

Moore, Norman H and Stephen W Pruitt (1987). "The market pricing of net operating loss carryforwards: Implications of the tax motivations of mergers". *Journal of Financial Research* 10.2, pp. 153–160.

Ohrn, Eric and Nathan Seegert (2019). "The impact of investor-level taxation on mergers and acquisitions". *Journal of Public Economics* 177, p. 104038.

Onji, Kazuki and Roger H Gordon (2021). "Taxes, corporate takeovers, and step transactions". Osaka University Discussion Papers in Economics and Business 21–16.

Ozcan, Yasin (2015). "Innovation and acquisition: two-sided matching in M&A markets". mimeo.

Plummer, Elizabeth and John R Robinson (1990). "Capital market evidence of windfalls from the acquisition of tax carryovers". *National Tax Journal* 43.4, pp. 481–489.

Politis, Dimitris N, Joseph P Romano, and Michael Wolf (1999). *Subsampling*. Springer Science & Business Media.

Price, Kenneth, Rainer M Storn, and Jouni A Lampinen (2006). *Differential evolution: a practical approach to global optimization*. Springer Science & Business Media.

Rhodes-Kropf, Matthew and David T Robinson (2008). "The market for mergers and the boundaries of the firm". *The Journal of Finance* 63.3, pp. 1169–1211.

Rhodes-Kropf, Matthew, David T Robinson, and Sean Viswanathan (2005). "Valuation waves and merger activity: The empirical evidence". *Journal of Financial Economics* 77.3, pp. 561–603.

Scholes, Myron S and Mark A Wolfson (1990). "The effects of changes in tax laws on corporate reorganization activity". *Journal of Business* 63.1.

Searle, Shayle R, George Casella, and Charles E McCulloch (1992). *Variance components*. NY: John Wiley & Sons.

Shackelford, Douglas A and Terry Shevlin (2001). "Empirical tax research in accounting". *Journal of Accounting and Economics* 31.1, pp. 321–387.

Todtenhaupt, Maximilian, Johannes Voget, Lars P Feld, Martin Ruf, and Ulrich Schreiber (2020). "Taxing away M&A: Capital gains taxation and acquisition activity". *European Economic Review* 128, p. 103505.

Vanistendael, Frans (1998). "Taxation of corporate reorganizations". *Tax Law Design and Drafting*. Ed. by Victor Thuronyi. Vol. 2. Washington, D.C.: International Monetary Fund, pp. 895–924.

Appendix

Derivation of the sum of match value inequality (2)

The following obtains by substituting $\tilde{u}_{i,j'}$ into (1):

$$V_a(i,j) - u_{i,j} \geq V_a(i,j') - V_b(i',j') - u_{i',j'} + V_b(i,j'). \tag{7}$$

Likewise, a similar inequality holds for (i', j'):

$$V_a(i',j') - u_{i',j'} \geq V_a(i',j) - V_b(i,j) - u_{i,j} + V_b(i',j). \tag{8}$$

Summing these inequalities and rearranging gives,

$$[V_a(i,j) + V_b(i,j)] + [V_a(i',j') + V_b(i',j')] \geq [V_a(i,j') + V_b(i,j')] + [V_a(i',j) + V_b(i',j)]. \tag{9}$$

The sum of match value inequality (2) follows directly.

Search areas used in full sample

To determine the search area, we run preliminary estimations with five different search area magnitudes (± 1, ± 5, ± 10, ± 20, ± 50), running the estimator 500 times using different starting values. From among 2,500 sets of estimates, we filter for runs exhibiting high fit to obtain a range of estimates. If the initial range of estimates is entirely negative or positive, we adjust the search area (e.g., ± 1) to ensure it spans zero. However, if this range is already very far from zero, this adjustment is not made. We report the final search area in Appendix Table A1

Estimation parameters

The objective function is numerically maximized using differential evolution implemented using R's DEoptim package. The number of parameter vectors (NP) is 100. The crossover probability (CR) is 0.6. The weighting factor is 0.8. The maximum number of iterations (itermax) is 500 with an allowance for exit after 200 consecutive iterations that did not improve the value of the objective function. The optimization strategy is the local-to-best (DE/local-to-best/1/bin). In the full sample, the optimization searches for the optimum in a range narrowed down from preliminary runs with wider search areas. In Appendix Table A5, the reported point estimates are the best-found

Table A1: Search areas for transaction samples

	1		2		3		4		5	
variables	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub
market distance	−5	1	−5	1	−15	−5	−5	1	−10	1
q distance	−5	1	−6	1	−15	−5	−6	1	−5	1
lcfgain	−1	15	−1	6	−10	2	−1	8	−10	1
fincon	−1	15	−10	1	−1	1	−1	8	−10	1
agegap	−1	1	−5	1	−10	1	−1	1	−10	1

Notes: This table shows search areas applied in the DE algorithm for each variable to obtain the estimate reported in Appendix Table A5. 'lb' and 'ub' refer to lower and upper bounds, respectively. Columns 1 to 5 represent transaction samples: 1. all acquisitions, 2. acquisitions with less than full post-transaction ownership, 3. mergers and acquisitions with 100% post-transaction ownership, 4. 100% acquisitions of majority-owned targets, and 5. toehold purchases.

maxima from 500 different starting population members. In the event of ties, a single set is randomly selected. The brackets contain 95% asymmetric confidence intervals computed using empirical distributions from subsampling with 1,000 iterations.

