Dynamic Effects of Dual-Class Shares: Theory and Evidence*

HYUNSEOB KIM† *Chicago Fed*

DORON LEVIT‡ *University of Washington*

RONI MICHAELY§ *University of Hong Kong*

February 9, 2026

Abstract

This paper analyzes the dynamic trade-off of dual-class shares. We argue that while these structures encourage firm-specific investment by mitigating a holdup problem, the value-added is eroded by private benefits rising with maturity. Using new U.S. data on dual-class firms from 1971–2022, we document five findings consistent with this model. Difference-in-differences estimates show that after dual-class recapitalizations, valuations rise initially but decline with firm maturity, while innovation increases persistently. These effects are concentrated in industries where investments are more likely to be firm-specific. Stock prices react positively to recapitalizations for young firms only. Investment by mature dual-class firms is less sensitive to q, and the voting premium increases with maturity. Taken together, our results support the dynamic treatment effects and yield new policy implications.

Keywords: Dual-class shares; dynamic trade-off; private benefits of control; sunset provisions; treatment vs. selection

JEL Classification: G34, D86, K22

*This paper subsumes an earlier version of the study, entitled "Sticking around Too Long? Dynamics of the Benefits of Dual-Class Voting." We appreciate comments from Felipe Aldunate, Lucian Bebchuk, Alon Brav, Camilo Garcia-Jimeno, Eitan Goldman, Todd Gormley, John Graham, Harald Hau, Rob Jackson, Wei Jiang, Andrew Karolyi, Dietmar Leisen, Tao Li, Chen Lin, Kate Litvak, Michelle Lowry, Alberto Manconi, Ernst Maug, Justin Murfin, Alexandra Niessen-Ruenzi, Thomas Philippon, Jean-Charles Rochet, Stefano Rossi, Benjamin Segal, Oliver Spalt, Margarita Tsoutsoura, Hannes Wagner, Fei Xie, Yishay Yafeh, Luigi Zingales, and seminar and conference participants from Ackerman Conference on Corporate Governance, Alberta, Bocconi, Bristol, Chicago Entrepreneurship Workshop, Chicago Fed (Finance and Micro), CELS (Emory), Cornell, Drexel Corporate Governance Conference, EFA (Bratislava), Erasmus Corporate Governance Conference, FinanceUC Conference (PUC de Chile), Goethe, Hebrew University of Jerusalem, Inquire UK, KAIST-Korea University, Mannheim, Massachusetts–Amherst, Melbourne, MFA, Seoul National University, SMU, Sun Yet Sen University Finance Conference, Tel Aviv University Finance Conference, Texas A&M, Universidad de los Andes Chile, University of Geneva GFRI, UIBE, and US Securities and Exchange Commission, financial support from the Cornell Institute for the Social Sciences, and excellent research assistance from Ikchan An, Roger Chen, Jee-Hun Choi, Sophie Croome, Daniel Gallego, Max Gillet, Syungjin Han, Hassan Ilyas, Kenny Jing, Julia Kharzeev, Dawoon Kim, Eric Kim, Seungsoo Kim, Yejun Kim, Boyao Li, Cindy Lu, Amani Moin, Pradeep Muthukrishnan, Ahmad Sabbagh, Brian Wickman, and Zakary Yudhishthu. All errors are our own.

†Economic Research Department, Chicago Fed. hyunseob.kim@chi.frb.org.

‡Foster School of Business, University of Washington. dlevit@uw.edu.

§Business School, University of Hong Kong. ronim@hku.hk.

1 Introduction

A rising share of newly public firms has adopted dual-class shares in the U.S. over the past several decades (Figure 1). Dual-class firms represented over 9% of U.S. stock market value in the recent decade—a more than fourfold increase from just over 2% in the early 1980s. Consequently, it has become increasingly important to understand how the allocation of voting rights—the choice between dual-class and single-class shares in particular—affects economic efficiency. When Meta's market value fell in 2022, for example, investors blamed its dual-class shares for the poorer performance among peers.[1] The SEC's then-Commissioner Robert Jackson had urged U.S. stock exchanges to consider limiting use of dual-class shares.[2] If dual-class shares lead to poor performance through increased agency costs, as is often argued, why are they widely used by young firms? What are the impacts of policies that limit their use? In this paper, we construct a new panel dataset on U.S. public firms' voting and ownership structures that spans a half-century to study these questions.

We develop a simple model to organize our empirical analysis. We argue that dual-class shares encourage the founder's firm-specific investment ("effort"), but the value-added from the investment is eroded by private benefits that rise with firm maturity—creating a dynamic trade-off. Crucially, we show private benefits are an important source of incentives for the founder's effort. Because dual-class shares protect the founder from losing these private benefits *ex post*, he has stronger incentives to exert firm-specific effort *ex ante*. In contrast, single-class shares limit the founder's ability to extract private benefits, which in turn weakens incentives—because shareholders cannot commit not to replace him once the effort is sunk. This holdup problem (e.g., Grossman and Hart (1986); Hart and Moore (1990)) is mitigated by dual-class shares. In essence, dual-class structures allocate control to the founder when his investment is particularly important for the firm's success. The model predicts a transitory dual-class premium: young dual-class firms have higher value than comparable single-class firms, ceteris paribus, but this initial value gap shrinks and ultimately reverses as firms mature.

We test this and other model predictions by using two empirical approaches. First, we analyze the dynamic effects of changes in governance structures on firm valuation and firm-specific investment using a stacked difference-in-differences event-study framework (e.g., Gormley and Matsa (2016); Cengiz et al. (2019)): (i) dual-class recapitalizations, where single-class firms adopt a dual-class structure, and (ii) dual-class stock unifications, where dual-class

[1] See, e.g., "Facebook and the Conglomerate Curse," *The Economist*, November 5, 2022. From the beginning through November of 2022, Meta's market value dropped by 65%, while the Nasdaq index declined by 28%.

[2] See https://www.sec.gov/news/speech/perpetual-dual-class-stock-case-against-corporate-royalty.

firms transition to a single-class structure. To this end, we carefully construct an event-study dataset that spans the 1971 to 2022 period drawing from public disclosure, news sources, *Moody's Manuals*, and CRSP–Compustat (see Section 2.1 for details). Recapitalizations and unifications are common: among about 1,100 unique firms that ever have dual-class shares, 226 and 347 firms have recapitalized and unified, respectively. We use Tobin's q as our measure of firm valuation (e.g., Cronqvist and Nilsson (2003); Gompers, Ishii, and Metrick (2010)) and patent outputs to proxy for the founder's firm-specific investment.[3]

We find that in the years leading up to a recapitalization or unification, neither Tobin's q nor patent output exhibits differential pre-trends. In the year firms recapitalize into a dual-class structure, Tobin's q rises by 0.35 relative to the control group of firms remaining single-class, and this valuation premium remains largely significant in the next five years. The premium then declines, and by ten years post-recapitalization, the valuation gap essentially disappears. Also consistent with our model's prediction, a recapitalization leads to a persistent increase in patent output—our proxy for founder effort. We find corresponding results for stock unifications—Tobin's q *declines* initially before reverting to its pre-unification level, while patent output experiences a persistent decline.

Second, we examine how the market's reaction to announcements of dual-class recapitalizations and stock unifications varies with firm maturity. We find that when single-class firms younger than 12 years ("young firms")—the sample median age for dual-class firms—announce a recapitalization, the average cumulative abnormal return (CAR) is 3.5%.[4] This positive market reaction early in the life cycle is consistent with the benefit of stronger founder incentives outweighing the cost of private benefits extraction for young dual-class firms. However, when single-class firms aged 12 years or older ("mature firms") announce a recapitalization, the average CAR is 4.6 percentage points lower. This finding suggests that the market views dual-class recapitalizations as increasingly value-reducing as firms mature, plausibly driven by increasing private benefits extraction.[5] Analogously, we find that the average CAR for the announcement

[3]This measure assumes that the founder's investment, as reflected in patent outputs, is not easily redeployable outside the firm, leading to a holdup problem. The firm-specificity of innovation likely stems from two sources: (i) the founder's attachment to the firm he has built, and perhaps more crucially, (ii) the economic costs of redeploying the product of his investment elsewhere. For example, moving a platform—such as one for social networking or ride sharing—whose value depends on network effects among both providers and users would be highly costly.

[4]To streamline the exposition, we refer to firms with ages less than the median as "young" and greater than or equal to the median as "mature" throughout the paper. Our empirical analysis employs alternative measures of firm maturity including a set of age indicators.

[5]Table 6 provides evidence for this key model premise—that private benefits of control increase with firm maturity. We measure private benefits using the voting premium (the price premium of the share class with superior voting rights) controlling for proxies of takeover-contest probability (e.g., Zingales (1995); Nenova (2003)). See Levit, Malenko, and Maug (2024) for a theoretical analysis showing that the voting premium may also reflect

of a dual-class stock unification is 5.0 percentage points higher for mature firms compared to young firms.

Crucially, our model provides a framework to assess whether the observed transitory dual-class premium reflects a treatment effect or selection. The model shows that if selection were driving the dynamics, both private benefits extraction and founder effort would be greater in *single-class* firms—the opposite to the prediction under a treatment effect. Intuitively, the founder opts for a single-class structure as a commitment device to behave when his expected private benefits are large (otherwise the share price and the value of his ownership stake would be too low).[6] Furthermore, under the selection hypothesis, if the negative effect of large private benefits initially outweighs the positive effect of strong incentives on single-class firms, firms where the founder selects a dual-class structure can exhibit higher initial valuations. As the incentive effect grows over time, this valuation gap shrinks, mimicking the transitory premium.

The set of evidence we find supports a treatment effect of dual-class shares rather than selection. First, the patent-output results described above indicate that founders of dual-class firms make greater firm-specific investment in innovation than their single-class counterparts. Second, to measure private benefits of control, we examine firms' investment decisions. The idea is that founders (and managers) have stronger preferences for preserving what they have invested in than shareholders do (e.g., Bertrand and Mullainathan (2003)). Consistent with private benefits being more pronounced in mature dual-class firms than in single-class firms, we find that investment by mature dual-class firms is less sensitive to Tobin's q—our measure of investment opportunities—than that of their single-class counterparts, particularly when firms are near the disinvestment threshold. Third, we find that the effects of dual-class recapitalization and unification on Tobin's q and innovation are concentrated in industries where investments are more likely to be firm-specific—consistent with the core mechanism of the model.

Given that private benefits tend to grow with maturity in dual-class firms, founders may have little incentive to unify dual-class shares as their firms mature. Consequently, some academics and policymakers argue that dual-class shares should include a sunset provision—under which the passage of time or the transfer of the founder's ownership triggers unification (e.g., Bebchuk and Kastiel (2017); Jackson (2018)). We analyze the impact of such provisions by comparing unifying dual-class firms with and without sunsets. We find that unification has no discernible effect on valuation for firms with sunsets, whereas it has the negative

heterogeneous preferences or beliefs among shareholders, and capture the marginal rather than the average benefit from additional voting rights.

[6] More precisely, the necessary condition is that private benefits are expected to grow faster for firms that choose single-class shares. See Proposition 3 for the detailed mechanism.

effect described above for firms without a sunset. A plausible interpretation is that sunset provisions weaken founders' *ex ante* incentives by reducing expected private benefits—which consequently mitigates the negative valuation effect of unification. Our analysis suggests that policymakers should also weigh the potential costs of sunset provisions—specifically, weaker incentives for firm-specific investments—which have been largely overlooked in the current debate on dual-class structures and their sunsetting.

This paper contributes to the literature on the allocation of voting rights. Empirical research has primarily examined the average cross-sectional relationship between voting rights, private benefits, and firm valuation (e.g., Claessens et al. (2002); Masulis, Wang, and Xie (2009); Gompers et al. (2010)).[7] On the theoretical side, DeAngelo and DeAngelo (1985) argue that an important advantage of dual-class shares is that they encourage managers to undertake firm-specific investment by mitigating the holdup problem (e.g., Williamson (1975)). However, to our knowledge, this incentive effect of dual-class shares has not been formalized in a model, nor has its relation to private benefits and shareholder value been characterized.

Our paper differs from this existing work in two key and related ways. First, we develop a theoretical model that integrates the incentive and private benefits effects of dual-class shares over the firm's life cycle, generating a rich set of *dynamic* predictions regarding founder effort, private benefits, and firm value. Second, we provide the long-run evidence consistent with these dynamic predictions by estimating the effects of dual-class recapitalization and unification using a difference-in-differences (DD) event-study framework. This analysis is enabled by our new panel and event datasets spanning 1971–2022. These datasets on dual-class firms will be available on the authors' websites.

A recent paper by Cremers, Lauterbach, and Pajuste (2024) documents evidence consistent with a transitory dual-class premium for IPO firms from 1980–2019 using pooled OLS. The main empirical difference is that we estimate the effects of dual-class shares using *changes* in governance structures in a DD framework with explicit control groups, whereas they examine the time-series evolution of firm valuations since IPO. Our DD design, including evidence on parallel trends, and our model jointly allow for a treatment effect interpretation of our estimates. In addition, we provide new evidence on the dynamics of founder effort (patents) and private benefits (voting premium and differential investment sensitivities), and announcement returns for dual-class recapitalizations and unifications—each of which helps interpret our

[7] An early exception is Cronqvist and Nilsson (2003), who estimate the within-firm effect of controlling shareholders' voting rights on valuation for Swedish firms using firm fixed effects. See, e.g., Adams and Ferreira (2008) for a review of this literature.

empirical findings through the lens of the model. Finally, we offer new policy implications that are different from those of existing work, including Cremers et al. (2024).

More broadly, this paper contributes to the growing literature on the dynamic effects of corporate governance, including work on boards of directors (e.g., Field, Lowry, and Mkrtchyan (2013)); Field and Lowry (2022)) and takeover defenses (e.g., Johnson, Karpoff, and Yi (2022)). Dual-class shares are arguably among the most effective mechanisms for consolidating control rights (e.g., Ruback (1988); Gompers et al. (2010)). Thus, dual-class structures provide a particularly powerful context in which to study the dynamic consequences of governance arrangements for firms and their controllers.

2 Data Construction and Descriptive Evidence

2.1 Constructing panel of dual-class firms including governance changes

We construct our panel dataset of dual-class firms in the United States from 1971 to 2022. We hand-collect information on dual-class firms drawing from several sources including *Moody's Manuals*, the SEC's EDGAR, and news articles, as well as merging to CRSP–Compustat data. For the years 1994–2002, we build upon data from Gompers et al. (2010)[8] but collect further information necessary for our key empirical analysis such as difference-in-differences event studies. The resulting dataset represents (i) the longest panel on U.S. dual-class firms covering more than a thousand unique dual-class firms over a half-century and (ii) the most extensive dataset on both dual-class recapitalization and unification events. We present several highlights of the dataset crucial for the empirical analysis below. Data Appendix G provides details of the data construction.

First, key variables in our dataset include voting rights, such as votes per share and rights to elect directors, and the number of shares outstanding by share class. Importantly, it provides comprehensive history of companies that we identify have ever been dual-class firms by collecting information on the firm's first dual-class year through either IPO or recapitalization, as well as the firm's last dual-class year through either delisting or unification.

Second, and related, for our difference-in-differences analyses, we carefully verify whether and when a dual-class firm has unified its share classes, or a single-class firm has recapitalized with dual-class shares. In the years that the SEC's EDGAR database is available (1994 to 2022), we collect this information from 10-K filings, and in earlier years we use *Moody's Manuals*, which list all equity classes of a firm and usually discuss recapitalization or unification, if it

[8]We thank Andrew Metrick for making the dataset on dual-class firms available on his website.

happened. We also collect the exact dates of these switches when available, as reported in the 10-K or Moody's, and verify ambiguous cases by searching for online company histories and news articles that document the governance changes. Finally, we cross-reference these switches by using the firm's shares outstanding by class and delisting codes from CRSP.

Third, to study short-term stock market reactions to dual-class recapitalizations, we collect their first public announcement dates from by combining information from Partch (1987) and Jarrell and Poulsen (1988) (covering 1971–84 and 1976–87, respectively) and our own news collection using Factiva and LexisNexis. If we find news articles on recapitalizations that precede those reported in the other papers, we use the earlier date as the event date. (The difference in event dates is typically within one to two days, although it can be up to 59 days.) Importantly, we exclude events that are confounded by announcements of other major corporate events, such as dividend declarations, mergers and acquisitions (M&As), and emergence from bankruptcy. For unification announcements, on which we are not aware of existing data,[9] we collect our own data by examining whether each dual-class firm's terminal year is due to a share unification using the sources mentioned above. We search for news articles that publicly announce these unifications first time in major news outlets using Factiva and LexisNexis.

2.2 Additional data sources

Patents. Another outcome variable of interest in our analyses is patent citations, which capture innovative output as a proxy for the effort that a firm's founder exerts. We combine two patent datasets: data compiled by Kogan et al. (2017), and data directly from the US Patent Office combined with a patent-firm linkage created by Ma (2025). We refer to these datasets as KPSS and SM, respectively.[10] We aggregate the patent-permno-level KPSS data at the firm (i.e., gvkey) level, combining patents across multiple permnos in a given year. However, KPSS is disproportionately missing dual-class firms, as the authors intentionally drop patents if they match with multiple permnos. To expand our data coverage, we use the SM data between 1975 and 2020 (the years in which it is most complete): for firms recorded as dual-class at any point in our data, if KPSS does not record patents for a given firm-year, we use the SM linkage to fill in this information. This combined linkage improves our measurement of patenting and citations for dual-class firms. In our empirical analysis, we use a forward-looking measure

[9] Ang and Megginson (1989) examine dual-class recapitalization and unification cases for U.K. firms in daily and monthly return event studies, and show positive effects of the former and mixed effects of the latter on stock price.

[10] We thank Noah Stoffman for making the KPSS patent dataset available, and Song Ma for sharing his patent-firm linkages.

that combines the number of citations one, two, and three years in the future as an outcome variable.

Ownership structure. We calculate ownership by corporate insiders, defined as the firm's executives and directors (e.g., Fabisik et al. (2021)), using the SEC's Form 3, 4, and 5 as provided by Refinitiv. Form 3 is the initial filing of ownership, Form 4 reports changes in ownership through transaction, and Form 5 is an annual summary of an insider's ownership. The database tracks the trading activity and resulting share ownership of insiders. We use the total shares held by an insider following the last transaction to capture his latest holdings in a given year. The resulting insider ownership data spans from 1988–2022. We forward- and back-fill an insider's holdings for a given firm in case a given insider-firm-year's ownership is missing.

We calculate institutional ownership for a given firm-institution using the SEC's Form 13F provided by Refinitiv. Form 13F is a mandatory report filed by institutional investors with assets under management at least $100 million. The dataset spans from 1980–2022 and records the total number of shares held by all reporting institutional investors in a given firm. Combining the insider ownership and ownership by institutions with at least 5% of a firm's equity ("block" institutional ownership), we calculate the Shapley value of ownership structure, a measure of the probability that incremental ownership is pivotal in a control contest (e.g., Rydqvist (1987)). We use it as an important control for the control-contest probability in the analysis of the voting premium in Section 4.6.

2.3 Sample selection

We merge the dataset on dual-class firms with CRSP–Compustat, from which we obtain information on stock prices and firm-level financials. We require that firm-year observations have the following variables constructed: book assets, Tobin's q, market leverage, research and development (R&D) expenses scaled by lagged book assets, asset tangibility, return on assets (ROA), payout ratio, sales growth rate, and SIC codes. We impute missing values of R&D to zero (for a similar adjustment see Brav et al. (2018)). A key variable for our analysis is firm age. Similar to Pástor and Veronesi (2003), we define the firm's birth (hence age is zero) as public firm as the first year in which the firm appears in Compustat with a valid share price. Appendix D shows the definitions of the variables.

We exclude firms in the financial (SIC 6000-6999), utilities (SIC 4900-4999), and unclassified (SIC 9900-9999) industries. To mitigate the influence of outliers, we exclude firm-years with book assets less than $20 million in 2000 constant dollars (adjusted using CPI) and Winsorize potentially unbounded variables at the 2^{nd} and 98^{th} percentiles. We obtain quantitatively

similar results by Winsorizing at an alternative level such as 1%, 3%, and 5%. These sample selection criteria lead to a final sample of 9,950 dual-class firm-years across 949 unique firms from 1971 to 2022. By adding 150,752 single-class firm-years, we have 160,702 firm-years in total.

INSERT FIGURE 2 HERE.

Figure 2 shows the number of dual-class firm-years with the aforementioned variables from 1971 through 2022, in comparison with the Compustat universe. The fraction of firms with dual-class shares ranged between 2.8% and 3.4% before the early 1980s, increased to 7.2% to 7.4% in the early 1990s, and has stayed between 6.0% and 8.0% since then. The rapid increase in the number of dual-class firms during the 1980s reflects many firms adopting takeover defenses in the period of heightened hostile takeover activities (e.g., Jarrell and Poulsen (1988)). Section 4 examines the effects of these dual-class recapitalizations, as well as unifications. The increase in number of dual-class firms in the last two years reflects a dual-class IPO "boom" in the Covid period.

INSERT TABLE 1 HERE.

Table 1 shows descriptive statistics for the samples of dual-class and single-class firm-years. Dual-class firms tend to be larger, older, and more highly levered, and to have higher ROA than single-class firms. They also have lower capital expenditures and R&D expenses than single-class firms.

2.4 Descriptive evidence: Valuation of dual-class firms

On average, Tobin's q for dual-class firms is significantly higher than that for single-class firms by 0.226. In Appendix C, we examine this relation in regressions with firm-level control and industry-by-year fixed effects. Appendix Table C1 shows that valuations are significantly higher for dual-class firms than otherwise comparable single-class firms in the cross-section. We then examine how this valuation gap evolves over time since firms' IPOs. Appendix Figure C1 shows that the difference in Tobin's q is significantly positive for most of the first decade post-IPO but declines through age 50, when the difference becomes significantly negative. This initial, descriptive evidence is consistent with a "transitory dual-class premium," as well as comparable estimates in Cremers et al. (2024). Nonetheless, these estimates are likely affected by both observable and unobservable heterogeneities across dual-class and single-class firms and over time. For a more credible testing of the model implication, therefore, we turn to the difference-in-differences analysis of changing governance structures in Section 4.

3 Theoretical Framework and Predictions

In this section, we present our theoretical framework and predictions that motivate our empirical analysis.

3.1 Setup

Consider a firm run by a risk-neutral founder. The firm has an infinite horizon and its total number of shares is normalized to one. An IPO is scheduled at time $t = 0$. Pre-IPO, the founder owns $\alpha \in (0, 1]$ shares of the firm. The founder retains a fraction $\Lambda \in [0, 1]$ of his ownership at the IPO, selling the rest to outside shareholders (referred to as "shareholders" hereafter).

Before the IPO, the founder exerts effort to implement his vision for the firm, which could transform it, thereby increasing its value. He chooses effort $e \geq 0$ at a private cost $\frac{1}{2}\kappa e^2$, where e is unobserved by shareholders and cannot be contracted upon. $\kappa > 0$ can be interpreted as the inverse of the founder's skills and ability to transform the firm. Given the effort, the firm's expected operating cash flow in period $t \geq 1$ is $y_t = e(1+g)^{t-1}$, where $g > 0$ is the firm's growth rate.

A key friction is an agency conflict between the founder and shareholders. The firm pays out all available cash flows to shareholders. Before dividends are paid out, however, the founder extracts a fraction $b_t \in [0, \phi]$ of the cash flows as private benefits, where $\phi \in (0, 1]$. These private benefits can be interpreted as reluctance to make personally costly decisions that benefit shareholders or investment in negative-NPV projects that generate perks to the founder. Thus, the paid dividend at period t is $(1 - b_t)\, y_t$, and the founder's utility from private benefits is $\gamma\, b_t\, y_t$, where $\gamma \in (\alpha\Lambda, 1)$. Assuming $\gamma < 1$ implies that private benefits consumption is inefficient and entails costs on both the founder and shareholders. Assuming $\gamma > \alpha\Lambda$ guarantees that the founder has incentives to extract private benefits even though it dilutes the value of his ownership. We assume b_t increases over time as follows:

$$b_t = \phi \left(1 - \frac{1}{(1+\tau)^{(t-1)}} \right). \tag{1}$$

Intuitively, the longer the founder holds office, the more powerful he becomes, enabling to extract greater private benefits (e.g., Graham, Leary, and Kim (2020)). Section 4.6 provides evidence that supports this key assumption using dynamics of the voting premium. Larger $\tau > 0$ implies the founder becomes entrenched more quickly.

Shareholders may fire the founder and appoint a new manager. We assume the majority of voting rights is owned by the founder (shareholders) under dual-class (single-class) shares.

Thus, under single-class shares (i.e., one share–one vote), shareholders can fire the founder at any point in time. Once fired, the founder can no longer consume private benefits (but retains his ownership). The new manager extracts no private benefits (i.e., $b_t \equiv 0$ for all t) but has no ability to transform the firm. (Alternatively, single-class shares could allow shareholders to monitor the firm and limit the founder's ability to extract private benefits.) In contrast, under dual-class shares, the founder has indefinite control of the firm and can never be fired. In Appendix A, we consider a richer set of governance structures including dual-class shares with a time-based sunset provision.

We denote the firm's stock price at time $t \geq 0$ by $p(t)$ and cost of capital by $r > g$. Under either governance structure, the stock price is the present value of all future dividends. The **timeline** is as follows:

1. IPO (period $t = 0$)
   (a) The founder sets up the firm's governance structure, which is observed by shareholders.
   (b) The founder chooses effort level e, which is unobserved by shareholders.
   (c) The founder sells $(1-\Lambda)\alpha$ shares at price $p(0)$.
2. Post-IPO (at each period $t \geq 1$)
   (a) If the firm has single-class shares and the founder still holds office, shareholders decide whether to fire the founder.
   (b) The firm produces operating cash flow y_t.
   (c) If the founder is retained, he extracts private benefits $b_t y_t$ and the firm pays out a dividend $y_t(1-b_t)$. If a new manager holds office, the dividend paid is y_t.

3.2 Analysis

3.2.1 Founder's incentives and treatment effect

The key trade-off is between the agency cost imposed by the founder and the "principal cost" imposed by shareholders. Single-class shares reduce the agency cost by limiting private benefits extraction, which increases over time. However, shareholders are tempted to exercise their voting power and fire the founder once his effort is sunk. This *ex post* expropriation risk diminishes his incentives to exert effort *ex ante*—a type of holdup problem.[11].

[11]The holdup is similar to that in Burkart, Gromb, and Panunzi (1997) in a context of large shareholders.

In general, the founder's motivation to exert effort stems from two sources. First, effort increases the likelihood that the firm transforms and thus increases the value of the founder's equity in the firm. Second, since the founder's expected private benefits of control increase with firm value, they also increase with his effort. As we formally show in Appendix A, both forces imply that the founder exerts more effort under dual-class than single-class shares, i.e., $e^*_{\text{dual}} > e^*_{\text{single}}$, in equilibrium. Thus, the principal cost under single-class shares is the inability of shareholders to commit not to fire the founder once effort is exerted and the firm is transformed.

Let $p^*_{\text{dual}}(t)$ and $p^*_{\text{single}}(t)$ be the equilibrium share prices at period $t \geq 0$ of a dual-class firm and a single-class firm, respectively. The following proposition presents our prediction on the effect dual-class shares on firm value over maturity.

Proposition 1. *There exist* $\underline{\phi} < 1$ *and* $\overline{\phi} > 0$ *where* $\underline{\phi} < \overline{\phi}$*, such that if* $\phi \in (\underline{\phi}, \overline{\phi})$ *then* $p^*_{dual}(t) > p^*_{single}(t)$ *if and only if* $t < T^*$ *for some* $T^* \in (0, \infty)$.

Proposition 1 shows that as long as the cost of private benefits extraction on shareholders is not too large (i.e., $\phi < \overline{\phi}$), the firm has higher value under dual-class than single-class shares when it is relatively young, but the value wedge shrinks and even reverses over time. (See Appendix A for explicit expressions for $\underline{\phi}$, $\overline{\phi}$, and T^*, and their dependence on other parameters.) We refer to this dynamic effect of dual-class shares on firm value as a **treatment effect**—i.e., when dual-class shares are randomly assigned to firms.

Intuitively, dual-class shares give the founder stronger incentives for effort since he expects to keep the position and to extract private benefits from control. However, the rising agency costs erode the added value from the founder's heightened incentives over time. If the agency cost is not too small (i.e., $\underline{\phi} < \phi$), a dual-class firm is valued at a discount relative to an otherwise identical single-class firm years after the IPO, i.e., $p^*_{\text{dual}}(t) < p^*_{\text{single}}(t)$ for $t > T^*$. Thus, Proposition 1 rationalizes the key empirical findings below.

3.2.2 Effects of changing governance structures, sunset provisions, and voting premium

The founder sets up a governance structure once and for all at the IPO. However, if a single-class firm unexpectedly recapitalize to be dual-class at period $t > 0$, then its share price will decrease by the present value of private benefits of control that the founder will extract. Similarly, if a dual-class firm unexpectedly unify its share classes at $t > 0$, then its share price will increase by the present value of private benefits which the founder would have otherwise extracted. We show in Appendix A that the abnormal return around the announcement of a dual-class stock recapitalization (unification) is negative (positive). Importantly, given our assumption that the

fraction of cash flows that go to private benefits, b_t, increases over time, the abnormal return is larger in absolute value for more mature firms. This prediction is consistent with the evidence we provide from event studies in Section 4.2.

