



# Test power properties of within-firm estimators of ownership and board-related explanatory variables with low time variation

Maria Boutchkova<sup>1</sup> · Diego Cueto<sup>2</sup> · Angelica Gonzalez<sup>1</sup>

Accepted: 5 May 2022 / Published online: 16 July 2022  
© The Author(s) 2022

## Abstract

Corporate governance research is often limited in its ability to employ within-firm estimators, which address time-invariant endogeneity, when the variables of interest exhibit low time variation (for example, ownership and board independence). The problem is further exacerbated if data for multiple points in time needs to be hand-collected. We offer simulation-based methodological guidance to improve the statistical power of within-firm estimators in the presence of low time variation. We illustrate the usefulness of our simulation results by replicating two influential studies on ownership and board independence and extending them with a within-firm estimator. Based on widely used databases as well as a novel granular database, we document the different degrees and nature of time variation of ownership and board independence across jurisdictions and subgroups by listed status, family control and complexity of ownership structure. Researchers can use our findings to optimize the hand-collection and pre-processing of governance data and thereby increase statistical power and/or to distinguish whether lack of significance is due to low time variation as opposed to absence of a true relationship between their governance variable of interest and the respective outcome.

**Keywords** Controlling ownership · Ultimate cash flow rights · Control rights · Ownership wedge · Board independence · Low time variation · Family firms · Within-firm estimator · Pyramid structures

**JEL Classification** G15 · G32

---

✉ Maria Boutchkova  
maria.boutchkova@ed.ac.uk

Diego Cueto  
dcueto@esan.edu.pe

Angelica Gonzalez  
angelica.gonzalez@ed.ac.uk

<sup>1</sup> University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh EH8 9JS, UK

<sup>2</sup> ESAN, Graduate School of Business, Av. Alonso de Molina 1652, Lima 33, Lima, Peru

## 1 Introduction

In studies explaining firm outcomes with governance variables, researchers commonly assume that they exhibit very little time variation. However, unless there is time variation in the data, there are limited ways to strengthen a claim of a causal effect, for example by using a within-firm estimator. We report power properties of the within-firm estimator for different degrees of time variation, length of time series and frequency of sampling of ownership and board independence. Our results help researchers determine whether they have sufficient power in a given empirical setting and how to increase it.

Typical research questions in finance and accounting consider how the controlling ownership stakes of different types of investors affect firm decisions (Banerjee and Homroy 2018; Larrain and Francisco 2013), how outcomes like cost of capital or performance depend on different degrees of conflict of interest among controlling stakeholders (Bertrand et al. 2002; Lin et al. 2011), or how board independence relates to performance or CEO turnover (Coles et al. 2008; Graham et al. 2020). Establishing a causal relationship between governance and these outcomes is very difficult because, among many other reasons, accounting for unquantifiable confounding factors is rarely possible. For example, a significant relationship between a governance variable and performance may not truly exist, but may appear as a result of unobservable firm characteristics (Lins 2003; Bennesen and Nielsen 2010). In a survey of the board literature, Hermalin and Weisbach (2003) argue that most board research as of that time had failed to address the endogeneity in board composition. More recent survey and editorial articles continue to highlight the problem of reliably addressing endogeneity in governance research (Adams 2017; Edmans and Holderness 2017). The state-of-the-art best practice employs empirical designs based on natural experiments. However, they are rare, or, if present, the setting is generally imperfect and requires additional strengthening tests that are often based on time variation (Atanasov and Black 2016, 2020).

From an empirical design perspective, the problem of insufficient time variation in ownership is exacerbated by the limitations of vendor-constructed databases that do not provide all ownership links and require extensive pre-processing and verification (Holderness 2009; Dlugosz et al. 2006), and/or the necessity to hand collect data. Board related variables bring fewer data processing challenges, but they are only easily available for listed firms and similarly may change slowly over time (Black et al. 2017). To facilitate research centered on governance variables with low time variation and/or limited availability, our solution consists of methodological guidance based on simulations and verified using real data from a variety of sources (including a rare granular ownership data source). Our simulations generate artificial data series that satisfy a number of constraints to imitate the underlying economic processes behind governance variables. The simulation results provide guidance on the amount of time variation required for sufficient statistical power of within-firm estimators, depending on a variety of institutional features. Our findings are useful to researchers in at least three ways. First, when a project requires the hand collecting of governance data with low variation, by following our simulation results a researcher can efficiently collect non-consecutive data points to ensure sufficient time variation. For example, if a researcher has hand collected two time observations of ownership (board independence) data four years apart for a sample of firms of interest and less than 60% (55%) of them change between the two periods, she could decide whether it is worthwhile collecting one additional time observation to increase the power properties of her tests. This additional time period would reduce the required proportion of changing firms

to around 42% (45%) to detect statistical significance if it was present. Second, suppose a researcher uses one of the widely available databases, for example board data from BoardEx or Execucomp, or any of the corporate ownership products of Bureau van Dijk (now part of Moody's), Thomson Reuters (now Refinitiv) or Standard & Poors. She can use our results to check whether the amount of time variation in her sample of interest is sufficient. This is important because, if she finds no significant relationship, knowing whether this may be due to the lack of time variation helps her determine the choice of additional tests to employ in the project. Or alternatively, if time variation is sufficient, helps strengthen her conclusions that a relationship is unlikely to exist. Third, if a researcher decides not to use within-firm estimators and substantiates this decision based on the lack of time variation in the data, our benchmark results help add a degree of formality to this reasoning.

To show the usefulness of our simulation findings and as a proof of concept, we replicate regressions from two existing studies—Lin et al. (2013) and Coles et al. (2008). Lin et al. (2013) examine how the divergence between cash flow and control rights of the controlling owner affects the proportion of public debt a firm holds. Coles et al. (2008) analyze the relationship between board independence, complexity of expertise required by a firm's board and Tobin's Q. Both studies originally do not use a within-firm estimator for their baseline results, motivated by a lack of time variation. We show modifications to their design, guided by our simulation results, where time variation is sufficient, and the more reliable within-firm estimator is able to detect statistically significant relationships as a result of the improved power. For example, in the case of Lin et al. (2013) using firms from the same jurisdiction as in the original paper with ten consecutive time observations (whereas they have on average four per firm) allows the within-firm estimator to detect statistical significance. Importantly, if having ten consecutive time observations is prohibitive, we show that sampling five years apart and having four time observations would be sufficient to find statistically significant results. The conclusions change when time variation is especially low (for example, for US firms). In this case neither more time observations per firm, nor sampling with gaps allows the detection of significance even though it exists.

Lastly, we provide evidence of the different patterns of time variation in governance variables. We employ commonly used data sources (BoardEx, Capital IQ, Thomson Reuters, CSMAR) plus a unique database (ICO by Statistics Canada) that provides a long time series, complete corporate structures and exact direct ownership stakes for all global multinationals that have any subsidiary operating in Canada. We document that the degree of time variation in ownership and board independence differs substantially depending on jurisdiction of control, listed status, family ownership and ownership structure complexity. Therefore, the stability of governance variables cannot always be taken for granted and should be tested for the particular dataset in question, as we do here. For example, the time variation in board independence is greater for non-listed than listed firms and for Anglo-Saxon<sup>1</sup> than US firms. Time variation in ownership is indeed modest for multinational firms controlled from Anglo-Saxon jurisdictions, and it is likely the case for Continental Europe; however, we find a considerable time variation in the population of Canadian

---

<sup>1</sup> By Anglo-Saxon jurisdictions we mean Australia, Ireland, New Zealand, UK and US (Esping-Andersen 1990). Canada is also considered an Anglo-Saxon country and we treat it as such in the board independence analysis. However, in our ownership analysis it represents a separate sample partly due to its hybrid common-civil law system, and partly because it covers the statistical population of firms as opposed to a subset of large multinationals. In the case of board independence, we treat the US as a separate subgroup for similar reasons of much wider data coverage in BoardEx.

firms. We further find that there is greater time variation in the ownership of family versus non-family-controlled firms belonging to pyramid structures, regardless of their control jurisdiction. In addition, the prevalence of complex pyramidal corporate structures is greater among family firms. Although the distinct features of family owners have been studied extensively (e.g., Almeida and Wolfenzon 2006; Anderson and Reeb 2003; Belenzon et al. 2019), the fact that they tend to adjust the ownership structures of their firms more frequently is a new finding. Across size groupings, unsurprisingly, firms belonging to larger structures change more frequently than in smaller structures. This greater variability for family firms is preserved when grouping by (i.e., controlling for) size and jurisdiction.

We further document the following patterns of time variation: (1) the averages of ownership and board independence change little over time; (2) the average proportion of firms that change over two consecutive years is relatively small for ownership and larger for board independence of listed firms; and (3) extended periods of stability are common for ownership and less so for board independence of listed firms, and it takes several years for a considerable proportion of firms to leave a period of stability<sup>2</sup>; (4) the standard deviation of ownership over time is greater for family firms and pyramids in univariate and multivariate analyses; (5) ANOVA and persistence regressions confirm the differing patterns of stability by family control, jurisdiction and ownership structure complexity.

The importance of considering corporate ownership and board independence in empirical research cannot be overstated. Corporate ownership is central to our understanding of firm decisions like growth (Belenzon et al. 2019), innovation (Aghion et al. 2013), mergers and acquisitions (Basu et al. 2009; Boateng et al. 2017), internationalization (Singla et al. 2017) and financing (Lin et al. 2011, 2013). Ownership and board independence shape incentives, preferences, and short-term versus long-term orientation of managers and boards in listed firms (Siegel and Choudhury 2012; Banerjee and Homroy 2018; Ellis et al. 2020) and of owner-managers in family firms (Deephouse and Jaskiewicz 2013); tip the balance of power among managers, and majority and minority owners (Aguilera and Crespi-Cladera 2016; Guo and Masulis 2015); allow for unique capabilities that different owners and directors bring to their firms (Rabbiosi et al. 2019; Edmans and Holderness 2017; Kim and Starks 2015), etc. The importance of quantifying governance variables more precisely (including ownership and board independence) is underscored by the ever-rising use of Environmental, Governance and Social (ESG) considerations in investment decision-making. They are already formally regulated in the UK and across Europe and gradually gaining importance in the US (Eccles and Klimenko 2019; Mooney 2018).<sup>3</sup> In turn, the presence of sizable controlling stakes versus more diversified holdings of different types of institutional investors has implications for monitoring incentives, shareholder activism and adjustments in executive pay structures, which ultimately affect firm decision making (Azar et al. 2018; Flammer and Bansal 2017; Lardon et al. 2019).

The finance and accounting literature abounds with studies that exploit key strategic changes to answer questions about firm outcomes. These settings often involve ownership structure adjustments (M&As, reorganizations, vertical and horizontal integration, new

<sup>2</sup> Stability is measured via the concept of “stable regime”, defined as a minimum length of time when no change occurs in a variable beyond a pre-specified interval, as in DeAngelo and Roll (2015).

<sup>3</sup> Ownership and board independence most obviously fit the governance aspect of ESG, but have strong relevance to the social aspect in terms of institutional development and social equity in control over productive assets. They are also relevant to the environmental aspect of ESG in terms of firm incentives to produce positive or negative environmental externalities.

product and market entry) and board composition changes (independence, minority representation, etc.). Our results open the way for time-variation-based methods of analysis where ownership or board independence is a key variable of interest. We help researchers understand the reason behind a finding of no significance of a within-firm estimator, highlight settings where time variation is likely higher and show researchers how to increase the power of within-firm estimators without necessarily having to collect a large number of time observations. This way we answer the call in several recent editorials and review articles highlighting the importance of addressing endogeneity if the quality of business research is to be improved (Semadeni et al. 2014; Reeb et al. 2012; Abdallah et al. 2015; Holmes et al. 2018).

The rest of the study is organized as follows: Sect. 2 outlines the background and justifies our simulation design; Sect. 3 describes our data sources and replicating regressions; Sect. 4 consists of the detailed analysis of time variation in ownership and board independence variables; and Sect. 5 concludes.

## 2 Background and simulation design

Corporate ownership and board composition governance variables exhibit little time variation, which limits their potential for clean causal inference. For example, the earliest ownership studies constructed hand-collected samples which were typically static because data was not available in electronic form even for listed firms (La Porta et al. 1999; Claessens et al. 2000; Faccio and Lang 2002, among many).<sup>4</sup> More recently, electronic sources improved coverage and technological advancements have allowed construction of time series datasets, but they have rarely been exploited to address endogeneity. For example, Villalonga and Amit (2009, p. 3083) explain: “Because these variables exhibit very little time-series variation, we abstain from using firm fixed effects,” while Banaliev and Eddleston (2011, p. 1065) acknowledge that “the standard fixed effects estimation is infeasible”. More recently, Pronobis and Schaeuble (2020, footnote 21) write “A closer look to the dataset reveals that our foreign ownership variables are relatively sticky. That makes it difficult to get a sufficient number of observations for which we can identify changes in foreign ownership.” On this basis, they acknowledge the subsequent lack of power: “Therefore, the results provided by estimating our change model have only limited explanatory power”. Similar examples abound in the literature on board independence. Coles et al. (2008) argue that including firm fixed effects is not appropriate in their setting because most of the variation in board size arises in the cross-section instead of in the time series. Choi et al. (2007) report that their results on the value of outside directors disappear once they introduce firm fixed effects. Black and Kim (2012) highlight the advantages of their Korean data set, as this provides enough time variation in the variable “outside directors” to make within-estimation feasible. Wintoki et al. (2012, p. 591) acknowledge that “[b]oard structure is highly persistent [, which] can reduce the power of any panel data estimator”. More recently, Frye et al. (2021) refer to the stickiness of board structure and in all their specifications they only use industry fixed effects, rather than a within-firm estimator.

<sup>4</sup> In the online appendix we summarize this literature in tabular form, covering seminal works from the empirical strategy, finance, economics and international business literature with a focus on controlling ownership published over the last thirty years. The papers are selected on the basis of high citation count and/or being recent.

Faced with this lack of time variation, a few of the papers employ a random effects estimation which does not rely exclusively on the time dimension of the data. However, it is not appropriate to consider this estimator as an alternative to fixed effects because the underlying assumption of the random effects estimator is exogeneity.<sup>5</sup> Empirically, Black et al. (2014) perform extensive testing and reject the equivalence of the fixed effects and random effects estimators in the context of multi-country governance studies. The fixed effects and first difference estimators are the ones alleviating time invariant endogeneity.<sup>6</sup> As Reeb et al. (2012, p. 214) highlight: “Evidence of causal relation with unit level fixed effects can be quite compelling [...]” Bliese et al. (2019, p. 9) go as far as to refer to the within-firm estimator as the “gold standard to which results from other analytic options are compared”.

There is growing scientific consensus that the most reliable causal inference methodologies are shock-based designs often referred to as quasi-natural experiments (Atanasov and Black 2016, 2020). These approaches often require at least two time points—before and after the shock—for a difference-in-difference (DiD) design (for example, Aguilera et al. 2017; Liu et al. 2015). Importantly, even when an instrument and/or a shock is available, the design may be imperfect and require additional strengthening measures to rule out alternative explanations (Atanasov and Black 2020). For example, one could include additional covariates and/or firm fixed effects in a DiD design to get closer to satisfying the parallel trends assumption. Furthermore, routinely required tests to support causal inference findings are placebo tests to show that a significant finding around the shock disappears around a different date. However, the absence of significance of the placebo test could be a mechanical consequence of the lack of time variation in the key explanatory variable and cannot be used as supporting evidence for the DiD findings.

Given the small amount of time variation in ownership and board independence and/or limited data availability, some researchers have opted to use data with gaps. For example, Franks et al. (2012) focus on family control at two points in time ten years apart, Basu et al. (2017) study blocks of any kind of owners five years apart, while Lin et al. (2011, 2012) construct ultimate ownership for four points in time two and three years apart. Wintoki et al. (2012) sample board independence every two years, while Boone et al. (2007) and Linck et al. (2008) sample every three years. There is no formal guidance in the literature on the gap length that ensures sufficient time variation in the variable of interest. Even in cases where electronic databases make uninterrupted time series readily available, extensive pre-processing and cleaning may preclude the researcher from examining every single consecutive observation. In addition, in the case of ownership, these series may not

<sup>5</sup> The key assumption for the validity of the random effects estimator is  $cov(x_{it}, a_i) = 0$ , i.e.,  $x$  should not be correlated with the unobservable firm-specific heterogeneity  $a_i$  (Wooldridge 2010). We thank an anonymous reviewer for suggesting the correlated random effects (CRE) estimator to us. If we were to treat ownership or board independence as one of the time-varying explanatory variables in the CRE setting, Wooldridge (2019) shows that the coefficient vector  $\hat{\beta}$  is identical to the within-firm estimator in the case of an unbalanced panel, while Mundlak (1978) has originally shown this result for the case of a balanced panel. Therefore, the simulation results would be mathematically equivalent if we were to use CRE instead of the within-firm estimator.

<sup>6</sup> There is one setting, where within-firm estimators are inappropriate due to its institutional specificity: CEO ownership. As Adams (2017) explains, using a fixed effect estimator to test the effect of CEO ownership on firm value will result in estimating the effect of CEO turnover instead of managerial incentives, because big changes in CEO ownership within a firm arise primarily from changes in the CEO.

reflect the complete set of ownership links, and additional data-collection is likely to be necessary.<sup>7</sup>

The first article to employ the fixed effects estimator in the study of managerial ownership and firm performance, using consecutive observations from a widely available database, found no significance (Himmelberg et al. 1999). However, one more recent extension (Kim and Lu 2011) and a comprehensive in-depth analysis (Fabisik et al. 2021), using a longer time period with greater time variation as a result of stock-based executive compensation, identified a significant relationship. Similarly, Graham et al. (2020) document that although there is large persistence in board independence, over longer horizons there are significant within-firm changes in board structure. This invites the question of the amount of time variation sufficient to identify a relationship, if it exists, as well as the number of observations necessary to detect it when data is not readily available, both of which we address in this work.

The time variation in most governance variables is shaped by the regulatory and institutional environment in which firms operate. In the case of ownership, the values naturally cluster at important threshold points: simple majority, supermajority and 100%. For board independence, different jurisdictions have gradually introduced legal minimums for listed firms. In addition, firm charters determine the frequency of replacing board members. There are also differences in terms of the range of values governance variables can take. For example, cash flow rights can be very close to 0, for firms held by multiple layers of subsidiaries, all the way to 1, for wholly owned firms. Board independence of listed firms is most likely to vary between 0.5 and 0.95. All these differences are paramount in both parts of our empirical design. In this section, we simulate separately the time variation in ownership and board independence as close as possible to their real-life characteristics, whereas in section four, we capture these differences by performing statistical analysis of real data from multiple angles (averages over time, cross-sectional and time-series standard deviation, year-to-year variation, persistence, etc.).