Subsampling procedure

Our prior is that an optimum in a subsample is probably near the point estimates obtained from the full sample. We therefore construct an initial search area centered around the point estimate with bounds plus and minus 110% of the absolute value of the point estimates. We also try bounds that are 150% of the point estimates. For each base-starting-area, we run the optimization algorithm 50 times.

In certain cases, the initial search area may not contain the optimal value for the subsample. This is signaled by multiple estimates converging to the search area boundaries. In these cases, we try another search area for that parameter, with the center anchored at the boundary value. In implementation, 'near the edge' is defined as 5% of the absolute value of the point estimate from a boundary, and the 'significant number' is defined as over 5% of estimates from the 50 runs. The span for the updated search area is 200% of the absolute value of the point estimate. The search areas for other parameters remain unchanged. For subsamples that meet the criteria, we run the optimization algorithm 20 times.

We also restrict the base search area for two reasons: to improve convergence and to refine the search. First, estimates from some subsamples are widely dispersed and reach both upper and lower bounds. Narrowing the search area facilitates convergence. Second, even if an estimate is well within a search area, further narrowing the search area allows the algorithm to locate a point estimate that achieves a higher model fit. For each subsample, we select a subset of the estimates that attain the highest fit for the subsample and obtain their ranges. We narrow the search area to the range of the subset with added 'buffers'. The lower bound is the minimum minus 50% of the average of absolute values of the maximum and minimum. For all subsamples, we run the algorithm 20 times applying the restricted search area.

In summary, we run estimation 120 times using three search areas (or 140 times using four search areas in case many estimates reach the boundary in the base search area) for each subsample. We then select an estimate that maximizes the number of correctly predicted inequalities.

The choice of block size

One practical issue in implementing subsampling is the choice of block size, or the fraction of the sample to be used in each iteration (Politis, Romano, and Wolf, 1999). While a broad range of block sizes is theoretically correct, confidence intervals are known to be sensitive to the choice of block size.

In a preliminary examination, we try a procedure that automatically selects a block size, referred to as the minimum volatility method by Politis, Romano, and Wolf (1999, pp. 199–200). Following the procedure in this method, we estimate confidence intervals for different block sizes, ranging from 4–20% of the number of markets in the base sample. We then compute rolling standard deviations (σ) of upper-bound and lower-bound for each parameters, using a window of ± 2 block sizes. To reduce a vector of parameters to one dimension, we compute the magnitudes $[\sum_{w=1}^{W}(\hat{\sigma_w})^2]^{\frac{1}{2}}$, where w denotes w-th parameter and W is the total number of parameters. The minimum volatility method selects the block size with the smallest magnitude. We find that standard deviations shrink quickly as block sizes increase. Since the automatic selection procedure prolongs estimation time without affecting sensitivity, we fix block size to speed up computation in our subsequent estimation. Block sizes are nearest integer numbers to 30% of the number of markets in the respective transactions.

Table A2: Description of transaction types

	Transaction sample	Description
1	Acquisition	A stock purchase by a corporate acquirer that results in a majority ownership of a target. Includes cases where effective control is gained with less than 50% ownership. Excludes cases where an acquirer initially holds more than 50% ownership.
2	Acquisition(<100%)	A subset of (1) where the acquirer's final direct ownership of less than 100%.
3	Acquisition(=100%) and Merger	A combined sample consisting of (6) and a subset of (1) with final direct ownership of 100%.
4	Full purchase of subsidiaries	A stock purchase resulting in 100% ownership of a subsidiary where the acquirer held a majority stake or effective control.
5	Toehold purchase	A stock purchase resulting in minority ownership of a target where the acquirer held no prior stake.
6	Merger	Combination of two or more entities into a single new entity. Includes integration via a holding company structure.

Notes: This table presents an abbreviated description of the primary transaction classification from the MARR Database. 'Effective control' refers to *de facto* control gained through methods such as indirect ownership or by appointing the acquirer's board members or employees to the target's corporate board.

Table A3: Months elapsed in multi-stage takeovers

From	To	N	Mean	SD	Min	Max
toehold	acquisition	107	32.20	34.30	0.30	191.90
toehold	100%sub	39	76.90	48.30	3.80	194.00
toehold	merger	25	33.40	31.70	3.10	100.80
acquisition	100%sub	134	42.40	36.00	1.20	209.20
acquisition	merger	18	28.10	22.60	2.00	88.50

Notes: This table shows time elapsed between transactions between same parties. The unit is month. Computed for a subset of transactions with overlaps across the transaction sample. This table omits one case of a merger following a full purchase of subsidiary; the elapsed time in that case is 2.3 month.