If we were to extend our model to allow the founder to exert effort post-IPO, the firm's productivity will increase upon a dual-class recapitalization due to the founder's heightened incentives. Thus, the abnormal return for young firms would be *positive*, to the extent that the loss from increased agency costs is relatively small for young firms. Similarly, the firm's productivity will decrease upon a unification due to reduced incentives. Then the abnormal return for young firms would be *negative*, again, to the extent that the gain from reduced agency costs is relatively small for young firms. Given the increasing fraction of cash flows that go to private benefits over maturity, the abnormal return decreases (increases) as firms mature post-recapitalization (-unification). These predictions are consistent with the evidence from dynamic difference-in-differences analysis of switchers in Section 4.1.

In Appendix A, we analyze dual-class shares with time-based sunset provisions, which allow shareholders to fire the founder only n periods after the IPO. Similar to the result for the price effect of an unexpected unifications, we show that the stock price reaction to an expected unification due to a sunset provision is also positive but smaller than that for cases without a sunset. The intuition for this result is that with a sunset, dual-class shares entail smaller private benefits consumption.[12] If the founder is allowed to exert effort post-IPO, then the key implication of our model would be that sunsets weaken the incentive effect of dual-class shares. Hence, stock unifications triggered by sunsets will have little effect on firm value, consistent with the finding in Figure A1.

Lastly, our model enables us to study how the difference between the value of voting and non-voting shares (referred to as the voting premium) changes over maturity. In our framework, the difference between the value of a voting share (which is held by the founder) and a non-voting share (held by shareholders) is the present value of the founder's future private benefits of control.[13] In Appendix A, we show that the voting premium is positive and increases over maturity, consistent with the evidenc in Section 4.6.

[12] In general, the magnitude of the positive stock price reaction is larger for sunset provisions with longer maturities (i.e., larger n). Intuitively, sunsets with longer maturities are more similar to having no sunset provision at all.

[13] Levit et al. (2024) study the determinants of the voting premium in a static model where both voting and non-voting shares are traded. They show a nuanced relationship between the voting premium and the conflict of interests between shareholders of the firm.

3.2.3 Founder's decision on governance structure and selection effects of dual-class shares

Our second result characterizes the optimal governance structure from the perspective of the founder at the IPO.

Proposition 2. *There exists $\phi^* > \overline{\phi}$ such that at the IPO the founder prefers dual-class shares over single-class shares if and only if*

$$\phi < \phi^*. \tag{2}$$

Moreover, ϕ^* increases in γ and decreases in τ.

To understand Proposition 2, recall that dual-class shares enable the founder to extract private benefits of control. If that was the only consideration, the founder would have always preferred dual-class shares. However, the founder also sells shares at the IPO, and thus must consider the effect of governance on IPO proceeds, which are determined by the value of shares from the perspective of shareholders.[14] Proposition 1 shows that, if $\phi < \overline{\phi}$, a dual-class firm is valued with a premium at the IPO and therefore, the founder faces no trade-off: dual-class shares enable him to both consume private benefits and maximize the share price at the IPO. However, if $\phi > \overline{\phi}$ (i.e., the agency friction is "severe"), then a dual-class firm is valued with a discount at the IPO. Now the founder faces a trade-off: choosing dual-class shares enables him to consume private benefits but leads to smaller IPO proceeds due to the larger expected agency costs. Proposition 2 shows that if the agency friction is not too severe (i.e., $\overline{\phi} < \phi < \phi^*$) and the resulting deadweight loss is small (i.e., large γ), and/or the founder becomes entrenched slowly over time (i.e., small τ), then the additional utility from the expected private benefits dominates, and the founder still prefers dual-class shares. In contrast, if the agency costs are too large (i.e., $\phi > \phi^*$, small γ, and/or large τ), then its negative impact on IPO proceeds dominates, and thus the founder prefers single-class shares.[15]

Proposition 2 highlights that the choice of the firm's governance structure is likely non-random and driven by factors that might also determine how firm value evolves over time as Proposition 1 characterizes. The model allows us to analyze whether the dynamic treatment effect of dual-class shares in Proposition 1 can also be attributed to a **selection effect**.

The next result shows that in general, a selection effect can generate the dynamic pattern in Proposition 1 only under specific conditions.

[14] While outside our model, another cost of lower IPO valuations is the smaller amount of capital the firm raises.

[15] Appendix A provides conditions under which the governance structure that maximizes firm value at the IPO involves dual-class shares with time-based sunset provisions. Since the founder trades off IPO proceeds with his private benefits of control, she generally has incentives to adopt sunsets with longer maturities than what is optimal for shareholders.

Proposition 3. *A selection effect can generate the dynamic pattern in Proposition 1 only if firms differ in parameter τ, the rate at which their founders become entrenched over time. Specifically, relative to a selected dual-class firm, a selected single-class firm has higher τ, its founder exerts more effort and extracts greater private benefits especially when the firm is young.*

Since τ is high for the firm, its founder expects to consume greater private benefits and in turn has stronger incentives for effort. This "effort effect" gives the selected single-class firm a higher value. On the other hand, the agency costs of the selected single-class firm grow faster. This "private benefits effect" gives the selected single-class firm a lower value. In Appendix B, we show there are cases in which the private benefits effect dominates for young firms, while the effort effect dominates for mature firms, which allows selection to generate the transitory dual-class valuation premium.[16] Intuitively, with high τ a marginal increase in the private benefits over time is relatively small for mature firms since the founder is already highly entrenched.

In addition, we show in the proof of Proposition 3 in Appendix B that when selection is due to other parameters than τ, the selection effect is either constant over time—i.e., the value of a selected single-class firm is always higher or lower than that of a selected dual-class firm, as in Figure 3—or the opposite from Proposition 1, i.e., dual-class firms' value is higher than single-class firms' when they are mature, not when they are young.

4 The Dynamic Effects of Dual-Class Shares

Our model posits that dual-class shares allocate control to the founder when his investment is crucial for the firm's success. The key mechanism is that dual-class structures protect the founder from losing private benefits *ex post*, providing him with strong incentives to exert firm-specific effort *ex ante*. This model has two major predictions. First, young dual-class firms have higher value relative to otherwise identical single-class firms but this premium shrinks as they mature (Proposition 1). We test this implication in Sections 4.1 and 4.2. Second, if this transitory dual-class premium is due to treatment (selection), then both the incentives for effort and private benefits are greater for dual-class (single-class) firms. In addition, under treatment (selection), the differences in effort and private benefits between dual-class and single-class firms grows (shrinks) as firms mature (Proposition 3). We test these implications in Sections 4.3 and 4.5.

[16] In other cases where the effort effect dominates the dynamic agency costs effect for young firms, selection cannot generate the dynamic pattern in Proposition 1.

4.1 Research design: Difference-in-differences event study

Our main empirical analysis uses a stacked difference-in-differences (DD) event-study design (e.g., Gormley and Matsa (2016); Cengiz et al. (2019)). For recapitalization, the treatment group includes firms that recapitalize with dual-class shares, whereas the control group includes firms that remain to be single-class. For unification, the treatment group includes firms that unify share classes, whereas the control group includes firms that remain to be dual-class. We consider alternative, matched control groups using pre-event firm characteristics below. The key identifying assumption for a treatment-effect interpretation of the DD estimates is parallel trends—which we test using pre-event observations. We estimate the following difference-in-differences equation for dual-class recapitalization and stock unification:

$$q_{ijth} = \alpha_{ih} + \alpha_{jth} + \sum_{\substack{\tau=-5 \\ \tau\neq-1}}^{15} \lambda_\tau d[t+\tau]_{ith} + \sum_{\substack{\tau=-5 \\ \tau\neq-1}}^{15} \delta_\tau Treat_{ih} \times d[t+\tau]_{ith} + \gamma' X_{it} + \epsilon_{ijth}, \qquad (3)$$

where q_{ijth} is Tobin's q for firm i in three-digit SIC industry j in year t of event-year cohort h; α_{ih} and α_{jth} represent firm i by cohort h and industry j by year t by cohort h fixed effects; $Treat_{ih}$ is an indicator variable equal to one if firm i experiences a recapitalization (unification), and zero if remains to be single-class (dual-class) in event-year cohort h; $d[t+\tau]_{ith}$ is an indicator variable equal to one if year t is $|\tau|$ years (-5 $\leq \tau \leq$ 15) before or after a recapitalization (unification) for firm i in cohort h, and zero otherwise. Year $t-1$ is the baseline year and thus $d[t-1]_{ith}$ is equal to zero by construction; X_{it} is a vector of control variables in Eq. (C.1); and ϵ_{ijth} represents random errors clustered at the firm level. We require that both the treated and control firms exist over the $[t-2,\ t+2]$ window, and exclude firm observations once they make another switch after year t. In essence, this stacked DD specification estimates the average effects of recapitalization (unification) across all event-year cohorts with the firm and industry-by-year fixed effects varying by event-year cohort. Our DD sample contains 186 recapitalizations and 224 unifications, representing a substantial share of about 1,100 unique dual-class firms.

Figure 4 presents the estimation results for recapitalization (Panel A) and unification (Panel B). First, both panels show no evidence of pre-trends. All estimates on $Treat \times d[t+\tau]$ for the pre-event period (i.e., τ = -5 through -2) are statistically insignificant and economically small. This result supports the key identifying assumption for a treatment-effect interpretation of the event-study estimates, and is inconsistent with alternative interpretations including anticipated switches reflected in q and changes in q affecting the switching decision—instances of reverse causality.

INSERT FIGURE 4 HERE.

Estimates in Panel A show that firms that recapitalize from single-class to dual-class shares experience an immediate increase in Tobin's q by 0.35, relative to a control group of firms that remain to be single-class throughout the event window. Post-recapitalization, the treated firms' q remains elevated until year $t+5$ but starts declining in year $t+6$. From year $t+10$ and onward, the differences in changes in treated and control firms' q are close to zero and statistically insignificant. These post-recapitalization dynamics are consistent with Proposition 1.

Turning to unifications in Panel B, the estimated dynamics are essentially a mirror image of the corresponding estimates for recapitalizations. Firms that unify share classes experience a *decrease* in Tobin's q relative to a control group of firms that remain dual-class. The unifying firms' q remains lower than the control firms' but the estimated treatment effects become less precise around year $t+7$. Eleven years after a unification, any discount in treated firms' q relative to control firms' essentially disappears. Again, these dynamic patterns are consistent with Proposition 1. The generally lower precision of the estimates for unifications likely owes to a much smaller number of firms in the control group—firms that remain to be dual-class.

One potential concern with this event study is that switching firms might have different characteristics than the control groups prior to treatment, which may bias the estimated treatment effects (e.g., Abadie (2005)). Appendix F shows the balance tables for these switching groups and employs a matching approach to address observable differences between the treatment and control groups. In this matched-sample DD, we find very similar results to the full sample event study, suggesting that differences in observed firm characteristics are not a major driver of our baseline results.

Appendix F reports the results of two additional robustness tests for this difference-in-differences analysis. The first test excludes dual-class firms that are equity carveouts, which we identify using data from Field and Lowry (2022) and John McConnell.[17] These carveouts differ from dual-class firms our model posits in that they are typically controlled other firms, instead of founders or their teams. For both types of events, the DD results are virtually identical even after excluding carveout firms. As a final robustness test, we perform the event-study analyses using the Sun and Abraham (2021) estimator, which avoids issues with two-way fixed effect models, such as not aggregating the treatment effects with appropriate weights. In our setting, the following facts mitigate this issue: (i) there are never-treated firms in the control group and

[17]We thank Michelle Lowry and John McConnell for sharing and making the datasets on equity carveouts available.

(ii) our treatment is not staggered (we use a stacked DD approach). Nonetheless, we find our DD analysis is robust to using this alternative estimator.

Overall, the DD event-study estimates show dynamic effects of recapitalization and unification on valuations consistent with the model predictions. The lack of pre-trends provides credence to a treatment-effect interpretation of the changes in Tobin's q after firms making changes to voting structure.

While the switcher firms allow us to examine treatment effects of dual-class shares, our model predicts similar dynamics post-IPO—the valuation of dual-class firms declines faster over time relative to that of single-class firms. Thus, in Table A1, column (1) we perform a within-firm analysis of Tobin's q after IPOs. We find that post-IPO, Tobin's q for dual-class firms declines significantly faster than that for single-class firms, controlling for firm and IPO cohort-by-industry-by-year fixed effects and firm-level controls. While the nature of this test does not allow for estimating a treatment effect of dual-class shares, it nonetheless provides evidence consistent with the model's prediction.

4.2 Market reactions to dual-class recapitalizations and stock unifications over maturity

As a complementary approach to testing the model prediction for dynamics of the dual-class premium, we examine how the market value of equity changes in the short run in response to dual-class recapitalizations and stock unifications. By using short-run stock returns, this approach sidesteps potential issues of valuation measures such as Tobin's q related to, e.g., the replacement cost of assets (see, e.g., Dybvig and Warachka (2015) and Bartlett and Partnoy (2020)). We describe our collection of announcement dates for these events in Section 2.1.

For the sample of recapitalization and unification announcements, we compute excess daily stock returns using a market-adjusted-return model: $\epsilon_{it} = R_{it} - R_{mt}$, where ϵ_{it} is the rate of abnormal return and R_{it} is the rate of stock return for firm i, and R_{mt} is the rate of return for the market portfolio, proxied by *vwretd* from CRSP on day t. We use cumulative abnormal returns (CARs) over the [-3, +3] window in the analysis, and find a qualitatively similar result by using alternative event windows such as [-2, +2] and [-2, +3]. To increase the sample size, for this analysis only we include firms in the financial, utilities, and unclassified sectors. Requiring the CAR provides a sample of 88 dual-class recapitalizations and 62 unifications announced between 1971 and 2022.

We estimate a regression in which the CAR is the dependent variable and the key independent variables include a $\mathbb{1}[\text{age} \geq 12]$ indicator and year fixed effects. The key identifying

assumption is that recapitalizing or unifying firms in a given year are otherwise comparable with each other. Later in this section, we examine whether selection to be dual- or single-class firms may explain our findings based on our model implications.

INSERT TABLE 2 HERE.

Table 2 presents the estimation results. The positive regression constants in Columns (1) and (2), 0.026 and 0.035 (significant at the 10% level), imply a positive value effect of dual-class recapitalizations for young firms. This positive effect is consistent with the incentive role of dual-class shares outweighing relatively small private benefits of control early in the firm's life. However, the negative coefficients on $\mathbb{1}[\text{age} \geq 12]$, -0.034 and -0.046 (significant at the 10% level), show that for mature firms, the value effect of recapitalization is lower by 3-5 percentage points. This key result is consistent with the model implication that a dual-class recapitalization has an increasingly negative effect on firm value over maturity, due to increasing private benefits, as well as the dyanmic DD estimates in the previous section. These results potentially explain the insignificant average effect of the recapitalization shown in previous research (e.g., Partch (1987); Jarrell and Poulsen (1988)).

Next, we examine the market's reaction to the announcement of dual-class stock unifications. The positive coefficient on $\mathbb{1}[\text{age} \geq 12]$ in column (4), 0.050 (significant at the 10% level), shows that the market perceives a 5 percentage-point higher value effect when mature firms unify dual-class shares compared to when young firms unify in the same year. The economic magnitudes of the effect of switching between dual-class vs. single-class structures on mature, relative to young, firms estimated using the two events are comparable—e.g., estimates on $\mathbb{1}[\text{age} \geq 12] =$ -0.046 and 0.050 in columns (2) and (4). This comparable magnitude suggests the estimates capturing a common economic force—such as the present value of private benefits of control—as in the model.

4.3 Technological innovation—evidence for incentive effects of dual-class shares

Sections 4.1 and 4.2 provide reinforcing evidence for a transitory dual-class premium in which dual-class firms have higher valuations than single-class firms when they are young but the valuation gap shrinks over maturity. In this section, we test the implications for the founder's effort by employing technological innovation as an outcome of the firm-specific effort. The idea is that under the treatment (selection) effect, the greater effort the founder of a dual-

(single-) class firm exerts, and greater firm-specific investments in general, will be reflected in greater innovative output the firm produces.[18]

We estimate a version of difference-in-differences Eq. (A1) in which the dependent variable is the number of citations the firm-year's patents generate in one, two, or three years ahead (e.g., Bernstein (2015)). To account for both the extensive- and intensive-margin effects, we use the inverse hyperbolic sine transformation. In addition, we examine the two margins separately by employing (i) indicator variables equal to one if the firm-year has any patents that generate citations, and zero otherwise; and (ii) the log of the number of the citations the firm-year's patents generate as the dependent variables.

Figure 5 presents the event-study estimates for patents. Both panels show no evidence of pre-trends, supporting a treatment-effect interpretation of the estimates. Panel A shows that post-recapitalization, the treated firms experience an increase in patent output, as measured by their citations, relative to firms that stay single-class, the control group. In contrast, Panel B shows that post-unification, the switching firms experience a decline in patent citations, relative to firms that stay dual-class. The effects are generally persistent over time.

INSERT FIGURE 5 HERE.

These results are consistent with a treatment effect in which the founder exerts greater firm-specific effort under dual-class than single-class structures—which in turn increases the probability of "transforming" the firm, as measured by generating greater innovative outputs. We further explore this implication by separately examining the extensive and the intensive margins of innovation in Appendix F. Following Chen and Roth (2024), we estimate upper and lower bounds on the intensive margin effect (Lee (2009)). We find that the positive effect of dual-class recapitalization on innovative output is largely driven by the extensive margin—i.e., the recapitalizing firms become innovative. For unifications, the decline in innovative output is driven by both extensive and intensive margin effects.

In Appendix F we also report these DD estimates on innovation output using the series of robustness tests that we run in Section 4.1 for q: using matched samples; excluding dual-class firms with equity carveouts; and using the Sun and Abraham (2021) estimator. Again, our results on firm innovation are consistent across all these tests.

Overall, the evidence on the dynamics of innovative output from the difference-in-differences analysis is consistent with a treatment effect of dual-class shares rather than selection. In particular, the results suggest that the heightened incentives provided by dual-class shares have

[18] Related, Acharya, Baghai, and Subramanian (2014) show evidence that wrongful discharge laws, which limit firms' ability to holdup, encourage employees' firm-specific innovative effort and as a result, spur innovation.

a larger impact on firms transitions into innovating (the extensive margin) than on the level of innovative output once firms become innovative (the intensive margin). Arguably, the former requires greater firm-specific investment, for which dual-class shares are well suited.

Finally, we conduct a within-firm analysis of innovative output for firms that go public with dual-class versus single-class share structures. Our model predicts that, relative to single-class firms, dual-class firms make increasing firm-specific investments over time following the IPO. We note that this test does not identify a treatment effect of dual-class shares on firm-specific investment. Nonetheless, the results in Table A1, column (2), show that post-IPO, dual-class firms exhibit increasingly higher innovative output over time relative to single-class firms, consistent with the model's predictions.

4.4 Specificity of investments and the effects of dual-class shares

Next, we test a core mechanism of our model: dual-class shares incentivize founders to undertake firm-specific investments by protecting them from the loss of private benefits. One implication is that the effects of dual-class shares are more pronounced when founders' investments are more firm-specific; otherwise, there is no holdup problem, and thus no effect of dual-class shares.

Given the lack of a readily available measure of the importance of firm-specific investments by founders, we proxy for it using a combination of the specificity of key production inputs: capital and labor. The idea is that the founder's investment is likely more specific when the capital and labor that he manages are more specific. Using this measure, we examine heterogeneity in the effects of firm switches on both Tobin's q and innovative output.

For capital, we employ an industry-level measure of asset redeployability from Kim and Kung (2017). Assets that are less widely-used across industries are considered more specific. For labor, we use an index of labor mobility from Barry et al. (2025), which measures how widely workers in a given industry move to other industries. If workers in an industry only move to a small set of other industries, they are considered more specific.[19] Finally, we take a weighted average of the standardized labor and capital measures using the labor share as the weight for labor and (1 – labor share) for capital. We construct an industry-level measure of the labor share, extending data from Elsby et al. (2013).[20]

[19]The asset redeployability measure is at the firm level, weighted-averaged across industry segments. We fill in missing firm-level measures using industry-level measures by BEA industry code, cross-walked to three-digit SIC codes. The labor mobility measure is at the 1990 Census industry code level, which we crosswalk to three-digit SIC codes following the approach in Autor et al. (2019).

[20]The data are from the Bureau of Labor Statistics KLEMS dataset. We measure the labor share for two-digit SIC industries from 1970 to 1986, and two-digit NAICS industries from 1987 to 2022.

Table 3 shows results for the difference-in-difference analysis for Tobin's q and patent citations, with the sample split at the median of investment specificity. Panel A shows recapitalizations and Panel B shows unifications. We find that the effects of recapitalization and unification are concentrated among firms in which the founder's specific investments are likely important — when their capital and labor are more specialized. In contrast, the effects are generally flat for firms with less specialized production inputs.

4.5 Sensitivity of investment to opportunities—evidence for private benefits of control

We now turn to testing the model implication for private benefits of control. The model shows that if the dynamic effects of dual-class shares are due to treatment (selection), then founders of dual-class firms consume greater (smaller) private benefits than those of single-class firms—and this gap increases (decreases) over maturity. We employ the extent to which firms' investment decisions are *insensitive* to investment opportunities as a measure of private benefits. The idea is that reluctance to adjust capital in response to changes in growth opportunities (relative to, e.g., a neoclassical benchmark) may reflect that founders (and managers) have stronger preferences for preserving what they have invested in than shareholders do (e.g., Morck et al. (1998); Bertrand and Mullainathan (2003)).

We thus estimate how the q-sensitivity of investment differs between dual-class and single-class firms over maturity. Following the extensive literature on investment, we use Tobin's q as a proxy for marginal q and also include cash flows in an investment equation (e.g., Kaplan and Zingales (1997); Hubbard (1998)). The resulting investment equation is:

$$Investment_{it} = \alpha_{i\times\text{Dual}} + \alpha_{t\times\text{Dual}} + \beta_1 q_{it-1} + \beta_2 q_{it-1} \times Dual_{it} + \beta_3 CF_{it} + \beta_4 CF_{it} \times Dual_{it} + \epsilon_{it}, \tag{4}$$

where $Investment_{it}$ is capital expenditures scaled by lagged book assets; $Dual_{it}$ is an indicator variable equal to one if firm i has dual-class shares in year t; $\alpha_{i\times\text{Dual}}$ and $\alpha_{t\times\text{Dual}}$ represent firm and year fixed effects interacted with $Dual_{it}$; q_{it-1} is beginning-year Tobin's q, a proxy for marginal q; CF_{it} is cash flow scaled by lagged book assets for firm i in year t; and ϵ_{it} represents random errors clustered at the firm level.

To examine whether dual-class and single-class firms exhibit different sensitivities over maturity, we also estimate Eq. (4) on subsamples of young (mature) firms with ages less than (greater than or equal to) 12 years, the median age. Given that the relevant private benefits

likely involve reluctance to cut capital, we estimate these sensitivities on samples of firms facing relatively weak demand conditions—when when they are likely near the disinvestment threshold—proxied by sales growth in the first quartile (e.g., Achyuta et al. (2013)), These results are robust to other cutoffs (see below). We estimate this equation both with and without industry-by-year-by-dual class fixed effects.

INSERT TABLE 4 HERE.

Table 4 reports the estimation results for Eq. (4). All columns show that investment is strongly positively correlated with q for single-class firms, as captured by baseline coefficients on q, both for young and mature firms. However, columns (2)-(3) show that mature dual-class firms have lower investment-q sensitivities than mature single-class firms, as captured by the negative and significant coefficients on $q \times Dual$ (at the 5% level), whereas there is no significant difference for young firms. To ensure that our results are robust to the cutoff for sales growth rates—we use the 25th percentile in Table 4—we show that these results hold for alternative cutoffs ranging 15th to 30th percentiles in Table A2.

Overall, these results are consistent with a treatment effect of dual-class shares but not with selection—which would predict single-class firms extract *more* private benefits of control than dual-class firms especially when they are young. Our tests strongly reject the hypothesis that $q \times Dual > 0$.

4.5.1 Implications for systematic risk and valuations of dual-class firms

In a neoclassical framework, the reduced sensitivity of investment near the disinvestment threshold represent higher downward adjustment costs.[21] Neoclassical asset-pricing models show that firms with high downward adjustments costs carry high systematic risk and exhibit characteristics of value firms (e.g., Zhang (2005); Cooper (2006)). The idea is that those firms generate lower cash flows when the price of risk is high and carry higher risk premiums. Therefore, based on the q-sensitivity results in Table 4, we hypothesize that dual-class firms have a higher HML ("value") factor loading than single-class firms particularly when they are mature. We test this hypothesis by estimating a Fama-French-Carhart (Fama and French (1993); Carhart (1997)) four-factor model using monthly zero-cost calendar-time portfolios that long dual-class firm stocks and short matched single-class firm stocks from 1976 through 2022. To avoid any mechanical effect of book-to-market ratios on HML factor loadings, for

[21] Specifically, one can interpret investment–q sensitivities as an inverse proxy for the convex portion of capital adjustment costs (see Abel and Eberly (1994) and Peters and Taylor (2017)).

each dual-class firm-year we find matched single-class firm-years in the same Fama-French 48 industry and year with book-to-market ratios within a [0.85, 1.15] bandwidth. We form the portfolios from the split samples of young and mature firms, defined at 12 years, the median age, as well as from the full sample. We begin the analysis in 1976 when there are a sufficient number of mature firms to form a portfolio.

INSERT TABLE 5 HERE.

Table 5 presents the estimation results for value- (Panel A) and equal-weighted (Panel B) portfolios. Column (1) in the equal-weighted portfolio shows that on average, the zero-cost portfolios that long dual-class and short matched single-class firms exhibit positive loadings on the HML factor (0.081), while we estimate a null loading on the HML factor for the value-weighted portfolio. For the value-weighted portfolio, the HML loading is considerably larger for mature firms than young firms, although it is not significant. In the equal-weighted portfolio, there are comparable loadings on the HML factor for portfolios with only young and only mature firms (0.107 and 0.093, respectively). These different factor loadings between dual-class and single-class firm stocks across maturity match the pattern shown in Table 4 that dual-class firms exhibit lower q-sensitivities than single-class firms especially when they are mature.

Therefore, the overall results in this section show that mature dual-class firms are likely riskier than mature single-class firms due to their higher downward adjustment costs, a form of private benefits of control. In addition to providing evidence for a treatment effect, the asset-pricing results suggest that increasing systematic risk, which works through discount rates, is another plausible channel for dual-class firms' relative value decline over maturity.

4.6 Voting premium: Private benefits of control increase with maturity

The previous section shows evidence that private benefits consumption is larger for dual-class firms *relative to* single-class firms particularly when they are mature. In this section, we test a related (but not identical) premise of our model—that private benefits of control in general increase over firm maturity. It is a crucial model premise, given that the growing private benefits are what lead the valuation premium for dual-class firms to shrink in the long run, as well as providing *ex ante* incentives to the founder early in the firm's life.

We use the voting premium—the difference between market prices of superior-voting shares relative to inferior-voting shares for dual-class firms—as a measure of private benefits of control. The model shows that the voting premium is the expected present value of future private benefits. In practice, the voting premium is also affected by other factors such as the

probability of a control contest, differences in liquidity across share classes, and heterogeneous beliefs among shareholders.[22] Therefore, our analysis controls for measures of other determinants of the voting premium, to the extent it is possible.

The analysis has to rely on a subsample of dual-class firms with both the superior- and inferior-voting shares traded and information on ownership structure. The resulting sample includes 1,188 dual-class firm-years (for 114 unique firms) from 1980 through 2022. To analyze the relation between firm maturity and the voting premium, we estimate the following regression:

$$VP_{it} = \alpha_i + \alpha_t + \sum_{n=2}^{4} \beta_n \mathbb{1}[\text{age in quartile}_n]_{it} + \gamma' X_{it} + \epsilon_{it}, \tag{5}$$

where VP_{it} is the voting premium for dual-class firm i in year t, computed as $(P_A - P_B)/(P_B - r P_A)$, where P_A (P_B) is the price of superior- (inferior-) voting shares and r is the ratio between the number of votes for the inferior- and superior-voting shares; α_i and α_t represent firm and year fixed effects; $\mathbb{1}[\text{age in quartile}_n]_{it}$ are indicator variables for the n^{th} firm age quartile, and firm age is defined using years since a firm's IPO; following Zingales (1995), X_{it} includes two proxies for the probability of a control contest: log market equity and the Shapley value of ownership structure, and the relative trading volume for firm i in year t; and ϵ_{it} represents random errors clustered at the firm level. When we include firm fixed effects, year fixed effect and firm age are collinear. To address this issue, we normalize year fixed effects following Deaton (1997).[23] To mitigate the influence of illiquidity and microstructure noise, we exclude firms with superior or inferior share prices below \$5. The coefficient of interest is β_n which we interpret as the relation between firm maturity and the private benefits. This interpretation is plausible partly because we control for the measures of the control contest probability, as well as relative trading volume.

INSERT TABLE 6 HERE.

