Boiling Down Boilerplate in M&A Agreements

Robert Anderson[*] *and Jeffrey Manns*[**]

Abstract

One of the most striking features of merger and acquisition agreements is the juxtaposition of numerous boilerplate provisions with a high level of "editorial churning," ad hoc edits that appear to be cosmetic rather than substantive. In our recent article, *The Inefficient Evolution of Merger Agreements*, 85 GEORGE WASH. L. REV. 57 (2017), we show empirically the high degree of "speciation" among merger agreements which hinders standardization both within law firms and across firms. We use a computer program to identify the "DNA fingerprints" of each precursor agreement that serves as a drafting template for over 12,000 public company merger agreements in a twenty-year data set. We document how on average over half of the text of merger and agreements are rewritten in the drafting process each time that a precedent is adapted into a new acquisition agreement, even though the substantive provisions of merger agreements have similar features.

In *Boiling Down Boilerplate* we focus on how editorial churning in acquisition agreements has shaped the evolution of boilerplate provisions and eroded both the standardization and meaning of these provisions. Scholars in other contractual contexts such as sovereign debt agreements have shown how the repetition of boilerplate text in the drafting process has produced pathologies known as the "black hole problem" and the "grey hole problem" as changes have led repeated provisions to lose all or part of their meaning. The problem arises from repetition that involves "rote usage" of boilerplate terms without critical examination of the terms, especially when that rote usage is combined with the "encrustation" of variations in the otherwise boilerplate terms.

[*] Associate Professor of Law, Pepperdine University School of Law.

[**] Associate Professor of Law, The George Washington University Law School. This paper is an early-stage draft which comes with all of the necessary caveats. We welcome any comments you may have, and please e-mail us at robert.anderson@pepperdine.edu and jmanns@law.gwu.edu.

We extend this research by examining the evolution of boilerplate provisions in the merger agreement context. We leverage our existing data set of over 12,000 public company merger agreements from 1994 to 2014 to create a comprehensive picture of changes in key boilerplate provisions over time. We do this by using a computer program to compute the "edit distance" or "Levenshtein distance" between boilerplate provisions in merger agreements. This approach allows us to measure the degree of textual similarity or dissimilarity based on the number of insertions and deletions (i.e., edits) in boilerplate provisions across agreements. We complement this quantitative approach with qualitative assessments of both the extent and consequences of boilerplate provisions diverging across merger agreements.

We show how it is possible to identify the paragraphs of acquisition agreements which serve as boilerplate and demonstrate both the degree and type of textual "drift" of these provisions over multiple generations. We construct "family trees" for boilerplate provisions by tracing the descendants of each "ancestor" provision. We demonstrate that common ancestors have progeny extending out in multiple directions which become increasingly dissimilar to each other over a few generations of acquisition agreements.

We show that incremental changes in boilerplate from generation to generation lead to rapid "speciation" of the terms. Small additions and deletions from boilerplate text lead to significant cumulative effects over multiple generations. We demonstrate that this textual "drift" takes place both within boilerplate that falls within a given chain of precedent, but also even more broadly for boilerplate provisions that have a common ancestor precedent, but evolve separately along different lineages of precedents. We also show spatially that the pattern of boilerplate "speciation" we discovered in entire merger agreements is replicated in individual boilerplate terms.

Our findings suggest that three inter-connected problems exist in the acquisition agreement context, which reinforce the black hole or grey hole hypothesis of the evolution of boilerplate having a potentially corrosive effect on the textual integrity and meaning of these terms. The core problem appears to be "rote usage." Lawyers recycle boilerplate without giving much, if any, thought to the meaning of this language. This lack of apparent reflection about the purpose of boilerplate provisions is coupled with the "encrustation"

and "abrasion" of words in boilerplate provisions, i.e. the idiosyncratic addition or removal of words in the drafting process. These idiosyncratic edits are then preserved in later generations of the provision, perpetuating drift away from the ancestor standard.

Small changes in legal language from one M&A deal to the next lead to significant distortions in boilerplate provisions over time, which reflects a broader problem of acquisition agreement "speciation" as the text of merger agreements becomes increasingly unrecognizable from the chain of precedent agreements. The process of ad hoc additions and deletions from boilerplate terms undermines the potential benefits of this (partial) standardization and may erode boilerplate language of its meaning.

Table of Contents

INTRODUCTION

"Boilerplate" provisions consist of standardized terms whose meaning is intended to be consistent from one transaction to the next or even from one transactional context to another. Boilerplate terms are ubiquitous in contracts and related transactional documents, yet some prominent scholars have highlighted how the nature of the legal drafting process may potentially undercut the standardization purpose, the wording, and the substantive meaning of these provisions.[1]

The appeal of boilerplate terms is that they offer the promise of partial standardization of otherwise non-standardized documents. In theory boilerplate terms heighten legal certainty and the universality of provisions among lawyers by providing uniform language whose meaning has stood the test of time (and often scrutiny by courts).[2] The use of boilerplate also offers the promise of drafting efficiency and reduces the likelihood of error if the provisions are consistently copied from precedent to subsequent agreement.

The problem lies in the drafting process of documents that are neither completely negotiated nor completely standardized. In such contexts, parties often recycle boilerplate provisions from precedents, a practice consistent one of the goals of standardization and the use of precedent. However, the penchant of lawyers to leave their mark on agreements through a myriad of additions and deletions to the text may come at a high price of eroding the substance of ostensibly standardized language. The interplay between mechanical recitation of boilerplate and high levels of edits is that boilerplate terms may morph into non-standardized language that loses its original

[1] *See* Stephen J. Choi, Mitu G. Gulati, & Robert E. Scott, *The Black Hole Problem in Commercial Boilerplate*, at 2-4, *available at* https://ssrn.com/abstract=2835681 (discussing the black hole problem in the context of the *pari passu* clause, a boilerplate provision in sovereign debt contracts); *see also* Christopher J. French, *The Illusion of Insurance Contracts*, 89 TEMPLE L. REV. (forthcoming 2016) (discussing the difficulties of determining the intent of drafters of standard form language in insurance contracts).

[2] *See* Marcel Kahan & Michael Klausner, *Standardization & Innovation in Corporate Contracting*, 83 VA. L. REV. 713, 719-20 (1997) (discussing the potential "learning benefits' of commonly used terms); Michael Klausner, *Standardization & Innovation in Corporate Contracting*, 81 VA. L. REV. 757, 783-84 (1995) (discussing the network benefits from familiarity with boilerplate terms).

meaning, even though the provisions continue to live on in agreements in increasingly fractured ways. These findings provide evidence substantiating the black hole or grey hole hypothesis of the evolution of boilerplate having a potentially corrosive effect on the textual integrity and meaning of these terms.[3]

The mergers and acquisitions context epitomizes this problem of unreflective copying of precedent provisions combines with ad hoc edits to individual clauses. Each agreement is based on a prior deal precedent, and drafters frequently incorporate sections of the prior deal without sufficient scrutiny about the degree to which idiosyncratic novelties have been introduced in the precedent document that may be inapplicable to the new deal. At the same time, high levels of "editorial churning" take place in the process of transforming each precedent into the current acquisition agreement. The result is a problem of "drafting drift." Boilerplate provisions live on from deal to deal, yet gradually shed their textual integrity and potentially lose their clear meaning as ad hoc edits are copied from deal to deal and new ad hoc edits are added at each stage.

We show how it is possible to identify the paragraphs of acquisition agreements which serve as boilerplate and demonstrate both the degree and type of textual "drift" of these provisions over multiple generations. We construct "family trees" for boilerplate provisions by tracing the decedents of each provisions backwards in a linear way to each prior precedent. Then we reverse the process to show how "ancestor" provisions have progeny extending out in multiple directions which become increasingly dissimilar to each other over a few generations of acquisition agreements.

Our study shows that incremental changes in boilerplate from one generation to the next lead to rapid "speciation" of the terms. Small additions and deletions from boilerplate text lead to significant cumulative effects over multiple generations. We demonstrate that this textual "drift" takes place both within boilerplate that falls within a given chain of precedent, but also even more broadly for boilerplate provisions that have a common ancestor precedent, but evolve separately along different lineages of precedents. Like the Big Bang, the heterogeneity of boilerplate text appears to increase in numerous

[3] *See* Choi, Gulati, & Scott, *infra* note 1, at 4 (discussing the need for empirical research on the extent of rote usage and encrustation in boilerplate provisions).

directions, which supports an "expanding universe" theory for boilerplate that undermines the textual integrity and the meaning of boilerplate terms. We also show spatially that the pattern of boilerplate "speciation" underscores the high impact of editorial churning in undercutting standardization of boilerplate.

