Abstract

This essay in celebration of Grossman and Hart (1986) (GH) discusses how the introduction of incomplete contracts has fundamentally changed economists’ perspectives on corporate finance and control. Before GH, the dominant theory in corporate finance was the tradeoff theory pitting the tax advantages of debt (relative to equity) against bankruptcy costs. After GH, this theory has been enriched by the introduction of control considerations and investor protection issues. This essay assesses how our understanding of corporate finance has been improved as a result and where the incomplete contracts perspective has failed.
It is a great pleasure to write this essay in celebration of Sandy Grossman’s and Oliver Hart’s classic 1986 article “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration”, and I want to take this opportunity to write a more personal account than is usually the case on such occasions. I have had the good fortune to be Oliver Hart’s student, first as an undergraduate in Cambridge from 1980 to 1981 and later as a Ph.D. advisee from 1982 to 1986 at the London School of Economics and MIT, just at the time when he and Sandy were writing their classic article.

To appreciate the importance and novelty of Sandy’s and Oliver’s contribution it is helpful to put it into context of their earlier research and the state of economic theory at the time. Unfortunately, as I only met Sandy much later, and only got to know his work like everyone else by reading his articles, my account here will be somewhat biased and will reflect more Oliver’s perspective as I understood it. It is fair to say that among the hottest research topics in economic theory in the early 1980s were General Equilibrium (GE) Theory, Rational Expectations, and Information Economics (Mechanism Design) broadly defined.

Both Oliver and Sandy had made several important contributions in these general areas. Oliver’s thesis was in the area of General Equilibrium theory with Incomplete Markets, and in the early 1980s he was working on General Equilibrium with Imperfect Competition. One central conceptual issue arising in GE with incomplete markets or imperfect competition is what the objective function of the firm is and to what extent there can be shareholder unanimity. Oliver and Sandy had written several major articles on this topic. They had also made important contributions on implicit labor contracts and involuntary unemployment, and on a general characterization of the Principal-Agent problem with Moral Hazard.

Much of this work appeared in the most prestigious journal at the time – Econometrica – and achieved the highest standards of mathematical rigor and generality. Indeed, both Sandy and Oliver were seen as belonging to an elite group of young theorists taking over the field of economic theory from the founding giants of mathematical economics like Kenneth Arrow, Frank Hahn, Gerard Debreu, and Robert Aumann. Certainly that is how I saw
things when I was starting my studies in economics. It is important to emphasize that there is virtually no hint in this earlier work of Sandy’s and Oliver’s of the ideas they were about to develop in their 1986 article, which by all appearances is a complete break from their earlier research both methodologically and conceptually.

What explains this sharp break? What prompted Oliver and Sandy to embark on this path-breaking endeavor? I have been asking myself this question, especially given that I was myself at the time struggling to master the complexities of GE theory, and mathematical methods, that they were in the process of abandoning. I can only speculate, but I can see mainly two motivations.

First, I was privileged to observe as a student of Oliver’s, that he had long had a deep interest in the theory of the firm, so much so that he encouraged his undergraduate students to read Coase (1937) as well as the early managerial theories of the firm of Penrose (1958), Baumol (1959), Cyert and March (1963), Williamson (1964), etc. One critique he made about these writings especially struck me: it is not that he was not persuaded by these ‘managerial perspectives’ of the firm, but that as theories he thought they failed because they were too detailed, descriptive, with too many different variables, to lend themselves as useful analytical tools. He was already thinking of how to formulate an analytically simple theory of the firm.

Second, the ultimately unsatisfactory outcome of Oliver and Sandy’s work on the Principal-Agent problem (Grossman and Hart, 1983) must have played an important role in prompting Sandy and Oliver to change direction. Why do I say unsatisfactory outcome? After all, as a piece of mathematical analysis, this is a brilliant achievement solving in one swoop both a fundamental non-existence problem in Principal-Agent problems, and a methodological flaw with the ‘first-order approach’ pointed out by James Mirrlees (1974, 1999). The unsatisfactory outcome is that the theory in full generality yields no robust predictions on the shape of optimal incentive contracts, and it produces no robust comparative statics results.