Table 6 presents the estimation results for Eq. (5). Column (1) shows that more mature firms are associated with higher voting premium in the cross-section. The estimate suggests that, relative to the youngest quartile of firms, voting premium for firms in the third and fourth quartiles of age increases by 4.6 and 6.2 percentage points. Similarly, when we control for firm and

[22] See, e.g., Lease et al. (1983); Zingales (1995); Nenova (2003); and Doidge (2004). Zingales (1995) shows both theoretically and empirically that the voting premium is determined by a combination of the probability of a control contest and the private benefits of control. Similarly, Nenova (2003) argues that the value of control-block votes is a lower bound for the private benefits to the controlling shareholder.

[23] Specifically, we normalize year fixed effects to have a zero mean over the sample period, making them orthogonal to the linear time trend. Firm age therefore captures the linear growth in the voting premium, while year fixed effects capture nonlinear deviations from that trend.

year fixed effects in column (2), voting premium increases by 4.5 and 9.4 percentage points for firms in the third and fourth quartile, respectively. The economic magnitudes are considerable, compared to the mean and standard deviation of the voting premium of 5.1% and 16.4%.[24] In column (3), we include year fixed effects, along with firm fixed effects, by normalizing one year fixed effect to zero. The magnitude of the voting premium for firms in the third and fourth age quartiles, relative to the youngest firms is comparable to those in columns (1) and (2). These results are robust when we exclude log market equity as a potentially endogenous variable to voting premium in Table A3. Overall, the evidence in this section shows that private benefits of control reflected in the value of votes increase over firm maturity, providing support for a key premise of our model and empirical analysis.[25]

5 Policy Implications and Discussions

Given the increasing private benefits associated with managing dual-class firms over maturity (see Tables 4 and 6), founders may have little incentive to unify dual-class shares as their firms mature, even if doing so may increase shareholder wealth (see Figure 4 and Table 2). This collection of evidence, at face value, appears to support the case for so-called "sunset provisions"—clauses in corporate statutes that automatically eliminate dual-class shares once a specific date is reached or a threshold event occurs (we discuss examples below). In fact, some academics and policy makers have argued that dual-class shares should be precluded altogether or, if they are used, should include a sunset provision (e.g., Bebchuk and Kastiel (2017); Jackson (2018); Cremers et al. (2024)). Our theoretical and empirical analysis suggests the following policy implications regarding dual-class shares and sunset provisions.

First, contrary to some existing view mentioned above, securities regulators and exchanges may not want to outright ban listing firms with dual-class shares (which was the case for the NYSE pre-1984, the Hong Kong and Singapore Exchanges until 2018, and the LSE's Main Market premier tier until 2021). Our analysis shows that an important benefit of dual-class shares is providing incentives for firm-specific investments—which recent research shows have become

[24] These moments are similar to those reported in previous research (e.g., Lease et al. (1983); Zingales (1995)). The coefficients on control variables have expected signs: the log market value of equity is negatively and the Shapley value is positively related with the voting premium. These estimates are consistent with the probability of a control contest being higher for smaller firms and for firms with higher Shapley values. In particular, the coefficients on the Shapley value (e.g., 0.055 in column (1)) suggest that the private benefits of control are about 5.5% of total equity value, similar to estimates in Zingales (1995) and Barclay and Holderness (1989).

[25] The voting premium analysis has to rely on a sample of dual-class firms with both classes of stock traded. If private benefits are greater (smaller) for these firms relative to the population of firms (including single-class ones), the estimated magnitude of private benefits are likely an upper (lower) bound.

increasingly important since about 2000 (Kankanhalli, Kim, and Kung (2025); Barry, Graham, Kim, and Kim (2025)).

Second, and related, any policies that mandate sunsets on dual-class shares should consider the negative consequences—namely reduced incentives for making value-adding yet difficult-to-contract-upon investments, such as innovation and specialized assets. Our model shows that dual-class shares provide a "dynamic trade-off" to shareholders between benefits from the founder's enhanced incentives early on and greater agency costs later. In particular, the model implies that sunset provisions weaken the incentive effect of dual-class shares. In Figure A1, we find evidence consistent with this implication—when dual-class firms with sunsets unify share classes, their valuations virtually do not change (Panel A). If dual-class shares with sunsets provide relatively little incentive to begin with, then their removal would have little impact.

Ultimately, which among dual-class shares with and without sunset provisions, and single-class shares is most suitable for a given firm would depend on both firm-specific and market-wide factors. For example, for firms where the deadweight loss from private benefits consumption is small (as captured by high γ in the model), all else equal, dual-class shares are preferred to single-class or dual-class shares with a sunset. Also, the fraction of (young) firms adopting dual-class shares may endogenously rise when external governance forces (e.g., activist investors, corporate raiders) in the economy become stronger, as captured by low ϕ in the model. Finally, an increase in the importance of the founder's firm-specific investment for the firms success—the key mechanism of our model—would lead to increased adoption of dual-class shares. These could explain the recent rise in dual-class IPOs as shown in Figure 1. Analyzing these heterogeneities in the cross-section and over time is an important topic for future research.[26]

6 Conclusion

In this paper, we make an explicit point that dual-class shares provide dynamic incentives to entrepreneurs using both theoretical and empirical analyses. The heightened incentives allow dual-class firm to be more innovative and to increase their value early in their lives, whereas increasing agency costs over maturity erode the value added from the heightened incentives. The private benefits of control, as measured by the voting premium and reluctance to cut capital and labor when investment opportunities are poor, are most pronounced for mature dual-class firms. New to the literature, our theoretical model provides an analytical

[26] See, e.g., Philippon (2006) and Gutiérrez and Philippon (2018) for the effects of varying agency frictions on firm outcomes in the aggregate economy.

framework to study the treatment and selection effects of dual-class shares and importantly, highlights their starkly different implications for the post-IPO evolution of private benefits, founders' effort, and firm value. Our empirical analysis that exploits within-firm switches between single-class and dual-class shares provides evidence that is consistent with treatment effects of dual-class shares, including parallel trends. Overall, the combination of the theory and evidence is consistent with treatment effects of dual-class shares rather than selection.

More broadly, dual-class shares can replicate other forms of deviation from one share–one vote such as pyramids and cross-ownership, which are widely used to consolidate corporate control outside the US (Bebchuk et al. (2000); Claessens et al. (2002); Faccio and Lang (2002); Becht et al. (2018)). In addition, our theoretical framework can be applied to analyze the treatment and selection effects of other mechanisms that consolidate corporate control, such as classified boards and takeover defenses. Therefore, the theoretical and empirical results in this paper have broader implications for dynamic effects of an array of corporate control and governance mechanisms.

References

Abadie, Alberto, 2005, Semiparametric Difference-in-Differences Estimators, *Review of Economic Studies* 72, 119.

Abel, Andrew B., and Janice C. Eberly, 1994, A unified model of investment under uncertainty, *American Economic Review* 84, 1369–84.

Acharya, Viral V., Ramin P. Baghai, and Krishnamurthy V. Subramanian, 2014, Wrongful discharge laws and innovation, *Review of Financial Studies* 27, 301–346.

Achyuta, Adhvaryu, A. V. Chari, and Siddharth Sharma, 2013, Firing costs and flexibility: evidence from firms' employment responses to shocks in India, *Review of Economics and Statistics* 95, 725–740.

Adams, Rene, and Daniel Ferreira, 2008, One share–one vote: The empirical evidence, *Review of Finance* 12, 51–91.

Ang, James S., and William L Megginson, 1989, Restricted voting shares, ownership structure, and the market value of dual-class firms, *Journal of Financial Research* 12, 301–318.

Autor, David, David Dorn, and Gordon Hanson, 2019, When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men, *American Economic Review: Insights* 1, 161–178.

Barclay, Michael J., and Clifford G. Holderness, 1989, Private benefits from control of public corporations, *Journal of Financial Economics* 25, 371–395.

Barry, John W., John R. Graham, Dawoon Kim, and Hyunseob Kim, 2025, Executive Mobility in the United States, 1920-2023, *Working paper* .

Bartlett, Robert, and Frank Partnoy, 2020, The misuse of tobin's q, *Vanderbilt Law Review* 73, 353–424.

Bebchuk, Lucian A., and Kobi Kastiel, 2017, The untenable case for perpetual dual-class stock, *Virginia Law Review* 103, 585–631.

Bebchuk, Lucian A., Reinier Kraakman, and George Triantis, 2000, Stock pyramids, cross-ownership, and the dual class equity: The mechanisms and agency costs of separating control from cash-flow rights, *Concentrated Corporate Ownership (R. Morck, ed.)* 295–315.

Becht, Marco, Yuliya Kamisarenka, and Anete Pajuste, 2018, Loyalty shares with tenure voting - A Coasian bargain? Evidence from the Loi Florange experiment .

Bernstein, Shai, 2015, Does Going Public Affect Innovation?, *The Journal of Finance* 70, 13665–1403.

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

Brav, Alon, Wei Jiang, Song Ma, and Xuan Tian, 2018, How does hedge fund activism reshape corporate innovation?, *Journal of Financial Economics* 130, 237–264.

Burkart, Mike, Denis Gromb, and Fausto Panunzi, 1997, Large Shareholders, Monitoring, and the Value of the Firm, *Journal of Political Economy* 112, 693–728.

Carhart, Mark M., 1997, On persistence of mutual fund performance, *Journal of Finance* 52, 57–82.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, 2019, The Effect of Minimum Wages on Low-Wage Jobs, *The Quarterly Journal of Economcis* 134, 1405–1454.

Chen, Jiafeng, and Jonathan Roth, 2024, Logs with Zeros? Some Problems and Solutions*, *The Quarterly Journal of Economics* 139, 891–936.

Claessens, Stijn, Joseph P. H. Fan, and Larry H. P. Lang, 2002, Disentangling the incentive and entrenchment effects of large shareholdings, *Journal of Finance* 57, 2741–2771.

Cooper, Ilan, 2006, Asset pricing implications of nonconvex adjustment costs and irreversibility of investment, *Journal of Finance* 61, 139–170.

Cremers, Martijn, Beni Lauterbach, and Anete Pajuste, 2024, The life cycle of dual-class firm valuation, *Review of Corporate Finance Studies* 13, 459–493.

Cronqvist, Henrik, and Mattias Nilsson, 2003, Agency costs of controlling minority shareholders, *Journal of Financial and Quantitative Analysis* 38, 695–719.

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

Deaton, Angus, 1997, *The analysis of household surveys: a microeconometric approach to development policy* (World Bank Publications).

Doidge, Craig, 2004, U.S. cross-listings and the private benefits of control: Evidence from dual-class firms, *Journal of Financial Economics* 72, 519–553.

Dybvig, Philip H, and Mitch Warachka, 2015, Tobin's q does not measure firm performance: Theory, empirics, and alternatives, *Working paper* .

Elsby, Michael, Bart Hobijn, and Ayşegül Şahin, 2013, The Decline of the U.S. Labor Share, *Brookings Papers on Economic Activity* 44.2, 1–63.

Fabisik, Kornelia, Rüdiger Fahlenbrach, René M. Stulz, and Jérôme P. Taillard, 2021, Why are firms with more managerial ownership worth less?, *Journal of Financial Economics* 140, 699–725.

Faccio, Mara, and Larry H. P. Lang, 2002, The ultimate ownership of western European companies, *Journal of Financial Economics* 33, 365–395.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Field, Laura, and Michelle Lowry, 2022, Bucking the trend: Why do IPOs choose controversial governance structures and why do investors let them?, *Journal of Financial Economics* 146, 27–54.

Field, Laury, Michelle Lowry, and Anahit Mkrtchyan, 2013, Are busy boards detrimental?, *Journal of Finance* 109, 63–82.

Gompers, Paul A., Joy Ishii, and Andrew Metrick, 2010, Extreme governance: An analysis of dual-class firms in the United States, *Review of Financial Studies* 23, 1051–1088.

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

Graham, John, Mark Leary, and Hyunseob Kim, 2020, CEO-board dynamics, *Journal of Financial Economics* 137, 612–636.

Grossman, Sanford J., and Oliver D. Hart, 1986, The costs and benefits of ownership: A theory of lateral and vertical integrationl, *Journal of Political Economy* 94, 691–719.

Gutiérrez, Germán, and Thomas Philippon, 2018, Ownership, concentration, and investment, *AEA Papers and Proceedings* 108, 432–437.

Hart, Oliver D., and John Moore, 1990, Property rights and the nature of the firml, *Journal of Political Economy* 98, 1119–1158.

Hubbard, R. Glenn, 1998, Capital-market imperfections and investment, *Journal of Economic Literature* 36, 1093–1225.

Jackson, Robert J., 2018, Perpetual Dual-Class Stock: The Case Against Corporate Royalty, *U.S. Securities and Exchange Commission* .

Jarrell, Gregg A., and Annette B. Poulsen, 1988, Dual-class recapitalizations as antitakeover mechanisms: The recent evidence, *Journal of Financial Economics* 20, 129–152.

Johnson, William C., Jonathan M. Karpoff, and Sangho Yi, 2022, The lifecycle effects of corporate takeover defenses, *Review of Financial Studies* 35, 2879–2927.

Kaplan, Steven N., and Luigi Zingales, 1997, Do investment-cash flow sensitivities provide useful measures of financing constraints?, *Quarterly Journal of Economics* 112, 169–215.

Kim, Hyunseob, and Howard Kung, 2017, The Asset Redeployability Channel: How Uncertainty Affects Corporate Investment, *The Review of Financial Studies* 30, 245–280.

Kogan, Leonid, Dimitris Papanikolau, Amit Seru, and Noah Stoffman, 2017, Technological innovation, resource allocation, and growth, *Quarterly Journal of Economics* 132, 665–712.

Lease, Ronald C., John J. McConnell, and Wayne H. Mikkelson, 1983, The Market Value of Control in Publicly-Traded Corporations, *Journal of Financial Economics* 11, 439–471.

Lee, David S., 2009, Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects, *The Review of Economic Studies* 76, 1071–1102.

Levit, Doron, Nadya Malenko, and Ernst G. Maug, 2024, The Voting Premium, *European Corporate Governance Institute Finance Working Paper No. 720/2021* .

Loughran, Tim, and Jay Ritter, 2004, Why has IPO underpricing changed over time?, *Financial Management* 33, 5–37.

Ma, Song, 2025, Technological Obsolescence, *The Review of Financial Studies* , forthcoming.

Masulis, Ronald W., Cong Wang, and Fei Xie, 2009, Agency Problems at Dualclass Companies, *Journal of Finance* 64, 1697–1727.

Morck, Randall K., David A. Stangeland, and Bernard Yeung, 1998, Inherited Wealth, Corporate Control and Economic Growth: The Canadian Disease, *NBER* .

Nenova, Tatiana, 2003, The value of corporate voting rights and control: A cross-country analysis, *Journal of Financial Economics* 68, 325–351.

Partch, Megan M., 1987, The creation of a class of limited voting common stock and shareholders wealth, *Journal of Financial Economics* 18, 313–339.

Pástor, L'uboš, and Pietro Veronesi, 2003, Stock valuation and learning about profitability, *Journal of Finance* 58, 1749–1789.

Peters, Ryan H., and Lucian A. Taylor, 2017, Intangible capital and the investment-q relation, *Journal of Financial Economics* 123, 251–272.

Philippon, Thomas, 2006, Corporate governance over the business cycle, *Journal of Economic Dynamics and Control* 30, 2117–2141.

Ruback, Richard, 1988, Coercive dual-class exchange offers, *Journal of Financial Economics* 20, 153–173.

Rydqvist, Kristian, 1987, *The Pricing of Shares with Different Voting Power and the Theory of Oceanic Games*, Ph.D. thesis, Stockholm School of Economics, Stockholm.

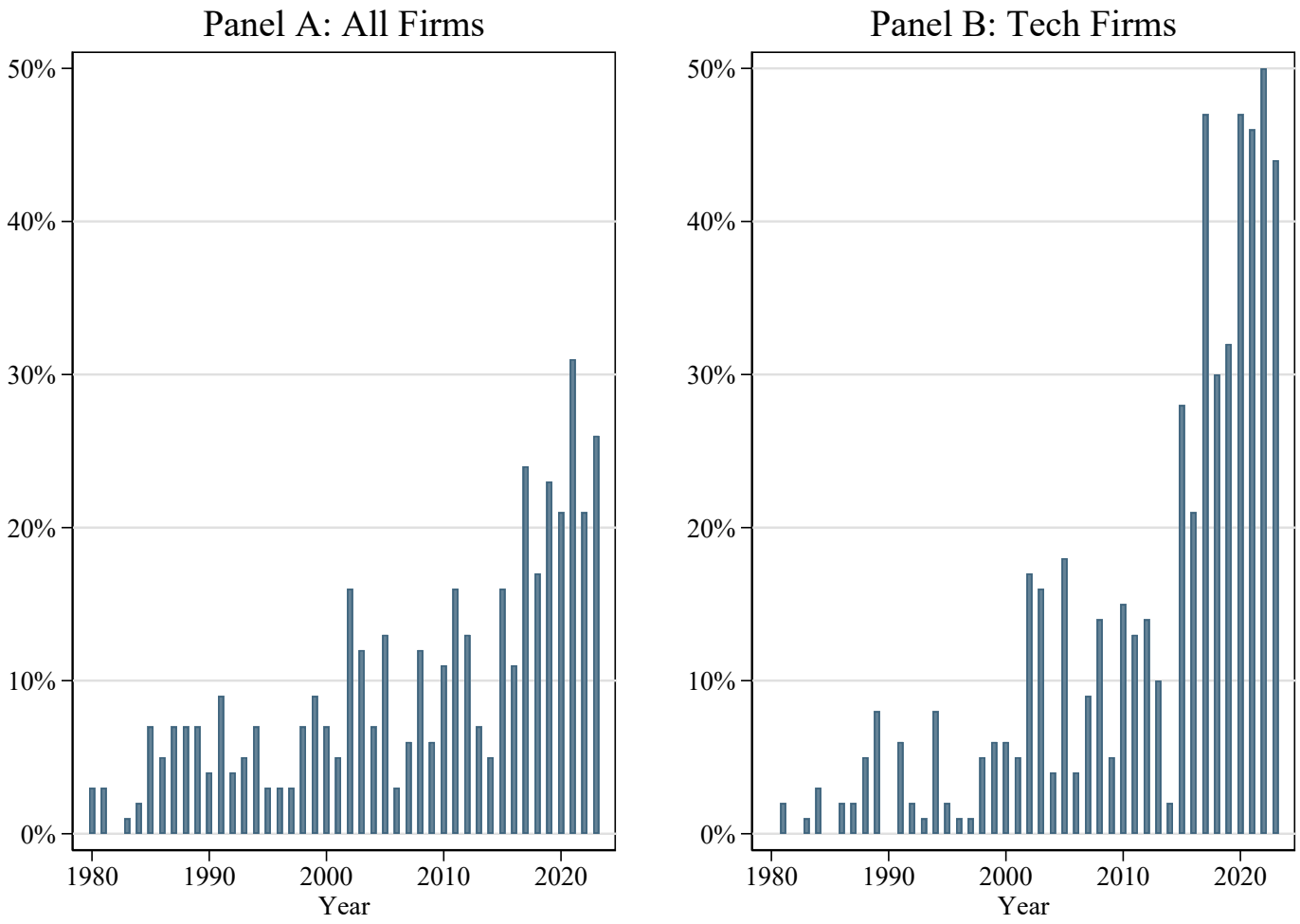
Sun, Liyang, and Sarah Abraham, 2021, Estimating dynamic treatment effects in event studies with heterogeneous treatment effects, *Journal of Econometrics* 225, 175–199.

Williamson, Oliver E., 1975, *Markets and Hierarchies: Analysis and Antitrust Implications* (The Free Press, New York).

Zhang, Lu, 2005, The Value Premium, *Journal of Finance* 60, 67–103.

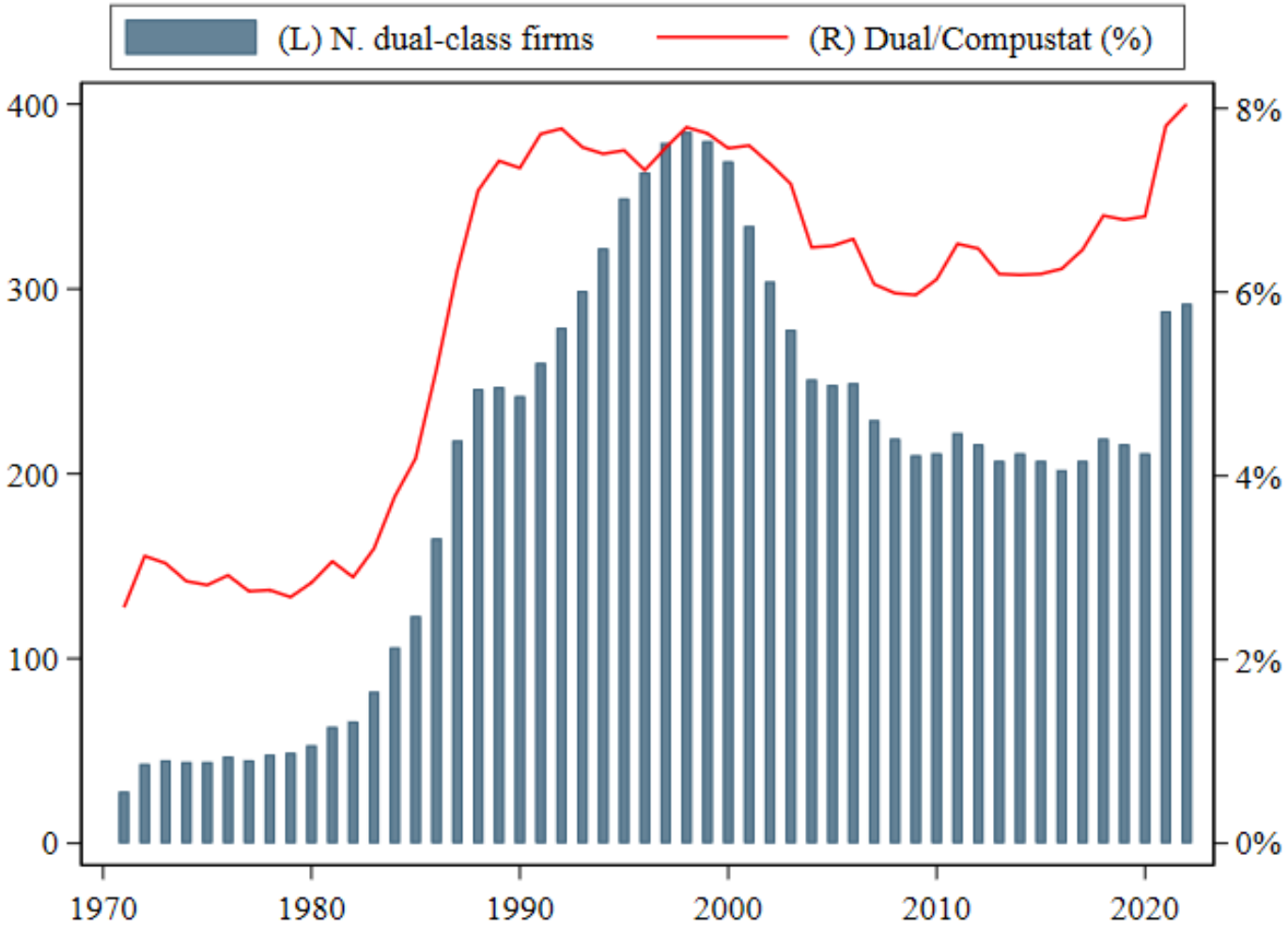
Zingales, Luigi, 1995, What Determines the Value of Corporate Votes?, *Quarterly Journal of Economics* 110, 1047–1073.

Figure 1.
Fraction of Dual-Class IPOs among the Universe of IPOs, 1980–2023



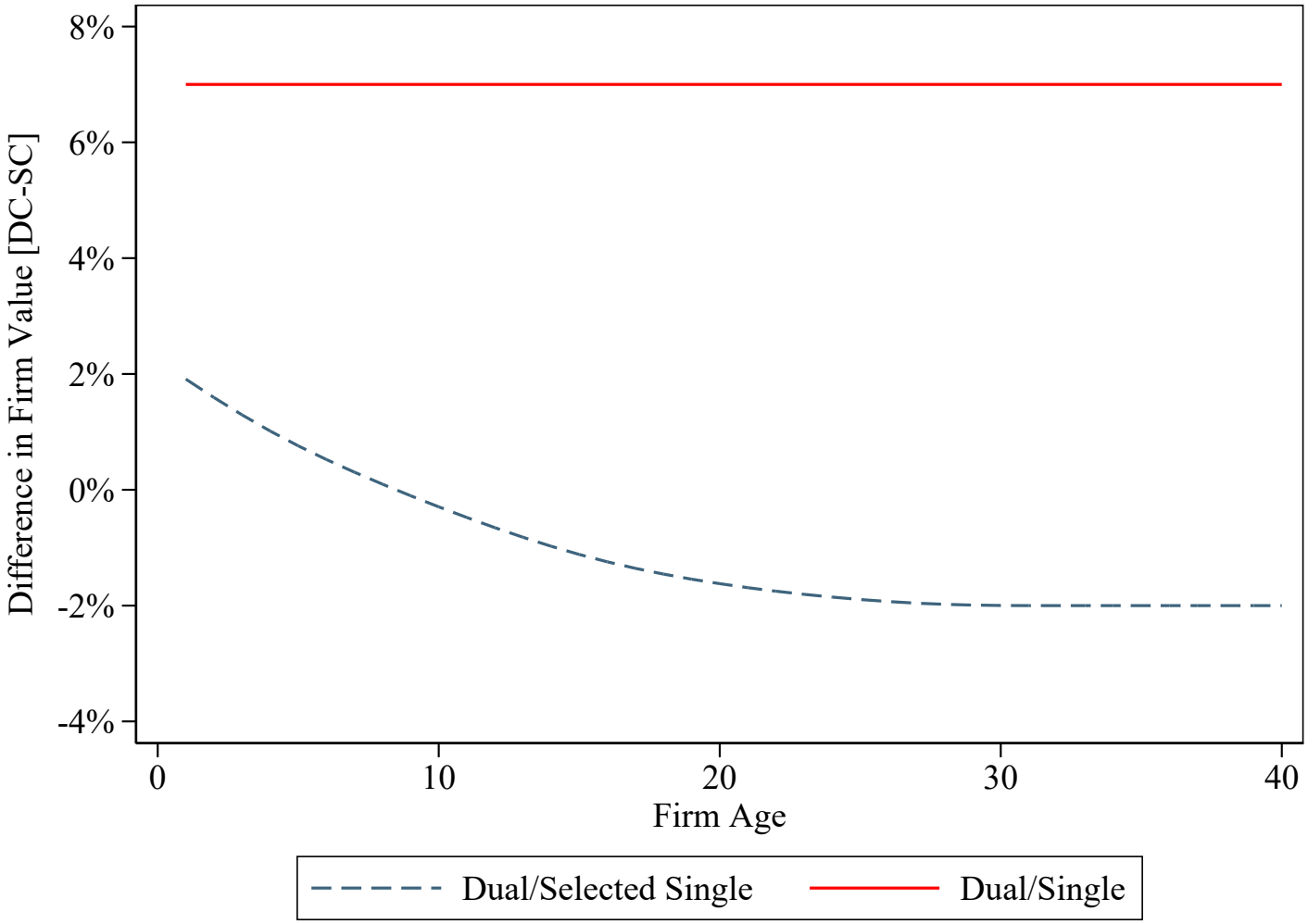
Notes: This figure shows the fraction of dual-class IPOs relative to the universe of all IPOs (Panel A) and IPOs in the technology sectors (Panel B) as definded in Loughran and Ritter (2004), plus SIC codes 3559, 3576, and 7389, from 1980 through 2023. The sample contains all offerings of common shares from SDC Platinum with an offer price of at least $5.00, excluding REITs, ADRs, closed-end-funds, units (including the special purpose acquisition companies, SPACs), and companies not listed on CRSP within 7 days of the IPO. The data for 2022 and 2023 are from Jay Ritter's website.