Our findings suggest that three inter-connected problems exist in the acquisition agreement context. The core problem appears to be "rote usage." Lawyers recycle boilerplate without giving much, if any, thought to the meaning of this language. This lack of apparent reflection about the purpose of boilerplate coupled with the "encrustation" and "abrasion" of words in boilerplate provisions, i.e. the idiosyncratic addition or removal of words in the drafting process, leads to significant divergences in the substance and meaning of boilerplate provisions over even a short number of generations. Lawyers appear to subsume boilerplate into the larger process of editorial churning that leads to rewriting of text throughout the acquisition agreement. The random variations that additions and deletions of boilerplate text introduce in the drafting process appear to be even more severe in the merger agreement context than in other contractual settings, leading to rapid drift away from the original boilerplate.

Part I will place this boilerplate study in the larger context of our research on the evolution of acquisition agreements and discuss the role of boilerplate provisions in precedent-based drafting. Part II will lay out our data and methodology as well as delineate the distinctive challenges of identifying boilerplate in non-standardized documents. Part III will lay out empirical evidence substantiating the high degree of textual drift within both lineages of boilerplate and the even more extensive drift between the divergent branches of boilerplate with a common precedent ancestor. Part IV will discuss some of the implications and shortcomings of this study that may merit further attention in future work.

I. BACKGROUND AND THEORY

A. Situating the Boilerplate Study in Our Larger Project

This Article builds on our larger project of systematically examining the evolution of public company merger agreements.[4] The challenge in assessing the drafting of acquisition agreements or individual provisions is that the public can see the end product, but is not privy to the process that creates the acquisition agreement. The Securities & Exchange Commission mandates disclosure of public company acquisition agreements, which provide a window to the end product of lawyering that the public often does not get to see in other areas of transactional law.[5] But we do not get to witness the process that leads to the formation of acquisition agreements.

For this reason it is difficult to assess directly the efficiency and evolution of the acquisition agreement drafting process or of particular provisions.[6] But the economic and legal stakes of mergers and acquisitions are so significant that it is important to gain a better understanding of what is taking place (or not taking place) in the legal drafting.[7] Our broad project is to reverse engineer the drafting process to identify potential inefficiencies and textual distortions by analyzing the evolution of public company acquisition agreements and particular provisions.[8]

---

4 *See* Robert Anderson & Jeffrey Manns, *The Inefficient Evolution of Merger Agreements*, 85 GEORGE WASH. L. REV. 57, 59-62 (2017).

5 *See* SEC, FORM 8-K, ITEM 1.01, at 4, https://www.sec.gov/about/forms/form8-k.pdf (requiring companies to disclose material definitive agreements outside of the ordinary course of business including merger agreements).

6 Ronald Gilson's seminal article highlighted the value added from transactional lawyers, while our study examines how efficient lawyers are in adding that value. *See* Ronald J. Gilson, *Value Creation by Business Lawyers: Legal Skills and Asset Pricing*, 94 YALE L.J. 239, 243, 254–55 (1984) [hereinafter Gilson, *Value Creation*] (observing that "the academic literature assume[s] that business lawyers *increase* the value of a transaction" and arguing that M&A lawyers add value by designing provisions in acquisition agreements that reduce transaction costs and increase mutual gain); *see also* Steven L. Schwarcz, *Explaining the Value of Transactional Lawyering*, 12 STAN. J.L. BUS. & FIN. 486, 487–88, 506–07 (2007) (using survey data from transactional lawyers and their clients to argue that lawyers add value to transactions primarily by reducing regulatory costs through legal expertise, rather than more broadly reducing transaction costs or adding reputational value).

7 Acquisition agreements delineate the rights and duties of parties for trillions of dollars of transactions each year. *See* Maureen Farrell, *2015 Becomes the Biggest M&A Year Ever*, WALL ST. J., http://www.wsj.com/articles/2015-becomes-the-biggest-m-a-year-ever-1449187101 (discussing how global mergers and acquisitions surpassed $4.3 trillion in 2015).

8 Other notable empirical works also examine changes in contractual provisions in other transactional contexts. *See, e.g.*, Stephen J. Choi & Mitu Gulati, *Innovation in*

Because of the SEC's disclosure rules, we can see the universe of the past twenty years of public company merger agreements (in our study from 1994 to 2014 which covers over 12,000 agreements).[9] Acquisition agreements are so complex and the legal stakes so high, that nearly every public company merger agreement is based on an earlier acquisition agreement that serves as its precedent.[10] Using computer textual analysis, we show how it is possible to identify the precedent which serves as the template for the drafting of each deal.[11] We leverage computer technology to lift the veil on the drafting process by showing how agreements are created and how both documents as a whole and individual provisions change in incremental ways over time.[12]

Boilerplate Contracts: An Empirical Examination of Sovereign Bonds, 53 EMORY L.J. 930, 932–34 (2004) (conducting empirical analysis of sovereign bond offerings to show that boilerplate provisions changed in response to significant shifts in the interpretation of key provisions, but only after an industry-wide delay which reflected the relocated of lawyers to change boilerplate provisions); MITU GULATI & ROBERT E. SCOTT, THE 3½ MINUTE TRANSACTION: BOILERPLATE AND THE LIMITS OF CONTRACT DESIGN 3–10 (2013) (using empirical data to show that once a boilerplate provision is in place it often becomes part of a transactional checklist regardless of its actual value-added); Jonathan C. Lipson, *Price, Path, & Pride: Third Party Closing Opinion Practice Among U.S. Lawyers (A Preliminary Investigation)*, 3 BERKELEY BUS. L.J. 59, 113–14 (2005) (using qualitative interviews to assess the logic behind lawyers' drafting third-party closing opinions).

9 *See* SEC, FORM 8-K, ITEM 1.01, at 4, https://www.sec.gov/about/forms/form8-k.pdf.

10 *See* SCOTT J. BURNHAM, DRAFTING AND ANALYZING CONTRACTS: A GUIDE TO THE PRACTICAL APPLICATION OF THE PRINCIPLES OF CONTRACT LAW 5–6 (3d ed. 2003) (discussing how attorneys "rarely start to draft on a blank slate. . . . [and generally] start with an existing contract or form").

11 *See* TINA L. STARK, DRAFTING CONTRACTS: HOW AND WHY LAWYERS DO WHAT THEY DO 335–36 (2007) (discussing the benefits of heightened efficiency and legal certainty from precedent-based legal drafting).

12 Two other notable empirical studies provide similar prisms for understanding the M&A drafting process, yet reach different conclusions that will be at the heart of our M&A panel's discussion at this conference. Professor Coates documents the growth in length of merger agreements over the past twenty years, which the author attributes to changes in legal risks and deal and financing markets, as well as the increase in "linguistic complexity' of these documents. *See generally* John C. Coates, *Why Have M&A Contracts Grown? Evidence From Twenty Years of Deals*, ECGI Working Paper No. 333/2016 (Nov. 2016). Professor Jennejohn argues that the complexity of M&A exposes acquisition agreements to multiple sources of path dependency which undercuts efforts at standardization. *See generally* Matthew Jennejohn, *Assymetric Standardization in M&A Agreements*, Mar. 25, 2017 (manuscript on file with authors). Other studies have examined the development of particular acquisition agreement provisions. *See, e.g.*, Afra Afsharipour, *Transforming the Allocation of Deal Risk Through Reverse Termination Fees*, 63 VAND. L. REV. 1161 (2010) (discussing attempts at reallocating deal risks through reverse termination fees that compensate target companies should the buyer walk away, and assessing the impact such attempts have on acquisition agreement drafting); William T. Allen, *Understanding Fiduciary Outs: The What*

By identifying the precedent for each deal, we are able to pinpoint the degree of edits from one deal to the next. We can analyze the overall extent of edits to highlight potential churning as well as identify individual changes within particular provisions. Additionally, we can assess the degree to which