In other words, the shape of the optimal incentive contract a Principal offers an Agent
is highly sensitive to the specific environment the parties find themselves in and is likely to be incredibly complex. Far more complex than anything one can see in reality, and far less robust than the fairly standard incentive contracts one typically observes. This concern is clearly voiced in the survey of Contract Theory by Oliver Hart and Bengt Holmstrom (1987) delivered at the World Congress of the Econometric Society of 1985:

“Indeed, the economic credibility of the contractual approach may be called into question when, as often happens, optimal contracts become monstrous state-contingent prescriptions. How are such contracts written and enforced?” [Hart and Holmstrom, 1987, pp 74]

and

“The main [concern] is its sensitivity to distributional assumptions. It manifests itself in an optimal sharing rule that is complex, responding to the slightest changes in the information content of the outcome \( x \). Such “fine-tuning” appears unrealistic. In the real world incentive schemes do show variety, but not to the degree predicted by the basic theory.” [Hart and Holmstrom, 1987, pp 90]

Interestingly, in response to this critique the contract theory literature has taken two separate directions, which even today remain in competition, and confusingly, sometimes overlap. The first direction taken by Grossman and Hart (1986) is pragmatic and simply limits contractual complexity by imposing plausible exogenous contract enforcement constraints. This is the *incomplete contracts* approach they pioneered. The second direction is fundamentalist and seeks to endogenously derive simple and realistic optimal contracts – such as e.g. linear incentive contracts – from a complex dynamic programming problem in which the Agent has a ‘rich’ action set. This is the approach pioneered by the optimal security design literature following Townsend (1979) and by Holmstrom and Milgrom (1987) in the context of dynamic Principal-Agent problems.

Both approaches have been important in shaping corporate finance theory post Grossman and Hart (1986). The incomplete contracts approach, which Philippe Aghion and I embraced
has allowed us to address issues of (contingent) control allocation and renegotiation which had been absent from previous discussions of corporate finance and financial contracting. My other work with David Scharfstein turns out to have contained both elements of the incomplete contracts approach and elements of the dynamic programming approach (for lack of a better term), as subsequent work by DeMarzo and Fishman (2007) and DeMarzo and Sannikov (2006) has highlighted.

Before discussing the incomplete contracts literature in corporate finance I will begin with a brief summary of the state of corporate finance theory pre GH-1986. I will then turn to a discussion of some of the main themes of the corporate finance theory literature post GH-1986 and evaluate the relative merits of the incomplete contracts and dynamic programming approaches. I shall argue that while both approaches have considerably improved our understanding of the dynamics of corporate finance they both still lack operational and practical relevance. Unlike the tradeoff theory inherited from Modigliani and Miller, which has huge practical relevance, the more modern incentive and incomplete contracting approaches have so far had a limited practical impact. I will conclude in suggesting a way forward in closing this important gap, based on recent work of mine with Hui Chen and Neng Wang.

1 Modigliani and Miller and the Foundations of Corporate Finance

Modern Corporate Finance practice is founded on Modigliani and Miller’s classic (1958) article on the cost of capital. In this article they lay out a methodology for valuing investments and argue that the value of a firm is independent of how the firm is financed. Economists have mostly focused on the famous Modigliani and Miller (MM) irrelevance theorem, stating that when capital markets are competitive and efficient, and when there are no taxes and bankruptcy costs, then the way a firm is financed is irrelevant. That is to say, under these conditions the value of a firm is independent of its liability structure, and only depends on
the value of its assets. The *MM irrelevance theorem*, already implicit in Arrow and Debreu’s (1954) proof of existence of a general competitive equilibrium, is a central proposition in general equilibrium theory, which pins down the objective function of a firm (and ensures ‘*shareholder unanimity*’).

In the real world, of course, some key conditions of the theorem are not met, in particular the presence of taxes and bankruptcy costs. Therefore, the way firms are financed in practice is not irrelevant to their value. Nevertheless, the Modigliani and Miller logic and their approach to determining a firm’s cost of capital has still proved to be of enormous practical value. I am referring here to their famous *weighted average cost of capital* (*WACC*), and their approach to valuation based on discounting a firm’s *free cash flow* using *WACC*. Combined with a suitable asset pricing model (most often in practice the Capital Asset Pricing Model, *CAPM*), and correcting for taxes, the discounted cash flow approach using *WACC* is today the most widely used approach to valuation and corporate investments (see Graham and Harvey, 2001).

