Title:

ANALYSING ECONOMIC REGULATION THROUGH INSTITUTIONS, FINANCE AND PUBLIC LAW

Candidate: Javier Tapia
University College London
PhD

DECLARATION OF AUTHENTICITY

‘I, Mr. Javier Tapia confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Signed: ……

Date: ……

The thesis analyses economic regulation primarily through the lenses of institutions, finance and public law. In terms of positive analysis, it focuses on post-privatisation developments in the UK.

The first chapter sets out the framework of analysis. Traditional models to study economic regulation are critically assessed. It is argued that these frameworks make only partial assessments of regulation under simplified assumptions. However, it is possible to find complementarities that may serve as fundamentals for further study.

Chapter II proposes to broaden the theoretical framework of analysis in three directions, with the aim of providing guidance on how to structure regulatory design in an interdisciplinary context.

The third chapter deals with financial implications of regulation. The case of regulation of the capital structure of utilities firms and the regulatory assessment of the cost of capital is specifically studied in order to exemplify the control of risks under regulatory practice.

In the fourth chapter the argument is extended to the analysis of the role of the State in infrastructure industries and the significant function that risk plays in economic regulation. It is argued that regulators should control focus their analysis on risk control and the avoidance of the provision of State guarantees as much as possible.

Finally, chapter V looks at the role of the courts and the judicial scrutiny of sector-specific economic regulators' decisions. Since the introduction of specialisation, judicial scrutiny is having an ever-increasing influence on substantive regulatory decisions in the UK. Despite this, some flaws in the regulatory design are highlighted, along with proposals to overcome them.

The thesis concludes with some implications for policy design and the analysis of economic regulation.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

My greatest debt is to three persons at UCL. Not only did they read and comment on many embarrassing drafts; they also had a profound impact on my understanding of economic regulation and competition law, and helped inspire many of my central concerns and claims. First and foremost, I would like to thank Tim Swanson, who went far beyond what one could expect from a supervisor and, in the process, became a mentor, role model and friend (he is also a very decent tennis player, I dare say). Tim constantly challenged my ideas, pushing me to think hard on the various issues revolving around law and economics. I feel extremely lucky to have been his supervisee. I hope one day I will be able to pay back all that he has taught and done for us – at least by teaching him how to prepare a proper barbeque! Secondly, Florian Wagner-von Papp – who acted as my second supervisor – was always very willing to assist me. His razor-sharp comments polished and improved my viewpoints. Finally Ioannis Lianos who, besides becoming a very good friend, acted in practice as a sort of third supervisor. He was extremely generous in providing me support, opportunities and new ideas, as well as introducing me to a vast number of people. It was my great pleasure to work with these phenomenal people and I look forward to doing so again in the near future.

Important amendments and new ideas were added after the viva. Professors Rick Rawlings and Martin Cave were kindly enough not only to comment on the overall approach of the thesis, but also provided detailed comments on several parts thereof. The final version owns much of its improvement to both of them. There were also several others – academics, practitioners and friends – who generously agreed to comment on several parts of this thesis at different stages of its development. Most of them went far beyond what I could have reasonably envisaged, setting aside large parts of their valuable time to discuss my ideas in great detail. It is therefore a pleasure to thank them by name: David Bailey, professor Paul Craig, Kieran Donohue, Shaun Larcom, professor Stephen Littlechild, George Lusty, Ruth McCarthy, professor Michael Pollitt, professor Tony Prosser, professor Rick Rawlings, Dr. Arad Reisberg, Aidan Robertson QC, Dr. Daniel Sokol, and Dr. Adam Scott OBE TD. I also benefited from more general conversations with professor Martin Cripps. Claudia Perez and Ruth McCarthy were extremely helpful in the last part of the process, dealing with the huge

task of correcting my unavoidable idiomatic infelicities. Indeed, none of these people bear any responsibility for my unfortunate errors, nor do they necessarily agree with my arguments.

The writing of this thesis was largely completed while I was working on a part-time basis at Ofgem, the UK energy regulator. I had the chance to work on both a policy team and in the legal department. This provided me with the priceless experience of being involved in regulatory practice within a common law system. Since I already had some familiarity with working in civil law regulatory regimes, this opportunity vastly enhanced my understanding of the regulatory complexities. At Ofgem, I especially benefited from the enormous experience of David Ashbourne, Michael Brocklehurst and John Wilson. More generally, however, many other colleagues helped me mould my views – without even noticing! They are too many to mention here without incurring the serious risk of embarrassingly forgetting someone, so I thank them all collectively.

I am also grateful to the participants of stimulating conferences in Toronto, Brussels, Trier, Rome, Istanbul and Las Palmas de Gran Canarias, where I had the opportunity to present some of the ideas developed in the thesis. Likewise, the reactions and advice from students enrolled at both the Law Faculty and the Public Policy school at UCL during the past three years were of immense benefit.

In addition, I would like to thank the scholarships provided by both the Chilean government and UCL, Faculty of Laws.

On a personal note, my beloved parents in Chile helped in many different ways during the gestation of this thesis. I apologise if working on it prevented me from demonstrating all the affection they truly deserve. And last, but certainly not least, my wife Carolina who accompanied me all these years in London, and offered me the so much needed companionship and love that kept me focused on life beyond this work. She also gave me the indispensable encouragement, care and empathy to complete it. For all of this: thank you. The two of us began this journey together; we are three at time of its completion. I cannot envisage a better way to embark on a new adventure.

INTRODUCTION

1. Overview

Recent times have witnessed a renewed interest in regulation. High-profile events such as financial crises, accounting scandals, climate change, and others, have provoked calls for more (or more effective) government action across different sectors. This contrasts with the previous dominant belief in the retreat of the State and its allegedly undesirable controls and interferences. Such belief reflected a true confidence in the deterministic reasoning that the market was capable of producing the fullest possible well-being achievable by individuals in a society. The pendulum seems now to be swinging back. Indeed, the need to avoid the excesses of public intervention in markets is recognised. However, there is an increasing acknowledgement that, ultimately, public policies (including regulatory policies) do not emerge from the action of the invisible hand, but they are deliberately constructed. That is, regulation is essentially a problem of design. When fashioning or reconfiguring a regulatory regime, the designer or policy-maker[1] is presented with a number of possibilities. The designer then makes a self-conscious option for particular contingent features, taking into account various constraints. Thus,

> [M]arket creation entails not merely a *negative* choice to get the state out of the way, but an affirmative act of *policy design*. All markets exist within institutional structures shaped by public policy: with any decision to decollectivize, privatize, or deregulate, policymakers must necessarily make a host of additional choices about how to configure, and perhaps even direct, the new market.[2]

This thesis deals with the basis of regulatory design. Undoubtedly, however, it does not aim to cover the entire field of regulation. The *summa divisio* of regulatory studies distinguishes between economic and social regulation. Only the former is the focus of the current research. Economic regulation 'typically refers to government-imposed restrictions on firms' decisions

[1] The two terms used interchangeably throughout this work.
[2] Jacobs & Teles (2007: 158) (emphasis on the original).

over price, quantity, and entry and exit'.[3] Indeed, market forces retain a significant role in spite of government interventions. But overall economic arguments mainly associated with 'market failures' and historical arguments associated with the political development of industries provided the basis for imposing governmental control on at least parts of them.[4]

The most widely accepted market failure is natural monopoly, which provides the rationale for regulating at least some segments of a number of 'utilities' industries – typically, electricity, telecommunications, water and gas. However, the terminology is somewhat ambiguous. For example, the concept of natural monopoly is not always clear: beyond the standard definition, its causes are debated amongst economists.[5] Likewise, the term 'utilities' is no more than a convenient nomenclature that lacks self-explanatory meaning.[6] In fact, on occasion economic regulation also extends to other industries that do not necessarily respond to a natural monopoly rationale (such as airports), or are not normally considered utilities (such as the rail industry). All these industries, utilities or not, are also commonly referred to as 'infrastructure industries'. As a result, any effort to make determinative categorisations normally ends up offering mere reasonable, common sense distinctions. Therefore, a precise definition is of no concern here. The focus on economic regulation is justified by the distinctive set of questions the subject poses.

As part of the reinvigorated general interest in regulation, the fundamentals of economic regulation in the UK – which is the main focus of this thesis – have also been affected.

[3] Viscusi *et al.* (2000: 207). They recognise that on occasion governments control other variables such as quality and investment.

[4] E.g. Breyer (1982: 15-35). On the basis of regulation, see Chapter I.

[5] The standard definition of natural monopoly indicates that the entire demand within the relevant market is satisfied at the lowest cost by one single firm, rather than two or more. However, the causes are debated. The classic literature emphasises economies of scale. See e.g. Kahn (1988: 123/I) ('a "natural monopoly" is an industry in which the economies of scale –that is, the tendency for average costs to decrease the larger the producing firm- are continuous up to the point that one company supplies the entire demand'). Other authors emphasise the 'subadditivity of costs': that the production of all combinations of outputs must be accomplished at least cost by a single firm. Subadditivity characterises the natural monopoly cost function, and it is generally due to *both* economies of scale and economies of scope (especially in the case of a multi-product firm). See e.g. Viscusi *et al.* (2000: 339-44) (noticing at 341 that 'subaddivity is the best way to define natural monopoly'). Finally, others stress the importance of sunk costs in long-term relations. See Gómez-Ibánez (2003: 9) (arguing that 'economies of scale are arguably less important than durable and immovable investments in establishing the barrier to entry in natural monopoly').

[6] See e.g. Troxel (1947: 4) (indicating that 'Regulation is sanctioned when a further breach in the solid front of the traditionally private enterprise seems tolerable'); Bonbright *et al.* (1987: 6) ('the term "public utility" is one of popular usage rather than of precise definition').

Recent events confirm this assertion. For example, the energy regulator, Ofgem, has now formally abandoned the well-known RPI-X formula for calculating the final price of gas and electricity, favouring an alternative model.[7] Despite being subject to important revisions and amendments throughout the years, RPI-X has been the model constantly applied to industries subject to economic regulation since privatisation – and indeed still remains in use in other sectors. Likewise, a recent review of competition and innovation in water markets (the 'Cave review') recommended changes to the regulatory and legislative frameworks of the industry to deliver more benefits to customers and the environment.[8] The review found weak incentives for innovation, with the framework focused more on productive rather than allocative and dynamic efficiency. Finally, the government is revising the role of many regulators with the prospect of altering the current institutional design.[9] The outcomes are due in early 2011. Together, these trends threaten to become a movement just as important as the widespread wave of privatisation-cum-regulation witnessed during the late 1980s and early 1990s.

The question has become a query for the extension of the revision of the basis of the system. As it is further explained in the thesis, this need comes from a reverse swing toward 'smarter' economic regulation (borrowing a term from broader regulatory studies) that has not been matched with new developments in regulatory theory. Unfortunately, traditional legal theories and economic approaches do not adequately provide for the type of remedies currently needed and sought by the regulatory practice.[10] Not only are they based on assumptions that do not take into account recent developments in the structure of the industries subject to economic regulation; they have also failed to consider the broadening of regulatory objectives and the consequential enlargement of actors involved in regulatory processes. The consequences are likely to be serious – practical expectations may not always be met if changes lack solid theoretical foundations and a deep knowledge of the subject.

[7] See Ofgem (2010a). The regulator is currently taking the first steps into the application of a new model (called RIIO) for the next price controls in transmission/transport and gas distribution.

[8] Cave (2009). See also CST (2009) (assessing and providing a solution on how to improve R&D in the water sector).