Figure 2.
Number and Fraction of Dual-Class Firms among Compustat Universe



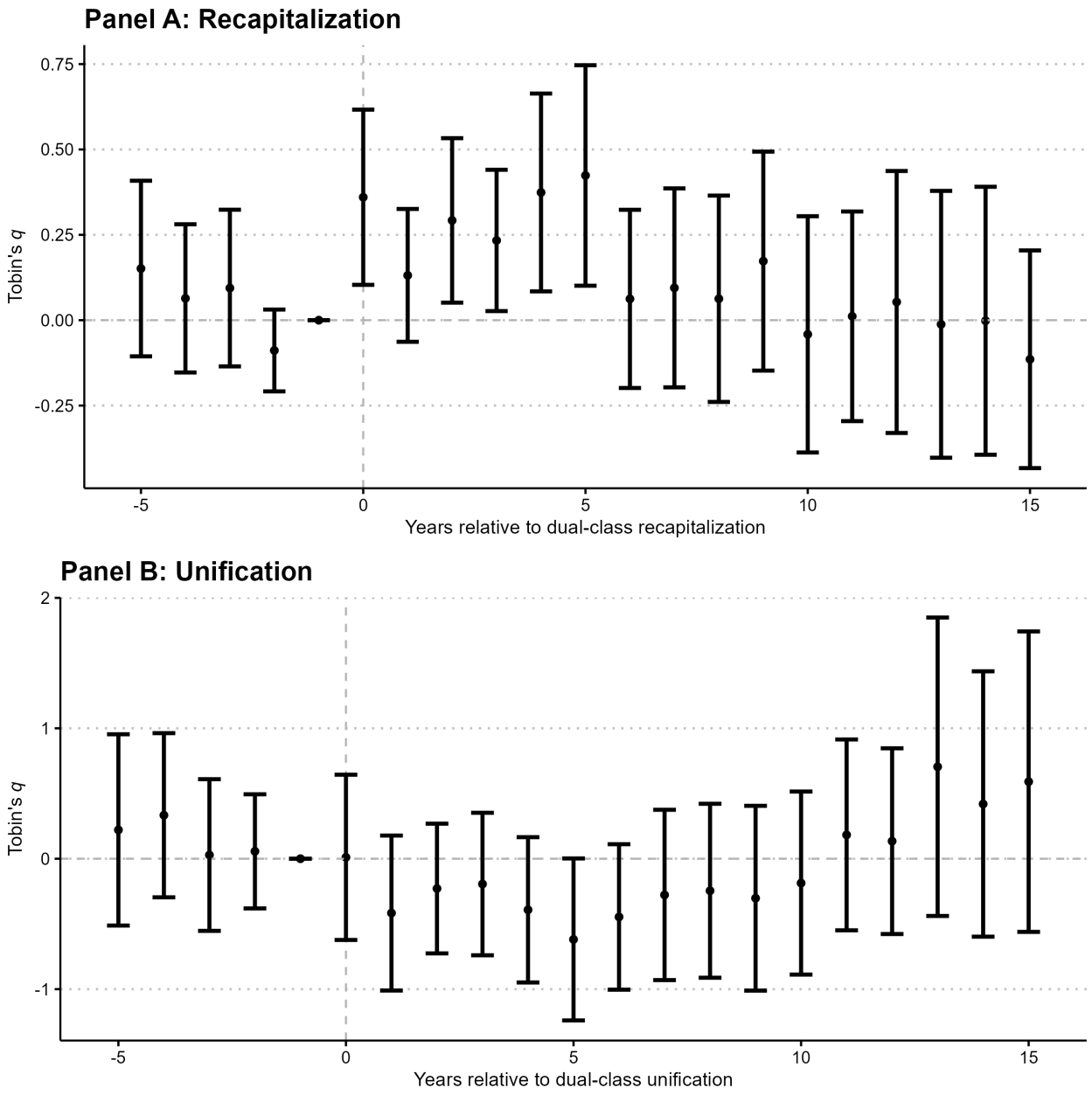
Notes: This figure shows the number (blue bars, left y-axis) and percent (red line, right y-axis) of dual-class firms relative to Compustat firms that satisfy the sample selection criteria in Section 2.3 from 1971 through 2022. Specifically, we exclude firm-years in the financial (SIC 6000-6999), utilities (SIC 4900-4999), and unclassified (SIC 9900-9999) industries, as well as firm-years with book assets less than $20 million in 2000 constant dollars (adjusted using the CPI).

Figure 3.
Theoretical Thought Experiment: Dynamics of Stock Prices of Dual-class vs. Single-Class Firms over Firm Age



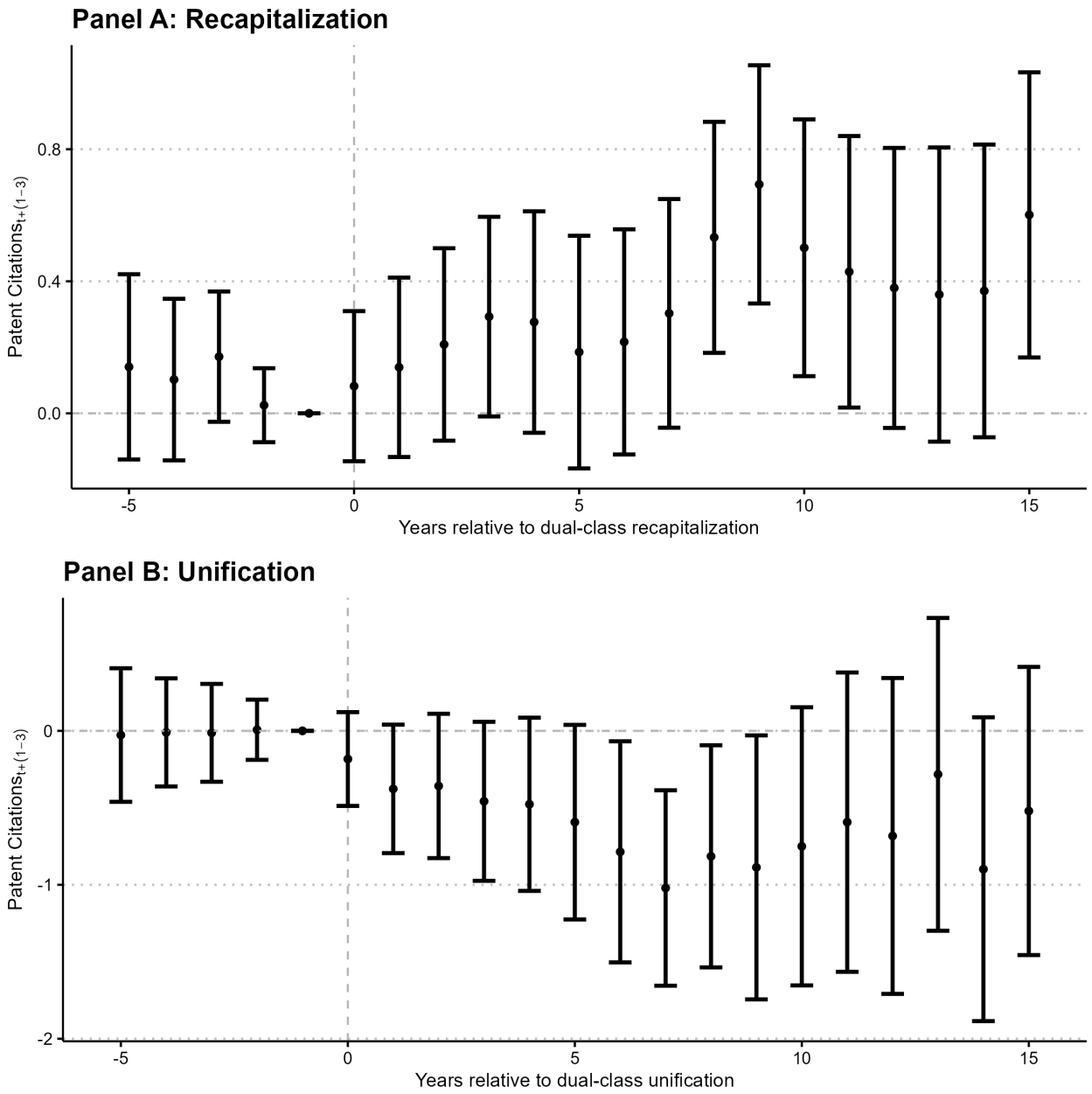
Notes: This figure presents dynamics of relative stock prices of two otherwise identical hypothetical firms but one with dual-class shares and the other with single-class shares as chosen by their respective founders at the IPO (age = 0). The blue dashed line plots the observed price premium of the dual-class relative to single-class firms. Thus, the dynamic pattern in the blue dashed line is due both to the selection and treatment effects of dual-class shares. The red solid line plots the hypothetical price premium by assuming that the single-class firm unexpectedly ends up switching to dual-class shares right after the IPO. Thus, the dynamic pattern in the red solid line is only due to the selection effect of dual-class shares but not to the treatment effect.

Figure 4.
Event Study of Dual-Class Recapitalization and Stock Unification: Tobin's *q*



Notes: This figure plots the dynamic stacked difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on Tobin's *q* from Eq. (A1). The markers on the lines show point estimates. The 90% confidence intervals in vertical lines are based on standard errors adjusted for sample clustering at the firm level. The event requires that both the treated and control groups of firms exist during the [t-2, t+2] window.

Figure 5.
Event Study of Dual-Class Recapitalization and Stock Unification: Innovative Output



Notes: This figure plots the dynamic stacked difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on innovative output from a version of Eq. (A1) in which the dependent variable is the inverse hyperbolic sine transformed number of citations the firm's patents receive one, two, or three years ahead. The markers on the lines show point estimates. The 90% confidence intervals in vertical lines are based on standard errors adjusted for sample clustering at the firm level. The event study requires that both the treated and control groups of firms exist during the [t-2, t+2] window.

Table 1. Descriptive Statistics on Dual-Class and Single-Class Firms

This table presents descriptive statistics on financial variables for dual-class and single-class firm-years merged with CRSP and Compustat from 1971 through 2022. "Book assets" is book value of assets; "Log book assets" is the natural log of book assets; "Age" is the number of years since an IPO (proxied by the first appearance in Compustat with a valid stock price); "Tobins *q*" is the ratio of the market value of capital to the book value of capital; "Sales growth" is the first difference in the natural log of sales; "ROA" is operating income before depreciation divided by lagged book assets; "Market leverage" is total debt divided by the sum of total debt and market equity; "Capex/assets" is capital expenditures divided by lagged book assets; "R&D" is research and development expenses divided by lagged book assets; "Tangibility" is net PP&E divided by book assets; "Payout ratio" is total payout including dividends and repurchases divided by market equity; and "Employment growth" is the first difference in the natural log of employment. The column "Dual – Single" shows differences between the means for dual-class and single-class firms. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively, based on standard errors adjusted for clustering at the firm level.

	Dual-class		Single-class		Dual – Single
Variable	Mean	Std Dev	Mean	Std Dev	Diff.
Book assets	3966.805	18872.27	3533.26	17343.25	433.5
Log book assets	6.344	1.791	5.756	2.016	0.588***
Age	15.053	12.985	13.257	11.85	1.796***
Tobin's q	2.351	2.808	2.125	2.187	0.226***
Sales growth	0.190	0.592	0.220	0.705	-0.029***
ROA	0.128	0.165	0.111	0.222	0.017***
Market leverage	0.293	0.261	0.265	0.249	0.028***
Capex/assets	0.072	0.092	0.083	0.106	-0.011***
R&D	0.024	0.065	0.042	0.091	-0.018***
Tangibility	0.290	0.213	0.313	0.242	-0.023***
Payout ratio	0.028	0.042	0.026	0.042	0.002*
Observations	9950		150752		

Table 2. Market Reactions to Dual-Class Recapitalizations and Stock Unifications Conditional on Maturity

This table examines the effects of dual-class recapitalizations (columns (1) and (2)) and stock unifications (columns (3) and (4)) on stock returns conditional on firm maturity. The dependent variable is the cumulative abnormal return from three days before to three days after the announcement of a dual-class recapitalization and unification. "$\mathbb{1}[\text{age} \geq 12]$" is an indicator variable equal to one if age is larger than or equal to 12 years, the sample median for dual-class firms, and zero otherwise. Standard errors are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable:	Cumulative abnormal return			
Event:	Dual-class recapitalization		Dual-class stock unification	
$\mathbb{1}[\text{age} \geq 12]$	-0.034*	-0.046 *	0.033*	0.050*
	(0.019)	(0.030)	(0.019)	(0.026)
Constant	0.026*	0.035*	0.003	-0.006
	(0.017)	(0.021)	(0.014)	(0.017)
Year fixed effects	N	Y	N	Y
R^2	0.035	0.178	0.046	0.355
Observations	88	88	62	62

Table 3. Effects of Recapitalization and Unification, by Specificity of Investments

This table shows the impacts of dual-class recapitalizations and unifications on firm value and innovation, as measured by Tobin's q and patent citations, split by the specificity of firm investments. Investment specificity is a combined measure of capital and labor redeployability. The capital redeployability measure is at the firm level weighted-averaged across industry segments, and we fill in missing firm-level measure using industry measures by BEA industry code, cross-walked to three-digit SIC codes. The labor mobility measure is at the 1990 Census industry code level, which we also crosswalk to three-digit SIC codes using crosswalks from Autor et al. (2019). Panel A shows recapitalizations and Panel B shows unifications, while Columns (1)-(2) show the results for Tobin's q and Columns (3)-(4) show the results for the inverse hyperbolic sine of patent citations. Columns (1) and (3) report each result for firms with above-median specificity, and Columns (2) and (4) report each result for firms with below-median specificity. Standard errors adjusted for clustering at the firm level are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Recaps

	Tobin's q		IHS(Cites)	
	(1)	(2)	(3)	(4)
Specificity:	High	Low	High	Low
$\mathbb{1}[-5 \leq t \leq -2]$	0.059	-0.159	0.120	0.162
	(0.154)	(0.161)	(0.201)	(0.155)
$\mathbb{1}[t = 5]$	0.634*	0.269	0.070	0.129
	(0.326)	(0.270)	(0.331)	(0.277)
$\mathbb{1}[t = 10]$	0.308	-0.095	0.583	0.523*
	(0.440)	(0.246)	(0.355)	(0.292)
$\mathbb{1}[t = 15]$	0.137	-0.353*	0.821**	0.538*
	(0.420)	(0.183)	(0.413)	(0.317)

Panel B: Unifs

	Tobin's q		IHS(Cites)	
	(1)	(2)	(3)	(4)
Specificity:	High	Low	High	Low
$\mathbb{1}[-5 \leq t \leq -2]$	-0.136	-0.032	-0.142	0.095
	(0.369)	(0.441)	(0.260)	(0.259)
$\mathbb{1}[t = 5]$	-1.227*	-0.225	-1.088**	-0.279
	(0.726)	(0.499)	(0.469)	(0.626)
$\mathbb{1}[t = 10]$	-0.586	-0.057	-1.540*	-0.029
	(0.707)	(0.805)	(0.814)	(0.734)
$\mathbb{1}[t = 15]$	0.872	-0.636	-1.030	-0.481
	(0.934)	(1.089)	(0.732)	(1.103)

Table 4. Investment-q Sensitivity for Dual-Class vs. Single-Class Firms over Maturity

This table examines the effects of dual-class shares and firm maturity on sensitivities of corporate investment to investment opportunities, as proxied by Tobin's q, using capital expenditure scaled by lagged assets as the dependent variable. "Dual" is an indicator variable equal to one if a firm-year has multiple classes of shares with differing voting rights, and zero otherwise. The full sample includes firm-years with sales growth in the first quartile of the distribution over 1971–2022. The "Young" firm sample includes firm-years with ages smaller than 12 years, the sample median for dual-class firms, whereas the "Mature" firm sample includes firm-years with ages larger than or equal to 12 years. "q" is Tobin's q as defined in Table 1. "Cash flow" is net income plus depreciation and amortization divided by lagged assets. All regressions includes "Cash flow" interacted with "Dual" as controls. Standard errors adjusted for clustering at the firm level are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
Dependent Variable:	Capex / assets		
Sample:	Full	Young	Mature
q	0.004***	0.004***	0.004***
	(0.000)	(0.001)	(0.000)
q × Dual	-0.001	0.001	-0.002**
	(0.001)	(0.002)	(0.001)
Cash flow controls	Y	Y	Y
Firm × Dual fixed effects	Y	Y	Y
Year × Dual fixed effects	Y	Y	Y
R^2	0.611	0.700	0.619
Observations	41,463	22,309	19,154
q × Dual × (Mature - Young)		-0.004	
t-statistic		1.87	

Table 5. Four-Factor Model Regressions for Dual-Class vs. Matched Single-Class Firm Stocks over Maturity

This table presents estimates of the asset-pricing factor loadings associated with dual-class shares conditional on firm maturity over the 1976 through 2022 period (564 months). It reports coefficient estimates and t-statistics from value- (Panel A) and equal-weighted (Panel B) calendar-time monthly zero-cost portfolio regressions with the sample excluding firms in the financial (SIC 6000-6999), utilities (SIC 4900-4999), and unclassified (SIC 9900-9999) industries. The portfolio longs stocks of dual-class firms and shorts stocks of matched single-class firms on Tobins q in the same Fama-French 48 industry. The stocks are allocated to two age groups ("Young" and "Mature") at 12 years, the median age of dual-class firms. "Alpha" is the estimate of the regression intercept. "MKT" is the estimate of the factor loading on the market excess return (the Fama-French RMRF). "SMB," "HML," and "UMD" are the estimates of the factor loadings on the Fama-French size and book-to-market factors, and the Carhart momentum factor, respectively. "R^2" is the R-squared from the regressions. "Obs." is the number of monthly portfolio observations.

Panel A: Value-weighted portfolio				Panel B: Equal-weighted portfolio			
	(1)	(2)	(3)		(1)	(2)	(3)
Sample:	Full	Young	Mature		Full	Young	Mature
Alpha	-0.038	0.308**	-0.100	Alpha	0.216***	0.300***	0.075
	(0.083)	(0.145)	(0.101)		(0.065)	(0.093)	(0.080)
MKT	-0.040**	-0.088***	0.000	MKT	-0.025*	-0.007	-0.035*
	(0.019)	(0.033)	(0.0229)		(0.015)	(0.021)	(0.018)
SMB	-0.236***	-0.344***	-0.120***	SMB	0.155***	0.069**	0.229***
	(0.028)	(0.050)	(0.034)		(0.022)	(0.032)	(0.027)
HML	0.001	-0.051	0.050	HML	0.081***	0.107***	0.093***
	(0.028)	(0.049)	(0.034)		(0.022)	(0.031)	(0.0268)
UMD	0.046**	0.920***	-0.013	UMD	-0.020	-0.029	-0.008
	(0.019)	(0.0338)	(0.023)		(0.015)	(0.022)	(0.019)
R^2	0.146	0.122	0.033	R^2	0.096	0.033	0.120
Obs.	564	564	564	Obs.	564	564	564

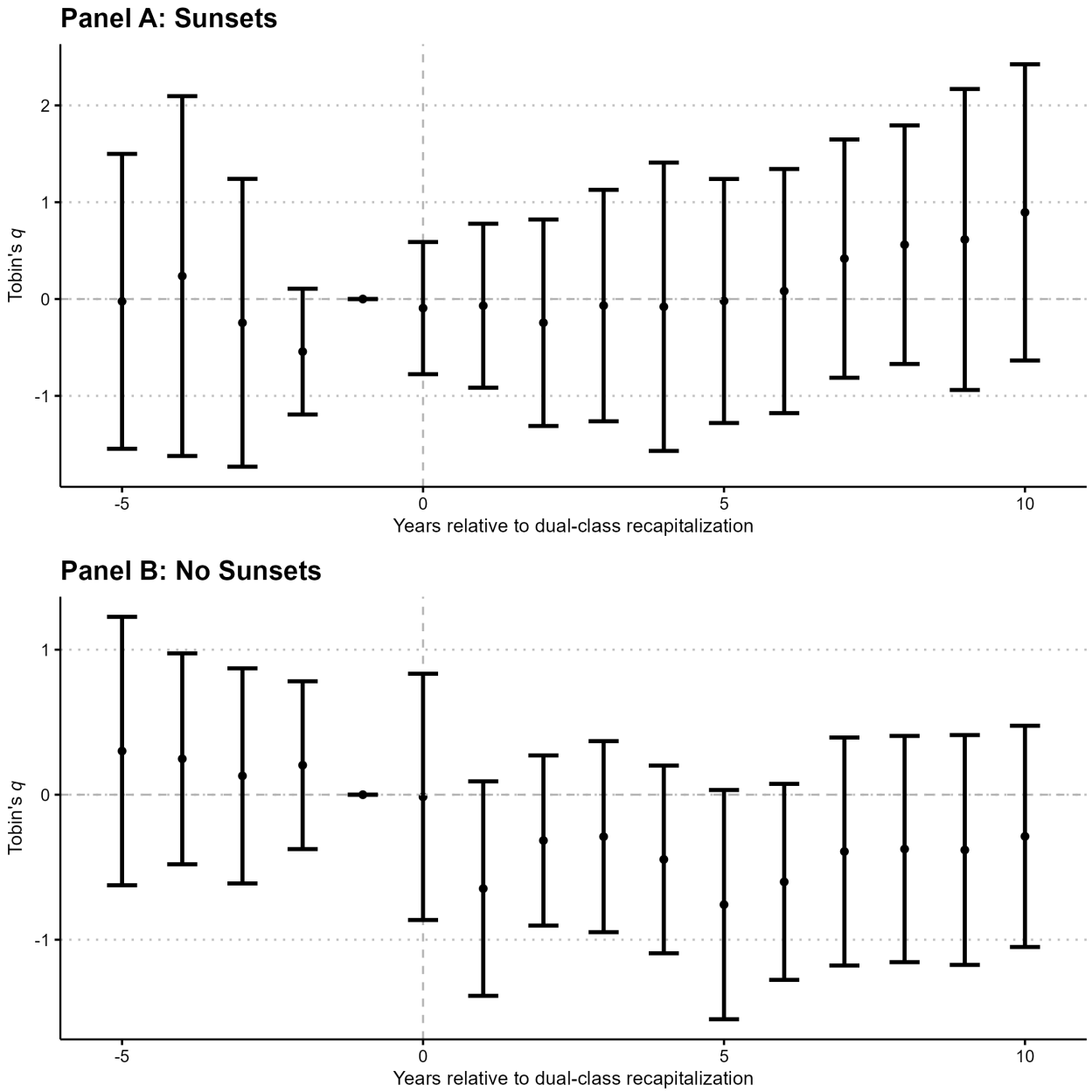
Table 6. Voting Premium for Dual-Class Firms over Firm Maturity

This table examines how the voting premium changes with firm maturity using a sample of dual-class firms for which information on stock prices and trading volume is available from CRSP for both inferior- and superior-voting shares and ownership structure information is available from 1980 through 2022. The dependent variable is the voting premium, computed as $(P_A - P_B)/(P_B - r\,P_A)$, where P_A (P_B) is the price of the superior- (inferior-) voting shares and r is the relative number of votes between inferior- and superior-voting shares. Firm age is measured as years since IPO and grouped into quartiles based on the distribution of firm ages in the sample. Firms in the first age quartile serve as the omitted group. "Log market equity" is the natural log of market equity scaled by the CPI; "Shapley value" is the Shapley value of the ownership structure; "Volume (sup. to inf.)" is the ratio of trading volume between superior- and inferior-voting shares. In column (2), we include year fixed effects using Deaton's normalization. In column (3), we use an alternative restriction that normalizes one additional year fixed effect to zero. Standard errors clustered at the firm level are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
Dependent Variable:	Voting premium		
$\mathbb{1}$[age in quartile 2]	0.031	0.006	0.008
	(0.021)	(0.014)	(0.015)
$\mathbb{1}$[age in quartile 3]	0.046**	0.045*	0.051*
	(0.023)	(0.023)	(0.027)
$\mathbb{1}$[age in quartile 4]	0.062***	0.094**	0.104**
	(0.023)	(0.037)	(0.048)
Log market equity	-0.011*	-0.030***	-0.029***
	(0.006)	(0.008)	(0.007)
Shapley value	0.055***	0.012	0.009
	(0.020)	(0.019)	(0.021)
Volume (sup./inf.)	0.031**	0.013	0.013
	(0.014)	(0.010)	(0.010)
Firm fixed effects	N	Y	Y
Year fixed effects	Y	Y	Y
R^2	0.097	0.538	0.538
Observations	1,188	1,188	1,188

4.

Figure A1.
Event Study of Dual-Class Stock Unification by Sunset Provisions: Tobin's *q*



Notes: This figure plots the dynamic stacked difference-in-differences estimates of the effect of stock unification on Tobin's *q* for firms with sunset provisions (Panel A) and without sunset provisions (Panel B). The event study includes 36 unification events for sunset firms and 149 unification events for firms without sunsets, in which both the treated and control groups of firms exist during the [t-2, t+2] window. The markers on the lines show point estimates. The 95% confidence intervals in vertical lines are based on standard errors adjusted for sample clustering at the firm level.

Table A1. Post-IPO Dynamics of Firm Valuation and Innovative Output

This table presents the results of estimating dynamics of Tobin's q and innovative output for firms that IPO with dual-class shares relative to firms that IPO with single-class shares. We estimate the following equation:

$$y_{ijt} = \alpha_i + \alpha_{jth} + \sum_{2}^{4} \lambda_n \mathbb{1}[\text{age in quartile}_n]_{it} + \sum_{2}^{4} Dual_i \times \mathbb{1}[\text{age in quartile}_n]_{it} + \gamma' X_{it} + \epsilon_{ijth},$$

where $\mathbb{1}[\text{age in quartile}_n]_{it}$ is an indicator variable equal to one if firm i's age at time t is in quartile n; y_{ijt} is either Tobin's q or IHS(cites) for firm i in three-digit SIC industry j in year t; α_i and α_{jth} represent firm i and industry j by year t by cohort h fixed effects; $Dual_{ih}$ is an indicator variable equal to one if firm i IPOs as a dual-class firm, and zero if as a single-class firm; X_{it} is a vector of control variables; and ϵ_{ijth} represents random errors clustered at the firm level. Firms that switch from one type of shares to the other are dropped from our sample upon switching. Column (1) reports results for q and Column (2) reports results for IHS(patent citations). Standard errors adjusted for clustering at the firm level are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1) Tobin's q	(2) IHS(cites)
$\mathbb{1}$[age in quartile 2]	-0.530***	-0.036
	(0.159)	(0.096)
$\mathbb{1}$[age in quartile 3]	-0.700**	0.468*
	(0.278)	(0.255)
$\mathbb{1}$[age in quartile 4]	-0.779*	1.282***
	(0.430)	(0.308)
Observations	100,021	119,263
Firm FE	Y	Y
SIC3 ×Year ×IPO Cohort FE	Y	Y

Table A2. Robustness of q-Sensitivity of Capital Investment to Different Sales Growth Percentile cutoffs

This table replicates columns (2)-(3) of Table 4, which shows that mature dual-class firms have relatively weaker investment sensitivity to Tobin's q, using different sales growth cutoffs. In Table 4, our sample consists of firms below the 25th percentile of sales growth to proxy for firms experiencing weak demand. The results below restrict the sample to the 15th, 20th, and 30th percentile of sales growth. Standard errors adjusted for clustering at the firm level are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Capex / assets					
Sample:	Young	Mature	Young	Mature	Young	Mature
q	0.003***	0.003***	0.003***	0.003***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
$q \times$ dual	0.005*	-0.003**	0.003	-0.003***	0.000	-0.002**
	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
SG percentile cutoff	15	15	20	20	30	30
Cash flow controls	Y	Y	Y	Y	Y	Y
Firm × Dual FE	Y	Y	Y	Y	Y	Y
Year × Dual FE	Y	Y	Y	Y	Y	Y
SIC3 × Dual × Year FE	N	N	N	N	N	N
R^2	0.756	0.673	0.726	0.646	0.684	0.600
Observations	14,823	10,054	18,656	14,514	25,720	24,037
$q \times$ Dual × (Mature - Young)	-0.007		-0.006		-0.003	
t-statistic	2.45		2.44		1.44	

Table A3. Voting Premium for Dual-Class Firms over Firm Maturity Excluding Market Equity

This table replicates Table 6 with log market equity excluded as a control. Standard errors adjusted for clustering at the firm level are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
Dependent Variable:	Voting premium		
𝟙[age in quartile 2]	0.037*	0.004	0.015
	(0.020)	(0.013)	(0.013)
𝟙[age in quartile 3]	0.056**	0.033	0.059**
	(0.022)	(0.021)	(0.026)
𝟙[age in quartile 4]	0.066***	0.054*	0.111**
	(0.022)	(0.032)	(0.049)
Shapley value	0.048**	0.014	0.003
	(0.020)	(0.021)	(0.022)
Volume (sup./inf.)	0.032**	0.017	0.015
	(0.014)	(0.012)	(0.011)
Firm fixed effects	N	Y	Y
Year fixed effects	Y	Y	Y
R^2	0.084	0.515	0.522
Observations	1,188	1,188	1,188

Appendix A Theory Appendix

This appendix provides a detailed analysis of the theoretical model presented in Section 3. Under governance structure $n \geq 1$, the founder can be fired only after period n. Thus, $n = 1$ corresponds to single-class shares, and $n = \infty$ to dual-class shares. Observe that at time t, the share price is

$$p(n,t) = E\left[\sum_{s=t+1}^{n-1} \frac{y_s(1-b_s)}{(1+r)^{s-t}} + \sum_{s=\max\{n,t+1\}}^{\infty} \frac{y_s}{(1+r)^{s-t}}\right] \tag{A.1}$$

and the founder's expected utility is

$$u_t(n) = E\left[\sum_{s=t+1}^{n-1} \frac{\alpha\Lambda y_s(e)(1-b_s)+\gamma b_s y_s(e)}{(1+r)^{s-t}} + \sum_{s=\max\{n,t+1\}}^{\infty} \frac{\alpha\Lambda y_s(e)}{(1+r)^{s-t}}\right] + \alpha(1-\Lambda)p(n,0) - \frac{1}{2}\kappa e^2. \tag{A.2}$$

Consider governance structure n. Once the founder exerts effort, her impact on the shareholder value is only through the consumption of private benefits. Therefore, shareholders fire the founder as soon as they can, at period $t = n$ (the founder has no incentives to resign voluntarily). Thus, $b_t = 0$ for all $t \geq n$. Since effort is not observed by outside shareholders, the proceeds from the IPO do not depend on the actual choice of effort, but only on the expectations of shareholders. Therefore, the founder chooses effort to maximize the value of the $\alpha\Lambda$ shares she retains in the firm post-IPO. Let us define

$$\eta \equiv \frac{1+g}{1+\tau} - 1. \tag{A.3}$$

The next result characterizes the equilibrium under governance structure n. All proofs are collected in Appendix B.

