

Incomplete Contracts and the Theory of the Firm¹

Sebastian Buhai*

-20th January 2003-

Introduction

The models within the theory of contracts can be divided into several sub-categories taking into consideration whether they describe a bilateral or a multilateral relationship, whether they are static or dynamic, whether they involve complete or incomplete contracts and so on. For our purpose, we shall focus below on the distinction between complete and incomplete contracts.

Assuming that contracts are complete is a very strong premise implying that all contingencies that may affect the contractual relationship are taken into account. Therefore while in an ideal world the parties would simply write down a binding contract specifying each agent's obligation in each conceivable eventuality, in reality such a contract is impossible to design. As Salanie (2000) points out, there might be several reasons why contracts would in fact be incomplete. In the first place, negotiating a contract in the real world is not void of costs; very often it mobilizes managers and lawyers. It comes by default then that at some point the cost of considering an improbable contingency outweighs the benefits of specifying it in a specific clause. Secondly, third parties are often unable or even unwilling to verify ex-post values taken by certain variables observed by the contracting parties. In other words the contracting parties, even when abstracting from the costs of negotiating, writing and legalizing the contract, are forced by their inherent bounded rationality to neglect some variables whose effect on the relationship they find difficult to evaluate. And finally, it is sometimes very cumbersome

¹ Research report partly based on material discussed by Patrick Bolton (Princeton University) in the Netherlands Network of Economics Workshop from December 2002, in Tilburg, the Netherlands

* Tinbergen Institute and Erasmus University Rotterdam, contact: TIA, Roetersstraat 31, 1018 WB Amsterdam; phone: +31 (0) 20 551 3561, email: buhai@tinbergen.nl

to assign a probability to some relevant events and thus to condition on such circumstances, which is again originating in our bounded rationality.

For all reasons mentioned above, contracts would typically consider a limited number of variables, either those that are the immediately relevant ones or those that are most easily verifiable by a court. That implies that many contingencies will have to be left out. If some of these omitted contingencies arise and they have an impact on the conditions of the relationship, it is clear that revisions and renegotiations will take place. In fact the contract per se is best seen as providing a suitable backdrop or starting point for such renegotiations rather than specifying the final outcome- Hart (1995).

Studying the incomplete contracts is still at a preliminary stage; different authors use different models which are often not comparable, all this when the general foundations of the theory have not been cleared out yet. Hart (1995) states that among the established theories, the transaction cost economics as created by Coase (1937) and Williamson (the survey in 1989 or previous papers) comes closest to our subject. And indeed as we will also see further in a short review of the classics, transaction cost economics builds on a world of opportunistic agents but deals with claims that are boundedly rational, hence contracts will be incomplete. However, albeit the emphasis given in the transaction cost literature to the costs of writing contracts and to the consequent incompleteness of contracts, less attention is paid to the idea that power or control are important, that institutional arrangements are in fact designed to allocate power among agents. We shall focus more on this aspect following a similar but non-identical sequence of ideas to the one presented in Bolton & Dewatripont (2002).

Contractual incompleteness and the theory of firms

A. Short review of the classics

In Hart (1995) the ex-post allocation of power following that contracts are incomplete is decisive. Oliver Hart does in fact argue that the contractual incompleteness and power can be used to understand a considerable number of economic institutions and arrangements, to shed light on some important economic phenomena. He reviews several implications of this approach, among them being the meaning of ownership, the boundaries of firms, financial securities, dispersed power. Unquestionably there is a considerable amount of specific research dedicated to each of these sectors and we will not make a purpose from trying to resume the current state of affairs therein. We shall rather sketch a brief overview of the main contributions to the link between the contractual incompleteness and the theory of the firms and afterwards analyze in more depth a simple model of property rights and residual control.

In an attempt to survey the development of the incomplete contracts models we need to take a trip back in time and perhaps start with the first formal model of incomplete contracts which is due to Simon (1951). Building on the internal-external transaction models of Ronald Coase², Simon (1951) compares two long-term contracts: a “sales contract”, in which the service to be provided by the seller is precisely specified, and an “employment relation contract”, where the service is left to the discretion of the buyer within some specified limits contained in the so-called “acceptance set”. The theory conceived by Herbert Simon did raise a lot of questions however, among the most relevant ones counting: “Why do the buyer and the seller have to agree to a trade before the uncertainty is realized?” or “Why can’t the buyer and seller agree to a contract specifying a state-contingent delivery plan rather than an “acceptance set”?”. There was obviously a need for further developing this approach and finding some reasonable answers.

