



European  
University  
Institute

Robert Schuman Centre for Advanced Studies

# **EUI Working Papers**

RSCAS 2007/16

Foundational Economic Theories for  
Political-Scientific Inter-Branch Studies

Yannis Karagiannis

**EUROPEAN UNIVERSITY INSTITUTE, FLORENCE**  
**ROBERT SCHUMAN CENTRE FOR ADVANCED STUDIES**

*Foundational Economic Theories for Political-Scientific Inter-Branch Studies*

**YANNIS KARAGIANNIS**

This text may be downloaded only for personal research purposes. Additional reproduction for other purposes, whether in hard copies or electronically, requires the consent of the author(s), editor(s).  
Requests should be addressed directly to the author(s).

If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper, or other series, the year and the publisher.

The author(s)/editor(s) should inform the Robert Schuman Centre for Advanced Studies at the EUI if the paper will be published elsewhere and also take responsibility for any consequential obligation(s).

ISSN 1028-3625

© 2007 Yannis Karagiannis

Printed in Italy  
European University Institute  
Badia Fiesolana  
I – 50014 San Domenico di Fiesole (FI)  
Italy  
<http://www.eui.eu/RSCAS/Publications/>  
<http://cadmus.eui.eu>

## **Robert Schuman Centre for Advanced Studies**

The Robert Schuman Centre for Advanced Studies (RSCAS), directed by Stefano Bartolini since September 2006, is home to a large post-doctoral programme. Created in 1992, it aims to develop inter-disciplinary and comparative research and to promote work on the major issues facing the process of integration and European society.

The Centre hosts major research programmes and projects, and a range of working groups and ad hoc initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration and the expanding membership of the European Union.

Details of this and the other research of the Centre can be found on:

<http://www.eui.eu/RSCAS/Research/>

Research publications take the form of Working Papers, Policy Papers, Distinguished Lectures and books. Most of these are also available on the RSCAS website:

<http://www.eui.eu/RSCAS/Publications/>

The EUI and the RSCAS are not responsible for the opinion expressed by the author(s).



## **Abstract**

Economic theories are increasingly popular in political science, and in particular in research on the relations between the legislative, the executive, and the judicial branches of government. Among these theories, principal-agent (‘PA’) and transaction cost economics (‘TCE’) feature particularly high in our research agenda. Yet, pushed by the view that “the content of ‘science’ is primarily the methods and rules” (King et al. 1994: 9), and working with limited resources, political scientists have tended to neglect careful theorizing. PA and TCE are taken off-the-shelf without much prior scrutiny, and past conceptual mistakes are perpetuated. This paper aims at introducing and explaining the real PA, positive agency, TCE, and incomplete contracts theories for the purposes of political analysis. In a companion paper, I show the serious mistakes perpetuated by political scientists, and I argue that, faced with a choice between those four economic theories, we should place our bets on a revised version of TCE.

## **Keywords**

Theory of delegation, political science, principal-agent models, transaction costs economics



## INTRODUCTION\*

Many political scientists are interested in inter-branch relations. I define inter-branch relations as the delegation, discretion, and control relations between the legislative and the executive branches of government. (This definition can be easily expanded to include the judiciary.) In that context, political scientists investigate the causes and consequences of delegation. Answers are sought by examining different empirical settings (such as the delegation of powers from the legislature to the executive branch of government within a state, or the delegation of powers from member states of an international organisation to the institutions of that organisation). Past failures to build an adequate theory of delegation and control have led political scientists to turn for inspiration to economic theories.

Ever since the mid-1980s, increasing numbers of political scientists attempt to explain delegation and control by referring to the economic theories of principal-agent and/or transaction cost economics. Within the more limited field of European studies (broadly defined), these theories have been used by, *inter alia*, Ballmann, Epstein & O'Halloran 2002, Bergman 2000, Blom-Hansen 2005, De Bièvre and Dür 2004, Doleys 2000, Egan 1998, Franchino 2000a, 2000b, 2002, 2004, 2007, Garrett 1992, Garrett and Weingast 1993, Gilardi 2001, Jun 2003, Kassim and Menon 2003, Lane 2002, Lane and Ersson 2003, Majone 1996a, 1996b, 1997b, 2001, Pierson 1996, Pollack 1997, 1998, 2002, 2003, 2006, Rasmussen 2005, Stone Sweet and Caporaso 1998, Strøm 2000, Tallberg 2000, 2002, Thatcher and Stone Sweet 2002, and Tsebelis and Garrett 2001.

Beyond European studies, courses and workshops on the relationship between these or similar research questions and these economic theories abound (e.g. RIAS, Harvard University, December 2002; WCIA, Harvard University, April 2003; APSA Conference, Philadelphia, August 2003; U.C. San Diego, September 2003; Birkbeck College, London, March 2004; European University Institute, Florence, May 2006; Duke University, March 2006). Americanist studies that refer to these theories run into the hundreds (Epstein & O'Halloran 1999: 28; Huber and Shipan 2006).

It may therefore be surprising to find that, as such, these economic theories do not say much about delegation and control. Three points need to be stressed from the outset.

- First, neither transaction cost economics ('TCE') nor principal-agent ('PA') theories can automatically generate hypotheses regarding the delegation of powers by politicians to a policy-making bureau. TCE can be adapted to examine whether a bureau will be integrated into the political system or not, but not whether there should be any delegation at all.<sup>1</sup> Similarly, PA theory can be used to characterise the set of feasible allocations of resources between politicians and the bureau given the participation, incentive compatibility, and coalition incentive compatibility constraints of bureaucrats – but not to determine whether politicians should rely on an in-house or an external bureau.
- Second, PA theories do not generate hypotheses regarding the possibility of *ex post* shirking or drifting by the supposed agent. PA theories can be used to describe how the principal should optimise the contract he offers to the agent (i.e. the allocation of resources), but they do not

---

\* I acknowledge the financial support of the European Commission under its Marie Curie Fellowships Programme, the hospitality of the Centre for Regulatory Governance, University of Exeter, and the Department of Political and Social Sciences, Universitat Pompeu Fabra. I thank Adrienne Héritier for her help on political-scientific points, Elia Marzal for her help on legal points, and Pascal Courty and Oliver Williamson for their help on specific economic points. Two anonymous referees provided extensive and useful comments, some of which go beyond my sphere of (in-)competence. Any remaining errors are entirely my own.

1 As will become apparent below, integration and delegation are different concepts, not different values of the same concept. For example, delegation occurs both in integrated and in non-integrated settings – unlike integration or non-integration, it is 'inevitable' (Huber and Shipan 2006: 256; see also McCubbins, Noll and Weingast 1987: 243).



focus on ex post elements that the principal cannot take into account ex ante. Referring to ex post maladaptations in the context of a PA analysis is abusive and misleading. (On the other hand, such ex post issues are captured by TCE.)

- And third, these two theories (PA and TCE) are neither the same nor complementary: they are based on radically different assumptions. One of these concerns the cognitive abilities of actors, and hence their resources. Where PA rests on the ‘classical’ behavioural assumption of perfect rationality, TCE rests on semi-strongly bounded rationality (Williamson 1996: 8-9). This, and other, different assumptions have far-reaching implications, which have often been overlooked in the literature.

Documentation of these claims is the topic of the present paper.

Thus, the theoretical foundations of the hypotheses advanced in the political-scientific literature, including the literature on the EU, are either ambiguous (where they are based on implicit extensions of the theories, or when they confuse the positive theory of agency with principal-agent theory) or plainly wrong (where they mix these theories, alone or with others, in an ad hoc fashion).<sup>2</sup> One primary aim of the present paper is to demonstrate why this is so, and why political-scientific models need to more carefully acknowledge the theoretical foundations of their hypotheses. That is, I will argue that the process of crudely translating economic theories into political science has come at a cost, one that affects both the internal consistency and the logical completeness of these theories.<sup>3</sup>

More specifically, I offer an extensive explication of the development, rationale, and assumptions of the theories of PA, positive agency (‘PTA’), TCE, and incomplete contracts theory (‘ICT’) – that is, four economic theories that political scientists have too long crudely lumped together under the heading ‘New Economics of Organization’ (Moe 1984). This explication may at times not make easy reading.<sup>4</sup> Yet, it is a necessary step aimed at (a) correcting various misunderstandings in the existing literature (especially in political science, and in Europeanist studies therein); (b) building more confidence in our models of inter-branch relations; and (c) clarifying the road for future research in these areas.<sup>5</sup>

---

2 I have encountered some difficulty in explaining this point to political scientists who are engaged in policy research and who use a ‘principal agent’ perspective. Most of my interlocutors did not perceive the combination of different theories as problematic, as long as theories generate falsifiable hypotheses. The first section in this paper is a digression aimed at explaining my views on this issue: theories and models have to be logically complete and internally consistent; ad hoc introductions of inconsistent hypotheses create both theoretical and methodological problems (e.g. an empirical observation cannot anymore be explained by the value of the independent variable, since it may also be explained by the underlying assumptions).

3 Internal consistency is achieved when the statements that compose a theory are mutually consistent. Logical completeness is achieved when the hypotheses deduced from a theory follow logically from the assumptions of a theory.

4 I believe this to be due not only to my own limitations (see the next footnote), but also to the complexity of this field in economics. Indeed, many economists do not seem to be themselves aware of all the differences between these theories, while the field is progressing very quickly. This is presumably the reason for all the reviewing, categorising and classifying work being done recently in this field (e.g. Williamson 1996, 1999; Malin and Martimort 2000; Chiappori and Salante 2000; Whinston 2001; Brousseau and Glachant 2002; Laffont and Martimort 2002; Martimort – forthcoming in the New Palgrave).

5 Of course, there are other reasons why I offer this detailed literature review here. They can be understood using a positive agency framework. You (the reader) represent for me a ‘societal principal’: you have financed my research (which I do not wish to stop in the future) and in return you have asked me to provide you with some knowledge on a specific topic. This is our implicit contract. Note, however, that you do not know as much as I know, either in terms of this topic, or in terms of the effort I have put in acquiring and producing this knowledge. What you need to do is offer me the right mix of incentives and punishments. What I need to do is (a) gain some assurance on your behalf that I will not bear the total risk of a potential final failure, thus (b) ensure that you take into account all of my effort, and thus (c) incur some bonding costs that allow me to proceed in tranquility. For me, providing you with this literature review is a bonding cost (which I can measure in terms of hours spent trying to understand, summarize and present complicated economic ideas and their relationship to political scientific works, and in terms of space dedicated to non-competition policy issues, which is my empirical field of research). For you, it is an assurance that I am not providing you with a poor-quality product, or with a

The starting point needs to be that the four aforementioned economic theories are neither identical, nor strictly linked in any way. References to one that appear in analyses based on the other are not only confusing, but, as far as they implicitly introduce different assumptions, plainly erroneous and misleading. This specification is necessary because the currently prevailing (confused and confusing) view seems to be that at least two of these theories (PA and TCE) are identical, or that one of them is only a special case of the other. For example, McCubbins, Noll and Weingast (1987: 247) wrote that *'monitoring and enforcement are not costless'*, thus hinting to the existence of transaction costs. In the next paragraph, however, they added that *'the problem of bureaucratic compliance has long been recognized as a principal-agent problem'*, thus mixing principal-agent with transaction costs. (Actually, Mcnollgast mixed more than these two theories, and thus more than a couple of inconsistent assumptions. They thus perpetuated the 'original sin' of Moe 1984.)

Similarly, Pollack writes of *'principal-agent models of delegation of powers'* (2003: 5) and ascertains that *'principal-agent models are quite successful in predicting the functions delegated to the Commission and the Court of Justice...'* (ibid. 9). Yet, as I will document below, PA theory is not able to generate such predictions regarding delegation of powers. In the absence of additional work in view of adapting these theories to our research questions, the claim that either PA, or PTA, or TCE, or ICT, can generate hypotheses regarding the delegation of powers to a public bureaucracy makes no sense.

In the same spirit, Epstein and O'Halloran make a long way in elucidating the meaning of TCE concepts for political science, and even point to the fact that PA cannot do the job of TCE (1999: 39). Yet, perhaps contradicting themselves, they also use PA as a *'building block'* for their political TC theory (ib. 27-9). This contradiction is furthered when they note that *'just as in economic situations, [transactions among political actors] may not take place due to the existence of transaction costs, such as ... principal-agent losses'* (ib. 44), and *'the problems of delegation stem mainly from Congress' principal-agent problems of oversight and control, which we describe as a political hold-up problem'* (ib. 49). These views are not only contradictory in themselves, but also in stark contradiction with the economic literature that they are meant to reflect. For example, Williamson has argued that *'ex post transaction costs are related to, but plainly differ from ... agency costs'* (1985: 21; see also Williamson 1996: 176; Jensen 1983; Eisenhardt 1989).

PA, PTA, TCE, and ICT were developed by economists studying the relationship between two contracting parties who seek to establish, or who already are, in a collaborative (i.e. risk-sharing) relationship.<sup>6</sup> TCE has been used to explain why actors choose to collaborate within a business firm,

(Contd.) \_\_\_\_\_

product that already exists in the market. (For the rent-efficiency trade-off that you are now facing, see the main text below. In any case, please do not skip this literature review. It is technical, but it bonds us. If you were an airline and I were an aircraft producer, would you prefer me copying a competitor's aircraft and thereby free-riding on his R&D investment, or would you rather make sure that I myself understand all the relevant bits of knowledge in jet technology?)

6 There is a second view, according to which such relationships were first theorised not by economists, but by social and political theorists. This historical argument runs as follows: PA theory is a theory of collaboration, or risk-sharing; risk-sharing is based on implicit or explicit contracts; the theoretical bases for the analysis of such contracts were thrown by social theorists, rather than by economists. Exponents of this view therefore argue that the contractarian model of politics stems from social theory and political science rather than from economics. For example, Buchanan (1986: 240 ff) traces that model to Hobbes' discussion of individuals who face 'total war' and hence contract with the sovereign, thereby losing part of their liberty and (possibly) democratic rights, but gaining order and security. He adds: *'The Hobbesian metaphor suggests ... that so long as the sovereign remains within the agreed and assigned limits of the initial contract ... 'democratic' attributes of the sovereign's decision-making would be out of place and, indeed, would be counter-productive.'* Similarly, Immergut (1998) has argued that the rational choice variant of the new institutionalism is traceable to Rousseau's theory of the influence of societal norms and institutions on individuals' preferences.

In my understanding, these views are correct, in the sense that they demonstrate the constant preoccupation of political thinkers with the costs of using the 'democracy infrastructure' or the dangers inherent in cooperative ventures (witness Rousseau's stag hunt game). However, it is also necessary to note that these old political theories were neither the older relevant theories in this area (see, for example, Bodin's *Six Livres sur la République*, or even Plato's speculations about the education of the Guardians, in the *Republic*), nor explicitly founded on the same conceptualisation of transaction costs as the modern economic theories.

or conversely, why actors who might benefit from such mutual exchange may actually not do so. Examples include many ‘non-standard’ business practices, such as vertical integration, or mergers. In such situations, the central question asked is: ‘*What purposes are served by supplanting classical market exchange ... by more complex forms of contracting (including nonmarket modes of economic organization)?*’ (Williamson 1985: 23) In other words, the central question raised by TCE concerns the existence and the boundaries of the firm. Hence, TCE aims at explaining different institutional outcomes – it is a theory of institutions.

PA, on the other hand, has been used to analyze the relationship between actors who seek to collaborate, and thereby develop a mutually beneficial short- or long-term trading relationship. Such actors must therefore define the terms of their collaboration, not only in terms of prices and quantities, but also in terms of their respective identities, the effort they put in the production of the goods they trade, and the ways in which they communicate. Examples include employers and employees, stockholders and chief executives, etc. In such situations, the central question asked is: What is the optimal structure of the contract that the principal can propose to the agent, given that the existence of an agent imposes a certain number of constraints upon the principal? In other words, the central question raised in PA concerns the optimization problem that the principal faces, given certain exogenously defined constraints. Hence, PA aims at explaining the effects of different institutions on the optimal contract – it is an institutionalist theory, but not a theory of institutions. (Note also that it would be wrong to equate PA with any strategic effort to establish co-operation.)

Having exhausted the two hands (TCE on the one hand, PA on the other), we can turn to one foot: PTA. The foot metaphor is useful, because that theory is actually a footnote to PA theory [<sup>7</sup>]: it is less of a theory and more of a reminding to PA analysts that the real-world is characterized by the ubiquity of agency costs (PA theory is a theory of perfect rationality and complete contracts). According to Eisenhardt (1989), PTA serves to test two hypotheses: (a) outcome-based compensations are more efficient than ex ante lump-sum payments because they limit agency costs; and (b) agency losses (defined as firm valuations plus monitoring costs) are proportional to information asymmetries. In what follows, I take a more detailed and nuanced view, but I agree with the implication of Eisenhardt’s presentation: this theory is not a theory that can adequately address the research questions asked by political scientists. Like with PA and TCE, considerable adaptations are needed in order to apply this view to political phenomena.