This approach to valuation is closely linked with a simple theory of the optimal capital structure of the firm: the *tradeoff theory* (see Miller 1977, Brennan and Schwartz, 1978, DeAngelo and Masulis, 1980, and Leland, 1994), which is still widely accepted. According to this theory, debt is a cheaper source of capital for the firm after tax if one ignores bankruptcy costs. If it were not for deadweight bankruptcy costs it would be efficient for the firm to finance itself only with debt. In the presence of bankruptcy costs, the optimum debt-equity ratio is determined by equating the marginal *tax shield benefit* of debt with the expected marginal *bankruptcy cost*.

The tradeoff theory assumes that the firm’s cash flow is observable and verifiable, and that any promised debt repayment satisfying the firm’s *limited liability* constraint is enforceable. Models of the tradeoff theory typically assume that Debt is a fixed claim (fixed coupon payments plus principal) which is independent of the firm’s realized earnings and equity is a claim entitling the owners to the firm’s free cash flow net of debt obligations. These
models also typically assume that the firm’s cash flow is fixed and reduce the firm’s financing problem to an optimal allocation problem of the firm’s cash flow to the holders of debt and equity, the tax authorities and nature (in the form of deadweight costs of bankruptcy).

This is a highly reductive theory of corporate finance, which misses many key aspects of corporate financial management. And, just as the CAPM, the tradeoff theory has found little empirical support (see e.g. MacKie-Mason, 1990, and Rajan and Zingales, 1995). Still, the tradeoff theory, especially in its ‘dynamic’ version following Leland (1994), retains a central place in corporate finance, mainly because it offers the most operational approach to the determination of the firm’s optimal capital structure and to the valuation of risky debt. True, the Leland (1994) model ‘predicts’ excessively high leverage ratios, but the more dynamic formulation in Goldstein, Ju, and Leland (2001) is able to predict reasonably accurate debt levels and credit spreads.

2 Grossman and Hart: Incomplete Contracts and Corporate Control

How is the corporate finance theory landscape changed post Grossman and Hart (1986)? One central dimension missing from the MM approach to corporate finance discussed above is ownership and control. The MM approach remains silent on how a firm’s cash flow is determined. It takes the cash flow as given and only asks how it should be allocated among different claimants and how it should be valued. In contrast, the incomplete contracts approach to corporate finance seeks to understand corporate control and how the exercise of control is affected by the firm’s choice of financing.

This is a richer and significantly more complex theory than the tradeoff theory. I will argue that it has yielded important new conceptual insights, but that it has so far only had a marginal operational impact. As much as chief financial officers (CFOs) feel that they need to understand the MM approach to corporate finance, they have so far not shown much
interest in the more modern theories of corporate finance that emphasize, adverse selection, moral hazard, incomplete contracts and control. It is not that they are unaware of the importance of these issues, but that the theory so far has offered little operational guidance. Before dealing with corporate finance post Grossman and Hart (1986) I must briefly discuss the ‘agency theories’ of corporate finance, which provide a first analysis of the endogenous determination of a firm’s cash flow.

2.1 Adverse Selection and Moral Hazard

Jensen and Meckling (1976) provide the first formal analysis of the incentive implications of a firm’s choice of capital structure. They start from Berle and Means’ (1932) observation that if the managers in control of a firm’s operations only own a small fraction of the company’s stock, they are likely to run the firm inefficiently, or at least not in shareholders’ best interests. They then suggest an argument in favor of debt financing, which has later become the main justification for leveraged buyouts (LBOs): by financing a firm with debt, managers are able to retain concentrated equity ownership in the firm, and thus are incentivized to invest in future cash flows that enhance shareholder value. However—they continue—too much debt financing creates its own incentive problems, excess risk-taking, so that the optimal leverage ratio for the firm trades off the moral hazard benefits in terms of better investment incentives for managers against the excess risk-taking incentives for shareholders of highly levered firms. Excess risk-taking by highly levered firms, has, alas, become an all too familiar notion after the financial crisis of 2008.

Another familiar cost faced by highly indebted borrowers is debt overhang, a concept first formally analyzed by Myers (1977). Highly indebted borrowers facing the risk of financial distress will pass up valuable investment opportunities or even sell assets at fire sale prices. To reduce the risk of debt overhang it may thus be desirable for firms not to borrow too much. Also, should a firm end up with too much debt on its books it may be desirable for borrower and lender to renegotiate or ‘restructure’ some of this debt. This latter observation
provides the main economic justification for the existence of bankruptcy reorganization and debt resolution procedures.