We begin by modelling the time series evolution of ownership that reflects its specific nature: being a particularly stable (“sticky”) variable, changing in a step-wise fashion and clustering in ranges with economic importance (for example, slightly above 50%). We employ transition probability matrices for this purpose.<sup>8</sup> A transition matrix  $T$  contains the probabilities  $p_{rc}$  of moving from state  $r$  (along the rows) to state  $c$  (along the columns). The states are ownership ranges that reflect important cutoffs used for significant decisions in corporate charters, for example, absolute majority (50% + 1 vote) or supermajority at two thirds and three quarters. We pick the minimum number of states that capture the important cutoffs and the typical patterns of ownership stakes as summarized in Faccio and Lang (2002), Claessens et al. (2000) and Holderness (2009):  $own < 0.05$ ;  $0.05 \leq own < 0.5$ ;  $0.5 \leq own < 0.51$ ;  $0.51 \leq own < 0.6$ ;  $0.6 \leq own < 0.75$ ;  $0.75 \leq own < 0.9$ ;  $0.9 \leq own < 1$ ;  $own = 1$ . For board independence, there are less likely to be common institutional cut-offs, because mandated minimums are jurisdiction-specific, not always codified, and often under a comply-or-explain regime. Thus, our simulations use data-driven quartiles (as in Graham

<sup>7</sup> See the online appendix for a comparison between the ownership links retrieved from the most popular vendor databases (the Bureau van Dijk family of databases) and those from the ICO database used here.

<sup>8</sup> We considered using a smooth parametric function, but it proved untenable in producing the stickiness we are trying to replicate.

et al. 2020), where the transition probabilities and bin cutoffs are determined by real-life global data in the BoardEx and CSMAR databases.<sup>9</sup>

We impose a structure on the ownership transition matrix generated in each iteration of the simulations satisfying a minimum number of rules consistent with stylized facts we already know from existing work and the nature of ownership data: (1) rows sum up to 1 (general property of transition probabilities); (2) the diagonal elements are very high (reflecting stickiness); (3) probabilities along the two off-diagonals (superdiagonal and subdiagonal) are relatively higher (reflecting the greater chance of moving to an adjacent state; for example, the probability of moving from bin 3 ( $0.5 \leq \text{own} < 0.51$ ) to bins 2 or 4 is higher than to other bins); and (4) the probabilities along the last column are also relatively higher (reflecting a tendency towards wholly owned subsidiaries (Nicodano and Regis 2019)).<sup>10</sup> This structure would hold for any one of the three controlling ownership measures used in the literature: ultimate cash flow rights, control rights or the ratio between the two (often referred to as wedge; for example in Faccio et al. 2011 and Lin et al. 2011, 2012 and 2013). It is likely that the ultimate cash flow rights variable exhibits greater time variation than control rights, because any rearrangement of the corporate structure will affect it, while control rights are based on a threshold of control and will only change if that threshold is reached. In this case, the wedge variable will co-vary with ultimate cash flow rights. Of course, there may be cases where both ultimate cash flow rights and control rights change to the same degree, whereby the wedge may appear more stable. To accommodate research applications with any of these ownership measures and a variety of settings from different jurisdictions where ownership may vary to a different degree over time, we define two degrees of stringency of the regularity constraints (2)–(4) above: one for a case of relatively higher time variation and another for lower time variation.

For board independence we define the data-driven quartile bins as follows:  $\text{board indep} < 0.5$ ;  $0.5 \leq \text{board indep} < 0.65$ ;  $0.65 \leq \text{board indep} < 0.8$ ; and  $0.8 \leq \text{board indep} < 0.95$ . These cutoffs are based on the actual distribution of board independence in the BoardEx database for global listed and non-listed firms and in the CSMAR database for Chinese listed firms. Transition probabilities and bin widths by jurisdiction, listed status and state ownership are given in the online appendix. In the context of board independence, time variation will depend on the regulatory and institutional setting in different countries with respect to listed, private and state-owned firms. For example, the Chinese Securities Regulatory Commission (CSRC) mandates that as of June 30, 2003, a minimum of one third of all board directors of a listed firm should be independent. While, among Anglo-Saxon countries and most of Europe, listing rules require that more than half of directors be independent (Papadopoulos 2019). To reflect this variety in the legal range of levels board independence can assume and thereby its scope for change, we again adopt a low and high time variation regime for the simulations of board independence.

The simulation analysis follows the standard power calculation framework outlined in Murphy et al. (2014). We generate artificial data to assess the power of hypothesis tests in a firm-fixed effects specification when using different time-series lengths and frequency of sampling, under the absence or presence of time-variant endogeneity. Our strategy is

<sup>9</sup> Refer to the online appendix for transition probabilities and bin widths of board independence for different jurisdictions and by listed and state ownership status. Ownership simulation results based on data-driven quartile bins are also available in the online appendix.

<sup>10</sup> Full details and the algorithm used in the execution of the simulations are available in the accompanying online appendix.

to generate pseudo-random samples by a process following a theoretical relationship of interest. Our focus is determining whether tests of statistical significance of the coefficient of interest are able to detect this relationship (by rejecting the null hypothesis). Intuitively, since our artificial data was generated under the alternative hypothesis, we should reject the null hypothesis most of the time. Conventionally, a test with a power of 80% or above is considered adequate (see Murphy et al. 2014).

More specifically, the true model underlying our generated data takes the form:

$$y_{it} = \alpha_i + \gamma_t + \beta X_{it} + \lambda W_{it} + \varepsilon_{it} \quad (1)$$

where  $\beta$  is the primary coefficient to be estimated,

$X_{it}$  is artificially generated ownership/board independence data based on a random transition probability matrix  $T$  following the minimal structure described above,

$W_{it}$  is artificially generated data for a generic control variable (like size, leverage, tangibility, etc.) that is correlated directly with  $X$  and with  $Y$  via its error term  $\varepsilon$ , where  $W_{it} = \pi X_{it} + \varepsilon_{it} + \xi_{it}$ . We choose relatively high levels for  $\pi$  and  $\lambda$  of 0.9 to generate a conservative case of high time-varying endogeneity.<sup>11</sup> In the set of results under time-varying exogeneity  $\pi = 0$  and  $\lambda = 0$ .

$y_{it}$  is a hypothetical outcome variable (e.g., firm value, operating performance, public-private debt ratio, foreign direct investment, export intensity, innovation, risk-taking, employee turnover, etc.),

$\alpha_i$  are firm-specific effects generated as  $\alpha_i \sim IIDN(0, 1)$ ,

$\gamma_t$  are time-specific effects,

$\varepsilon_{it}$  and  $\xi_{it}$  are Gauss-Markov disturbances generated as  $\varepsilon_{it} \sim IIDN(0, 1)$  and  $\xi_{it} \sim IIDN(0, 1)$ .

Under time-varying endogeneity we also allow the explanatory variable to be related to  $\varepsilon_{it}$  through the transition probabilities  $p(X_{rcit} | \psi, f_t)$ . We follow Bazzi et al. (2017). The parameter vector  $\psi$  contains all static parameters that govern the transition probabilities, while  $f_t$  captures the dynamic elements that depend on  $\varepsilon_{it}$ . Making  $X$  and  $W$  dependent on the disturbance term of the outcome and correlated with each other is realistic and creates correlation with the error, which violates the zero conditional mean assumption and leads to biased estimates.<sup>12</sup>

<sup>11</sup> The parameters  $\pi$ ,  $\lambda$  and the implicit coefficient of 1 in front of  $\varepsilon$  in the  $W$  equation all contribute to endogeneity. It propagates via several channels: through the key variable of interest  $X$ , ownership or board independence, as well as the control variable  $W$  and the correlation between them. We have selected moderately high levels of endogeneity (close or equal to 1—around the upper range of values we explore for  $\beta$ , namely 0.4, 0.6 and 0.8) to show a situation where statistical power is substantially affected. The time-variant exogeneity case serves as a lower bound where power is not affected. We experimented with levels of these parameters  $\in (0, .9)$ , and the results, as expected, fall between the two cases we show here (exogeneity and moderately high endogeneity). For even more extreme levels of the endogeneity parameters that result in highly positive or highly negative bias, again as expected, power increases because we always reject the null as the estimated  $\beta$  is very far away from 0. In these cases different degrees of time-variation do not play a role in not finding statistical significance if it is present.

<sup>12</sup> We thank an anonymous reviewer for suggesting that we add an additional control variable that is also endogenous. This setup is realistic and introduces more channels for endogeneity to propagate. When endogeneity is introduced only through the transition probabilities of  $X$ , which is stable by construction, endogeneity has little scope to pass on. Another suggestion by the anonymous reviewer was to model the endogeneity in  $X$  through measurement error instead of the transition probabilities. The results are consistent with the ones presented here. The online appendix provides details on these alternative approaches.

We simulate data under (1) using three different values for beta,  $\beta = 0.4, 0.6, 0.8$ , for two transition matrices reflecting high and low time variation in  $X$ . We then estimate regression model (1) on a subsample of the generated data using different gaps and lengths of time. The null and alternative hypothesis are:

$$H_0 : \beta = 0$$

$$H_a : \beta \neq 0$$

Table 1 reports the proportion of times out of 1000 iterations that the null hypothesis is (correctly) rejected. If we find a rejection rate for hypothesis  $H_0$  of at least 80%, the amount of time variation is sufficient for statistical power purposes (shown in bold). In Panel A of Table 1 we simulate data that mimics a situation in which the researcher has available consecutive data points for 10, 15 and 20 years shown along the x-axis. The different values of the theoretical beta coefficient in (1), i.e.,  $\beta = 0.4, 0.6$  and  $0.8$ , are given along the y-axis. The reported results are based on a conservatively small sample of 500 firms, given the limitations of hand-collecting data. The first and third sections summarize the results where strict exogeneity holds; i.e.,  $cov(X_{it}, \varepsilon_{it}) = 0$  in (1), and the second and fourth sections summarize the results where ownership is endogenous; i.e.,  $cov(X_{it}, \varepsilon_{it}) \neq 0$  in (1) and  $cov(W_{it}, \varepsilon_{it}) \neq 0$ . Results for the low (high) time variation case are at the left (right) of each table.

We find that for the simulations exploiting consecutive time observations (10, 15 and 20), whether based on relatively high or low time variation, if there was a true relationship between  $X$  and  $y$ , it should be uncovered by hypothesis tests under time-varying exogeneity of  $X$  (first and third sections of Panel A Table 1). However, under time-varying endogeneity (second and fourth sections of Panel A Table 1) a weak theoretical relationship ( $\beta = 0.4$ ) cannot be detected for any time series length. In the case of  $\beta = 0.6$  and low time variation, power is insufficient for ownership, but for board independence, 20 years of data overcomes this.<sup>13</sup>

Since in many cases collecting ten consecutive years of data may not be practical, we repeat our simulations using different time-series lengths and frequency of sampling. In Panels B and C we consider data with gaps—every 2–5 years—(along the y-axis) for two, three or four time observations (along the x-axis). For example, the coordinate (two time obs, every three years) is consistent with collecting data in 2010 and 2013. The different values for the beta coefficient in Eq. (1) are represented vertically. The first section of Panels B and C shows results where strict exogeneity holds, whereas the second section shows results for the presence of time-varying endogeneity.

In Panel B (ownership), for the high time variation case (right-hand side), we find that the 80% power threshold is reached for a theoretical  $\beta = 0.8$  and any number of time points and gaps in sampling, while for  $\beta = 0.4$ , two or three time observations less than four years apart are no longer sufficient. The results are somewhat stronger in the right-hand side of Panel C (board independence), with additional cases of sufficient power for two time observations every three years under  $\beta = 0.6$ , as well as three time observations every three years under  $\beta = 0.4$ . For low time variation (left-hand side), in Panel B (ownership),

<sup>13</sup> We stress that our simulation results have no bearing on addressing endogeneity. Our decision to run the simulations under exogeneity and in the presence of endogeneity is only intended to diagnose the ability of regression analysis to detect a relationship, if it exists, under these two sets of assumptions.

**Table 1** Simulation results

Panel A: Statistical power of within-firm estimations using consecutive time observations

		Exogenous time variation						Endogenous time variation						
		Low			High			Low			High			
		Number of time observations						Number of time observations						
		10	15	20	10	15	20	10	15	20	10	15	20	
<i>Ownership</i>														
Theoretical beta coefficient	$\beta = .8$	<b>.988</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	.8	<b>.812</b>	<b>.978</b>	<b>.998</b>	<b>1</b>	<b>1</b>	<b>1</b>
	$\beta = .6$	<b>.962</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	.6	.238	.47	.608	<b>.848</b>	<b>.982</b>	<b>.996</b>
	$\beta = .4$	<b>.892</b>	<b>.998</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	.4	.052	.082	.164	.14	.244	.348
<i>Board independence</i>														
Theoretical beta coefficient	$\beta = .8$	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	.8	<b>.93</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
	$\beta = .6$	<b>.99</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	.6	.34	.614	<b>.844</b>	.736	<b>.962</b>	<b>.992</b>
	$\beta = .4$	<b>.82</b>	<b>.98</b>	<b>1</b>	<b>.99</b>	<b>1</b>	<b>1</b>	.4	.098	.13	.19	.118	.224	.342

Panel B: Statistical power of within-firm estimations when sampling ownership data with gaps

		Exogenous time variation						Endogenous time variation							
		Low			High			Low			High				
		2	3	4	2	3	4	2	3	4	2	3	4		
Sampling every 2, 3, 4 and 5 years	$\beta = .8$	5	.414	.786	<b>.948</b>	<b>.984</b>	<b>1</b>	<b>1</b>	5	.185	.408	.616	.695	<b>.972</b>	<b>.998</b>
		4	.352	.708	<b>.894</b>	<b>.992</b>	<b>1</b>	<b>1</b>	4	.179	.339	.548	.658	<b>.943</b>	<b>.999</b>
		3	.296	.568	<b>.864</b>	<b>.944</b>	<b>1</b>	<b>1</b>	3	.126	.278	.456	.569	<b>.911</b>	<b>.995</b>
	$\beta = .6$	2	.214	.436	.672	<b>.874</b>	<b>1</b>	<b>1</b>	2	.114	.207	.353	.461	<b>.84</b>	<b>.967</b>
		5	.26	.564	.77	<b>.892</b>	<b>.996</b>	<b>1</b>	5	.066	.116	.148	.199	.36	.546
		4	.238	.444	.704	<b>.826</b>	<b>.988</b>	<b>1</b>	4	.068	.106	.139	.171	.337	.511
	$\beta = .4$	3	.19	.364	.658	.76	<b>.994</b>	<b>1</b>	3	.058	.101	.125	.151	.305	.46
		2	.148	.25	.426	.674	<b>.946</b>	<b>1</b>	2	.08	.073	.101	.145	.235	.375
		4	.144	.286	.442	.586	<b>.902</b>	<b>.978</b>	5	.042	.052	.066	.053	.095	.106
		4	.14	.216	.39	.472	<b>.856</b>	<b>.964</b>	4	.048	.052	.051	.06	.084	.099
		3	.12	.202	.318	.4	.794	<b>.956</b>	3	.041	.041	.039	.06	.072	.091
		2	.094	.14	.224	.334	.658	<b>.888</b>	2	.055	.048	.068	.061	.073	.083

Panel C: Statistical power of within-firm estimations when sampling board independence data with gaps

		Exogenous time variation						Endogenous time variation							
		Low			High			Low			High				
		2	3	4	2	3	4	2	3	4	2	3	4		
Sampling every 2, 3, 4 and 5 years	$\beta = .8$	5	<b>.953</b>	<b>1</b>	<b>1</b>	<b>.995</b>	<b>1</b>	<b>1</b>	5	.708	<b>.966</b>	<b>1</b>	<b>.971</b>	<b>1</b>	<b>1</b>
		4	<b>.939</b>	<b>.999</b>	<b>1</b>	<b>.995</b>	<b>1</b>	<b>1</b>	4	.646	<b>.948</b>	<b>.997</b>	<b>.933</b>	<b>.999</b>	<b>1</b>
		3	<b>.895</b>	<b>.995</b>	<b>1</b>	<b>.984</b>	<b>1</b>	<b>1</b>	3	.516	<b>.898</b>	<b>.984</b>	<b>.886</b>	<b>1</b>	<b>1</b>
	$\beta = .6$	2	<b>.81</b>	<b>.985</b>	<b>1</b>	<b>.95</b>	<b>.995</b>	<b>1</b>	2	.382	.731	<b>.939</b>	.792	<b>.989</b>	<b>1</b>
		5	.78	<b>.977</b>	<b>1</b>	<b>.945</b>	<b>.998</b>	<b>1</b>	5	.347	.663	<b>.895</b>	.683	<b>.952</b>	<b>.99</b>
		4	.743	<b>.976</b>	<b>.999</b>	<b>.924</b>	<b>.998</b>	<b>1</b>	4	.306	.633	<b>.849</b>	.607	<b>.928</b>	<b>.994</b>
	$\beta = .4$	3	.656	<b>.934</b>	<b>.996</b>	<b>.883</b>	<b>.998</b>	<b>1</b>	3	.265	.521	.749	.526	<b>.871</b>	<b>.978</b>
		2	.615	<b>.885</b>	<b>.985</b>	.76	<b>.95</b>	<b>1</b>	2	.202	.375	.617	.427	.763	<b>.948</b>
		5	.435	.765	<b>.916</b>	.666	<b>.937</b>	<b>.994</b>	5	.098	.191	.301	.185	.343	.475
		4	.441	.728	<b>.905</b>	.628	<b>.917</b>	<b>.985</b>	4	.104	.175	.252	.157	.309	.468
		3	.371	.659	<b>.863</b>	.576	<b>.887</b>	<b>.986</b>	3	.1	.131	.196	.137	.263	.389
		2	.36	.545	.785	.435	.69	<b>.89</b>	2	.077	.102	.154	.115	.228	.348

power is only sufficient for  $\beta=0.8$ , with four time observations sampled three years apart or more. In the left-hand side of Panel C (board independence), power is a lot higher—being sufficient in all cases under  $\beta=0.8$  and for three and four time observations under  $\beta=0.6$  and even under  $\beta=0.4$  for four time observations every three years or more. The results in Panel B show that collecting ownership data with gaps does not generally lead to good power properties when time variation is low, except when the theoretical relationship is strong and there are at least four time observations. By contrast, in the case of board independence (Panel C), even for weaker theoretical strength under relatively low time variation, sampling every three years or more and having at least four time observations provides sufficient power.

In the right-hand sections of Panels B and C we present results under time-varying endogeneity. We find that endogeneity leads to worse power properties than under exogeneity for both the low and high time variation cases. For ownership (Panel B), power is only sufficient under  $\beta=0.8$ , high time variation and three or more time observations, while for board independence (Panel C), we see cases of sufficient power for low time variation and  $\beta=0.6$ . We caution that the power measure is not as informative in the presence of endogeneity since the t-statistic at its basis has a biased numerator.

To sum up, if a relationship between  $X$  and  $y$  exists, consecutive data of ownership or board independence covering at least ten years should detect it when using a within-firm estimator under exogeneity. Under time-varying endogeneity, however, a weaker theoretical relationship cannot be detected even with 20 years of data and large time variation. Under the practical approach to collecting data with gaps in the presence of time-varying endogeneity, hypothesis tests are more likely to detect a relationship with ownership (board independence) if there are at least three time observations for a relatively strong (even for less strong) theoretical relationship and time variation is high (or even when time variation is low).