## 1. THE ROLE OF THEORY AND CAREFUL THEORISING

Having mentioned the confusion that characterises political-scientific works that use these theories, a trade-off may arise between the goals of (a) substantive knowledge accumulation at the level of specific hypotheses (empirical knowledge), and (b) rigor and progress at the level of theories (theoretical advancement). Aiming for lower-level empirical knowledge alone means adopting existing hypotheses and testing them in a new empirical setting – hopefully a setting that is methodologically interesting. However, it also means replicating past theoretical mistakes, and not being able to proceed scientifically (e.g. not being able to proceed by describing and solving a game, which then generates testable hypotheses). On the other hand, aiming for higher-level theoretical advancement means correcting past mistakes, but at the same time rejecting many existing hypotheses and insights about the motivations of political actors – and hence losing sight of the existing literature and of currently dominant explanations of institutions. The way that I propose to proceed, then, is by delimiting the scope of the economic theories of PA, PTA, TCE, and ICT (this paper), and then

---

7 I use the word ‘footnote’ in the same way as Alfred Whitehead in his famous dictum that ‘[t]he safest general characterization of the European philosophical tradition is that it consists of a series of footnotes to Plato.’ PTA (like many other branches of economics), is a footnote to Kenneth Arrow. Obviously, footnote writers can carry names such as Aristotle, Descartes, Locke, Hume and Kant. (Note also the end of the preceding paragraph in Whitehead’s text: ‘ultimately nothing rests on authority; the final court of appeal is intrinsic reasonableness.’)

examining how these theories have been (inaccurately or mistakenly) used in political science (accompanying paper).

The remaining of this section is dedicated to a reminding of (a) the role of theory in political science, and in the social sciences in general; and (b) the criteria for the selection of one theory among many competing ones. These points are raised here because of two reasons. First, faced with the aforementioned problems posed by theoretical inquiries, readers who are interested only in the specifics of a certain policy may be tempted by the idea of rejecting theoretical approaches altogether. Yet, theoretical difficulties should not discourage scientific inquiry: they do not warrant a return to a-theoretical, case-specific and non-generalisable knowledge. Second, the market for theoretical ideas in political science is far from perfectly competitive. Bottlenecks, sunk costs, increasing returns to scale and important transaction costs combine to make the search for the best theory to apply in specific empirical researches prohibitively expensive. This has resulted in path-dependent adoptions of the same theories, and, by the same token, path-dependent perpetuation and aggravation of the same theoretical mistakes.

It is useful to begin the discussion on causal theories by returning to specific research questions. As mentioned above, the ones I have in mind here concern (a) the reasons why legislators delegate policy-making powers to bureaucrats; and (b) the consequences of such acts of delegation in terms of the subsequent relationship between legislators and bureaucrats. These two research questions can be asked at higher or lower levels of abstraction and generality. For example, at a lower level of abstraction we may be interested in the reasons why Member States delegate powers to the European Commission, or in the subsequent effects of such delegation in terms of the politics of the EU. At a more general level (but with greater loss of information), we may be interested in the reasons why sovereign states delegate powers to international agents, or to the subsequent effects of such delegation in terms of the politics of the corresponding international organisation.

When going from a lower to a higher level of abstraction, specific cases become decreasingly well-defined, and thus decreasingly differentiated. In other words, when moving towards higher levels of abstraction, increasing amounts of information are lost. For this reason, the theory that is chosen to answer the research questions needs to be capable of covering many different cases, while at the same time being capable of differentiating among such cases. Thus, a theory needs to be both as general as possible, and able to sustain hypotheses that are as specific as possible. This point can be illustrated by reference to an example, such as my own first research question. This concerns the causes of delegation of competition powers in the EU to the European Commission. What is needed is not only a theory that can explain delegation generally, but also a theory that can differentiate between the specific values taken by 'delegation' in this specific context (i.e. delegation by sovereign states to an international executive, delegation of one or more functions but not of other functions, delegation to a multi-agent organization and not to a single-purpose organization, etc). Conversely, a theory that can only explain the value taken by 'delegation' in only one specific case (one observation) and not in other cases, is a poor theory. Thus, if the first goal is to learn something about the politics of EU competition policy (which is itself a more general goal than, say, trying to learn something about the decision-making process followed within the Commission in a specific antitrust case), the research needs to be firmly based on a specific and explicitly stated theory of more general relevance.

This problem, I believe, lies at the root of many inaccuracies in the existing literature: driven by the desire to refer to theories that are as generally applicable as possible, some analyses have (a) lost the ability to explain the specific cases that they investigate, and (b) have mixed many different theories, without acknowledging the corresponding mix of different assumptions.

Yet, theoretically-founded research is always better than a-theoretical research. Indeed, a theory serves to (1) focus the research on those elements that are thought to be significant, and eliminate those elements that are thought to be less significant; (2) explicate and assess the plausibility of assumptions about elements that we do not directly observe; (3) describe the mechanisms that render

the links between different observations plausible; (4) derive testable hypotheses; (5) locate the most interesting explanatory factors, and thus enhance comparison; and (6) build an identifiable body of cumulative research. In the absence of a theoretical referent (which, to take again an example from my own research, is admittedly a very frequent situation in the legal and political literature on EU competition policy), an analysis cannot claim anything more than what is directly observed. This means that a-theoretical analyses cannot make any truly causal claims, and that their findings cannot be generalised to any other case than the one that they describe. In other words, in the absence of a theory, even ‘similar’ cases are not similar ... unless a theory tells us they are. Of course, a counter-argument against explicit theorising could be that similar cases are ‘obviously’ similar. Yet, the answer to this counter-argument is that ‘obvious’ depends on theoretical assumptions, and that it is always an improvement to make these explicit.

But, *the need* for a theory does not answer the question of *which* theory should be chosen. This question is particularly important because, by definition, when choosing a theory we do not yet know how well that theory fits with the empirical data. Given that there are several social-scientific theories to which one could refer in order to explain political, social, economic and legal issues, choosing *the best* theory would not only involve impossible amounts of theory-comparing work, but it would also equate the beginning and the end of research. Additional choice criteria are therefore needed.

Commonly used criteria for the selection of a theory include (a) fit with the research questions, (b) internal consistency, (c) logical completeness, (d) falsifiability, (e) parsimony, (f) plausibility, and (g) current development in conceptual accuracy and empirical testing. However, these criteria raise two problems. First, their use presupposes an accurate *ex ante* knowledge of different theories. Such knowledge is rarely available to researchers in the real world: it is more usually gained while working with a theory, and its acquisition is often perceived by empirical researchers (and disciplines) as counter-productive. Second, although a theory is supposed to increase the scientific content and method of research, these criteria themselves are not entirely scientific. For example, parsimony, plausibility and current development are only partly scientific criteria.

Whatever the magnitude of these problems, however, they do not justify a rejection of either theory in general, or of conscious attempts to justify one’s choice in favour of a particular theory. One way to deal with this problem is to be aware of its existence and to clearly state the theoretical assumptions of a research, so that subsequent work can determine whether the best choice was indeed made. The starting points are therefore that (a) any research is implicitly or explicitly founded on a number of assumptions regarding the phenomena that it examines, and (b) making these assumptions explicit is better than keeping them implicit.

According to King, Keohane and Verba, ‘*a model is a simplification of, and approximation to, some aspect of the world*’ (1994: 49). But, again, how should one decide which simplifications and approximations are appropriate, and which are not? (We sometimes encounter speakers who, when asked about how they chose their theoretical framework, reply with an ‘authority argument’, like: ‘X and Y did it, so that is why I do it, too.’ My goal here is to think about ways to go beyond that answer ... which I have sometimes given myself.) What are the relevant aspects of the political reality that models of delegation, agency, and EU policy-making should capture, and what are the irrelevant ones that can be excluded? I posit that the appeal of the economic theories of PA, PTA, TCE, and ICT consists of their demonstrable qualities of logical completeness, internal consistency, and (perhaps) falsifiability. What remains to be seen, is which theory is more *relevant* for our science of inter-branch relations.

Yet, one of our main scientific motivations must be to find ways to further improve the accuracy of existing theories. In Europeanist studies, for example, we already possess a considerable syllabus of PA and TCE works – the best of which include (by virtue of either originality or/and sophistication or/and publishing success) Franchino 2000a, 2000b, 2004, 2007; Majone 1996a, 1996b, 2001; and Pollack 1997, 2002, 2003, 2006. Although these authors have done an invaluable job in renewing the

study of the EU by translating Americanist models, they have perpetuated and/or aggravated the inaccuracies of the Americanist models, which were themselves translations of economic models into political science. One of the main arguments in this paper is this: some important economic insights were lost in translation. We should stop just getting ‘inspiration’ from the economic literature – this leads us to mixing up different theories that are based on different implicit assumptions. Rather, we should try to understand these theories, in all their complexity – including in their differences. In other words, if economists had delegated to political scientists the task of applying economic theories to political analysis, then political scientists have not fulfilled their contractual obligations to the benefit of either discipline.

Although I believe that economic theories carry a great promise for political-scientific analysis (and especially for inter-branch relations), I also concur with Jean Tirole’s argument that the economic theories of transaction costs and agency are not readily applicable to political science (Tirole 1986). After all, the two disciplines continue to be taught separately, and this is probably due more to the fact that the object of their studies somehow differs, than to some kind of market rigidity that does not allow them to merge (for the case of TCE, see Karagiannis 2007; for a different view based on a more general perspective, see Ordeshook 1990). In addition, several economic concepts, such as ‘markets’, ‘transactions’, ‘gains from exchange’, ‘efficiency’, ‘hierarchy’, ‘incentives’, etc, need to be translated into politically relevant and significant concepts. Thus, in order to use economic theories correctly, we need to adapt them to our specific political-scientific objects of inquiry. To put it boldly, if political scientists delegate to economists the task of analysing political phenomena using their existing models, economists will almost certainly not fulfil their contractual obligations successfully either.<sup>8</sup>

## **2. TRANSACTION COSTS AND PRINCIPAL-AGENT: DIFFERENT ORIGINS, DIFFERENT FOCUS, DIFFERENT NATURE.**

In this section I present the different origins of the economic theories of transaction costs (‘TCE’), principal-agent (‘PA’), and positive theory of agency (‘PTA’). TCE is rooted in the work of Coase (1937), while PA and PTA are rooted in the work of Berle and Means (1932). Later, PA theory became known as normative principal-agent theory, and generated the theories of contracts, incentives and implementation. (The latter does *not* correspond to implementation studies in policy studies, though it could usefully inform it). As the next section shows, these different origins are still relevant, in the sense that they still produce important differences, and they still generate heated debates between economists. Yet, political scientists have often confused these three theories. They have sought to add theoretical rigor to their works, but they have relied only on an intuitive understanding of what have become pop-theories. A more correct application can only be based on their thorough understanding – including a thorough understanding of their differences and their different origins.

The origins of TCE [and, more generally, of the new institutional economics (‘NIE’)] can be traced back to Roland Coase’s famous 1937 article ‘The Theory of the Firm.’ What troubled Coase in his efforts to understand the post-Depression economy was that, whereas mainstream economics assumed that competition, acting through a system of prices, would do all the necessary co-ordination to stabilise the economy, there actually *was* a factor of production whose function was to co-ordinate: management. The co-ordinating function of management had been unjustifiably neglected by neo-classical economists who focused exclusively on the (competing) co-ordinating function of the market

---

8 Yet another way to express the same idea is that political economy is not a bit of politics and a bit of economics, but economics and politics. If it were the former, the analyst might be able to pick and choose the doses according to convenience. Economics-inclined analysts would be heavier on formalization, and politics-inclined analysts would be heavier on law, institutions, and history. Political economy, as I understand it, is best represented by Bates et al. (1998; but see Elster 2000) and Boix (2003).

mechanism. Yet, the existence of management hinted to the fact that it was actually an economising device. It allowed the firm not only to reduce expenses, but, more fundamentally, to exist. In its absence, there would be no firms, but just a myriad of inter-personal market exchanges. (We shall see below that this is precisely the view of PA, PTA, and ICT.) From these observations, Coase came to the following conclusion: *'The main reason why it is profitable to establish a firm would seem to be that there is a cost in using the price mechanism. The most obvious cost of organising production through the price mechanism is that of discovering what the relevant prices are.'* (1937: 390)

Coase's famous insight means that using the market is costly, and therefore that economic agents who are given the chance to do so will organise production in view of minimising market-generated transaction costs ('TCs'). Note in the quotation from Coase that *'discovering ... relevant prices'* is not the *only* TC, but just *'the most obvious.'* Indeed, other TCs include negotiating and drafting contracts, executing simultaneously myriads of separate but inter-related contracts, monitoring deliveries and quality, settling disputes, paying brokers' commissions, paying taxes and duties, etc. (see also Coase, 1994; 7-8) More generally then, TCs fall into five broad categories:

- (a) search and information costs;
- (b) bargaining costs;
- (c) costs of executing multiple contracts;
- (d) monitoring and enforcement costs; and
- (e) public policy incentives to avoid the market.<sup>9</sup>

Obviously then, TCs are not just ubiquitous in the market: they are an inherent characteristic of the market. They are 'the costs of running the economic system' (Arrow 1970: 48), 'the economic equivalent of friction in physical systems' (Williamson 1985: 19). (Note that any theory that does not acknowledge the presence of such TCs may be doing so either because it relies on relevant empirical findings – in which case it may still be a TC theory – or because it excludes TCs by assumption – in which case it is not a TC theory.) It follows that, the more market transactions an economic actor performs, the more his TCs will increase. Hence, whenever possible, rational profit-maximising actors will seek to minimise these costs by avoiding the market, subject only to the constraints created by their own managerial resources (Coase's 'decreasing returns to the entrepreneur function'). They will thus first try to identify all potentially significant TCs, compare these costs of relying on the market with the costs of producing themselves the goods they need, and then choose the least costly option. In short, if using the market is costly, rational economic actors will not do so if there is a less costly alternative.

That 1937 article raised all the fundamental questions that came to define the research agenda of economic theorists working in the fields of transaction costs, the theory of the firm and organizational theory – but not, as we shall see, principal-agent models. Three main research questions were posed: (1) why does a firm emerge at all in a specialized exchange economy? (i.e. the question of the existence of the firm); (2) what are the forces which determine the size of the firm? (i.e. the question

---

9 Coase also referred to an additional category of TCs, that of protecting trade secrets. When a business (firm X) relies on the market, it relies on external, contracted agents (A). By collaborating with A, X reveals to them (some of) its trade secrets, i.e. its instruments for obtaining an advantage over its competitors (firms Y, Z). Yet, because A are not hired (i.e. they are not part of firm X), they may subsequently be contracted by Y or Z, to whom they may supply these trade secrets. X is thus at a disadvantage relatively to Y or Z, which may aim at free riding on the knowledge accumulated by X. That disadvantage and the costs of limiting it, are due to the existence of a market. X therefore has an incentive to hire A on a permanent basis, and therefore to destroy the market for A.

Protecting trade secrets is a very interesting source of TCs. Yet, its use in political science is unclear. In many political circumstances, an actor may actually wish the others to operate in exactly the same way as he does, rather than the opposite. Important exceptions include security policies, party organizations, trade instruments, etc. It might be productive to study the extent of such TCs in EU politics.

of the boundaries of the firm); and (3) when do ‘diminishing returns to management’ occur? (i.e. the internal organization of the firm).

Despite its insightfulness, Coase’s theory took a long time to influence mainstream economics (Coase 1972). By and large, the competitive market continued to be viewed as a costless exchange mechanism that was conducive to maximum efficiency, even if it needed the protection afforded by courts. Despite the prominence achieved by Herbert Simon’s concepts of organizational decision-making under uncertainty, imperfect information, and ‘satisficing’, most economists did not acknowledge that exchanges and cooperation could be costly because of the ubiquity of TCs.

Coase’s first contribution came in 1937. Five years earlier, however, Berle and Means had come to a similar (not identical!) conclusion regarding corporate finance (Berle and Means 1932). They had argued that the separation in American firms between ownership and control, the empowerment of managers compared to the (increasingly diluted) power of shareholders, and the ensuing emergence of that powerful new social class (managers of big business) posed several threats for the American economy and democracy. First, these factors made the corporate world rely on additional, unnecessary, and potentially wasteful exchange relationships; these relationships worked at the advantage of managers, who thus became increasingly unaccountable, thus damaging the interests of shareholders. Second, they made the American polity rely on a new self-serving and powerful social class, which, through its monopolisation motives and potential, posed a big threat to the competitive organisation of the economy. And third, they ultimately represented a threat to US democracy, similar to that identified by Thomas Jefferson in his industrial policy argument against James Madison. (Jefferson had championed a decentralised organisation of the economy along competitive lines with small owner-controlled units of production, while Madison had championed the creation of large industrial structures.)<sup>[10]</sup>

More specifically, among other claims, Berle and Means argued that the dilution of capital and the delegation of powers from (a) shareholders to boards of directors, and (b) boards of directors to managers made such exchange relationships work to the benefit of managers’ private interests rather than to those of the shareholders. In other words, shareholder dispersion created managerial discretion, and managerial discretion worked against the interests of shareholders, the economy and the polity. In addition, they specified the causal mechanism producing this effect: under such arrangements, shareholders need to spend considerable resources in gathering information about managers’ effort and activities, in negotiating managers’ contracts and duties, and in enforcing their views on managers. Further, this effect was reinforced by the dilution of ownership in the hands of ever increasing numbers of (relatively) small shareholders. This led to unwarranted managerial strategies, to inefficient monopolisation, and thus to the dangers identified at the systemic level.