The next author on this line is Oliver Williamson with his successive publications in 1975, 1979 and 1985. Williamson stays in the tradition of Simon, although by then a rival stream in the literature had been born (we shall tackle this approach latter in this section); he expands on Simon’s work adding transaction costs in markets as being mainly

² Most papers of Ronald Coase, including the 1937 one, were reprinted in Coase(1988)

bargaining costs; in particular he argues that investment in “relationship specific assets” increases the rents over which the parties may bargain. Williamson also introduces the bounded rationality concept justifying why one cannot specify everything in advance in a contract. Another worth to recall innovation are the negotiations procedures set up to reduce the parties’ incentives to pursue their self interest, the so-called “governance structures” in the work of Oliver Williamson. Given these governance structures, in essence a firm could provide “low-powered” incentives to the management reducing the intensity of any dispute. Exactly this easy to draw conclusion was the one raising further the question” Then what limits the size of a firm?”.

Klein, Crawford and Alchian (1978) continued Williamson’s ideas emphasizing the hold-up problem close to the relation specific investments and opportunism concepts pioneered by Williamson. Klein et al (1978) point out that the hold-up problem leads to underinvestment and arrive at the conclusion that the most appropriate institution to address this hold-up issue is the joint ownership. In this respect, they analyze at length one of the current most cited empirical examples used in incomplete contracting, the relationship between General Motors and Fisher Bodies—GM turns out to have been effectively held up by Fisher Bodies until 1926 when it decided to buy the car parts maker.

One cannot abandon this line of thought that started with Coase and Williamson’s intuition without mentioning Alfred Chandler. Chandler’s contribution materialized in 3 books in 1962, 1977 and 1990. Chandler started by investigating the rise of the modern corporation. He stipulated that decentralization in the so-called M-form or multidivisional form is the best organizational feature of a modern enterprise and an alternative to the overload of the top management. Chandler (1990) also stresses the importance of technology in general and of the returns to scale in particular as important determinants of the size of the firms.

Surely not everybody agreed with the ideas of Coase and Simon. Alchian and Demsetz (1972) protested that there is no meaningful distinction between internal and external

transactions and that the notion of authority is empty, being unclear for instance what is the source of an employer's authority over an employee. In other words there is no difference between a "sales" and "an employment relation" contract as seen by Simon (1951), an employer asking something from an employee and a customer ordering a grocer to supply different goods being equivalent notions. In this approach, the firm is designed to address incentive problems such as the moral hazard in the team of workers; while in general the incentive problem theory is not connected to the theory of the firm, Alchian and Demsetz (1972) make this connection stressing that the employer's role is to supervise the employees and making sure that they all work, in other words the employer as the "residual claimant" of the profits has the incentive to monitor the activity of his workers.

A sequel of the work in Alchian and Demsetz (1972) is the paper by Jensen and Meckling (1976). The authors further extend therein the theory developed by Alchian and Demsetz investigating for instance why is it that the manager is not always the sole owner of the firm. And here we are faced for the first time with the incomplete contracts and the portfolio choice of debt and equity the manager is faced with. The intuition is rather straightforward: since the manager has the highest incentives to maximize profits if he is a full residual claimant he will tend to raise funds in terms of debt and hold all the equity. Jensen and Meckling also expand in this respect on the arising conflict between the shareholders and the creditors, given that in their view the firm is a "nexus of contracts: between the manager and every other stakeholder.

B. A simple model of property rights and residual control

In an ideal world we would write complete contracts and the allocation of property rights would only matter for distributive purposes, without anything to say inasmuch as efficiency is concerned. Salanie (2000) gives the simplest example in this sense showing that the theory of complete contracts does not intersect in any way with the determination of the firm boundaries: if a manufacturer produces computers sold by a certain dealer, it

is socially indifferent whether the first firm buys the second or viceversa or whether the two firms remain legally distinct when a complete contract between them can be written. The boundaries of firms are nonetheless not random. In the Anglo-Saxon legal tradition the property rights are such that when an unforeseen contingency occurs, the owner has the right to decide how the good should be used. This has been labeled “residual rights of control” over the respective asset, terminology introduced in Grossman and Hart (1986). The owner also gets exclusive privileges on all possible income streams that have not been shared ex-ante in a contract. Certainly agreement with respect to this latter feature of ownership does not strike as guaranteed; a distinct segment of this literature, pioneered by Alchian and Demsetz (1972) and Jensen and Meckling (1976), define the owner as the residual claimant only of the cash-flow and not of all revenues generated by the owned asset, thus emphasizing even more the importance of the departure of the incomplete contracts approach from the standard incentive theory. We will however not be concerned with this second approach in the hereby paper.