and the Why of an Anomalous Concept, 55 BUS. LAW. 653 (2000) (discussing the role of fiduciary outs in providing an "escape hatch" to targets to consider unsolicited higher offers from third-party bidders); Thomas W. Bates & Michael L. Lemmon, *Breaking Up Is Hard to Do? An Analysis of Termination Fee Provisions and Merger Outcomes*, 69 J. FIN. ECON. 469 (2003) (arguing that deals with target termination fees entail greater premiums for target shareholders and higher completion rates than deals without such provisions); Albert Choi & George Triantis, *Strategic Vagueness in Contract Design: The Case of Corporate Acquisitions*, 119 YALE L.J. 848 (2010) (arguing that before closing the deal, the intentional vagueness of material adverse change ("MAC") clauses creates more efficient incentives for the seller, rather than more precise and less costly proxies); Yair Y. Galil, *MAC Clauses in a Materially Adversely Changed Economy*, 2002 COLUM. BUS. L. REV. 846 (discussing how unclear judicial interpretations of the contours of MAC clauses and material adverse effect ("MAE") clauses cast a shadow over merger deals); Ronald J. Gilson & Alan Schwartz, *Understanding MACs: Moral Hazard in Acquisitions*, 21 J.L. ECON. & ORG. 330 (2005) (using economic modeling to analyze the role that MAC and MAE clauses play in the structure of the standard acquisition agreement and the incentive effects for acquirers and targets); Sean J. Griffith, *Deal Protection Provisions in the Last Period of Play*, 71 FORDHAM L. REV. 1899 (2003) (discussing the significance of Delaware's judicially created limitations on deal protection provisions meant to resolve the conflicting incentives of the acquirer's and target's management when facing last minute third-party bids); Claire A. Hill, *Bargaining in the Shadow of the Lawsuit: A Social Norms Theory of Incomplete Contracts*, 34 DEL. J. CORP. L. 191 (2009) (arguing that the legal terms in acquisition agreements are intentionally ambiguous to deter litigation and incentivize negotiators to close the deal); Robert T. Miller, *Canceling the Deal: Two Models of Material Adverse Change Clauses in Business Combination Agreements*, 31 CARDOZO L. REV. 99 (2009) (advocating a judicial framework for interpreting MAC clauses that places the burden of material changes on targets and the burden of immaterial changes on acquirers during the closing period); Robert T. Miller, *The Economics of Deal Risk: Allocating Risk Through MAC Clauses in Business Combination Agreements*, 50 WM. & MARY L. REV. 2007 (2009) (arguing that the reciprocal allocations of deal risk in MAC clauses serve to further efficiency in transactions by decreasing the likelihood that parties will exercise termination rights); Brian JM Quinn, *Optionality in Merger Agreements*, 35 DEL. J. CORP. L. 789 (2010) (arguing that reverse termination fees that are equal in size to termination fees inefficiently leave targets exposed to more risk from exogenous events); Christina M. Sautter, *Rethinking Contractual Limits on Fiduciary Duties*, 38 FLA. ST. U. L. REV. 55 (2010) (advocating contractual limits on fiduciary outs to allow target company managers to sidestep fiduciary duties to make merger recommendations on third-party bids during the closing period); Alan Schwartz & Robert E. Scott, *Contract Interpretation Redux*, 119 YALE L.J. 926 (2010) (arguing for interpretative default rules in construing MAC clauses); Andrew A. Schwartz, *A "Standard Clause Analysis" of the Frustration Doctrine and the Material Adverse Change Clause*, 57 UCLA L. REV. 789 (2010) [hereinafter Schwartz, *Standard Clause*] (arguing that MAC clauses transform conventional default rules by (1) allowing a contractual exit in cases of frustration of secondary purposes or partial loss of value and (2) shifting exogenous risk from the acquirer to the target); Eric L. Talley, *On Uncertainty, Ambiguity, and Contractual Conditions*, 34 DEL. J. CORP. L. 755 (2009) (arguing that MAE clauses are a tool for allocating the risk of market uncertainty present while negotiating the acquisition agreement).

a particular section of the agreement remains consistent from deal to deal (e.g. a truly boilerplate provision) or is the focal point of drafting activity.

This approach allows us to show empirically that a high level of "editorial churning" takes place as merger agreements appear riddled with edits that are cosmetic and unnecessary.[13] The drafting of every acquisition agreement necessarily entails deal-specific edits, as well as reflects a fusion of the vision for the agreement from both parties as they seek to frame or reframe terms to their advantage.[14] Additionally, innovations are taking place in acquisition agreements in a more episodic fashion in response to exogenous events. But the fact that over half of the text of merger agreements is routinely rewritten from one deal to the next suggests that there is a high level of inefficiency in the precedent selection and drafting process that cannot be explained away in terms of substantive changes in acquisition agreements.

Our initial study demonstrated that public merger agreement terms are not based off a common "form" agreement, but rather are the product of a highly path-dependent "evolution" over many generations. This point is true even within large law firms where drafts are based on prior agreements rather than standardized form language. The absence of even firm-specific forms has led to haphazard and inconsistent lawyering, as lawyers add significant amounts of deal-specific edits to each deal and inadvertently retain deal-specific information from prior deals.

Professor Coates (in a paper being presented at this conference) questions our claim of editorial churning. He argues that the high level of edits reflects the incremental increase in length of acquisition agreements, which he attributes to changes in legal risks and deal and financing markets, as well as the increase in "linguistic complexity' of these documents.[15] This interpretation of our results is possible because our initial paper reflected a macro-view of editorial churning in assessing the extent of word changes from precedent to the final deal in each of the 12,000 agreements, and we

[13] *See* Jill Schachner Chanen, *Merger Mayhem: Lawyers are Sometimes to Blame When M&As Fail*, ABA J., Apr. 2004 (discussing how lawyers potentially endanger M&A deals through protracted negotiations concerning the fine print of the deal).

[14] *See* Avery Katz, *The Strategic Structure of Offer and Acceptance: Game Theory and the Law of Contract Formation*, 89 MICH. L. REV. 215, 277 (1990) (discussing the tradeoffs between standardization and customization in contractual drafting).

[15] *See generally* John C. Coates, *Why Have M&A Contracts Grown? Evidence From Twenty Years of Deals*, ECGI Working Paper No. 333/2016 (Nov. 2016).

did not engage in fine-grained analysis of particular provisions to test our hypothesis of drafting inefficiency. For this reason our original paper did not directly support the black hole or grey hole theory because it is possible that Professor Coates was right and that legal drafters were simply integrating paragraphs or whole sections of text as they engaged in innovative lawyering.

In this paper we address Professor Coates' critique by examining the degree of changes in individual boilerplate provisions from deal to deal. We find that the same editorial churning and drift are apparent when provisions are examined on a clause-by-clause basis as when agreements are examined as a whole. Our data shows that boilerplate provisions like virtually every other part of acquisition agreements are drifting over time due to incremental changes in each agreement which have cumulative effects over multiple generations. Haphazard editing takes place throughout virtually every part of acquisition agreements which afflicts even ostensibly standardized boilerplate language and potentially erodes the text and meaning.

We demonstrate that this textual "drift" takes place both within boilerplate that falls within a given chain of precedent, but also even more broadly for boilerplate provisions that have a common ancestor precedent, but evolve separately along different lineages of precedents. We also show spatially that the pattern of boilerplate "speciation" underscores the high impact of editorial churning in undercutting standardization of boilerplate. These findings reinforce the black hole or grey hole theories as "rote usage" of boilerplate without an appreciation of its significance fuels "encrustation" and "abrasion" of words in boilerplate provisions, i.e. the idiosyncratic addition or removal of words in the drafting process. The combination of these haphazard additions and deletions to the text of ostensibly standardized terms leads to significant divergences in the substance and meaning of boilerplate provisions over even a short number of generations.

B. The Nature of Boilerplate

The term boilerplate frequently carries a pejorative connotation among lawyers and the public for dubious reasons. Lawyers often act as if they are only adding value when they are making changes, which helps to explain some of the pathologies of legal drafting. Non-lawyers may regard standardized provisions as "legalese" that serves no substantive purpose. Self-interest may motivate this dismissive view of boilerplate provisions as

lawyers understandably want to believe that their intervention adds value to their clients in the drafting process.

The irony is that virtually all legal drafting is based on precedents whose purpose is to reduce the need for reinventing the wheel in the drafting process. The carryover of boilerplate language from deal to deal is designed to standardize legal terms, reduce uncertainty, and to economize the use of lawyers' time and resources. Boilerplate terms are ubiquitous in contracts and related transactional documents because of the advantages of consistency for efficiency and universality in meaning.[16] The additional appeal of boilerplate terms is that they offer the promise of partial standardization of otherwise non-standardized documents. This virtue is especially important in the acquisition agreement context since these agreements are highly non-standardized.

The problem is that the shortcomings of the drafting process expose boilerplate to a variety of distortions that may undermine the meaning, text, and purpose of provisions. The paradox of boilerplate is that the virtue of these standardized provisions may also expose some of its limitations. The advantage of standardization may become a downside if the provision is repeatedly carried over from agreement to agreement, but the parties lose sight of its meaning through rote repetition.

The related concern is that repetition without understanding coupled with haphazard editing may lead to "drafting drift" as lawyers make edits on standardized provisions (often for good but deal-specific reasons) and others copy them without identifying the ad hoc changes to the boilerplate provision. Boilerplate provisions matter offer the promise of drafting efficiency and reduces the likelihood of error only if the provisions are consistently copied from precedent to subsequent agreement, unless there is a deal-specific reason for making changes. As we will highlight, the problem is that the haphazard nature of the editing process means that the substance of ostensibly standardized language may change over time in inadvertent ways. We will show both the problem of "encrustration" of additional terms being added to boilerplate and "abrasion," the delegation of terms both pose

[16] *See* Marcel Kahan & Michael Klausner, *Standardization & Innovation in Corporate Contracting*, 83 VA. L. REV. 713, 719-20 (1997) (discussing the potential "learning benefits' of commonly used terms); Michael Klausner, *Standardization & Innovation in Corporate Contracting*, 81 VA. L. REV. 757, 783-84 (1995) (discussing the network benefits from familiarity with boilerplate terms).

equal threats to the substance and meaning of boilerplate provisions.

C. Unpacking the Transactional Drafting Process

Transactional drafting is a remarkably important engine for the economy and the bottom line of law firms,[17] yet the process that lawyers use when they draft agreements has not attracted much theorizing despite its importance to private ordering. We set out to shed light on the opaque legal drafting process by examining the final transactional outputs that are made public and try to reengineer the drafting process by matching each agreement to its likely precedent and in turn linking each boilerplate provision to its immediate ancestor. For this reason it is important to explain how it is possible to understand the precedent selection and drafting process by examining these public end products.