Just like Coase’s 1937 contribution defined the fields that later came to be known as TCE, the theory of the firm, and organizational theory, Berle and Means’ 1932 contribution defined the fields that came to be known as PA and PTA. The concept of conflicting incentives was identified and defined, and its implications became the main research questions of the resulting body of works. Among these, the most fundamental question concerned the optimal structure of the PA contract. Intriguingly, just as Coase’s contribution was not integrated into economics for four decades, the contribution of Berle and Means took a considerable time to make its impact, too.<sup>11</sup> Even in their own

---

10 On the normative nature of the debate in 19th century America and in post-World War II sociology, see the absorbing account of Mizruchi 2004.

11 These delays have been identified in the literature (e.g. Coase 1972; Williamson 1996), and evidently point to some inefficiency in the functioning of the academy. They could be studied profitably using a PA, PTA, or TCE perspective. If (a) society (or research funding bodies) has an implicit (or explicit) contractual relationship with academics; (b) academics do not necessarily have the same motivations as the less informed members of society (or of funding bodies); and (c) monitoring the output of academics is costly, then academics may be allowed to pursue inefficient strategies. A PA perspective should focus on the optimization of academic research contracts from a societal perspective. A PTA



sub-discipline (corporate finance), the dominant view throughout most of the post-war decades was that of Modigliani and Miller (1958, 1963). According to Modigliani and Miller's 'theory of irrelevance', a firm's corporate and financial structure made no difference to its total value: the separation of ownership and management (i.e. the reliance on an exchange mechanism) did not seem to raise any particularly relevant issues, and nor did the choice between equity and different sorts of debt. Where Berle and Means had raised the issues of adverse selection, moral hazard and the resulting costs associated with correcting these flaws, Modigliani and Miller argued that such issues were irrelevant. With hindsight gained from theoretical developments and recurring corporate scandals, it is surprising to find that Modigliani's and Miller's theory remained dominant until the 1980s.

**Box 1: The original basic differences between Transaction Cost Economics and Normative Principal-Agent theory.**

	<b>TRANSACTION COST ECONOMICS</b>	<b>PRINCIPAL-AGENT</b>
<b>Origins</b>	Coase 1937	Berle and Means 1932
<b>Empirical Focus</b>	Vertical integration of firms	Separation of ownership and control
<b>Theoretical Focus</b>	Positive analysis of institutions	Normative analysis of contract
<b>Research questions</b>	When do firms produce to their needs, and when do they procure in the market?	What are the organizational and public policy ramifications of the separation of ownership and control?
<b>Institutions</b>	Institutional arrangements need to be explained	The effects of institutions need to be explained
<b>Contracts</b>	Contracts may be complete or incomplete, but the presence of transaction costs makes incompleteness more likely	Contracts are complete, because actors are rational and there are no transaction costs.

### **3. THE PATH TO FULL DIFFERENTIATION BETWEEN THE THREE ECONOMIC THEORIES: PA, POSITIVE AGENCY AND TC ECONOMICS IN THE LAST 40 YEARS.**

The rediscovery of TCE and PA started in the 1960s – and it started rather confusingly (albeit to the benefit of economics in general). The first steps were made in works that sought to combine Coasian TCs with the Knightian concept of moral hazard. More specifically, these works conceptualised moral hazard as a potential TC that resulted from the existence of conflicting incentives: each actor seeks to maximize his own utility function; since actors are in the market for different reasons, they have different utility functions; this means that, when they enter in a contractual relationship, each will try to 'pull' the exchange towards his own goals; unless the others are aware of this, they will fall pray to

(Contd.) \_\_\_\_\_

perspective should focus on the ways real-world actors seek to minimize agency losses. A TCE perspective should focus on a comparative study of the determinants of such failure. (e.g. Guston 2000, who goes as far as to refer to the '*integrity of research*'.) Note also that the 'opportunism' of academics may not be due to cynicism, but merely to specific incentives which are incompatible with societal goals. For example, the measurement of performance by publishing scores, or the measurement of doctoral research performance by years-in-writing standards distorts academics' incentives towards such specific 'deliverable', and away from time-consuming literature reviews and/or dissenting theoretical views.

his morally hazardous behaviour; hence, they will either try to be fully aware of this, or they will avoid entering into the contractual relationship.

Very soon, however, the existence of moral hazard was dissociated from TC. As we shall see below, the literature split into two branches: one that carried two sub-branches called PA theory, and PTA, and one that carried TCE. The first sub-branch of the first branch (PA theory, i.e. the formal, ‘normative’ principal-agent theory associated with Arrow, Akerlof, Dewatripont, Holmström, Laffont, Martimort, Mirrlees, Pauly, Spence, and Tirole) worked under the assumptions of perfect rationality and ‘transaction costlessness’. It focused on the causes and the effects of conflicting incentives and asymmetric information, and sought to (a) address the issue of efficient risk-sharing, and hence (b) provide a full characterization of the set of implementable allocations under conditions of imperfect information. Because the traditional topic of economics was markets, these theorists analysed ‘horizontal’ contracts, i.e. situations where there is exchange but no formal hierarchy. The central themes and historical developments of this branch of the literature (PA theory) can be split into four overlapping phases: (1) work on the first consequences of moral hazard: in the presence of moral hazard, market allocations under uncertainty will not be unconstrained Pareto optimal (Arrow, Akerlof, Pauly, etc – see the reviews and references in the text below); (2) work on the implications of the first consequences: the properties of the second-best allocations need to be systematically analysed (Ross, Mirrlees, Holmström, Tirole, Grossman and Hart, etc); (3) work on methodologies for such systematic analysis: the Revelation Principle, the rent-efficiency trade-off, the Informativeness Principle (Green, Laffont, Myerson, Dasgupta, Hammond, Maskin, Harris, Raviv, Holmström); and (4) work on extensions and reviews: multi-agents, common agency, renegotiation-proofness, dynamics, etc (Tirole, Laffont, Martimort, Dewatripont, etc).

The second sub-branch of the first branch (the ‘positive’ theory of agency, PTA, associated with Fama, Jensen<sup>12</sup> and Meckling), by contrast, retained Berle and Means’ original focus and later became associated with the analysis of vertical relationships within an organisation. (This is probably a path-dependent development which is independent of the original focus of the theory. Note also that vertical relationships in this theory are not to be confused with Coasian hierarchies and authority: the positive theory of agency relies explicitly on the assumption that all risk-sharing is voluntary). Although positive agency theory has always shared the normative theory’s concern for the causes and effects of conflicting incentives, this sub-branch has sought to explain empirical differences in the performance of different monitoring and bonding technologies within an organization (as expressed in the ex ante financial valuation of firms).

The second branch carries an altogether different theoretical approach, TCE. TCE starts with the assumptions that (a) actors may be opportunistic [due to (c) below], but they are also boundedly rational, (b) TCs are a ubiquitous feature of markets, (c) market transactions are plagued by incomplete contracts, which generates both ex ante and ex post contracting problems, and (d) market transactions are plagued by lock-in among trading partners. TCE is concerned with many contracting issues, chief among which is the determinants of the boundaries of the firm. It is thus based more directly on Coase than on Berle and Means (whose work is only a secondary source of influence), and it is associated with the works of Williamson (1968, 1971, 1975, 1983, 1985, 1996), Klein, Crawford and Alchian (1978), and Joskow (1985). Halfway between, on the one hand, PA and PTA, and, on the other hand, TCE, Grossman, Hart, and Moore, planted incomplete contracts theory (‘ICT’).

In order to understand the far-reaching differences between these economic theories, it is useful to present the fundamental works in each one. I first present the more demanding normative PA theory, then the more accessible PTA, and then TCE and the property rights critique addressed to it, which is

---

12 In fact, Jensen (1983) made the distinction between ‘principal-agent theory’ (the normative, formal theory) and the ‘positive theory of agency’ (the theory that political scientists call principal-agent theory). We will come back to the sources and the effects of this confusion in political science in the next section.

usually labelled ICT. I conclude with a summary presentation of the differences between these theories.

### ***3.1. Normative Principal-Agent theory: asymmetric information, frictionlessness and ex ante optimization***

Starting with his path-breaking 1963 article, Kenneth Arrow re-introduced into economics the concepts of information asymmetry, adverse selection, moral hazard and risk-bearing. Arrow (who did not explicitly study the agency relationship, but who raised the same questions) thus made the most important contribution to PA theory since Berle and Means. More specifically, he started exploring the ramifications of information asymmetries in the health- and life-insurance markets, i.e. settings where the seller of insurance schemes cannot directly observe the characteristics of buyers of such schemes, and where the buyers have an incentive to overuse the insurance scheme. He thus posed the problem of the optimal allocation of risk-bearing under conditions of un-revealed private information (adverse selection and moral hazard): how can actors choose the socially optimum amount of reimbursement claims, so that no other choice can make every individual better off? His theoretically-driven answer was that ‘the competitive allocation of risk-bearing is guaranteed to be viable only if the individuals have attitudes of risk-aversion.’ (Arrow 1964: 91)[13] The policy implication was that policy needs to shape individuals’ attitudes accordingly. Rational actors should be able to resolve all the problems raised by the existence of private information by drawing appropriate ex ante solutions. Note four points here: Arrow (a) took an explicitly normative approach; (b) worked under the assumption of actors’ perfect rationality; (c) did not include the possibility of the existence of transaction costs, even when he discussed public policy implications; and (d) focused on optimal ex ante incentive alignments. These, I will show below, are fundamental differences between PA and TCE.

Building on Arrow, and using the unlikely example of the second-hand market for cars (as well as insurance and credit markets), Akerlof (1970) analysed the effects of adverse selection and moral hazard and thereby established the basics of the theory of asymmetric information. In Akerlof’s market, some second-hand cars are in good working order (the ‘cherries’ – good-looking when you see them and sweet when you use them), while others have hidden defects (the ‘lemons’ – good-looking when you see them, but bitter when you use them). Yet, because of the existence of asymmetric information and the costs involved in correcting this condition, buyers don’t know which cars are lemons and which are cherries. What they do know, however, is that car-dealers have an incentive to sell lemons packaged as cherries. One intermediary result of this interaction between divergent incentives, asymmetrical information and transaction costs, is that the market price of even cherries will eventually decrease (because, under uncertainty, demand will decrease for any kind of second-hand car). The final result is that sellers of cherries are less inclined to sell their cars, and even a competitive market will only be filled with bad cars. The inefficiency does not only reside in the preponderance of lemons in the market, but also (and mainly) in the disappearance of the market for cherries. In PA terms (which Akerlof did not use), the impossibility of correcting the informational asymmetries between a principal (the prospective buyer, or the seller of a cherry car) and an agent (the car dealer) may lead to the collapse of the market through backwards deductions performed by the rational actors. In other words, unless the agent can offer a strong commitment to a minimum effort, there may be no exchange at all. But, the ability of the agent to make such a commitment is contingent upon the ability of the principal to not renegotiate the terms of the contract once the agent has made his investment.

---

13 Arrow later expressed doubts about the validity of his original insight. Normative PA eventually came to work most usually under the behavioral assumption of risk-neutral principal(s) and risk-averse agent(s).

The far-reaching implications of Akerlof's model were soon acknowledged by other economists, such as Arrow (again), Zeckhauser, Spence and Pauly,<sup>14</sup> who developed an important body of work on insurance problems. (In general, PA theory advances through applications of the same basic theoretical and methodological insights to different institutional contexts, with institutional differences being defined exogenously.) In this context, issues of commitment and renegotiation became more tractable, because the relationship between the principal and the agent develops over a longer period of time. However, these theorists did not extend their framework to incomplete contracts: due to its normative orientation, PA theory maintained its focus on complete contracts. It was shown that, due to adverse selection problems that pervaded the market for medical insurance, the insured were more likely to suffer a loss than the uninsured – again, the 'bad' won over the 'good'. Indeed, the seller of an insurance scheme is often unable to accurately distinguish between higher- and lower-risk customers. For example, a life-insurer may be unable to distinguish between smokers (who had an incentive to hide their addiction in order to not incur the corresponding premium price) and non-smokers. The insurer thus has to cover his risk by charging the same premium to both groups. This means that the insurance scheme on offer will be a relatively better deal for those covered by a smoker (e.g. his family) than for those covered by a non-smoker, and it will thus bias the customer base of the insurer in favour of smokers. But, if the insurer discovers this (i.e. if he learns that the mortality rates are higher among his customers than among the general population), he will further increase his premiums. As a result, non-smokers will prefer to not buy a life insurance, even though they were initially disposed to do so. In view of this market failure, some economists proposed a compulsory public insurance policy (Akerlof 1970; Arrow 1970), which would indirectly re-establish the rights of 'non-smokers'. Others seemed more sceptical regarding the substantive argument, but contributed to the development of the theoretical perspective (Pauly 1974). The normative inclination of PA theory, and the description of a frictionless (transaction-costs-less) world that were evident in Arrow's work were thus reinforced by this new focus on the design of 'mechanisms' (i.e. optimal game forms proposed by the principal).

In the meanwhile, an important development occurred when Ross (1973) put the foundations of modern contract theory with his first contribution to formal PA theory. Based on a view similar to Alchian and Demsetz's (1972) property rights approach and to Williamson's (1971, 1973) internal organization costs, Ross argued that much of the previous literature on moral hazard and informational asymmetries was actually concerned with contracts, and in particular with agency problems under conditions of uncertainty. His formal PA theory became very influential, most obviously because it raised a new question that complemented the one relating to the determinants of delegation. Simply put, Ross asked: what happens after delegation, and how is this taken into account at the time of delegation? Efforts to find the answer led to the creation of a new branch of economic theory, now called contract theory, incentives theory, implementation theory, or simply PA.

The basic theoretical apparatus of the PA methodology was completed with the demonstration of the so-called Revelation Principle (Green and Laffont 1977; Myerson 1979; Dasgupta, Hammond and Maskin 1979; and Harris and Raviv 1979). The Revelation Principle states that for any indirect revelation mechanism to which the agent reports his type, there is a direct mechanism with the same outcome. Using such a direct mechanism does not involve any loss of generality, and computations that go on within the mind of an agent who participates in a non-direct mechanism now become part of the direct mechanism.

The simplest way to understand this fundamental principle is by going back to an initial real world situation analysed by PA, i.e. a situation characterised by conflicting utility functions, incomplete

---

14 This does not mean, however, that Akerlof had been pushing an open door: before being published in the *Quarterly Journal of Economics*, the 'Market for Lemons' had been rejected by *The American Economic Review*, the *Review of Economic Studies*, and the *Journal of Political Economy*. The explanation accompanying these rejections was that the topic was irrelevant and/or that the analysis was wrong. (Akerlof 2001)

information and adverse selection. Both the principal and the agent will voluntarily accept to trade only if their expected payoff from participating (trading) exceeds the costs of their participation. In a game of incomplete information, the payoff-relevant private information of a player is called his 'type', and each player has at least two possible types. In such situations, the players need to discover each other's type; they thus need to agree upon a 'revelation mechanism' (i.e. a form of game) for communicating their type to one another; and thus they need to compensate each other for their participation and incentive compatibility costs. Evidently, by backwards deduction these decisions fall upon the principal, rather than upon the agent. In a revealing equilibrium of such a revelation mechanism each player is satisfied that his incentive and participation constraints (i.e. the conditions that need to be fulfilled in order for that player to not mimic another type than his real type) are met. The Revelation Principle is the idea that to any equilibrium corresponds an associated revelation mechanism that allows the players to truthfully report their types, subject to their participation and compatibility constraints being met. [Note the absence, by definition, of any factors that might impede the automatic application of the Principle, including the absence of (a) bounded rationality; (b) ex post communication bottlenecks; (c) costs related to the prevention of collusion between agents (or between the agent and external, competing principals); (d) agency costs imposed by the principal upon his own principals; (e) renegotiation costs; and (f) enforcement costs – in short, the absence of transaction costs.<sup>15</sup>]

Using the Revelation Principle, PA analysts were/are able to proceed to the analysis of the optimization problem of the principal, i.e. to stipulate an objective function for the organization and to calculate the effects of the trade-off between the achievement of allocative efficiency and the cost of insuring incentive compatibility – the so-called efficiency-rent trade-off.<sup>16</sup> They were/are thus able to

---

15 Extensive work in the 1980s and 90s relaxed some (but not all) of these assumptions: on the most important point (renegotiation costs due to the impossibility of full commitment) see Dewatripont 1988; Hart and Tirole 1988; Laffont and Tirole 1993: chapter 9. Interestingly, Hart and Tirole (1988) found that non-commitment may lead to the standard Coasian durable good model, and hence (in terms of the analysis) to the need to incorporate transaction costs.