We shall use as example to portray these concepts a very simple case with two separate firms or productive assets, where parties write incomplete contracts³. While often the classical instance used in this respect is the “Printer-Publisher” example drawn from Grossman and Hart (1986), we will take here another one, still very pertinent and straightforward to understand, introduced in Holmstrom and Tirole (1991):

Let us consider a computer manufacturer M and his dealer D that have agreed on an initial benchmark contract. One of the parties decides to make a specific investment; herein M decides to invest in R&D of an innovation that will increase the market value of its computers with an expected variable v that may take two values: 2 and 4. The innovation entails however a random cost c for the dealer, who would have to modify his product line and market routine, with c taking the values 1 or 3. We also have that M can influence the probability distribution of v by making a specific investment x that costs him x^2 but ensures that $P\{v=4\}=x$, where $P\{v=4\}$ denotes the probability that v takes the

³ A by now celebrated general model with more than 2 agents, with complementary investments, is discussed in Hart and Moore(1990), while the corresponding general framework using substitutable investment is in depth analyzed in Bolton and Whinston(1993)

value 4. On the same line, D can ensure that $P\{c=1\}=y$ by making a specific investment y that costs him y^2 . We assume that the four variables v , c , x , and y , once realized, are observed by both parties but are non-verifiable. Therefore there is no contract signed ex-ante. For a given allocation of property rights, the following game shall subsequently take place:

- i) M and D choose the values for x and respectively y , paying the associated costs.
- ii) M and D observe the values taken by v and c
- iii) M and D renegotiate to further decide whether or not the innovative computer is to be put on sale and, if so, how to share the surplus

One can notice immediately that out of the four possible cases: $\{v=2, c=3\}$, $\{v=2, c=1\}$, $\{v=4, c=3\}$ and $\{v=4, c=1\}$, in the latter 3 cases where $v > c$ the innovation will be sold. Only in the first case having probability $P\{v=2, c=3\}=(1-x)(1-y)$ we have an exception. We can calculate immediately the expected gross social surplus, that is,

$$S(x,y)=E\{\max(v-c,0) \mid x,y\}=0+(2-1)P\{v=2,c=1\}+(4-3)P\{v=4,c=3\}+(4-1)P\{v=4,c=1\}$$

$$S(x,y)=(1-x)y+x(1-y)+3xy$$

thus $S(x,y)=xy+x+y$

Maximizing in x and y the net expected social surplus W , that is the expected social surplus minus the associated costs, $W=S(x,y)-x^2-y^2$, having $S(x,y)$ calculated above, we immediately get $x=y=1$ and $W=1$. The social optimum level consists thus in making the maximal investments so as to obtain the most favorable case $\{v=4, c=1\}$ when the innovation is sold with probability 1.

What is much more interesting nevertheless is to consider now three separate cases and see what will be the outcomes in each of them:

- The *nonintegrated* case: M and D are distinct legal entities. We assume that when renegotiating M and D decide to share the gross expected social surplus equally, thus settling for the Nash bargaining solution. In this case M chooses her investment x as the maximizing value for $S(x,y)/2-x^2$ and similarly D chooses y to maximize $S(x,y)/2-y^2$.

Straightforward to notice, the Nash equilibrium gives $x=y=1/3$ and $W=5/9$. In conclusion in this non-integrated case here both parties *underinvest* (knowing that they can only get half of the revenues subsequent to their investments).

- The *downstream integration* case: M buys D . Of course in this case D will not invest at all since he has no rights on the income streams that the investment would generate and would solely bear costs if investing. Consequently $y=0$ and $c=3$ to start with. M only has to choose the level of her investment this time by finding the maximum in x to $S(x,0)-x^2$. The solution is $x=1/2$ and thus $W=1/4$. In comparison to the previous case, here M *overinvests* and D *underinvests*. We get a lower social surplus than before.