Although some areas of law practice have developed standard forms that underpin most negotiated documents, other areas such as M&A have not coalesced on standard forms. For this reason our starting point is the fact that virtually every M&A drafting process starts with the selection of one or more precedents from past deals.[18] Except in very extraordinary circumstances no law firm would start off a deal from scratch because lawyers and their clients value legal certainty and drafting from scratch would be prohibitively costly.[19] Building deal documents off of terms and boilerplate provisions that have been used repeatedly in past deals mitigates the parties' risk exposure and significantly reduces the time and money that lawyers (and their clients) must invest in the drafting process.[20] This fact is

[17] Thirty-two percent of the revenue of Am Law 50 law firms comes from transactional law, and transactional law practices play a disproportionate role in both the profitability and prestige of these firms. *See* PEER MONITOR, THOMSON REUTERS, RISE OF THE TRANSACTIONALS: HOW TRANSACTIONAL PRACTICES ARE INCREASINGLY ASSUMING LEADERSHIP FOR LAW FIRM GROWTH 1 (2015), https://peermonitor.thomsonreuters.com/wp-content/uploads/2015/06/Transaction-Practices-Spotlight_2015.pdf; *see also* GEORGETOWN LAW CTR. FOR THE STUDY OF THE LEGAL PROFESSION & PEER MONITOR, THOMSON REUTERS, 2015 REPORT ON THE STATE OF THE LEGAL MARKET 4 (2015), http://www.law.georgetown.edu/academics/centers-institutes/legal-profession/upload/FINAL-Report-1-7-15.pdf (documenting how transactional law has been a driver of law firm growth).

[18] *See* FIONA BOYLE, DEVERAL CAPPS, PHILIP PLOWDEN & CLARE SANDFORD, A PRACTICAL GUIDE TO LAWYERING SKILLS 153–54 (Cavendish 3d ed. 2005) (discussing the role of precedents in legal drafting).

[19] *See* SCOTT J. BURNHAM, DRAFTING AND ANALYZING CONTRACTS: A GUIDE TO THE PRACTICAL APPLICATION OF THE PRINCIPLES OF CONTRACT LAW 5–6 (3d ed. 2003) (discussing how attorneys "rarely start to draft on a blank slate. . . . [and generally] start with an existing contract or form").

[20] *See* STARK, *supra* note 11, at 335–36*;* Claire A. Hill, *Why Contracts Are Written in*

what makes this study possible. If we can identify the likely beginning point for a deal then we can potentially compare that with the end product of the final agreement to deduce what took place in the drafting process.

Because there is no standardized template for acquisition agreements either across the legal world (that firms actually use) or within law firms, the precedent selection is the first step in time-intensive editing and negotiation processes. Each agreement delineates the deal terms that lay out the details and structure of the transaction, which entails a balance of deal-specific information and standard legal terms. Lawyers must adapt the deal terms to the present transaction, which involves decustomizing the terms specifically crafted for the last transaction and recustomizing the terms specifically crafted for the present transaction.[21] Some decustomization is obvious (such as names and dates), but other terms may appear generic but were actually crafted to address particular issues concerning the precedent deal. This process of decustomization and recustomization of a precedent may then repeat itself if the document from the present transaction is subsequently chosen as the precedent for a later transaction. This dynamic process of copying and editing documents creates the potential for an evolutionary process. We explore the precedent selection, decustomization, and recusotmization process in turn.

Since precedent selection can reduce legal uncertainty, minimize time and resources invested in the deal, and offer strategic advantages to the buyer, one would imagine that precedent selection would be a, if not *the* focal point in the drafting process. But the precedent selection step appears to be the step that lawyers devote the least time to.

In an M&A transaction counsel for the acquirer typically choose the precedent that forms the starting point for the drafting process.[22]

"Legalese", 77 CHI.-KENT L. REV. 59, 63 (2001) (explaining why drafting from a prior agreement speeds the drafting process, increases certainty, and decreases cost).

[21] *See* Marcel Kahan & Michael Klausner, *Path Dependence in Corporate Contracting: Increasing Returns, Herd Behavior and Cognitive Biases*, 74 WASH. U. L.Q. 347, 354–55 (1996) (discussing the tradeoffs from customizing agreements rather than relying on standard terms).

[22] *See* Scott Austin, *Acquirers Back in the Game, but VCs Advised to Tread Carefully*, WALL ST. J. (Apr. 29, 2009, 4:36 PM), http://blogs.wsj.com/venturecapital/2009/04/29/acquirers-back-in-the-game-but-vcs-

In theory lawyers attempt to choose a precedent that is high quality, favors the client's side, and comes from a prior transaction with similar characteristics.[23] But rather than searching systematically for a precedent that is the best fit for the current deal, we showed in prior work that lawyers often choose precedents that they are familiar with based on prior involvement by the lawyer or the lawyer's law firm.[24] While this familiarity bias may lead to the selection of precedent that are not as on point as other documents that are publicly available, in theory this bias towards a firm's own precedents should reduce the need for idiosyncratic additions or deletions.[25] But in practice, this familiarity bias appears to do little to stop the legion of cosmetic changes that take place in most drafting processes.

Any negotiated precedent document will have deal-specific terms that the lawyers will need to "decustomize" for the present transaction. Although the lawyers will successfully identify many or even most of the customized terms that need to be removed, there will likely be "slippage" in any drafting process as non-standard customized terms remain in the finished product that are artifacts from prior deals. The idea of slippage derives from the notion that lawyers likely select precedents they are familiar with as a heuristic strategy for dealing with complexity. Few lawyers can be familiar with every textual variant of every term, and therefore rely on familiar precedents to compensate for bounded rationality. Assuming this is correct and that such slippage does occur, some inappropriate edits will slip through the review process.

The fact that the precedent may have been selected based on familiarity has two potentially offsetting implications for the degree of slippage. On the one hand, the familiarity makes the lawyers more likely to identify and remove customized terms (at least those that

advised-to-tread-carefully/ (discussing norm for acquirers to make the first draft of the merger agreement).

23 *See* ROBERT A. FELDMAN & RAYMOND T. NIMMER, DRAFTING EFFECTIVE CONTRACTS: A PRACTITIONER'S GUIDE 1-20 (2d ed. 2005) (discussing basic strategies in drafting contracts); JAMES C. FREUND, ANATOMY OF A MERGER: STRATEGIES AND TECHNIQUES FOR NEGOTIATING CORPORATE ACQUISITIONS 26–27 (1975) (discussing how the power to make the first draft gives the drafter leverage over other parties because defaults matter).

24 *See* Robert Anderson & Jeffrey Manns, *The Inefficient Evolution of Merger Agreements*, 85 GEORGE WASH. L. REV. 57 (2017).

25 *See* STARK, *supra* note 11, at 335–36 (discussing the benefits of heightened efficiency and legal certainty from precedent-based legal drafting).

favor the other side). However, the same selection based on familiarity rather than correspondence to the business terms of the present transaction means that document will often have superfluous terms that are inapplicable to the present transaction.

The failure to identify and decustomize terms from prior transactions may propagate throughout generations of a precedent or a particular provision. Subsequent users of the text may ascertain that a text is substantively non-standard, but may respond to that substantive deviation not by reversing the non-standard edits, but by changing other text in the agreement to respond to the non-standard terms. In this way, the edits may tend to accumulate and even increase over time as subsequent users attempt to adjust to the edits, creating a document that is a kludge. This process, like the game of telephone, can produce a product that differs significantly from the starting point or even becomes a black hole or a grey hole.

The necessary conditions for this process are (1) multiple generations and (2) slippage. Both of these conditions are characteristic of the M&A drafting process, in which precedents are used rather than standard forms. Neither of these conditions is present when standard forms are used because parties revert to original forms instead of copying new negotiated documents

II. DATA AND METHODOLOGY

A. Methodology

Our study compiled a data set of 12,407 merger agreements filed with the SEC between 1994 and 2014 and performed a word-for-word comparison of each of these documents.[26] The computer script visited each URL contained in the Archive Indices of the SEC EDGAR Database and collected the full text of each acquisition agreement.[27] We excluded any document whose title did not contain "merger" or "reorganization" to ensure that we were not including any non-acquisition agreements.[28] We also excluded duplicative agreements,

[26] *See Archive Indices of the SEC EDGAR Database*, SEC, http://www.sec.gov/cgi-bin/edgar_archive_indices [https://perma.cc/USL4-V94J] (last modified Apr. 28, 2014).

[27] Exhibit 2 is the exhibit where merger agreements are filed, along with any other "plan of acquisition, reorganization, arrangement, liquidation or succession." *See* 17 C.F.R. § 229.601(b)(2) (1995). Such agreements can also be filed under Exhibit 10, but primarily when they relate to other companies, such as subsidiaries.