In light of our simulation results, we can revisit some existing work where time variation can be deduced. Donelli et al. (2013) report explicitly an average proportion of 6–7% of firms changing each year over a 20-year period for their Chilean data. This degree of time variation clearly corresponds to our low variation case (lower left corner of section two of Panel A Table 1). In their Table 8 they report various outcome regressions, whereby ownership changes are significant only when firm fixed effects are not included. There could be two possible explanations: (1) either time-invariant endogeneity leads to bias in the OLS estimator and erroneously shows significance, or (2) the relationship indeed exists but the low time variation in their data prevents them from detecting it with a within-firm estimator. Our simulation results show lack of power for the same 20-year period, with theoretical relationship strength below  $\beta=0.8$  and low time variation, as is the case in their data, and therefore provide support for explanation 2. Thus, the conclusion of Donelli et al. (2013) regarding a weak theoretical relationship between ownership and real outcomes is supported by our work.

In Panels A (ownership) and C (board independence) of Fig. 1 we summarize the simulation results under time-varying endogeneity and sampling data with gaps to produce a relationship between a measure of time variation (the proportion of firms changing between one time point and the next) and statistical power. These graphs show the predicted marginal probabilities of rejecting  $H_0$  based on a logit specification using the data

from all simulations.<sup>14</sup> The logit model regresses power on frequency of sampling, the proportion of firms that change between the first and next period, true beta, number of time observations, and starting year of sampling.<sup>15</sup> We show statistical power (along the y axis) as a function of the proportion of firms that change between two time periods (along the x axis) for three different strengths of the theoretical relationship being tested, sampling gaps and number of time observations. Suppose a researcher has collected two time observations five years apart for a sample of firms, as in Basu et al. (2017). She can compute the proportion of firms that change between the two points in time; then she can compare it to the required 63% (for  $\beta=0.8$ ) or 86% (for  $\beta=0.6$ ) for sufficient statistical power of at least 80% (bottom row of Panel A of Fig. 1). If she observes lower proportions, she could decide to collect an extra time observation, which reduces the required range to 40% and 60% respectively.

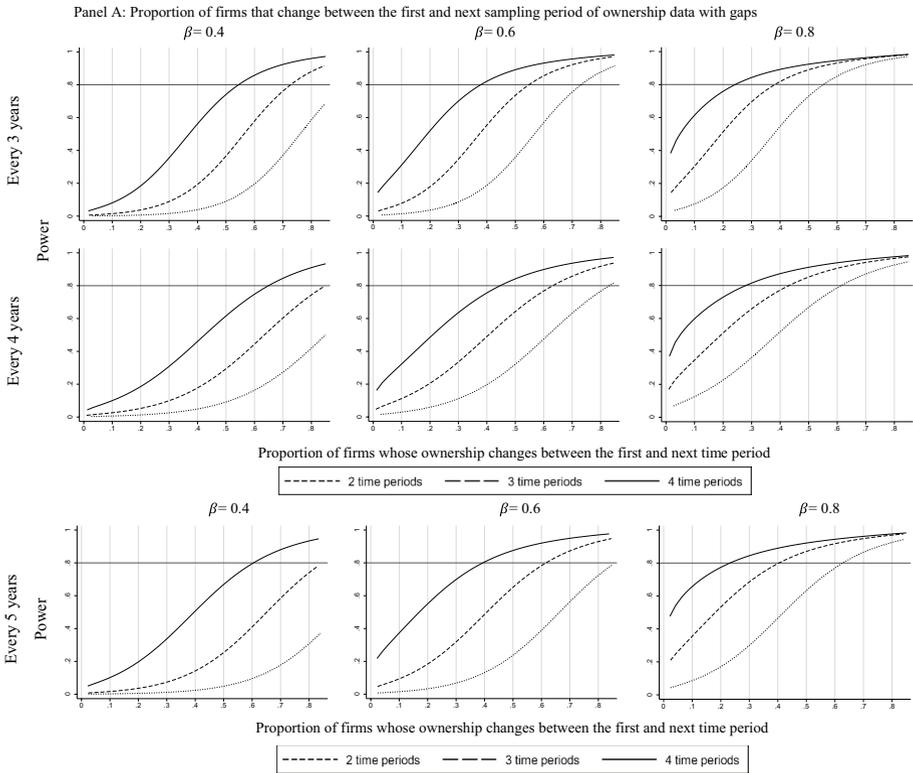
Two of the studies we surveyed allow us to deduce the proportion of observations that change in their sample but only over the whole period of analysis: in Lin et al. (2011), 21% of firms exhibit changes in ownership over a 13-year period, while in Lin et al. (2013) the proportion is 39% over a 10-year period. In the first paper, the authors collect ownership data four years apart. To see how these degrees of time variation map with our results, in Panel B of Fig. 1 we present a subset for sampling every four years as a function of the proportion of firms that change between the first and last year of data (instead of the following year, as in Panels A and C). We note that Lin et al. (2011) opt for difference regressions not on their full data but only on the observations that do change. The 13 years of data sampled every four years corresponds to the right-most graph in Panel B. For 17% of firms changing (close to the 21% in Lin et al. (2011)) between the first and last period, sufficient power is only present for a theoretical beta coefficient of 0.8. Therefore, a presumed lack of power on a full first difference estimation in their case is consistent with a relatively lower theoretical beta.

For board independence under endogeneity (Panel C of Fig. 1) we show sampling every two years, consistent with the design in Wintoki et al. (2012), and every four years for contrast. For low theoretical strength and two time periods power is very low, but for four time observations we find that power is enough once 45% of the firms change between the first and next sampling period. For higher theoretical strength this proportion drops to 10%. We examine these findings further in the next section, where we replicate two influential studies where ownership and board independence are the key variables of interest.

Figure 1 may help researchers in several ways. Even when consecutive years of data are easily available in electronic form from data vendors, careful empirical design will benefit from comparing the degree of time variation in the data (relevant to a research question of interest) to our benchmark results. For example, if the researcher finds no significant relationship and the amount of variation in their data is low relative to our benchmark results, the decision not to use within-firm estimators can be substantiated with more formality. When data collection is a hurdle, for example when board data for private firms is not readily available, starting with two non-consecutive time observations of data and measuring the proportion of firms that change allows the researcher to decide whether collecting an additional time observation of data is worthwhile. Furthermore, a researcher starting a new project could look up an existing study that uses data with similar characteristics

<sup>14</sup> We thank Vlado Atanasov for suggesting these visualizations of the simulation results.

<sup>15</sup> The logit regression results are available from the authors upon request.



**Fig. 1** Statistical power as a function of time variation under time-variant endogeneity of ownership and board independence

(jurisdiction, family control, listed status, etc.), check the proportion of firms that change in that study, and, coupled with our power graphs, choose a sampling time gap that provides sufficient statistical power.

### 3 Data and replicating regressions

Next, we check whether our simulation findings play out in real-life data by replicating two highly cited studies focusing on ownership (Lin et al. 2013) and board independence (Coles et al. 2008). We use several data sources for ownership and board independence data. Our most comprehensive source (ICO) provides a long time series coverage of detailed controlling ownership for a large number of firms. It is distinctive from most of the popular ownership databases in the following ways. First, it provides the complete chain of ownership links from any firm operating in Canada (above a certain size threshold<sup>16</sup>) to its ultimate

<sup>16</sup> The size thresholds have changed from above CAD15mil in revenues or CAD10mil in assets before 2006 to above CAD80mil in revenues or CAD200mil in assets after 2006. For exact details on disclosure requirements and documentation of the ICO database see the accompanying online appendix.

Panel B: Proportion of firms that change between the first and last period of ownership data with sampling every 4 years

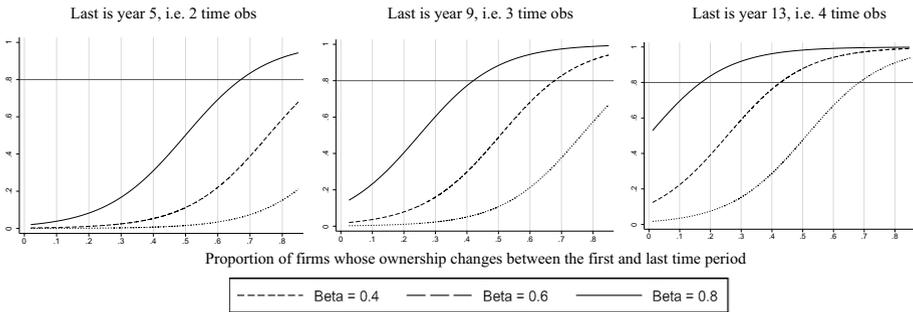


Fig. 1 (continued)

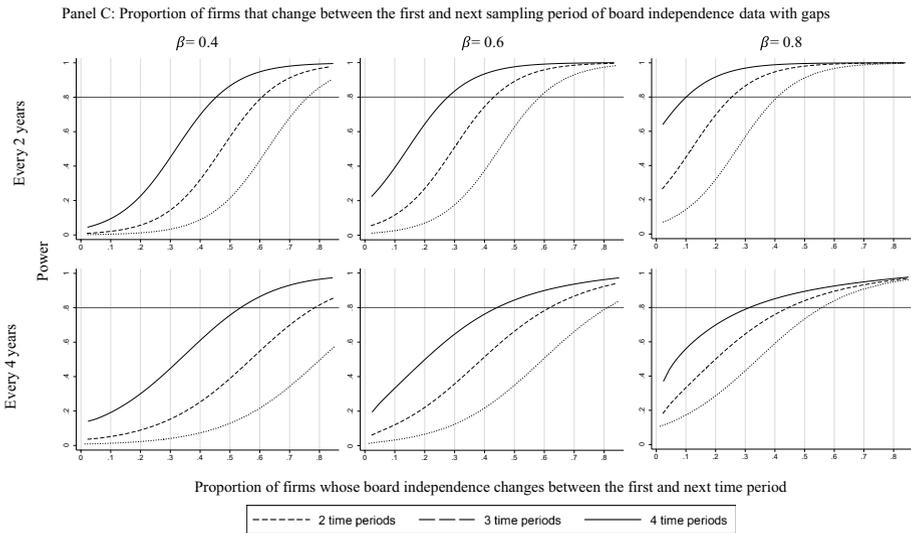


Fig. 1 (continued)

owner. This disclosure is mandated every year under the Corporations Returns Act (CRA). By contrast, the family of ownership databases compiled by Bureau van Dijk (BvD)—Osiris, Orbis, Amadeus, Fame, etc.—update their historical data only if new information becomes available, which means that it has uneven sampling gaps by design. We summarize all differences between our data source and BvD in the online appendix. Importantly, CRA covers multinationals as long as they have a subsidiary in Canada, which allows us to observe all cross-border and foreign ownership links that lead to the ultimate controlling owner and therefore we have a large number of global companies. One of the most valuable features of this data source, which we exploit in the next section, is that it covers all non-listed firms, as opposed to the typical database coverage, where only some non-listed firms self-select to provide disclosure. We have extensive coverage for Anglo-Saxon firms (19,082, of which 1369 are listed) and somewhat less from Continental Europe (6269, of which 308 are listed) and Rest of the World (3019, of which 137 are listed)—refer to Fig. 2

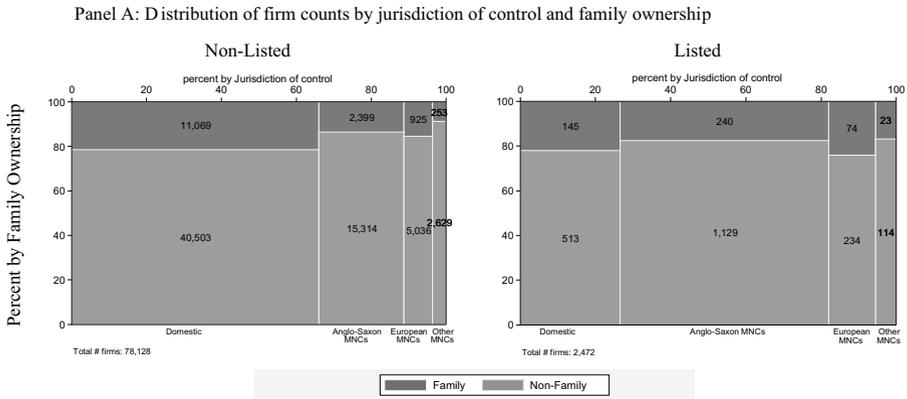
for full coverage of the data and further splits by family control and ownership structure complexity. We perform an extensive hand-collection of family control status, apply a verification algorithm with multiple alternative sources and exploit the long time series to detect inconsistencies and outliers. Last, we compute three ownership variables—cash flow rights of the ultimate owner, control rights of the ultimate owner and the ratio between the two—using computational tools from graph theory and formalized in Almeida et al. (2011). To ensure greater representativeness for the rest of the world, we augment all analyses with the CSMAR database covering all listed firms in China and offer further splits by state ownership.

The raw direct ownership stakes we start with come from the Intercorporate Ownership Database (ICO) compiled by Statistics Canada for the period 1995–2019. The CRA specifies a penalty of fines and/or prison for failure to disclose ownership information. In ICO, all firms are attributed to belong to structures referred to as “enterprises” based on common control.<sup>17</sup> The two groups of firms (domestic Canadian and multinationals) differ in that the domestic dataset includes a large majority of private and small firms, while the multinational (MNC) dataset represents firms that are larger and more likely to be listed. For the replicating regressions we are limited only to listed firms, for which there is financial data (in either Capital IQ or Thomson One) and for which we can construct an outcome variable (public-to-private debt ratio and firm performance as proxied by Tobin’s Q) plus the control variables in each of the two studies we replicate. We compute three ownership variables: the ultimate cash flow rights of the controlling owner ( $ucfr$ ), the control rights of the controlling owner ( $cr$ ) and the ratio between the two ( $wedge$ ).  $ucfr_i$  is the proportion of one unit of disbursement from firm  $i$  that is received by the controlling owner, while  $cr_i$  is the critical control threshold, which is shown to be equivalent to the concept of the weakest link (as used in La Porta et al. 1999, Claessens et al. 2000, and Faccio and Lang 2002) when cross-shareholdings and multiple links are absent but can also be computed for more complex structures. We follow Almeida et al. (2011) in the construction of  $ucfr$  and  $cr$ . In the remainder of the text, for brevity, we use the term “ownership” synonymously with any one of the three variables. Full definitions of ownership and board independence variables are given in Appendix 1, while summary statistics of all variables used in the two replications are in Appendix 2.

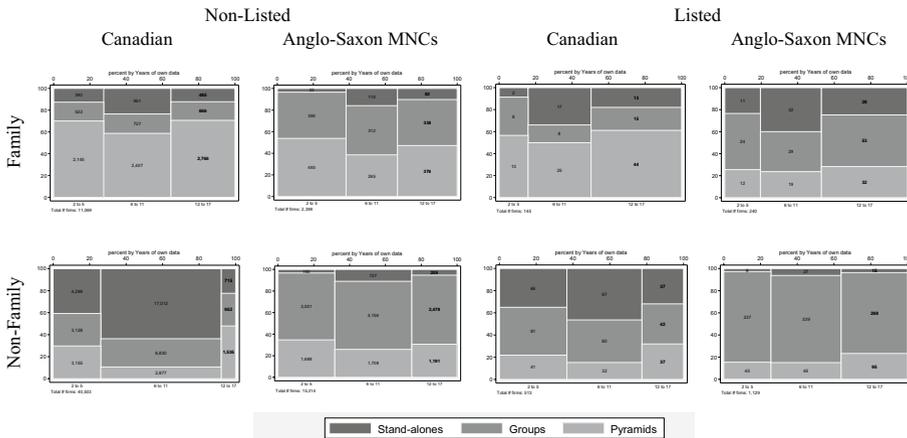
We collect board independence data from BoardEx, which is the most comprehensive board composition database, covering more than 20,000 companies globally. We separate firms into four jurisdiction groups in descending order of coverage: US (12,559, of which 7342 are listed), Anglo-Saxon (7228, of which 5258 are listed), Continental Europe (3532, of which 2865 are listed) and Rest of the World (5716, of which 4993 are listed). Again, we augment the East-Asian coverage by analyzing board independence of all Chinese listed companies in the CSMAR database (2613, of which 1000 are state-owned). For the illustrative regression on board independence, we use the same data as in the original study—US listed firms—and retrieve the required financial variables from Compustat.

We begin with a quasi-replication of Lin et al. (2013). They study the choice between public and private debt among Western European and East-Asian companies. Lin et al. (2013) use consecutive time observations for a 10-year period (2001–2010), whereas in two earlier related studies (Lin et al. 2011, 2012) they construct ultimate ownership for

<sup>17</sup> Variables and concept definitions are given in Appendix 1. The attribution to common control structures is done by Statistics Canada as detailed in the accompanying online appendix. For example, the assignment of control rules is based on voting rights, option grants and interlocking directorships.



Panel B: Distribution of firm counts inside each of the left four rectangles in Panel A by number of years of data and ownership structure complexity (stand-alone, group or pyramid)



**Fig. 2** Data coverage of the ICO database. This figure shows the coverage of the ICO data for productive firms only (excluding financials, holding companies and charities –exact definition in Appendix 1). The same figure for non-productive firms is in the online appendix.

four points in time that are two and three years apart. Although Lin et al. (2013) have a 10-year period, their unbalanced panel provides on average four time observations per firm (9783 firms and 43,273 observations in their Table 4, which contains the specification we use). Our ownership data cover 25 years and allow us to select firms that have sufficient number of consecutive time observations so that we are able to examine different time series properties.

We adopt their specification<sup>18</sup>:

$$\text{debt choice} = \gamma_t + \eta_j \gamma_t + \beta_1 \text{wedge} + \beta_2 \text{zscore} + \beta_3 \text{wedge} * \text{zscore} + \sum_{k=4}^{10} \beta_k X_k + \sum_{q=1}^n \theta_q \text{ind}_q + \varepsilon \quad (2)$$

Our extensions employ the within-firm estimator:

$$\text{debt choice} = \alpha_i + \gamma_t + \eta_j \gamma_t + \beta_1 \text{wedge} + \beta_2 \text{zscore} + \beta_3 \text{wedge} * \text{zscore} + \sum_{k=4}^{10} \beta_k X_k + \varepsilon \quad (3)$$

where we suppress firm-year subscripts for brevity; *debtchoice* is the ratio of public to private debt;  $\alpha_i$  are firm fixed effects, which we use in our extensions but which are not present in Lin et al. (2013);  $\gamma_t$  are year effects,  $\eta_j \gamma_t$  are country  $\times$  year effects;  $X_k$  is a vector of control variables: cash-flow rights, leverage, tangibility, size, profitability and Tobin's Q, all defined exactly as in the original paper (*ind<sub>q</sub>* are industry fixed effects, which are absorbed by the firm fixed effects in our extensions).

We use a sample of similar but not identical firms because the data source in Lin et al. (2013) is a proprietary database (ORBIS by BvD—refer to the online appendix for the differences in precision between the BvD family of databases and ICO). However, we apply the exact same regression specification as in their Table 4, given in Eq. (2) above, plus within-firm estimators with different frequencies of sampling, given in Eq. (3). We show results for three jurisdictional samples in Table 2: All firms (Panel A), Continental Europe and East Asian firms—CEEA (Panel B) and US & UK firms (Panel C). We examine how the results may differ for firms not analyzed in Lin et al. (2013)—US firms, which exhibit lower time variation in ownership than those in Western Europe and East-Asia.

Column (1) of Panels A, B and C in Table 2 reports the same OLS estimator used by Lin et al. (2013), while in column (2) we show the within-firm estimator using the full data. In columns (3) and (4) of Panels B and C we limit the sample to contain four time observations with sampling frequency every two and every five years. To maintain comparability between the two sampling frequencies in Panels B and C, we keep the number of time observations the same, which means that in the five-years-apart case we need at least 16 years of data (including the first and last year).