16 Although it is useless to enter here into the most technical details of economic theory, this efficiency-rent trade-off is important for our purposes, too. It lies at the heart of PA theory, and political PA theory (if there is one to come) should duly take it into account. This trade-off can be understood as follows:

**INFORMATIONAL ASSYMETRY AND ADVERSE SELECTION:** A principal ('P') delegates the production of an output to an agent ('A'). P benefits from consuming the good produced by A, while A incurs a cost in producing it. However, before agreeing on the terms of the contract that binds them, only A knows the marginal cost of producing the good. P can therefore only try to calculate the risk taken by A when accepting to trade.

**FIRST BEST:** In the absence of incomplete information, P would propose to A a take-it-or-leave-it offer, so that the marginal cost of producing the good would equal the marginal benefit of consuming it. P would define the output target and a lump-sum compensation. At the margin, A would then be indifferent between producing the good or not. [Equivalently, the terms of trade would be efficiently established by the market.]

**EFFECTS OF ASYMMETRIC INFORMATION:** Unlike in the first best scenario, a relatively efficient A will claim higher costs (i.e. will mimic the type of a relatively inefficient agent) and will ask for a lower target output for that sum, i.e. he will seek to gain an informational rent (to reveal his true type only with adequate compensation). P will therefore have to overcome this information asymmetry: (a) he will need to gain information from A, and thus pay an informational rent that satisfies the incentive compatibility constraint; (b) he will need to ensure A's participation, and thus induce the participation of A even if A is relatively inefficient. Thus, P will either have to pay an informational rent, or decrease the output requested to a relatively less efficient A (so that A reveals his true type, since the resulting downward distortion requested by P to the least efficient A will induce the most efficient A to reject that offer).

**THE TRADE-OFF:** Hence, if P allows for the payment of an informational rent, he can be sure to collaborate with the most efficient A; if he does not allow for such a payment, he runs the risk of collaborating with a least efficient A. Resources will flow towards the most efficient A only if P accepts to pay a price for this allocative efficiency.

The stated objective of PA theory is to find the structure of the optimal contract given these constraints. The theory does so without reference to either transaction costs or to the possibility of ending up with an incomplete contract – the focus is exclusively on the effects of incomplete information under different institutional configurations.

obtain a full characterization of the set of implementable allocations within an organization under informational constraints.

Leaving aside the technical foundations and implications of the Revelation Principle (which relate to the trade-offs involved in ensuring the truthful cooperation of the agent), we can concentrate on what is of more direct relevance for the purposes of political-scientific research in inter-branch relations. First, the Revelation Principle is based on the assumptions that the principal can (a) perfectly reconstruct the strategies of the agent; (b) perfectly reflect these strategies in the mechanism that he proposes; and (c) perfectly commit himself to that mechanism, or else integrate his limited commitment into his initial offer (subject to paying a further price in order to satisfy the agent's renegotiation-proofness constraint). In other words, PA theory continues to work under the assumptions of perfect rationality, limited uncertainty regarding the future, and thus complete contracts. (This is most probably due to Arrow's legacy.)

And, second, the Revelation Principle renders the boundaries of the firm irrelevant. As Malin and Martimort (2000) explicate, the optimal direct revelation mechanism may be implemented in many different, equivalent and indirect ways. For example, the agent may report his information first, and the principal may then choose the particular output target and compensation (in which case the agent has some freedom of action but no real authority). Alternatively, the principal may offer a nonlinear price first, and the agent may then choose within this 'menu' his most preferred option (in which case the agent does have some amount of real authority). In both cases, however, the Revelation Principle can lead to the production of an optimal output. By the same token, the Revelation Principle cannot serve to explain the allocation of authority within the firm, nor indeed whether the agent owns his own productive unit or is part of the same organization as the principal. According to Malin and Martimort, this amounts to an 'irrelevance theorem' regarding ownership: '*... incentive theory [i.e. normative PA] has nothing to say about such things as the distribution of authority within an organization, the limits of the firm, the separation between the public and the private spheres of the economy, and more generally nothing to say about organizational forms and designs.*' (p. 129)

In addition to studying the effects of adverse selection, PA theory is also interested in the effects of moral hazard. The archetypical analysis of moral hazard posits that: (a) The good to be traded between an agent and a principal may be of different qualities and these qualities are observable, contractible and enforceable; (b) The quality of the good depends on the effort put by the agent in its production, and that effort is costly (i.e. the agent's utility function is not similar to that of the principal); (c) The agent may therefore choose the effort he will put in producing the good, or to not produce at all; (d) Neither the principal nor a court of law can directly observe and measure the effort put by the agent in the production of the good; and (e) The first-best solution (i.e. under complete information) would be to specify in the contract the value of the agent's effort (e.g. with a flat payment that would be independent from the realized quality of the good). Such a contract would at the same time force the agent to exert the first-best level of effort, and allow him to be fully insured against uncertainty on his final performance.<sup>17</sup> However, condition (d) makes the first-best unfeasible.

For these reasons, the principal must find a form of contract that reduces the negative effects of his incomplete information, while at the same time satisfying the agent's incentive and participation constraints. As with adverse selection, the principal faces an optimization problem. The first step of the analysis thus consists in describing the set of feasible incentive contracts implementing a given level of effort. Here, two analytical options are available, depending on whether the agent is risk-neutral or risk-averse:

---

<sup>17</sup> This condition is particularly important, for it shows that the 'first-best' condition satisfies both the principal and the agent. Most political-scientific accounts that claim to be based on this theory do not acknowledge this condition. (As I will show below, this is most probably because their claim to be based on PA is false; the alternative hypothesis is that political scientists do base their work on PA, but have not understood it.)

- a) With a risk-neutral agent, the principal can offer a ‘sell-out’ contract such that the incentives of the agent are aligned with the social incentives. This is achieved by having the agent pay a bond for the right to supply the principal, and having the principal pay an amount that depends on the quality realized. Under these conditions, the agent bears the whole responsibility for a bad performance (this is why this is feasible only with a risk-neutral agent). Note, however, that this is not feasible when institutions impose limited liability for the agent. With limited liability, the agent cannot be held totally responsible for bad performances; he is thus able to extract a limited liability rent. The principal can therefore offer only rewards (not punishments), and this makes him worse-off. This rent creates a rent-efficiency trade-off similar to the trade-off under adverse selection.
- b) With a risk-averse agent, the principal must offer the agent both insurance (i.e. a fixed payment that is independent of the realized quality of the good), and incentives (i.e. a link between payment and the realized quality of the good). The resulting trade-off between insurance and incentives is called the incentive-intensity principle. In addition (for technical reasons that need not be presented here), higher levels of performance are more informative about the agent’s effort than lower levels. And finally, the optimal contract must only use informative signals, but not uninformative ones. Uninformative signals increase the probability of the agent taking more risk, but without a corresponding need to relax incentives. [This is Holmström’s (1979) Informativeness Principle.]

By the mid-1980s, PA theory had reached a critical mass. At that moment, Grossman and Hart (1983, 1984, 1986) extended the theory to the analysis of incomplete contracts. Up to that point, PA had assumed that contracts can only be complete, in the sense that they foresee all relevant future contingencies, or at least that they foresee rules that allow to reach a common understanding of the nature of future contingencies. (Because the work of Grossman and Hart was presented as a constructive critique of TCE, I review it under that heading.)

Before offering some preliminary comments on this theoretical literature, let us note that the basic insights and methodology of the Revelation and the Informativeness Principles have been extended to analyse various environments. The most important extensions for our purposes are those that consider environments with multiple principals, with multiple agents, and with dynamic relations. Let us turn very briefly to each one of these.

Regarding the presence of multiple principals, note that the simple Revelation Principle methodology is based on the assumption that the single principal can costlessly prevent communication between his agent and outside (potential) principals. When this assumption is relaxed (which can only be done by allowing for transaction costs, and hence taking one element of the model exogenously), agents may seek to communicate with principals who have conflicting preferences. (This condition, which results in excessive rent extraction to the benefit of the agent, is sometimes described as a ubiquitous one in public policy-making.) Under such conditions, PA theory cannot use the Revelation Principle to describe the set of equilibrium allocations. Various scholars have thus developed a Taxation Principle (Martimort 1992; Martimort and Stole 1999a, 1999b, Peters 1999 – Martimort and Stole, and Peters are described in Malin and Martimort 2000), according to which non-cooperating principals cannot offer an efficient direct mechanism that is an equilibrium. The only possible equilibrium is when (a) the common agent ‘taxes’ the principals by retaining all powers, and (b) the common agent does not communicate any information that could induce the principals to compete between them.

Regarding the presence of many agents (‘common agency’), note that the simple Revelation Principle methodology is again based on the assumption that the principal can costlessly prevent communication between his agents. When this assumption is relaxed (which, again, can only be done by allowing for transaction costs), agents may seek to collude, either between themselves (e.g. workers against the interests of managers), or with their supervisors (e.g. workers and managers against the interests of shareholders). Such collusion may have a negative impact on the efficiency of the

organization. When this is so, the principal will draft his offer subject to one additional constraint, the coalition incentive compatibility constraint ('CICC'). The addition of CICC in the principal's calculus will guarantee that possible coalitions do not gain from colluding against the principal (Tirole 1986, 1992; Laffont and Martimort 1997). The Revelation Principle can accommodate this additional constraint.

Finally, regarding dynamic relationships, these can take many forms. The starting point is the assumption behind the use of the Revelation Mechanism that the principal can fully commit to the mechanism he proposes to the agent. The first-best is characterised by full commitment: the contract is complete, and it foresees large penalties in case of renegotiation. In the real world, however, the execution of a contract may reveal information about the existence of a Pareto-efficient re-negotiation. In the presence of limited abilities to commit (which are taken exogenously), the principal will foresee these, and he will accordingly offer an initial contract that takes renegotiation-proof constraints into account (i.e. adding these to the participation and incentive-compatibility constraints of the standard Revelation Principle).

The most relevant features of PA theory for our purposes are thus the following:

- 1) PA theory is based on the assumptions of perfect rationality and TC-lessness. PA analysis cannot be performed if these assumptions are relaxed to include bounded rationality, communication problems, inefficacious court ordering, etc. Empirical analyses of ex post contracting problems that cannot be comprehensively resolved ex ante violate the basic assumptions of the theory and are non-sensical.
- 2) Participation is voluntary: both in adverse selection and in moral hazard, the agent has an outside option, which is non-participation and non-production respectively. It is the existence of these options that create a set of constraints that the principal must satisfy in order to proceed to optimization. Empirical analyses that attempt to use PA without focusing on these constraints and on the resulting optimization problem of the principal are not principal-agent analyses.
- 3) PA theory does not treat the location of the agent as relevant: the agent may be part of the principal's organization, or he may own his own production unit. In both cases, the optimization problem facing the principal is the same. Attempting to explain either the location of the agent, or the consequences of his location place the analysis outside the PA framework.
- 4) PA theory does not attempt to explain the existence of multi-agent organisations: where these exist, they are defined exogenously, and their existence is usually attributed conjecturally to the need to (a) share common resources, (b) produce public goods, (c) internalise production externalities, or (d) enjoy information economies of scale.
- 5) Nevertheless, the existence of more than one agent is relevant in many respects: it affects the implementation concept proposed by the principal, the corresponding incentive feasibility, and ultimately the principal's payoff. Similarly, the theory admits that the optimal multilateral contract is very sensitive to the structure of information. The prevailing view here is that, unless (a) the multiple agents are risk-neutral and not risk-averse, (b) the multiple agents do not know each-other's type, and (c) the principal can condition one agent's compensation on another agent's report, the principal will not be able to fully extract the rent from all agents (Martimort 2006).
- 6) This theory takes into account the allocational effects of multiple and competing agents: when the multiple agents have types that are independently distributed, they will be able to extract rents from the principal. However, the distribution of these rents among the agents depends on the externalities that one agent's tasks may exert on another agent's tasks.
- 7) However, the theory admits that, in competitive environments, agents face a lesser incentive to report untruthfully (i.e. to overstate their costs), since competition among them creates a threat that such 'overstaters' will be excluded from production. This means that agents will either collude or loose rents. But collusion, too, can be avoided by the principal, for example by delegating more powers to lower levels of the hierarchy (Tirole 1986).



- 8) Thus, PA is not the ‘pop-theory’ usually presented by political scientists. It does not tell us anything about the causes of delegation, nor about ex post contracting issues – at least not directly. Analysing the relationship between one or more principals and one or more agents, all of whom are boundedly rational, all of whom seek to economize on TCs, all of whom supposedly have different goals and are engaged in incessant fights between themselves subject to institutional rules, are not conditions that correspond to this theory. PA is a very specific body of normative theoretical work, whose adoption requires a strict respect of its underlying assumptions (and a lot of mathematics!).

### ***3.2. Positive agency theory: incomplete contracts, agency losses, and periodical bargaining over ownership issues***

As mentioned above, the pioneering insights of Berle and Means did not immediately find their way into economics (or political science). These insights regarding the causes and consequences of the separation of ownership and control of large US corporations were revived only four decades later, in the work of Jensen and Meckling (1976). Jensen and Meckling argued that managers will not act to maximize the returns to shareholders unless appropriate governance structures are implemented; however, they explicitly mentioned that their positive theory of agency differed from the normative focus of principal-agent theory: the latter was concerned with the optimal solution to agency problems, while they were concerned with the effects of existing solutions (at 308). Because their work is the undisputed foundational work of modern agency theory, it is useful to present it in some detail.

Jensen and Meckling sought to combine the theories of (1) property rights, (2) agency, and (3) finance to develop a theory of ownership structure of the firm. Their theory aimed at explaining the value of several dependent variables, chief among which were:

*1. why an entrepreneur or manager in a firm which has a mixed financial structure (containing both debt and outside equity claims) will choose a set of activities for the firm such that the total value of the firm is less than it would be if he were the sole owner, and why this result is independent of whether the firm operates in monopolistic or competitive product or factor markets; 2. why his failure to maximize the value of the firm is perfectly consistent with efficiency. (at 305)*

The theoretical starting points were two: (1) the theory of the firm was actually a theory of markets that treated the firm as a black box (i.e. as a profit-maximizing production function), and not a theory of how conflicting interests inside the firm were brought into equilibrium so as to yield this result; and (2) the developing literatures on property rights and agency costs were different, but similar and complementary. The problem was therefore to use these theories in order to develop a genuine theory of the firm, while retaining the assumption that all individuals are resourceful, evaluative, maximizing and rational. Thus, Jensen and Meckling started with the following five assumptions<sup>18</sup>:

- 1) Costs and rewards are allocated among the participants in any organization according to a contractual (explicit or implicit) specification of property rights. These property rights create claims, and it is these claims that structure the incentives faced by the parties. (An obvious implication of this is that relations are set in a strategic context.)
- 2) Relationships within an organization can be understood as involving a contract between one or more principals and one agent, each having a different utility maximizing function.<sup>19</sup> Jensen and Meckling analyse the relationship between equity holders (principals) and managers

---

<sup>18</sup> These are the assumptions that are most relevant for our purposes. Jensen and Meckling introduce many more technical assumptions, only some of which are reported in the fifth point here.

<sup>19</sup> Jensen and Meckling are interested in hierarchical relationships within an organization, but clearly point to the fact that the agency problem exists in any kind of cooperative venture, whether vertical (as within some governmental authorities and bureaus) or horizontal (as between some governmental authorities or bureaus).

(agents). The principals can limit ‘aberrant activities of the agent’ with appropriate incentives and costly monitoring; the agent may benefit from guaranteeing his allegiance to the principals’ goals by incurring bonding costs. Both of these activities are costly: optimization does not equal zero divergence between the agent’s actions and the maximization of the principals’ welfare – there is always a residual cost due to inevitable shirking. Hence: Agency costs = monitoring expenditures + bonding costs + residual loss. Crucially, all players are assumed to act under a budget constraint. For example, the principals monitor the agent subject to their budget constraint.