Next to these theories of capital structure based on moral hazard problems, Myers and Majluf (1984) have proposed a powerful theory based on asymmetric information or adverse selection. They argue that the firm as an issuer of claims on the firm’s cash flow has to overcome investors’ suspicion that the firm may be trying to sell overpriced claims. This is easiest to do if the firm is able to issue safe debt (say, senior, short-term, collateralized, debt like repos), for issuer and investors alike ought to be able to easily value such a safe fixed-income instrument. If safe debt is unavailable the next easiest claim to value is risky debt, or possibly convertible debt, and the hardest claim to value may be equity. That is, equity may be the hardest claim to value as it may be the most information sensitive claim. Which is why, Myers and Majuf—under their pecking order theory—propose that equity claims should only be issued as a last resort and that firms should first use internally generated funds, then safe debt and then if needed risky debt.

These theories are an important advance over the simple tradeoff theory and considerably improve our understanding of corporate finance. However, it should come as no surprise that the pure agency theories of corporate finance encounter similar conceptual difficulties as the Principal-Agent theory à la Grossman-Hart (1983). Indeed, a central problem with agency theories of corporate finance, as Dybvig and Zender (1991) have emphasized, is that it is perfectly possible to separate the choice of capital structure of the firm from the problem of optimally incentivizing a firm’s manager by directly designing an optimal compensation contract for the CEO. Under an optimal incentive contract, which admittedly may be very complex and sensitive to the special circumstances a firm finds itself in, it is possible to address directly all relevant moral hazard and adverse selection issues, so that capital structure choice again becomes irrelevant or is determined by the tradeoff theory. In sum, the firm’s capital structure is relevant for managerial incentives, and therefore also for corporate control, only if incentive and financial contracting is limited by enforcement
constraints; that is, only if contracts are incomplete.

2.2 Incomplete Contracts and Corporate Control

Contingent Control Allocation

This last observation was the starting point of our analysis with Philippe in Aghion and Bolton (1992). Having learnt from Oliver and Sandy the relevance of incomplete contracts for the theory of the firm’s boundaries, we simply adopted their pragmatic approach to modeling incomplete financial contracts, and proceeded to develop a theory of capital structure choice which included the dimension of corporate control. The general idea sounds simple enough, but as we painfully learned, its execution was far from obvious.

We started by assuming that not all actions available to a manager and not all states of the world were describable ex ante in a contract or verifiable. Our thinking was that if some future action choices could not be specified in a contract, this would give rise to a problem of control, as some actions (and states of the world) left out of the initial contract would have to be determined ex-post. This would then raise the question: who is charged with taking these actions? In other words, who is in control? But, this was only the starting point, as the next question was why it was relevant at all who is in control? In our bilateral contracting problem with an entrepreneur and a financier allocations of control to one or the other party would have to lead to different outcomes for control to matter at all. Different outcomes presupposed different objectives, which could neither be fully aligned through ex ante contracting nor through ex post renegotiation.

Our first approach to this second question was to assume that the entrepreneur and financier had different beliefs about which investments were preferable. Based on casual observation, we assumed that the entrepreneur was generally more optimistic about the success of risky investments than the financier. This difference of opinions combined with the assumption of limited liability (and limited wealth of the entrepreneur) naturally gives rise to conflicting objectives, which cannot be fully aligned through contracting. It also gives
rise to a plausible contingent control allocation.

The entrepreneur seeks to keep control in the states of the world where the difference of opinions is largest; that is, where she is likely to be much more optimistic about success than the financier. And, she is willing to give up control in states of the world where differences of opinion are smaller. To the extent that differences of beliefs are likely to increase as the venture’s prospects improve, this contingent control allocation could be implemented through debt financing, whereby the financier gains control in the event of default and otherwise the entrepreneur retains control. It could also be implemented through staged transfers of control under a Venture Capital (VC) contract.