When using the full data with a OLS specification with industry dummies (column (1) of Panels A, B and C), we confirm Lin et al. (2013) findings that the use of public debt is lower when firms have a higher wedge and z-score, but the interaction of the two counteracts this effect. However, in column (2) of Panel A, using the within-firm estimator results in lack of significance. This could be attributed to the low time variation in ownership of US firms, which represent around half of the sample.

The within-firm estimator in column (2) of Panels A, B and C only detects significance for the sub-sample of CEEA firms. At the bottom of Panels B and C in Table 2 we report the proportion of firms that change between the first and second sampling period. In column (3) Panel B Table 2 only 14.6% of CEEA firms change and the wedge effect is not

<sup>18</sup> In their robustness section Lin et al. (2013) use a within-firm estimator (a change regression) to address time invariant endogeneity. However, they estimate it only on the firms that exhibit ownership changes, which represent 30% of their sample. This approach is imperfect because, the excluded observations would have contributed to finding no relationship between  $X$  and  $y$  since  $X$  does not vary while  $y$  does. Thus, the omission would likely lead to bias.

detected, but five years apart (column (4)), 35.7% of the firms change and the effect is revealed.

For US & UK firms (Panel C of Table 2), however, the potential effect of the wedge cannot be detected by the within-firm estimator (columns (2), (3) and (4)). The low degree of time variation for the US & UK sample is evident in that only 10.8% of firms change in two consecutive years, while the proportion increases slightly to 13.4% and 17.1% respectively for two and five years apart. Both sets of results in Panels B (CEEA firms) and C (US & UK firms) are consistent with our simulation findings. In particular, having ten years of data is sufficient to detect an effect if it exists and is relatively strong (second section of Panel B Table 1); however, sampling five years apart with four time observations requires at least 24% of firms to change (left-most graph on the bottom row of Panel A Fig. 1).

Next, we replicate the regressions in Table 5 of Coles et al. (2008). They study the effect of firm complexity on firm value as measured by Tobin’s Q in requiring a variety of rich expertise from the board. Their specification is:

$$Q = \gamma_t + \beta_1 insider\ frac + \beta_2 outsiders + \beta_3 advice + \beta_4 advice * outsiders + \sum_{k=5}^{11} \beta_k X_k + \sum_{q=1}^n \theta_q ind_q + \epsilon \tag{4}$$

Our extensions employ the within-firm estimator:

$$Q = \alpha_i + \gamma_t + \beta_1 insider\ frac + \beta_2 outsiders + \beta_3 advice + \beta_4 advice * outsiders + \sum_{k=5}^{11} \beta_k X_k + \epsilon \tag{5}$$

where we suppress firm-year subscripts for brevity;  $Q$  is Tobin’s Q;  $\alpha_i$  are firm fixed effects, which we use in our extensions but are not present in Coles et al. (2008);  $\gamma_t$  are year effects;  $insiderfrac = 1 - board\ independence$ ;  $outsiders$  is the log of the number of independent directors;  $advice$  is a dummy variable equal to 1 if the firm-year observation ranks above the median by the first principal component of firm complexity based on the number of segments, size and leverage;  $X_k$  is a vector of control variables: R&D dummy, standard deviation of returns, profitability plus its lag, intangible assets and CEO ownership, all defined as in the original paper; and  $ind_q$  are industry fixed effects, which are absorbed by the firm fixed effects in our extensions. Here we require three time observations four years apart, which means a minimum of nine years of data plus one for lagged profitability.

We begin by mimicking the number of observation in Coles et al. (2008) in Panel D of Table 2. We are unable to replicate their number of firms, since they are not reported in the original study. We show that the within-firm estimator only detects the significance on insider fraction for the subset of firms with high time variation in column (4). This implies that the sample Coles et al. (2008) were working with likely had low time variation. Next, we exploit our full sample, which is larger than Coles et al. (2008), despite the fact that we limit the time period to ten years as is the case in their paper. This is due to the better coverage in BoardEx, which begins in 1999, while the Execucomp data in Coles et al. (2008) covers the period 1992–2001, but is sparser. The OLS and within-firm estimators in columns (1) and (2) of Panel E Table 2 detect the effect of the proportion of insiders as found by Coles et al. (2008) in their Table 5 column (3). When we sample with gaps using three time observations, power decreases but is still enough to detect the effect in column (3) given that 34.8% of firms change between the first and next sampling period (consistent with the dashed line in the top right graph of Panel C in Fig. 1, which requires more than 30% of firms changing). In column (4), however, 41.2% of firms changing four years apart

no longer provides sufficient power, as the bottom right graph of Panel C in Fig. 1 requires more than 45% of firms to change.

Our replicating regressions verify that the within-firm estimator can detect a relationship between a sticky variable of interest even in a short panel as long as the sampling is performed as far apart as necessary to capture sufficient time variation. This is important, because within-firm estimators (even if imperfect) are more reliable in establishing causality than alternative approaches in the absence of shocks or valid instruments.<sup>19</sup>

## 4 Empirical analysis

Last, we perform six types of time variation analysis to investigate how the four variables we consider map along our simulation findings (Table 1 and Fig. 1) depending on jurisdiction, listed status, family or state control and ownership structure complexity. Three of the analyses are of descriptive nature (averages over time, year-to-year variation and stable regimes), while the other three analyses are regression-based (analysis of variance, regressions showing the determinants of the time-series standard deviation of ownership and board independence and persistence regressions). We show additional sub-sample results along all six types of analyses in the online appendix.

The granular ownership data in ICO with a long time dimension allows for the most detailed sample splits, the dominance of state-owned firms in China provides an additional interesting dimension, while BoardEx shows the time variation patterns in board independence by jurisdiction and listed status. The composition of the ICO database is presented in Panel A of Fig. 2. The group of MNCs controlled from Anglo-Saxon countries is large and provides relative jurisdictional homogeneity. Therefore, our baseline analysis focuses on the two largest groups: Canadian firms and Anglo-Saxon MNCs. We present results for the two remaining groups—Continental Europe and Rest of the World—in the accompanying online appendix (the right-most two rectangles of each graph in Panel A of Fig. 2).<sup>20</sup> Anglo-Saxon countries are united by a type of capitalism characterized by a lower degree of government intervention and greater reliance on free market mechanisms (Esping-Andersen 1990). We focus on productive<sup>21</sup> firms that are part of either a pyramid or a group corporate structure.<sup>22,23</sup> This approach may appear too restrictive at first glance, but we emphasize that the granularity of the data means that firms which usually would be classified as stand-alone, here fall in the group or pyramid type. In particular, listed or large private firms, even if non-family-controlled, almost always have subsidiaries (often wholly owned and organized in flat structures in

<sup>19</sup> In the ANOVA analysis in the following section we show that year, industry, size deciles and the interactions thereof capture a tiny proportion of the variation in ownership and board independence.

<sup>20</sup> The results for the subsamples of Continental Europe and Rest of the World are more consistent with the Anglo-Saxon MNC than the Canadian subsample, but any conclusions based on them would be based on a much smaller sample size.

<sup>21</sup> We refer to non-financial firms as ‘productive’ throughout the text, tables and graphs. Financial firms include all types of financial intermediaries, investment funds, trusts and holding companies.

<sup>22</sup> Results for financials are available in the online appendix. Similarly, some of our baseline analyses focus on firms with at least 12 years of data, in these case we show results for firms with less than 12 years of data in the online appendix as well.

<sup>23</sup> We use an established definition of a pyramid without the restriction that it must contain a listed firm (consistent with Almeida et al. 2011 and Faccio et al. 2010).

**Table 2** Illustrative regression results

Panel A: All firms—specification as in Lin et al. (2013) Table 4

	All data; OLS	All data; Fixed effects
Wedge	-0.083** (0.035)	-0.019 (0.028)
Z-score	-0.017** (0.007)	0.011* (0.006)
Z-score × Wedge	0.117*** (0.007)	-0.004 (0.006)
Cash-flow rights	0.008 (0.023)	0.013 (0.016)
Leverage	-0.048** (0.022)	-0.031** (0.016)
Tangibility	-0.019 (0.019)	0.002 (0.020)
Log assets	-0.007*** (0.002)	-0.009*** (0.002)
Profitability	-0.002 (0.015)	-0.010 (0.007)
Q	-0.001 (0.006)	-0.003 (0.004)
Industry effects	Yes	No
Country × time effects	Yes	Yes
Obs	34,016	34,016
Firms	2044	2044
Adj/Overall R-sq	0.058	0.032

Panel B: Continental Europe and East Asian firms (CEEAs)—specification as in Lin et al. (2013) Table 4

	All data; OLS	All data; Fixed effects	Every 2 years; Fixed effects	Every 5 years; Fixed effects
Wedge	-0.269** (0.104)	-0.174** (0.085)	-0.122 (0.130)	-0.183** (0.087)
Z-score	-0.231*** (0.032)	-0.112*** (0.026)	0.023 (0.029)	-0.031 (0.023)
Z-score × Wedge	0.094*** (0.032)	0.054** (0.026)	0.100*** (0.030)	0.150*** (0.024)
Cash-flow rights	0.089* (0.046)	0.094** (0.046)	0.049 (0.088)	-0.023 (0.047)
Leverage	0.084 (0.056)	-0.017 (0.053)	-0.086 (0.102)	0.015 (0.070)
Tangibility	0.018 (0.061)	0.037 (0.059)	0.019 (0.125)	-0.004 (0.078)
Log assets	-0.017*** (0.005)	-0.015* (0.008)	-0.005 (0.016)	-0.023** (0.010)
Profitability	-0.002 (0.034)	-0.040 (0.034)	0.090* (0.051)	-0.004 (0.032)

**Table 2** (continued)

Panel B: Continental Europe and East Asian firms (CEEA)—specification as in Lin et al. (2013) Table 4

	All data; OLS	All data; Fixed effects	Every 2 years; Fixed effects	Every 5 years; Fixed effects
Q	−0.029* (0.018)	−0.032* (0.016)	−0.024 (0.032)	−0.004 (0.025)
Industry effects	Yes	No	No	No
Country × time effects	Yes	Yes	Yes	Yes
Obs	15,896	15,896	3,443	3,145
Firms	938	938	955	972
Prop. firms changing	0.123	0.123	0.146	0.357
Adj/Overall R-sq	0.161	0.052	0.004	0.140

Panel C: US and UK firms—specification as in Lin et al. (2013) Table 4

	All data; OLS	All data; Fixed effects	Every 2 years; Fixed effects	Every 5 years; Fixed effects
Wedge	−0.102** (0.046)	0.009 (0.032)	0.111 (0.068)	0.043 (0.041)
Z-score	−0.015** (0.008)	−0.002 (0.007)	0.002 (0.012)	0.003 (0.009)
Z-score × Wedge	0.120*** (0.008)	−0.007 (0.007)	−0.016 (0.012)	−0.004 (0.009)
Cash-flow rights	0.035 (0.038)	−0.006 (0.024)	−0.065 (0.052)	−0.045 (0.029)
Leverage	−0.045 (0.029)	−0.021 (0.019)	−0.066 (0.042)	−0.032 (0.030)
Tangibility	−0.037 (0.028)	−0.006 (0.028)	−0.087 (0.075)	−0.004 (0.044)
Log assets	−0.004 (0.002)	−0.004 (0.003)	−0.011 (0.008)	−0.003 (0.004)
Profitability	−0.007 (0.020)	−0.004 (0.011)	0.013 (0.027)	0.012 (0.017)
Q	0.001 (0.008)	0.003 (0.005)	0.018 (0.013)	0.008 (0.008)
Industry effects	Yes	No	No	No
Country × time effects	Yes	Yes	Yes	Yes
Obs	17,531	17,531	3,663	3,428
Firms	1029	1029	1042	1,070
Prop. firms changing	0.108	0.108	0.134	0.171
Adj/Overall R-sq	0.062	0.011	0.021	0.022

**Table 2** (continued)

Panel D: Board independence specification on subsamples with low and high time variation—specification as in Coles et al. (2008) Table 5

	Low time var.; OLS	High time var.; OLS	Low time var.; Fixed effects	High time var.; Fixed effects
Insider fraction	0.120* (0.068)	0.537*** (0.121)	0.239 (0.182)	0.392** (0.175)
Log(Outsiders)	0.045** (0.022)	0.080* (0.043)	-0.003 (0.069)	0.058 (0.080)
Advice dummy	-0.039 (0.061)	-0.290*** (0.090)	-0.127 (0.133)	-0.213 (0.175)
Advice dummy = 1 # Log(Outsiders)	-0.031 (0.033)	0.100** (0.045)	-0.016 (0.068)	0.071 (0.082)
R&D dummy	0.000 (0.012)	0.030** (0.014)	-0.038 (0.044)	-0.065 (0.076)
Risk	-0.205*** (0.032)	-0.218*** (0.035)	-0.119*** (0.038)	-0.138*** (0.045)
Profitability	0.021 (0.064)	-0.046 (0.077)	0.011 (0.057)	-0.036 (0.071)
Lagged Profitability	0.044 (0.064)	0.033 (0.077)	0.101 (0.071)	0.048 (0.077)
Intangible assets	-0.039* (0.022)	0.114*** (0.028)	0.080 (0.090)	0.461 (0.299)
Insider ownership	0.398*** (0.055)	-0.328* (0.180)	-0.122 (0.085)	-0.114 (0.251)
Industry dummies	Y	Y	N	N
Year dummies	Y	Y	Y	Y
Obs	6775	6716	6775	6716
Firms	464	406	464	406
Adj/Overall R-sq	0.026	0.032	0.014	0.019

Panel E: Board independence specification on full data and sampling with gaps—specification as in Coles et al. (2008) Table 5

	All data; OLS	All data; Fixed effects	Every 2 years; Fixed effects	Every 4 years; Fixed effects
Insider fraction	0.100** (0.042)	0.160** (0.069)	0.144* (0.081)	0.032 (0.112)
Log(Outsiders)	0.090*** (0.019)	0.074** (0.032)	0.119** (0.052)	-0.026 (0.047)
Advice dummy	-0.013 (0.040)	-0.219*** (0.067)	-0.047 (0.098)	-0.184* (0.098)
Advice dummy = 1 # Log(Outsiders)	-0.041* (0.021)	0.039 (0.035)	-0.048 (0.055)	0.037 (0.053)
R&D dummy	-0.003 (0.007)	-0.079*** (0.020)	-0.111*** (0.033)	-0.021 (0.032)

**Table 2** (continued)

Panel E: Board independence specification on full data and sampling with gaps—specification as in Coles et al. (2008) Table 5

	All data; OLS	All data; Fixed effects	Every 2 years; Fixed effects	Every 4 years; Fixed effects
Risk	−0.214*** (0.019)	−0.098*** (0.018)	−0.102*** (0.030)	−0.224*** (0.044)
Profitability	−0.050 (0.064)	0.020 (0.040)	0.170** (0.078)	−0.024 (0.069)
Lagged Profitability	0.099 (0.064)	0.121*** (0.035)	0.129* (0.076)	0.147** (0.062)
Intangible assets	0.041*** (0.014)	0.235*** (0.091)	0.521*** (0.159)	−0.214 (0.146)
Insider ownership	−0.138 (0.126)	−0.156 (0.113)	−0.319 (0.306)	−0.081 (0.139)
Industry dummies	Y	N	N	N
Year dummies	Y	Y	Y	Y
Obs	27,315	27,315	7,521	7,521
Firms	2507	2507	2507	2507
Prop. firms changing	0.274	0.274	0.348	0.412
Adj/Overall R-sq	0.028	0.014	0.007	0.027

Panels A, B and C report quasi-replications of Table 4 in Lin et al. (2013), who use data from Western European and East Asian countries. Our ownership data is from ICO, merged with financials from Capital IQ and Thomson Reuters. The dependent variable is public debt ratio. The reported standard errors in parentheses below coefficients are clustered by firm. In Panels D and E we show replications of Coles et al.'s (2008) Table 5 with robust standard errors as in their specification. The dependent variable is Tobin's Q. We use data on US listed firms as in the original paper. Panel C shows subsample splits with the same number of observations as in the original paper by low and high time variation in board independence. Panel E columns (1) and (2) show OLS and within-firm estimators on the full data, while columns (3) and (4) show sampling every two and four years respectively. Summary statistics of the variables used in replications are given in Appendix 2. All variable definitions are as in the original papers. Significance symbols and thresholds used: \* for  $p < 0.1$ ; \*\* for  $p < 0.05$  and \*\*\* for  $p < 0.01$

the case of US or UK control) and therefore will be classified as either group or pyramid (the bottom two groups in Panel B of Fig. 2). Our definition of a stand-alone firm applies to small firms without any links to other legal entities. Stand-alone firms exhibit almost no time variation in ownership (results available in the online appendix). While all member firms of a corporate structure are included in the calculation of ultimate cash flow rights, in the subsequent analyses we are only interested in productive firms, because the non-productive ones are most often shell companies or individual trusts without any business activity. This still leaves in the analysis the productive firms that may have an ultimate owner that is a financial institution.

These filtering steps have a balancing role in that the differences in size, industry and listed status between Canadian firms and Anglo-Saxon MNCs become less pronounced. In particular, the variety of sizes and industry composition represented in the structures are not significantly different between the Canadian and Anglo-Saxon MNC samples (industry and size splits are presented in the online appendix). The proportions of listed firms remain

different, whereby 21% of the Canadian enterprises contain a listed firm, while 55% of the Anglo-Saxon MNCs do.<sup>24</sup>

Starting with the total counts of productive firms in the database shown at the bottom of each graph in Panel A of Fig. 2 (78,128 non-listed and 2472 listed), we focus on the pyramids and groups with Canadian (28,116) and Anglo-Saxon (17,620) control. After dropping the firms that disappear from the database in 2006 due to a higher size threshold for mandatory disclosure, we have 6033 Canadian and 4848 Anglo-Saxon firms split by family control and listed status in Panel B of Fig. 2.

In Fig. 3 we show anecdotal examples of the time variation in *ucfr* for two firms each from the Canadian (Panel A) and Anglo-Saxon MNC (Panel B) subsamples. The two Canadian firms are Celestica Inc, and Indigo Inc.—two publicly listed firms that are part of the G. W. Schwartz group. The steep drop in *ucfr* in 2001 for Indigo happens when it was acquired by G. W. Schwartz. It was initially added in the group structure with a direct stake of 0.44, which later varies as Indigo moves further down the layers of pyramiding. Similarly, the stake in Celestica, which is a subsidiary of the investment holding company arm of the group (Onex Corp.), varies as the intermediate firms in the chain of control up to the ultimate owner are reshuffled.

In Panel B of Fig. 3 we show another two listed firms controlled by US-based MNCs: Sears Canada, controlled consecutively by Sears-Roebuck and Co, and ESL Investments Inc. (the investment arm of hedge fund investor Eddie Lampert) until its bankruptcy in 2017. The second is Kronos Inc., controlled by Contran Inc.—the private holding company of the late Dallas billionaire Harold Simmons. We see that both firms maintain a relatively stable *ucfr* slightly above 0.5 in the period 1997–2004.

These examples illustrate typical economic processes behind the time variation in ownership. We now turn to a detailed analysis of the patterns of these changes for different groups of firms over time.