- The *upstream integration* case: M is bought by D . In this situation M does not have any incentive to invest resulting in the parameter values $x=0$ and $v=2$. D will obtain y by maximizing $S(0, y)-y^2$ whence $y=1/2$, resulting in $W=1/2$. As we saw, here M *underinvests* and D *overinvests*. The social surplus is higher in this case than in the second one, but still lower than in the non-integrated case.

Summing up, we have observed on the one hand that in the example analyzed the (net) social surplus is smaller if firms are integrated, which means that the costs of integration outweigh its benefits. Nonetheless, as shown for instance in Hart (1995) this conclusion is very sensitive to the exact specification of the model and in particular to the values of the parameters. In fact Hart(1995) derives optimal conditions for either case of the allocation of property rights, in function of whether the assets owned by the two firms are complementary or independent, whether the human capital in the two enterprises is essential or not and so on. For our purpose, the model just analyzed above simply shows that given incomplete contracts, some allocations of property rights can be more efficient than others⁴.

On the other hand, we have observed that for the example from Holmstrom and Tirole(1989) that we analyzed, we always had underinvestment and not efficient

⁴ As Hart(1995) also mentions there might be of course other reasons for integration or non-integration, apart efficiency considerations. For instance firms might integrate horizontally to raises prices to the consumers and they might integrate vertically to anticipate and prepare in front of rival purchasers and suppliers.

investment, whatever the allocation of property rights. This is a hot debate in the literature and penetrates even farther in the dispute over the foundations of contracting with unverifiable information. In this particular instance one view is promoting that when contracts are incomplete the parties will tend to underinvest in specific assets, as shown for instance in the papers by Grout(1984) or Hart and Moore(1988); the other view sustains that under certain assumptions there exist simple contracts that would implement the socially optimal levels of specific investment, with the papers by Chung (1991) or the more recent ones by Aghion, Dewatripont and Rey (1994) and Noldeke and Schmidt (1995) being the most cited. Without entering in too much detail, the Aghion et al(1994) paper shows that it is decisive to consider how the *status quo* is defined and what is the allocation of bargaining power at the renegotiating stage. The contractual status quo is the agreement that parties fall back on when they cannot agree on renegotiation terms; while in our example discussed above this agreement was given a priori, Aghion et al (1994) assume that it can be chosen arbitrarily. Another difference is that in our model the surplus generated was shared equally between the parties, while Aghion et al(1994) follow the standard assumption in the economics of contracts literature where one of the parties has all bargaining power and thus can get the whole surplus at the renegotiating stage. The debate is still on since the defenders of the “underinvestment” results argue that there are many cases when the characteristics of the good(s) to be traded are so many that it would be impossible to describe it previously and therefore to write such “specific performance contracts” that can specify a given level of trade that parties can require ex post, albeit it is efficient or not.

To render matters even more exciting, there is a relatively new stream in the literature that considers direct externalities, id est any direct effect of the buyer’s investment on the costs of the seller or viceversa, of the seller’s investment on the buyer’s valuation. Che and Hausch (1998) allow for this and obtain that *no initial contract* may be the optimal contract, which seems a sensational result. Still obtaining that ex-ante describability may not matter, but this time since the first-best can be achieved without specifying ex-ante any explicit message contingent-outcome, is the general result presented in Maskin and Tirole (1999). If renegotiation can be prevented, using complicated message games will

allow the parties to achieve exactly the same payoffs as when contracts are complete. In principle the essence of this reasoning is based on the fact that as long as all variables are obtained by all parties (and the literature so far has only considered incomplete contracts framework with symmetric information), their nonverifiability can be avoided in a way by playing revelation games taken from the literature on subgame-perfect equilibria. In particular, the hold-up problem can be solved in the Maskin and Tirole (1999) context without underinvestment ever arising.