[28] This approach eliminates agreement types that may overlap such as "Contribution

intra-firm reorganizations, reincorporations in other states, and private company acquisitions.[29] We also eliminated older plain-text agreements for which paragraph demarcations are unreliable,[30] resulting in a focus of a subset of our database on agreements filed after 2001.

The key to our analysis is the use of a computer program to engage in a word-for-word comparison of each agreement to every other agreement in the data set. The underlying premise is that a document retains substantial word-for-word similarity to its precedent document even after a high degree of edits. The same logic applies for our comparison of boilerplate provisions from precedent to subsequent agreement across numerous generations. The computer program calculated the "edit distance" (also known as the Levenshtein distance) between each pair of agreements.[31] Edit distance is a method for measuring the extent of textual similarity or dissimilarity based on the number of insertions and deletions (i.e., edits) that differentiate two documents.[32] The concept is analogous to the traditional "blacklining" or "redlining" process of comparing two documents with one another which is routinely used in transactional law drafting. The difference in our approach is that we are seeking to assess quantitatively the degree of difference between each agreement in order to determine which agreement form the most likely precedent

Agreement," "Stock Purchase Agreement," "Asset Purchase Agreement," "Transaction Agreement," "Share Exchange Agreement," "Arrangement Agreement," and the like. Although these agreements certainly contain overlapping language, this study focused on documents that were clearly public company acquisition agreements. Very short documents that are less than 15,000 characters were also eliminated because these agreements likely did not address the complex issues raised in larger public company acquisitions. Mutual holding company conversions were also excluded.

29 Near duplicates were defined as those documents filed within 100 days of each other and having 97% or more similarity to one another. Most of these were the identical document, but some were amended and restated versions of the same document. Many of the documents contained extraneous text such as attachments to the main merger agreement. To remove this text, this study disregarded text following the first occurrence (if any) of "In witness whereof," which typically signals the end of a merger agreement.

30 The paragraph demarcations are unreliable because paragraphs are separated with carriage returns but so are page breaks, making it ambiguous in many cases whether particular text is separated by a new paragraph or a new page. The HTML documents have tags indicating new paragraphs and therefore do not suffer from this problem.

31 *See* DAN GUSFIELD, ALGORITHMS ON STRINGS, TREES, AND SEQUENCES: COMPUTER SCIENCE AND COMPUTATIONAL BIOLOGY 215–16 (1997) (discussing the Levenshtein distance).

32 *See id.*, at 215–16.

for a subsequent agreement.

The computer program allowed us to engage in this comparative analysis for each agreement in our database. As a result, we were able to identify the likely precedent document for each merger agreement in the database by determining which document had the smallest length-normalized pairwise edit distance (among those with earlier dates than the given document). This finding provides us with a window for seeing the starting point and the end result for the drafting of each acquisition agreement, so that we can establish quantitatively the degree of edits in each transaction. We then compared the individual paragraphs in the descendant agreement to its ancestor agreement to determine the source for each paragraph in the descendant.

B. The Backdrop of Evidence of Inefficiency in M&A Agreements

In theory lawyers and clients ought to crave the legal certainty that comes from building on precedents and provisions whose language has stood the test of time (and of courts). Additionally, one might expect that acquisition agreements would have significant textual similarity since they generally follow similar broad outlines of categories of provisions. But contrary to these plausible hypotheses we found that there was little evidence of standardization among merger agreements. Not only was there significant divergence in the text from each agreement to its precedent, but also there was remarkable diversity in the merger templates that were being used. Table I highlights the remarkably small degree of commonalities among merger agreements based on word-for-word comparisons of both full documents and compressed documents that omit the most common words.

TABLE I. SIMILARITY DISTRIBUTION OF THE DATA

	Full Documents
More than 30% Similar	0.5%
25–30% Similar	3.8%
20–25% Similar	44.7%
15–20% Similar	40.3%
10–15% Similar	7.4%
Less than 10% Similar	3.4%
Median	19.9%
Mean	19.5%
*Based on a sample of 50,000 random comparisons drawn from the documents.	

The most striking finding is that only 4.3% of agreements were more than 25% similar despite the fact that they had nearly identical provisions and subject matter. These findings suggest that the world of acquisition agreements is remarkably diverse even though these agreements deal with similar categories of information, and each agreement is based on a precedent. The mean and median degree of similarity of documents were less than twenty percent which suggests that there is only a small core of standardization that cuts across the agreements.

Another step in our empirical analysis entailed examining the degree of divergence between each precedent and the resulting agreement. While we showed earlier the remarkable degree of diversity among acquisition agreements, the most telling evidence of inefficiency is the high level of editorial churning in the drafting process. Figure 1 shows the percentage of the textual similarity between documents and their precedents, assessed at the whole document level.

FIGURE I. SIMILARITY OF DOCUMENTS TO PRECEDENTS

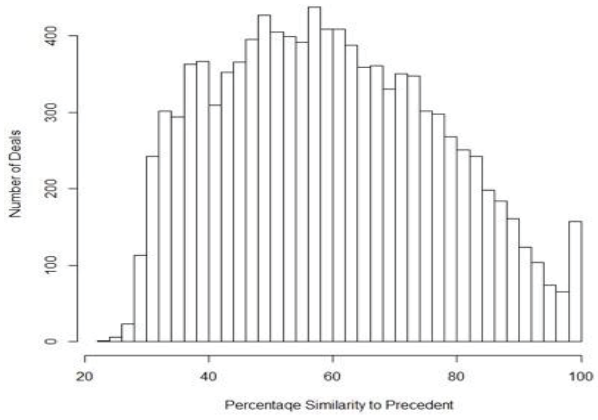


Figure I highlights the high degree of editorial churning that takes place during the drafting process. While there are significant outliers on both ends of similarity and dissimilarity, the largest number of acquisition agreements have approximately 50% similarity to their nearest precedent. Some acquisition agreements do have 80% or greater similarity with their precedents. But most of these documents are repeat-player acquisitions involving the acquirer which means they have limited applicability to the broader pattern of precedent selection.

It would be challenging to assess the precise degree of inefficiency because deal-specific edits are an essential part of every deal and the degree of edits will necessarily vary based off of the transaction. But we can extrapolate the amount of time that lawyers are investing in the drafting process to put the potential degree of editorial churning in context. From 1994 to 2014 the median number of words in an acquisition agreement increased from about 21,000 words to approximately 39,400 words a year. The rate of increase was just over 900 words a year as there was a remarkable "accretion effect" that led to a near doubling in the length of the average acquisition agreement. Table II highlights the dramatic increase in the length of merger agreements over time.

TABLE II. AVERAGE LENGTH OF ACQUISITION AGREEMENTS FROM 1994-2014

	Average Number of Words
1994	21,013.5
1995	22,435.5
1996	21,110.0
1997	21,653.0
1998	22,582.0
1999	23,850.0
2000	24,685.0
2001	25,601.0
2002	26,186.5
2003	26,697.0
2004	27,378.0
2005	29,116.0
2006	30,360.0
2007	31,992.0
2008	33,134.0
2009	35,344.5
2010	35,941.0
2011	37,467.5
2012	36,736.0
2013	37,614.0
2014	39,403.0

Of course some of this additional word count may be justified by legal responses to exogenous events or other substantive developments in the architecture of acquisition agreements. But while Professor Coates has pointed to some degree of substantive changes over this period,[33] in reality there appears to be little substance to justify this dramatic expansion in the length of merger agreements.

The evidence of a consistently high level of editing suggests that lawyers are doing an ineffective job of engaging in precedent selection and document design throughout the drafting process. Some deals may require more edits because of the distinctive nature of the deal, but "revolutionary" deals are few and far between and the degree of editorial churning that routinely occurs in the deal process suggests that there is inefficiency in the precedent selection and document design.

[33] *See generally* John C. Coates, *Why Have M&A Contracts Grown? Evidence From Twenty Years of Deals*, ECGI Working Paper No. 333/2016 (Nov. 2016).