In our analysis a firm is considered to be family-controlled if its ultimate owner is reported as a single entity describing an individual, a family, a group of related individuals or a group of related families, as in Claessens et al. (2002).<sup>25</sup> We take great care to assign family status with as much precision as we can by verifying the data in at least five other sources (as described in the accompanying online appendix).<sup>26</sup>

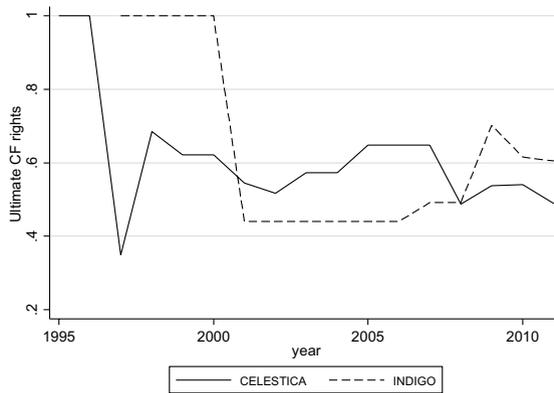
<sup>24</sup> The descriptive statistics and analysis in this section is done at the firm level, where multiple firms belong to the same enterprise (corporate structure). Here we give an idea of the proportion of enterprises having a listed firm because the disclosure requirements pertaining to listed firms likely have relevance to the entire structure under common control. The balancing achieved by focusing on the complex structures is a side product of eliminating stand-alone firms, whose ownership largely does not change over time.

<sup>25</sup> In defining family ownership, we include spouses and children, but not cousins. For example, the Bronfman family shows up in the data with two separate and distinct groups: the Charles and Edgar Bronfman group (the sons of the patriarch Samuel), which controlled Seagram until 2001; and the Peter and Edward Bronfman group (the sons of Samuel's brother Allen), which went through several metamorphoses: Edper investment company, Brookfield Asset Management, Brascan Ltd. and presently Partners Ltd.

<sup>26</sup> Firms are considered non-family-controlled if we cannot find evidence otherwise. This approach means that any measurement error in the coding of the variable family control will be against finding a difference between family and non-family firms, because presumably a number of cases that are family-controlled may have been left in the non-family category. This, however, is not a problem for us, since any differences we document will in fact be even higher, were the measurement error eliminated. State control is another important type of ownership, but because we observe very few enterprises under state control in the ownership data from ICO, we exclude them from the baseline analysis.

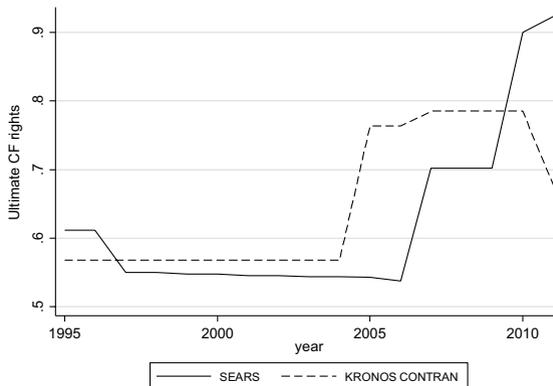
## Panel A: G.W. Schwartz

We present the time variation in ultimate cash flow rights of two listed firms that are part of the G.W. Schwartz pyramid: Celestica and Indigo Books.



## Panel B: Sears and Conran

Time variation in ultimate cash flow rights of two listed firms that are part of two foreign-controlled groups: the first is Sears Canada, controlled consecutively by Sears-Roebuck and Comp, and ESL Investments Inc. (the investment arm of hedge fund investor Eddie Lampert) until its bankruptcy in 2017. The second is Kronos Inc. controlled by Conran Inc. – the private holding company of the late Dallas billionaire Harold Simmons.



**Fig. 3** Examples of ownership changes over time

We study the patterns of time variation in ownership across jurisdictions, family and non-family status, and ownership structure complexity. These groupings are guided by the extensive literature on family firms, ownership structures and institutions. Existing literature reveals a number of reasons why family owners make ownership change decisions differently than non-family ones. For example, inheritance planning is only relevant for family owners and therefore some changes in the ownership structure of family firms are motivated by the distribution of wealth among descendants (Villalonga and Amit 2009; Tsoutsoura 2015). Family owners have a longer-term investment horizon (Bertrand and Schoar 2006) and are more risk-averse (Faccio et al. 2011), which affects the timing and

size of their ownership stake decisions. Importantly, family business groups are likely to be organized as pyramids (Almeida and Wolfenzon 2006) or employ other control enhancing mechanisms (Masulis et al. 2011).

Board independence may vary more in mature markets with stricter regulatory or institutional shareholder scrutiny. Among 26 countries which undertook corporate governance reforms in the beginning of the 2000s, only three mandated a minimum board independence threshold, while most employed a comply-or-explain model (Kim and Lu 2013). At the same time the range of values board independence can take depends directly on the current legal minimum in the respective jurisdiction—one third being most predominant in emerging markets versus one half in developed economies (Papadopoulos 2019). In addition, non-listed firms are likely to not be subject to any minimum board independence requirements, while state-owned firms may have even stricter limits (for example 90% in Sweden and 80% in Vietnam (OECD 2018)).

#### 4.1 Summary statistics and standard deviation regressions

We begin the analysis of time variation of ownership and board independence with a summary graph of yearly averages across all firms, which we show in Fig. 4. In Panel A, the cash flow rights of the ultimate owner exhibit great stability. Starting from the bottom, the group of Canadian family firms are held with the lowest stake on average, which changes the most year to year. The average ownership stake of the non-family Canadian and family Anglo-Saxon MNC groups is always close to 0.9 and fluctuates somewhat. The group with the least time variation in its average is that of Anglo-Saxon MNC non-family firms. In addition, the average of this group is always above 0.95, which suggests that the vast majority of these firms are wholly owned. The top line in Panel B shows very low variation in the average board independence of US listed firms, with no discernible adjustment following the Sarbanes–Oxley Act. On the other hand, the rest of the Anglo-Saxon countries display the gradual increase in board independence consistent with the implementation of legal minimums or comply-or-explain codes. The majority of non-listed firms start being covered by BoardEx in 2007 and show much lower average levels of independence, around 0.5, with slightly more adjustments over time for the US subsample. Overall, Fig. 4 points to little time variation in time averages within groupings but different degrees of time variation across type of firm and jurisdiction. Corresponding graphs for control rights and wedge are shown in the online appendix.

This broad view of time variation does not tell us what proportion of the observations are responsible for the movements in the average, nor whether the movements have a common source for all firms or are firm-specific; we analyze these possibilities in the following sections.

In Table 3 we present three types of summary statistics: overall means, cross-sectional standard deviations (equal to the standard deviation of the time averages for all firms) and average time standard deviations (computed over time and then averaged across firms) of ownership and board independence by subsamples. Not surprisingly, the average ownership (standard deviation) in a pyramid is lower (higher) than in a group. However, even for pyramids the means are relatively high, which means that the typical firm is wholly owned (Panel A of Table 3). For Chinese listed firms in Panel B the means are much lower, reflecting the absence of wholly owned firms. The time standard deviations are always lower than the cross-sectional ones, reflecting the sticky nature of the variables. In Panel C of Table 3 mean board independence is higher for listed than non-listed firms for all

jurisdictions, while the cross-sectional standard deviation is much higher for non-listed firms. Time standard deviations are always low as expected for a stable variable. In China, both private and state-owned firms have mean board independence very close to the legal minimum, with both cross-sectional and time standard deviations also very low, suggesting that within-firm estimators are unlikely to detect effects in Chinese data. Tests for differences in variances by subgroups are almost always statistically significant (shown in the online appendix). Corresponding tables of summary statistics and tests for differences in variances for control rights and wedge are available in the online appendix.

To test whether time variation is statistically significantly different in a multivariate setting, in Table 4 we present regressions of the time-series standard deviation of ownership and board independence on indicators for the different groupings in Table 3. The dependent variable is constructed over 4-year rolling windows for each firm. We confirm that family-controlled firms have significantly higher ownership time-series standard deviation. Across jurisdictions, Canadian firms exhibit statistically significantly higher standard deviation of ownership than Anglo-Saxon MNCs. The greater time variability of family firms is maintained through different size groupings (Panel B of Table 4). Board independence of listed firms has statistically significantly higher time-series standard deviation across jurisdictions (Panel C). The size decile analysis in Panel D reveals that the lower time variability in board independence for US listed firms is driven by the two smallest deciles.

Overall, the summary statistics and test results in Tables 3 and 4 suggest that the amount of time variation in ownership and board independence differs across jurisdictions, family control and listed status.

## 4.2 Analysis of year-to-year variation

In the next two sections we would like to see how our granular data maps onto the generalized simulation results in Fig. 1, where we show the relationship between power and proportion of firms that change between the first and next (last) period. We start with year-to-year changes and the firms responsible for these changes. In Table 5 Panel A (Panel B) we report the proportion of Canadian (Anglo-Saxon MNC) firms that change in each year. We use two definitions of change: any change  $> 0$  and changes  $> 0.05$  (as in Donelli et al. 2013), and we report non-family and family firms separately.<sup>27</sup> We also analyze the yearly proportions of changes by industry and size, but for brevity they are only reported in the online appendix. Corresponding tables with very similar results for control rights and wedge are also shown in the online appendix.

In both Panels A and B of Table 5 we note the lack of systematic spikes or drops in the proportion of firms that change in particular years or periods. The only exception is a jump in the period 2005–2006, which coincides with the changes in disclosure rules and size thresholds in CRA.

Table 5 Panel A (B) tells us that at a minimum of two years of data (that allows the computation of a change) on average we have only 13.8% (7.1%) of Canadian (Anglo-Saxon) family firms, and only 11.1% (4.3%) of Canadian (Anglo-Saxon) non-family firms, exhibiting a change in ownership. The numbers for changes greater than 5% are comparable to the results in Donelli et al. (2013) for Chile.

<sup>27</sup> Fahlenbrach and Stulz (2009) use a 2.5% cut-off for the change in managerial ownership.

Over 17 years, the proportion of firms that change at least once becomes 60.6% (45%) and 36.9% (22.8%) respectively.

In both Panels C and D of Table 5 we see a very different picture for board independence. The minimum proportion of firms that change their board independence in any year is 45%, with the overall proportion of firms that change at least once becoming 100% over the entire period. Therefore, we uncover a large difference in the time variation of ownership and board independence when it comes to the proportion of firms changing each year.

Consider the findings in Table 5 relative to the generalized simulation results in Fig. 1. The minimum average proportion of firms that change ownership between two consecutive calendar years among all subsamples is 4.3%, which accumulates to 22.8% for the full period. Distributing this accumulated change equally over the period implies a rough estimate at year 13 of 21%, which would provide sufficient statistical power for four time observations and  $\beta=0.8$ , corresponding to the right-most graph of Panel B of Fig. 1. For the other subsamples time variation is higher, and they would provide sufficient power even for a shorter time series: the 13.8% average year-to-year proportion of changing firms for the Canadian family subsample would translate to 42.1% in year 9 and have sufficient power for  $\beta$  close to 0.6.

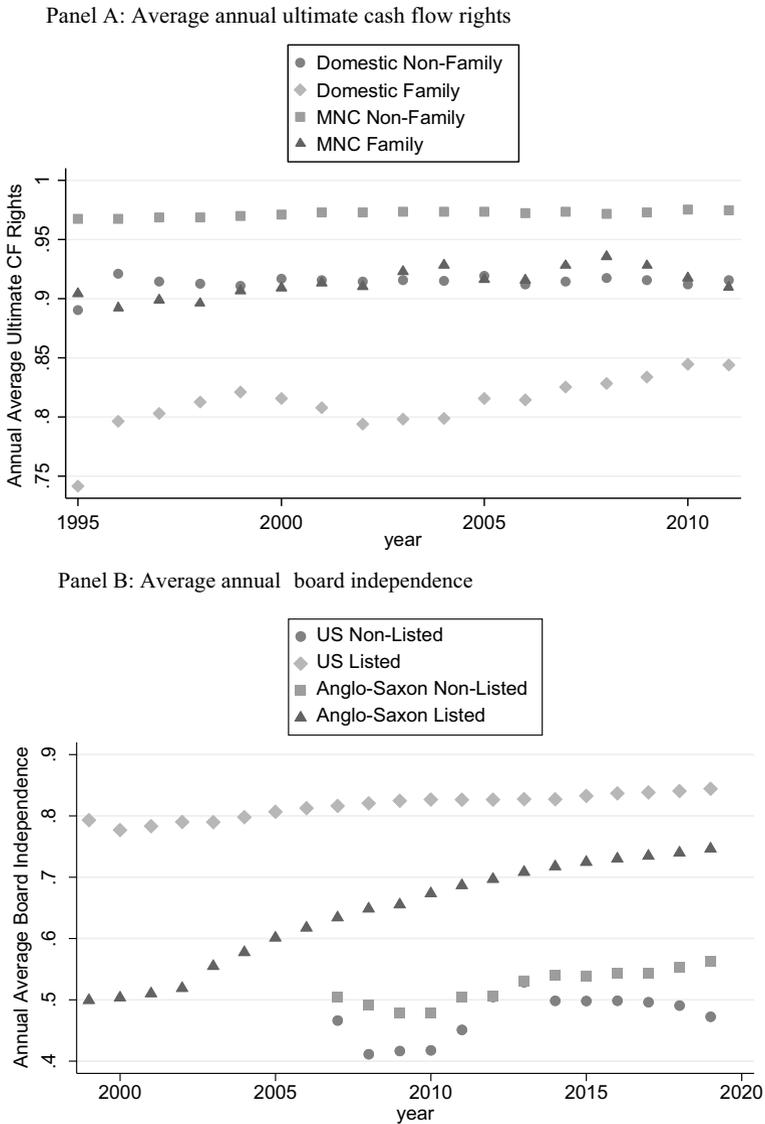
Realistically, in empirical research involving sticky data it is much more likely to have a few time observations and be in the situation where the proportion of firms that account for the firm-specific time variation is relatively low. This highlights the local nature of an estimator based on time variation for a persistent variable and necessitates a closer look at the changing firms. We examine the kind of firms that change by industry and size (results in the online appendix). The industry and size splits reveal that the most changeable firms are larger and that in all categories family firms change more than non-family firms.

### 4.3 Analysis of stable regimes

Finally, we note that the year-to-year changes in the previous section are calendar year specific. Therefore, in Table 6 we introduce the concept of a stable regime that is calendar year independent. We follow DeAngelo and Roll (2015), who examine the stability of capital structures in US listed firms over long horizons. A stable regime is defined as a period of  $t$  years, where ownership does not change outside a predetermined bandwidth. For example, in Fig. 3 we see that Indigo is in a stable regime where no changes occur to its wholly owned status from 1997 until 2001, then it enters another stable regime at ~45% until 2006 and subsequently it experiences changes every year. Indigo will then be part of the proportion of firms that are in a stable regime for at least five years, while for Celestica, the longest stable regime is three years (~65% in 2005–2007).

We examine minimum lengths of stable regimes from 3 to 16 years and changes of 0%, 5% and 10%. For example, when we require the change to be greater than 10% for the stable regime to end, Celestica will be in one for nine years (1999–2007). However, when we require the change in ownership to be less than 5% or 0%, Celestica is never in a stable regime. We limit the analysis to firms that have at least 12 years of data to make the computed proportions in each year be driven only by changes and not by the number of available firms in the sample.

In Panel A of Table 6 we report the proportion of Canadian non-family firms with stable ownership regimes. Comparing the overall reduction in stability between Canadian non-family (Panel A) and family firms (Panel B), we note that the stability of family firms drops faster. From the second row of Panel A (Panel B) we note that although 98.2% (99.1%) of



**Fig. 4** Cross-section over time. The ownership sample in Panel A consists of productive firms for which we have data for at least 12 years, while for board independence in Panel B we require a minimum of five years of data. The coverage of BoardEx for non-listed firms is very limited before 2007.

non-family (family) firms do not change their ownership structure within three years, by year 7 we see a decrease to 89.9% (84.4%), and this proportion further decreases to 67.2% (56.9%) by year 12 and 63.0% (51.6%) by year 16.

When we relax our definition of stable regime by increasing the bandwidth to 10%, we still find a clear reduction in the proportion of firms that follow stable regimes as time increases. The family effect is evident across jurisdictions, i.e., ownership stability dissipates faster for family firms.