Lemma 1. *Under governance structure n, the founder exerts effort*

$$e^*(n) = e^*(\infty) - \left(\gamma - \alpha\Lambda\right)\frac{\phi}{\kappa}\psi(n-1), \tag{A.4}$$

where

$$e^*(\infty) = \frac{1}{\kappa}\frac{\alpha\Lambda}{r-g} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi}{\kappa}\frac{g-\eta}{r-\eta} \tag{A.5}$$

and

$$\psi(n) = \frac{1}{r-g}\left(\frac{1+g}{1+r}\right)^n - \frac{1}{r-\eta}\left(\frac{1+\eta}{1+r}\right)^n. \tag{A.6}$$

The share price at time t is

$$p^*(n,t) = e^*(n)\begin{cases} \frac{(1+g)^t}{r-g} - \phi\left[\frac{(1+g)^t}{r-g}\left(1-\left(\frac{1+g}{1+r}\right)^{n-1-t}\right) - \frac{(1+\eta)^t}{r-\eta}\left(1-\left(\frac{1+\eta}{1+r}\right)^{n-1-t}\right)\right] & \text{if } t+1 \leq n-1 \\ \frac{(1+g)^t}{r-g} & \text{if } t+1 \geq n \end{cases} \tag{A.7}$$

Assuming $\tau > 0$ ensures $r > g > \eta$. Therefore, $e^*(n)$ increases in n, that is, the founder exerts more effort when she has more job security. In particular, let

$$e^*_{\text{single}} \equiv e^*(1) = \frac{1}{\kappa}\frac{\alpha\Lambda}{r-g} \tag{A.8}$$

and

$$e^*_{\text{dual}} \equiv e^*(\infty) = e^*_{\text{single}} + \frac{\gamma - \alpha\Lambda}{r-g}\frac{\phi}{\kappa}\frac{g-\eta}{r-\eta}. \tag{A.9}$$

Thus, $e^*_{\text{dual}} > e^*_{\text{single}}$, that is, the founder exerts more effort under dual-class shares.

Next, we compare the performances of two governance structures over time: single-class and dual-class shares. As a special case, Lemma 1 implies

$$p^*_{\text{single}}(t) \equiv p^*(1,t) = \frac{e^*_{\text{single}}}{r-g}\left(1+g\right)^t \tag{A.10}$$

and

$$p^*_{\text{dual}}(t) \equiv p^*(\infty,t) = \left(1-\phi\right)\frac{e^*_{\text{dual}}}{r-g}\left(1+g\right)^t + \phi\frac{e^*_{\text{dual}}}{r-\eta}\left(1+\eta\right)^t. \tag{A.11}$$

To ease the exposition, define

$$\Gamma \equiv \phi\frac{g-\eta}{r-\eta}, \tag{A.12}$$

and notice $\Gamma \in (0,1)$. The next result shows that the share price of dual-class firms is higher than that of single-class firms during the early years after the IPO.

Lemma 2. *For a given firm, $p^*_{dual}(0) > p^*_{single}(0)$ if and only if*

$$\frac{\gamma}{\alpha\Lambda} > \frac{2-\Gamma}{1-\Gamma}. \tag{A.13}$$

If in addition $\frac{\gamma}{\alpha\Lambda} \geq 1 + \frac{\phi}{1-\phi}\frac{1}{\Gamma}$, then $p^*_{dual}(t) > p^*_{single}(t)$ for all t. Otherwise, if $\frac{\gamma}{\alpha\Lambda} < 1 + \frac{\phi}{1-\phi}\frac{1}{\Gamma}$ then

$$p^*_{dual}(t) > p^*_{single}(t) \Leftrightarrow t < T^* \tag{A.14}$$

where

$$T^* \equiv \frac{\ln\left[\phi\frac{r-g}{r-\eta}\frac{e^*_{dual}}{e^*_{single} - e^*_{dual}(1-\phi)}\right]}{\ln[1+\tau]}. \tag{A.15}$$

Proposition 1 follows directly from Lemma 2 by defining $\overline{\phi} \equiv \frac{r-\eta}{g-\eta} - \frac{\alpha\Lambda}{\gamma-\alpha\Lambda}\frac{r-\eta}{g-\eta}$ (notice $\frac{2-\Gamma}{1-\Gamma} < \frac{\gamma}{\alpha\Lambda} \Leftrightarrow \phi < \overline{\phi}$) and $\underline{\phi} \equiv 1 - \frac{\alpha\Lambda}{\gamma-\alpha\Lambda}\frac{r-\eta}{g-\eta}$ (notice $\frac{\gamma}{\alpha\Lambda} < 1 + \frac{\phi}{1-\phi}\frac{1}{\Gamma} \Leftrightarrow \phi > \underline{\phi}$). Also observe that $r > g \Rightarrow \underline{\phi} < \overline{\phi}$, $\underline{\phi} < 1$, and $\overline{\phi} > 0$.

Next, we analyze the founder's optimal choice of governance structure.

Lemma 3. *Under governance structure $1 \leq n \leq \infty$, the share price at the IPO is*

$$p^*(n,0) = e^*(n)\left[\frac{1}{r-g} - \phi\left(\frac{g-\eta}{(r-g)(r-\eta)} - \psi(n-1)\right)\right], \tag{A.16}$$

and the founder's expected utility at the IPO is

$$u^*(n) = \alpha(1-\Lambda)p^*(n,0) + \frac{1}{2}\kappa e^*(n)^2. \tag{A.17}$$

The founder prefers dual-class shares over single-class shares (i.e., $u^*(\infty) > u^*(1)$) if and only if

$$\frac{\gamma}{\alpha\Lambda} > H(\Gamma) \tag{A.18}$$

where

$$H(\Gamma)\equiv-\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}+\sqrt{\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda}}\in\left(1,\frac{2-\Gamma}{1-\Gamma}\right). \tag{A.19}$$

The governance structure that maximizes $p^*(n,0)$, which we denote by n^*_{sh}, satisfies:

(*i*) *If $\frac{\gamma}{\alpha\Lambda}\leq 2$ then $n^*_{sh}=1$, i.e., single-class shares.*

(*ii*) *If $\phi\leq\frac{1}{2}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}\frac{r-\eta}{g-\eta}$ then $n^*_{sh}=\infty$, i.e., dual-class shares.*

(*iii*) *In all other cases, $2\leq n^*_{sh}<\infty$, i.e., dual-class shares with a sunset provision, where n^*_{sh} uniquely solves*

$$\psi(n^*)\leq\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g}-\psi(n^*-1)\leq\psi(n^*-2). \tag{A.20}$$

The governance structure that maximizes the founder's utility at the time of the IPO satisfies $n^*_{founder}\geq n^*_{sh}$.

Proposition 2 follows from Lemma 3 and the following observations: (i) $H(\Gamma)$ is an increasing function of Γ; (ii) $H(\Gamma)<\frac{2-\Gamma}{1-\Gamma}$; (iii) $\frac{2-\Gamma}{1-\Gamma}<\frac{\gamma}{\alpha\Lambda}\Leftrightarrow\phi<\overline{\phi}$; and (iv) Γ is an increasing function of ϕ. This also proves that ϕ^* increases in γ. In the proof of Proposition 3, we show that Γ increases in τ, which proves that ϕ^* decreases in τ.

Additional predictions of the model

We consider the following additional predictions of the model:

1. **Unexpected unification of dual-class shares.** If an unexpected unification happens to a dual-class firm at time t, then the share price changes from $p^*_{\text{dual}}(t)=(1-\phi)\frac{e^*_{\text{dual}}}{r-g}(1+g)^t+\phi\frac{e^*_{\text{dual}}}{r-\eta}(1+\eta)^t$ to $p^*_{\text{uni}}(t)=\frac{e^*_{\text{dual}}}{r-g}(1+g)^t$. That is, the effort level does not change, but the founder stops extracting private benefits from control. The abnormal return to the announcement of the unification, defined as $R_{\text{uni}}(t)\equiv\frac{p^*_{\text{uni}}(t)}{p^*_{\text{dual}}(t)}-1$, is given by

$$\begin{aligned}R_{\text{uni}}(t) &= \frac{\frac{e^*_{\text{dual}}}{r-g}(1+g)^t}{(1-\phi)\frac{e^*_{\text{dual}}}{r-g}(1+g)^t+\phi\frac{e^*_{\text{dual}}}{r-\eta}(1+\eta)^t}-1 \\ &= \frac{1}{1-\phi+\phi\frac{r-g}{r-\eta}\frac{1}{(1+\tau)^t}}-1.\end{aligned} \tag{A.21}$$

   Notice $R_{\text{uni}}(t)$ increases with t, and $R_{\text{uni}}(t)>0\Leftrightarrow 1>\frac{r-g}{r-\eta}\frac{1}{(1+\tau)^t}$ which holds given that $g>\eta$.

2. **Unexpected unification of dual-class shares with post-unification effort**. Consider the following simple variant of the baseline model that involves post-IPO effort and unexpected unification/recapitalization. There exists a future date $n>0$ such that for any $t<n$ the firm remains immature: the founder exerts no effort, the firm's cash flows are normalized to zero, and its corporate governance is irrelevant (i.e., the founder cannot divert resources). At $t=n$ the firm adopts dual-class shares (i.e., recapitalization) with an exogenous probability $\xi\in(0,1)$, and single-class shares (i.e., unification) otherwise. Then, the founder chooses effort e at a cost of $\frac{1}{2}\kappa e^2$. From this point onward, the firm's cash flow is given by $y_t=e(1+g)^{t-n}$. Importantly, the rate of extraction of private benefits by the founder is b_t as given by (1). Thus, as n grows larger (i.e., recapitalization/unification occurs when the firm is more mature), the relative importance of agency cost at the recapitalization/unification event is higher. For reasons similar to those in the baseline model we expect that:

- Founder's effort, i.e., firm's productivity, would increase (decrease) upon a dual-class recapitalization (unification), since the founder expects to be able to extract more (fewer) private benefits.
- The abnormal returns around a recapitalization (unification) event decreases (increases) with n, as for small n the agency cost b_n is larger. The abnormal return will be positive (negative) for small n to the extent that the loss (gains) from increased (decreased) agency costs is small relative to the gains (losses) from increased (decreased) productivity.

Notice that the baseline model is a special case where $n=0$.

3. **Expected unification of dual-class shares with a sunset provision $n \geq 2$.** Upon unification, the value of the firm is expected to change from $p^*(n, n-2)$ to $p^*(n, n-1)$. Based on Lemma 1,

$$p^*(n,e,n-1) = \frac{e\left(1+g\right)^{n-1}}{r-g} \tag{A.22}$$

$$p^*(n,e,n-2) = \frac{e\left(1+g\right)^{n-2}}{r-g} - e\phi\frac{\left(1+g\right)^{n-2}-\left(1+\eta\right)^{n-2}}{1+r} \tag{A.23}$$

and

$$\frac{p^*(n,e,n-1)}{p^*(n,e,n-2)} = \frac{\frac{e\left(1+g\right)^{n-1}}{r-g}}{\frac{e\left(1+g\right)^{n-2}}{r-g} - e\phi\frac{\left(1+g\right)^{n-2}-\left(1+\eta\right)^{n-2}}{1+r}} = \frac{1+g}{1-\frac{r-g}{1+r}\phi\left(1-\frac{1}{(1+\tau)^{n-2}}\right)} > 1+g. \tag{A.24}$$

Moreover, $\frac{p^*(n,e,n-1)}{p^*(n,e,n-2)}$ increases in n. Thus, the price reaction to such a unification is positive, since the growth in firm value at the unification is larger than its normal growth rate g. Moreover, the fact that $\frac{p^*(n,n-1)}{p^*(n,n-2)}$ increases in n implies that the price reaction to the unification increases when the horizon of the sunset provision is longer.

4. **Expected unification of dual-class shares with a sunset provision $n \geq 2$ and post unification effort**.Suppose that at date n, when the sunset provision kicks in, the founder decides again whether to exert effort. If effort is exerted, the founder incurs a personal cost of $\epsilon > 0$ and y_t continues to grow at a rate of g as before. If no effort is exerted, then the firm stops growing and $y_t = y_n$ for all $t > n$. Upon unification at time n, the founder can expect to be fired and won't be able to consume additional private benefits. Thus, upon time n unification, exerting effort increases the share price by

$$e\frac{\left(1+g\right)^{n-1}}{r-g} - e\frac{\left(1+g\right)^{n-1}}{r} = \frac{e\left(1+g\right)^{n-1}}{r-g}\frac{g}{r} > 0.$$

If

$$\epsilon > \alpha\Lambda\frac{e\left(1+g\right)^{n-1}}{r-g}\frac{g}{r} \tag{A.25}$$

then the founder will not exert any effort upon time n unification. Relative to the baseline model, for a given pre-IPO effort e, the stock price upon time n unification would be lower by $\frac{e\left(1+g\right)^{n-1}}{r}$. Thus, while the unification eliminates the agency costs, it also inflicts principal costs since shareholders cannot commit not to fire the founder, thereby weakening the his incentives to exert effort. This observation could explain the evidence that stock unifications triggered by sunsets will have small effect on firm value.[27]

[27]Notice that there is a complementarity between pre-IPO and post-IPO effort in this setting. If the founder does not expect to exert effort upon unification, then he would have weaker incentives to exert effort pre-IPO. Indeed, given governance structure n, the founder expects the effect of pre-IPO effort to last only n periods, and hence the incentives are weaker. At the same time, as can

5. **The voting premium.** The expected value to the founder at period $t \geq 1$ gross of the the proceeds from the IPO and the cost of effort is

$$\begin{aligned}
& \alpha\Lambda \sum_{s=t+1}^{\infty} \frac{y_s^*(e^*_{\text{dual}})(1-b_s)}{(1+r)^{s-t}} + \gamma \sum_{s=t+1}^{\infty} \frac{y_s^*(e^*_{\text{dual}})b_s}{(1+r)^{s-t}} \\
= \; & \alpha\Lambda p^*_{\text{dual}}(t) + \gamma \sum_{s=t+1}^{\infty} \frac{e^*_{\text{dual}}\left(1+g\right)^{s-1}\phi\left(1-\frac{1}{(1+\tau)^{(s-1)}}\right)}{(1+r)^{s-t}} \\
= \; & \alpha\Lambda p^*_{\text{dual}}(t) + \gamma\phi e^*_{\text{dual}} \sum_{s=t+1}^{\infty} \frac{\left(1+g\right)^{s-1} - \left(1+\eta\right)^{s-1}}{(1+r)^{s-t}} \\
= \; & \alpha\Lambda p^*_{\text{dual}}(t) + \gamma\phi e^*_{\text{dual}} \left(\left(1+g\right)^{t} \sum_{s=1}^{\infty} \frac{\left(1+g\right)^{s-1}}{(1+r)^{s}} - \left(1+\eta\right)^{t} \sum_{s=1}^{\infty} \frac{\left(1+\eta\right)^{s-1}}{(1+r)^{s}} \right) \\
= \; & \alpha\Lambda p^*_{\text{dual}}(t) + \gamma\phi e^*_{\text{dual}} \left(\frac{\left(1+g\right)^{t}}{r-g} - \frac{\left(1+\eta\right)^{t}}{r-\eta} \right) \\
= \; & \alpha\Lambda p^*_{\text{dual}}(t) + \gamma\phi e^*_{\text{dual}} \left(1+g\right)^{t} \left(\frac{1}{r-g} - \frac{1}{r-\eta}\frac{1}{(1+\tau)^{t}} \right).
\end{aligned}$$

Thus, the voting premium, which is proportional to $\frac{\gamma}{\alpha\Lambda}\phi e^*_{\text{dual}}\left(1+g\right)^{t}\left(\frac{1}{r-g} - \frac{1}{r-\eta}\frac{1}{(1+\tau)^{t}}\right)$, is positive and increasing in t, as required.

be seen from (A.25), larger e, i.e., larger pre-IPO effort, increases the founder's incentives to exert effort post-unification. Intuitively, allowing the firm to continue its growth has a larger impact if the firm is already successful. This complementarity can give rise to multiple equilibria, and either increase or decrease the founders' incentives to exert effort pre-IPO.

Appendix B Proofs

Proof of Lemma 1. Shareholders will fire the founder as soon as they can, that is, at time $t = n$. The founder chooses e to maximize

$$\begin{aligned}
u(e,n) &= \sum_{s=1}^{n-1}\frac{\alpha\Lambda y_s(e)(1-b_s)+\gamma b_s y_s(e)}{(1+r)^s}+\sum_{s=n}^{\infty}\frac{\alpha\Lambda y_s(e)}{(1+r)^s}+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2\\
&= (\gamma-\alpha\Lambda)\sum_{s=1}^{n-1}\frac{y_s(e)\,b_s}{(1+r)^s}+\alpha\Lambda\sum_{s=1}^{\infty}\frac{y_s(e)}{(1+r)^s}+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2\\
&= (\gamma-\alpha\Lambda)\left[\sum_{s=1}^{\infty}\frac{y_s(e)\,b_s}{(1+r)^s}-\sum_{s=n}^{\infty}\frac{y_s(e)\,b_s}{(1+r)^s}\right]+\alpha\Lambda\sum_{s=1}^{\infty}\frac{e\left(1+g\right)^{s-1}}{(1+r)^s}+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2\\
&= (\gamma-\alpha\Lambda)e\phi\left[\sum_{s=1}^{\infty}\frac{\left(1+g\right)^{s-1}-\left(1+\eta\right)^{s-1}}{(1+r)^s}-\sum_{s=n}^{\infty}\frac{\left(1+g\right)^{s-1}-\left(1+\eta\right)^{s-1}}{(1+r)^s}\right]\\
&\quad+\alpha\Lambda\frac{e}{r-g}+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2\\
&= (\gamma-\alpha\Lambda)e\phi\left[\frac{1}{r-g}-\frac{1}{r-\eta}-\sum_{s=n}^{\infty}\frac{\left(1+g\right)^{s-1}}{(1+r)^s}+\sum_{s=n}^{\infty}\frac{\left(1+\eta\right)^{s-1}}{(1+r)^s}\right]\\
&\quad+\alpha\Lambda\frac{e}{r-g}+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2\\
&= (\gamma-\alpha\Lambda)e\phi\left[\frac{g-\eta}{(r-g)(r-\eta)}-\left(\frac{1+g}{1+r}\right)^{n-1}\frac{1}{r-g}+\left(\frac{1+\eta}{1+r}\right)^{n-1}\frac{1}{r-\eta}\right]\\
&\quad+\alpha\Lambda\frac{e}{r-g}+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2\\
&= (\gamma-\alpha\Lambda)e\phi\left[\frac{g-\eta}{(r-g)(r-\eta)}-\psi(n-1)\right]+\alpha\Lambda\frac{e}{r-g}+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2.
\end{aligned}$$

Thus

$$u(n)=\left[e^*_{\text{dual}}-(\gamma-\alpha\Lambda)\frac{\phi}{\kappa}\psi(n-1)\right]e\kappa+\alpha(1-\Lambda)p_0-\frac{1}{2}\kappa e^2$$

and

$$e^*(n)=e^*_{\text{dual}}-(\gamma-\alpha\Lambda)\frac{\phi}{\kappa}\psi(n-1).$$

Notice, $\psi(n-1)$ decreases in n and hence $e^*(n)$ increases in n.

The share price is

$$\begin{aligned}
p^*(n,e,t) &= \sum_{s=t+1}^{n-1} \frac{y_s(e)(1-b_s)}{(1+r)^{s-t}} + \sum_{s=\max\{n,t+1\}}^{\infty} \frac{y_s(e)}{(1+r)^{s-t}} \\
&= -\sum_{s=t+1}^{n-1} \frac{y_s(e)\,b_s}{(1+r)^{s-t}} + \sum_{s=t+1}^{n-1} \frac{y_s(e)}{(1+r)^{s-t}} + \sum_{s=\max\{n,t+1\}}^{\infty} \frac{y_s(e)}{(1+r)^{s-t}} \\
&= -\sum_{s=t+1}^{n-1} \frac{y_s(e)\,b_s}{(1+r)^{s-t}} + \sum_{s=t+1}^{\infty} \frac{y_s(e)}{(1+r)^{s-t}} \\
&= \frac{e\left(1+g\right)^t}{r-g} - \sum_{s=t+1}^{n-1} \frac{y_s(e)\,b_s}{(1+r)^{s-t}}
\end{aligned}$$

If $t+1 \geq n$ then

$$p^*(n,e,t) = \frac{e\left(1+g\right)^t}{r-g}.$$

If $t+1 \leq n-1$ then

$$\begin{aligned}
p^*(n,e,t) &= \frac{e\left(1+g\right)^t}{r-g} - \sum_{s=t+1}^{n-1} \frac{y_s(e)\,b_s}{(1+r)^{s-t}} \\
&= \frac{e\left(1+g\right)^t}{r-g} - \left[\sum_{s=t+1}^{\infty} \frac{y_s(e)\,b_s}{(1+r)^{s-t}} - \sum_{s=n}^{\infty} \frac{y_s(e)\,b_s}{(1+r)^{s-t}}\right] \\
&= \frac{e\left(1+g\right)^t}{r-g} - e\phi\left[\sum_{s=t+1}^{\infty} \frac{\left(1+g\right)^{s-1} - \left(1+\eta\right)^{s-1}}{(1+r)^{s-t}} - \sum_{s=n}^{\infty} \frac{\left(1+g\right)^{s-1} - \left(1+\eta\right)^{s-1}}{(1+r)^{s-t}}\right] \\
&= \frac{e\left(1+g\right)^t}{r-g} - e\phi\left[\begin{array}{c} \left(1+g\right)^t \sum_{m=1}^{\infty} \frac{(1+g)^{m-1}}{(1+r)^m} - \left(1+\eta\right)^t \sum_{m=1}^{\infty} \frac{(1+\eta)^{m-1}}{(1+r)^m} \\ -\sum_{s-n+1=1}^{\infty} \frac{(1+g)^{s-n+n-1}}{(1+r)^{s-n+n-t}} + \sum_{s-n+1=1}^{\infty} \frac{(1+\eta)^{s-n+n-1}}{(1+r)^{s-n+n-t}} \end{array}\right] \\
&= \frac{e\left(1+g\right)^t}{r-g} - e\phi\left[\begin{array}{c} \left(1+g\right)^t \sum_{m=1}^{\infty} \frac{(1+g)^{m-1}}{(1+r)^m} - \left(1+\eta\right)^t \sum_{m=1}^{\infty} \frac{(1+\eta)^{m-1}}{(1+r)^m} \\ -\frac{(1+g)^{n-1}}{(1+r)^{n-1-t}} \sum_{m=1}^{\infty} \frac{(1+g)^{m-1}}{(1+r)^m} + \frac{(1+\eta)^{n-1}}{(1+r)^{n-1-t}} \sum_{m=1}^{\infty} \frac{(1+\eta)^{m-1}}{(1+r)^m} \end{array}\right] \\
&= \frac{e\left(1+g\right)^t}{r-g} - e\phi\left[\frac{\left(1+g\right)^t}{r-g}\left(1-\left(\frac{1+g}{1+r}\right)^{n-1-t}\right) - \frac{\left(1+\eta\right)^t}{r-\eta}\left(1-\left(\frac{1+\eta}{1+r}\right)^{n-1-t}\right)\right].
\end{aligned}$$

Overall,

$$p^*_{\text{sunset}}(n,e,t) = e\begin{cases} \frac{(1+g)^t}{r-g} - \phi\left[\frac{(1+g)^t}{r-g}\left(1-\left(\frac{1+g}{1+r}\right)^{n-1-t}\right) - \frac{(1+\eta)^t}{r-\eta}\left(1-\left(\frac{1+\eta}{1+r}\right)^{n-1-t}\right)\right] & \text{if } t+1 \leq n-1 \\ \frac{(1+g)^t}{r-g} & \text{if } t+1 \geq n \end{cases}$$

and

$$\begin{aligned} p^*(n,e,0) &= p^*(n,e) \\ &= \frac{e}{r-g} - e\phi\left[\frac{g-\eta}{(r-g)(r-\eta)} - \left(\frac{1+g}{1+r}\right)^{n-1}\frac{1}{r-g} + \left(\frac{1+\eta}{1+r}\right)^{n-1}\frac{1}{r-\eta}\right] \\ &= e\frac{1}{r-g}\left(1-\phi\frac{g-\eta}{r-\eta}\right) + e\phi\psi(n-1). \end{aligned}$$

Also notice that

$$\begin{aligned} p^*(n,e,n-1) &= \frac{e(1+g)^{n-1}}{r-g} \\ p^*(n,e,n-2) &= \frac{e(1+g)^{n-2}}{r-g} - e\phi\frac{(1+g)^{n-2}-(1+\eta)^{n-2}}{1+r} \end{aligned}$$

and

$$\frac{p^*(n,e,n-1)}{p^*(n,e,n-2)} = \frac{\frac{e(1+g)^{n-1}}{r-g}}{\frac{e(1+g)^{n-2}}{r-g} - e\phi\frac{(1+g)^{n-2}-(1+\eta)^{n-2}}{1+r}} = \frac{1+g}{1-\frac{r-g}{1+r}\phi\left(1-\frac{1}{(1+\tau)^{n-2}}\right)} > 1+g.$$

Moreover, $\frac{p^*(n,e,n-1)}{p^*(n,e,n-2)}$ increases in n, and thus, the price reaction to an expected unification increases in n. □

Proof of Lemma 2. Notice

$$\begin{aligned} p^*_{\text{dual}}(t) - p^*_{\text{single}}(t) &= \frac{e^*_{\text{dual}}}{r-g}(1+g)^t - \phi e^*_{\text{dual}}\left[\frac{(1+g)^t}{r-g} - \frac{(1+\eta)^t}{r-\eta}\right] - \frac{e^*_{\text{single}}(1+g)^t}{r-g} \\ &= \frac{e^*_{\text{dual}}(1-\phi) - e^*_{\text{single}}}{r-g}(1+g)^t + \frac{\phi e^*_{\text{dual}}}{r-\eta}(1+\eta)^t \end{aligned}$$

and

$$\begin{aligned} p^*_{\text{dual}}(0) &> p^*_{\text{single}}(0) \Leftrightarrow \\ \frac{e^*_{\text{dual}}(1-\phi) - e^*_{\text{single}}}{r-g} + \frac{\phi e^*_{\text{dual}}}{r-\eta} &> 0 \Leftrightarrow \\ \frac{\gamma}{\alpha\Lambda} &> 1 + \frac{1}{1-\phi\frac{g-\eta}{r-\eta}} = \frac{2-\Gamma}{1-\Gamma}. \end{aligned}$$

Also notice that if

$$e^*_{\text{dual}}(1-\phi) - e^*_{\text{single}} \geq 0 \Leftrightarrow \frac{\gamma}{\alpha\Lambda} \geq 1 + \frac{\phi}{1-\phi}\frac{1}{\Gamma}$$

then $p^*_{\text{dual}}(t)-p^*_{\text{single}}(t)$ is positive and increases in t for all $t\geq 0$. In this case, $T^*=\infty$. Suppose $\frac{\gamma}{\alpha\Lambda}<1+\frac{\phi}{1-\phi}\frac{1}{\Gamma}$. Then,

$$\begin{aligned}
p^*_{\text{dual}}(t) &> p^*_{\text{single}}(t)\Leftrightarrow\\
\frac{e^*_{\text{dual}}(1-\phi)-e^*_{\text{single}}}{r-g}(1+g)^t+\frac{\phi e^*_{\text{dual}}}{r-\eta}(1+\eta)^t &> 0\Leftrightarrow\\
\frac{e^*_{\text{dual}}(1-\phi)-e^*_{\text{single}}}{r-g}+\frac{\phi e^*_{\text{dual}}}{r-\eta}\frac{1}{(1+\tau)^t} &> 0\Leftrightarrow\\
t &< T^*=\frac{\ln\left(\phi\frac{r-g}{r-\eta}\frac{e^*_{\text{dual}}}{e^*_{\text{single}}-e^*_{\text{dual}}(1-\phi)}\right)}{\ln(1+\tau)},
\end{aligned}$$

where $\frac{e^*_{\text{dual}}}{e^*_{\text{single}}-e^*_{\text{dual}}(1-\phi)}=\frac{1}{\frac{1}{1+(\frac{\gamma}{\alpha\Lambda}-1)\Gamma}-(1-\phi)}$.