[9] E.g. in August 2010 the Government commissioned a review of the economic regulator for the water industry, Ofwat, which includes its objectives, governance, boundaries of responsibility, statutory duties, effectiveness of the current arrangements, etc. Likewise, in July 2010 the Government launched the Ofgem Review, which will explore whether any changes are needed to the regulatory framework to enable achievement ofgovernmental goals.

[10] Rose-Ackerman & Rossi (2000:1437).

Without said knowledge, the danger lies in that 'the new round of regulation will be rooted not in new research and new thinking, but rather in old ideas that are conveniently dusted off and reused in the absence of anything better'.[11]

Indeed, this is not a call to demolish the intellectual foundations of economic regulation – traditional frameworks of analysis have made numerous important advances that must be judged within their proper context.[12] It is more a recognition that, in time, some assumptions need to be replaced by others informed by recent experience, and the study of the subject needs to be redefined and moved forward based on an improved understanding.

In this context, if one were to rethink the overall basis of economic regulation, what should be the focus of regulatory design? Established law and economics tradition instructs that the regulatory regime (as any other legal rule) should aim to create incentives for firms to behave efficiently and allocate risks efficiently amongst them, taking into account the relevant trade-offs.[14] However, this is insufficient. Economic considerations are not the only ones that are embedded in regulatory regimes – even in areas known as 'economic regulation'. There is a need to recognise explicitly that social, moral or political values also play an important role. From this recognition, several questions for regulatory design arise. To name a few, what should be the institution in charge of dealing with conflicting regulatory goals? How conflicts should be resolved? Likewise, practical developments point towards a rethinking of the traditional roles of stakeholders in the regulatory process. Again, regulatory design needs to determine what implications of this change are. Finally, collaboration is an important characteristic that affects the relations between different actors in the regulatory process. Nonetheless, it is a feature that has been neglected by traditional analysis. Overall, there seems to be good reasons to expand the framework of analysis of economic regulation at least considering these three crucial aspects.

[11] Balleisen *et al.* (2010: 1).

[12] '[R]egulation cannot properly be said either to have "failed" or "succeeded" in an overall historical sense. Instead, individual regulatory experiments and episodes must be judged against a standard true to the particular historical moment' (McCraw, 1984: 308).

[14] E.g., Polinsky (1983).

There are important consequences arising from the expansion of the analytical framework. The first one is related to regulatory risk. From the point of view of economic regulation, risk can be seen as two-fold. First, it can be seen as a specific concept applicable to the financial aspects of economic regulation. With notable exceptions, theoretical accounts tend not to consider these specificities, which are left to financial treatments or even to practical developments. However, a deep understanding of the financial aspects of regulation is crucial. Not only finance is at the core of regulatory practice, but it also influences the way regulatory incentives are received by the firms. Second, risk can be seen from the perspective of the regulatory regime as a whole. In this sense, it is a concept related to the broader incentives and goals of the system and, as such, it has been subject to a number of studies by commentators. As will be shown, both perspectives are influenced by goals that go beyond economic considerations, are the product of interactions between several actors and are the outcome of at least partially collaborative activities.

A second consequence is that the role of both regulators and courts has become less apparent. Unlike previous stages of development of economic regulation, where objectives were somewhat clear and generally economic-oriented[15], the plurality of objectives and actors has opened further areas of concern that demand both new regulatory tools and new ways to deal with them. Even 'core' regulatory issues, such as the determination of the cost of capital or the assessment of risk, possess new angles that challenge regulators and policy-makers alike. Nonetheless, it seems like regulators keep acting proactively only in the face of traditional concerns (which appear reasonably clear and manageable), whilst adopting a much more reactive role when confronting new ones. Courts, on the other hand, have traditionally played a minor role in economic regulation – at least in the UK. However, the increase in their specialisation has placed them in a privileged position that allows them to surmount regulatory inertia.

Certainly, a thesis can merely start addressing the aforementioned and other pressing issues, and hence leaves many crucial questions unanswered. However, as the number of recent sectoral revisions initiated by regulators and the government alike suggests, the problems

[15] Although 'political compromises had to be made, and the legislation…contained a number of duties mixing economic and social rationales' (Prosser, 2005a: 68).

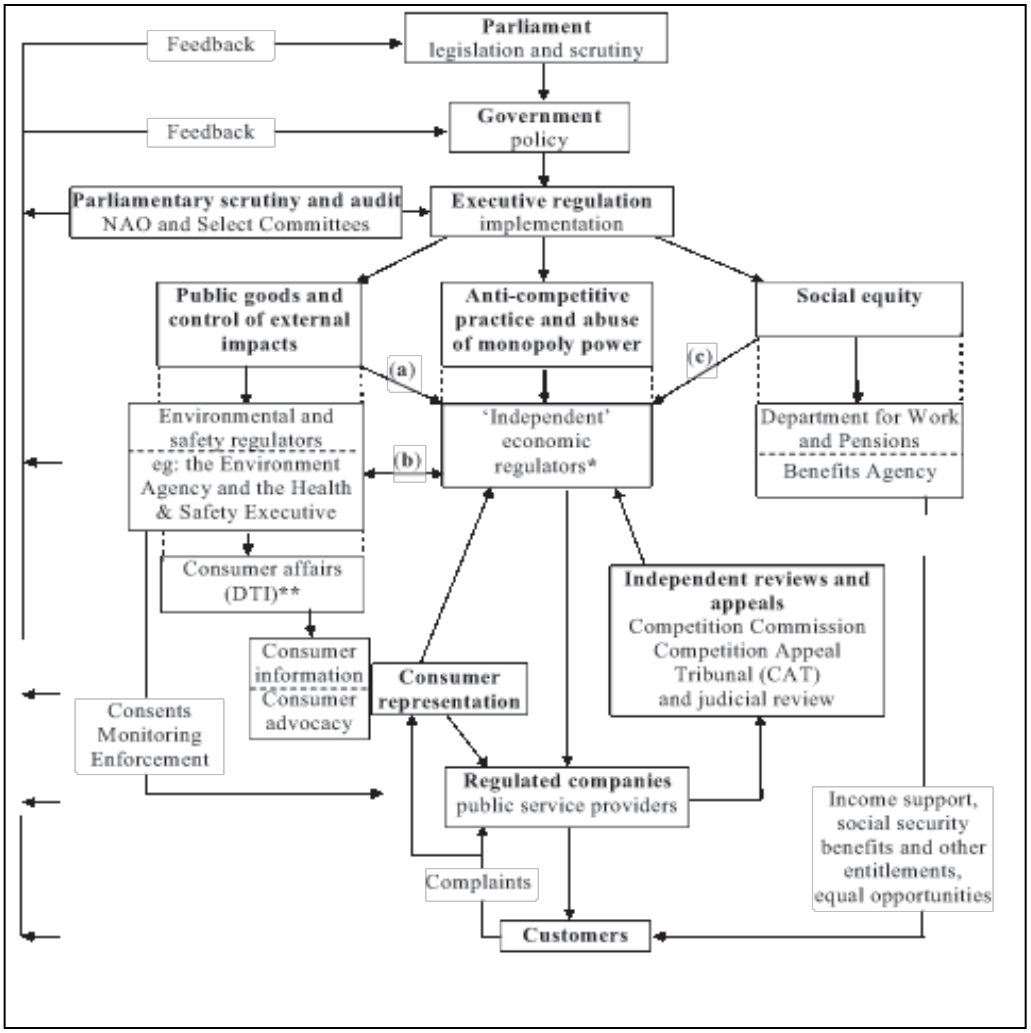
encountered at the current stage are not the same as those that were common under the old trend of integrated monopolies and/or State-owned firms, nor during the first ages of privatisation. To a large extent government failures and deregulation of markets remain at the core of the current thinking of many academics and practitioners in this area. The framework of analysis advanced here aims to push the boundaries of current thinking, move away from traditional views, and perhaps partially shape the intellectual environment in which policy-making is framed and implemented.

2. Scope

As mentioned, in terms of positive analysis the thesis focuses on post-privatisation developments in economic regulation in the UK. The evolution and key aspects of the industries subject to economic regulation are emphasised, as well as the main characteristics of some administrative procedures. In particular, the thesis deals mainly with the work of sectoral economic regulators of water (Ofwat); energy (Ofgem)[17]; communications (Ofcom), but only with respect to the regulation of telecommunications; rail (ORR); and airports (CAA). The work of the Competition Commission (CC) and the Competition Appeal Tribunal (CAT) in the area of economic regulation is also included. Figure 1 summarises the institutional design of the UK regulatory framework in economic regulation.

[17] Legally, the energy regulator is the Gas and Electricity Markets Authority (GEMA). Ofgem, which is the office that supports GEMA, lacks legal status. In this thesis Ofgem and GEMA are used interchangeably.

Figure 1: *The UK regulatory framework*



Source: HL (2004a: Appendix 5).

Most case studies and examples included in the thesis have been selected from the aforementioned industries. In addition, some examples have also been taken from the London Underground rail network. The underground is an industry not traditionally considered within the field of economic regulation, although it is normally included within the railways sector in the UK.[18] However, both its regulatory characteristics and recent developments are sufficiently close to the infrastructure industries to justify its inclusion at least as an illustration of some points. None of these industries is studied in detail, however. It is true

[18] The UK rail industry is formed by four sectors: the mainline network, the underground, light rails and tramways, and the so-called minor and heritage railways.

that every industry possesses certain underlying characteristics (economic or of another nature) that make it rather unique and different from others. Moreover, these characteristics often shape governmental action. However, for most industries subject to economic regulation the differences seem subtler rather than radical. For that reason, the thesis emphasises commonalities. This is also consistent with the normative claim explained in the next part.

3. Approach

As noted by others, regulation is a broad subject that calls for a multidisciplinary approach that encompasses a number of different perspectives.[19] Recognising that inter-disciplinarity is difficult to achieve, this thesis is an effort to study economic regulation primarily through the lenses of institutions, finance and public law. Given space constraints and the need to remain focused, the appraisal of economic regulation, even within these three sub-disciplines, is necessarily incomplete. From the institutional perspective, institutional design is the topic that underlies the analysis carried out in this work and encompasses all parts thereof. Its normative claim is that only by undertaking comparative institutional analysis is it possible to properly structure the institutional design of regulation.[20] The focus is on governance architectures and the processes and formal rules that channel political and economic activity – I have not dealt in detail with informal norms. The financial assessment is largely centred on specific parts of price controls, particularly those related to the capital structure and the cost of capital. Finally, the public law area focuses on a specific part of accountability; namely, the judicial scrutiny of regulatory decisions.

However, it is not possible to achieve a full understanding of regulatory evolution without also paying attention to other disciplines. In addition to the three core perspectives, political science, sociological and psychological literature help to further clarify the various complexities of the subject. Also, on occasions, partial historical accounts of the evolution of

[19] E.g. Baldwin *et al.* (2010: 11-13); Baldwin & Cave (1999: 1).

[20] By comparative institutional analysis, I utilize Oliver Williamson's conceptualization as 'an examination of the comparative costs of planning, adapting, and monitoring task completion under alternative governance structures' (WILLIAMSON, 1985: 2).

economic regulation in the UK shall serve as background. Last, but certainly not least, the modern study of economic regulation requires emphasising, at least implicitly, the importance of competition. Particularly in the UK context, competition law and regulation have been two areas practically inseparable since privatisation. However, within the broad subject of competition law, one aspect that is not considered in the thesis is merger control. There are two reasons for this. On the one hand, merger is a vast subject in itself, and hence would require a separate treatment. On the other, developments in merger control and regulation have taken rather opposite directions. Whereas regulation has moved from being centred exclusively on competition towards considering a plurality of social, economic and environmental objectives; merger control has moved from public interest considerations (which included a plurality of objectives) to a competition-centred analysis.