**Table 3** Summary statistics

Panel A: CF Rights of non-listed and listed Canadian firms and Anglo-Saxon multinationals with operations in Canada

	Non-Family						Family									
	Pyramid			Group			Pyramid			Group			Stand-Alone			
	Mean	Cross-sect. St. Dev	Ave Time St. Dev	Mean	Cross-sect. St. Dev	Ave Time St. Dev	Mean	Cross-sect. St. Dev	Ave Time St. Dev	Mean	Cross-sect. St. Dev	Ave Time St. Dev	Mean	Cross-sect. St. Dev	Ave Time St. Dev	
Canadian																
Non-listed	0.843	0.219	0.066	0.955	0.119	0.015	0.709	0.291	0.123	0.123	0.291	0.127	0.074	0.885	0.167	0.111
Listed	0.931	0.124	0.043	0.968	0.095	0.030	0.634	0.263	0.166	0.166	0.263	0.106	0.086	0.843	0.230	0.151
Anglo-Saxon																
Non-listed	0.914	0.162	0.033	0.992	0.043	0.007	0.858	0.220	0.074	0.074	0.220	0.065	0.033	0.978	0.056	0.031
Listed	0.943	0.139	0.019	0.997	0.027	0.003	0.830	0.226	0.084	0.084	0.226	0.049	0.044	0.841	0.212	0.041

Panel B: CF Rights of listed firms in China

	Private				State-owned			
	Mean	Cross-sect. St. Dev	Ave time St. Dev	Firm count	Mean	Cross-sect. St. Dev	Ave time St. Dev	Firm count
China—listed firms	0.332	0.149	0.059	1613	0.368	0.151	0.066	1000

Panel C: Board independence of non-listed and listed firms covered by BoardEx

	Non-listed				Listed			
	Mean	Cross-sect. St. Dev	Ave Time St. Dev	Firm Count	Mean	Cross-sect. St. Dev	Ave Time St. Dev	Firm Count
US	0.549	0.347	0.066	5217	0.812	0.094	0.048	7342
Anglo-Saxon	0.528	0.311	0.085	1970	0.675	0.185	0.075	5258
Cont Europe	0.653	0.353	0.048	667	0.784	0.166	0.055	2865
Rest of the World	0.644	0.309	0.066	723	0.679	0.208	0.054	4993

**Table 3** (continued)

Panel D: Board independence of listed firms in China

	Private			State-owned				
	Mean	Cross-sect. St. Dev	Ave time St. Dev	Firm count	Mean	Cross-sect. St. Dev	Ave time St. Dev	Firm count
China—Listed firms	0.375	0.044	0.025	1613	0.364	0.041	0.031	1000

Panels A and B present summary statistics of firm ownership (CF rights) and board independence by subsamples. Cross-sectional standard deviation is calculated as the st. dev of the time means of all firms, while average standard deviation over time is the mean of the time st. dev-s of each firm. The summary statistics for Control rights and Wedge, as well as CF Rights for Continental Europe and Rest of the World, are presented in the online appendix. The ownership sample consists of productive firms only (corresponding to the overall firm counts in Panel B of Fig. 2). Panel C shows the same measures for board independence of the firms covered by the BoardEx database (containing the most accurate and comprehensive global board and relationship intelligence collected and curated over 20 years), while Panel D reports on Chinese listed firms covered by the CSMAR database

**Table 4** Regressions of time-series standard deviation of ownership and board independence

Panel A: Ownership full sample						
	CF		Control		Wedge	
Family	0.007***		0.003**		0.005***	
	(0.002)		(0.001)		(0.001)	
Ang-Sax	-0.011***		-0.007***		-0.005***	
	(0.001)		(0.001)		(0.001)	
Family # Ang-Sax	0.001		0.000		-0.000	
	(0.002)		(0.002)		(0.002)	
Pyramid	0.020***		0.010***		0.015***	
	(0.001)		(0.001)		(0.001)	
Family # Pyramid	0.005**		0.004**		0.013***	
	(0.002)		(0.002)		(0.002)	
Ang-Sax # Pyramid	-0.009***		-0.005***		-0.007***	
	(0.002)		(0.001)		(0.001)	
Family # Ang-Sax # Pyramid	0.004		-0.006***		0.002	
	(0.003)		(0.002)		(0.003)	
Obs	136,903		136,903		136,903	
Firms	12,253		12,253		12,253	
Adj R-sq	0.053		0.028		0.056	
Panel B: Ownership by jurisdiction and size						
	CA CF	Ang-Sax CF	CA Control	Ang-Sax Control	CA Wedge	Ang-Sax Wedge
Size groups = 2	0.006***	0.005***	-0.001	0.002***	0.007***	0.004***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Size groups = 3	0.012***	0.006***	0.005**	0.003***	0.011***	0.004***
	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Size groups = 4	0.032***	0.011***	0.017***	0.001	0.020***	0.010***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Family	0.007***	0.010***	0.003**	0.007***	0.009***	0.009***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Size groups = 2 # Family	0.007**	0.000	0.005**	-0.007***	0.004*	0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Size groups = 3 # Family	0.016***	0.023***	0.007**	-0.010***	0.013***	0.027***
	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)	(0.004)
Size groups = 4 # Family	-0.000	-0.002	0.006*	-0.008***	0.018***	-0.000
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Obs	75,103	61,800	75,103	61,800	75,103	61,800
Firms	7369	6015	7369	6015	7369	6015
Adj R-sq	0.041	0.032	0.029	0.006	0.057	0.043
Panel C: Board Independence full sample						
	Board Independence					
US	0.014***					
	(0.004)					
Ang-Sax	0.023***					

**Table 4** (continued)

Panel C: Board Independence full sample		Board Independence
		(0.004)
Rest of the World		0.020*** (0.006)
Listed		-0.010*** (0.003)
US # Listed		-0.020*** (0.004)
Ang-Sax # Listed		-0.006 (0.004)
Rest of the World # Listed		-0.011* (0.006)
Obs		145,626
Firms		18,537
Adj R-sq		0.043
Panel D: Board Independence by jurisdiction and size		
	US	Ang-Sax
Size Decile = 1 # Listed	-0.036*** (0.009)	0.026*** (0.008)
Size Decile = 2 # Listed	-0.049*** (0.012)	0.006 (0.007)
Size Decile = 3 # Listed	-0.010 (0.014)	-0.009 (0.009)
Size Decile = 4 # Listed	0.013 (0.009)	-0.005 (0.008)
Size Decile = 5 # Listed	0.006 (0.010)	-0.027*** (0.010)
Size Decile = 6 # Listed	0.002 (0.011)	-0.012 (0.010)
Size Decile = 7 # Listed	0.007 (0.011)	-0.002 (0.009)
Size Decile = 8 # Listed	0.006 (0.012)	0.001 (0.009)
Size Decile = 9 # Listed	0.006 (0.008)	-0.004 (0.007)
Obs	54,726	34,111
Firms	5849	4318
Adj R-sq	0.092	0.077

We present regressions of the time-series standard deviation of ownership and board independence on indicators for jurisdiction, family control, ownership structure complexity, listed status and size deciles. Time-series standard deviation is calculated over rolling windows of 4 years for each firm. The ownership sample consists of productive firms only (corresponding to the overall firm counts in Panel B of Fig. 2). Panel A shows full sample analysis, while in Panel B we run the regressions separately by jurisdiction. Size groups in the ownership data are proxied by the number of subsidiaries in each structure (group 1 for 10 subsidiaries or less; group 2 for 11–25 subsidiaries; group 3 for 26–50 and group 4 for 51 or more), a measure which we confirm has a high correlation with total assets for listed firms. Panels C and D show the corresponding regressions for board independence of the firms covered by the BoardEx database. Size deciles in the board data are formed based on total assets, where we have merged the BoardEx data with CapitalIQ using ISIN codes. Significance symbols and thresholds used: \* for  $p < 0.1$ ; \*\* for  $p < 0.05$  and \*\*\* for  $p < 0.01$

**Table 5** Year-to-year variation in ownership

Panel A: Proportion of changes in Ownership of Canadian Firms

Year	Non-family		Family		Test <i>p</i> -val	
	Change > 0	Firm Count	Change > 0	Firm Count	Diff change > 0	Diff change > 0.05
1996	0.109	2096	0.191	2544	0.000	0.000
1997	0.097	2222	0.124	2740	0.003	0.004
1998	0.099	2328	0.133	2939	0.000	0.000
1999	0.090	2495	0.136	3151	0.000	0.041
2000	0.083	2619	0.125	3372	0.000	0.002
2001	0.111	2638	0.154	3480	0.000	0.080
2002	0.073	2640	0.096	3555	0.001	0.391
2003	0.089	2627	0.112	3547	0.004	0.790
2004	0.087	2691	0.097	3604	0.154	0.003
2005	0.103	2750	0.199	3631	0.000	0.000
2006	0.165	2496	0.139	3505	0.005	0.000
2007	0.166	2416	0.197	3468	0.002	0.289
2008	0.144	2256	0.113	3413	0.001	0.000
2009	0.154	2218	0.118	3312	0.000	0.000
2010	0.121	2140	0.164	3222	0.000	0.246
2011	0.083	2042	0.102	3155	0.022	0.141
Average 1996–2011	0.111	2417	0.138	3290	0.003	0.866
Proportion of firms	0.450	3414	0.606	4126	0.000	0.000
Proportion of obs	0.105	40,576	0.132	54,844	0.000	0.734

**Table 5** (continued)  
 Panel B: Proportion of changes in Ownership of Anglo-Saxon Firms

Year	Non-family		Family		Test <i>p</i> -val			
	Change > 0	Change > 0.05	Firm Count	Change > 0	Change > 0.05	Firm Count	Diff change > 0	Diff change > 0.05
1996	0.018	0.014	3372	0.050	0.019	639	0.000	0.351
1997	0.029	0.024	3562	0.069	0.045	685	0.000	0.002
1998	0.030	0.022	3714	0.065	0.051	727	0.000	0.000
1999	0.026	0.022	3938	0.089	0.085	764	0.000	0.000
2000	0.032	0.022	4098	0.041	0.022	785	0.238	0.924
2001	0.034	0.029	4076	0.083	0.056	804	0.000	0.000
2002	0.031	0.028	4120	0.025	0.022	807	0.322	0.332
2003	0.027	0.024	4122	0.062	0.059	818	0.000	0.000
2004	0.033	0.029	4132	0.042	0.042	824	0.182	0.039
2005	0.042	0.037	4167	0.059	0.034	843	0.024	0.718
2006	0.081	0.069	4050	0.098	0.087	804	0.118	0.073
2007	0.085	0.081	3826	0.132	0.111	772	0.000	0.006
2008	0.065	0.058	3695	0.107	0.095	756	0.000	0.000
2009	0.062	0.055	3473	0.095	0.088	739	0.001	0.001
2010	0.055	0.043	3274	0.079	0.076	720	0.012	0.000
2011	0.035	0.015	3172	0.033	0.028	673	0.736	0.022
Average 1996–2011	0.043	0.036	3799	0.071	0.058	760	0.001	0.005
Proportion of firms	0.228	0.182	4867	0.369	0.315	1110	0.000	0.000
Proportion of obs	0.041	0.034	63,870	0.068	0.055	12,737	0.000	0.000

**Table 5** (continued)  
 Panel C: Proportion of changes in Board Independence of US Firms

Year	Non-listed		Listed		Test <i>p</i> -val	
	Change > 0	Change > 0.05	Change > 0	Change > 0.05	Firm count	Firm count
2000			0.986	0.963	1498	
2001			0.650	0.415	1823	
2002			0.574	0.296	1888	
2003			0.766	0.580	3212	
2004			0.623	0.380	3834	
2005			0.585	0.320	4173	
2006			0.522	0.271	4339	
2007	0.958	0.946	0.514	0.247	4396	0.000
2008	0.770	0.721	0.501	0.230	4168	0.000
2009	0.660	0.598	0.462	0.219	4047	0.000
2010	0.664	0.597	0.465	0.206	4007	0.000
2011	0.717	0.669	0.457	0.197	3957	0.000
2012	0.712	0.646	0.459	0.186	3894	0.000
2013	0.670	0.593	0.490	0.219	3876	0.000
2014	0.681	0.593	0.494	0.221	3938	0.000
2015	0.709	0.623	0.504	0.223	3947	0.000
2016	0.598	0.497	0.496	0.187	3710	0.000
2017	0.609	0.511	0.481	0.173	3505	0.000
2018	0.618	0.511	0.500	0.163	3307	0.000
2019	0.638	0.531	0.521	0.160	3121	0.000
Average 2000– 2019	0.732	0.669	0.553	0.293	3532	0.000

Table 5 (continued)

Panel C: Proportion of changes in Board Independence of US Firms									
Year	Non-listed			Listed			Test $p$ -val		
	Change > 0	Change > 0.05	Firm count	Change > 0	Change > 0.05	Firm count	Diff change > 0	Diff change > 0.05	Diff change > 0.05
Proportion of firms	1.000	1.000	3428	1.000	1.000	5956			
Proportion of obs	0.669	0.585	25,725	0.533	0.267	70,722	0.000	0.000	0.000
Panel D: Proportion of changes in Board Independence of Anglo-Saxon Firms									
Year	Non-listed			Listed			Test $p$ -val		
	Change > 0	Change > 0.05	Firm count	Change > 0	Change > 0.05	Firm count	Diff change > 0	Diff change > 0.05	Diff change > 0.05
2000				0.750	0.709	673			
2001				0.709	0.655	831			
2002				0.660	0.618	924			
2003				0.692	0.636	1197			
2004				0.673	0.602	1519			
2005				0.645	0.577	1860			
2006				0.617	0.542	2137			
2007	0.944	0.907	107	0.570	0.501	2372	0.000	0.000	0.000
2008	0.796	0.746	181	0.527	0.433	2450	0.000	0.000	0.000
2009	0.749	0.656	215	0.494	0.399	2415	0.000	0.000	0.000
2010	0.706	0.616	245	0.534	0.430	2546	0.000	0.000	0.000

**Table 5** (continued)

Panel D: Proportion of changes in Board Independence of Anglo-Saxon Firms

Year	Non-listed		Listed		Test <i>p</i> -val	
	Change > 0	Change > 0.05	Firm count	Change > 0	Change > 0.05	Firm count
2011	0.733	0.666	326	0.512	0.407	2628
2012	0.785	0.707	525	0.502	0.382	2660
2013	0.730	0.648	708	0.513	0.396	2834
2014	0.660	0.589	890	0.502	0.371	2944
2015	0.740	0.676	1260	0.501	0.379	3082
2016	0.608	0.547	1247	0.461	0.315	2985
2017	0.622	0.540	1232	0.443	0.301	2862
2018	0.617	0.531	1213	0.451	0.306	2750
2019	0.604	0.512	1138	0.436	0.290	2590
Average 2000–2019	0.565	0.517	465	0.560	0.462	2213
Proportion of firms	1.000	1.000	1295	1.000	1.000	4143
Proportion of obs	0.671	0.595	9307	0.534	0.428	44,707

This table covers productive firms for which we have data for at least 12 years and which have a group or a pyramid ownership structure (corresponding to the overall firm counts in Panel B of Fig. 2; results for the remaining firms are in the online appendix). In panels C and D data for non-listed firms begins only in 2006 and is initially very sparse

**Table 6** Stable regimes in ownership

Panel A: Stable Ownership Regimes of Canadian Non-Family Firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
(1) = 0	0.971	0.955	0.938	0.903	0.866	0.821	0.778	0.729	0.690	0.612	0.609	0.606	0.581	0.567	13	12.09
(2) < = 0.05	0.982	0.975	0.962	0.938	0.899	0.862	0.824	0.786	0.744	0.672	0.671	0.664	0.636	0.630	13	12.66
(3) < = 0.10	0.982	0.975	0.964	0.944	0.916	0.882	0.844	0.803	0.764	0.689	0.687	0.682	0.657	0.652	13	12.84
Decreases in line (2)	0.008	0.012	0.025	0.025	0.039	0.037	0.037	0.039	0.042	<b>0.072</b>	0.001	0.006	0.028	0.006		
# of firms	3697	3697	3697	3697	3697	3697	3697	3697	3697	3697	3123	2709	2273	1945		
Panel B: Stable Ownership Regimes of Canadian Family Firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
(1) = 0	0.967	0.926	0.875	0.813	0.751	0.687	0.629	0.580	0.539	0.485	0.469	0.441	0.439	0.437	11	10.79
(2) < = 0.05	0.991	0.971	0.945	0.899	0.844	0.775	0.722	0.671	0.620	0.569	0.549	0.519	0.517	0.516	12	11.81
(3) < = 0.10	0.993	0.984	0.965	0.918	0.862	0.810	0.764	0.707	0.653	0.603	0.584	0.556	0.554	0.550	13	12.19
Decreases in line (2)	0.020	0.026	0.026	0.047	0.054	<b>0.070</b>	0.052	0.052	0.050	0.051	0.021	0.029	0.002	0.002		
# of firms	4542	4542	4542	4542	4542	4542	4542	4542	4542	4542	4036	3619	3146	2733		
Panel C: Stable Ownership Regimes of Anglo-Saxon Non-Family Firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
(1) = 0	0.983	0.975	0.967	0.949	0.933	0.908	0.889	0.864	0.830	0.778	0.767	0.752	0.737	0.734	14	13.58
(2) < = 0.05	0.984	0.978	0.972	0.961	0.951	0.932	0.914	0.896	0.861	0.815	0.805	0.795	0.780	0.783	15	13.92
(3) < = 0.10	0.984	0.978	0.972	0.962	0.952	0.934	0.917	0.899	0.864	0.819	0.810	0.800	0.786	0.788	15	13.96
Decreases in line (2)	0.006	0.006	0.011	0.011	0.009	0.019	0.018	0.018	0.035	<b>0.046</b>	0.011	0.009	0.015	-0.003		

Table 6 (continued)

Panel C: Stable Ownership Regimes of Anglo-Saxon Non-Family Firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
# of firms	4354	4354	4354	4354	4354	4354	4354	4354	4354	4354	3807	3332	2912	2493		
Panel D: Stable Ownership Regimes of Anglo-Saxon Family Firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
(1)=0	0.977	0.958	0.929	0.896	0.866	0.833	0.800	0.754	0.709	0.654	0.612	0.592	0.559	0.518	13	12.38
(2)<=0.05	0.977	0.963	0.937	0.914	0.889	0.861	0.830	0.796	0.758	0.705	0.660	0.644	0.617	0.581	14	12.84
(3)<=0.10	0.977	0.964	0.938	0.914	0.891	0.870	0.839	0.803	0.767	0.713	0.670	0.659	0.629	0.591	14	12.93
Decreases in line (2)		0.014	0.026	0.023	0.025	0.027	0.032	0.034	0.037	<b>0.054</b>	0.044	0.016	0.027	0.036		
# of firms	915	915	915	915	915	915	915	915	915	915	801	728	671	575		
Panel E: Stable Regimes in Board Independence of US Listed firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
(1)=0	0.932	0.712	0.501	0.345	0.238	0.163	0.121	0.092	0.063	0.047	0.036	0.025	0.014	0.007	5	5.28
(2)<=0.05	0.999	0.977	0.937	0.872	0.796	0.704	0.611	0.519	0.443	0.370	0.316	0.270	0.239	0.200	10	10.25
(3)<=0.10	0.999	0.992	0.980	0.945	0.897	0.830	0.755	0.685	0.619	0.547	0.503	0.462	0.425	0.388	12	12.16
Decreases in line (2)		0.022	0.040	0.065	0.077	0.092	<b>0.093</b>	0.093	0.076	0.073	0.054	0.045	0.032	0.038		

Table 6 (continued)

Panel E: Stable Regimes in Board Independence of US Listed firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
# of firms	2939	2939	2939	2939	2939	2939	2939	2939	2939	2939	2732	2465	2201	1929		
Panel F: Stable Regimes in Board Independence of Anglo-Saxon Listed firms																
	3y	4y	5y	6y	7y	8y	9y	10y	11y	12y	13y	14y	15y	16y	Median # of years	Mean # of years
(1)=0	0.911	0.687	0.472	0.324	0.227	0.175	0.136	0.111	0.097	0.089	0.086	0.098	0.078	0.032	4	5.44
(2) < =0.05	0.964	0.845	0.682	0.548	0.432	0.349	0.277	0.232	0.195	0.168	0.156	0.149	0.119	0.071	6	7.06
(3) < =0.10	0.984	0.911	0.801	0.679	0.557	0.469	0.390	0.322	0.274	0.240	0.217	0.210	0.180	0.131	7	8.20
Decreases in line (2)	0.119	<b>0.163</b>	0.134	0.116	0.083	0.072	0.045	0.037	0.037	0.027	0.012	0.007	0.030	0.049		
# of firms	1717	1717	1717	1717	1717	1717	1717	1717	1717	1717	1513	1260	1024	779		

This table covers productive firms for which we have data for at least 12 years (corresponding to the overall firm counts in Panel B of Fig. 2; results for the remaining firms are in the online appendix). Ownership is ultimate cash flow rights of the controlling owner. We report the proportion of firms with the longest stable regime (based on each ownership range specified in the rows) having lasted at least as many years as specified in the columns

The median and mean number of years in a stable regime reported in the last two columns of Table 6 also show that ownership stability is lower for family firms in both the Canadian and Anglo-Saxon MNC sample, as is board independence stability for Anglo-Saxon versus US firms.

On row four of all panels in Table 6 we report the year-to-year decrease in stability under the 0.05 bandwidth. In bold we have shown the years with the biggest drop. Using this metric, we find yet again that the biggest drop happens sooner (year 8) for family than non-family firms (year 12) in the Canadian sample. In Panels E and F of Table 6 we note that for board independence the largest drop in stability for US firms happens in year 9 (equal to the drop in year 10), while for Anglo-Saxon firms it is in year 5. Both groups begin with very high proportions of three-year stability, just as in the case of ownership, but the reductions are much larger, and by year 16 the proportion of firms still in a stable regime has dropped to 0 on row 1, whereas for ownership that proportion stays around or above 50% (Panels A–D of Table 6).