As simple and plausible as this solution seemed to us, the contract theory community at the time was not ready to accept two departures from orthodoxy in the same paper: incomplete contracts and differences of opinion. We received almost unanimous advice to change the model and do away with differences of opinion. So, instead of modeling differences in objectives arising from different beliefs, we modeled them as arising from the presence of private benefits: we assumed that the entrepreneur derives both financial returns and private benefits from the venture, while the investor derives only financial returns. In a way, this new model is more general, as differences of opinion can be mapped into financial returns and a particular form of ‘private benefits’, but vice-versa, it is not always possible to transform an objective function combining financial rewards plus private benefits into an objective function with no private benefits but differences of beliefs.

Our model delivers predictions on the separation of cash-flow and control rights that are consistent with common contractual clauses in VC contracts, as Kaplan and Stromberg’s (2003) study of Venture Capital contracts revealed. It also delivers predictions on cash inventory management under investor control, showing that it may be optimal for the entrepreneur to accumulate cash reserves that may be used to induce the investor to choose an investment with high private benefits for the entrepreneur. While these are valuable qualitative insights, the model remains in many ways too abstract and general to be an op-
erationally useful analytical tool. Part of the difficulty lies in the somewhat vague notion of private benefits. The other difficulty is that the enforcement limits on financial contracts are exogenously imposed in a somewhat arbitrary fashion. There is also the conceptual difficulty revealed by Maskin and Tirole (1999) that when an action or state of the world is observable to the contracting parties but not to a court (or judge) it can still be made verifiable through a suitable revelation mechanism.

**Limited Commitment**

For all these reasons it is not completely surprising that most of the subsequent literature on incomplete contracts in corporate finance has focused on narrower models in which private benefits are associated with some form of ‘stealing’ or cash-flow diversion. The pertinent image here is that of the cashier who is able to discreetly lift a few bills from the cash till. This literature also downplays the ‘observable but not verifiable’ distinction, and focuses on what is now generally referred to as a ‘limited commitment’ problem that the borrower faces due to her limited ability to commit to repay a loan.

In Bolton and Scharfstein (1990) we explore such a model of limited commitment, where the firm’s realized earnings are private information. To elicit truthful reporting of high realized earnings the financier must then offer a ‘carrot’ to the firm. In our model this carrot takes the form of allowing the firm access to new loans that are necessary to continue operating the business when the firm repays its old loans. Oliver Hart and John Moore (1994, 1998) also explore a model of limited commitment, in which, however, repayment of old loans is elicited with a ‘stick’: the threat of liquidating the firm if it does not repay its loan quickly enough.

Perhaps the main conceptual innovation of limited commitment models is a better understanding of the mechanics of debt default. Under the MM approach, default is assumed

---

1Another related approach to private benefits proposed by Holmstrom and Tirole (1997, 1998) is not directly related to stealing, but indirectly through shirking. By shirking the manager obtains private benefits in the form of reduced effort costs and ‘diverts’ (i.e. does not produce) the financial returns promised to investors.
to occur when the firm is *insolvent*; that is, when debt liabilities exceed the value of the firm’s assets (see e.g. Merton, 1974). This is a natural assumption if financial contracts are perfectly enforceable, for if the firm were to default when it is still solvent then debtholders could simply enforce payment by seizing the firm’s assets. In contrast, under limited commitment it is possible to separate default from insolvency and to distinguish between *liquidity* and *strategic* defaults.

The former is a situation where the firm is forced to default due to a cash shortage and the latter one where the firm chooses to default and force a debt restructuring because it is in its interest even though it is able to service the existing debt. Allowing for strategic default is a major conceptual breakthrough because it draws attention to an important practical aspect of debt design that is completely absent from the MM approach, namely the protections that creditors require in the form of seniority, collateral, security interests, and covenants. These are all protections that increase the likelihood that creditors will be repaid by reducing the probability of a strategic default and by increasing the creditors’ bargaining position in a future debt renegotiation. Under the MM approach and the tradeoff theory, all the firm’s debts should be junior unsecured debt. Indeed, these debts would give the firm all the tax shields it wants and would minimize bankruptcy costs, as these debts are relatively easy to restructure or *dilute* with new debt issues.

**Debt Structure**

In Bolton and Scharfstein (1996) we show how the threat of strategic default can be mitigated by having a well-protected, dispersed debt structure, which is difficult to renegotiate simply because it is more difficult to bring many people around a bargaining table. We also show that depending on the risk of a liquidity default, the firm may or may not want to structure its debt to make it difficult to restructure. If the risk of a liquidity default is high the firm may be better off with a debt structure that is easy to restructure, while if the risk of a liquidity default is low the firm may prefer to have debt that is hard to renegotiate.