Furthermore, other concepts used throughout this thesis are contested in the literature – some even widely acknowledged. Nonetheless, for the most part ambiguous or disputed concepts have no consequences for the treatment of the topics in the subsequent chapters. One exception is perhaps the 'regulatory decision'. This term, often repeated throughout the thesis, is used very broadly as referring to 'any action or inaction of the designate regulatory body or another government entity that affects the economic interests of participants in the sector'.[21] As such, the decision can take different forms, and be either formal or informal. Also, the objectives of regulation are not specified; this is a task that depends on each regulatory framework. For purposes of analysis, however, it may be considered that the overall objective of any privatised regulatory regime is to attract investment[22], and that such objective is placed above others instrumental to investment attraction. Beyond these clarifications, the context should suffice to elucidate the meaning of other contested concepts.

[21] Brown *et al.* (2006: 151).

[22] Recognising long-term investment as the main reason to privatise, see Rose-Ackerman & Rossi (2000: 1437 and note 6); Levy & Spiller (2006).

4. Outline of chapters

Chapter I sets out the general framework of analysis by contextualising economic regulation in a broader theoretical and practical perspective. It begins by describing the general changes in the structure of industries subject to economic regulation, comprising both the ownership pattern and the competitive nature of the services. The changes are now quite well-known, and include privatisation, the advance of competition and diversification of services. The chapter then shows how these advances have been coupled with the expansion of the number of actors involved in the regulatory process. Afterwards, the chapter introduces the main traditional frameworks used in the literature for analysing utilities regulation – namely, principal-agent models, institutional analysis and public-choice accounts. It is argued that these traditional frameworks make only partial assessments of the regulatory relation under simplified assumptions. Three stand out: the focus on a single regulatory objective (efficiency, achieved through competition); the preference for the market as an institution of choice (normally within the market-hierarchy dichotomy); and the use of neoclassical economic ideas (especially those associated with the rational 'economic man' and the disregard for cognitive and behavioural constraints). However, it is possible to find ways for complementarities between the three main groups of theories that may serve as fundamentals for an improved and more up-to-date analytical framework.

Chapter II proposes an expanded analytical approach, which draws upon previously developed theoretical frameworks, but explicitly incorporates some facets absent in traditional models. The approach comprises three central elements: the acknowledgment of the multiple objectives that economic regulation may pursue, the recognition of the plurality of actors that participate in the regulatory process, and the central role of cooperation. These elements jointly address the main features of regulation identified in the first chapter and further shown in the next three chapters. The last part assesses the consequences – particularly, that the application of the framework results in a different theoretical approach to regulation and a more 'horizontal' view.

Chapters III, IV and V are devoted to specific aspects of the UK regulatory regime. Chapter III uses the case of the regulation of the capital structure of utilities firms to exemplify how firms respond to the incentives of the regulatory regime in different ways. An understanding of the 'old' interactions between capital structure, the cost of capital, and economic incentives within regulated environments is critical to assess how the system has performed and provided avenues for improvement. The chapter begins by describing the changes in the capital structure of UK utilities firms since privatisation and providing possible explanations for them. It is argued that the most plausible justifications are those that go beyond financial approaches and take into account specific features of the regulatory regime. Then, the main regulatory responses to the changes in firms' capital structures are critically assessed. The different means of influencing the capital structure are a response to various concerns related to high gearing. It is submitted that some of these concerns are unfounded and that the effectiveness of various aspects of the regulation of capital structure and the setting of the cost of capital are highly questionable. Chapter III concludes by considering the role of equity and debt in industries where revenues are controlled and to a large extent guaranteed. Two technical appendices explain how regulators calculate the cost of capital.

The capital structure of the regulated firms affects the distribution of risk between firms, taxpayers and consumers, and ultimately influences the price paid by the latter. Chapter IV addresses the topic of risk and regulation in a more general manner. Regulators should focus their tasks on risk control (especially the risk of default, against which they have a duty to protect firms), on the one hand; and the avoidance of the provision of State guarantees and the production of unintended effects as much as possible, on the other. The chapter defines four types of risks and presents the rules for their allocation. Subsequently, how they operate in the different regulatory regimes used in UK utilities is explained. The second section of the chapter (the longest) deals with the avoidance of State guarantees. Here, it is shown why guarantees provide negative incentives not only to the parties of the regulatory relation, but also to third parties. Likewise, the chapter also asserts that despite the fact that unintended effects largely depend on regulatory design, they are ubiquitous in utilities regulation and therefore a key regulatory task is to avoid them as much as possible. The chapter concludes with some considerations regarding the 'insurance' implications of economic regulation.

The final chapter (V) analyses judicial scrutiny of sector-specific economic regulators' decisions. The chapter is divided into three sections. The first one provides a background with the most salient institutional features of the UK regime since privatisation. This part concludes that the main characteristics of the regime are specialisation and the existence of various forms of appeals and standards of review. These characteristics are revised in the following two sections. Specialisation is assessed both empirically and theoretically. It is shown that unlike the first stages of privatisation, where judicial review played a modest role in UK regulation and courts were deferential to regulators, currently judicial review is having an ever-increasing influence on substantive regulatory decisions in the UK, mainly due to the creation of the Competition Appeal Tribunal (CAT). The CAT is a specialized tribunal whose characteristics of expertise, experience and 'object-specificity' have encouraged a general movement towards less deferential approaches. At the same time, however, the standard of review seems to push towards more deference. The result is a rather flawed design that impacts negatively on regulatory policy, induces procedural formality and may increase regulatory risk. As a way forward, the last section proposes to allow the CAT to apply a more intrusive standard of review.

Overall, the interdisciplinary approach adopted in this work may serve as basis not only for a future agenda of theoretical and empirical study of regulation, but also for practical application.

CHAPTER I

ECONOMIC REGULATION IN CONTEXT

> 'There has grown up a spirit of resistance… to the interference of government, merely as such, and disposition to restrict its sphere of action within the narrowest bounds'.[23]
>
> - John Stuart Mill

OVERVIEW

The foundations of any institutional framework of analysis need to grow out of a deep and sophisticated understanding of the structure of the industries at issue and how that structure evolves. On the one hand, it is vital to appreciate how the relation between regulators, firms and other actors develops in order to explain the basis of any regime. Although every industry is undoubtedly unique, a number of commonalities allow the study of the general features of the industries subject to economic regulation considering a rather universal perspective. On the other hand, it must be acknowledged that regulation is an evolving phenomenon. True, normally the structure is relatively stable. However, it can crucially change over time as the industry evolves, perhaps simply because over time competition tends to eliminate weaker organizational structures.[24] In any event, the understanding of both the general structure of industries subject to economic regulation and how it progresses sets the stage for institutional analysis and regulatory design.

During the last three decades, the analysis of structure has been dominated by the emphasis on the benefits of competition and a preference for the market as the main vehicle to achieve efficiency. Three theories stand out. Respectively, they stress the role of interests with the aim of creating awareness of the dangers of their influence; offer a view on how regulation

[23] Stuart Mill (2004 [1848]: 728).
[24] Alchian (1950: 213-14).

should be organised considering informational asymmetries and propose ways to design appropriate incentives to overcome them; and endeavour to understand how institutions influence performance. Despite their different centres of attention, with more or less prominence most of these frameworks of analysis consider the market as the benchmark institution, whereas the State is largely deemed to be a secondary institution of choice. Also, the studies depart from competition and efficiency as the prominent, if not exclusive, role to be pursued in regulation. They normally assume that agents have rational motivations and perfect optimizing abilities.

Each of these three common characteristics is controversial. First, the State is not confined to establishing a set of rules giving directions; rather, it uses a wider range of more flexible strategies to influence particular activities of firms. It is necessary to pay attention to how these strategies influence the regulatory environment, so that they can be adapted for the need of constantly evolving markets. This means the State plays a role far more important than theories assume. Secondly, interactions between agents are strategic and may result in an asymmetrical distribution of power incompatible with notions of efficiency. Lastly, the models normally (albeit, not always) ignore cognitive problems derived from limited rationality and/or agents' behavioural constraints and downplay the role of collaboration. For all of these reasons, the three groups of explanations for regulation are non-comprehensive frameworks for the analysis of economic regulation.

To some extent, each traditional framework of analysis is a snapshot of one moment in the evolution of the structure of industries subject to economic regulation. That structure is in constant flux. That the frameworks centre on competition and the benefits of markets says as much about the advance of these theories as of the historical development of the industries in analysis. During the last three decades, in many settings private firms have been operating in increasingly competitive markets. The developments are now notorious. They normally comprise of the breaking-up of former monopolies, changes to ownership patterns and the alleged 'relinquishment' of State control. However, as later literature has stressed, they also include the broadening of the concept of regulation to include a wide variety of actors and policies that go beyond the direction of the State but stop short of 'getting rid' of it.

This chapter describes both the main explanations of economic regulation and its evolution. For purposes of clarity, it seems convenient to invert the order of exposition. The exposition begins by summarising the main developments. Then, the second part surveys the main bodies of theoretical research that have provided explanations and/or solutions relevant to economic regulation, and subjects them to critical appraisal.

I. ECONOMIC REGULATION AS AN EVOLVING PHENOMENON

A. The evolution of the structure

For many years the structure of markets subject to economic regulation presented a stable structure. Despite important differences across sectors, the structure was generally similar due to reminiscences of their historical characterisation as natural monopolies.[25] The main features of that market structure have been extensively studied.[26] Fundamentally, these industries evolved as vertically integrated firms that were either state-owned or privately owned (although usually the first one was the case) and controlled every component of the market. In the electricity sector, for example, utilities usually developed their activities across generation, transmission, distribution and retail supply. Also, the telecommunications industry developed under a model in which the same utility controlled both the upstream and downstream segments of the market. Usually *de facto*, such integrated companies enjoyed exclusive franchises (monopolies) over a certain geographic area and provided services to different types of consumers: residential, commercial and industrial. The causes for such a structure mix economic arguments mainly associated with 'market failures' and historical arguments associated with the political development of the industries.[27] As HELM and JENKINSON explain:

> State ownership 'resolved' the conflict of interests between the private and public good; vertical integration ensured that customers bore the risk of upstream sunk investments; and

[25] In contrast, see Posner (1999 [1969]) (stating that natural monopoly is not a necessary justification for regulation).

[26] For a good and comprehensive overview, see Kessides (2004).

[27] Hovenkamp (1984: 1266) (rejecting the pure historical view that monopoly is nothing more than politics).

> monopoly prevented the destructive competition which was widely thought to have pervaded the industries in the 1920s and 1930s.[28]

Similarly, the long-established role assigned to economic regulators (a function normally in the hands of government) was analogous across most industries subject to economic regulation. It comprised at least one, but normally most, of four vital functions.[29] First, they controlled entry and exit. That is, they made the central determination of whether a company would be permitted to enter the industry, a function usually matched with the authority to control exit. Secondly, regulators controlled quantity – i.e., the amount of the product or service offered to the market, or even its nature. Thirdly, regulators controlled prices. They reviewed and approved prices that firms could charge to users of their services. Finally, regulators controlled profits: they determined the maximum level of profits that firms could earn. Each of these functions fitted the structure of the regulated industries, but the pattern of ownership normally determined the regulatory 'style'. Whilst in the US context these functions were carried out by sectoral regulatory agencies, in the UK and most other countries ministerial regulation was the norm.