- 3) Contrary to Coase (1937), activities within the firm are not governed by authority, but by contracts that vehicle voluntary exchanges (Alchian and Demsetz 1972). Organizations (including governmental bodies and states) are therefore legal fictions (i.e. artificial constructs that are legally treated as individuals) that serve as a nexus for a set of voluntary contracting relationships. In addition, firms are also characterized by the existence of divisible residual claims on their assets and cash flows which can generally be sold without permission of the other contracting individuals.
- 4) The starting condition is one where the manager of the firm is also its owner. He/she has to decide whether to continue with this financial structure (which allows him/her to enjoy various non-pecuniary benefits, but which comes at a cost in terms of availability of capital), to appeal to the equity market (including the extent of that appeal; such an appeal relaxes the financial constraints of the firm, but comes at a cost in terms of non-pecuniary benefits), or to take on debt (which, contrary to equity, does not dilute ownership).
- 5) Among a number of permanent assumptions, we find the following: (a) no trade credit is available; (b) all outside equity shares are non-voting; (c) no outside owner gains utility from ownership in any way other than through its effect on his wealth or cash flows; (d) dynamic and multiperiod aspects are ignored; and (e) there exists a single manager with ownership interest in the firm.

These assumptions have several far-reaching consequences. The five most relevant for our purposes of inter-branch analyses are the following:

- a) *‘Viewed this way, it makes little or no sense to try to distinguish between things that are ‘inside’ the firm (or any other organization) from those things that are ‘outside’ of it. There is in a very real sense only a multitude of complex relationships (i.e. contracts) between the legal fiction (the firm) and the owners of labor, material and capital inputs and the consumers of output.’* (section 1.5)
- b) *‘the personalization of the firm implied by asking questions such as ‘what should be the objective function of the firm?’ or ‘does the firm have a social responsibility?’ is seriously misleading. The firm is not an individual. It is a legal fiction which serves as a focus for a complex process in which the conflicting objectives of individuals (some of whom may ‘represent’ other organizations) are brought into equilibrium within a framework of contractual relations. In this sense the ‘behavior’ of the firm is like the behaviour of a market, that is, the outcome of a complex equilibrium process.’* (section 1.5)
- c) Equity and debt generate different agency costs, and so do different degrees of equity or debt. When a wholly-owned firm is managed by the owner, the optimum mix between of the various pecuniary and non-pecuniary (e.g. prestige, appointment power, level of employee discipline, personal relations, slack, purchase of inputs from friends, etc) benefits is achieved when the marginal utility of expenditure is equal for each non-pecuniary item and equal to the marginal utility of an additional dollar of after-tax purchasing power (wealth). *With equity*, the owner-manager bears only a fraction of the cost of his non-pecuniary benefits: the larger the proportion of outside equity, the greater the agency costs that it generates. On the other hand, equity markets anticipate these effects, hence prospective minority shareholders deduct agency costs from the price they are willing to pay for equity, and hence the owner-manager bears the entire wealth

effects of these costs. In addition, as the fraction of the manager's own ownership falls, monitoring becomes more rigorous (or bonding expenditures increase), perquisites diminish, and the value of the firm increases. *With debt*, the owner-manager can increase his non-pecuniary benefits. However, this happens because debt generates perverse incentive effects: the owner-manager engages in investments which promise very high payoffs if successful, even if they have a very low probability of success. *'If they turn out well, he captures most of the gains; if they turn out badly, the creditors bear most of the costs.'* For that reason, *'we don't find many large firms financed almost entirely with debt-like claims.'* (section 4.1) [<sup>20</sup>]

- d) The probability distribution of future cash flows depends on the capital (i.e. ownership) structure of the firm, because that structure determines the existence and the level of agency costs. This result is not only contrary to the Modigliani-Miller irrelevance theorem (see above); it is also contrary to the view that cash flows depend on structure because that structure determines bankruptcy costs. The existence of agency costs will determine the ownership structure of the firm; and given the perverse effects of debt on agency costs, agency costs will balance the structure in favour of equity.<sup>21</sup>
- e) The level of activity of the firm, the operational risk assumed by managers, and the agency costs resulting from the separation of ownership and control do not depend on the competitive environment of the firm, but on the incentive structure that results from the allocation of property rights. In particular, whether a firm is a monopolist or not, equity- and bond-holders will face an inherent incentive to protect their wealth.

Crucially, the method followed consists of two steps: (A) an identification and explanation of agency costs, given assumptions 1-3 above; and (B) an explanation of how these agency costs affect organizational structures. In other words, it is the explanation of why and how the corporate form (i.e. an organizational form that is a nexus of contracts between voluntarily collaborating individuals who can exercise their right to exit without preconditions) generates agency costs that leads to a theory of the ownership (or capital) structure of the firm. In the absence of the first step of the analysis, it is impossible to proceed to the second step. Note three important points here: first, in the theory, the first step is made by the agent (it is the agent who decides whether he wants any principals and, to a lesser extent, what powers they should have over him); second, the theory depends on the existence of competitive and transparent financial markets that value firms (including agency costs) *ex ante* (there are no *ex post* costs that financial valuation cannot pre-establish); and third, it is implicitly assumed that any (improbable) *ex post* dispute can be efficiently resolved through public ordering (the judicial system).

The most important result of Jensen and Meckling for our purposes is that the initial owner-manager ultimately bears the entire wealth effects of the (expected) agency costs. Thus, *'although the separation of ownership from control attenuates profit incentives, that is anticipated at the time separation occurs and is fully reflected in the price of new shares ... The future therefore holds no surprises; all of the relevant contracting action is packed into ex ante incentive alignments.'* (Williamson 1985: 27) The result reached by Jensen and Meckling is less surprising (though innovative at its time) when one considers the assumptions underlying this theory, particularly regarding the rationality of actors (including in their anticipatory actions and reactions), the perfect functioning of financial markets, and the limitation of the utility of ownership to cash flows.

The next main step in the development of agency theory was the work of Fama and Jensen (1983). This important and accessible contribution is probably the single most influential article in political-

---

20 The difference between equity and bonds is important – and possibly for political scientific analyses, too. However, it is enough to present it here without going into the details of finance theory, since these details do not obviously reflect similar differences in politics.

21 One of the referees points out that Jensen's initial conclusions on this point were subsequently revised, given the ubiquity of leveraged buyouts (LBOs) characterized by a high level of debt.

scientific works that claim (misleadingly) to be based on ‘principal-agent’ theory. It is therefore necessary to present its assumptions, its logic and its conclusions in some detail, too.

Fama and Jensen started by expressly referring to Berle and Means, and then defined their goal as follows: ‘*we are concerned with the survival of organizations in which important decision agents do not bear a substantial share of the wealth effects of their decisions.*’ Given this dependent variable, Fama and Jensen proceeded by setting out the following four sets of assumptions and definitions:

- 1) Organizations are nexuses of contracts (as in Jensen and Meckling 1976). The two most central contracts specify (a) the nature of residual claims, and (b) the allocation of the steps of the decision process among agents. ‘*These contracts distinguish organizations from one another and explain why specific organizational forms survive.*’ (at 302) Regarding (a), concentrating residual claims to one group of actors reduces monitoring costs, adjustment costs, and production costs, and hence increases the survival probability of the organization. Regarding (b), there are four identifiable steps of the decision process: (A) initiation, (B) ratification, (C) implementation, and (D) monitoring. Steps (A) and (C) correspond to ‘decision management’ functions, while (B) and (D) correspond to ‘decision control’ functions.
- 2) Residual claims can be categorised as open or close, according to the restrictions that they impose on their owners. Open claims are those that do not impose any obligations on their owners and which therefore allow unrestricted risk sharing (e.g. common stocks of large publicly-listed corporations). Closed claims impose restrictions, in that they can only belong to internal decision agents (e.g. stocks of small family firms). The crucial difference between these two systems lies in the number of risk-bearing actors. (Note how residual claims correspond to residual risk-bearing: if everything goes well, claims are positive; if everything goes badly, they are negative.)
- 3) Agency costs arise because contracts are not costlessly written and enforced, and because output is never perfectly aligned with risk-bearers’ preferences - see Jensen and Meckling (1976). Agency problems arise when decision managers are not the major residual claimants and therefore do not bear a major share of the wealth effects of their decisions. In such cases, effective control procedures are necessary. Hence, residual claimants must separate ‘*decision control*’ (which they may exercise themselves) from ‘*decision management*’ (which they may delegate).
- 4) Organizations are distinguished according to their complexity. ‘Noncomplex’ means that specific information (i.e. detailed information that is costly to transfer among actors) relevant to decisions is concentrated in one or a few actors. ‘Complex’ means that specific information relevant to decisions is diffused. [There is a rough (but imperfect) equivalence between an organization’s size and its complexity.] Fama and Jensen ‘*take it as given that optimal organizations in some activities are noncomplex.*’

The main question regards the determinants of the combination or separation of (a) residual claims, (b) decision control, and (c) decision management, or any two of these. When it is more efficient to combine (b) and (c) (e.g. in small noncomplex organizations), control becomes relatively costly, and this makes it preferable to combine (a), too. However, this comes at a cost, since it sacrifices the benefits of unrestricted risk-sharing and specialization. On balance, then, combination occurs only where the benefits of unrestricted risk-sharing and specialization of decision functions are less than the costs that would be incurred to control the resulting agency problems. This is indeed the case where small noncomplex organizations are characterized by trust, by the absence of demands for a wide range of specialized decision agents, by low total risk of net cash flows to be shared, and by low demands for wealth from residual claimants. (See also Fama and Jensen 1983b.)

Separation of (a) and (b), on the other hand, occurs when it is more efficient to separate (c) from (a): the agency problems that result from the efficient separation of decision management from residual risk-bearing are best confronted by separating decision control from decision management.

The separation of decision management from residual risk-bearing may be due to the fact that the corresponding organizations are complex, i.e. to the fact that decision-specific knowledge is diffused and costly to transfer.<sup>22</sup> In that case, the agency problems of diffuse decision management can be reduced by separating the management (initiation and implementation) and control (ratification and monitoring) of decisions.

The institution of the separation of decision and risk-bearing functions survives (as opposed to other organizational arrangements, or competing organizations with such different arrangements) because of two reasons: (a) it allows to benefit from the specialization of management and risk-bearing; and (b) its potentially negative effects are mitigated by the creation of control mechanisms. Crucially, Fama and Jensen did not deny the relevance of (a), but placed more emphasis on (b). Hence, the question of the survival of such organizations was answered by pointing to the separation of the ratification and monitoring of decisions from initiation and implementation of the decisions.

Several points need to be stressed regarding Fama and Jensen's contribution. First, delegation may be a central concern to their theory, but this is not a theory of delegation. Note, for example, that the separation between (a) and (c) is the starting, not the ending, point of the separation between (a) and (b). Complexity is taken as given. This is most probably due to the fact that Fama and Jensen were implicitly referring only to large corporations engaged in the production of technological or complex financial products. In such cases, complexity is a matter of technology, and it can therefore only be defined exogenously. Second, these scholars were interested in the survival of organizations in competitive environments (which, as a research question, differs from the reasons for their creation of such organizations, the degree of separation of functions, and the consequences of such separations); in other words, their argument did not concern environments where the organization has some degree of monopoly power.<sup>23</sup> The functionalist explanation of organizational survival was inherently linked to the existence of such competitive pressures<sup>24</sup>. Third (and related), Fama and Jensen argued that the case under investigation (the survival of organizations where decision is separated from risk-bearing) is relevant for many categories of organizations, including non-profit organizations; nevertheless, they explicitly limited their analysis to (a) private-sector organizations that (b) depended on voluntary contracting. Fourth, the model was based on an understanding of the organization as a nexus of contracts, and among these contracts one of the two most important ones concerned the nature of residual claims, i.e. the definition of the residual risk-bearers. It was precisely that explicit definition of the residual risk-bearers (i.e. the definitional absence of the availability of blame-shifting strategies) that motivated the principals in their control of the agents.

The positive theory of agency immediately raised the interest of management scholars (Foss 2005; Eisenhardt 1989; Donaldson and Davis 1991) and public management scholars (Lane 2003). Of particular interest is the work of Lex Donaldson and James Davis, who sought to test the empirical

---

22 In their own words, 'Most organizations characterized by separation of decision management from residual risk bearing are complex. ... Again, we take it as given that the optimal organizations in some activities are complex.'

23 In their own words, '*the form of organization that delivers the output demanded by customers at the lowest price, while covering costs, survives.*' (at 302)

24 This point is particularly important for political-scientific 'principal-agent' models. [Such models are evidently more influenced by Fama and Jensen's PTA (and, to a lesser but unspecified extent, by Williamsonian TCE) than by the economic PAt literature.] In these models, delegation of powers to a bureaucratic agent is said to be explained by reference to functionalist reasons (i.e. searching for causes by looking at effects). This may be attributable to Fama and Jensen's functionalist 'natural selection' approach, whereby (1) relatively efficient types are discovered by looking at types that survive, (2) survival is explained by looking at organizational differences between surviving and disappearing types, and (3) these differences are assumed to be the determinants of relative efficiency. Note, however, that Fama and Jensen did not describe delegation as their dependent variable: they assumed that some delegation would occur exogenously. In addition, their functional logic was strictly limited to the point mentioned above. This is evident from their non-functional distinction between profits and survival: '*producing outputs at lower cost is in the interests of residual claimants because it increases net cash flows, but lower costs also contribute to survival by allowing products to be delivered at lower prices.*' (at 305)

validity of agency theory against the opposite hypotheses generated by stewardship theory. As mentioned above, agency theory in the Berle & Means and Jensen & Meckling tradition stipulates that shareholders' interests will be safeguarded only where the chairmanship of the board of directors is separated from the chief executive officer ('CEO') function. This separation serves to introduce an essential element of control upon otherwise opportunistic CEOs. Stewardship theory, by contrast, stipulates that (a) there is no general problem of executive motivation and incentive compatibility constraints, and thus, (b) a more critical factor for shareholder returns is a correctly designed organization structure which allows the CEO to take effective action. By granting real authority to the CEO, shareholders provide him/her with the ability to quickly and adequately respond to changing circumstances, thus allowing the uninterrupted pursuit of value creation. According to Donaldson and Davis, these two theories correspond to different '*models of man*' (at 51), i.e. different fundamental assumptions about the nature of the social world. (The authors of this study found evidence against agency theory and some evidence in favor of stewardship theory.)

Positive agency theory is therefore not merely a simplification of the normative PA theory, but also a footnote to that theory. It is a simplification in so far as it adopts a less formal style, while accepting many of the assumptions of PA (it accepts the analysis of relationships as contracts; it accepts the existence of opportunism and different utility functions, and it accepts the view that asymmetric information poses coordination problems – however, it does not seek to demonstrate any of these elements). It is also an important footnote in so far as it is also founded on the idea of costly transactions – this in itself represents a significant departure from PA theory. In addition, where PA theory assumes a risk-neutral principal and a risk-averse agent, most works in positive agency theory make the opposite assumption: they define a risk-averse principal and a risk-neutral agent. Finally, the goal of agency theory is to specify the mechanisms used by real-world organizational institutions in order to reduce agency loss, such as incentive schemes, monitoring and sanctions.

### 3.3. *Transaction<sup>25</sup> costs economics: friction, incomplete contracts, lock-ins, and governance*

TCE is an elaboration on (mainly) Coase's and (secondarily) Berle and Means' original insights on the role of organizational factors in economics. Despite some obvious similarities with PA and PTA (i.e. the institutional focus and the opportunism assumption), TCE economics differs in some fundamental ways, which pertain to assumptions, analysis and results. It should therefore be contrasted with these theories. TCE cannot enter a PA analysis as an ad hoc consideration, for PA is explicitly based on the assumption that there are no TCs. Conversely, PA concepts should not enter a TCE analysis on an ad hoc basis. Needless to say, the same holds for TCE, PTA, and ICT.

As mentioned above, Coase (1937) had raised the questions of the determinants of the scope of the firm, and argued that transactions will be organized in the firm ('integration') when the cost of doing this is lower than the cost of doing it in the market.<sup>26</sup> This idea was resurrected by Williamson, who is by far the most important writer in TCE (Williamson 1963, 1964, 1967, 1968, 1973, 1975, 1983, 1985, 1996 – but see also Klein, Crawford and Alchian 1978, and Joskow 1985, 1987). Let us then present a summary of Williamsonian TCE, turning first to its origins and some of its main developments, and then to a summary presentation of its main points for our purposes.

Williamson (1963, 1964) started by focusing on the nature of the firm, and by attempting to treat it as a governance structure (i.e. an organizational construction) rather than, as in the neoclassical mode, as a production function (i.e. rather than as a technological construction). For this reason, he first

---

25 A transaction occurs '*when a good or service is transferred across a technologically separable interface.*' (Williamson 1985: 1) In politics, this is the case whenever a political entrepreneur shares his idea with someone who will help him, whenever a legislative committee sends a bill to the floor, whenever the legislature delegates powers to the bureaucracy, etc.