In Bolton and Freixas (2000, 2006) we apply these ideas further and distinguish between
expensive (due to intermediation costs) bank relationship-lending, which is flexible and easy to restructure, and cheaper bond issues, which are, however, more difficult to restructure, and derive a partial equilibrium of the financial system with co-existence of a banking sector and securities markets. In this equilibrium riskier firms rely on bank lending as an important source of funding, because they value the flexibility it offers, while safer firms rely more on bond financing. This model lends itself, in particular, to an analysis of monetary policy through the ‘lending channel’.

In a series of related articles Diamond and Rajan (2000, 2001, and 2005) also build on the idea of debt dispersion to counteract a threat of strategic default, to develop a limited commitment theory of banking, bank fragility, and monetary policy. In the process they substantially upgrade the classic theory of banking as liquidity transformation of Diamond and Dybvig (1983): while in Diamond and Dybvig banks offer demand deposits to savers as a liquidity service, in Diamond and Rajan, banks offer demand deposits as a disciplining device, to facilitate ‘exit’ by disgruntled investors should the bank make bad loan decisions, or be a ‘weak’ debt collector. In their theory, bank fragility becomes a commitment device in a world of limited commitment, helping banks make more efficient loans than non-intermediated lenders who are vulnerable to ex-post strategic default and debt renegotiation.

Limited commitment models of debt point to the importance of debt renegotiation, a topic that has mostly been ignored by the MM approach. One could argue, of course, that as a first approximation this is a valid omission. However, recent empirical studies suggest that debt renegotiation is an important issue in practice. For example, Roberts and Sufi (2009) find that the vast majority of corporate long-term debt (over 90%) is renegotiated before maturity in response to changes in the firm’s environment. They also find that debt design reflects the parties anticipation of future renegotiation and attempts to allocate bargaining power on a state contingent basis. Similarly, Rauh and Sufi (2010) highlights that firms’ debt structures vary systematically with underlying firm risk, and that riskier firms have more complex and more collateralized debt structures.
Interestingly, a small, more recent, MM-based literature has incorporated elements of strategic default into their models. Thus, Anderson and Sundaresan (1996) consider a problem of debt valuation in a dynamic setting in which the borrower can strategically default. Not surprisingly, they find that both the default frontier and debt design are modified relative to the Merton (1974) model. Similarly, Mella-Barral and Perraudin (1997) analyze a continuous-time model of debt à la Leland (1994) with strategic default and find that the pricing of debt is significantly affected by the possibility of strategic debt renegotiation. Their model, in particular, provides more accurate estimates of credit spreads than other structural models without strategic default.

**Non-Exclusivity**

This first generation of debt structure models takes a comprehensive contracting approach to the design of debt structure: the number of creditors, and the seniority structure are optimally determined ex-ante in a multilateral (incomplete) contract with the borrower. The implicit premise in this literature is that the debt structures we observe are efficient from an ex-ante perspective. A more recent, second generation of debt structure models, relaxes the assumption of ex-ante comprehensive contracting and adds another dimension of contractual incompleteness, namely that debt structure is the equilibrium outcome of a debt contracting game with non-exclusivity.

The notion of non-exclusivity refers to the fact that a borrower may be able to borrow from a second set of creditors without the agreement from the first set of lenders. The analysis of equilibrium debt structures under non-exclusivity can be formulated as a common agency game with externalities (see Bernheim and Whinston, 1985, 1986a, 1986b and Segal, 1999). Not surprisingly, in the presence of externalities the equilibrium outcome of the contracting game will generally be inefficient. In the context of a corporate debt structure problem, when new debts are piled onto old debts, the expected payoff of old creditors is affected, as the new debts may increase the probability of default and reduce the recovery value of old debts in default. Since new creditors do not take account of this externality on
old creditors, there tends to be too much debt in an equilibrium with non-exclusivity (see e.g. Bizer and DeMarzo, 1992, for an early analysis of borrowing with non-exclusivity from multiple lenders). This is why non-exclusivity is a major concern for creditors, and to the extent possible creditors will attempt to protect themselves against future lending by the firm through various forms of debt covenants in the debt contract.