C. The Challenge of Delineating Boilerplate

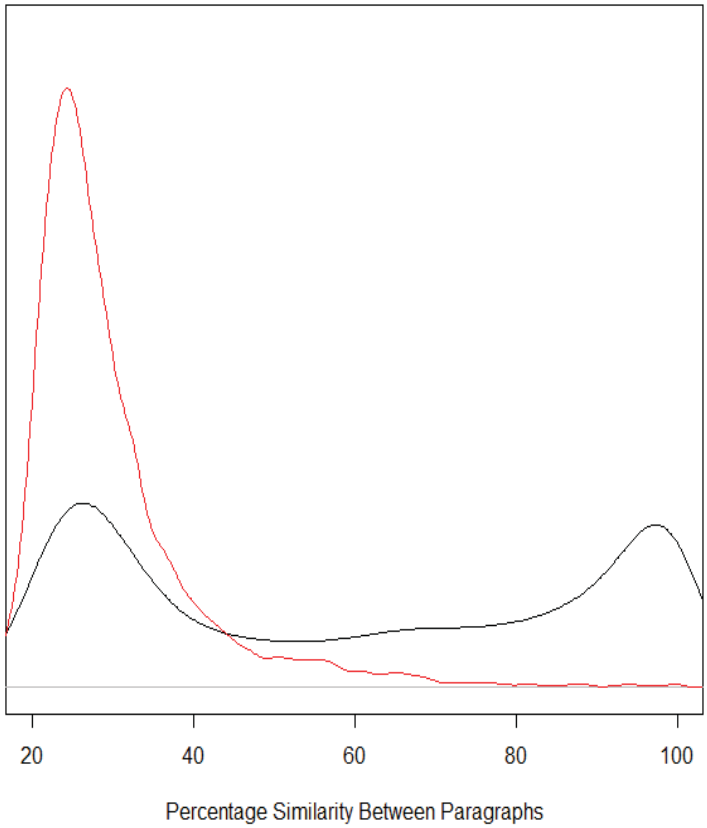
Having provided evidence of underinvestment in precedent selection and high levels of editorial churning, we turn to the more granular question of whether the textual changes reflect the insertion of new provisions and paragraphs or whether editorial churning is pervasive throughout the document. This question is key to testing the black hole or grey hole theory for boilerplate provisions, as we would expect to see the highest degree of standardization in acquisition agreements at the paragraph or provision level for boilerplate text if it were being substantively embraced.

We compared the individual paragraphs in each agreement to its precedent agreement to determine the source for each paragraph in the agreement. Our challenge was defining what constitutes boilerplate because of the high degree of edits throughout the agreements. Some types of clauses are readily classified as "boilerplate" based on their subject matter (e.g., Governing Law, Entire Agreement, Waiver of Jury Trial). Many other clauses, however, do not fall neatly into the boilerplate or non-boilerplate category, raising the question of how to define boilerplate text. This question is a threshold issue for empirical analysis of boilerplate in any context, as text is reused in many ways, even in fully negotiated paragraphs (e.g., jargon phraseology). But only certain types of reused text reaches a sufficiently high degree of standardization to qualify as boilerplate.

To answer this question, we turn to the data itself. The following figure presents the distribution of the percentage similarity of paragraphs to their nearest ancestor paragraph in the precedent agreement, denoted by the black line.[34] For comparison purposes, the red line shows the percentage similarity between paragraphs and the nearest paragraph from a merger agreement chosen at random (i.e. not the precedent document).

[34] This data was generated by drawing ten paragraphs at random from each descendant and computing the normalized edit distance to the closest paragraphs in the immediate ancestor for each such paragraph.

Figure II: Distribution of Similarities Between Paragraphs



The Figure makes it clear that there is a bimodal distribution of similarity to precedent paragraphs, with a fairly strong bifurcation between boilerplate provisions and fully negotiated provisions. The right hump in the black line is the relatively common boilerplate provisions that are 70%-100% similar to their precedents. The left hump in the black line is the negotiated provisions that are not much more similar than clauses from random agreements. Interestingly, in the middle there is a range between 40%-70% similar where the moderately negotiated provisions in the precedent agreements are only slightly more similar to each other than the agreements chosen at random.

The black line in the Figure makes it clear that merger agreements contain a large number of boilerplate paragraphs, a large number of fully negotiated paragraphs, and relatively few paragraphs in between these extremes. This point validates the idea that boilerplate paragraphs differ qualitatively, not just quantitatively,

from negotiated (generally deal-specific) provisions. Relying on the Figure, we set the threshold for boilerplate copying at 70% similarity and up, which captures most of the agreements in the boilerplate category. Thus, our analysis focuses on the degree of continuity or evolution from precedent to the next between paragraphs that are 70% or more similar to one another.

Our interest is in the paragraphs that are copied over multiple generations to determine the degree and type of drift from the original ancestor over time. To examine this drift, we construct a "family tree" for the set of paragraphs. We begin by taking the set of paragraphs that have no descendants, which often (but not always) come from later agreements near the end of our data coverage. We then trace the copying history of each of those paragraphs back in time, finding its ancestor, the ancestor of the ancestor, and so on. Because each paragraph has only one immediate ancestor, these lineages do not branch as they are traced back in time.

We then reverse the direction, constructing a family tree for each ancestor by tracing the descendants of each ancestor over time using the reverse-lineages just constructed. Accordingly, we follow the evolution of each lineage of paragraphs from the original ancestor to all of its direct and indirect descendants. This creates a tree-like structure for each ancestor. Some ancestors have many branches (and branches of branches), while others have a single lineage through time.

We then drew a sample of 28,717 ancestor paragraphs from the total set of 202,422 ancestor paragraphs to analyze, and included all descendants of those ancestor paragraphs. The following table presents some descriptive statistics about the boilerplate paragraphs analyzed in the study.

Table III. Descriptive Statistics

	Median	Mean	Standard Deviation
Words per Paragraph	68	88.8	69.7
Number of Edits per Paragraph per Generation	4	14.4	24.4
Number of Generations per Lineage	2	2.9	1.4

The boilerplate paragraphs tend to be relatively short with a median length of 68 words and tend to have relatively few edits from one generation to the next, with a median number of words edited of just 4 (about 6% of the median paragraph length). The lineages also tend to be very short, with the median lineage only 2 generations long (meaning that the median paragraph was copied only once). This latter result occurs because most merger agreements themselves are copied only once (if at all), with the descendant never being copied again. This means that there are a lot of "dead ends" in the evolutionary process. In some cases we find that paragraphs are copied over many generations of agreements. In other cases, paragraphs are copied once and then become "extinct." Having established reasonable parameters for what constitutes boilerplate provisions in the M&A context and a framework for identifying family trees for these provisions, we turn to our analysis of the degree of drift in boilerplate provisions over time.

III. RESULTS

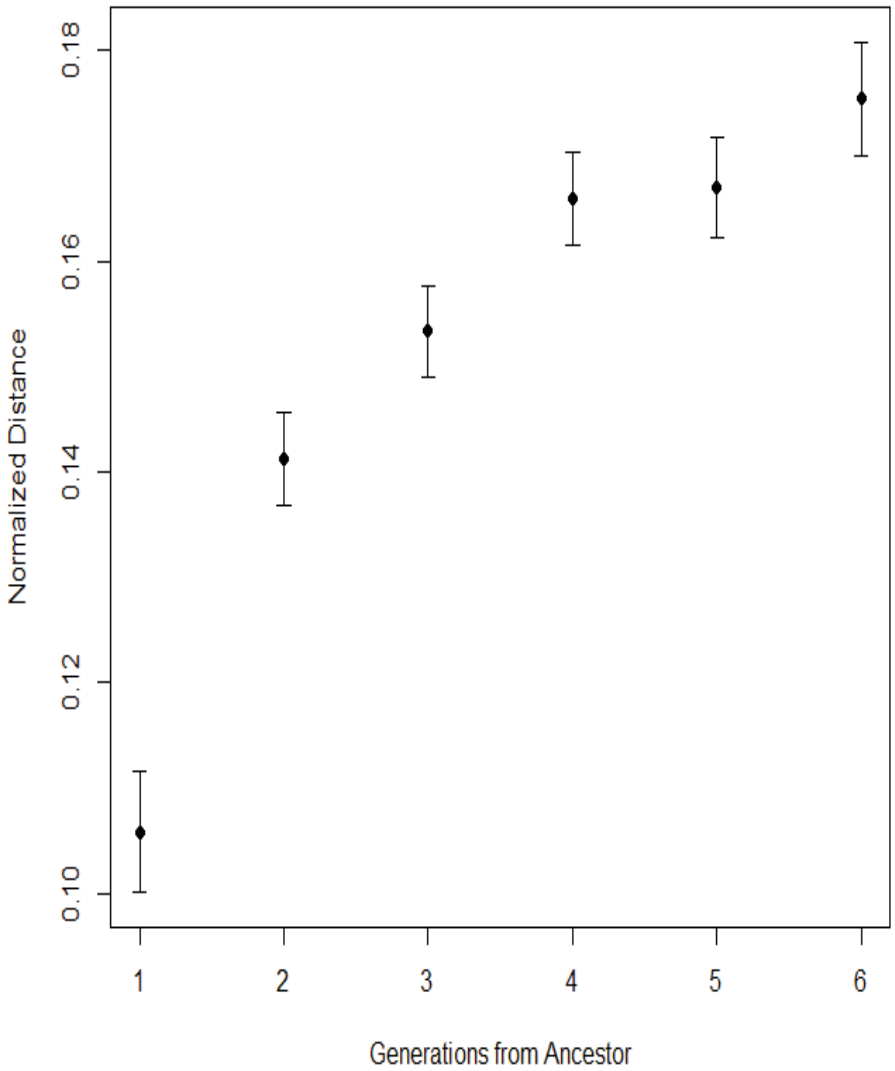
In this Part we examine the evolution of boilerplate clauses from three different perspectives. First, we examine the extent of drafting drift over generations between the original boilerplate provision and direct chains of descendants. Second, we examine the astounding variety of clauses produced by a single boilerplate ancestor which evolve separately along different lineages of precedents. Third, we examine the geometry of the relationships among the clauses to illustrate spatially the high degree of divergence in boilerplate provisions over time.