26 Note here that there is a notable difference between integration and 'delegation'; the latter occurs both in market and in integrated settings, and its causes and implications are not topics that are specific to TCE.

analyzed managerial discretion in the firm, and the effects of such discretion on the firm's behavior. He systematically referred to Coase, Knight, Berle and Means, and Simon, and laid the foundations of the concepts of incomplete contracts, asset-specificity, and adaptation. Indeed, because the most obvious example of the need for adaptation was the employment contract [to which Coase had originally referred, Simon (1951) had already analysed in terms of incomplete contracts, and Berle and Means (1932) had ... failed to place at the heart of financial economics], Williamson analysed such 'vertical' contracts, i.e. situations where the exchange involves a command-and-control function in return for a remuneration (in Simon's original phrasing, '*authority relationships*', Simon 1951: 293). Williamson posited that management (i.e. the employees of shareholders) was responsible for three inter-related decisions: (a) prices and maximum profits, (b) marketing expenditures, and (c) own remuneration. Each of these decisions could serve as the basis for decisions in the other two. The question then, was which of the three decisions, (a), (b), or (c), would prevail over the other two. Building on the work of Baumol (1959), Williamson showed that rational managers might not only display a preference for (b) and (c), but might actually be able to impose their decisions on shareholders who preferred (a). (Note the similarity with the conclusions of Berle and Means.) His proof consisted in showing through careful case studies that managerial discretion mattered: in commercial crises that took the form of sharp decreases in demand, the typical reaction of managers in those firms where management enjoyed discretion consisted in a disproportionate cut of expenses compared to all the other outlays. This falsified the profit-maximising hypothesis (e.g. Friedman 1953), which predicted that in times of crisis all outlays would be reduced proportionately. By the same token, these two articles served to revitalise interest in governance structures and agency problems.

In a subsequent article on competition policy, Williamson (1968) argued that non-standard (i.e. apparently anti-competitive hierarchical governance) corporate behavior, such as franchise restrictions and mergers, should not be automatically attributed to monopolization strategies, but rather to efforts to minimize TCs. Thus, although some franchise agreements and/or mergers may indeed lead to some efficiency losses (higher prices), such effects are almost always outweighed by efficiency gains (lower TCs).<sup>27</sup> This set the stage for the development of what Williamson has called the TCE approach of the efficiency branch of contract (1985: 26-9). In particular, the main insights of the 1968 paper were to lead to the full explication of the consequences of transaction-specific assets and bilateral monopoly.

In subsequent work (1973, 1975), Williamson attempted a synthesis of two literatures of the emerging new-institutionalist approach: on the one hand, the 'horizontal' literature on possible market failures and, on the other hand, the 'vertical' literature on governance problems within the firm. He argued that '*concern with the study of market failures should be expanded to include 'institutional failures' (of internal organization, political, and judicial types) more generally [because] substantially the same factors that are ultimately responsible for market failures also explain failures in internal organization.*' (1973: 316) In other words, Williamson argued that the same concepts and models that can be used to study contracts *between* firms can be used to study relationships *inside* firms, and vice versa (1975, chapter 1). Where external procurement is plagued by incomplete contracts, lock-ins among trading partners, increasing opportunism, and correspondingly high efficiency losses, internal procurement mediates these problems, but creates issues of bureaucratic costs and low-powered incentives. He therefore implicitly conceptualized discrete market exchange and centralized hierarchical organization as defining the extremes of a continuum. Between these two extremes lie an infinite number of possible institutional arrangements (quasi-markets). His focus thus turned to the determinants of particular institutional arrangements, where the dependent variable could take a very large number of values. Measured by the degree of institutionalization of the solution, these could be any of the following (or anything between these values): no exchange, market exchange, quasi-market

---

<sup>27</sup> Note that this result may be inconsistent with Williamson's previous work, which, as mentioned above, emphasized managers' non-efficiency motivations and hence their inefficient strategies. This apparent inconsistency became one of Williamson's motivations in further research.

exchange, quasi-hierarchy exchange, hierarchy exchange, and complete hierarchical contract. (Specifications regarding the fundamental assumption of forbearance prevailing in internal allocational disputes came later.)

Note here that Williamson did not make explicit use here of the developing PA theory. First, he explicitly referred to political and judicial institutional failures, thus differentiating his economics from the perfectly rational, frictionless, complete, and *ex ante* contracting world of PA. Second, his research questions now concerned the determinants of the *boundaries* of the firm (and the *relative advantages of different institutional structures*), not solely the *internal organization* of the firm. (Although it is true that PA theory might in principle be used to study both issues, it is also true that it cannot differentiate between the two.) Williamson's goal being to develop a theory that endogenised the formation of both institutions and governance structures, he had to adopt a larger and different lens.

More specifically then, Williamson posed again the classic Coasian question of why some firms choose to rely more on the market (i.e. have more open boundaries and therefore face higher market TCs), while others choose to rely more on hierarchy (i.e. have more defined boundaries and therefore face lower market TCs but higher internal organization costs). In order to answer this question he referred to the concepts of asset-specificity, time inconsistency, hold-up, and private ordering. Let us turn now to these fundamental concepts of TCE.

Like Coase (1937, 1972), Williamson argued that hierarchy could be explained by reference to its role in transaction costs-economizing, and that it would be expanded until TCs of managing transactions within the firm met TCs of mediating transactions through the market. Yet, his first original contribution was to add that there exists a fundamental difference between environments characterized by competition and environments characterized by asset specificity and thus bilateral monopoly. In competitive environments, firms may be subject to both bounded rationality and to opportunism. But, in the absence of asset specificity, they have no continuing interests in the identity of one another: competition makes discrete market contracting efficient, even in the presence of bounded rationality and opportunism. In the presence of asset specificity, however, the pair-wise identity of the parties matters, and thus external competition cannot lead to efficiency. The only circumstance in which asset specificity might not lead to considerable inefficiencies is where the judicial system (court ordering) is efficient. However, in the presence of bounded rationality, opportunism, and transaction costs (i.e. the assumptions of TCE), this cannot be easily achieved, and hence asset specificity usually leads to inefficient outcomes. Contractual safeguards (e.g. exclusivity contracts, or integration), then, are a legitimate (efficient, not monopolization-minded) way to re-establish some measure of efficiency: they are '*safeguards to protect investments*'. (Williamson 1975, 1983, 1985: 32)

It is important to note that, although neo-classical economics had made some inroads into this idea, Williamson (1983, 1985) offered a totally original interpretation. Economists had long recognized that firms might not always be able to buy their inputs, or sell their outputs, on competitive markets. For example, bilateral monopolies (where the specific output of a firm could serve as an input of only one other firm) might impede that. However, they had argued that this would be rare, since most markets would be characterized by the existence of substitutable products and clients. Thus, neo-classical economists recognized that non-standard contracting and/or integration might amount to benign strategies, but only under the very rare circumstances under which such asset specificity prevailed. Williamson's original insight about asset specificity accepted this view, but also added an entirely new justification for non-standard contracting and/or integration. He acknowledged the existence of neo-classical exogenous (or, 'technological') dependencies, and, based on Klein et al. (1978), he went further to argue that dependency could also be endogenous (or, 'contractual'). Dependency could result from inter-temporal, contractual factors: '*what begins as a large numbers supply condition frequently is transformed into a small numbers exchange relation during contract execution and at contract renewal intervals.*' (1996: 26). Thus, '*decisions to integrate are rarely due*



*to technological determinism but are more often explained by the fact that integration is the source of transaction cost economies.*' (1985: 87; the same applied to non-standard contracts) Only through non-standard forms of contracts and/or integration can the firm put an end to endless bargaining and re-contracting.

The problems created by the Williamsonian dependency are best illustrated by the 'hold-up' metaphor (to which Williamson refers more commonly as 'expropriation hazards' – see also Joskow 1985): when contracts are incomplete and incentives differ, and therefore when there is a potential problem of time-inconsistent incentives, the dependent party has to ensure that the other's opportunistic behaviour does not generate a hold-up. Hence, once such contractual dependency arises and the ensuing hold-up risk becomes too important, integration represents a rational, economizing and non-monopolistic solution. Hence, according to Williamson, asset specificity is one of the most important of the critical dimensions on which transactions differ:

*Transaction cost economics ... maintains that the most critical dimension for describing transactions is the condition of asset specificity. Parties engaged in a trade that is supported by nontrivial investments in transaction-specific assets are effectively operating in a bilateral trading relation with one another. Harmonizing the contractual interface that joins the parties, thereby to effect adaptability and promote continuity, becomes the source of real economic value.* (1985: 30; see also 1996: 26)

The penultimate important conceptual innovation of Williamson (at least for our purposes) was the link that he established between (a) asset specificity, (b) the problem of adaptation, and (c) credible commitments. The combination of asset specificity and bilateral dependency means that adaptations may become problematic. This is because, if asset-specificity is sought and a relationship-specific investment is to be made, the investor requires a credible commitment by his partner that the investment will actually be used (i.e. the investor needs to be insured because his costs are sunk). Yet, such a credible commitment will impede adaptation. And yet, adaptation is the central problem of economic organization (and, as Williamson repeatedly emphasized by referring to and correcting Machiavelli, political organization, too). Therefore, the parties (or at least one of them) face a trade-off between investment and flexibility. This trade-off is resolved either by referring disputes to an efficient court system (the existence of which is axiomatically denied), or by agreeing to exchange 'hostages'. [Note here that the court system is not assumed away – it is only thought of as too costly to become the main dispute resolution mechanism (Williamson 1996: 122-23); thus, as the court system becomes comparatively inefficient, parties seek to establish governance structures ('private ordering').] In other words, the parties need to institute reciprocal trading relations that bond them: *'An alternative way [to arbitration] by which to protect contracts against expropriation is to expand the contractual relation. One way of accomplishing this is for the buyer and seller to devise a mutual reliance relation.'* (1996: 132)

Finally, and crucially, Williamson defended the idea that internal *'hierarchy is its own court of ultimate appeal.'* (1991, 1996: 98) This means that the contract law applicable to internal organizational disputes is that of forbearance (i.e. courts of law will forbear, in the sense that they will refuse to hear disputes between one internal division and another). In other words, being inside or outside a hierarchy matters, because being inside it forces one to adopt cooperative conflict-mitigating strategies, while being outside it allows one to be heard by a court of law. In economics, that means that firms differ from markets, because the employees of a firm will have to resolve their conflict within the firm; on the contrary, inter-firm relations will be regulated by the courts. To the best of my knowledge, this important issue has not been addressed by political scientists – at least not by those working on inter-branch relations. And yet, to the extent that it constrains our definition of different organizational actors, it is of vital importance to political science. (See Scharpf 1997: 52-60)

Williamson's conceptual innovations met with some skepticism [see the review of Grossman and Hart (1986) below] but were also strongly supported by empirical research. In particular, the pioneering empirical work of Joskow (1985, 1987) on contractual transactions between coal suppliers

and electric utilities showed that (a) it is unnecessary to refer only to the simple dichotomy between vertical integration and the market, since long-term contracts allow us to gain many interesting insights; (b) asset specificity considerations were indeed an important factor affecting the structure of vertical relationships in such markets; and (c) the level of relationship-specific investments determines both the length of commitments to the terms of future trade at the contract execution stage, and (negatively) the level of ex post bargaining. (*Coal market transactions are interesting to focus on because there is considerable variation in the duration and structure of vertical relationships between buyers and sellers.* Joskow 1987: 168) This relatively easy-to-grasp proposition has not been empirically tested in political science.

Thus, Williamson and Joskow made a strong point in favour of the TCE approach, and, within the remit of this approach, against interpreting all non-competitive market outcomes (such as high profits, vertical integration, etc) as attributable to monopoly. Furthering his 1968 insights, and again contradicting the dominant neo-classical view (according to which high profits and/or a choice in favour of hierarchy was only attributable to monopolization), Williamson argued that such outcomes almost always have a more benign explanation: transactions cost-economizing strategies.

Regarding the main characteristics of TCE, they can be summarized as follows:

- 1) The ultimate goal of TCE is to explain institutions and governance structures – that is the dependent variable. Like in PTA, the underlying assumption is that competition works as a mechanism of natural selection of the most efficient institutions.<sup>28</sup>
- 2) Like PA, PTA, and ICT, TCE adopts a contracting orientation, in which both ex ante incentive alignments (PA and PTA) and property rights (ICT) matter. Contracts may be explicit or implicit, and thus every exchange relation qualifies. Unlike PA and PTA, however, TCE views contracting as problematic. The problem is not situated at the level of incentive alignments (which can be achieved ex ante via imaginative contracts), but at the levels of TCs (which contracts cannot resolve, and which thus create both ex ante and ex post problems) and asset specificity. Consequently, one of the originalities of TCE lies in its greater emphasis on the contract execution stage<sup>29</sup>: *‘bargaining is pervasive’* (Williamson 1985: 29).
- 3) The behavioral attributes of human agents are opportunism (like in PA and PTA) but also bounded rationality (contrary to the perfect rationality assumption of PA and PTA). Opportunism means that actors are *‘self-interest seeking with guile.’* Bounded rationality means that behavior is *‘intendedly rational, but only limitedly so.’* (Williamson 1985: 30, 45-50; 1996: 8-9, 12, 36-7, 42-4, 173, 224-5) Obviously, self-interest may not be egoistic or immoral: in certain procuring transactions even the Red Half-Moon may be self-interested.
- 4) The operational research question of TCE is whether transactions are completed harmoniously, or whether there is instead an economic equivalent to mechanical friction, which would render transactions inefficient. Contrary to the PA and PTA, this question refers to both post-contracting and ante-contracting issues. It thus refers both to institutions (the original contract), and to governance structures (the ex post adaptation mechanisms).
- 5) The basic tool of TCE is a micro-analytic focus on the organizational factors that affect transactions: asset specificity and the resulting TCs (or, more generally, the degree of hazard inherent in a transaction) are the independent variables. But, merely assuming the existence of TCs is not enough: TCs have to be operationalized by identifying the salient features of

---

28 It is impossible to dissociate economic theories from their reliance on the idea that actual or potential competition acts as a natural selection mechanism (the only exception being the quasi-theory of path dependence). This may be a consequential obstacle to their direct application to (some) political environments. Perhaps it is also an indication of the scope that different theories of institutional change should try to acquire (i.e. theories of institutional change should specify the nature, operation and extent of the mechanism of institutional selection on which they rely).

29 Any study of post-contracting relationships can therefore benefit from performing a TCE analysis, rather than in engaging in non-sensical translations of PA or PTA concepts.

different market, hierarchical, quasi-market organizational forms. Different organizational forms generate different TC levels, and these levels are best studied comparatively rather than as absolute values. Such TC may include ex ante and ex post costs. Note, however, that TCE acknowledges the presence of additional factors that may explain institutional and governance-related decisions. In particular, *'economizing takes place with reference to the sum of production and transaction costs, whence the tradeoffs in this respect must be recognized.'* (Williamson 1985: 22) This means that trying to minimize TCs may result in higher production costs, and vice versa. (This is reminiscent of, but clearly different from, the rent-efficiency trade-off in PA theory.)

- 6) Ex ante TCs include drafting, negotiating, and safeguarding costs. Depending on the level of drafting and negotiating costs, the contract may be more or less complete (drafting complete contracts is costly, but drafting incomplete contracts may generate frictions in the future – whence a trade-off that is resolved by the degree of bounded rationality, and hence by the complexity of the envisaged transaction). Depending on the level of safeguarding costs, the parties may agree to pool their property rights (which is not unproblematic), or to fashion signals of credible commitments.
- 7) Ex post TCs are due to the limitations of legal centralism (see the next point below). They include (a) maladaptation costs incurred when transactions drift out of alignment in relation to the 'shifting contract curve'; (b) haggling costs incurred if bilateral efforts are made to correct ex post misalignments; (c) setup and running costs associated with the governance structures to which disputes are referred; and (d) bonding costs of effecting secure commitments.
- 8) Unlike PA and PTA, TCE assumes that court ordering (i.e. the use of the judicial system) is not a frictions-reducing mechanism, but a costly mechanism in its own right. Access to the court system, however impartial the latter may be, is costly. This is the source of ex post TC (see the previous point above): if courts could costlessly resolve any post-contracting issue, there would not be ex post TC (which is a defining assumption of PA theory, and which is rejected by TCE).
- 9) Acknowledging the existence of TC, and TC-economizing strategies and institutions, amounts to questioning the neoclassical skepticism towards nonstandard modes of economic organization, such as customer and territorial restrictions, tie-ins, block-booking, franchising, vertical integration, etc. Under conditions of uncertainty, bounded rationality, opportunistic behavior, and bilateral monopoly, such governance institutions do not amount to monopolization strategies, but should be interpreted as safeguards to protect investments.
- 10) The method is comparative: *'Transaction cost analysis supplants the usual preoccupation with technology and steady-state production (or distribution) expenses with an examination of the comparative costs of planning, adapting, and monitoring task completion under alternative governance structures.'* (Williamson 1985: 2, emphasis in the original) Contrary to PA and PTA, the focus here is not only on normative, but also on positive issues. The basic unit of empirical analysis is the individual transaction. Comparative analysis is therefore based on the view that (a) different governance structures represent different adaptive capacities and thus different levels of TCs; (b) different transactions differ in their attributes; and thus (c) TCs are economized by assigning transactions to governance structures in a discriminating way. (Williamson 2002: 441(v))

For our purposes, the three most crucial points of Williamsonian TCE are these:

First, the firm is not analyzed as a production function that consumes inputs in order to produce outputs in an otherwise unspecified way. On the contrary, that black box of neoclassical economic theory is opened up in order to be analyzed as a governance structure. Institutions and governance structures are then explained by reference to their transaction cost-economizing roles, rather than by reference to the production costs and profits of the firm (though overall efficiency is a necessary

requirement, in the absence of which an organizational form will be supplanted by competition). Even more simply put, what matters is the harmonizing effect that institutions and governance structures may have on transactions, rather than the cost (pecuniary, opportunity or political) of producing a good.