Notice, $1+\frac{\phi}{1-\phi}\frac{1}{\Gamma}>\frac{2-\Gamma}{1-\Gamma}\Leftrightarrow r>g$, which holds by assumption. Therefore, there are cases where $1+\frac{\phi}{1-\phi}\frac{1}{\Gamma}>\frac{\gamma}{\alpha\Lambda}>\frac{2-\Gamma}{1-\Gamma}$, that is, $p^*_{\text{dual}}(0)>p^*_{\text{single}}(0)$ and $T^*\in(0,\infty)$. □

Proof of Lemma 3. The founder's expected utility given n is

$$\begin{aligned}
u(n) &= \left[e^*_{\text{dual}}-(\gamma-\alpha\Lambda)\frac{\phi}{\kappa}\psi(n-1)\right]e^*(n)\kappa+\alpha(1-\Lambda)p^*_{\text{sunset}}(n)-\frac{1}{2}\kappa e^*(n)^2\\
&= \alpha(1-\Lambda)p^*(n,0)+\frac{1}{2}\kappa e^*(n)^2.
\end{aligned}$$

Therefore, $u^*(\infty)>u^*(1)$ if and only if

$$\begin{aligned}
\frac{\kappa}{2}\left[e^{*2}_{\text{dual}}-e^{*2}_{\text{single}}\right]+\alpha(1-\Lambda)\left[p^*_{\text{dual}}(0)-p^*_{\text{single}}(0)\right] &> 0\Leftrightarrow\\
\frac{1}{2}\left[e^{*2}_{\text{dual}}-e^{*2}_{\text{single}}\right]+\frac{1-\Lambda}{\Lambda}\left[e^*_{\text{single}}e^*_{\text{dual}}(1-\Gamma)-e^{*2}_{\text{single}}\right] &> 0\Leftrightarrow\\
\frac{1}{2}\left[\left[1+\Gamma\left(\frac{\gamma}{\alpha\Lambda}-1\right)\right]^2e^{*2}_{\text{single}}-e^{*2}_{\text{single}}\right]+\frac{1-\Lambda}{\Lambda}\left[\left[1+\Gamma\left(\frac{\gamma}{\alpha\Lambda}-1\right)\right]e^{*2}_{\text{single}}(1-\Gamma)-e^{*2}_{\text{single}}\right] &> 0\Leftrightarrow\\
\frac{1}{2}\left[\left[1+\Gamma\left(\frac{\gamma}{\alpha\Lambda}-1\right)\right]^2-1\right]+\frac{1-\Lambda}{\Lambda}\left[\left[1+\Gamma\left(\frac{\gamma}{\alpha\Lambda}-1\right)\right](1-\Gamma)-1\right] &> 0\Leftrightarrow\\
\frac{1}{2}\left(\frac{\gamma}{\alpha\Lambda}-1\right)\left(2+\Gamma\left(\frac{\gamma}{\alpha\Lambda}-1\right)\right)+\frac{1-\Lambda}{\Lambda}\left[(1-\Gamma)\left(\frac{\gamma}{\alpha\Lambda}-1\right)-1\right] &> 0\Leftrightarrow\\
\left(\frac{\gamma}{\alpha\Lambda}-1\right)^2+\left(\frac{\gamma}{\alpha\Lambda}-1\right)2\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)-\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda} &> 0\Leftrightarrow\\
\frac{\gamma}{\alpha\Lambda} &> H(\Gamma)
\end{aligned}$$

as required. Notice $H(\Gamma)>1\Leftrightarrow$

$$\sqrt{\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda}}>1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\Leftrightarrow\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda}>0,$$

which always holds. Also, $H(\Gamma) < \frac{2-\Gamma}{1-\Gamma} \Leftrightarrow$

$$\begin{aligned}
\sqrt{\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda}} &< 1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}+\frac{1}{1-\Gamma} \Leftrightarrow \\
\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda} &< \left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+2\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)\left(\frac{1}{1-\Gamma}\right)+\left(\frac{1}{1-\Gamma}\right)^2 \Leftrightarrow \\
\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda} &< 2\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)\left(\frac{1}{1-\Gamma}\right)+\left(\frac{1}{1-\Gamma}\right)^2 \Leftrightarrow \\
0 &< \frac{2}{1-\Gamma}+\frac{2}{\Gamma}+\frac{1}{\Lambda}\frac{1}{\Gamma}+\left(\frac{1}{1-\Gamma}\right)^2,
\end{aligned}$$

which always holds. Moreover, notice that

$$\frac{\partial H(\Gamma)}{\partial \Gamma} = \frac{1}{\Lambda}\frac{1}{\Gamma^2} - \frac{2\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)\frac{1}{\Lambda}\frac{1}{\Gamma^2}+\frac{2}{\Gamma^2}\frac{1-\Lambda}{\Lambda}}{2\sqrt{\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda}}}$$

and

$$\begin{aligned}
\frac{\partial H(\Gamma)}{\partial \Gamma} &> 0 \Leftrightarrow \\
1 &> \frac{\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)+1-\Lambda}{\sqrt{\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda}}} \Leftrightarrow \\
\sqrt{\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda}} &> \left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)+1-\Lambda \Leftrightarrow \\
\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+\frac{2}{\Gamma}\frac{1-\Lambda}{\Lambda} &> \left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)^2+2(1-\Lambda)\left(1+\frac{1}{\Lambda}\frac{1-\Gamma}{\Gamma}\right)+(1-\Lambda)^2 \Leftrightarrow \\
0 &> 2(1-\Lambda)-2\frac{1-\Lambda}{\Lambda}+(1-\Lambda)^2 \Leftrightarrow \\
2 &> \Lambda,
\end{aligned}$$

which always holds given the assumption $\Lambda \in (0,1)$.

Next, observe that

$$p^*(n+1,e)-p^*(n,e) = e\phi\left[\psi(n)-\psi(n-1)\right] < 0$$

and

$$e^*(n+1)-e^*(n) = \left(\gamma-\alpha\Lambda\right)\frac{\phi}{\kappa}\left[\psi(n-1)-\psi(n)\right] > 0.$$

For $n \geq 1$,

$$\begin{aligned}
& p^*(n+1,e^*(n+1))-p^*(n,e^*(n)) \\
= \;& p^*\left(n+1,e^*(n)+\left(\gamma-\alpha\Lambda\right)\frac{\phi}{\kappa}\left[\psi(n-1)-\psi(n)\right]\right)-p^*(n+1,e^*(n))+e^*(n)\phi\left[\psi(n)-\psi(n-1)\right] \\
= \;& p^*\left(n+1,\left(\gamma-\alpha\Lambda\right)\frac{\phi}{\kappa}\left[\psi(n-1)-\psi(n)\right]\right)+e^*(n)\phi\left[\psi(n)-\psi(n-1)\right] \\
= \;& \left[\frac{1}{r-g}\left(1-\phi\frac{g-\eta}{r-\eta}\right)+\phi\psi(n)\right]\left(\gamma-\alpha\Lambda\right)\frac{\phi}{\kappa}\left[\psi(n-1)-\psi(n)\right]+e^*(n)\phi\left[\psi(n)-\psi(n-1)\right] \\
= \;& \left(\left[\frac{1}{r-g}\left(1-\phi\frac{g-\eta}{r-\eta}\right)+\phi\psi(n)\right]\frac{\gamma-\alpha\Lambda}{\kappa}-e^*(n)\right)\phi\left[\psi(n-1)-\psi(n)\right] \\
= \;& \left(\frac{\gamma-2\alpha\Lambda}{\kappa}\frac{1}{r-g}-2\frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi}{\kappa}\frac{g-\eta}{r-\eta}+\left(\gamma-\alpha\Lambda\right)\frac{\phi}{\kappa}\left[\psi(n-1)+\psi(n)\right]\right)\phi\left[\psi(n-1)-\psi(n)\right].
\end{aligned}$$

Thus,

$$p^*(n+1,e^*(n+1))>p^*(n,e^*(n)) \Leftrightarrow \psi(n)+\psi(n-1)>\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g}.$$

Notice that $\Psi(n)\equiv\psi(n)+\psi(n-1)$ is a positive, decreasing function, where $\Psi(1)=2\frac{(g-\eta)}{(r-g)(r-\eta)}>0$ and $\lim_{n\to\infty}\psi(n)=0$. Thus

1. If $\Psi(1)\leq\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g} \Leftrightarrow \frac{\gamma}{\alpha\Lambda}\leq 2$ then $\Psi(n)<\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g}$ for all $n>1$ and hence, $p^*(n+1,e^*(n+1))<p^*(n,e^*(n))$ for all n, that is, $p^*(n,e^*(n))$ is a decreasing function of n. Therefore, $n^*=1$, that is single-class shares are optimal. Notice that $\frac{\gamma}{\alpha\Lambda}\leq 2$, and hence, this must imply $\Gamma>\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}\Rightarrow\frac{\gamma}{\alpha\Lambda}<1+\frac{1}{1-\Gamma}$, that is, $p^*_{\text{dual}}(0)<p^*_{\text{single}}(0)$.

2. If $\lim_{n\to\infty}\Psi(n)>\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g} \Leftrightarrow \phi<\frac{1}{2}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}\frac{r-\eta}{g-\eta}$ then $\Psi(n)>\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g}$ for all $n>1$ and hence, $p^*(n+1,e^*(n+1))$ $p^*(n,e^*(n))$ for all n, that is, $p^*(n,e^*(n))$ is an increasing function of n. Therefore, $n^*=\infty$, that is dual-class shares are optimal. In this case, it must be $\gamma-2\alpha\Lambda>0$, and hence, this must imply : $\Gamma<\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}\Rightarrow$ $\frac{\gamma}{\alpha\Lambda}>1+\frac{1}{1-\Gamma}$, that is, $p^*_{\text{dual}}(0)>p^*_{\text{single}}(0)$.

3. If $\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g}\in(0,\Psi(1))$ then the optimal n is $2\leq n^*<\infty$. Notice that this is equivalent to requiring

$$\phi>\frac{1}{2}\frac{\frac{\gamma}{\alpha\Lambda}-2}{\frac{\gamma}{\alpha\Lambda}-1}\frac{r-\eta}{g-\eta} \text{ and } \frac{\gamma}{\alpha\Lambda}>2,$$

and it satisfies

$$\Psi(n^*)\leq\frac{2\frac{g-\eta}{r-\eta}-\frac{1}{\phi}\frac{\gamma-2\alpha\Lambda}{\gamma-\alpha\Lambda}}{r-g}\leq\Psi(n^*-1).$$

Finally, recall $u(n)=\alpha(1-\Lambda)p^*(n,0)+\frac{1}{2}\kappa e^*(n)^2$. Since $e^*(n)$ is increasing in n, $\arg\max_n u(n)\geq n^*_{\text{sh}}$, that is, the founder always benefits from a longer sunset since this way she gets more private benefits and higher IPO proceeds. □

Proof of Proposition 3. According to Lemma 3 and Proposition 2, firm i chooses dual-class shares if and only if $\frac{\gamma_i}{\alpha_i\Lambda_i}>H_i(\Gamma_i)\Leftrightarrow\phi_i<\phi_i^*$. Suppose firm $i=d$ chose to be a dual-class firm (i.e., $\phi_d<\phi_d^*$) and firm $i=s$ chose to be a signal-class firm (i.e., $\phi_s>\phi_s^*$). We study whether the selection effect can stem from one-dimensional heterogeneity, where firms differ in at most one variable from the set $\{g,r,\tau,\phi,\alpha,\Lambda,\gamma\}$. We consider two versions

of this thought experiment. First, we assume firms that chose to adopt dual-class shares are unexpectedly forced to adopt single-class shares after the initial choice. Second, we assume firms that chose single-class shares are unexpectedly forced to adopt dual-class shares.

1. **Dual-class firms are forced to be single-class firms.** In this case, firm d is forced to be single-class in spite of its choice to be dual-class. Under this assumption, and according to Lemma 1, the founder of firm $i \in \{d,s\}$ chooses the effort level

$$e^*_{\text{single,i}} = \frac{1}{\kappa_i}\frac{\alpha_i \Lambda_i}{r_i - g_i}$$

and the share price at time t is

$$p^*_{\text{single,i}}(t) = \frac{e^*_{\text{single,i}}}{r_i - g_i}\left(1+g_i\right)^t.$$

To replicate the pattern in Proposition 1, we ask whether there exists $T^* \in (0,\infty)$ such that $p^*_{\text{single,d}}(t) > p^*_{\text{single,s}}(t)$ if $t < T^*$ and $p^*_{\text{single,d}}(t) < p^*_{\text{single,s}}(t)$ otherwise.

Observe that an increase in one of the parameters in $\{g, r, \tau, \phi, \alpha, \Lambda, \gamma, \kappa\}$ will lead to an increase in $p^*_{\text{single,i}}(t)$ for all t (in the case of g, α, and Λ), a decrease in $p^*_{\text{single,i}}(t)$ for all t (in the case of r and κ), or no change at all (in the case of τ, ϕ, and γ). Therefore, if the selection is driven by heterogeneity in only one of these parameters, it must be either $p^*_{\text{single,d}}(t) < p^*_{\text{single,s}}(t)$ for all t or $p^*_{\text{single,d}}(t) > p^*_{\text{single,s}}(t)$ for all t — i.e., the selection effect cannot stem from one-dimensional heterogeneity in this case.

2. **Single-class firms are forced to be dual-class firms.** In this case, firm s is forced to be dual-class in spite of its selection to be single-class. Under this assumption, and according to Lemma 1, the founder of firm $i \in \{d,s\}$ choose the effort level

$$\begin{aligned}
e^*_{\text{dual,i}} &= e^*_{\text{single,i}} + \frac{\gamma_i - \alpha_i\Lambda_i}{r_i - g_i}\frac{\phi_i}{\kappa_i}\frac{g_i - \eta_i}{r_i - \eta_i} \\
&= e^*_{\text{single,i}} + \frac{\gamma_i - \alpha_i\Lambda_i}{r_i - g_i}\frac{\phi_i}{\kappa_i}\frac{\left(1+g_i\right)^{\frac{\tau_i}{1+\tau_i}}}{r_i - g_i + \left(1+g_i\right)^{\frac{\tau_i}{1+\tau_i}}} \\
&= e^*_{\text{single,i}} + \frac{\gamma_i - \alpha_i\Lambda_i}{r_i - g_i}\frac{\phi_i}{\kappa_i}\frac{1+g_i}{1 + r_i + \frac{r_i - g_i}{\tau_i}}
\end{aligned}$$

where $\eta_i = \frac{1+g_i}{1+\tau_i} - 1$, and the share price at time t is

$$\begin{aligned}
p^*_{\text{dual,i}}(t) &= \left(1-\phi_i\right)\frac{e^*_{\text{dual,i}}}{r_i - g_i}\left(1+g_i\right)^t + \phi_i\frac{e^*_{\text{dual,i}}}{r_i - \eta_i}\left(1+\eta_i\right)^t \\
&= e^*_{\text{dual,i}}\left(1+g_i\right)^t\left[\frac{1-\phi_i}{r_i - g_i} + \frac{\phi_i}{r_i - g_i + \left(1+g_i\right)^{\frac{\tau_i}{1+\tau_i}}}\frac{1}{\left(1+\tau_i\right)^t}\right] \\
&= \frac{1}{\kappa_i}\frac{\alpha_i\Lambda_i}{r_i - g_i}\left[1 + \frac{\gamma_i - \alpha_i\Lambda_i}{\alpha_i\Lambda_i}\phi_i\frac{1+g_i}{r_i + \frac{r_i - g_i}{\tau_i} + 1}\right]\left(1+g_i\right)^t\left[\frac{1-\phi_i}{r_i - g_i} + \frac{\phi_i}{r_i - g_i + \left(1+g_i\right)^{\frac{\tau_i}{1+\tau_i}}}\frac{1}{\left(1+\tau_i\right)^t}\right].
\end{aligned}$$

To replicate the pattern in Proposition 1, we ask whether there exists $T^* \in (0,\infty)$ such that $p^*_{\text{dual,d}}(t) > p^*_{\text{dual,s}}(t)$ if $t < T^*$ and $p^*_{\text{dual,d}}(t) < p^*_{\text{dual,s}}(t)$ otherwise.

Observe that an increase in one of the parameters in $\{r, g, \alpha, \Lambda, \gamma, \kappa\}$ will lead to either an increase in $p^*_{\text{dual,i}}(t)$ for all t (in the case of g, α, Λ and γ) or a decrease in $p^*_{\text{dual,i}}(t)$ for all t (in the case of r and κ). Therefore, if the selection is driven by heterogeneity in only one of these parameters, it must be either

$p^*_{\text{dual,d}}(t) < p^*_{\text{dual,s}}(t)$ for all t, or $p^*_{\text{dual,d}}(t) > p^*_{\text{dual,s}}(t)$ for all t — i.e., the selection effect cannot stem from one-dimensional heterogeneity of these parameters.

It's left to consider τ and ϕ.

- Consider first the effect of ϕ. Recall firm i chooses dual-class shares if and only if $\phi_i < \phi_i^*$. Therefore, it must be $\phi_d < \phi_s$, and $p^*_{\text{dual,d}}(t) > p^*_{\text{dual,s}}(t)$ if and only if

$$\begin{aligned}
&\left(e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi_d}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\right)\left(\frac{1-\phi_d}{r-g} + \frac{\phi_d}{r-g+\left(1+g\right)\frac{\tau}{1+\tau}}\frac{1}{(1+\tau)^t}\right) \quad > \\
&\left(e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi_s}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\right)\left(\frac{1-\phi_s}{r-g} + \frac{\phi_s}{r-g+\left(1+g\right)\frac{\tau}{1+\tau}}\frac{1}{(1+\tau)^t}\right) \quad \Leftrightarrow \\
&\left(\begin{array}{c}\frac{1}{r-g}\left(e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi_d}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\right) \\ +\left(e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi_d}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\right)\left(\frac{1}{r-g+\left(1+g\right)\frac{\tau}{1+\tau}}\frac{1}{(1+\tau)^t} - \frac{1}{r-g}\right)\phi_d\end{array}\right) \quad > \\
&\left(\begin{array}{c}\frac{1}{r-g}\left(e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi_s}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\right) \\ +\left(e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{\phi_s}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\right)\left(\frac{1}{r-g+\left(1+g\right)\frac{\tau}{1+\tau}}\frac{1}{(1+\tau)^t} - \frac{1}{r-g}\right)\phi_s\end{array}\right) \quad \Leftrightarrow \\
&\frac{\gamma-\alpha\Lambda}{r-g}\frac{1}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\left(\frac{1}{r-g+\left(1+g\right)\frac{\tau}{1+\tau}}\frac{1}{(1+\tau)^t} - \frac{1}{r-g}\right)\left(\phi_d^2-\phi_s^2\right) \quad > \\
&\left[\frac{1}{r-g}\frac{\gamma-\alpha\Lambda}{r-g}\frac{1}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}} + e^*_{\text{single}}\left(\frac{1}{r-g+\left(1+g\right)\frac{\tau}{1+\tau}}\frac{1}{(1+\tau)^t} - \frac{1}{r-g}\right)\right]\left(\phi_s-\phi_d\right).
\end{aligned}$$

  Thus $p^*_{\text{dual,d}}(t) > p^*_{\text{dual,s}}(t)$ if and only if

$$\frac{r-g+\left(1+g\right)\frac{\tau}{1+\tau}}{r-g}\frac{e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{1}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\left(\phi_s+\phi_d-1\right)}{e^*_{\text{single}} + \frac{\gamma-\alpha\Lambda}{r-g}\frac{1}{\kappa}\frac{1+g}{1+r+\frac{r-g}{\tau}}\left(\phi_s+\phi_d\right)} > \frac{1}{(1+\tau)^t}.$$

  This means that either the inequality has the same sign for all t, or that it holds only for large t. Either way, it does not replicate the dynamic effect Proposition 1 describes.

- Next, consider the effect of τ. Recall firm i chooses dual-class shares if and only if $\frac{\gamma_i}{\alpha_i\Lambda_i} > H_i(\Gamma_i)$, where $H(\cdot)$ is an increasing function. Notice that Γ can be rewritten as $\Gamma = \phi\frac{1+g}{1+r+\frac{r-g}{\tau}}$, which is an increasing function of τ. Therefore, $\tau_d < \tau_s$. By example, we show that it's possible to generate the dynamic pattern in Proposition 1. In particular, it holds for the following parameter values: $r = 0.5$, $g = 0.4$, $\phi = 0.9$, $\alpha = 0.4$, $\Lambda = 0.4$, $\kappa = 1$, $\gamma = 0.305$, $\tau_s = 2$, and $\tau_d = 1$. It can be verified that with these parameters, firm i chooses dual-class shares if and only if $1.9062 > H(\tau)$, where $H(1) = 1.897$ and $H(2) = 1.909$. Therefore the $\tau = 1$ firm selects dual-class and the $\tau = 2$ firm selects single-class shares. We can now plot the difference between these two firms' share prices from $t \in [0, 10]$.

Finally, recall $e^*_{\text{dual,i}} = e^*_{\text{single,i}} + \frac{\gamma_i-\alpha_i\Lambda_i}{r_i-g_i}\frac{\phi_i}{\kappa_i}\frac{1+g_i}{1+r_i+\frac{r_i-g_i}{\tau_i}}$. Thus, all else equal, $e^*_{\text{dual,i}}$ increases in τ_i. Moreover, notice $\frac{\partial b_t}{\partial \tau} = \frac{\ln(1+\tau)\phi}{(1+\tau)^{t-1}} > 0$, which completes the proof.

□

Multi-dimensional unobserved heterogeneity

In this section, we allow firms to differ in more than one dimension. We first analyze the selection effect under the assumption that firms that choose dual-class shares are forced to adopt single-class shares. Under this assumption,

$$\begin{aligned} p^*_{\text{single,d}}(t) &> p^*_{\text{single,s}}(t) \Leftrightarrow \\ p^*_{\text{single,d}}(0)\left(1+g_d\right)^t &> p^*_{\text{single,s}}(0)\left(1+g_s\right)^t \Leftrightarrow \\ t\ln\left[\frac{1+g_d}{1+g_s}\right] &> \ln\left[\frac{p^*_{\text{single,s}}(0)}{p^*_{\text{single,d}}(0)}\right]. \end{aligned}$$

Thus, the dynamic pattern in Proposition 1 can be generated only if $g_d \neq g_s$. If $g_d > g_s$ then

$$p^*_{\text{single,d}}(t) > p^*_{\text{single,s}}(t) \Leftrightarrow t > \frac{\ln\left[\frac{p^*_{\text{single,s}}(0)}{p^*_{\text{single,d}}(0)}\right]}{\ln\left[\frac{1+g_d}{1+g_s}\right]},$$

which implies that young (mature) selected dual-class firms underperform (outperform) selected single-class firms, which is the opposite from the claim in Proposition 1 and our empirical findings. Thus, it must be $g_d < g_s$. In this case,

$$p^*_{\text{single,d}}(t) > p^*_{\text{single,s}}(t) \Leftrightarrow t < \frac{\ln\left[\frac{p^*_{\text{single,s}}(0)}{p^*_{\text{single,d}}(0)}\right]}{\ln\left[\frac{1+g_d}{1+g_s}\right]}.$$

Notice the right hand side of the above inequality is positive if and only if $p^*_{\text{single,d}}(0) > p^*_{\text{single,s}}(0)$. Thus, conditional on $p^*_{\text{single,d}}(0) > p^*_{\text{single,s}}(0)$, the dynamic pattern in Proposition 1 requires the selected dual-class firms to have lower growth but higher valuation at the IPO. Recall $p^*_{\text{single,i}}(0) = \frac{1}{\kappa_i}\frac{\alpha_i\Lambda_i}{\left(r_i-g_i\right)^2}$ and that a firm chooses dual-class shares if and only if $\frac{\gamma_i}{\alpha_i\Lambda_i} > H_i(\Gamma_i)$. Thus, selected dual-class firms have, all else equal, lower $\alpha_i\Lambda_i$ and lower $H_i(\Gamma_i)$, that is, lower $\Gamma_i = \phi\frac{\frac{\tau}{1+\tau}}{\frac{1+r}{1+g}-\frac{1}{1+\tau}}$ and hence lower g. In principle, there are cases in which the selected dual-class firm has smaller g and $\alpha\Lambda$, but still $p^*_{\text{single,d}}(0) > p^*_{\text{single,s}}(0)$ (e.g., $\kappa_d < \kappa_s$, that is, the founder of a dual-class firm is very skilled but the venture has low growth). Importantly, however, to generate the dynamic pattern in Proposition 1, it must be that firms differ in at least two dimensions: growth rate, and some other unobserved characteristic.

Next, we analyze the selection effect under the assumption that firms that chose single-class shares are forced to adopt dual-class shares. Under this assumption, $e^*_{\text{dual,i}} = e^*_{\text{single,i}} + \frac{\gamma_i-\alpha_i\Lambda_i}{r_i-g_i}\frac{\phi_i}{\kappa_i}\frac{g_i-\eta_i}{r_i-\eta_i}$, and the share price can be written as

$$p^*_{\text{dual,i}}(t) = e^*_{\text{dual,i}}\left(1+g_i\right)^t\left[A_i + \frac{B_i}{\left(1+\tau_i\right)^t}\right]$$

where $A_i = \frac{1-\phi_i}{r_i-g_i} > 0$ and $B_i = \frac{\phi_i}{r_i-\eta_i} > 0$. Then,

$$\begin{aligned} p^*_{\text{dual,d}}(t) &< p^*_{\text{dual,s}}(t) \Leftrightarrow \\ e^*_{\text{dual,d}}\left(1+g_d\right)^t\left[A_d + \frac{B_d}{\left(1+\tau_d\right)^t}\right] &< e^*_{\text{dual,s}}\left(1+g_s\right)^t\left[A_s + \frac{B_s}{\left(1+\tau_s\right)^t}\right] \Leftrightarrow \\ \left(\frac{1+g_d}{1+g_s}\right)^t &< \frac{e^*_{\text{dual,s}}\left[A_s + \frac{B_s}{\left(1+\tau_s\right)^t}\right]}{e^*_{\text{dual,d}}\left[A_d + \frac{B_d}{\left(1+\tau_d\right)^t}\right]}. \end{aligned}$$

Notice that in the limit of $t \to \infty$, $p^*_{\text{dual,d}}(t) < p^*_{\text{dual,s}}(t)$ holds only if $g_d \leq g_s$. Indeed, if $g_d > g_s$ then $\lim_{t\to\infty}\left(\frac{1+g_d}{1+g_s}\right)^t = \infty$ but $\lim_{t\to\infty} \frac{e^*_{\text{dual,s}}\left[A_s + \frac{B_s}{(1+\tau_s)^t}\right]}{e^*_{\text{dual,d}}\left[A_d + \frac{B_d}{(1+\tau_d)^t}\right]} = \frac{e^*_{\text{dual,s}} A_s}{e^*_{\text{dual,d}} A_d} > 0$. Therefore, the dynamic pattern in Proposition 1 requires $g_d \leq g_s$. Thus, we consider the following two cases.

1. Suppose $g_d = g_s$. Then,

$$p^*_{\text{dual,d}}(0) > p^*_{\text{dual,s}}(0) \Leftrightarrow 1 > \frac{e^*_{\text{dual,s}}[A_s + B_s]}{e^*_{\text{dual,d}}[A_d + B_d]}$$

   and

$$\lim_{t\to\infty} p^*_{\text{dual,d}}(t) < \lim_{t\to\infty} p^*_{\text{dual,s}}(t) \Leftrightarrow 1 < \frac{e^*_{\text{dual,s}} A_s}{e^*_{\text{dual,d}} A_d}.$$

   Combined, we require

$$\frac{e^*_{\text{dual,s}}[A_s + B_s]}{e^*_{\text{dual,d}}[A_d + B_d]} < 1 < \frac{e^*_{\text{dual,s}} A_s}{e^*_{\text{dual,d}} A_d} \Leftrightarrow \frac{A_s + B_s}{A_d + B_d} < \frac{e^*_{\text{dual,d}}}{e^*_{\text{dual,s}}} < \frac{A_s}{A_d}.$$

   With enough degrees of freedom, this condition can hold in principle.

2. Suppose $g_d < g_s$. Then, as in the case where firms that choose dual-class shares are forced to adopt single-class shares, to generate the dynamic pattern in Proposition 1, it must be that firms differ in at least two dimensions: growth rate (where the selected dual-class firm has lower), and some other unobserved characteristic.

INSERT FIGURE 3 HERE.

Appendix C Relations between Dual-Class Shares and Valuation

This appendix describes how we estimate the relation between dual-class shares and firm valuation, as measured by Tobin's q, building upon a large body of work that examines the relation between deviations from one share–one vote and firm valuation (e.g., Cronqvist and Nilsson (2003); Gompers et al. (2010)). We estimate the following equation:

$$q_{ijt} = \alpha_{jt} + \beta_1 Dual_{it} + \gamma' X_{it} + \epsilon_{ijt}, \tag{C.1}$$

where q_{ijt} is Tobin's q for firm i in three-digit SIC industry j in year t; α_{jt} represents industry j by year t fixed effects; $Dual_{it}$ is an indicator variable equal to one if firm i has dual-class shares in year t; X_{it} is a vector of control variables including log book assets, market leverage, R&D expenses scaled by lagged book assets, asset tangibility, sales growth rates, payout ratio, and ROA; and ϵ_{ijt} represents random errors clustered at the firm level.