Instead of conclusion

Several decades ago, in his 1937 seminal paper, Ronald Coase was asking the crucial questions :“ What is a firm? What determines the boundaries of the firm”? We have to admit that at this time we still do not have one definite answer. What we do have are several competing theories that attempt to elucidate the behavior of the market agents in practice. One of the best candidates, albeit few yet existing controversies within its framework, seems to consist in employing the methodology of incomplete contracts. In its very essence, this theory holds that economic institutions such as firms arise when and where people write incomplete contracts and where the consequent allocation of power and control is therefore important. What traditional approaches cannot explain, incomplete contracting or property rights theory pretends to have answers for. It appears throughout the literature that this perspective does have satisfying answers as far as the boundaries of the firms are concerned and can clarify to a certain extent the meaning of the asset ownership; it further looks as if contractual incompleteness does a successful job in understanding firms’ financial decisions, particularly as regards the nature of debt and equity; it also seems more than reasonable to model the capital structure decisions of companies using such an approach while the allocation of voting powers among an enterprise’s shareholders also finds a resolution using this framework. In other words, paraphrasing Bolton & Dewatripont (2002), the unified theory of ownership and financial structure of the firm based on the incomplete contracting methodology, has all the chances of becoming a classic itself sooner or later.

References

- Aghion, P., M. Dewatripont and P. Rey “Renegotiation Design with Unverifiable Information”, *Econometrica*, 62:257-82, 1994
- Alchian, A. and H. Demsetz “Production, Information Costs, and Economic Organization.”, *American Economic Review*, 62(5):777-95, 1972
- Bolton, P. and M. Dewatripont “Introduction to Contract Theory”, Part IV, in “Incomplete Contracts and the Theory of the Firm” by P. Bolton, lecture notes NAKE Workshop, Tilburg, 9-13 December 2002
- Bolton, P. and M. Whinston “Incomplete Contracts, Vertical Integration, and Supply Assurance”, *Review of Economic Studies*, 60:121-48, 1993
- Chandler, A. D. “Strategy and Structure”, New York: Doubleday and Company, 1962
- Chandler, A. D. “The Visible Hand: The Managerial Revolution in American Business”, Cambridge, Mass.: Harvard University Press, 1977
- Chandler, A. D. “Scale and Scope: The Dynamics of Industrial Capitalism”, Cambridge, Mass.: Harvard University Press, 1990
- Che, Y.-K. and D. Hausch, “Cooperative Investments and the Value of Contracting”, *American Economic Review*, 1998
- Chung, T. Y. “Incomplete Contracts, Specific Investments and Risk Sharing”, *Review of Economic Studies*, 58:1031-42, 1991
- Coase, R. “The Nature of the Firm”, *Economica*, 4, 386-405, 1937
- Coase, R. “The Firm, the Market, and the Law”, Chicago: University of Chicago Press, 1988
- Grossman, S. and O. Hart “The Cost and Benefits of Ownership: A Theory of Vertical and Lateral Integration”, *Journal of Political Economy*, 94: 691-719, 1986
- Grout, P. “Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach”, *Econometrica*, 52:449-60, 1984
- Hart, O. “Firms, Contracts and Financial Structure”, Clarendon Press, Oxford, 1995
- Hart, O. and J. Moore “Incomplete Contracts and Renegotiation”, *Econometrica*, 56:755-85, 1988

Hart, O. and J. Moore “Property Rights and the Nature of the Firm”, *Journal of Political Economy*, 98: 1119-58, 1990

Holmstrom, B. and J. Tirole “The theory of the Firm”, *Handbook of Industrial Organization*, vol.1, North-Holland, 1991

Jensen, M. and W. Meckling “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure”, *Journal of Financial Economics*, 1976

Klein, B., R. Crawford and A. Alchian “Vertical Integration, Appropriable Rents and the Competitive Contracting Process”, *Journal of Law and Economics*, 21:297-326

Maskin, E. and J. Tirole “Unforeseen Contingencies and Incomplete Contracts”, *Review of Economic Studies*, 66:83-114, 1999

Noldeke, G. and K. Schmidt “Option contracts and renegotiation: A solution to the hold-up problem”, *Rand Journal of Economics*, 26: 163-79, 1995

Salanie, B. “The Economics of Contracts”, The MIT Press, Cambridge, 2000

Simon, H. “A Formal Theory of the Employment Relationship”, *Econometrica*, 19: 293-305, 1951

Williamson, O. E. “Markets and Hierarchies: Analysis and Antitrust Implications”, New York, The Free Press, 1975

Williamson, O. E. “Transaction Cost Economics: The Governance of Contractual Relations”, *Journal of Law and Economics*, 22:233-61, 1979

Williamson, O. E. “The Economic Institutions of Capitalism”, New York: Free Press, 1985

Williamson, O. E. “Transaction cost economics”, *Handbook of Industrial Organization*, vol. 1. Amsterdam: North Holland, 1989