The process of privatisation-cum-deregulation experienced in large segments of different markets across many countries (including the UK), brought about a redefinition of the regulatory setting. First, in many countries the pattern of ownership evolved towards privatisation.[30] Secondly, regulatory functions were separated from central government and put in the hands of myriad independent regulatory bodies.[31] Finally, competition was deemed the best approach to achieve efficiency and provide long-term benefits for society – market

[28] Helm &Jenkinson (1998: 1).

[29] See Kearney & Merrill (1998: 1359-61). See also Troxel, (1947); Viscusi *et al.* (2000: 298-300); and Spulber (1989: 23, referring to 'old style' regulation).

[30] The literature analysing privatisation 'exploded' during the late 80s and early 90s. Explanations for the new trend came from legal, economic and political perspectives. See among others: Vickers & Yarrow (1988) (analysing privatisation from an economic perspective); MacAvoy *et al.* (1989) (presenting studies of the privatisation trend in the UK, Canada and the US); Jones *et al.* (1990) (developing a quantitative approach to make divestiture decisions); Graham & Prosser (1991) (assessing constitutional constraints of privatisation in the UK context); Bös (1991) (providing an economic analysis of privatisation); Targetti (1991) (collecting essays about privatization across Europe); Ramanadham (1993) (containing studies on privatisation in 25 countries); Bishop *et al.* (1994) (a collection of essays about privatisation in the UK); Moran & Prosser (1994) (containing political essays analysing the privatisation process in Europe); Feigenbaum & Henig (1994) (introducing a triple typology of privatisation strategies to explain the phenomenon: pragmatic, tactical and systemic); Savas (2000) (analysing causes and results of privatisation generally).

[31] See, e.g., Thatcher & Stone Sweet (2002).

forces would suffice to induce improvements of quality of service and cost reductions.[32] Most industries saw key changes in their underlying features. The configuration may be summarised in what JOSKOW, in the context of markets for power, has called the 'textbook' architecture of desirable features for restructuring, regulatory reform and the development of competitive markets: private ownership, vertical separation, competitive wholesale markets, competitive retail market, regulation of network services, and creation of incentives for quality of service.[33]

Closely following a neoliberal approach, many scholars extol the virtues of markets and contrasted them with regulatory outcomes.[34] Although only a handful of industries became predominantly unregulated[35], the label of 'deregulation' – that is, the relinquishment of the government's regulatory powers – was generally in vogue. As a consequence of this process, the relationship between firms and regulators and their respective roles changed. Because of deregulation, many authors envisaged a much more limited role for regulators, largely centred on the creation and maintenance of competitive markets. This view is summarised in the now classic legal account of the US evolution made by KEARNEY and MERRILL, who argued that:

> Under the new paradigm, the regulator plays a far more limited role. Instead of comprehensively overseeing an industry in order to protect the end-user, its principal function is to maximize competition among rival providers, in the expectation that competition will provide all the protection necessary for end-users. Specifically, the regulator is expected to intervene only when there is some reason to conclude that a regime of market-based transactions will not suffice to advance competition […] the agency becomes a limited-jurisdiction enforcer of antitrust principles, applying a version of the 'essential facilities' doctrine in a single industry.[36]

[32] Kearney & Merrill (1998).
[33] Joskow (2008).
[34] The belief extended way beyond regulation. See Fukuyama (1992: esp. at 211) (stating that 'at the end of history, there are no serious ideological competitors left to liberal democracy').
[35] In the US, for instance, airlines became deregulated after the Airline Deregulation Act of 1978 was passed. Whilst keeping a standard of 'just, reasonable and non-discriminatory' rates, the act eliminated the duty of carriers to file their rates in tariffs with a government regulator (49 U.S.C. app. §1551[a][2][A] [1988]).
[36] Kearney & Merrill (1998: 1361).

Likewise, it is now widely acknowledged that the intellectual basis of the UK privatisation process can be found in the economic thought of the Austrian school.[37] For that school, competition is a discovery process rather than a state. The economic problem is seen not as one of efficient allocation of resources, as neoclassical economists thought, but one of continuous discovery of preferences, techniques and resources.[38] Accordingly, 'The danger embedded in any regulatory regime, according to the Austrian position, is not so much that it may distort the allocation of resources (the standard neoclassical view) but that it may fail to simulate, and instead stifle (or channel along wholly superfluous routes) the discovery process of the open market'.[39] Repeating the oft-quoted passage of LITTLECHILD in his report that set the basis for the regulation of British Telecom (BT) (the first utility to be privatised and regulated):

> Competition is indisputably the most effective means – perhaps the only effective means – of protecting consumers against monopoly power. Regulation is essentially a means of preventing the worst excesses of monopoly; it is not a substitute for competition. It is a means of "holding the fort" until the competition arrives. Consequently, the main focus of attention has to be on securing the most promising conditions for competition to emerge, and protecting competition from abuse. It is important that regulation in general […] does not prejudice the achievement of this overall strategy.[40]

However, after only a decade or so the triumph of competition already appeared somewhat overstated. The role of regulators remained relevant and certainly much more complex than the relatively simple task of an antitrust court. Regulators are not simply limited to 'setting the background rules that help define an industry structure'.[41] Often, they also play the game. Whilst it remains true that the central tasks of regulators may have changed slightly, in many

[37] Stephen Littlechild, widely credited for being the 'architect' of the telecommunications privatisation regime, was an admirer of Austrian ideas. See, e.g., Littlechild & Owen (1980). Others have also acknowledged this point. E.g. Parker (2009: 276): 'as a free market economist of an "Austrian" persuasion Littlechild was instinctively against controlling profits'.

[38] Hayek, for example, considered 'competition systematically as a procedure for discovering facts which, if the procedure did not exist, would remain unknown or at least would not be used'. See Hayek (2002).

[39] Burton (1997: 160).

[40] Littlechild (1983: 7, para 4.11). By the same token, see Beesley & Littlechild (1997 [1983]: 31), stressing that the regulatory regime 'is inappropriate if competition is not expected to emerge … It is a temporary safeguard, not a permanent method of control'.

[41] Kearney & Merrill (1998: 1361).

cases major interventions (even greater than those needed before) are still required.[42] There may be various reasons for this. The complexity of traditional areas of regulation has sharply increased. For example, there can be little doubt that regulators still play a large discretionary role during many stages of price controls, which remained in place and became more and more burdensome.[43] More importantly, however, new regulatory objectives (sometimes remotely related with the traditional core regulatory areas) have enlarged the scope of regulatory concerns, as will be described in more detail in the final chapter.

B. Beyond the State (but with the State)

Besides the appreciation of the importance of the State that arises from the evolution of the structure, there has been an increasing recognition that the main regulatory interactions are not centred exclusively on the State (or the regulator) and firms. Rather than being a two-tier hierarchy, a number of other entities also interact strategically in a specific domain of regulated activities, giving shape to the outcomes. The incorporation of various types of participants is captured in the view of the regulatory process as a 'stakeholder approach' – considering stakeholders as 'any group or individual who can affect or is affected by the achievement of the firm's objectives'.[44] The main insight of such an approach is that no group has priority over the other. The evolution of an industry is a function of the changes in the relative strategic position of any of those on whom the firm's activities have an impact.

Both the recognition that markets are not the antithesis of regulation, on the one hand, and that regulation encompasses a plurality of actors, on the other, have been well-captured in contemporary political science literature devoted to the study of regulation from a more general perspective (i.e., beyond economic regulation). From the 'rise of the regulatory state'[45] to 'decentred regulation'[46] and 'nodes of governance'[47], different characterisations

[42] Vogel (1996).

[43] In the UK,such a situation was clear almost from the outset of the privatisation process. In the context of the telecommunications industry, e.g., Armstrong *et al.* (1994: 170) explain: 'In practice the DGT's (currently Ofcom) ability to change the license, for example, when reviewing the X factor, gives him a much more active role'.

[44] See Freeman (2010: 25 and 52). See also Freeman *et al.* (2010).

[45] Majone (1994). See also Scott (2004: 145); and Moran (2003).

have tried to capture and typify what is, largely, the same phenomenon. Recently, the developments have been summarised in the most acute concept of 'regulatory capitalism'.[48] As it so precisely synthesises, markets are merely one more regulatory mechanism that joins a plethora of other state and non-state forms of regulation, forming what BRAITHWAITE calls a 'hybridity between the privatisation of the public and the publicisation of the private'. Regulation is now diffuse and fragmented. The concept of 'governance' entails a wide set of control activities and policies that go beyond the direction of the State, but that reserve it a dominant place.[49]

C. The lack of a comprehensive framework to analyse economic regulation

To summarise the previous description, the advances of the structure can be viewed chronologically in three stages: 'monopoly', 'transition' and 'normalisation'.[50] On passing from one stage to the other, some form of regulation is replaced by another.[51] The regulatory role of the State, however, remains equally important. On the other hand, there is recognition of the need to understand regulation beyond the State. It is possible to apply a number of alternative modes of governance that mainly depend on the characteristics of the industry at issue.

Economic regulation, however, seems to have been excluded from this story – or at least lags way behind it. To a non-expert observer, it may seem as if the analysis has remained frozen. In 1993 LAFFONT & TIROLE indicated that '[i]t is fair to say that despite much theoretical progress in the recent past, the field of regulation resembles that of industrial organization of the 1970s'.[52] Almost 20 years later, the situation does not seem much improved. Most of the work produced during the past two decades in the field remains largely technical and

[46] Black (2001).
[47] Shearing & Wood (2003).
[48] Braithwaite (2008). See also Levy-Faur (2005) (coining the term) and Jordana (2005).
[49] Stone-Sweet (1999) defines governance as the processes through which the rule systems in place in any human community are adapted, on an ongoing basis, to the needs and purposes of those who live under them.
[50] The classification comes from Alexiadis & Cave (2010: 504-5).
[51] See Bonbright *et al.* (1987: 6) (arguing that 'no industries are permanent utilities but evolve through a life cycle of no regulation, then some regulation, then great regulation, and finally "ashes to ashes" back to no regulation).
[52] Laffont & Tirole (1993: xvii).

problem-specific (focusing primarily on the welfare implications of pricing and cost rules). Despite the fact that in practice it is possible to perceive the same kind of phenomena underlying the 'regulatory capitalism' claim (i.e., the growth of non-state regulatory governance without necessarily reducing the importance of the State), the analysis is still attached to classical models of hierarchical regulation.[53] At the same time, the literature that has tried to overcome hierarchy simply does not take into account the technical developments made by studies working at the more 'micro-analytical' level, opting instead for providing mere 'high level' solutions. As submitted in the next part, there is no comprehensive, up-to-date framework for analysing economic regulation.

II. TRADITIONAL FRAMEWORKS FOR ANALYSING ECONOMIC REGULATION

The history of efforts to explain economic regulation is long. Particularly over the last thirty-five years or so, there has been no shortage of explanations for why the State regulates. There are myriad varieties of regulatory theory even if only economic regulation is considered. Yet these theories are often conflicting or aim to answer different sets of sub-questions. To a large extent, each theory seems to answer the main question from a different perspective, so it remains unclear what the focus should be. That there has been so much debate and so many different approaches to the same phenomena in the first place suggests that the subject is either not completely understood and/or the analysis remains partial.