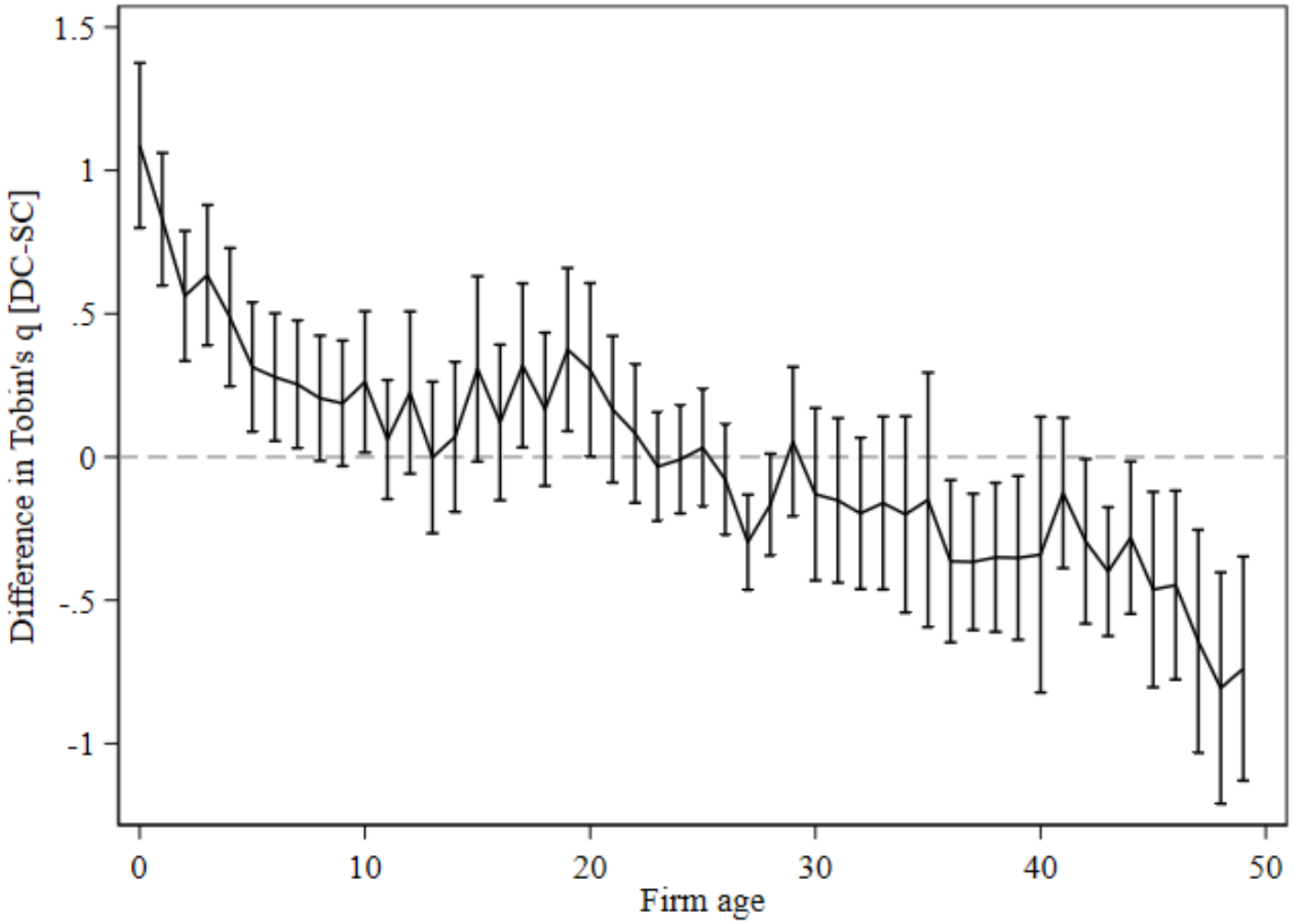
Table C1 presents the results of estimating Eq. (C.1). Figure C1 presents the results of estimating a version of Eq. (C.1) that interacts *Dual* with $\sum_{k=0}^{50} \mathbb{1}\big[\text{age} = k\big]$, where $\mathbb{1}\big[\text{age} = k\big]$ is an indicator equal to one if firm age $= k$ ($0 \leq k \leq 50$), and zero otherwise. We restrict the estimation sample to firms with ages less than or equal to 50 (sample size is too small to allow for precise estimation for ages over 50).

Table C1. Relation between Dual-Class Shares and Valuation in the Cross-Section

This table examines the average relation between dual-class shares and firm valuation as measured by Tobins q. "Dual" is an indicator variable equal to one if a firm-year has multiple classes of shares with differing voting rights, and zero otherwise. Definitions of the other variables are as in Table 1. Standard errors are adjusted for clustering at the firm level, and t-statistics are in parentheses. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)
Dependent Variable:	Tobin's q	
Dual	0.255***	0.272***
	(3.26)	(3.77)
Log book assets		0.016**
		(2.32)
Market leverage		-1.962***
		(-44.06)
R&D		6.606***
		(29.03)
Tangibility		-0.326***
		(-4.57)
Sales growth		0.141***
		(11.05)
Payout ratio		-2.435***
		(-15.24)
ROA		1.085***
		(12.41)
SIC3 × year fixed effects	Y	Y
R^2	0.219	0.312
Observations	160,702	160,702

Figure C1.
Relation between Dual-Class Shares and Valuation over Maturity



Notes: This figure plots the differences in Tobin's q between dual-class and single-class firms over their age. To construct the graph, we first estimate a version of Eq. (C.1) that interacts *Dual* with $\sum_{k=0}^{50} \mathbb{1}[\text{age} = k]$, where $\mathbb{1}[\text{age} = k]$ is an indicator equal to one if firm age $= k$ $(0 \leq k \leq 50)$, and zero otherwise. The line plots the coefficients on *Dual* $\times\, \mathbb{1}[\text{age} = k]$ along with 95% confidence intervals (in vertical lines).

Appendix D Variable definitions

This appendix provides definitions of firm-level financial variables from CRSP and Compustat.

- *Log book assets* is the natural log of book assets.
- *Age* is the number of years since an IPO (proxied by the first appearance in Compustat with a valid stock price).
- *Tobin's q* is the ratio of the market value of capital to the book value of capital, where the market value is market equity + book debt (proxy for market debt), and the book value is book equity + deferred taxes + book debt.
- *Tobin's q (GIM)* is the ratio of the market value of assets to the book value of assets, where the market value is book assets + market equity – book equity – deferred taxes.
- *Sales growth* is computed as the first difference of the natural log of sales.
- *ROA* is operating income before depreciation divided by lagged book assets.
- *Market leverage* is total debt divided by the sum of total debt and market equity.
- *R&D* is research and development expenses divided by lagged book assets.
- *Tangibility* is net property, plant, and equipment divided by book assets.
- *Payout ratio* is total payout, including dividends and repurchases divided by market equity.
- *Capex/assets* is capital expenditure divided by lagged book assets.
- *Employment growth* is computed as the first difference of the natural log of employment.
- *Cash flow* is net income plus depreciation and amortization divided by lagged book assets.

Appendix E Dual-class status change case studies

E.1 Dual-class Recapitalization

Alberto-Culver (1986 Recapitalization)

In April 1986, shareholders of the beauty product manufacturer Alberto-Culver approved a recapitalization of the firm's stock, creating a Class A stock, which controlled 10 votes per share, and a Class B stock, which controlled one vote per share. Founder Leonard H. Lavin served as both CEO and board chairman at this point, effectively controlling the firm.

In public statements, the company described the recapitalization as a mechanism to "keep a near majority of the voting power of the companys stock in the hands of the family of the company's founder" while also enabling the company to "raise additional equity or make acquisitions for stock without significantly diluting the voting power of shares held by the current holders."[28]

Dow Jones (1984 Recapitalization)

Dow Jones & Company shareholders approved a recapitalization in April 1984, creating a Class B common stock that would have 10 votes per share. Dow Jones had long been controlled by the Bancroft family, which the company sought to cement through the recapitalization. In a statement, the Dow Jones board said "the purpose [of the recapitalization] is to try to assure the long-term future operation of the Wall Street Journal and Down Jones' other publications and services under the same quasi-public-trust philosophy that Clarence Barron and his descendants [i.e., the Bancroft family] have followed throughout the company's history." Another widely-discussed reason for the recapitalization was preventing hostile takeovers of Dow Jones. [29]

E.2 Dual-class Unification

Alberto-Culver (2003 Unification)

In October 2003, beauty product manufacturer Alberto-Culver – the same firm described above as recapitalizing in 1986 – unified its stocks to a single class of common stock, with one vote per share. President and CEO Howard B. Bernick, the son-in-law of founder Leonard H. Lavin, suggested that converting to a single-class would create "alignment of the voting rights of the company's shareholders with their economic interests."[30]

Founder Lavin had stepped down as CEO in 1993, and stepped down as chairman in October 2004 at age 84. Lavin's successor Bernick ultimately left his CEO position in 2007, four years after the unification, marking an end to the Lavin-Bernick familys control of the company. Soon after the unification and Bernick's departure, Alberto-Culver sold off substantial parts of its business, and in 2010 Unilever bought the entire firm.[31] The sale concludes several years of speculations in the market that the company might not be able to maintain its independence while remaining competitive in an increasingly consolidating beauty-product industry.

Badger Meter (1999 Unification)

The firm Badger Meter, which produces water meters, announced in 1999 that it would convert all of its Class B stocks, which had ten votes per share, to common stock with a single vote per share. The previous year, longtime chairman James O. Wright (also the son of previous longtime CEO Charles Wright) stepped down as chairman of the firm. Wright had first followed his father as Badger Meter's CEO in 1952, serving as both CEO and chairman until 1981, and chairman thereafter until his retirement in 1998 at age 77. Shortly after

[28] "Alberto-Culver Co. Board Clears 4-for-3 Stock Split," Dow Jones News Service, February 24, 1986; "Founder grew Alberto Culver with'iron fist"', Reuters, September 27, 2010

[29] "Dow Jones Declares 50% Stock Dividend, Raises Dividend 20%," Dow Jones News Service, January 18, 1984; "Dow Jones Insures Family Control," The New York Times, Jan. 20, 1984

[30] "Alberto-Culver to Convert to Single Class of Stock," PR Newswire, October 22, 2003

[31] "Split Decision", Chicago Magazine, June 25, 2007; "Founder grew Alberto Culver with 'iron fist'", Reuters, September 27, 2010

Wright's departure from management, professional managers James Forbes and Richard Meeusen took office as chairman and CEO, respectively.

The company's statements described longstanding benefits of dual-class shares while the firm was relatively young and the founder was still in charge — but suggested that the value of these benefits no longer justified dual-class shares:

> Since 1986 [when Badger recapitalized with dual-class shares], the Class B common stock has provided a degree of control while we took the steps to build shareholder value. Over the past 13 years, the company has made significant investments in technology, new products, sales and marketing personnel, and manufacturing equipment and systems. As a result of this long-term growth strategy, sales and earnings have improved significantly, we have generated cash and we are financially strong. With this solid foundation in place, we believe it is in the best interest of our shareholders to revert to a single class of common stock that is generally more attractive to the investing public.[32]

[32] "Badger Meter Announces Elimination of Class B Stock and Stock Buyback," Business Wire, August 13, 1999

Appendix F Recapitalization and Unification Difference-in-Differences Event Studies, Robustness Tests

F.1 Matching

In Section 4.1, we estimate difference-in-difference (DD) sevent studies on the impact of dual-class recapitalization and stock unification on Tobin's q. Firms that switch from dual-class to single-class shares (or vice versa) may select into treatment on observables. While our event studies show a lack of pre-trends for both types of events, a related concern is that firms that recapitalize or unify have different observable characteristics than their respective control groups, hence the estimated treatment effects are biased. In fact, Panel A of Tables F1, F2, F3, and F4 show that some firm characteristics are not well-balanced in the event studies one year before treatment occurs.

To address this concern, we perform the event studies with matched samples. We match treated firms with a single control firm based on the Mahalanobis distance over a parsimonious set of imbalanced variables. For recapitalizations, we match on log(assets) and an indicator variable for positive R&D expenses, while for unifications we match on age and sales growth. We require that the Mahalanobis distance between the treated firm and the matched firm is in the 60th percentile of match distance among all potential controls. Additionally, we require that the matched firm have the same industry as the treated firm — either the same 3-digit SIC code, or the same 2-digit SIC code if we cannot generate a match with the 3-digit code. Finally, we trim matches based on imbalance in selected variables to achieve full balance across firm characteristics: for recapitalizations, we trim matches with imbalanced ROA and age, while for unifications, we trim matches with imbalanced leverage. Panel B of Tables F1-F4 show the resulting balance for the matched samples, with no statistically significant difference between treatment and control groups.

Using our matched sample, we estimate a version of Eq. (A1) in which we further include matched group fixed effects and cluster standard errors at the match group level, instead of the firm level, to account for correlation within matched groups.

Figure F2 plots the matched sample event study coefficients with Tobin's q as the outcome. The results are similar to the full sample event study results in Figure 4. As with the full sample, the pre-event estimates are not significantly different from zero in both event studies. Panel A shows that recapitalizing firms have an immediate increase in Tobin's q of 0.40, which declines and becomes consistently insignificant in year $t+6$ and beyond. This is very similar to the full sample estimates of recapitalization in Figure 4. Panel B of Figure F2 shows an immediate fall of 0.55 in Tobin's q after unification. This decline is largely statistically significant until year $t+6$, and in later years we estimate no difference in q.

Figure F3 plots the equivalent event study results with innovative output (measured as the inverse hyperbolic sine of patent citations) as the outcome. Again, the results after matching are qualitatively similar to the full sample results, as plotted in Figure 5. These results are reassuring that the results from Section 4.1 are not driven by differences in pre-event firm characteristics for unifying or recapitalizing firms.

Table F1. Recapitalizations: Matched and Full Sample Balance Tables, Tobin's q

This table presents descriptive statistics for control and treatment groups in two different recapitalization event studies: one with the full sample, and one with the matched sample. In both cases, the control group consists of single-class firms. Variables are defined as described in Table 1, and *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively, based on standard errors adjusted for clustering at the firm (Panel A) and match group (Panel B) levels.

Variable	Control Mean	Treated Mean	Diff. Mean
Panel A: Full Sample, Tobin's q			
Log(assets)	5.213	5.774	-0.561***
Age	12.901	13.647	-0.747
Tobin's q	1.869	1.962	-0.093
Market leverage	0.287	0.258	0.029**
R&D	0.017	0.034	-0.017***
Tangibility	0.330	0.332	-0.002
Sales growth	0.221	0.158	0.063
Payout ratio	0.022	0.028	-0.006**
Return on assets	0.181	0.141	0.041***
Observations	171	84,152	-
Panel B: Matched Sample, Tobin's q			
Log(assets)	5.182	5.208	-0.026
Age	12.529	11.471	1.057
Tobin's q	1.839	1.752	0.087
Market leverage	0.293	0.260	0.033
R&D	0.017	0.016	0.002
Tangibility	0.330	0.338	-0.008
Sales growth	0.168	0.153	0.015
Payout ratio	0.023	0.027	-0.004
Return on assets	0.177	0.162	0.015
Observations	157	157	-

Table F2. Unifications: Matched and Full Sample Balance Tables, Tobin's q

This table presents descriptive statistics for control and treatment groups in two different unification event studies: one with the full sample, and one with the matched sample. In both cases, the control group consists of dual-class firms. Variables are defined as described in Table 1, and *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively, based on standard errors adjusted for clustering at the firm (Panel A) and match group (Panel B) levels.

Variable	Control Mean	Treated Mean	Diff. Mean
Panel A: Full Sample, Tobin's q			
Log(assets)	6.197	6.432	-0.235
Age	12.514	16.761	-4.248***
Tobin's q	2.426	2.107	0.318
Market leverage	0.291	0.283	0.008
R&D	0.023	0.019	0.004
Tangibility	0.328	0.294	0.034*
Sales growth	0.187	0.125	0.062**
Payout ratio	0.028	0.029	-0.001
Return on assets	0.139	0.141	-0.002
Observations	185	6,539	-
Panel B: Matched Sample, Tobin's q			
Log(assets)	6.222	6.317	-0.095
Age	12.650	12.825	-0.175
Tobin's q	2.567	2.308	0.259
Market leverage	0.293	0.254	0.039
R&D	0.026	0.034	-0.008
Tangibility	0.315	0.290	0.025
Sales growth	0.182	0.136	0.046
Payout ratio	0.028	0.025	0.003
Return on assets	0.132	0.132	-0.001
Observations	160	160	-

Table F3. Recapitalizations: Matched and Full Sample Balance Tables, Innovative Output

This table presents descriptive statistics for control and treatment groups in two different recapitalization event studies: one with the full sample, and one with the matched sample. In both cases, the control group consists of single-class firms. Variables are defined as described in Table 1, and *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively, based on standard errors adjusted for clustering at the firm (Panel A) and match group (Panel B) levels.

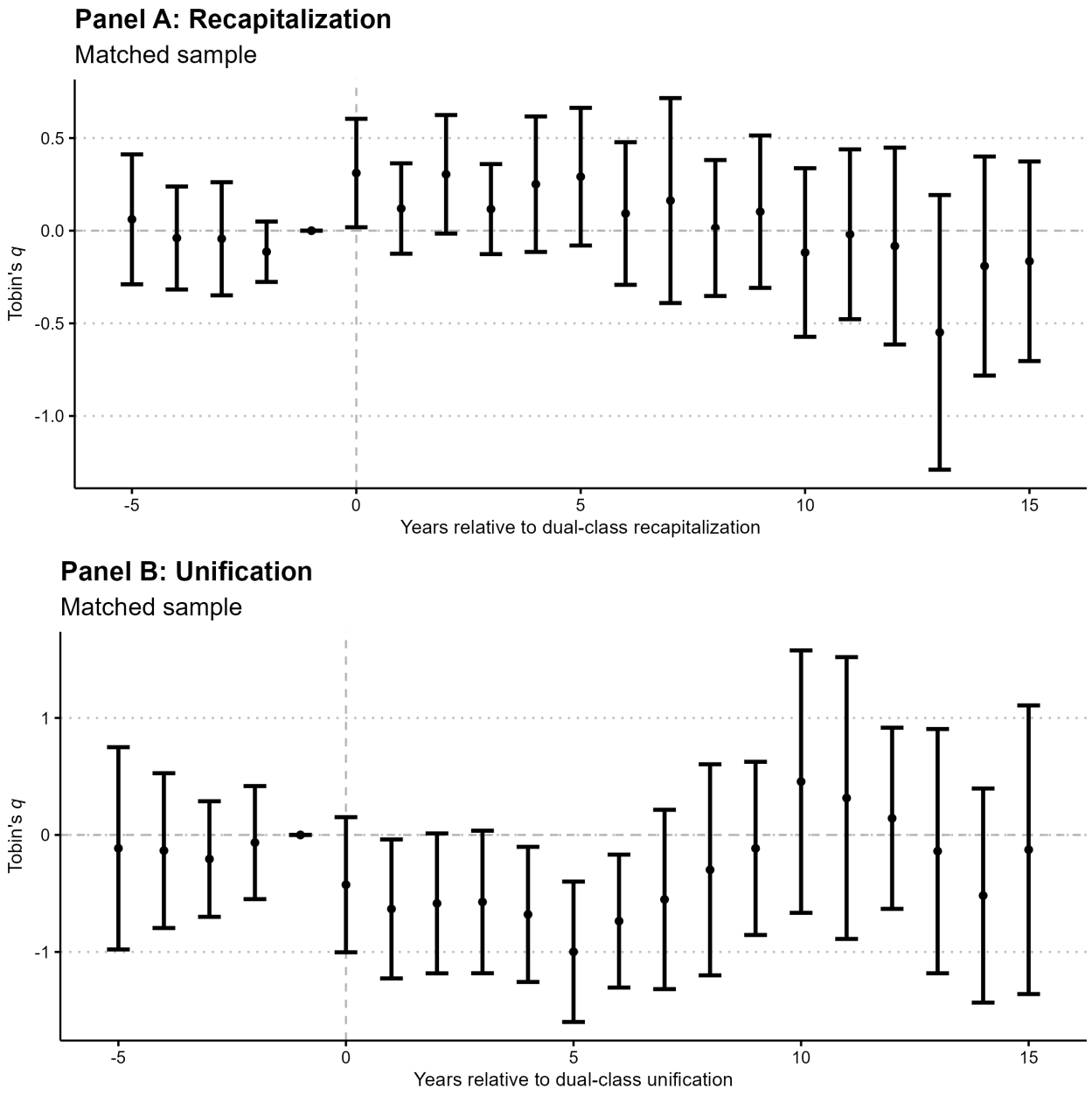
Variable	Control Mean	Treated Mean	Diff. Mean
Panel A: Full Sample, Innovative Output			
Log(assets)	5.264	5.836	-0.572***
Age	12.823	13.854	-1.032
Tobin's q	1.949	1.991	-0.042
Market leverage	0.295	0.261	0.034**
R&D	0.021	0.035	-0.013***
Tangibility	0.337	0.328	0.009
Sales growth	0.231	0.158	0.074
Payout ratio	0.023	0.028	-0.005*
Return on assets	0.178	0.138	0.040**
IHS citations	1.763	2.362	-0.599
Citations >0	0.333	0.420	-0.087
Observations	186	91,985	-
Panel B: Matched Sample, Innovative Output			
Log(assets)	5.240	5.256	-0.016
Age	12.676	11.585	1.091
Tobin's q	1.853	1.688	0.165
Market leverage	0.294	0.261	0.033
R&D	0.020	0.021	-0.000
Tangibility	0.338	0.341	-0.003
Sales growth	0.165	0.147	0.018
Payout ratio	0.024	0.026	-0.003
Return on assets	0.177	0.158	0.020
IHS citations	1.769	2.058	-0.289
Citations >0	0.335	0.335	0.000
Observations	176	176	-

Table F4. Unifications: Matched and Full Sample Balance Tables, Innovative Output

This table presents descriptive statistics for control and treatment groups in two different unification event studies: one with the full sample, and one with the matched sample. In both cases, the control group consists of dual-class firms. Variables are defined as described in Table 1, and *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively, based on standard errors adjusted for clustering at the firm (Panel A) and match group (Panel B) levels.

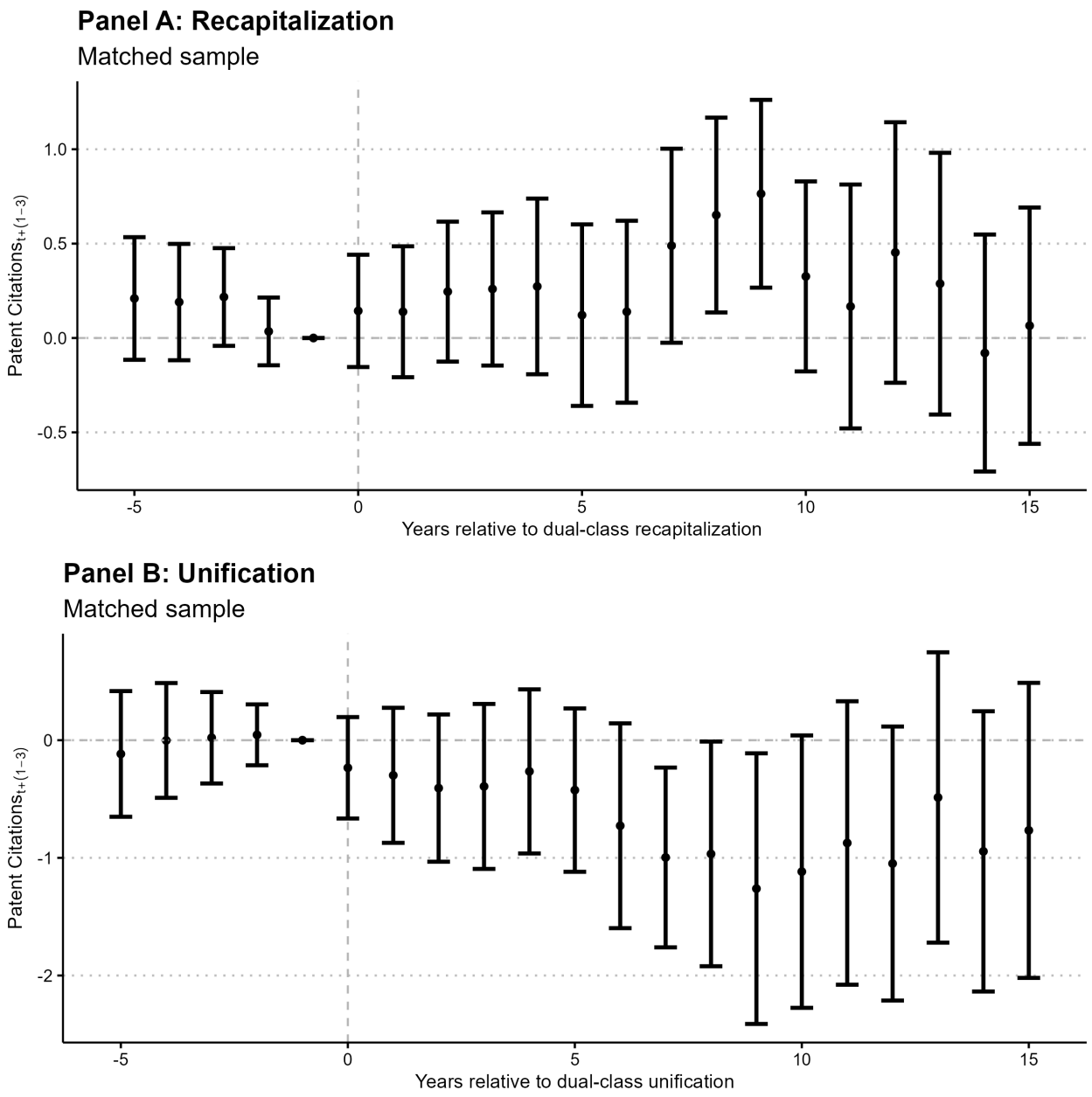
Variable	Control Mean	Treated Mean	Diff. Mean
Panel A: Full Sample, Innovative Output			
Log(assets)	6.263	6.472	-0.209
Age	11.746	16.720	-4.974***
Tobin's q	2.493	2.127	0.366*
Market leverage	0.321	0.285	0.036*
R&D	0.026	0.021	0.005
Tangibility	0.323	0.288	0.035**
Sales growth	0.187	0.129	0.058**
Payout ratio	0.026	0.029	-0.003
Return on assets	0.126	0.140	-0.014
IHS citations	2.027	1.875	0.152
Citations >0	0.388	0.378	0.011
Observations	224	7,243	-
Panel B: Matched Sample, Innovative Output			
Log(assets)	6.260	6.377	-0.117
Age	11.968	12.126	-0.158
Tobin's q	2.592	2.558	0.035
Market leverage	0.302	0.269	0.033
R&D	0.030	0.033	-0.003
Tangibility	0.309	0.287	0.021
Sales growth	0.185	0.146	0.040
Payout ratio	0.026	0.021	0.006
Return on assets	0.120	0.136	-0.016
IHS citations	1.769	2.058	-0.289
Citations >0	0.335	0.335	0.000
Observations	190	190	-

Figure F2.
Matched Sample Event Studies: *q*



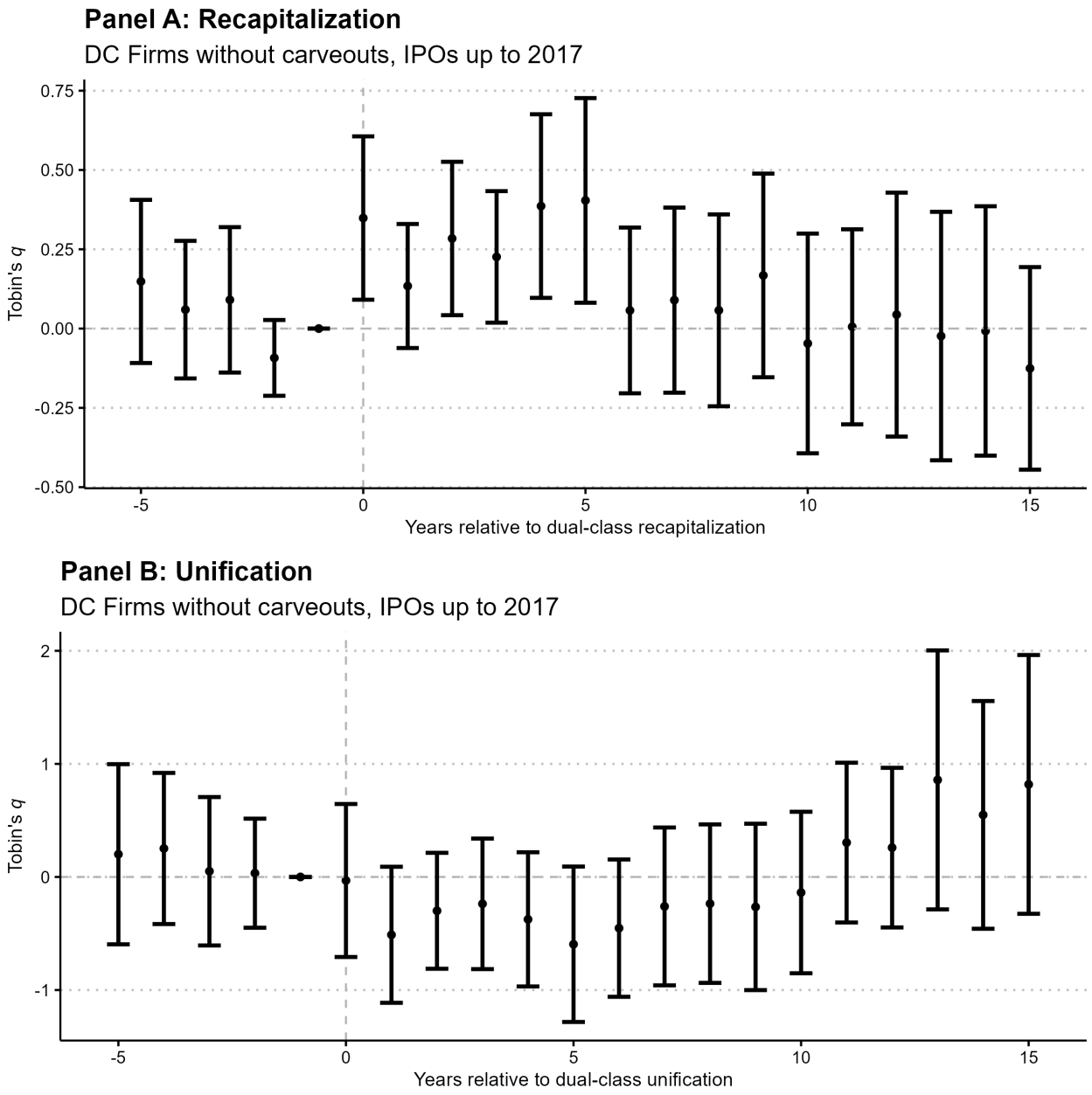
Notes: This is the analogous plot to Figure 4, using our matched sample and a version of Eq. (A1).