The stable regimes analysis shows that time variation accrues faster for some kinds of firms (Canadian family firms in the case of ownership and Anglo-Saxon firms for board independence). While within seven years close to 90% of non-family firms remain in stable ownership regimes, five more years reduces this proportion to two thirds. Stable regime tables for control rights and wedge are presented in the online appendix.

#### 4.4 Analysis of variance and persistence regressions

In Table 7 we show disaggregation of the variation in ownership and board independence by firm fixed effects, year fixed effects, industry-year, industry-size deciles and firm-time interactions, following DeAngelo and Roll (2015) and Lemmon et al. (2008). The first column shows the adjusted R-sq of each model, while the remaining columns report the relative proportion of variation explained by each set of fixed effects. Exact regression specifications and methodological details are presented in the online appendix. The proportion of explained variation in ownership accounted for by firm-time interactions is substantial in cash flow rights: 0.1325 for Canadian non-family firms and 0.1756 for family firms; 0.1814 for Anglo-Saxon MNC non-family firms and 0.2808 for Anglo-Saxon MNC family firms. The results are similar for control rights, wedge and board independence. For comparison, industry-time interactions are ten times less important for all groups. The analysis-of-variance results tell us that over a longer period of 25 years the firm-specific time varying component of ownership is important—accounting for a just under a fifth of all explained variation for family Canadian firms and close to a third for family Anglo-Saxon MNC firms. Therefore, the within-firm estimator should not be dismissed on the grounds of lack of firm-specific time variation. In contrast, the common-to-all-firms time series variation, industry-specific time variation and the one captured by industry-size deciles are all negligible. Existing studies often argue that using industry, time, size deciles and their interactions is a way to address the inability to use firm fixed effects. Our results, however, show that this approach does not come close to capturing the variation accounted for by firm fixed effects.

The persistence regressions in Table 8 follow Graham et al. (2020), where ownership and board independence are regressed on their initial level. The statistically significant and large coefficients confirm the stability we have documented in the previous analyses, as well as the fact that it differs across groups. For example, in Panel A we see that cash flow rights for Canadian non-family firms change by between 0.65 and 0.7 for a unit level

change in the initial value, while this is only between 0.44 and 0.47 for family firms. The initial level coefficients for board independence in Panel B similarly show different degree of stability across jurisdictions. The lower persistence for US listed firms is evident in line with the time-series standard deviation results in Table 4.

#### 4.5 Summary of empirical analysis

In the last part of this study we show the different time variation patterns in ownership and board independence manifesting in different subgroups by jurisdiction and type of control. In year averages we find that both variables vary little for most subgroups (Fig. 4), and the standard deviation over time is lower than that computed cross-sectionally (Table 3). The year-to-year analysis shows that the proportion of firms that change between any two consecutive calendar years is different across subgroups (Table 5), with dramatically higher year-to-year variation for board independence than ownership. On the other hand, the stable regimes are calendar time independent and indicate that a relatively shorter period is required for time variation in ownership to accrue for Canadian family firms and longer for non-family firms and Anglo-Saxon MNCs, as well as for board independence of Anglo-Saxon versus US firms (Table 6). We confirm these descriptive findings in multivariate analyses. In particular, the time-series standard deviation of ownership is higher for family firms and pyramid structures, but lower for Anglo-Saxon firms (Table 4). In terms of listed status, the time-series standard deviation of board independence of non-US firms is higher relative to non-listed ones. Variance decomposition and persistence regressions (Tables 7 and 8) similarly demonstrate the different patterns of time variation revealed in the descriptive findings. Therefore, we confirm that the time variation in ownership and board independence exhibits different patterns across subgroups of firms, with the subsamples with higher time variation corresponding to sufficient power in the simulation results.

### 5 Conclusion

Recent methodological and editorial guidance calls for improving causal inference and ruling out alternative explanations in all sub-disciplines of business research. This can be done either through a source of exogenous variation or through careful and exhaustive tests that convincingly support the baseline results. Very often these approaches rely on time variation. For example, existing studies where controlling ownership or board independence is an explanatory variable often find it impossible to employ within-firm estimators (that address time-invariant endogeneity) because of the lack of time variation. Based on theoretically and empirically grounded rules, we simulate artificial data reflecting the evident stickiness of such variables in existing work and analyze the power properties of within-firm estimators under different degrees of time variation. The results provide guidance on key elements of the empirical design of governance research. Responding to common challenges, such as difficulty of data collection, we derive relationships between lengths of time and gaps in data collection on the one hand and statistical power on the other. Our results are useful for a variety of research settings in any jurisdiction and can be employed as a benchmark to assess whether time variation is sufficient to detect a relationship if it exists.

When data needs to be hand collected or extensively pre-processed, a researcher may begin by examining the proportion of firms that change between two time points. Our simulations suggest that for a range of strengths of the theoretical relationship of interest, when

**Table 7** Analysis of variance

Panel A: Cash flow rights – Canada

	Adj R-sq	Proportion of explained variation in ownership accounted for by:				
		Firm FEs	Year FEs	Ind*Year FEs	Ind*Size decile FEs	Firm*Time FEs
(1)	(2)	(3)	(4)	(5)	(6)	
<b>Non-family</b>						
(1)	0.7883	0.9880	0.0034			
(2)	0.7906	0.9590	0.0012	0.0050		
(3)	0.7972	0.8749	0.0046		0.0114	
(4)	0.9003	0.5204	0.0004			0.1325
<b>Family</b>						
(1)	0.7123	0.9674	0.0200			
(2)	0.7194	0.9026	0.0049	0.0144		
(3)	0.7281	0.7681	0.0214		0.0217	
(4)	0.8473	0.5137	0.0034			0.1756

Panel B: Cash flow rights—Anglo-Saxon

	Adj R-sq	Proportion of explained variation in ownership accounted for by:				
		Firm FEs	Year FEs	Ind*Year FEs	Ind*Size decile FEs	Firm*Time FEs
(1)	(2)	(3)	(4)	(5)	(6)	
<b>Non-family</b>						
(1)	0.7205	0.9730	0.0047			
(2)	0.7316	0.8931	0.0010	0.0180		
(3)	0.7346	0.7957	0.0051		0.0192	
(4)	0.8696	0.4698	0.0007			0.1814
<b>Family</b>						
(1)	0.5975	0.9346	0.0147			
(2)	0.6225	0.7913	0.0121	0.0675		
(3)	0.6780	0.6688	0.0142		0.1148	
(4)	0.8256	0.3937	0.0021			0.2808

Panel C: Board Independence

	Adj R-sq	Proportion of explained variation in board independence accounted for by:				
		Firm FEs	Year FEs	Ind*Year FEs	Ind*Size decile FEs	Firm*Time FEs
(1)	(2)	(3)	(4)	(5)	(6)	
(1)	0.6915	0.9632	0.0600			
(2)	0.6967	0.8562	0.0162	0.0135		
(3)	0.6895	0.7139	0.0528		0.0089	
(4)	0.8319	0.6166	0.0080			0.1885

This table reports adjusted R-squares and proportions of explained variation in ownership accounted for by firm and year dummies (1), and three sets of interactions: industry\*year (2), industry\*size deciles (3) and firm\*time (4). The estimated Eqs. (1)–(4) are given in the online appendix. We show the Type III partial sum of squares for each effect in the model and then normalize it by the sum across the effects, forcing columns 2 through 6 on each row to sum to one. The ownership sample consists of productive firms only

**Table 8** Persistence regressions

Panel A: CF rights		CA Non-F	CA Non-F	CAF	CAF	Ang-Sax Non-F	Ang-Sax Non-F	Ang-Sax F	Ang-Sax F	Cont EU Non-F	Cont EU Non-F	Cont EU F	Cont EU F
Initial level	0.703 <sup>***</sup> (0.008)	0.647 <sup>***</sup> (0.008)	0.464 <sup>***</sup> (0.011)	0.438 <sup>***</sup> (0.010)	0.551 <sup>***</sup> (0.018)	0.511 <sup>***</sup> (0.018)	0.440 <sup>***</sup> (0.032)	0.425 <sup>***</sup> (0.033)	0.540 <sup>***</sup> (0.027)	0.524 <sup>***</sup> (0.029)	0.353 <sup>***</sup> (0.035)	0.384 <sup>***</sup> (0.042)	
Year FEs	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	
Year × Ind × size decile FEs	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	
Obs	196,678	196,653	77,761	77,638	63,369	63,070	12,211	11,686	19,610	19,141	5787	5251	
Firms	12,361	12,361	5770	5767	4143	4140	902	898	1203	1198	374	366	
Adj R-sq	0.543	0.573	0.303	0.367	0.450	0.490	0.283	0.373	0.446	0.508	0.225	0.443	

Panel B: Board independence		US Private	US Listed	US Listed	Ang-Sax Private	Ang-Sax Private	Ang-Sax Listed	Ang-Sax Listed	Cont EU Private	Cont EU Private	Cont EU Listed	Cont EU Listed	Rest World	Rest World	Rest World	
Initial level	0.885 <sup>***</sup> (0.007)	0.679 <sup>***</sup> (0.026)	0.669 <sup>***</sup> (0.015)	0.622 <sup>***</sup> (0.017)	0.882 <sup>***</sup> (0.010)	0.768 <sup>***</sup> (0.154)	0.717 <sup>***</sup> (0.011)	0.643 <sup>***</sup> (0.045)	0.905 <sup>***</sup> (0.032)	0.868 <sup>***</sup> (0.044)	0.738 <sup>***</sup> (0.026)	0.724 <sup>***</sup> (0.026)	0.878 <sup>***</sup> (0.023)	0.858 <sup>***</sup> (0.028)	0.865 <sup>***</sup> (0.010)	0.857 <sup>***</sup> (0.010)
Year FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year × Industry FEs	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Year × Ind × size decile FEs	N	Y	N	Y	N	Y	N	Y	N	N	N	N	N	N	N	N
Obs	21,101	6643	68,784	45,463	7800	258	44,065	5830	1977	1885	26,019	25,957	2339	2199	33,829	33,765
Firms	2472	690	5349	3045	934	32	3724	577	248	245	1956	1954	297	288	3561	3560
Adj R-sq	0.784	0.608	0.478	0.544	0.773	0.828	0.602	0.522	0.794	0.804	0.562	0.569	0.766	0.790	0.753	0.756

This table shows regressions of ownership and board independence on their starting levels with different sets of fixed effects. Standard errors clustered by firm are reported below coefficients. Panel A uses ICO data for productive (excluding financial, non-profits and government bodies) firms for which we have data for at least 12 years. Family control (denoted as F vs Non-F in column headings) is defined in the Appendix. In ICO Canada has the largest coverage and is shown separately from the remaining Anglo-Saxon countries. Panel B uses BoardEx data where US has the largest coverage and is shown separately from the remaining Anglo-Saxon countries. Significance symbols and thresholds used: \* for  $p < 0.1$ ; \*\* for  $p < 0.05$  and \*\*\* for  $p < 0.01$

sampling ownership data three years apart, a minimum of between 25 and 55% of firms in the data should change between two neighboring time points for sufficient statistical power. For board independence, sampling data two years apart requires between 15 and 50% of firms changing.

Even when a project uses an electronic data source, which does not present data collection challenges, our findings are useful in judging whether a finding of no significance is due to lack of power as opposed to ownership or board independence being unrelated to the outcome of interest.

We illustrate the simulation findings with quasi-replications of seminal studies of ownership (Lin et al. 2013) and board independence (Coles et al. 2008). In the case of ownership, a unique granular database (ICO) allows the exact computation of variables for consecutive years over a relatively longer time period and a large number of firms. For board independence, we use the most popular source with the widest coverage (BoardEx). Based on our replications, we confirm existing results that firms controlled by an Anglo-Saxon jurisdiction often have insufficient variation in ownership for statistical power purposes; however, non-Anglo-Saxon firms exhibit more time variation that allows the use of within-firm estimators. For board independence of US listed firms, a relationship is detectable only if the time series is sufficiently long or, for a shorter series with gaps, if the proportion of firm changing is sufficient. We further establish that there are indeed different degrees of time variation in governance variables across type of control, jurisdiction and complexity of ownership structure, which supports the usefulness of the simulation findings. When researchers use our findings for empirical design decisions and data collection strategies, they should condition these choices on the specific characteristics of their data as we do here.

Beyond methodological guidance addressing a particular problem with governance data, our findings open new avenues for future research. The fact that the amount of time variation in ownership and board independence differs across jurisdictions, listed status, family control, etc., invites a natural question about the drivers of this greater time variation.<sup>28</sup> At least three possible theoretical explanations can be explored. Greater decision making flexibility for single/concentrated versus multiple/dispersed owners (Bhaumik et al. 2010) is consistent with more frequent adjustments in governance mechanisms in response to arising technological and competitive opportunities or regulatory and commodity pricing shocks (Graham et al. 2020). The joint pursuit of business and personal objectives by decision makers can manifest in governance changes associated with personal inheritance and tax planning (Carney et al. 2014; Tsoutsoura 2015) or reputational concerns (Belenzon et al. 2019; Deephouse and Jaskiewicz 2013). If certain types of controllers are more likely to expropriate outside shareholders, they may opportunistically adjust their own exposure to the cash flows generated by a firm they control in the expectation of future negative performance (Bertrand et al. 2002).

Our generalized simulation results can be used to revisit many of the research questions in existing literature, where adding one more time observation or sampling further apart can allow time-variation-based methods to be employed. Further, by being able to exploit the time dimension in the data, many new research projects are more likely to become viable. In terms of new avenues of research where time variation in ownership can

---

<sup>28</sup> It is beyond the scope of this study to establish what the drivers of time variation are. In a companion paper, we focus on regulatory and commodity price shocks that have affected particular industries and/or jurisdictions to explore different channels.

be useful, consider the adoption of peer-to-peer technology that has democratized access to startup and secondary equity financing. This recent phenomenon can inform the long-lasting debate on whether business groups or other forms of concentrated corporate control are compensating for institutional weakness or hampering innovation and growth (Khanna and Yafeh 2007; Morck 2005). Researchers can revisit this question by studying the effect of changes in ownership patterns in industries that are relatively less costly for new entrants (software services, call centers) in jurisdictions with different degrees of institutional weakness, on firm-level innovation, growth and value added. Similarly, time-varying ownership data can be used to test whether the advent of block-chain technology, and the unprecedented degree of transparency of business transactions (including in ownership stakes) it allows, reduce the incentives to maintain complex ownership structures (Yermack 2017).

In terms of board independence, we know from Graham et al. (2020) that it exhibits long-term dynamics for US listed firms. Given regulatory and institutional constraints within other jurisdictions and different incentives for private as opposed to listed firms, further theories of bargaining and/or dynamic contracting become testable in data with manageable time series lengths.

## Appendix 1 Variable definitions of ownership variables

Variable/concept	Definition	Source
Enterprise	A collection of connected firms as recognized by Statistics Canada in the compilation of ICO. Refer to the accompanying online appendix. Note that we use this term differently from its general dictionary meaning by referring to all the different types of structures in our data as enterprises	ICO by Statistics Canada definition
Family controlled group/firm	A dummy variable equal to 1 if an enterprise/firm has an ultimate owner reported as a single entity describing an individual, a family, a group of related individuals or a group of related families	ICO by Statistics Canada, Bureau van Dijk, EDGAR, SEDAR, Times and Forbes Billionaire lists
Pyramid	A dummy variable equal to 1 if an entity belongs to a group of connected legal entities, where if the ultimate owner is an individual, there must be at least one corporation that is a parent to a non-wholly-owned subsidiary. Where the ultimate owner is non-family controlled, then it must be a parent to a non-wholly-owned subsidiary. Often, researchers require that at least one of the entities in the group be publicly listed. We do not make this requirement because our data covers privately held entities and their ownership structure represents interest in itself	Author definition consistent with Almeida and Wolfenzon (2006) and Khanna and Yafeh (2007)

Variable/concept	Definition	Source
Apex	A dummy variable equal to 1 if an entity is the apex of an enterprise. An apex is the entity at the top of a corporate ownership structure, which exercises effective control over all other entities therein. The types of apexes we encounter are either families/individuals or companies for which control cannot be assigned to a single family/individual	ICO
Group	A <i>group</i> represents connected legal entities under common control. Groups are either flat structures, where the controlling owner holds all subsidiaries directly or could be layered, but all subsidiaries are wholly owned	Author definition consistent with Almeida and Wolfenzon (2006) and Khanna and Yafeh (2007)
Stand-alone firm	A firm that is not a subsidiary or is directly held by its controlling owner in case of an individual or family	Author definition
Number of firms	The number of firms in a business group	ICO
Control jurisdiction	A categorical variable taking the value 0 if the control jurisdiction of an enterprise is Canada, and 1 if it is an Anglo-Saxon country	ICO, Bureau van Dijk
Productive	A dummy variable equal to 1 if an entity belongs to an industry that is not banking, finance or insurance (we do include utilities as a productive sector). It equals 0 for financial firms in banking, finance or insurance. Importantly, many of the financial firms are classified as holding companies (part of the financial group in the NAICS classification) without any productive activity. Non-profits are also included in the financial category as often corporate structures contain legal entities with a charity status that in practice serves as a holding company. We exclude entities classified as government bodies from the data and only retain them if they are the ultimate owner of a structure for classification purposes	ICO and Thompson Financial

Variable/concept	Definition	Source
Anglo-Saxon country	Anglo-Saxon jurisdictions are Australia, Ireland, New Zealand, UK and US (Esping-Andersen 1990). Canada is also considered an Anglo-Saxon country but here it represents a separate sample partly due to its hybrid common-civil law system, partly because it covers the population of Canadian firms as opposed to a subset of large global multinationals. In the case of board independence data, we treat the US as a separate subgroup for similar reasons of data coverage in BoardEx	
Ultimate cash flow rights	Ultimate cash flow rights ( $ucfr_i$ ) equals the proportion of one unit of disbursement from firm $i$ that is received by the apex. $ucfr_i$ is the $i$ -th entry of the vector: $ucfr = f' (I_n - A)^{-1}$ , where $I_n$ is the identity matrix with dimensions $n \times n$ . We treat the $ucfr$ of the apex in itself as 1. See Almeida et al. (2011)	ICO by Statistics Canada and author calculations
Control rights	Control rights ( $cr_i$ ) is defined in Almeida et al. (2011) as the critical control threshold which is shown to be equivalent to the concept of the weakest link (as used in La Porta et al. 1999, Claessens et al. 2000, and Faccio and Lang 2002) when cross-shareholdings and multiple links are absent but can also be computed for more complex structures	“
Wedge	The ratio of ultimate cash flow rights and control rights ( $ucfr_i / cr_i$ )	“
Board independence	The ratio of independent directors out of the total number of directors	BoardEx

## Appendix 2 Summary statistics of variables used in replications

Below are the summary statistics for all variables used in the replications reported in Table 2. All variable definitions are as in the original papers. In the case of Lin et al. (2013), we use ownership data compiled by us from ICO and merged with financials from Capital IQ and Thomson Reuters, resulting in 2044 firms. In the case of Coles et al. (2008), who analyze US listed firms, we collect board variables from BoardEx (covering 1999–2019) and merge with financials from Compustat, resulting in 2507 firms in the replication.