Second, when a strong bilateral interdependence exists between two parties in an incomplete contractual relationship, appropriate safeguards (e.g. vertical integration, incentive alignment mechanisms, arbitration mechanisms, or reciprocities) enable the parties (or at least one of them) to protect their specific investments against the potential hold-up that the other's opportunistic behavior could generate.

And third, the central prediction of TCE is that:

*as market transactions become characterized by increasing levels of contractual incompleteness, inter-dependence, and quasi-rents, the likelihood of integration should increase. (...) In a typical study, some measure of lock-in, such as the specificity of the product procured or investments made, is related to the choice of whether to integrate. The strong association that this literature has found between specificity and integration has made TCE one of the great success stories in industrial organization over the last 25 years. (Whinston 2001: 184-85)*

Shortly after being developed, TCE started being criticized from two sides. On the one hand, traditionalist antitrust specialists took an 'anti-bigness' approach similar to that of Berle and Means (1932). They maintained that integration might have some positive efficiency effects, but that Williamson had greatly exaggerated their magnitude, while minimizing the magnitude of their negative effects (e.g. Pitofsky 1979, Turner, quoted in Williamson 1985: 19). Williamson refers to this line of criticism as the 'monopoly branch of contract' (1985: 23-6). On the other hand, some economists took a sympathetic but critical view of TCE, arguing that it could perhaps explain the importance of TCs in market settings, but could definitely not explain (a) how integration could change the scope for opportunism, and therefore (b) how integration was more economizing than the market (Grossman and Hart 1986). It is to this second criticism of TCE, branded 'property rights theory' or 'incomplete contract theory', that I now turn.

Note preliminarily that Epstein and O'Halloran (1999, which is probably the best political-scientific work using these economic theories to explain delegation) refer to both TCE and the 'property rights paradigm' (p. 42). In the context of describing and explaining the importance of controlling the physical assets involved in production for the purposes of avoiding hold-ups, they write:

*The more usual solution to the hold-up problem is for the upstream and downstream firms simply to merge ... The advantage of this hierarchically structured firm is that one person controls all physical assets of both companies, and she can then engage in relationship-specific investments that would not otherwise be made for fear of the hold-up problem. ... The importance of controlling the physical assets involved in production is the keystone of the property rights approach to vertical integration. (note 7: In fact, this is the principal advantage of the property rights paradigm, as traditional transaction cost approaches have difficulty explaining exactly why mergers can solve the hold-up problem.. (at 49)*

What is puzzling in this quote is that Epstein and O'Halloran introduce the property rights paradigm in their 'transaction cost politics approach to policy making' (the title of their book). I show here that this is theoretically impossible, because these two theories rely on different assumptions. In particular, the property rights paradigm to which Epstein and O'Halloran refer works under the explicit assumptions that (a) there are no ex post TCs, (b) there are only limited instances of uncertainty, (c) there are no information asymmetries, (d) contractual incompleteness stems mainly from limitations of the judicial system, and (e) limitations notwithstanding, the court system is an

efficient dispute resolution mechanism. These assumptions, as we just saw, are not compatible with a TCE framework.<sup>30</sup>

Grossman and Hart (1986 – the two leading exponents of the property rights approach to which Epstein and O’Halloran refer) started by expressing their sympathetic view on TCE, and by acknowledging the presence of ex ante TCs and incomplete contracts (by the same token, they differentiated themselves from the mainstream of PA theory, which did not envisage this possibility). They also differentiated themselves from PTA, defining the firm as being composed of its assets, rather than as a nexus of contracts between voluntarily participating individuals.<sup>31</sup> On this basis, Grossman and Hart argued that (a) Williamson had exaggerated the benefits of integration by assuming that integration automatically turns a hostile supplier into a docile employee, (b) TCE did not make any distinction between the activities carried out via contract between separate owners and the activities carried out in a single ownership unit, and thus did not provide an answer to the question of the boundaries of the firm, and (c) TCE was unable to distinguish in terms of ‘integration’ between in-house employees and exclusive agents. Evidently, this criticism revolves around the issue of forbearance. (And, to the extent that forbearance has not yet been convincingly demonstrated in public policy environments, Grossman and Hart point to a crucial issue facing political scientists.)

Grossman and Hart relied on the concepts developed in the property rights literature (e.g. Alchian and Demsetz 1972). They thus distinguished between the situation where a contractor uses the production facilities that the firm provides to him, and the situation where the contractor uses his own production facilities. *‘If the firm owned [all the] important assets of the independent agents, then we would say that such a company had the same degree of integration as a company in which the retail sales force was composed of ‘employees’.*’ (at 694)<sup>[32]</sup> On this basis, they made the following five assumptions:

- 1) There is, ex ante, a competitive market in identical potential trading partners. This market determines the ex ante division of surplus between the principal and the agent. Optimization involves maximizing the principal’s benefit subject to the agent’s participation constraint (unlike in PA models, there is neither uncertainty nor asymmetric information, and thus no incentive compatibility constraint);
- 2) The payment method used by the principal to compensate the agent (whether he is an employee or an independent contractor) is some function of the observable states of the nature and the observable performance of the parties to the contract;
- 3) Contracts are incomplete in the sense that (a) ex post residual rights, (b) ex ante investments, and (c) ex post benefits to the respective parties, cannot be totally foreseen, specified and contracted ex ante. This is because (a) depends, by definition, on the realization of the state of nature, because (b) are either too complex to be described, or not verifiable by a court of justice, and because (c) depends both on the realization of the state of nature and on nonverifiable actions, such as effort. (Note that incompleteness means that some assets are not contracted upon, and hence that their owners retain full control over them);

---

30 Both Grossman and Hart (1986) and Hart (1995), and Williamson (1985) have repeatedly emphasized these differences. Epstein and O’Halloran refer to all these authors, but do not seem to have paid particular attention to their comments on this issue.

31 In addition, Grossman and Hart noted that *‘control or ownership is never absolute. For example, a firm that owns a machine may not be able to sell it without the permission of the lenders for which the machine serves as collateral.’* (at 694) They thus recognized that, in spite of ownership giving full rights, ownership itself may be diluted.

32 This important point can serve to clarify the difference between the concepts of integration and delegation. Integration means acquisition of property rights over assets, while delegation relates to contractual compensation. Delegation occurs independently of integration.

- 4) Integration in itself does not make any new variable observable to both parties<sup>33</sup>; and
- 5) Integration in itself (a) does not change the cost of writing down contracts, and hence cannot be associated with complete contracting; but (b) does change who has control over those provisions not included in the contract.

Grossman and Hart proceeded by building a two-periods model which applied equally to vertical and horizontal relationships. In the ex ante period, both firms make relationship-specific investments; in the ex post period, some further (partly unforeseeable, unforeseen and thus non-contractible) production decisions are taken and the benefits from the relationship are realized. The incompleteness of the contract regarding (some) ex post decisions and benefits makes it necessary to allocate residual rights of control ex ante.

In the ex post period, the state of the world is revealed, whatever information asymmetries existed disappear, further decisions are taken<sup>34</sup>, and the parties can renegotiate and recontract on further actions and residual rights of control. [Such renegotiations and recontracting are assumed to be costless; ex post costlessness is justified by the fact that the state of nature is now realized, and hence by the fact that (a) both parties now know what they need, and (b) they now know what they need equally well.] This possibility amounts to ex post efficient allocation, and renders the initial allocation of property rights ex post irrelevant. Crucially, however, the distribution of the ex post surplus (as opposed to its amount) is not insensitive to property rights: if the agent owns his production facilities, he may refuse to continue the relationship unless he extracts more surplus – or he may be prepared to give up his right for a side payment as part of a renegotiation. In turn, this will affect ex ante investment decisions on non-contractible assets.

Finally, assuming that the parties allocate ownership rights in a way that minimizes ex ante investment distortions, Grossman and Hart argued that the usual argument that integration can only expand the feasible set fails: integration distorts the incentives of the principal, and (given the unspecified nature of his residual rights) he cannot commit to intervene only selectively in the agent's operations. Hence, integration leads to overinvestment by the integrating firm, and underinvestment by the integrated one; it therefore imposes costs as well as benefits. The only solution is not Williamsonian integration, but a clearer specification of property rights. Phrased as a hypothesis, Grossman and Hart argued that residual rights of control over assets (or actions) determine who own which assets (or actions).

## CONCLUSION

This a detailed presentation of four economic approaches to organization (PA, PTA, TCE, and ICT) demonstrates that, in spite of some common concerns, these economic theories are neither substitutes nor complements. The fact that these theories are used to frame different research questions does not mean that they can be used sequentially in a Cartesian deconstruction, nor that elements of one can be introduced in an ad hoc fashion in an analysis based on another one. In fact, each theory builds on a different set of assumptions that differ from those of the other theories. Economists such as Arrow,

---

33 This was a direct criticism of Arrow (1975), who had argued that vertical integration economizes on communication costs, but who had not explained why (he had not explained why the incentives facing opportunistic agents change with integration). As mentioned several times in the text, it also affected TCE and its assumption of forbearance.

34 For our purposes, it is necessary to note that each party takes such further decisions as a unitary actor: the choice of such decisions is (from an organizational point of view) unproblematic, because any subordinate can take it. In Grossman and Hart's words, '*no special skills are required ... [S]ince there are many subordinates available, none is in a position to refuse to carry out the owner's wishes or to argue about terms.*' (at 699) Note two things here: (1) This leaves the question of the existence of the firm unanswered (if no special skills are required, the corresponding decision could be sub-contracted externally without risk); and (2) This competition among subordinates is not necessarily found in all political settings.

Fama, Grossman, Hart, Jensen, Laffont, Martimort, Tirole and Williamson have been particularly attentive to these points, and there are no obvious reasons why political scientists should not take example.

The upshot is that a complete and consistent theory cannot be based on more than one of these theories. An actor can be assumed to be either rational or irrational, but not both at the same time. (Similarly, most political scientists would want to assume that all actors are either rational, or that all are boundedly rational.) The same holds for transaction costs (they can be assumed to be relevant or irrelevant), the existence or absence of competition in factor markets, the completeness or incompleteness of contracts, etc.

A strong personal experience can be reported to illustrate the point. Having finished the review that you just read, I asked 'X' (an eminent empirically-oriented political scientist who specializes in issues of delegation and control) to comment on the main parts of this paper. As X himself confessed, he thought this paper was counter-productive. Specifically, his comment was that there already exists a significant body of political-scientific literature, which proves that excellent empirical work can be done without even noting all these ('uncertain', I think his word was) complications. But this confirms one of my two main points! Much of what X read was indeed very far away from his current research-design political-scientific interests. X was right, but only in the (unacknowledged by him) sense that much of PA, PTA, TCE, and ICT has (almost) nothing to do with the existing framing of delegation and control issues.

Now consider what happened in the end. X stopped reading this paper, apologized for so doing, and went back to his own work. In that work (on which I was kindly asked to comment), X continued to make plenty of references to 'principals', 'agents', 'principal-agent', 'delegation to international organizations', 'transaction costs', 'runaway agents', Coase, Williamson, Epstein and O'Halloran, etc. This confirms my second main point! Because of various resource constraints and informational asymmetries, the dynamics of the political-scientific 'Lemon market' win over those of the 'Cherry market': the necessary exchange between good economists and good political scientists never actually occurs. (I also think that X forgot to mention in his text that he was not referring to Williamson, but to what he has heard about him. This rather cynical view of mine is consistent with most of the theories examined here.)

Of course, politics differs from economics, both in substance, and in theories and methods. It is therefore necessary to accept the fact that inter-disciplinarity should not come at the expense of some amount of pluralism. In addition, confusing as they are, the works of eminent scholars such as Epstein and O'Halloran, Franchino, Majone, Moe, etc, have done a great deal in terms of empirical knowledge, theoretical sophistication and methodological explication. Nevertheless, political science as a discipline still lags behind economics. This may be due to several factors, including ones that we may never be able to act upon. But it may also be due to the fact that we have not paid enough attention to the theoretical equivalents of our methodological rules for validity and reliability, namely (a) logical completeness, (b) internal consistency, and (c) relevance. In the more limited context of rational choice approaches to the issues of delegation of powers and subsequent relationships, we have not yet distinguished carefully between the different elements of the theories that we apply.

The box below is aimed at summarizing the main characteristics of the economic theories of PA, positive agency and TC, and at highlighting their important differences. (I do not include ICT because there is no identifiable political-scientific literature that claims to be based on that theory.) A companion paper entitles 'Economic Theories and the Science of inter-Branch Relations' makes more inroads into (a) the specific ways these theories have been used (and often abused) by political scientists; (b) the necessity to turn our attention to more theory and – if operating with limited resources – less methodology; (c) the trade-offs involved in choosing one among these four economic theories; and (d) the relative appeal of opting for a TCE approach – which does not mean that there are no opportunity costs in so choosing.

**Box 2: A synthetic comparison of the main characteristics of the economic theories of PA, positive agency, and transaction costs.**

	<b>PA</b>	<b>Positive agency</b>	<b>Transaction costs</b>
<b>Intellectual origins</b>	Arrow	Berle & Means	Coase, Berle & Means, Simon
<b>Research question</b>	Optimal contract	Agency costs and risk-sharing	Institutions of governance
<b>Unit of analysis</b>	The contract	The agent	The transaction
<b>Temporal focus</b>	Ex ante	Ex ante	Ex ante and ex post
<b>Behavioral assumptions</b>	Perfect rationality and Opportunism	Rationality (varying) and Opportunism	Bounded rationality and Opportunism
<b>Contracts</b>	Complete but renegotiable subject to penalties	Incomplete with periodical renegotiations	Incomplete with constant renegotiations and bargaining
<b>Organizations</b>	Irrelevant	Open nexus of voluntary contracts	Authoritarian or cooperative hierarchies
<b>Problem</b>	Asymmetric information; optimization under rent-efficiency trade-off	Asymmetric & complex information, definition of rights and agency costs	Transaction costs and asset specificity
<b>Dispute resolution mechanisms</b>	No disputes	Disputes and efficient court system	Disputes and costly court system; need for governance structures
<b>First best</b>	Unattainable due to the rent-efficiency trade-off	Unattainable due to agency costs	Unattainable due to bounded rationality and transaction costs

Yannis Karagiannis  
 Marie Curie Fellow, Centre for Regulatory Governance, University of Exeter, UK  
 SPS, European University Institute, Florence, Italy  
 i.karagiannis@exeter.ac.uk  
 ykaragiannis@yahoo.co.uk



**REFERENCES**

- Akerlof, G.A. (1970), 'The Market for Lemons: Quality Uncertainty and the Market Mechanism.' *Quarterly Journal of Economics* 84(3), 488-500.
- (2001), 'Writing the "The Market for 'Lemons'": A Personal and Interpretative Essay.' <http://nobelprize.org/economics/articles/akerlof/index.html>.
- Alchian, A., and H. Demsetz (1972), 'Production, Information Costs, and Economic Organization.' *American Economic Review* 62(4), 772-95.
- Arrow, K. (1963), 'Uncertainty and the Welfare Economics of Medical Care.' *American Economic Review*, 53(5), 941-973.
- (1964), 'The Role of Securities in the Optimal Allocation of Risk-Bearing.' *Review of Economic Studies* 31(2), 91-96.
- (1970), *Essays in the Theory of Risk-Bearing*. Amsterdam: North Holland Publishing Co.
- (1975), 'Vertical Integration and Communication.' *Bell Journal of Economics* 6(2), 173-83.
- Ballmann, A., D. Epstein, and S. O'Halloran (2002), 'Delegation, comitology, and the separation of powers in the European Union.' *International Organization* 56(3), 551-74.
- Bates, R., A. Greif, M. Levy, J-L. Rosenthal, and B. Weingast (1998), *Analytic Narratives*. Princeton: Princeton University Press.
- Baumol, W. (1959), *Business Behavior, Value and Growth*. New York: Macmillan.
- Bergman, T. (2000), 'The European Union as the next step of delegation and accountability.' *European Journal of Political Research* 37(3), 415-29.
- Berle, A., and G. Means (1932), *The Modern Corporation and Private Property*. New York: Macmillan. [ed. 1969, New York: Harcourt, Brace & World]
- Blom-Hansen, J. (2005), 'Principals, agents, and the implementation of EU cohesion policy.' *Journal of European Public Policy* 12(4), 624-48.
- Boix, C. (2003), *Democracy and Redistribution*. New York: Cambridge University Press.
- Brouseau, E. and M. Fares (2000), 'Are Incomplete Contract Theory and New-Institutional Economics Substitutes or Complements?' In C. Ménard (ed.), *Institutions, Contracts, and Organizations: Perspectives from New Institutional Economics*. Cheltenham Glos: Edward Elgar, 399-421.
- Buchanan, J. (1986), *Liberty, Market, and State: Political Economy in the 1980s*. New York: New York University Press.
- Chiappori, P. and B. Salante (2000), 'Testing Contract Theory: A Survey of Some Recent Work.' Mimeo, University of Chicago.
- Coase, R. (1937), 'The Theory of the Firm.' *Economica* 4(16), 386-405.
- (1972), 'Industrial organization: a proposal for research.' In V. Fuchs (ed.), *Policy Issues and Research Opportunities in Industrial Organization*. New York: NBER, 59-73.
- (1992), 'The Institutional Structure of Production.' *American Economic Review* 82(4), 713-19.
- (1994), 'How Should Economists Choose?' in *Essays on Economics and Economists*. Chicago: University of Chicago Press, chapter 1.
- (2000), 'The Acquisition of Fisher Body by General Motors.' *Journal of Law and Economics* 43(1), 15-31.