The efficiency of corporate debt structures and corporate borrowing thus depends to a large extent on the protection offered by debt covenants. The effectiveness of debt covenants, in turn, depends on how easy they are to enforce and how comprehensive an exclusion they provide. In short, the area where the issue of (endogenous) contractual incompleteness perhaps matters most, when it comes to debt contracts, is the design and enforcement of debt covenants.

As we argue in Ayotte and Bolton (2011), a critical distinction between property rights and contractual rights lies at the heart of the non-exclusivity problem. Following legal scholarship we define a property right as a right that is enforceable both against the parties to a contract and third parties (future potential lenders), while a contractual right is only enforceable against the parties to the contract. Property law limits which rights can be enforced against third parties. The property rights of creditors come mainly in the form of security rights on collateral that has been perfected (i.e. liens on assets that have been registered, and for which, therefore, third parties have been notified). All other debt covenants, whether they are negative pledge clauses, limitations on new investments, or acceleration clauses are only contractual rights against the borrower. In other words, they can only be enforced through legal actions against the borrower (e.g. through injunctions) and they have no force in bankruptcy against new lenders.

The implication is that debt covenants are costly to enforce because they require continuous monitoring of the borrower by the lender. As we argue, the reason why property rights law is structured in this way is to provide basic protections to new lenders against expropriation by old lenders. Indeed, debt covenants can be hard to find in a lengthy debt contract,
and if all covenants were enforceable against new lenders, these lenders would face potentially huge expropriation risk, which could lead to severe credit rationing in equilibrium.

Given that most covenants are costly to enforce, lenders concerned about non-exclusivity prefer to rely on the property rights offered by collateral and security interests. For many borrowers, however, such as financial firms, there are too few tangible assets that can be used as collateral. For these borrowers, covenants may also be too costly to enforce. As a result, these borrowers may be constrained in maintaining highly inefficient debt structures. These could take the form of excessively short-term debt, as Brunnermeier and Oehmke (2010) have argued in the context of corporate borrowing (and Bolton and Jeanne, 2009, in the context of sovereign borrowing). Note that according to this theory of debt structure, bank fragility arising from short-term liabilities may be an inefficient equilibrium outcome caused by non-exclusivity and not necessarily an optimal outcome to discipline bank lending, as in Diamond and Rajan (2000).

**Equity Structure**

Just as they open the way to a theory of debt structure, incomplete contracting and limited commitment models of corporate finance also provide a foundation for a theory of equity ownership structure. Thus, Fluck (1998) shows how equity can emerge as an open-ended claim receiving a regular dividend payment in a self-enforcing equilibrium in a limited commitment environment in which firm managers can divert cash. Similarly, Myers (2000), and more recently Lambrecht and Myers (2010), develop a theory of dividend payments as a way of pre-empting a hostile takeover. Incomplete contracting theories of equity structure can also be divided into efficient equity structure design theories and inefficient equilibrium equity structure theories.

Among the former theories, Admati, Pfleiderer and Zechner (1994) and Bolton and Von Thadden (1998) argue that concentrated ownership in the hands of a large block-holder may be an optimal ownership structure when monitoring of management is important. An alternative theory of limited controlling block size and optimal *managerial entrenchment*
(implemented e.g. through poison-pills) by Burkart, Gromb, and Panunzi (1997) and Pagano and Roell (1998) is that managers’ discretion needs to be protected to some extent to give them optimal incentives to originate new investment opportunities.

Among the latter theories, Bebchuk (1999) and Shleifer and Wolfenzon (2002) argue that ownership concentration arises as an inefficient equilibrium outcome driven by the blockholder’s inability to commit not to divert cash from the firm and desire to protect valuable private benefits of control.

3 An Assessment

As the above brief discussion of the corporate finance theory literature post Grossman and Hart (1986) suggests, the introduction of limited commitment into the MM framework has substantially enriched our understanding of corporate control, corporate debt structure, leverage, and equity structure. As a positive theory of corporate finance, the limited commitment theory offers many new predictions, some of which have been borne out in empirical studies.