Some theories centred on explaining regulation mainly by stressing the role of interests, with the aim of creating awareness of the dangers of their influence. However, the influence that interests exert has been overplayed, with many studies showing that it is not sufficient to explain changes in regulatory settings. Others argue that the focus should be on external constraints, and endeavour to understand how institutions influence performance. Yet they fall short of providing a solution to the regulatory problem. A final popular theory aims to offer a view on how regulation should be organised. These studies look at the regulator-regulatee relation, mainly by highlighting the problems that arise from informational asymmetries and proposing ways to design appropriate incentives to overcome them. But

[53] There are some exceptions: Emmons (2000) (recognising the role of governments beyond 'deregulation').

these views completely downplay the role of cooperation and overemphasise rational motivations for individual and collective action. Clearly, none of these explanations, any more than other sub-varieties that have been put forward, are fully satisfactory. Each contains some truth; but they do not provide a comprehensive form for the analysis of economic regulation as they supposedly aim to do so.

This section surveys the main groups of explanations, generally called public interest theories, public choice analysis (or private interest approach), principal-agent analysis, and institutional approaches. It also presents the main criticisms that have been directed against them. The description is biased toward the aspects that will be used in the next chapter as the basis of a proposal on an alternative framework of analysis.

A. The confusing 'public interest' idea

1. British origins, American developments

The so-called 'public interest theory' of regulation is the most traditional approach applied to government interventions. According to this idea, regulation seeks the benefit of the public at large. That is, regulation would make society better off. Although this idea is as old as the concept of government intervention and, as such, it has been present since long ago in the political (and philosophical) areas, its legal origins are judicial. However, as this revision shows, judicial views and political approaches have not always developed in line. This is clear from the study of the public interest idea from a comparative perspective. Whilst judicially the concept has British precedents, it is in the US where the public interest idea obtained more judicial and political support. Therefore, it is relevant to at least analyse briefly the main developments in both countries.

The legal discussion regarding public interest in 'public utilities' regulation seems to have its origins in an old British case: *Allnutt v. Inglis* (1810).[54] The case was about monopoly pricing. A monopoly granted with a parliamentary licence had denied service to a customer

[54] *Allnutt v. Inglis*, 12 East 530 (1810).

reluctant to pay the price because the latter was allegedly 'unreasonable'.[55] The court based its decision in Lord Matthew HALE's previous work *The Portibus Maris* (1787), and held that if there was only one public wharf in a port, the duties it charges for wharfage, carnage or other services should be 'reasonable' and 'moderate'. If a wharf, crane or other facilities – Lord HALE concluded – enjoy a monopoly licensed or chartered by the King, they were 'affected with a public interest' and therefore the business become *juris publici* rather than *juris private* – which means that the interest of the users outweighed those of the firm. This reasoning has its indirect base on 'common callings', according to which those who provide services available to the general public (i.e., exercise a common calling) may charge only a reasonable rate for it.[56] Since the object of the licence in *Allnutt* was to encourage trade and commerce, the activities were hence affected with the public interest and the monopolist had an obligation not to charge excessive tolls, but only a 'reasonable' price for the provision of the services.

The argument in *Allnutt* is an archetypical example of the nineteenth century confusion (or lack of clear boundaries) between law and morality. The recourse to the 'reasonability' seems similar to the common law moral idea of *just price*, which implies a sort of rejection of 'high prices' based on reasons of fairness, distributional inequality and accumulation.[57] Underlying the morals of just price was the idea that in order to keep a fair balance between users and companies, these were entitled only to 'moderate gains' instead of 'excessive gains', which were deemed illegal. The law merely adopted such ideas and applied them to several areas,

[55] At issue was whether the London Dock Company, a firm that possessed a monopoly to receive certain wines, could lawfully exclude from the docks a cargo owner who had refused to pay their schedule of charges.

[56] For an illuminated account of the origins and evolutions of common and 'public' callings, see Arterburn (1927). He explains how a duty to use care originally applied to all trade and callings, but that no remedy was available for those who engage in such activities *occasionally* (as opposed to *common* practice). A duty to serve originated later on due to some conditions (i.e. the Black Death) and also applied to all trade and callings. The obligation of charging reasonable prices evolved due to economic and social conditions, to counteract potential abuses from some monopolistic industries, which became known as common or public callings. There were common carriers, common innkeepers, and common millers, amongst others. Therefore, as Craig (1991) summarises, the origin of the term common calling was 'simply a service that was available to the public generally'.

[57] The origin of the idea of just price lies in the work of Aristoteles. Later on, the concept was importantly influenced by Christian beliefs (such as 'Lend freely, hoping nothing in return' or that one should do onto others what one is expecting others do onto oneself) and particularly the work of Thomas Aquinas, who argued that just price was either the market price or the price fixed by the authority, in both cases limited to an amount that could keep the social position of the merchant.

including the law of trade, price controls[58] and the law of usury[59]. It is not strange, then, that there is just a small distance between this reasoning and Lord HALE's idea of a 'reasonable' and 'moderate' charge for a business so as to keep the balance between the utility of the firm and the utility of the users.

From a political perspective, the public interest idea influenced the way the English institutional system later developed. However, developments did not follow directly from *Allnutt*'s precedent or Lord Hale's reasoning. The ad-hoc legislation of public utilities that emerged from the nineteenth century onwards may have 'obviated at least some of the need for direct, common law regulation of price'.[60] At that time, several specialised bodies took over the control of certain industries, some of which were affected with public interest (e.g., water, gas, lightning, railways, and roads). Economic functions in these industries were performed by large corporations with special statutory authority that enabled them to carry out their tasks. Within their geographic areas, these corporations possessed something akin to what is nowadays known as market power. Some bodies, quite similar to modern 'regulatory agencies', were tasked with the duty of regulate the rates to be charged for the services and supervise the statutory 'tariffs'.[61] In time, 'modern' agencies were even tasked with the decision of cases according to the public interest test.[62] Likewise, the public interest idea was embedded into in many legislative acts regulating public utilities. Several statutes adopted the notion that proponents of regulation should act as agents for the public interest.[63]

[58] Price controls have been used in the British economy since long ago. For example, in 1623 James I ordered that corn prices should be lowered in times of scarcity. Only the Ricardian ideas developed in the nineteenth century would reverse some of these controls.

[59] As early as in the sixteenth century, several criminal offenses were enacted – all of them under the general idea of outlawing excessive gains in trade. For example, manipulation of market prices was severely punished, as well as *ingrossing* (intermediating between the farmer and the consumer), *regrating* (reselling goods in the same market) and *forestalling* (manipulation of market prices before the goods arrived to the market).

[60] Craig (1991: 540). He also advances an alternative, partial explanation for the denial of Lord HALE's arguments: 'Blackstone did not take up Hale's observation in this regard' (*Ibid*).

[61] Indeed, the term 'regulatory agencies' is misleading. Supervision at that time was normally departamental.

[62] *E.g.* the Competition Commission and the Monopolies and Mergers Commission that preceded it. As Wilks (1999: 1) confirms, '[t]he test which the Commission is asked to apply is that of 'the public interest''.

[63] For example, both the Electricity Act 1989 and the Telecommunications Act 1984 made various explicit references to the public interest as standard to decide certain matters.

Public interest, therefore, was much more successfully embedded in the system as a political discourse rather than a legal (or economic) concept. Moreover, it may seem that regulation obviated the need for direct judicial control. The problem was that the content of the public interest idea in the UK was never fully clarified.[64] Indeed, this shortcoming might have been overcome if the judiciary would have decided, ultimately, how the public interest should be best served in each dispute. However, for the reasons further explained in Chapter V, at least the English judiciary has traditionally restrained itself from intervening in regulatory affairs unless the regulator has exceeded its mandate (*ultra vires*), so the concept of public interest remained largely confined to regulatory interpretation.

In the US, the public interest idea also permeated early judicial approaches. Developments started with two well-known cases decided by the Supreme Court during the nineteenth and early twentieth century – *Munn v. Illinois* (1876) and *Nebbia v. New York* (1934).[65] In *Munn* – the seminal case in this area – the Supreme Court decided about the allegedly unlawful operation of a grain elevator and warehouse, due to the lack of licensing to operate and the higher rates it charged (above the state regulations). The reasoning of the majority (delivered by Waite C.J.) draws directly on the jurisprudence on English courts. The Supreme Court accepted regulation on the grounds that property rights were not absolute when the effect of the enjoyment of a service had public consequences. The interest of the community at large prevails over the individual and hence the Government, derived from its sovereignty, could use its police power to regulate private property – even establishing maximum charges – when the good of the public deserved it.

The effect of *Munn* was to effectively segregate public utilities from general corporations. After the case, the analysis of the state power to regulate rates needs to make the distinction between firms affected with a public interest and those that were not.[66] This meant a logical move from old charter-based to statute-based rate regulation of railroads and some other utilities: rates were not considered to be the product of an agreement anymore, but they were unilaterally imposed by the legislature.[67] Certainly, the legislature could impose

[64] *See* section II.A.2 further below.
[65] *Munn v. Illinois*, 94 U.S. 113 (1876); *Nebbia v. New York*, 291 U.S. 502 (1934).
[66] Hovenkamp (1991: 129).
[67] It took only two decades to make the transformation. For a history, see Hovenkamp (1991).

'confiscatory' rates – i.e. rates that do not give the firms a fair rate of return. However, as the Court indicated, 'We know that [rate regulation] is a power which may be abused; but that is not argument against its existence'.[68] As corollary, the Court affirmed the view that the extent of government powers in this area is not controlled by the judiciary, but by citizens through elections.[69]

Only decades later a new case, *Nebbia*, would give a definitive precedent regarding the content of the phrase 'affected with public interest'. In this decision, the Supreme Court understood the concept of public interest as nothing more than being 'subject to the exercise of police power', recognising that government inherently has an unquestionable power to promote general welfare. Therefore, the Fourteenth Amendment of the US Constitution allows the government to introduce regulation for the public welfare, as long as it accomplishes such end by means consistent with due process. As Justice Roberts indicated,

> [T]he guarantee of due process, as has often been held, demands only that the law shall not be unreasonable, arbitrary, or capricious, and that the means selected shall have a real and substantial relation to the object sought to be attained. It results that a regulation valid for one sort of business, or in given circumstances, may be invalid for another sort, or for the same business under other circumstances, because the reasonableness of each regulation depends upon the relevant facts.[70]

The Court gave the government an ample sphere of attribution in economic policy and a wide margin of appreciation of the circumstances. As long as due process is respected (i.e., as long as the there is a proper legislative purpose and the means to achieve it are not arbitrary, capricious or unreasonable), the government was free to adopt its own policies to promote public welfare.[71] As KAHN has stated, '[a]s far as the United States Constitution is concerned,

---

[68] *Munn*, 94 U.S. 113, 134.

[69] 'For protection against abuses by legislatures the people must resort to the polls, not to the courts' (*Ibid.*).

[70] *Nebbia v. New York*, 291 U.S. 502, 525. By the same token, the Court held that 'It is clear that there is no closed class or category of businesses affected with a public interest, and the function of courts in the application of the Fifth and Fourteenth Amendments is to determine in each case whether the circumstances vindicate the challenged regulation as a reasonable exertion of governmental authority or condemn it as arbitrary or discriminatory...' (291 U.S. 502, 536).