Panel A Lin et al. (2013)

Variable names	Mean	Std. Dev
Debt choice	1.122	1.272
Wedge	0.948	0.178
Z-score	3.140	1.727
Cash-flow rights	0.855	0.276
Leverage	0.477	0.215
Tangibility	0.338	0.253
Log assets	7.956	3.537
Profitability	0.104	0.256
Q	1.193	0.805

## Panel B Coles et al. (2008)

Variable names	Mean	Std. Dev
Ln (Board size)	2.130	0.309
Insider fraction	0.186	0.105
Insider ownership	0.009	0.047
R&D dummy	0.587	0.492
Advice dummy	0.658	0.474
Q	2.917	0.491
St Dev returns	0.132	0.094
Profitability	0.393	0.266
Intangible assets	0.768	0.242

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11156-022-01074-8>.

**Acknowledgements** We would like to thank Heitor Almeida, Seth Armitage, Vlado Atanasov, Leonidas Barbopoulos, Mike Burkart, Giovanni Cerulli, Jo Danbolt, Simon Doering, Espen Eckbo, Alex Edmans, Mara Faccio, Matti Keloharju, Jason Leonard and Jack Lothian at Statistics Canada, Brian Main, Bill Megginson, Peter Moles, Randall Morck, Talis Putnins, Sergei Sarkassian, Ben Sila, Andy Snell, John Wald, Pauline Weetman, participants at the FMA European Conference in Kristiansand, participants at the International Workshop on Computational Economics and Econometrics: The Socio-Economics of Networks in Rome, seminar participants at the University of Exeter, University of St. Andrews, and Heriot-Watt University.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abdallah W, Goergen M, O'Sullivan N (2015) Endogeneity: how failure to correct for it can cause wrong inferences and some remedies: endogeneity. *Br J Manag* 26(4):791–804. <https://doi.org/10.1111/1467-8551.12113>
- Adams RB (2017) The ABCs of empirical corporate (governance) research. *Corp Gov Int Rev* 25(6):461–464. <https://doi.org/10.1111/corg.12229>
- Aghion P, Van Reenen J, Zingales L (2013) Innovation and institutional ownership. *Am Econ Rev* 103(1):277–304. <https://doi.org/10.1257/aer.103.1.277>
- Aguilera RV, Crespi-Cladera R (2016) 'Global corporate governance: on the relevance of firms' ownership structure. *J World Bus* 51(1):50–57. <https://doi.org/10.1016/j.jwb.2015.10.003>
- Aguilera RV, Desender KA, Lamy MLP, Lee JH (2017) The Governance impact of a changing investor landscape. *J Int Bus Stud* 48(2):195–221. <https://doi.org/10.1057/s41267-016-0043-y>
- Almeida HV, Wolfenzon D (2006) A theory of pyramidal ownership and family business groups. *J Finance* 61(6):2637–2680. <https://doi.org/10.1111/j.1540-6261.2006.01001.x>
- Almeida H, Park SY, Subrahmanyam MG, Wolfenzon D (2011) The structure and formation of business groups: evidence from Korean chaebols. *J Financ Econ* 99(2):447–475. <https://doi.org/10.1016/j.jfineco.2010.08.017>
- Anderson RC, Reeb DM (2003) Founding-family ownership and firm performance: evidence from the S&P 500. *J Finance* 58(3):1301–1327. <https://doi.org/10.1111/1540-6261.00567>
- Atanasov V, Black B (2016) Shock-based causal inference in corporate finance and accounting research. *Crit Finance Rev* 5(2):207–304. <https://doi.org/10.1561/104.000000036>
- Atanasov V, Black B (2020) The trouble with instruments: the need for pretreatment balance in shock-based instrumental variable designs. *Manag Sci* 67(2):1270–1302. <https://doi.org/10.1287/mnsc.2019.3510>
- Azar J, Schmalz MC, Tecu I (2018) Anticompetitive effects of common ownership. *J Finance* 73(4):1513–1565. <https://doi.org/10.1111/jofi.12698>
- Banalieva ER, Eddleston KA (2011) Home-region focus and performance of family firms: the role of family vs non-family leaders. *J Int Bus Stud* 42(8):1060–1072. <https://doi.org/10.1057/jibs.2011.28>
- Banerjee S, Homroy S (2018) Managerial incentives and strategic choices of firms with different ownership structures. *J Corp Finance* 48:314–330. <https://doi.org/10.1016/j.jcorpfin.2017.10.001>
- Basu N, Dimitrova L, Paeglis I (2009) Family control and dilution in mergers. *J Bank Finance* 33(5):829–841. <https://doi.org/10.1016/j.jbankfin.2008.09.017>
- Basu N, Paeglis I, Toffanin M (2017) Reading between the Blocks. *J Corp Finance* 45:294–317. <https://doi.org/10.1016/j.jcorpfin.2017.04.017>
- Bazzi M, Blasques F, Koopman SJ, Lucas A (2017) Time-varying transition probabilities for Markov regime switching models. *J Time Ser Anal* 38(3):458–478. <https://doi.org/10.1111/jtsa.12211>
- Belenzon S, Chatterji AK, Daley B (2019) Choosing between growth and glory. *Manag Sci*. <https://doi.org/10.1287/mnsc.2019.3296>
- Bennedsen M, Nielsen KM (2010) Incentive and entrenchment effects in European ownership. *J Bank Finance* 34(9):2212–2229. <https://doi.org/10.1016/j.jbankfin.2010.02.007>
- Bertrand M, Schoar A (2006) The role of family in family firms. *J Econ Perspect* 20(2):73–96. <https://doi.org/10.1257/jep.20.2.73>
- Bertrand M, Mehta P, Mullainathan S (2002) Ferreting out tunneling: an application to Indian business groups. *Q J Econ* 117(1):121–148. <https://doi.org/10.1162/003355302753399463>
- Bhaumik SK, Driffield N, Pal S (2010) Does ownership structure of emerging-market firms affect their outward FDI? the case of the indian automotive and pharmaceutical sectors. *J Int Bus Stud* 41(3):437–450. <https://doi.org/10.1057/jibs.2009.52>
- Black B, Kim W (2012) The effect of board structure on firm value: a multiple identification strategies approach using Korean data. *J Financ Econ* 104(1):203–226. <https://doi.org/10.1016/j.jfineco.2011.08.001>
- Black B, Gledson A, de Carvalho V, Khanna WK, Yurtoglu B (2014) Methods for multicountry studies of corporate governance: evidence from the BRIKT countries. *J Econom* 183(2):230–240. <https://doi.org/10.1016/j.jeconom.2014.05.013>
- Black B, Gledson A, de Carvalho V, Khanna WK, Yurtoglu B (2017) Corporate governance indices and construct validity. *Corp Gov Int Rev* 25(6):397–410. <https://doi.org/10.1111/corg.12215>
- Bliese PD, Schepker DJ, Essman SM, Ployhart RE (2019) Bridging methodological divides between macro- and microresearch: endogeneity and methods for panel data. *J Manag*. <https://doi.org/10.1177/0149206319868016>

- Boateng A, Bi XG, Brahma S (2017) The impact of firm ownership, board monitoring on operating performance of Chinese mergers and acquisitions. *Rev Quant Finance Account* 49(4):925–948. <https://doi.org/10.1007/s11156-016-0612-y>
- Boone AL, Field LC, Karpoff JM, Raheja CG (2007) The determinants of corporate board size and composition: an empirical analysis. *J Financ Econ* 85(1):66–101. <https://doi.org/10.1016/j.jfineco.2006.05.004>
- Carney M, Gedajlovic E, Strike VM (2014) Dead money: inheritance law and the longevity of family firms. *Entrep Theory Pract*. <https://doi.org/10.1111/etap.12123>
- Choi JJ, Park SW, Yoo SS (2007) The value of outside directors: evidence from corporate governance reform in Korea. *J Financ Quant Anal* 42(4):941–962. <https://doi.org/10.1017/S0022109000003458>
- Claessens S, Djankov S, Lang LHP (2000) The separation of ownership and control in East Asian corporations. *J Financ Econ* 58(1–2):81–112. [https://doi.org/10.1016/S0304-405X\(00\)00067-2](https://doi.org/10.1016/S0304-405X(00)00067-2)
- Claessens S, Djankov S, Fan JPH, Lang LHP (2002) Disentangling the incentive and entrenchment effects of large shareholdings. *J Finance* 57(6):2741–2771. <https://doi.org/10.1111/1540-6261.00511>
- Coles JL, Daniel ND, Naveen L (2008) Boards: does one size fit all? *J Financ Econ* 87(2):329–356. <https://doi.org/10.1016/j.jfineco.2006.08.008>
- DeAngelo H, Roll R (2015) How stable are corporate capital structures? *J Finance* 70(1):373–418. <https://doi.org/10.1111/jofi.12163>
- Deephouse DL, Jaskiewicz P (2013) Do family firms have better reputations than non-family firms? An integration of socioemotional wealth and social identity theories: corporate reputation in family firms. *J Manag Stud* 50(3):337–360. <https://doi.org/10.1111/joms.12015>
- Dlugosz J, Fahlenbrach R, Gompers P, Metrick A (2006) Large blocks of stock: prevalence, size, and measurement. *Corp Gov* 12(3):594–618. <https://doi.org/10.1016/j.jcorpfin.2005.04.002>
- Donelli M, Larrain B, Francisco Urzúa I (2013) Ownership dynamics with large shareholders: an empirical analysis. *J Financ Quant Anal* 48(2):579–609. <https://doi.org/10.1017/S0022109013000173>
- Eccles RG, Klimenko S (2019) The investor revolution. *Harvard business review*, no. May–June. <https://hbr.org/2019/05/the-investor-revolution>
- Edmans A, Holderness CG (2017) Blockholders: a survey of theory and evidence. In: *The handbook of the economics of corporate governance*, vol 1. Elsevier, pp 541–636. <https://doi.org/10.1016/bs.hecg.2017.11.002>
- Ellis J, Guo L, Mobbs S (2020) How does forced-CEO-turnover experience affect directors? *J Financ Quant Anal* 56(4):1163–1191. <https://doi.org/10.1017/S0022109020000307>
- Esping-Andersen G (1990) *The three worlds of welfare capitalism*. Princeton University Press, Princeton
- Fabisik K, Fahlenbrach R, Stulz RM, Taillard JP (2021) Why are firms with more managerial ownership worth less? *J Financ Econ* 140(3):699–725. <https://doi.org/10.1016/j.jfineco.2021.02.008>
- Faccio M, Lang LHP (2002) The ultimate ownership of western European corporations. *J Financ Econ* 65(3):365–395. [https://doi.org/10.1016/S0304-405X\(02\)00146-0](https://doi.org/10.1016/S0304-405X(02)00146-0)
- Faccio M, Lang LHP, Young L (2010) Pyramiding vs leverage in corporate groups: international evidence. *J Int Bus Stud* 41(1):88–104. <https://doi.org/10.1057/jibs.2009.33>
- Faccio M, Marchica M-T, Mura R (2011) Large shareholder diversification and corporate risk-taking. *Rev Financ Stud* 24(11):3601–3641. <https://doi.org/10.1093/rfs/hhr065>
- Fahlenbrach R, Stulz RM (2009) Managerial ownership dynamics and firm value. *J Financ Econ* 92(3):342–361. <https://doi.org/10.1016/j.jfineco.2008.06.005>
- Flammer C, Bansal P (2017) Does a long-term orientation create value? Evidence from a regression discontinuity. *Strateg Manag J* 38(9):1827–1847. <https://doi.org/10.1002/smj.2629>
- Franks J, Mayer C, Volpin P, Wagner HF (2012) The life cycle of family ownership: international evidence. *Rev Financ Stud* 25(6):1675–1712. <https://doi.org/10.1093/rfs/hhr135>
- Frye MB, Pham DT, Zhang R (2021) Board monitoring and advising trade-offs amidst economic policy uncertainty. *Financ Rev*. <https://doi.org/10.1111/fire.12286>
- Graham JR, Kim H, Leary M (2020) CEO-board dynamics. *J Financ Econ* 137(3):612–636. <https://doi.org/10.1016/j.jfineco.2020.04.007>
- Guo L, Masulis RW (2015) Board structure and monitoring: new evidence from CEO turnovers. *Rev Financ Stud* 28(10):2770–2811. <https://doi.org/10.1093/rfs/hhv038>
- Hermalin BE, Weisbach MS (2003) Boards of directors as an endogenously determined institution: a survey of the economic literature. *FRBNY Economic Policy Review*
- Himmelberg CP, Glenn Hubbard R, Palia D (1999) Understanding the determinants of managerial ownership and the link between ownership and performance. *J Financ Econ* 53(3):353–384. [https://doi.org/10.1016/S0304-405X\(99\)00025-2](https://doi.org/10.1016/S0304-405X(99)00025-2)
- Holderness CG (2009) The myth of diffuse ownership in the United States. *Rev Financ Stud* 22(4):1377–1408. <https://doi.org/10.1093/rfs/hhm069>

- Holmes RM, Hoskisson RE, Kim H, Wan WP, Holcomb TR (2018) International strategy and business groups: a review and future research Agenda. *J World Bus* 53(2):134–150. <https://doi.org/10.1016/j.jwb.2016.11.003>
- Khanna T, Yafeh Y (2007) Business groups in emerging markets: paragons or parasites? *J Econ Lit* 45(2):331–372
- Kim EH, Lu Y (2011) CEO ownership, external governance, and risk-taking. *J Financ Econ* 102(2):272–292. <https://doi.org/10.1016/j.jfineco.2011.07.002>
- Kim EH, Lu Y (2013) Corporate governance reforms around the world and cross-border acquisitions. *J Corp Finance* 22:236–253. <https://doi.org/10.1016/j.jcorpfin.2013.05.005>
- Kim D, Starks LT (2015) Board heterogeneity of expertise and firm performance. In: 27th annual conference on financial economics and accounting paper
- La Porta R, Lopez-De-Silanes F, Shleifer A (1999) Corporate ownership around the World. *J Finance* 54(2):471–517. <https://doi.org/10.1111/0022-1082.00115>
- Lardon A, Beuselinck C, Deloof M (2019) Does stable ownership create value? Evidence from the global financial crisis. *Rev Quant Finance Account* 52(2):573–642. <https://doi.org/10.1007/s11156-018-0719-4>
- Larrain B, Francisco UI (2013) Controlling shareholders and market timing in share issuance. *J Financ Econ* 109(3):661–681. <https://doi.org/10.1016/j.jfineco.2013.03.013>
- Lemmon ML, Roberts MR, Zender JF (2008) Back to the beginning: persistence and the cross-section of corporate capital structure. *J Finance* 63(4):1575–1608. <https://doi.org/10.1111/j.1540-6261.2008.01369.x>
- Lin C, Ma Y, Malatesta P, Xuan Y (2011) Ownership structure and the cost of corporate borrowing. *J Financ Econ* 100(1):1–23. <https://doi.org/10.1016/j.jfineco.2010.10.012>
- Lin C, Ma Y, Malatesta P, Xuan Y (2012) Corporate ownership structure and bank loan syndicate structure. *J Financ Econ* 104(1):1–22. <https://doi.org/10.1016/j.jfineco.2011.10.006>
- Lin C, Ma Y, Malatesta P, Xuan Y (2013) Corporate ownership structure and the choice between bank debt and public debt. *J Financ Econ* 109(2):517–534. <https://doi.org/10.1016/j.jfineco.2013.03.006>
- Linck J, Netter J, Yang T (2008) The determinants of board structure. *J Financ Econ* 87(2):308–328. <https://doi.org/10.1016/j.jfineco.2007.03.004>
- Lins KV (2003) Equity ownership and firm value in emerging markets. *J Financ Quant Anal* 38(1):159. <https://doi.org/10.2307/4126768>
- Liu Yu, Miletkov MK, Wei Z, Yang T (2015) Board independence and firm performance in China. *J Corp Finance* 30:223–244. <https://doi.org/10.1016/j.jcorpfin.2014.12.004>
- Masulis RW, Pham PK, Zein J (2011) Family business groups around the world: financing advantages, control motivations, and organizational choices. *Rev Financ Stud* 24(11):3556–3600. <https://doi.org/10.1093/rfs/hhr052>
- Mooney A (2018) ESG wake-up call for pension laggards. *Financial times*, 15 October 2018. <https://on.ft.com/2SBrzje>
- Morck R (2005) How to eliminate pyramidal business groups the double taxation of inter-corporate dividends and other incisive uses of tax policy. In: *Tax policy and the economy*, vol 19. MIT Press, pp 135–179. <http://www.nber.org/chapters/c0167>
- Mundlak Y (1978) On the pooling of time series and cross section data. *Econom J Econom Soc*. <https://doi.org/10.2307/1913646>
- Murphy KR, Myers B, Wolach AH (2014) *Statistical power analysis: a simple and general model for traditional and modern hypothesis tests*, 4th edn. Routledge, New York
- Nicodano G, Regis L (2019) A trade-off theory of ownership and capital structure. *J Financ Econ* 131(3):715–735. <https://doi.org/10.1016/j.jfineco.2018.09.001>
- OECD (2018) Professionalising boards of directors of state-owned enterprises: stocktaking of national practices
- Papadopoulos K (2019) Corporate governance in emerging markets. *ISS*
- Pronobis P, Schaeuble J (2020) Foreign ownership and audit fees in European listed firms. *Eur Account Rev*. <https://doi.org/10.1080/09638180.2020.1830819>
- Rabbiosi L, Gregorič A, Stucchi T (2019) ‘Diaspora ownership and homeland firms’ internationalization’. *J Int Manag*. <https://doi.org/10.1016/j.intman.2018.12.002>
- Reeb D, Sakakibara M, Mahmood IP (2012) From the editors: endogeneity in international business research. *J Int Bus Stud* 43(3):211–218. <https://doi.org/10.1057/jibs.2011.60>
- Semadeni M, Withers MC, Certo TS (2014) The perils of endogeneity and instrumental variables in strategy research: understanding through simulations: research notes and commentaries. *Strateg Manag J* 35(7):1070–1079. <https://doi.org/10.1002/smj.2136>

- Siegel J, Choudhury P (2012) A reexamination of tunneling and business groups: new data and new methods. *Rev Financ Stud* 25(6):1763–1798. <https://doi.org/10.1093/rfs/hhs008>
- Singla C, George R, Veliyath R (2017) Ownership structure and internationalization of Indian firms. *J Bus Res* 81:130–143. <https://doi.org/10.1016/j.jbusres.2017.08.016>
- Tsoutsoura M (2015) The effect of succession taxes on family firm investment: evidence from a natural experiment. *J Financ* 70(2):649–688. <https://doi.org/10.1111/jofi.12224>
- Villalonga B, Amit R (2009) How are U.S. family firms controlled? *Rev Financ Stud* 22(8):3047–3091. <https://doi.org/10.1093/rfs/hhn080>
- Wintoki MB, Linck JS, Netter JM (2012) Endogeneity and the dynamics of internal corporate governance. *J Financ Econ* 105(3):581–606. <https://doi.org/10.1016/j.jfineco.2012.03.005>
- Wooldridge JM (2010) *Econometric analysis of cross section and panel data*, 2nd edn. MIT Press, Cambridge
- Wooldridge JM (2019) Correlated random effects models with unbalanced panels. *J Econom* 211(1):137–150. <https://doi.org/10.1016/j.jeconom.2018.12.010>
- Yermack D (2017) Corporate governance and blockchains. *Rev Finance* 21(1):7–31. <https://doi.org/10.1093/rof/rfw074>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.