A. The Drift of Lineages over Time

The theory developed above suggests that slippage in the drafting process will have cumulative effects that will distort the boilerplate text. If drafters are unable to identify all non-standard edits embedded in a precedent document (or simply fail to invest time in checking for consistency), some of the edits from previous transactions will be retained in addition to the edits added for the present transaction. As a result, each generation of a paragraph will tend to differ more from the original provision than the last generation. Thus, if slippage is occurring we should observe paragraphs drifting farther from their original ancestors as the number of generations between the drafts increases.

To examine whether paragraphs drift over time, we examined all lineages with at least six generations and compared the text of each descendant paragraph at each generation to the original ancestor. The results of this comparison are presented in Figure 2 below, with point estimates and 95% confidence intervals denoted by the points and bars, respectively.

Figure III: Distance From Ancestor by Generation

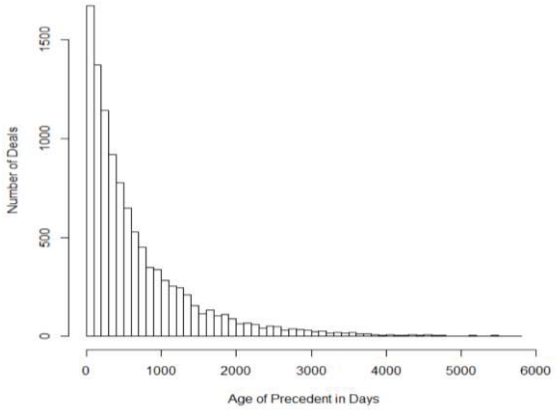


The amount of average overall drift from one generation to the next is remarkable considering that most pairs of paragraphs in our analysis only have very slight edits (or even none in any given

generation). Small changes have cumulative effects over multiple generations, however, eventually producing a descendant that is quite different from its ancestor in terms of the text. Over a long enough time horizon the substance of these provisions may be incrementally transformed which may undermine the purpose of having standardized text.

To put this finding in context, the median number of days between a document and its precedent is 423.5, meaning that the median precedent document is just over a year old.

FIGURE IV. AGE OF PRECEDENT DOCUMENTS



As a result, the boilerplate evolution we witnessed over six generations covers roughly six years, which reinforces the story that there is rapid "speciation" of boilerplate that parallels the broader editorial churning in merger agreements, albeit at a somewhat slower pace of change compared to the agreements as a whole.

A. The Heterogeneity of Descendants Over Time

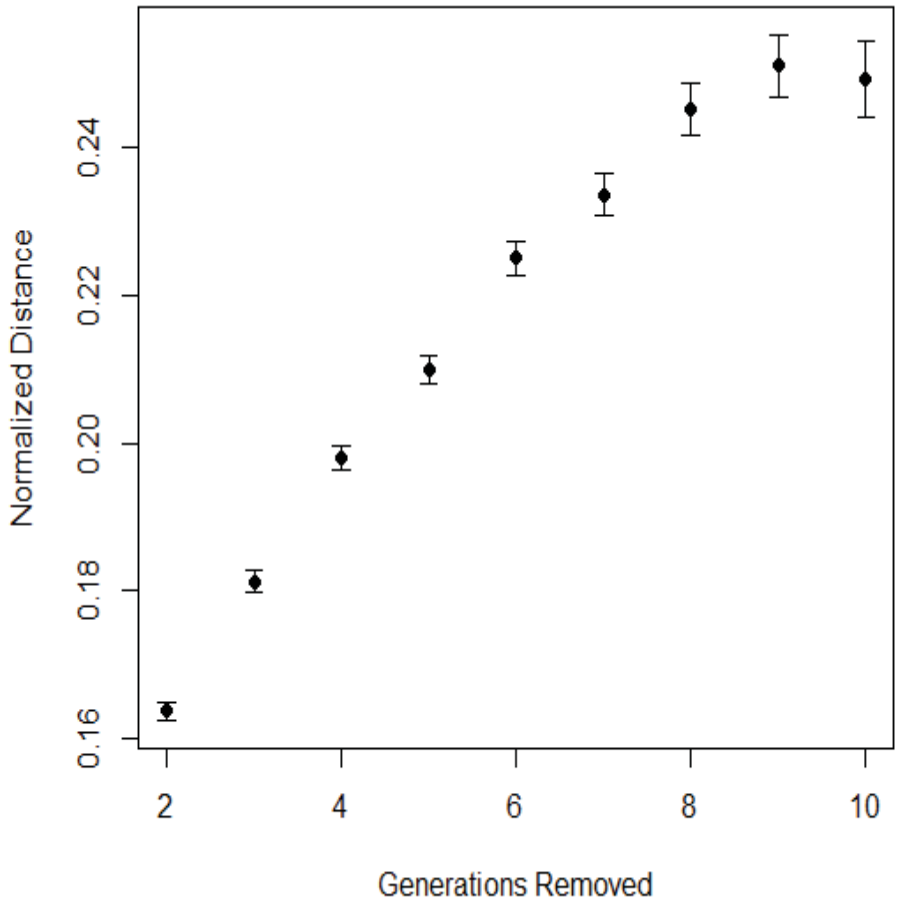
This subpart examines the heterogeneity of descendants of a particular ancestor over time. In other words, to what extent does a single ancestor produce a variety of descendant clauses? Although this question is closely tied to the drift of each lineage over time examined in subpart A, the two issues are distinct. It is entirely possible that different lineages from the same ancestor could drift rapidly over time as in subpart A, yet not diverge from one another. This result would occur, for example, where the various lineages were responding in tandem to external shocks, such as changes in the economic or

regulatory environment which, if true, would be consistent with Coates' thesis of change being driven primarily by innovation. If, instead, the lineages from the same ancestor diverge rapidly, the explanation might more plausibly be attributed to editorial churning rather than rational adaptation.

Because our aim is to examine multiple lineages from the same ancestor, we exclude clauses that only had one or fewer descendants. For each "family tree" descended from an ancestor, we compute the diversity among the ancestor's descendants according to the number of generations to connect them. So for a sibling pair of paragraphs descended from a common parent the number of generations is two (one up from one sibling to the parent and one down to the other sibling). For a grandchild to its "uncle" paragraph the distance would be three (two generations up to the grandparent and one down to the uncle paragraph. For each such generational "distance" we then compute the average normalized edit distance among the paragraphs at that distance to assess the overall heterogeneity by number of generations removed.

The following Figure sets forth the mean heterogeneity of descendants of the same ancestor by number of generations separating texts.

Figure V: Distance Among Descendants By Generations



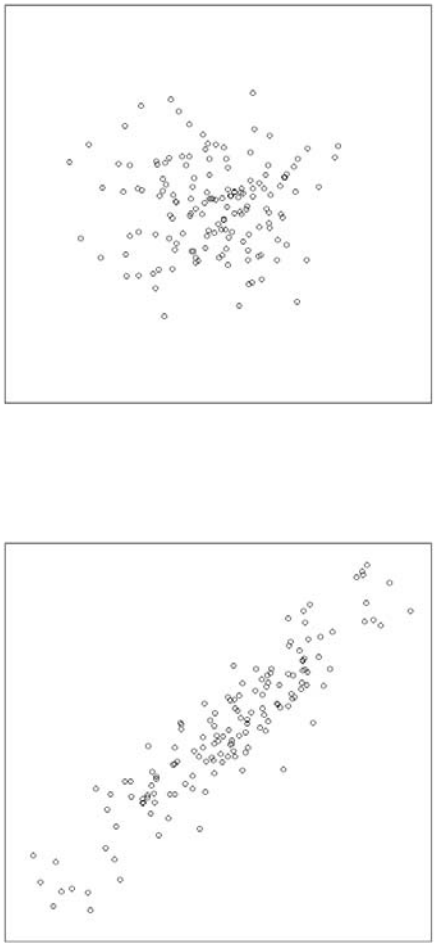
The average distance among descendants from the same ancestor increases with generational separation just as the distance from an ancestor increases over the generations. Comparing this Figure to Figure 2 above shows that the process of "drift" is not only away from ancestors, but also away from other lineages descended from the same ancestor. Indeed, all paragraphs are moving away from one another in an "expanding universe" of clauses akin to the "Big Bang" theory on the scale of boilerplate. This finding is consistent with widespread editorial churning that is haphazard in nature, rather than being driven by a response to exogenous legal events or any attempt at innovation.

C. Geometry of the Clauses

The previous sections have shown that virtually all boilerplate paragraphs are moving away from one another in terms of their textual similarity. In this subpart, we attempt to characterize the evolution of merger agreement clauses in terms of the geometrical shapes of the clauses in a high-dimensional space. Although it is unusual to think of contract clauses in terms of their shapes, the concept of difference or distance lends itself to such a graphical interpretation.

Consider the following figures derived from simulated data.