[71] 'The courts are without authority either to declare such [economic] policy, or, when it is declared by the legislature, to override it' (291 U.S. 502, 538).

there is no longer any distinction between public utilities and other industries'.[72] Fundamentally, however, *Nebbia* supported a similar view that in *Munn* (which is, in turn, similar to the British principle of *ultra vires*) and it has ever since become one of the main foundations of the Supreme Court's approach to regulation: normally, the Supreme Court would uphold regulatory determinations.[73] As in English law, judiciary self-restraint to review the merits of a decision became the rule.[74]

But as in the UK, in US politics the public interest idea developed in a rather different direction. Whilst the idea helped the judiciary to solve controversies on regulation, in politics it was a goal to achieve by regulation. The public interest idea dominated the debate between the 1890s and the 1930s, having its heyday during the so-called 'Progressive era' until well after the 'New Deal'.[75] This period of time coincides with the dominance of railways as the main regulated industry. Therefore, the main target of the 'theory' was to introduce federal railroad regulatory policy, because the public had to be protected against powerful monopolies that were too large to be reached by state legislation.[76] However, the idea was soon applied more widely. Politically, public interest became part of an ideological posture about the government – its aims and means to achieve them.[77] From a regulatory (and antitrust) perspective, the essence of the public interest concept both under the Progressivism and the New Deal was straightforward: it meant the protection of medium and small businesses against the action of large trusts. As a consequence, strong business regulation was one of the main policies adopted at that time.

Political ideas permeated the academic field. The public interest idea was adopted and developed by authors such as Charles Francis Adam Jr., Louis Brandeis and James M. Landis.[78] Nonetheless, their work does not reflect any special theoretical construction of the

[72] Kahn (1988: 8/I).
[73] *See generally*, Kahn (1988: 3-11/I).
[74] ON English law, *see* Chapter V.
[75] Sunstein (1990: 19): '[T]he rise of modern social and economic regulation should be seen as part of the New Deal reformation of the American constitutional structure'.
[76] Hovenkamp (1991: 132).
[77] Undoubtedly, considered as 'discourse' in regulatory politics and political ideology, the public interest idea is certainly deeper that this simple description. It may even be regarded as the 'proper role' of the State. A broader description, however, exceeds the limits of this work.
[78] *See* McCraw (1984: 1-79); Landis (1938); and Cushman (1941).

role of public interest in regulation.[79] Rather, their work is a legal articulation of the political thinking prevailing at that time concerning the role of the State and the need to control governmental power. For these authors, the link between administrative regulation and public interest was straightforward. Since public interest was obtained by the action of administrative agencies, it was necessary to develop arguments about how the agencies should be organised and how they should accomplish their goals. However, there was no specific explanation (or forecasting) of what public interest is or how it should be applied.

As in English law, the public interest idea in the US was much more successful as a political discourse rather than a legal concept. The vagueness of the idea would become one of its main criticisms.

2. Criticisms

The main substantive criticisms to the public interest idea of regulation are centred on the concept of public interest. First, it has been argued that the concept is extremely vague.[80] Indeed, despite the common occurrence of the concept, it is highly difficult to find a definition thereof. When such definition is given, it is usually too ample to be used as workable hypothesis.[81] The same criticism applies to both the judicial and political approaches. Consider *Allnutt*, for example. The idea of public interest in the decision resembles the idea of just price, according to which society's stakes in its own benefit are balanced against 'excessive gains' by traders. Therefore, price regulation could have been equally applicable to similar situations by analogy, even if services were not provided on a

[79] *See* Mitnick (1980: 97): 'In older literature it is only rarely that something like an explicit "Public Interest Theory" of regulation is put forth. This is largely a construct imposed by later work'.

[80] *E.g.*, Levine & Forrence (1990).

[81] On English law, an example is *R v Bedfodshire* (24 L.J.Q.B. 84), where Campbell C.J. defined a matter of public interest in the following terms: *'[It] does not mean that which is interesting as gratifying curiosity or a love of information or amusement; but that in which a class of the community have a pecuniary interest, or some interest by which their legal rights or liabilities are affected'*. As seen, any economic interest, legal rights or liabilities that may be affected are a matter of public interest. Equally ample was the task of the former Monopolies and Mergers Commission: 'the Commission does not apply a legal test but makes a judgement about 'the public interest'. This judgement is made anew in each individual case. There is no formal guidance, no informal guidelines and the Commission is not bound by precedent' (Wilks, 1999: 17).

basis of 'reasonableness' and 'moderation', as the case required.[82] On the other side, politically the idea of public interest seems roughly analogous to 'public good', 'common good' or 'general welfare'. However, in this sense the notion seems to overlap with a number of other theories and thoughts, such as the theory of public good, pluralist thought, functionalism, the [old] conviction that the societal whole is greater than the sum of the parts, etc. All in all, in practice the public interest seems to be no more than '... an expression of a socially-based right of public access to scarce resource'.[83]

The second line of criticism is essentially a specification of the former. According to some authors, the public interest idea would have similarities with the welfare economics rationale for regulation. The link between the public interest idea and welfare economics would have been initiated by the work of STIGLER and POSNER, who 'attempt[ed] to create a benchmark for the newly created Chicago Theory or [sic] regulation'.[84] However, the link is at least weak and at most completely misguided. There are several reasons for this. First, a modern and coherent theory of natural monopoly had not yet emerged in the nineteenth century, when the idea that some services were 'affected with the public interest' took shape.[85] Conversely, the debate remained highly politicised precisely due to the lack of consensus on a dominant economic model. As mentioned, some authors (particularly political economists) were becoming aware that certain industries – especially railroads – were subject to 'market failures' and that, consequently, competition was not operating efficiently. But to link the ideas of welfare economics and public interest is going too far.

In addition, as HANTKE-DOMAS has convincingly demonstrated, even if there were similarities between both the public interest idea and welfare economics, that would not prove or deny any connection between them.[86] In fact, in economics the concept of public interest is absent from the concept of market failure. From a legal and political perspective, the idea of public interest does not seem to go beyond a political discourse.

---

[82] The decision does not shed light on the 'reasonableness' of the price, nor the 'reasonability' of the utility that both consumers and the company are supposedly entitled to earn.

[83] Prosser (1997: 24). He mentions that other cases follow similar reasoning – e.g. *Harris v. Packwood*, 3 taunt. 263 (1810) and *Minister of Justice for the Dominion of Canada v. City of Levis*, A.C. 505 (PC) (1919).

[84] Hantke-Domas (2003: 188). The link between public interest and welfare economics would have been first suggested by Posner (1974) and further developed by Joskow and Noll (1981) (*Ibid*).

[85] Hovenkamp (1991: 131).

[86] Hantke-Domas (2003).

The third strand of criticisms arises from the consideration of regulation as the (necessary) exercise of public power by the government to cure market failures.[87] Based on this premise, the public interest idea would be both a positive theory about what *motivates* regulation and a normative theory about what *should motivate* regulation – it uses normative analysis to generate a positive theory by saying that regulation is supplied in response to the public's demand for the correction of a market failure or for the correction of highly inequitable practices (e.g., price discriminations).[88] Thus, in reality the theory puts forth the hypothesis that regulation occurs when it should occur (!). In this sense, the theory is plainly incomplete. Lacking in the analysis is a description of the mechanism that allows the public to bring the result about (i.e., how the potential for a net social welfare gain generates public demand for regulation). The analysis is also silent or arbitrary on the motivations of policy-makers: do they pursue their own conception of public interest or that of their superiors or even constituents?

All of these criticisms make the public interest a rather weak 'theory' (if there is such theory at all) to be considered as basis for analysing economic regulation. As detailed in the next section, in time the vagueness of the concept would lead to the development of the capture hypothesis and other theoretical elaborations that can be placed under the common label of 'public choice analysis'.

B. Public choice analysis

Generally speaking, public choice analysis refers to the wide group of traditional theories that aim to explain regulation through the lenses of political science and neoclassical economics. A non-exhaustive description of this group includes accounts associated with the so-called

[87] As seen, some early judicial decisions, especially in the US, adopted such view explicitly.

[88] For this reason the theory has also been called 'Normative Analysis as a Positive Theory'. As Joskow and Noll (1981: 36) states: 'The essence of this normative analysis as a positive theory is that one begins an analysis of a regulatory process with the assumption that its purpose is to maximize some universal measure of economic welfare, such as consumer surplus or total surplus'. *See also* Viscusi et al. (2005).

'Virginia School'[89]; public interest explanations; private interest accounts (associated with the so-called 'Chicago School')[90]; and the 'capture' hypothesis[91]. These analytical approaches constituted the first formal accounts of regulation and have influenced regulatory thinking ever since the late 70's. Indeed, there are vast divergences between each of them. However, because these theories share many of the same fundamental assumptions about markets, regulation, and politics; and also because only their most salient features are of importance here, it is appropriate to consider them together.[92]

Public choice studies focus on incentives and consequences of regulation. The key common assumption is one of perverse government. It is argued that governments frequently produce policies that benefit select interest groups, but that are detrimental to overall public welfare. As STIGLER famously put it: 'the problem of regulation is the problem of discovering when and why an industry (or other group of like-minded people) is able to use the State for its purposes'.[93] The answer, as it was normally posited, was that the state provides a favourable treatment in exchange of valuable political goods – i.e. votes and money. As a result of this rent-seeking behaviour[94], the government acts as a perverse mechanism for transfers of wealth.

On this basis, a great part of the subsequent public choice literature focused on demonstrating how powerful interests groups might influence elected officials or bureaucrats. This is not a view strange to the UK.[95] In fact, somewhat surprisingly, there are some similarities between public choice accounts of interest groups exerting influence over bureaucrats, on the one hand, with the early views that Herbert MORRISON remarks about socialisation of public enterprises in the UK, on the other. In sharp contrast with public choice accounts, MORRISON advocated public firms pursuing the public interest (an abstract notion he never made clear). However, he stated similarly that the board of a public firm should not be a board of 'full-time technicians' because 'the Board meetings would tend to be meetings of departmental

[89] *E.g.*, Downs (1957); Buchanan & Tullock (1962).
[90] *E.g.*, Stigler (1971); Posner (1971); Posner (1974); Pelztman (1976); Becker (1983). For a compilation of Chicago School studies on regulation and other areas, see Stigler (1988).
[91] *E.g.* Bernstein (1955); Kolko (1965).
[92] In the same sense, Leight (2010).
[93] Stigler (1971: 4).
[94] Tullock (1967). The concept, however, was first coined by Krueger (1974).
[95] For an assessment of the influence of pressure groups in the UK, see Grant (2000).

officers, each concerned to argue the case of his own department and the others not desiring to appear critical of a colleague'.[96] Furthermore, the Morrisonean model, which was applied in the UK more or less invariably (but with formal attempts to modify it) almost until the beginning of privatisation, advocated a firm operating at arm's length of the government, but subject to the ministerial power to make key decisions on policy and economic control.[97] This structure resulted in ministers preferring to deal directly and secretly with the industries, arguably leaving large space for interest groups to exert pressures on government officials.

More recently, public choice accounts have been employed to explain the privatisation process carried out during the 1980s. As will be explained in more detail in chapter III, the encouragement of wider employee share ownership as a means to alter the balance of political power has been singled out as one of the main causes of the movement. Some authors argue that as a consequence of the number of trade-offs made and the need to 'buy' support from different interest groups, the government ended up adopting policies that would not have been in the interest of consumers.[98]

At the same time, a second group of studies began the analysis of the causes and consequences of the intervention of governments in different markets – usually arguing that intervention has a harmful effect on consumer welfare.[99] One important strand is the so-called capture hypothesis.[100] This hypothesis developed as a sort of synthesis of all perverse interventions in markets and why regulation, whatever its manifestation, produces flaws in the policymaking process. On the one hand, it became the main theoretical support for the conclusion that regulation tends to be excessive and the regulator is prone to passing regulations which benefit industry.[101] But this idea is weak. For example, it does not consider the wide variety of regulatory tools available for regulators. Capture is strongly related with discretionary modes of regulation (*e.g.*, command and control), which give regulators substantially more leeway to affect market outcomes. Conversely, more market-based

[96] See Morrison (1933: 179).
[97] Graham & Prosser (1991: 10).
[98] E.g. Veljanovski (2010: 90).
[99] For an early survey, see Joskow & Rose (1989).
[100] So pervasive has been the influence of the concept, that on many occasions public choice theories are called indistinctly 'capture theory'.
[101] For early works in utilities, see Gerwig (1962); Demsetz (1968); MacAvoy (1971).

instruments should reduce the scope for influences, since direct regulation 'retreats' and only provides incentives for behaviour.[102] A framework that explicitly takes into account the whole range of regulatory instruments can assume, more realistically, that capture is less important than many studies in the public choice tradition have claimed.