- Dasgupta, P., P. Hammond, and E. Maskin (1979), 'The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility.' *Revue of Economic Studies* 46(2), 185-216.
- De Bièvre, D., and A. Dür (2004), 'Delegation and Agency Control in European and American Trade Policy.' Paper presented at the ECPR Joint Sessions, Uppsala, April 13-18.
- Dewatripont, M. (1988), 'Commitment through Renegotiation-Proof Contracts with Third Parties.' *Review of Economic Studies* 55(3), 377-89.
- Dixit, A. (1997), 'Power of Incentives in Private versus Public Organizations.' *American Economic Review* 87(2), 378-82.
- Doleys, T. (2000), 'Member states and the European Commission: Insights for the new economics of organization.' *Journal of European Public Policy* 7(4), 532-53.
- Donaldson, L., and J. Davis (1991), 'Stewardship Theory or Agency Theory: CEO Governance and Shareholder Returns.' *Australian Journal of Management* 16(1), 49-64.
- Egan, M. (1998), 'Regulatory strategies, delegation and European market integration,' *Journal of European Public Policy* 5(3), 485-506.
- Eisenhardt, K. (1989), 'Agency theory: an assessment and review.' *Academy of Management Review* 14(1), 57-74.
- Elster, J. (2000), 'Rational Choice History: A Case of Excessive Ambition.' *American Political Science Review* 94(3), 685-95.
- Epstein, D. and S. O'Halloran (1996), 'Divided Government and the Design of Administrative Procedures: A Formal Model and Empirical Test.' *Journal of Politics* 58(3), 373-97.
- (1999), *Delegating Powers: A Transaction Cost Politics Approach to Policy Making Under Separate Powers*. New York: Cambridge University Press.
- (2006), 'Delegation Games and Sovereignty.' Paper prepared for presentation at the Workshop on Delegating Sovereignty, Duke University, 3-4 March 2006.
- Fama, E., and M. Jensen (1983), 'Separation of Ownership and Control.' *Journal of Law and Economics* 26(2), 301-26.
- (1983b), 'Agency Problems and Residual Claims.' *Journal of Law and Economics* 26(2), 327-49.
- Foss, N. and P. Klein (2005), 'The Theory of the Firm and Its Critics: A Stocktaking and Assessment.' DRUID Working Paper 05/03, [http://www.druid.dk/wp/pdf\\_files/05-03.pdf](http://www.druid.dk/wp/pdf_files/05-03.pdf)
- Franchino, F. (2000a), 'Control of the Commission's Executive Functions: Uncertainty, Conflict and Decision Rules.' *European Union Politics* 1(1), 63-92.
- (2000b), 'The Commission's Executive Discretion, Information, and Comitology.' *Journal of Theoretical Politics* 12(2), 155-81.
- (2002), 'Efficiency or credibility? Testing the two logics of delegation to the European Commission.' *Journal of European Public Policy* 9(5), 677-94.
- (2004), 'Delegating Powers in the European Community.' *British Journal of Political Science* 34(2), 269-93.
- (2007), *The Powers of the Union: Delegation in the EU*. Cambridge: Cambridge University Press.
- Friedman, M. (1953), 'The Methodology of Positive Economics', in *Essays in Positive Economics*. Chicago: University of Chicago Press, chapter 1.

- Garrett, G. (1992), 'International Cooperation and Institutional Choice: The European Community's Internal Market.' *International Organization* 46(2), 533-60.
- Garrett, G. and B. Weingast (1993), 'Ideas, Interests, and Institutions: Constructing the European Community's Internal Market.' In J. Goldstein and R. Keohane (eds.), *Ideas and Foreign Policy*. Ithaca NY: Cornell University Press, ch. 7.
- Gibbons, R. (1999), 'Taking Coase Seriously.' *Administrative Science Quarterly* 44(2), 145-57.
- Gilardi, F. (2001), 'Principal-Agent Models Go to Europe: Independent Regulatory Agencies as Ultimate Step of Delegation.' Paper presented at the ECPR General Conference, Caterbury (UK), 6-8 September.
- Green, J, and J-J. Laffont (1977), 'Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods.' *Econometrica* 45(2), 427-38.
- Grossman, S., and O. Hart (1983), 'An Analysis of the Principal-Agent Problem.' *Econometrica* 51(1), 7-45.
- (1986), 'The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration.' *Journal of Political Economy* 94(4), 691-719.
- Guston, D. (2000), *Between Politics and Science: Assuring the Integrity and Productivity of Research*. Cambridge: Cambridge University Press.
- Harris, M., and A. Raviv (1979), 'Optimal Incentive Contracts with Imperfect Information.' *Journal of Economic Theory* 20(2), 231-59.
- Hart, O. (1995), *Firms, Contracts, and Financial Structure*. Oxford: Clarendon Press.
- Hart, O., and J. Tirole (1988), 'Contract Renegotiation and Coasian Dynamics.' *Review of Economic Studies* 55(4), 509-40.
- Hart, O. and J. Moore (1990), 'Property Rights and the Nature of the Firm.' *Journal of Political Economy* 98(6), 1119-58.
- Holmström, B. (1979), 'Moral hazard and observability.' *Bell Journal of Economics* 10(1), 74-91.
- Huber, J., and C. Shipan (2006), 'Politics, Delegation, and Bureaucracy.' In B. Weingast and D. Wittman (Eds.), *The Oxford Handbook of Political Economy*. Oxford: Oxford University Press, 256-72.
- Immergut, H. (1998), 'The Theoretical Core of the New Institutionalism.' *Politics & Society* 26(1), March 1998, 5-34.
- Jensen, M. (1983), 'Organization theory and methodology.' *Accounting Review* 58(2), 319-39.
- Jensen, M., and W. Meckling (1976), 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure.' *Journal of Financial Economics* 3(4), 305-60.
- Joskow, P. (1985), 'Vertical Integration and Long-Term Contracts: The Case of Coal Burning Electric Generating Plants.' *Journal of Law, Economics and Organization* 1(1), 33-80.
- (1987), 'Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets.' *American Economic Review* 77(1), 168-85.
- Jun, H-W. (2003), 'Catching the Runaway Bureaucracy in Brussels: Euro-Parliamentarians in Budgetary Politics,' *European Union Politics* 4(4), 421-45.
- Karagiannis, Y. (2007), 'Economic Theories and the Science of Inter-Branch Relations.' Robert Schuman Centre for Advanced Studies Working Papers Series, 2007/04.

- Kassim, H., and A. Menon (2003), 'The principal-agent approach and the study of the European Union: promise unfulfilled?' *Journal of European Public Policy* 10(1), 121-39.
- King, G., R. Keohane, and S. Verba (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton (N.J.): Princeton University Press.
- Klein, B., R. Crawford, and A. Alchian (1978), 'Vertical Integration, Appropriable Rents, and the Competitive Contracting Process.' *Journal of Law and Economics* 21(2), 297-326.
- Knight, F. (1921[1965]), *Risk, Uncertainty and Profit*. New York: Augustus M. Kelley.
- Laffont, J-J. and D. Martimort (1997), 'Collusion under Asymmetric Information.' *Econometrica* 65(4), 875-911.
- (2002), *The Theory of Incentives: The Principal-Agent Model*. Princeton (N.J.): Princeton University Press.
- Laffont, J-J. and J. Tirole (1993), *A Theory of Incentives in Regulation and Procurement*. Cambridge, Mass: MIT Press.
- Lane, J-E. (2002), 'Democracy in the European Union: What is the Democratic Deficit?' Lee Kuan Yew School of Public Policy Working Paper SPP-20-02, National University of Singapore.
- Lane, J-E., and S. Ersson (2003), 'Regional Coordination Mechanisms: Major Similarities and Differences.' Lee Kuan Yew School of Public Policy Working Paper SPP-49-03, National University of Singapore.
- McCubbins, M., R. Noll, and B. Weingast (1987), 'Administrative Procedures as Instruments of Political Control.' *Journal of Law, Economics, and Organization* 3(2), 243-77.
- Majone, G. (ed.) (1996a), *Regulating Europe*. London: Routledge.
- (1996b) 'Temporal Consistency and Policy Credibility: Why Democracies Need Non-Majoritarian Institutions,' Robert Schuman Centre for Advanced Studies, EUI Working Papers, RSC No. 1996/57.
- (1997a), 'From the Positive to the Regulatory State: Causes and Consequences of Changes in the Mode of Governance.' *Journal of Public Policy* 17(2), 139-167
- (1997b), 'Independent Agencies and the Delegation Problem: Theoretical and Normative Dimensions.' In B. Steunenberg and F. van Vught (eds.), *Political Institutions and Public Policy*, Kluwer, Dordrecht, 139-56.
- (2001), 'Two Logics of Delegation: Agency and Fiduciary Relations in EU Governance.' *European Union Politics*, 2(1), 103-21.
- Malin, E., and D. Martimort (2000), 'Transaction Costs and Incentive Theory.' *Revue d'Economie Industrielle* 92(2/3), 125-48.
- Martimort, D. (1992), 'Multi-Principaux avec Anti-Sélection.' *Annales d'Economie et de Statistique* 28(1), 1-38.
- Mizruchi, M. (2004), 'Bearle and Means revisited: The governance and power of large U.S. corporations.' *Theory and Society* 33(5), 579-617.
- Modigliani, F. and M.H. Miller (1958), 'The Cost of Capital, Corporation Finance and the Theory of Investment.' *American Economic Review* 48(3), 261-97.
- (1963), 'Corporate Income Taxes and the Cost of Capital: A correction.' *American Economic Review* 53(3), 433-43.
- Moe, T. (1984), 'The New Economics of Organization'. *American Journal of Political Science* 28(4), 739-77.

- Moravcsik, A. (1998), *The Choice for Europe: Social Purpose and State Power from Messina to Maastricht*. Ithaca: Cornell University Press.
- Myerson, R. (1979), 'Incentive Compatibility and the Bargaining Problem.' *Econometrica* 47(1), 61-73.
- Ordeshook, P. (1990), 'The emerging discipline of political economy,' In J. Alt and K. Shepsle (Eds.), *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press, 9-30.
- Pauly, M. (1974), 'Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection.' *Quarterly Journal of Economics* 88(1), 44-62.
- Pierson, P. (1996), 'The Path to European Integration: A Historical Institutional Analysis.' *Comparative Political Studies* 29(2), 123-63.
- Pitofsky, R. (1979), 'The Political Content of Antitrust.' *University of Pennsylvania Law Review* 127(3), 1051-79.
- Pollack, M. (1997), 'Delegation, Agency and Agenda Setting in the European Community.' *International Organization* 51(1), 99-135.
- (1998), 'The Engines of Integration? Supranational Autonomy and Influence in the European Union.' In W. Sandholtz and A. Stone Sweet (eds.), *European Integration and Supranational Governance*. New York: Oxford University Press, 217-49.
- (2000), 'International Relations Theory and European Integration.' EUI Working Papers, RSC No. 2000/55, <http://people.cas.sc.edu/coate/Readings/Pollack.pdf>
- (2002), 'Learning From the Americanists (Again): Theory and Method in the Study of Delegation.' *West European Politics* 25(2), 200-19.
- (2003), *The Engines of European Integration: Delegation, Agency, and Agenda Setting in the EU*. Oxford: Oxford University Press.
- (2006), 'Principal-Agent Analysis and International Delegation: Red Herrings, Theoretical Clarifications, and Empirical Disputes.' Paper prepared for presentation at the Workshop on Delegating Sovereignty, Duke University, 3-4 March 2006.
- Rasmussen, A. (2005), 'EU Conciliation Delegates – Responsible or runaway agents? – Principal-agent analysis and the study of delegation.' *West European Politics* 28(5), 1015-34.
- Ross, S. (1973), 'The Economic Theory of Agency: The Principal's Problem.' *American Economic Review* 63(2), 134-39.
- Scharpf, F. (1997), *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research*. Boulder CO: Westview Press.
- Simon, H. (1951), 'A Formal Theory of the Employment Relationship.' *Econometrica* 19(3), 293-305.
- Spence, M., and R. Zeckhauser (1971), 'Insurance, Information, and Individual Action.' *American Economic Review* 61(2), 380-87.
- Stone Sweet, A., and W. Sandholtz (1998), 'Integration, Supranational Governance, and the Institutionalization of the European Polity.' In Sandholtz, W. and Stone Sweet, A. (eds.), *European Integration and Supranational Governance*, Oxford University Press, Oxford, 1-26.
- Strøm, K. (2000), 'Delegation and Accountability in Parliamentary Democracies.' *European Journal of Political Research* 37(3), 261-289.
- Tallberg, J. (2000), 'The Anatomy of Autonomy: An Institutional Account of Variation in Supranational Influence.' *Journal of Common Market Studies* 38(5), 834-64.

- (2002), 'Delegation to Supranational Institutions: Why, How, and With What Consequences?' *West European Politics* 25(1), 23-46.
- Tirole, J. (1986), 'Hierarchies and Bureaucracies: On the Role of Collusion in Organizations.' *Journal of Law, Economics and Organization* 2(2), 181-214.
- (1992), 'Collusion and the Theory of Organizations.' In J-J. Laffont (ed.), *Advances in Economic Theory*. Cambridge: Cambridge University Press, 151-206.
- (1994), 'The Internal Organization of Government.' *Oxford Economic Papers*, New Series, 46(1), 1-29.
- Tsebelis, G., and G. Garrett (2001), 'The Institutional Foundations of Intergovernmentalism and Supranationalism in the European Union.' *International Organization* 55(2), 357-90.
- Whinston, M. (2001), 'Assessing the Property Rights and Transaction-Cost Theories of Firm Scope.' *American Economic Review* (Papers and Proceedings), 91(2), 184-88.
- Whitehead, A. (1927 [Rep. 1979]), *Process and Reality*, New York, Free Press
- Williamson, O.E. (1963), 'Managerial Discretion and Business Behavior.' *American Economic Review* 53(5), 1032-57.
- (1964), *The Economics of Discretionary Behavior: Managerial objectives in a theory of the firm*. Englewood Cliffs (NJ): Prentice-Hall.
- (1967), 'Hierarchical Control and Optimum Firm Size.' *Journal of Political Economy* 75(2), 123-38.
- (1968), 'Economics as an Anti-Trust Defense: The Welfare Trade-Offs.' *American Economic Review* 58(1), 18-36.
- (1971), 'The Vertical Integration of Production: Market Failure Considerations'. *American Economic Review* 61(1), 112-23.
- (1973), 'Markets and Hierarchies: Some elementary considerations'. *American Economic Review* 63(2), 316-25.
- (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.
- (1983), 'Credible Commitments: Using Hostages to Support Exchange.' *American Economic Review* 73(3), 519-40.
- (1985), *The Economic Institutions of Capitalism*. New York: Free Press.
- (1991), 'Comparative Economic Organization: The Analysis of Discrete Structural Alternatives.' *Administrative Science Quarterly* 36(2), 269-96.
- (1996), *The Mechanisms of Governance*. New York: Oxford University Press.
- (2000), 'The New Institutional Economics: Taking Stock, Looking Ahead.' *Journal of Economic Literature* 38(3), 595-613.
- (2002), 'The Lens of Contract: Private Ordering.' *American Economic Review* 92(2), 438-43.
- Wilson, R. (1969), 'The Structure of Incentives for Decentralization under Uncertainty.' In G. Guilbaud (ed.), *La Décision: Agrégation et Dynamique des Ordres de Préférence*. Paris: Centre National de Recherche Scientifique, 287-307.
- Zeckhauser, R. (1970), 'Medical Insurance: A Case Study of the Trade-off between Risk Spreading and Appropriate Incentives.' *Journal of Economic Theory* 2(1), 10-26.