Figure VI. A Spherically Distributed Cluster and an Ellipsoidal Cluster



The figure on the left has a well-defined center (or standard) with a point cloud around it. The figure on the right is elongated. These two figures have approximately the same average distance to the nearest point.

The top plot is similar to what one would expect from documents based on a standard form. Although parties can negotiate terms in standard forms, parties will reverse deal specific edits in subsequent uses of the form, meaning that documents based on forms will have a (hyper-)spherical distribution. This fact does not necessarily mean the documents are only lightly edited. As in Figure VI, documents can be close to the center or far from it; the key is that there is a center to which documents tend to revert because of the use of a form or standardized language. Similarly, documents not based off forms can change slowly and incrementally, but end up very far from where they began. The distinction between the two shapes is whether the past editing history of the document shapes its future form. The same logic applies to analysis of particular provisions of boilerplate language which serve as loci of standardization within acquisition agreements.

With this background, we now examine the data from the merger agreements. Although there are many methods that could examine whether the underlying data have a spherical or non-spherical structure, we use the eigenvalues derived from a multidimensional scaling of the distance matrices for sets of related paragraphs (those derived from a common ancestor).[35] If the "clouds" of points representing documents are roughly spherical, then the eigenvalues should be close to one another. If, on the other hand, the clouds have a linear structure to them, at least one of the eigenvalues will tend to be significantly larger than the other ones.

Indeed, we find that very few of the "family trees" of boilerplate agreements have the spherical structure we would expect from documents based on a standard form. The first eigenvalue accounts for a median of .65 of the variation of all the eigenvalues (obtained by dividing the first eigenvalue by the sum of the eigenvalues). This suggests that a small number of eigenvectors (or

[35] We used the cmdscale( ) function in R, which performs classic multidimensional scaling.

even one) can account for most of the variation, indicating that our data have a structure that deviates markedly from a spherical shape.

One important implication of these findings is that there is no "center" or "standard" to most boilerplate paragraphs. One might expect that the ancestor paragraphs of a set of descendants would be the "center" of the descendants, and indeed that would be the case if the ancestor were used as a "form." But each paragraph's form is transient without fixed referents to which it reverts in subsequent generations. The ends of these elongated "point clouds" bear little resemblance to each other, meaning that new forms of clauses are constantly arising in a process similar to speciation.

IV. DISCUSSION

The results confirm that the clause-by-clause evolution of merger agreements mirrors the overall evolution of agreements. The changes that are introduced at each generation of a document's evolution tend to be preserved in subsequent generations, causing the text to drift significantly over time. This finding has a number of implications for the drafting process as well as the emerging literature on black holes and grey holes in contract law, which are explored in this Part.

A. Where Does the Evolutionary Process Lead?

The analysis in this paper provides support for the rote use and encrustation processes that lead to black holes and grey holes in contracts. We also identified equally significant evidence of "abrasion" as deletions shaped the evolution of boilerplate terms, even though over time these provisions, like acquisition agreements as a whole, tended to increase by length year by year. The data suggests a strong role for slippage in the drafting process that may lead toward unconsidered and ultimately unintended variations in documents.

But the main empirical conclusion of this analysis, that edits in one generation are often passed down to subsequent generations, could have other interpretations. For example, it is possible that the edits improve the document and are retained as part of a process of evolution toward better agreements. Our analysis cannot definitively resolve the question of whether the cumulative edits over many generations have effects on M&A boilerplate that is positive, negative,

or neutral. That would require a more find-grained, qualitative analysis of individual boilerplate terms to attempt to assess the legal implications of textual changes over time. The fact that the descendants of a common ancestor boilerplate term diverge from one another is strong evidence of random drift rather than conscious improvement. The random drift and consequent speciation may lead to black holes or grey holes as language becomes unmoored from accepted interpretations. We would need to conduct further qualitative research to analyze in a selective fashion the degree to which the textual evolution of particular provisions has transformed the substantive meaning of boilerplate.

B. The Consequences of Drift and Speciation

The drift and speciation characteristics of the evolutionary process may lead to black holes or grey holes of boilerplate losings its meaning. But this erosion of meaning does not necessarily occur in the majority of the cases, at least over a small number of generations. However, the longer the number of generations of drift from the original boilerplate, the more likely that the meaning of boilerplate will evolve over time in tandem with the increasing level of textual changes. Since lawyers typically choose precedents that are approximately a year old, it would be possible to extend out the number of generations in future studies and to combine that with fine-grained qualitative analysis of the meaning of particular provisions.

The result of rapid speciation that is certain to occur, however, is an undermining of the value of network effects that come through standardization of language. Our study stipulated that boilerplate consisted of text that had a seventy percent or more degree of similarity to a paragraph in its immediate precedent document, which is a high degree of similarity given the non-standardized nature of acquisition agreements. But our empirical analysis has shown that the degree of drift effectively undercuts the emergence of truly standardized boilerplate language in M&A agreements, at least in the sense of a standardized form that we see in other areas of contracts. This fact imposes significant costs on market participants.

The first type of cost of non-standardization is the easiest to see. The extra unnecessary effort expended in the drafting process as lawyers introduce random edits, another set of lawyers using the

precedent attempt to compensate for those edits, and so forth occurs generation after generation. This phenomenon is the "editorial churning" we identified was occurring on an entire document basis in our previous article. In this paper, we show that the same churning is occurring on a clause-by-clause level, which serves as evidence of inefficiency.

The more important costs of the lack of standardization, however, come through impairment of the network effects that arise through standardization of boilerplate in other contractual contexts. As ancestors change through encrustation and abrasion and rote repetition of encrusted texts, the value of the network effects decline. As ancestors split into multiple descendant species based on divergent lineages of precedents and provisions, the network effect value declines further. In this paper we have shown that both trends occur in the M&A boilerplate context. The text both drifts from its original version and splits into multiple lineages, each of which drifts away from the ancestor and away from each other.

C. Caveats

Our work provides evidence for the rote use and encrustation phenomena in the context of merger agreements. The results have a number of limitations, however, as detailed in this section.

Limited Number of Generations

While our database of acquisition agreements covers 1994 to 2014, the nature of the SEC's pre-2001 document format makes it difficult to engage in paragraph for paragraph comparisons for this earlier period. For this reason we plan to collect additional merger documents from earlier periods for future work to ensure that we have more comprehensive coverage of earlier generations of agreements. To extend our database further in time will require collection and processing of paper documents from the SEC archives. Our hope is that this additional data will make it easier to substantiate both the extent of document drift and the degree of erosion of the meaning of boilerplate provisions over time.

Missing or Misidentified Precedents

The starting point for our analysis is the identification of the likely precedent for each public company acquisition agreement in our data base. We only look for precedent clauses within the documents

determined to constitute the precedent documents. It is possible the precedent documents are not the actual precedent documents because the actual precedents are not available in the dataset. It is also possible that the precedent clauses are not found because the clause was copied from an agreement other than the precedent for the whole document (i.e., a clause was swapped from a different precedent).

While we would tend to discount the probability of either of these possibilities in most cases, we do recognize that much of the "innovation" in acquisition agreements occurs from copying the innovations of first movers in other acquisition agreements. For this reason if an exogenous legal shock arises, it is quite possible that lawyers will take advantage of SEC-mandated transparency and the absence of intellectual property protection to copy and paste relevant provisions from an agreement that is not the precedent for the current deal. This issue is more significant for our study of boilerplate than our broader study of the evolution of acquisition agreements because of our ability to identify the likely precedent for each agreement with a high degree of probability based on the degree of similarity.

But we should not overstate the risk of opportunistic copying of innovations skewing our boilerplate analysis. In the case of swapped-in language from another precedent, we would expect to find large edit distances. As shown in Figure II, a typical merger agreement clause does not find close matches in another random merger agreement, even for boilerplate provisions. Therefore, we would expect that we would typically not even identify the swapped-in clause as boilerplate for the purposes of our study. For this reason the boilerplate provisions that are the focus of our study are much more likely to have continuity from one precedent to the next. While numerous edits take place throughout acquisition agreements and boilerplate provisions, the empirical evidence suggests that piecemeal editing rather than transplantation of terms from other precedents is the norm.

CONCLUSION

Our study shows that the high levels of "editorial churning" that take place in the process of transforming each precedent into the current acquisition agreement affect agreements on a clause-by-clause basis, not just an entire document basis. Boilerplate provisions live on from deal to deal,

yet gradually shed their textual integrity and potentially lose their clear meaning as they evolve over generations of lineages.

We show that incremental changes in boilerplate from one generation to the next lead to rapid "speciation" of the terms. We demonstrate that this textual "drift" takes place both within boilerplate that falls within a given chain of precedent, but also even more broadly for boilerplate provisions that have a common ancestor precedent, but evolve separately along different lineages of precedents. Our findings reinforce the black hole or grey hole concern that rote usage, combined with encrustation and abrasion of terms may distort the degree of standardization and meaning of boilerplate over even a short number of generations. We plan on building on this study for future research that is larger in scope and duration and also integrates qualitative assessments of the evolution of particular boilerplate provisions over time.