Table A4: Summary statistics on match characteristics

		1		2		3		4		5	
variable	stat.	real	cf	real	cf	real	cf	real	cf	real	cf
geo_dist	mean	1.94	2.77	1.81	2.60	1.52	2.75	1.41	2.20	1.43	1.94
	sd	2.39	2.78	2.30	2.56	2.20	2.72	2.22	2.49	2.13	2.43
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	max	8.87	14.17	8.87	13.68	9.59	14.19	9.46	10.58	13.48	14.19
	N	216	1422	104	589	120	657	100	537	727	4729
mkt_dist	mean	2.37	3.04	2.59	3.13	2.36	2.99	2.46	3.11	2.78	3.10
	sd	1.48	0.89	1.50	0.92	1.37	0.81	1.40	0.68	1.40	0.99
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	max	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
	N	216	1422	104	589	121	665	100	537	729	4763
q_dist $\times 10$	mean	7.03	7.81	7.41	8.07	5.22	6.18	5.85	6.48	9.91	10.13
	sd	8.91	9.25	9.46	9.08	6.73	6.54	6.66	6.90	13.46	13.01
	min	0.00	0.00	0.15	0.03	0.00	0.00	0.05	0.00	0.00	0.00
	max	61.34	67.68	52.33	60.42	41.85	47.66	36.82	44.49	82.97	82.97
	N	216	1422	104	589	119	648	100	537	727	4751
lcfgain $\times 100$	mean	0.08	0.14	0.35	0.17	−0.03	0.07	0.21	0.18	0.03	0.29
	sd	2.34	1.77	2.75	1.52	2.55	0.85	0.70	0.77	5.73	3.21
	min	−14.51	−34.33	−13.14	−16.50	−17.56	−6.78	−1.50	−3.03	−95.68	−64.47
	max	19.47	30.67	19.47	12.83	14.11	10.34	4.95	6.74	47.74	116.59
	N	216	1422	104	589	120	664	100	537	729	4763
fincon	mean	1.15	1.14	1.12	1.11	1.19	1.17	1.12	1.13	1.29	1.26
	sd	0.51	0.45	0.51	0.48	0.53	0.44	0.63	0.60	0.60	0.58
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	max	4.03	4.22	4.03	3.89	4.82	4.82	3.95	3.98	4.08	5.54
	N	203	1322	99	552	116	630	98	524	663	4381
agegap	mean	0.63	0.75	0.60	0.74	0.49	0.57	0.65	0.71	0.70	0.78
	sd	0.81	0.81	0.68	0.75	0.75	0.73	0.73	0.72	0.73	0.76
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	max	4.49	4.44	3.76	3.93	4.47	4.45	4.50	4.13	4.20	4.56
	N	216	1422	104	589	121	665	100	537	728	4757

Notes: This table shows summary statistics on firm-pair characteristics for actual matches and for counterfactual transactions ("real" and "cf"). Columns 1 to 5 represent transaction samples: 1. all acquisitions, 2. acquisitions with less than 100% post-transaction ownership, 3. acquisitions with 100% post-transaction ownership combined with mergers, 4. complete purchases of majority-owned targets, 5. toehold purchases. q-dist and lcfgain is multiplied by 10 and 100 respectively to improve visibility.

Table A5: Matching maximum score estimates from various takeover transactions

	1	2	3	4	5
variables	all acquisitions	acquisitions < 100%	acquisitions = 100%	100% subsidiaries	toehold purchase
market distance	-4.50	-3.0	-10.07	-2.16	-3.5
	$[-8.91, -1.68]$	$[-7.13, -1.1]$	$[-23.1, -3.613]$	$[-4.84, -0.761]$	$[-8.9, -2.91]$
q distance	-1.87	-4.3	-12.99	-5.04	-1.2
	$[-4.66, -0.53]$	$[-9.59, -1.3]$	$[-34.3, -4.310]$	$[-10.27, -0.483]$	$[-3.2, -0.16]$
lcfgain	4.87	3.4	-2.83	5.90	-9.9
	$[-0.25, 12.35]$	$[0.22, 8.1]$	$[-6.2, -0.177]$	$[-0.43, 14.72]$	$[-23.4, -0.32]$
fincon	11.07	-7.5	-0.65	4.11	-3.7
	$[-1.40, 23.18]$	$[-19.07, 0.5]$	$[-1.2, -0.056]$	$[0.49, 9.447]$	$[-9.5, 0.10]$
agegap	-0.69	-3.8	-8.76	-0.31	-3.9
	$[-1.93, -0.14]$	$[-10.80, -0.8]$	$[-24.4, -2.923]$	$[-0.89, 0.055]$	$[-9.5, -1.23]$
no. ineq	626	275	307	263	2100
no. markets.	31	17	19	16	105
subsample size	9	5	6	5	32
pct. correct	0.84	0.85	0.89	0.89	0.72

Notes: Columns 1 to 5 represent transaction samples: 1. all acquisitions, 2. acquisitions with less than 100% post-transaction ownership, 3. acquisitions with 100% post-transaction ownership combined with mergers, 4. complete purchases of majority-owned targets, 5. toehold purchases. The coefficient on geographical distance, which is omitted from the table, is normalized to -1.