As a normative theory, the limited commitment approach also offers a useful framework to assess legal interventions shaping equity and debt structures. However, normative analyses based on limited commitment models are often constrained by the lack of realism of these models with respect to the core assumption of non-verifiability of cash flows, investments, and states of nature. In reality, both investments and earnings are partially verifiable. Moreover, the contracting parties can spend resources to make them more verifiable. Also, the contractual incompleteness of debt contracts in practice is endogenous and is not simply the outcome of a technological or institutional constraint.

As a result, the contracts, financial structures, and legal rules observed in reality sometimes bear only a distant resemblance with the contracts derived in the theory. While a good theory inevitably leads to such simplifications and abstractions, these, of course, also
make a normative analysis more difficult. Still, relative to the MM framework of complete markets (with or without asymmetric information), the introduction of incomplete contracts and limited commitment has considerably enriched our understanding of corporate finance practice and corporate law.

The one major weakness of limited commitment and agency theories of corporate finance, however, is that they are not operational. Unlike the MM approach (and the tradeoff theory), agency and limited commitment models do not offer a methodology that practitioners can use. This is why thirty five years after the publication of Jensen and Meckling (1976), agency, information, and control issues still remain marginal and esoteric topics confined to advanced corporate finance classes. These issues are often treated more like an afterthought, something that is mentioned as a caveat following a systematic and thorough valuation exercise based on the MM approach.

As a result, agency issues are often ignored in practice simply because there is no simple quantitative methodology available to handle them. My view, therefore, is that before we pursue further refinements of the theory to put it on stronger foundations we need to make more progress on making the theory more operational, even if this means taking shortcuts. The structural models following Leland (1994) offer one direction, but they need to be augmented to introduce agency costs and limited commitment.

I have recently been involved in one effort in that direction in my work with Hui Chen and Neng Wang (Bolton, Chen and Wang, 2010, 2011). Basically what we do is build a corporate finance problem around a continuous-time, stochastic, version of the neoclassic q theory of investment à la Hayashi (1982) (which assumes MM neutrality) by adding a reduced-form cost of external financing. Granted, this cost could be derived from first principles in a model along the lines of DeMarzo, Fishman, He, and Wang (2010), but our point is that this would make the model essentially non-operational.

As is easy to see, an external cost of financing creates a role for corporate cash-balances and risk-management along the lines suggested by Froot, Scharstein and Stein (1993). In-
deed, in our dynamic model the critical state variable is the firm’s cash-to-capital ratio (a variable that is easy to construct and track from a firm’s balance sheets). When this ratio is very high the firm behaves like a financially unconstrained firm, and when it is low the firm engages in various forms of dynamic hedging, underinvests, possibly sells assets at fire-sale prices, and as a last resort raises costly external financing. There is basically no conceptual innovation in this model.

However, the model easily lends itself to a quantitative analysis, and by carefully calibrating the key parameters of the model (which can all be easily observed or estimated) one can provide concrete prescriptions to firms on how much they should invest, how they should manage their cash balances, how much they should engage in dynamic hedging, and how they should finance their investments. This is only a start and this is a highly simplified model. Still, it is a richer and more realistic model than the dynamic tradeoff-theory model, and it can provide a quantitative methodology that allows practitioners to take account of agency and limited commitment problems.

4 Conclusion

By introducing a way of modeling incomplete contracts and by proposing a simple theory of the firm based on the allocation of residual rights of control, Sandy Grossman’s and Oliver Hart’s 1986 article has opened the way for formal economic theory to address important issues that had almost exclusively been left to corporate law, management, accounting, and sociology of organization scholars. At the same time, their article has drawn attention to a largely neglected issue in economics, namely limits to contracting that arise from contract enforcement constraints (as opposed to asymmetric information and incentive constraints). By emphasizing contract enforcement constraints, their article has helped ground the more abstract and general economics of contracts literature in a more institutionally realistic context. Even though it is now 25 years since the publication of their article, this process
is still under way and far from complete. There is still too little communication, for my
taste, between legal scholars and economists. Still, by taking a bold pragmatic step and
introducing somewhat ad hoc (but plausible) constraints on contracting, Sandy Grossman
and Oliver Hart have profoundly changed the field of contract and institutional economics.
They have made it more relevant and rescued the field from “monstrous state-contingent
prescriptions”.

References

Sharing, and Financial Market Equilibrium,” Journal of Political Economy 102, 1097-
1130.


MacMillan Co.

NBER working paper No. 7203.


24