Figure F3.
Matched Sample Event Studies: Innovative Output



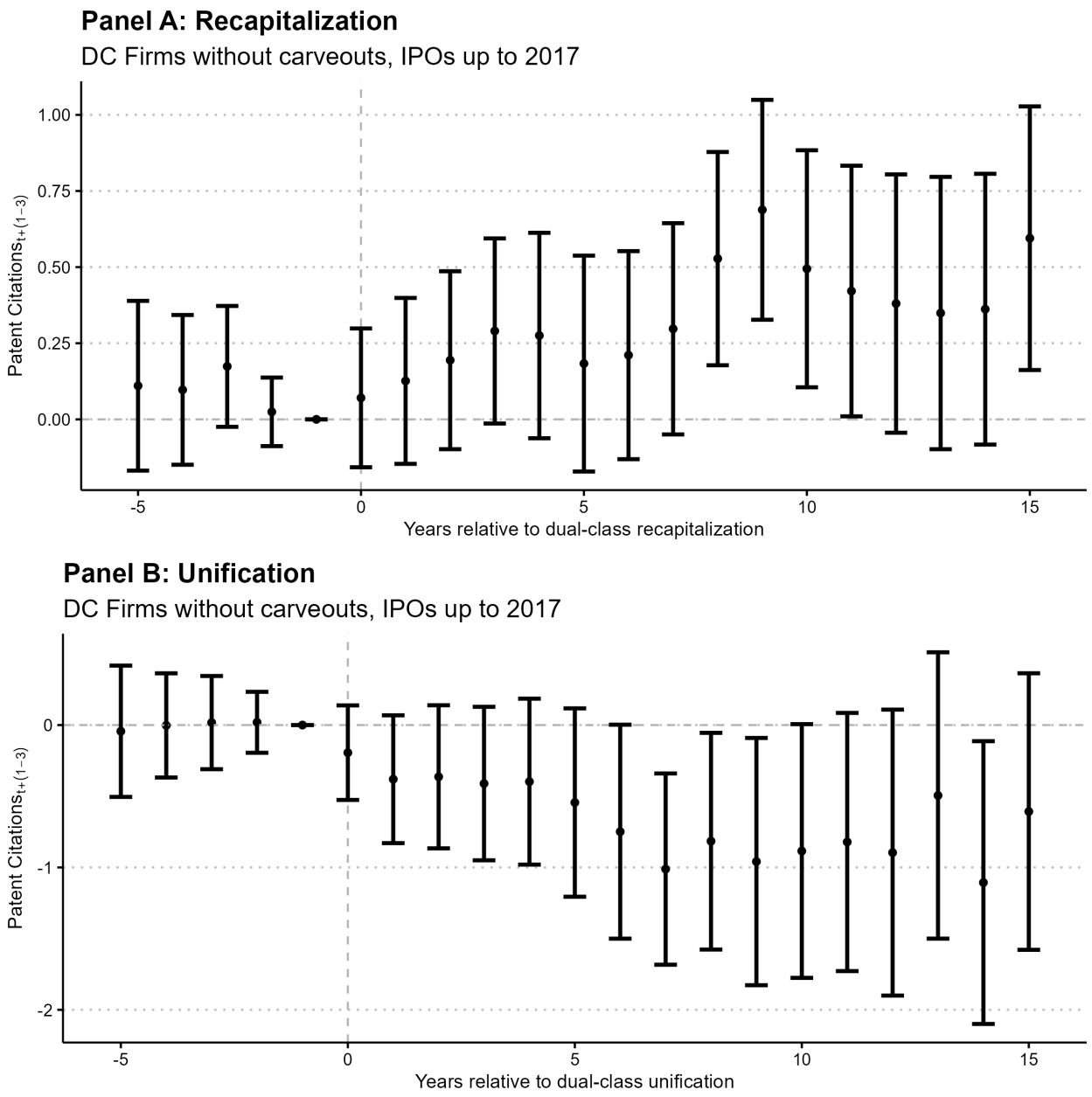
Notes: This is the analogous plot to Figure 5, using our matched sample and a version of Eq. (A1).

Figure F4.
Tobin's *q*: Robustness to carveout exclusion



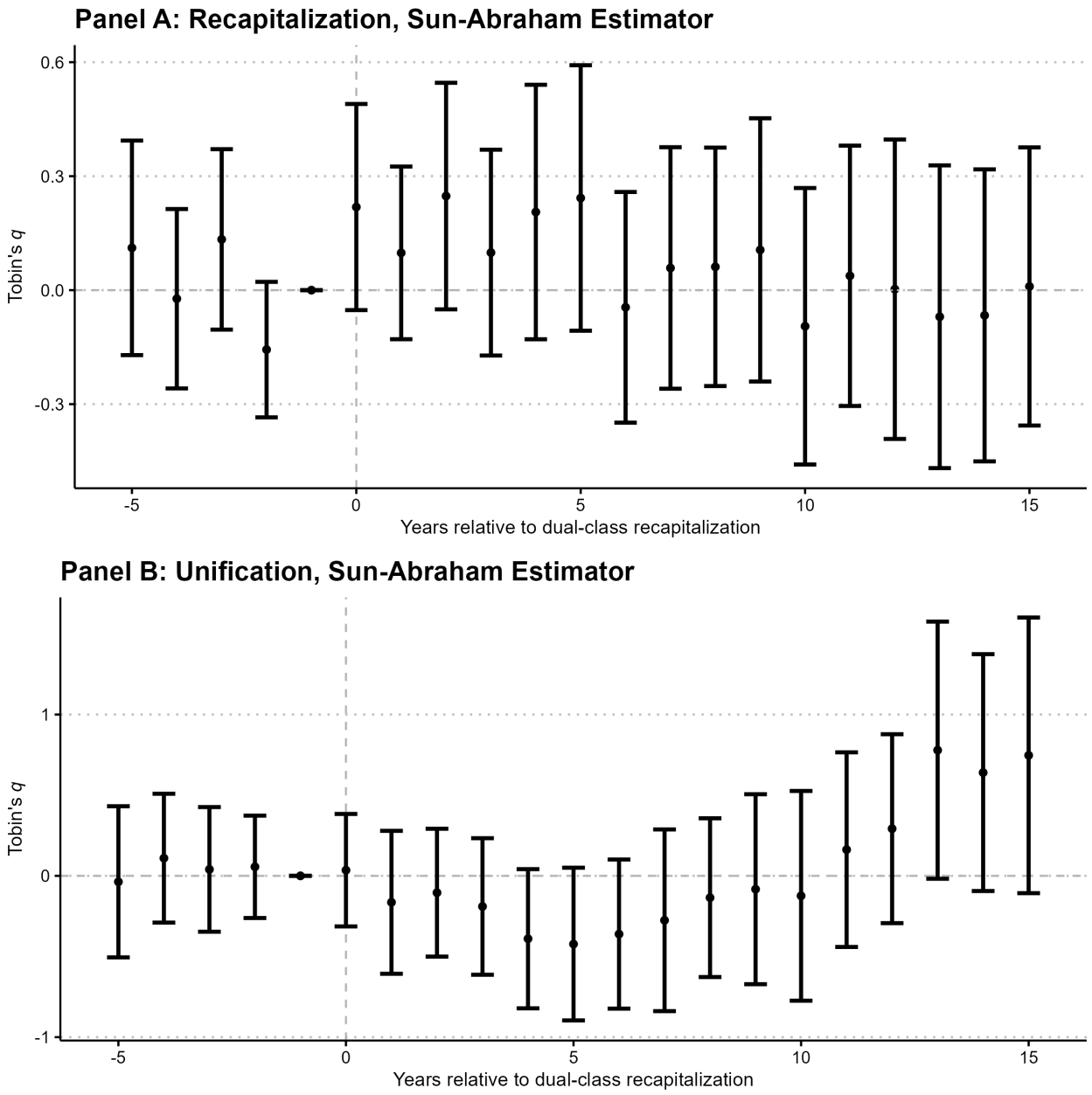
Notes: This figure plots the dynamic difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on Tobin's *q*, removing dual-class firms with a carveout. Because our carevout data only covers firms with IPOs through 2017, we drop all firms with IPOs after 2017 in this regression.

Figure F5.
Innovative Output: Robustness to carveout exclusion



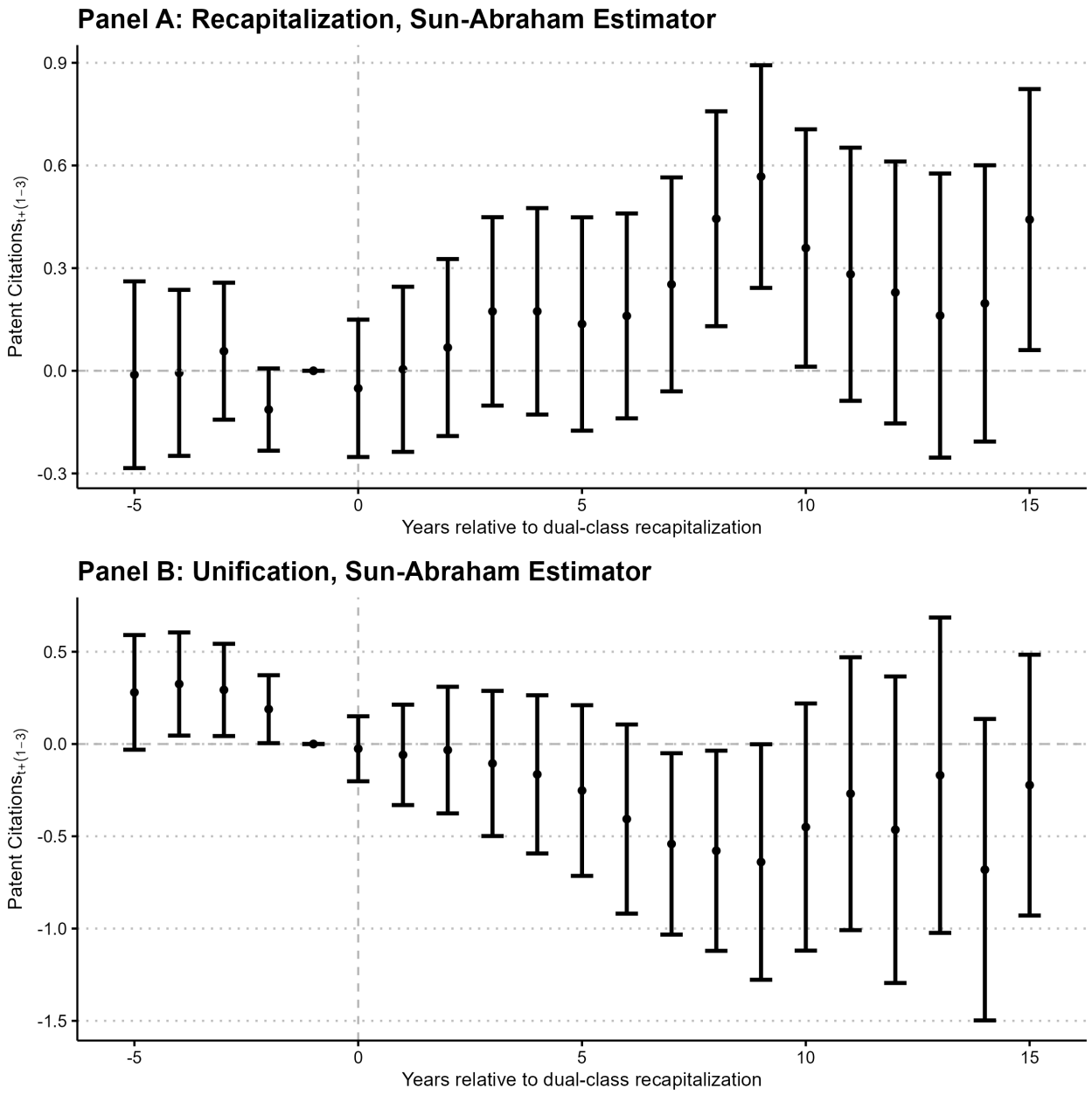
Notes: This figure plots the dynamic difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on innovative output, removing dual-class firms with a carveout. Because our carevout data only covers firms with IPOs through 2017, we drop all firms with IPOs after 2017 in this regression.

Figure F6.
Tobin's *q*: Robustness to Sun-Abraham estimator



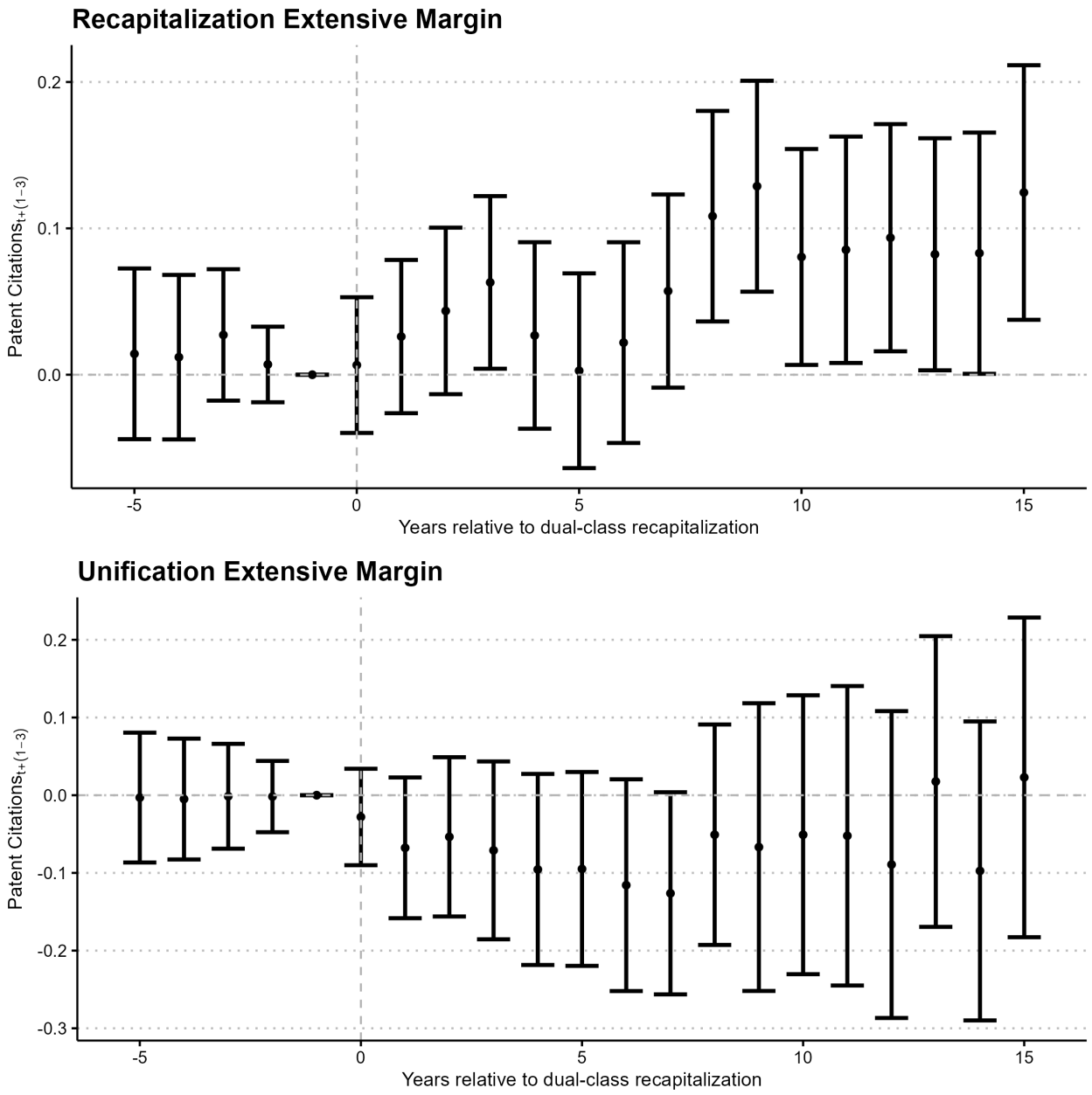
Notes: This figure plots the dynamic difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on Tobin's *q*, using the event-study estimator from Sun and Abraham (2021).

Figure F7.
Innovative Output: Robustness to Sun-Abraham estimator



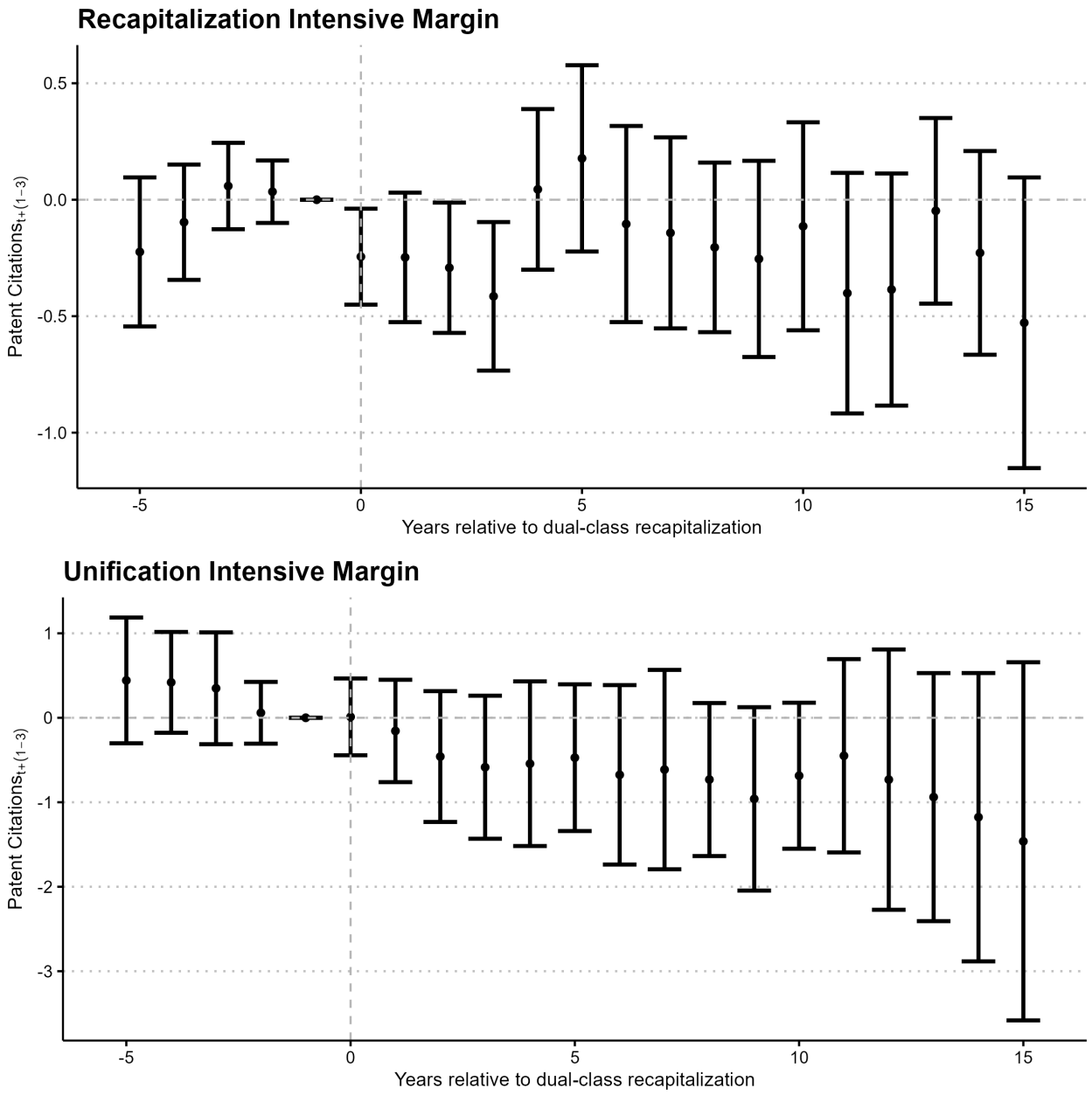
Notes: This figure plots the dynamic difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on innovative output, using the event-study estimator from Sun and Abraham (2021).

Figure F8.
Innovative Output: Extensive Margin



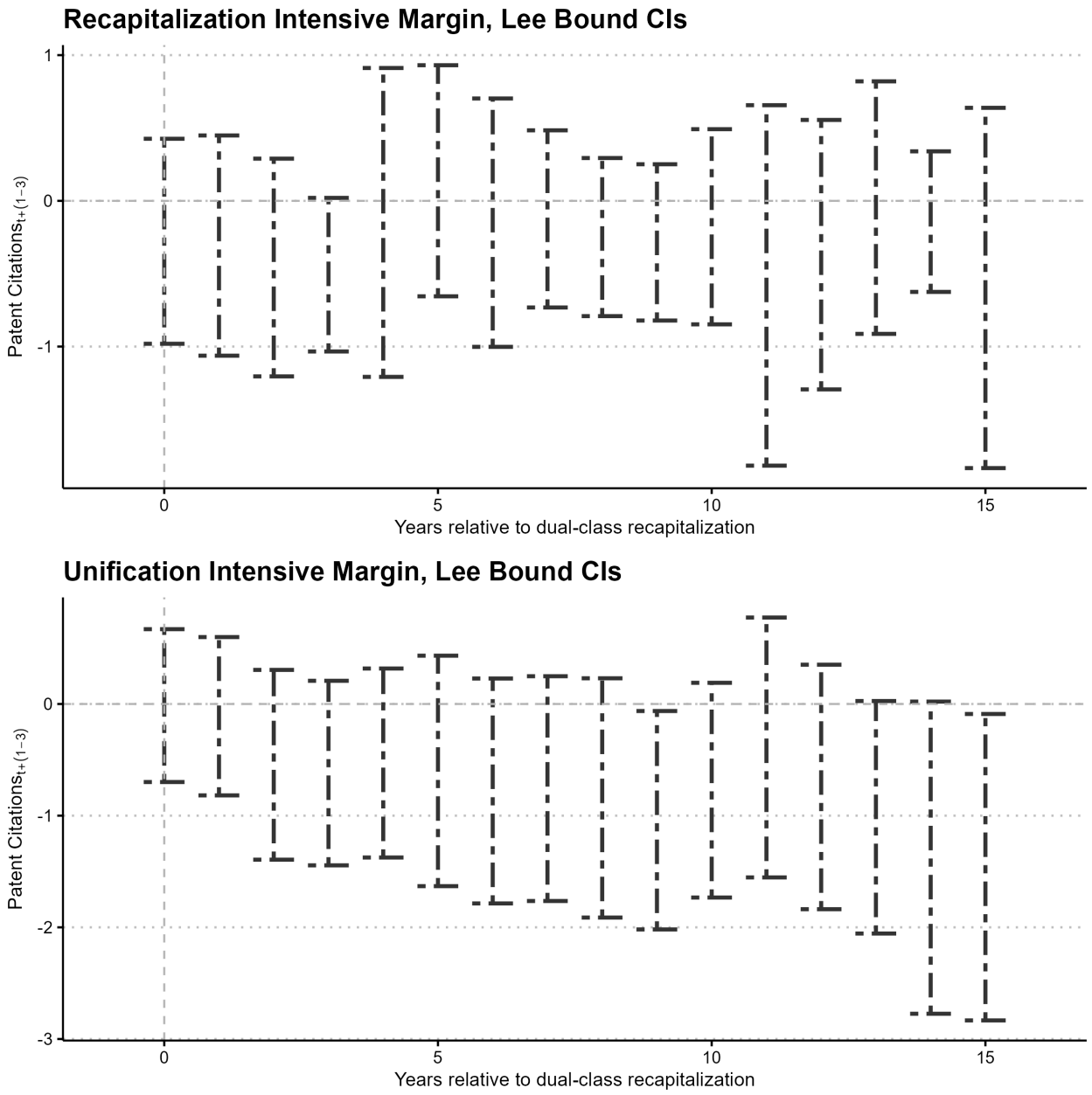
Notes: This figure plots the dynamic difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on innovative output, showing only the extensive margin of citations (i.e., an indicator variable for whether a firm has patent citations in a given year).

Figure F9.
Innovative Output: Intensive Margin



Notes: This figure plots the dynamic difference-in-differences estimates of the effect of dual-class recapitalization (Panel A) and stock unification (Panel B) on innovative output, showing only the intensive margin of citations as measured by log(cites).

Figure F10.
Innovative Output: Lee Bound confidence intervals Intensive Margin



Notes: This figure estimates bounds on the intensive margin of citations, following the trimming procedure detailed in Lee (2009). Within each event-year, we drop firms from either the treatment group or the control group until each group has the same portion of firms with non-zero citations. This generates bounds by either trimming the firms with the highest or lowest citations in a given event-year. We plot only the "widest" version of the standard errors: that is, the upper bound relative to the upward-trimmed estimate, and the lower-bound relative to the downward-trimmed estimate.

Appendix G Data Appendix

G.1 Identifying dual-class firms

We identify candidate dual-class firms by comparing a given firm's number of shares outstanding obtained from CRSP and Compustat. CRSP provides the number of shares outstanding at the security level (i.e., for each class of shares), whereas Compustat provides the corresponding number at the firm level (i.e., the sum across all classes of shares issued by the firm). Thus, a significant difference between the two numbers indicates that the firm might have multiple classes of shares. If the numbers of shares from CRSP and Compustat differ by more than 2%, we place those firm-years into a candidate set (following the approach introduced in Gompers, Ishii, and Metrick (2010)). We supplement this set with a data set of dual-class IPOs from Jay Ritter's website. We then hand-check whether firms in this candidate sample have multiple share classes and for confirmed dual-class firms, we collect information on the number of votes and shares outstanding with two data sources. For 1994–2022, we use the annual report (Form 10-K) and proxy statement (DEF 14A) taken from the SEC's EDGAR database for each firm-year, except for those covered by the GIM data set. Specifically, we determine whether a firm has dual-class shares by examining descriptions of voting rights and shares outstanding for multiple classes of common shares. The SEC EDGAR does not provide information in electronic form for 1971–93, and so we use Moody's Manuals (the Capital Stock section) during the period.

A notable fraction of firms with more than one class of common shares has the same number of votes across classes (e.g., one vote per share). We therefore determine whether these firms have disproportional voting rights between share classes by manually examining security filings from the EDGAR (such as the DEF 14A, 424B, S-1, 10-K/Q) and Moody's. We find that there are three possible reasons that these firms have multiple share classes, the first two of which represent differential voting rights:

(1) Different classes have differing voting rights for director election. In a typical case in this category, one class has the right to elect two-thirds of the directors, and the other class, one-third. For these cases, we define the class with greater director election right as superior. There are 435 dual-class firm-years in this category.

(2) Some firms use specific formulas to calculate the numbers of votes for different classes. A typical example involves a "superior" class of common stocks with the number of votes per share equal to the number of "holdings units" in a limited liability company that a small group of shareholders own. An "inferior" class typically carries one vote per share. These cases are rare, with only 24 firm-years involved.

(3) The final category includes cases for which a dual-class structure appears to be set up for reasons other than giving disproportional voting rights. For example, Triple-S Management has issued Class B common stocks as a capital asset for tax purposes. In other cases, non-US firms restrict ownership of one class of common stocks to citizens of specific countries.[33] Given that these cases do not presentation a deviation from one share–one vote, we define them as "non-dual-class" and drop them from the analysis.[34] This category includes 202 firm-years.

[33] For example, Grupo Iusacell, S.A. de C.V. restricts its Class B common stock ownership to Mexican citizens only, whereas their Class A common stock has no ownership restrictions.

[34] An alternative is to treat these firms as *de facto* single-class firms, which gives a very similar result.