On the other hand, the capture hypothesis also has practical implications. For instance, it may have led policymakers to over-emphasise the independence of the regulator from the regulated firms – considering these as a powerful interest group that can exert a dominant negative influence on the regulator. Regardless of how sensible that policy is, the underlying assumption has never been clear. In fact, many of the authors considered adherents to the public choice theory expressly built their models assuming that different interest groups can exert opposite influence on government officials, with a vast numbers of outcomes either positive or negative.[103] Utilities regulation provides an example. Within the scenario of deregulated industries, capture loses much of its relevance. As more firms are active in the market, a single powerful firm is unlikely to dominate the regulatory process and is therefore unlikely to have strong influence on the outcome. Indeed, the firm may not comply passively with the regulatory rules and intend to modify the outcome. But to immediately infer capture from its behaviour is logically inconsistent.

In sum, the capture hypothesis is theoretically and practically feeble. It does not consider the multiplicity of regulatory tools and approaches available to regulators, nor does it take into account the array of actors involved in the regulatory process. Both aspects are crucial for a more comprehensive perspective. The concept of capture is imprecise, and only tells 'an extreme story – one that probably does not characterize most regulated industries'.[104]

More generally, public choice theory has also received many criticisms. First, many authors pointed out the empirical and theoretical weaknesses of the thesis according to which groups influence regulators for their own benefit.[105] Early works based this critique on deregulation. It was argued that public choice arguments were unable to explain deregulatory movements

[102] See Grabosky (1995a).
[103] Peltzman (1976); Becker (1983).
[104] Rossi (2005: 236).
[105] E.g., Quirk (1981).

that were opposed by large firms. For some, deregulation may well have been the result of the 'force of ideas'.[106] More recently, commentators have added to this claim indicating that there are many ambiguities in the public choice assessment and that the theory is impossible to falsify.[107] In addition to these arguments, others have singled out the defects in the theoretical foundations of public choice accounts: they would place inordinate attention on particular aspects of substantive regulation, rather than understanding the mechanisms of change underpinning a regulatory regime *before* deriving a theory of it.[108]

More important for this chapter is a second line of criticisms that can be classified under the loose heading of cognitive and motivational failures. On the one hand, principal-agent models (described in the next section) have shown that many regulations may have been the result of informational asymmetries on historical data or the state of technology, for example, rather than paying favours to regulated firms.[109] This suggests the idea of informational deficiencies in the cognitive process. Moreover, arguably one might apply the insights of behavioural economics and other disciplines and submit that some regulations have simply been the result of a different set of cognitive biases – namely, not only the inability to grasp information, but also more serious failures in the subjective assessment, with data being processed to heuristic rules.[110] On the other hand, in public choice models it is unclear what the motivational factors of the agents are. True: there is a clear claim that maximisation of utility is the impulse that underlies the behaviour of all actors (i.e., voters, governments and regulators). However, it is difficult to find the actual wealth function. In many cases motivation is constructed on *ad-hoc* basis. In addition, some have indicated that the concept of rationality underlying the models is completely ambiguous.[111] And, finally, others have pointed out that the motivation may be circular: 'In one sense, the model is self-evidently true: almost all behaviour serves personal

[106] E.g., Derthick & Quirk (1985). For a 'follow-up', see Landy *et al.* (2007). Contrast Peltzman (1989) (explaining deregulation from a public choice perspective).

[107] See Leight (2010: 237-8). See also Rossi (2005: 55-6) (stating that 'at best, empirical support for regulatory agency capture is mild').

[108] Priest (1992: 323) (criticising the traditional 'theories of regulation' because they largely deflect attention from the 'dynamic relationship' between the parties of the regulatory relation).

[109] Although many authors that follow the P-A approach supported the capture hypothesis. E.g. Laffont & Martimort (1999); Martimort (1999).

[110] Kahneman *et al.* (1982); Jolls *et al.* (1998); Sunstein (2000); Korobkin & Ulen (2000); Gilovich *et al.* (2002).

[111] Green & Shapiro (1994).

"interests," somehow defined, and thus is self-interested'.[112] The same claim underlies all these arguments: in public choice accounts of regulation, the central motivational factor of the agents remains unclear.

C. Principal-agent models

1. Conceptualisation

A third group of explanations for regulation uses the so-called 'principal-agent' (P-A) framework. P-A models describe the 'simple situation where a principal, or company, delegates a task to a single agent through a contract – the essence of management and contract theory'.[113] The principal contracts with the agent aiming to maximise his own profit, but facing some degree of informational disadvantage and being subject to the constraints imposed by the agent's own incentives. On the one hand, information asymmetries are at the heart of the P-A relation. Typically, the agent is considered to have some private information that the principal cannot verify. That information may be about the agent's own skills or preferences, or of another type. Most commonly, private information is said to take the form of 'moral hazard' (hidden action) or 'adverse selection' (hidden information). Whatever the private information is, it allows the agent to know (presently or in the future) something the principal does not, or take secret actions. This knowledge provides the agent with a crucial advantage: s/he can perform certain tasks too costly or complicated to be completed by the own principal or which may be unverifiable by the principal, or whose outcome may not be verified. On the other hand, frictions may arise if the agent's incentives and goals are not perfectly aligned with those of the principal. The reasons for the misalignment may vary. A central one (relevant to the subsequent analysis of economic regulation) is the limited ability of the agent to commit, but there are also many others.[114] All in all, informational

[112] Wilson (1980: 361).
[113] Laffont & Martimort (2002). See also Laffont (2003) (collecting essays on the area).
[114] For other reasons, see Sappington (1991).

asymmetries and conflicting incentives may result in inefficiencies, which are likely to be harmful for the principal.[115]

The principal is aware of the agent's informational advantage and his own potential handicap. For that reason, the central aim of P-A models is motivational: how to develop the right incentives to control the agent and induce him to perform exactly as the principal wants.[116] The principal intends to devise a method to overcome its own unfavourable position and incorporate protections within the contract.[117] The approach of the theory is 'vertical': the principal is allowed, and has the contractual ascendance, to impose a mechanism of incentives-alignment over the agent in a top-down process.

To a great extent, this approach is akin to the delegation problem much studied by political science literature – i.e. the problem of overseeing and controlling power delegated to independent regulatory agencies. In this context, there are two different, but interconnected P-A relations in regulation. On the one hand, 'principal are the elected officials who pass legislation and delegate their authority to bureaucratic agents who must develop an appropriate pattern of implementation actions'.[118] On the other hand, each firm is assumed to enter into a relation as agent of the regulator (which represents consumers). The regulator needs to find the better mechanism to overcome the strategic informational advantage of the firm and create the right incentives for it to outperform the regulatory targets. Whilst a great part of early political science literature focused on the first P-A relation and their possible solutions[119], current regulatory economics literature tends to emphasise more on the second one – which is also the focus of the next part.

[115] In utilities regulation and some other areas (like corporate governance), for instance, expropriation is a highly probable outcome of the agency problem. See Schleifer & Vishny (1997) (analysing the P-A model in the context of corporate governance).

[116] Sappington (1991: 45): 'The central concern is how the principal can best motivate the agent to perform as the principal would prefer, taking into account the difficulties of monitoring the agent's activities'.

[117] See Ross (1973) for a technical analysis.

[118] Eisner *et al.* (2006: 25).

[119] E.g. McCubbins & Schwartz (1984); McNollGast (1987) and (1989).

2. P-A models in economic regulation

Because of the presence of asymmetries of information between regulators and firms, the P-A model was also applied to economic regulation.[120] With the purpose of finding the most efficient mechanism of incentive-alignment between firms and regulators, P-A models started to rely on mechanism-design theory.[121] This theory aims to devise the best way to communicate both principal and agents using overt actions or encoded messages under conditions of 'privacy-preserving'. Social interactions between the principal and the agent focus on the design of incentive compatibility constraints for agents and mechanisms of preferences/information revelation. The designer acknowledges that agents adopt their decisions in a self-regarding manner: agents will reveal their true preferences and positions only to the extent compatible with their hidden motives. That is, the approach is based on a confrontational premise. Also, given the informational asymmetries, the outcome of the models is necessarily a second-best contractual result. In the end, in mechanism design models all the actions aim to attain the individual's own goals. If there is an advancement of the goals of another individual, this is only a by-product.

Mechanism-design theory provided the theory of regulation with powerful mathematical and economic tools to overcome agency concerns – indeed, tools far more sophisticated than those applied in previous models. Its approach relies heavily on game theory. Thus, the regulatory process started to be modelled as a non-cooperative game of incomplete information, in order to compare the different set of equilibrium outcomes of the game. The performance of different results can be analysed relative to some theoretical optimum. Making use of the so-called 'revelation principle', new regulatory models were able to derive optimal schemes without resorting to *ad-hoc* assumptions (commonly employed in previous models).[122] Thus, where traditional approaches to regulation had naïvely assumed complete

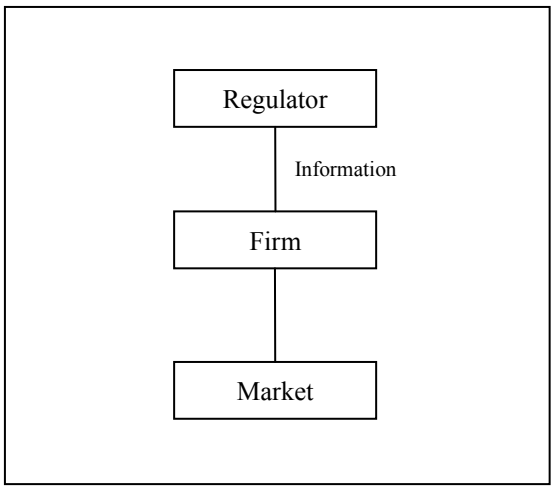
[120] The first formulation of regulation as an agency problem is due to Loeb & Magat (1979).

[121] See Hurwicz (1973); Hurwicz & Reiter (2006). For a summary of the mechanism-design approach in different disciplines, see Nobel Prize (2007).

[122] According to the 'revelation principle', any equilibrium outcome of an arbitrary mechanism can be replicated by an incentive-compatible direct mechanism—i.e., a mechanism where the dominant strategy of a player is to report his/her private information truthfully. Therefore, the principle implies that an optimal outcome can always be found within a sub-class of direct mechanisms, despite the fact that the set of all possible

information between the parties, new models of optimal regulation focused explicitly upon cases when the full-information outcome is not possible.[123] This way of analysing the regulatory relation came to be known as the 'new economics of regulation', which, as any other P-A approach, considered the relation from a vertical perspective (Figure 2).

Figure 2: *P-A simple approach to regulation*



The new economics of regulation was a highly successful movement.[124] Its theoretical solutions were extended in many directions, and studies tested empirically the effect of regulation on firm's behaviour in the observed practice – generally with satisfactory results. The analysis was applied to both State-owned companies and private firms.[125] In both cases

mechanisms is vast. The task is then to find the direct mechanism that maximises a given goal function (e.g., social welfare), subject to an incentive-compatibility constrain. For the most general version of the revelation principle, see Myerson (1979) and (1982). Amongst others, well-known economic models of regulation in this line are Baron & Myerson (1982) (where the level of costs is given to the firm, but the marginal cost is exogenous and unobserved by the regulator); Baron & Besanko (1986) (extending the analysis of Baron and Myerson to dynamic models with commitments); and Laffont & Tirole (1986) (where the marginal cost is endogenous and can be observed by the regulator, but it is influenced by two unobserved factors: firm's cost-reducing effort and the state of nature).

[123] See Laffont & Tirole (1993: 1) (defining moral hazard as 'endogenous variables that are not observed by the regulator', and adverse selection as the problem arising 'when the firm has more information than the regulator about some exogenous variables') (emphasis omitted).

[124] For a survey of the main results of the new economics of regulation, see Laffont (1994).

[125] *See*, e.g., Vickers & Yarrow 1988: 92 *et seq.*) (applying the P-A model to the relation between public or private owners and managers, and also between firms and the regulatory authority); Martimort (2006) (proposing a unified theoretical framework based on incentive regulation to discuss the costs and benefits of

principals and firms may be assumed to be engaged in a relation whereby the principal does not know how the agent is acting and must devise a mechanism to align their objectives. This became apparent after the widespread wave of privatisation in utilities sectors, when the change in ownership structure altered the nature of the incentives faced by firm.[126] Generally speaking, the effects on wealth arising from regulatory decisions are fundamentally less relevant for State-owned companies than for private firms.

In public-owned firms, the principal is endogenous to the firm. Therefore, concerns are primarily focused on how the government can discover what its managers are doing with the company – that is, managerial concerns. Unlike private firms, the variety of actors who directly or indirectly control public firms (i.e., government officials, managers, unions, political parties, etc.) do not directly benefit from profits in the form of dividends. On the contrary, on many occasions benefits may arise from inefficiencies that increase costs, such as raises in salary, employment of more work-force than necessary, the adoption of lax working conditions, etc. In addition, recognised political forces usually present in utilities regulation give rise to demands that, under certain conditions, may be responded to with populist measures that produce political dividends – e.g., setting prices below cost or subsidising the public firm.[127] The effects on wealth are secondary. This incentive structure produces three main effects. First, it leads public-owned companies to *work* with higher costs than efficient ones. At the same time, incentives may prevent the introduction and use of mechanisms of comparison with some standard of efficiency.[128] Finally, there is also a 'ratchet effect': if the company's past-performance is used by the government to set future goals, agents will have limited incentives to improve performance, because every effort will 'ratchet up' future objectives.[129]

---

privatization); Laffont & Tirole (1993: 34) (arguing that 'a more rigorous and realistic approach must adhere to the discipline of the broader principal-agent theory').

[126] The analysis adopts a regulatory perspective. This does not mean that from other perspectives, e.g., managerial, agency problems might be similar in both privately- and publicly-owned companies.

[127] The influence of political forces in the public utilities regulation is widely recognised. See, among others, Derthick & Quirk (1985) and Landy *et al.* (2007: 9) ('old-fashioned group-based politics never died away').

[128] *See* Leibenstein (1966).

[129] A technical analysis is presented by Laffont & Tirole (1993: 375 *et seq.*).

By sharp contrast, under private ownership the principal in the regulatory relation (i.e., the regulator) is exogenous to the firm. However, the effects on wealth arising from regulatory decisions of the former are crucial for the latter – at least if profit-maximisation is assumed. Take the extreme case of monopolies: every increase in the tariff will be reflected in more profits. This creates an incentive structure that leads private firms to *declare* costs higher than the real ones in order to increase the regulated tariff – even if the firm works with costs close to the optimum.[130] The problem for the regulator is that the firm is placed in an advantaged position given its deeper knowledge of its own costs and demand functions. The firm may manipulate information to achieve prices closer to monopolistic ones. The key regulatory task, then, becomes how to incentivise the firm to reveal the true information about its costs.

One of the main contributions of the new economics of regulation was to place informational concerns at the centre. This approach emphasised the use of incentive-based or performance-based regulation to overcome those concerns. Incentive-based regulation included a wide variety of price mechanisms – not only the widely popular price-caps, but also profit sharing and revenue sharing mechanisms and various forms of yardstick competition.[131] Many studies showed that, overall, these mechanisms were more apt to deal with informational problems than 'old' ratemaking schemes such as cost of service (CoS) regulation.[132]As a result, incentive regulation provided positive incentives for cost-reducing investments.[133]

However, two caveats are worth considering. The first is that the focus on informational constraints was not a by-product of incentive-based regulation. In fact, many concerns regarding CoS schemes were already dominated by informational aspects. Customers' ability to perceive and control tariffs greatly diminished under these ratemaking schemes. In addition, the regulator must usually rely upon information given by the firm, limiting the control that it can exert over the regulated firm. The focus on information has, therefore, historical roots. What incentive regulation did –with the input of P-A models– was to openly

[130] Note the clear difference. The public company has an incentive to work with costs higher above the efficiency line. The private firm, even if working with the same level of (inefficient) costs, will tend to declare costs even higher.

[131] See e.g. Baldwin & Cave (1999: esp. ch. 17).

[132] For traditional and detailed accounts of CoS regulation, see Troxel (1947), Bonbright *et al.* (1988) and Kahn (1988).

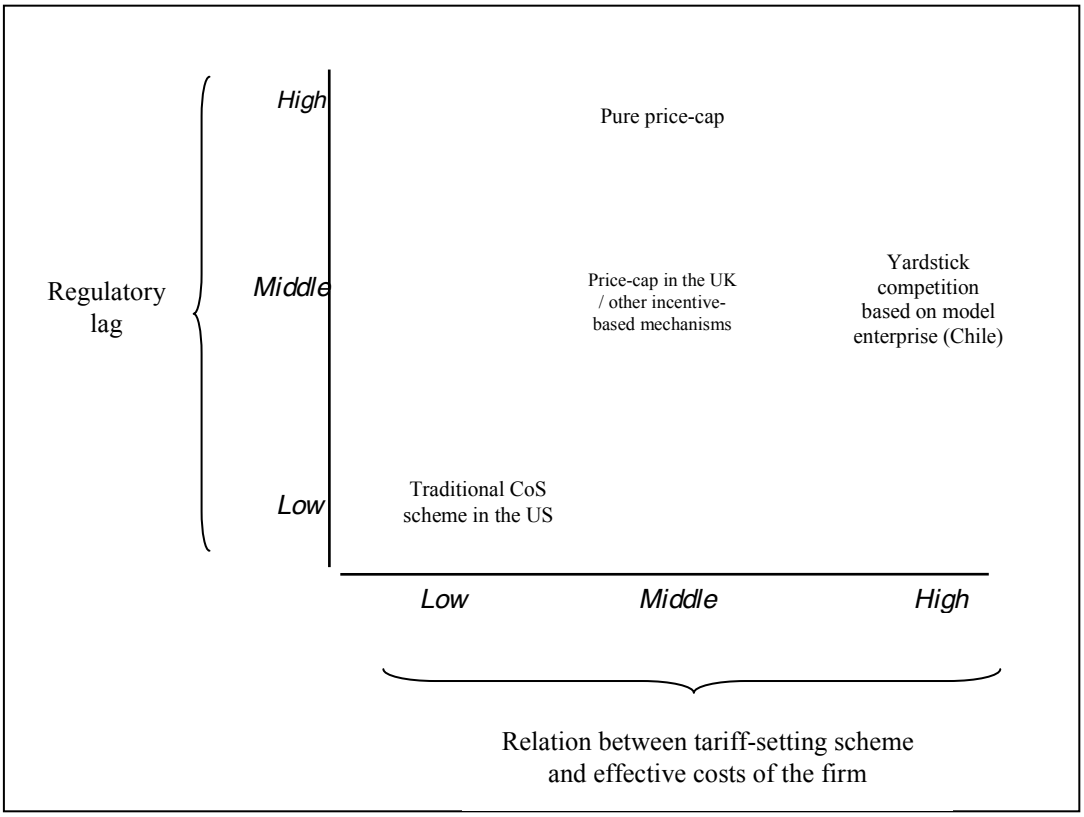
[133] Among the main economic studies are: Cabral & Riordan (1989); Biglaiser & Riordan (2000); and Roques & Savva (2009).

(and technically) describe the distortions caused by the lack of information and the need to limit firms' excessive rent. The second caveat is that incentives seem to work differently for infrastructure investment.[134] Whilst CoS regulation may provide strong incentives for developing new infrastructure due to the reduced risk, price cap regulation and other mechanisms may weaken the incentives to invest, because the regulatory lags are shorter than the life of the assets.[135]

The analysis showed that despite the fact that all schemes are imperfect information-revealing mechanisms, some of them are more effective at surmounting the informational constraints that the regulator faces. This means that different regulatory contracts have different 'powers' – i.e. 'the link between the firm's…prices and its costs or profit performance'.[136] As detailed in Figure 3, this link depends on the lag between price controls. On these grounds, P-A analysis confirmed the advantages of, for example, price-caps *vis-à-vis* regulation based on CoS schemes.

[134] See generally Armstrong & Sappington (2006).
[135] See *infra* Chapter IV, note 419 and accompanying text.
[136] Laffont & Tirole (1993: 11).

Figure 3: Power of different regulatory ratemaking schemes



The P-A framework of analysis fostered new important developments and influenced the design of post-privatisation regulatory frameworks. Many – if not most – of those frameworks were introduced or later reformed with the P-A problem and informational concerns in mind (including those for UK utilities). The vast majority of tools introduced in the frameworks were also based upon outcomes extracted from that approach in response to the demand for information gathering.

3. Criticisms

In the context of regulation, P-A models and mechanism design theory can be criticised on both theoretical and practical grounds.

On the theoretical side, it can be argued that the regulatory relation is much more complex and diverse than the rather simplistic portraits shown by P-A models and mechanism design theory. To a large extent, P-A models are built on the basis of the *homo economicus* that constituted the basis of mainstream, neoclassical economics.[137] Models usually assume agents with well-ordered and transitive preferences. Moreover, their beliefs, stimulus, and 'tastes' are all pre-fixed prior to social interaction. These individualist, solipsistic agents are solely concerned with their own actions and driven and conditioned by their own individual motivational factors. Single-mindedly, they pursue the maximisation of their own welfare, leaving no space for other non-private motivations different from mere self-interest. This constraint leads the models to envisage no means of detecting misrepresentations that conduce to unintended effects. As such, the models were open to a wide range of criticisms arising against the thin view of rational actors.

The criticisms, ubiquitous across different disciplines, are well-known and therefore it is not necessary to repeat them here in great detail. Generally they fall under one of two groups. On the one hand, works on different disciplines have shown that the predictions made under very simplistic assumptions did not match observable behaviour. The most recognized body of literature on this area is behavioural economics.[138] Beyond the insights of KAHNEMAN and TVERSKY and their followers, however, there are other strands of literature that emphasise cognition and apply it to institutional design without using the behavioural economics framework. Among them, psychological works on cognitive coherent-based reasoning[139], legal literature on social norms and trust[140], political science work on common property regimes[141], and jurisprudence[142] and sociological[143] literature, have all shown the importance of cognitive failures of rationality and the predictive strength of traditional models.

[137] In spite of its limited assumptions, it is commonly argued that models are justified by the quality of their predictions, not the plausibility of their assumptions. See Friedman (1953) (defending mainstream economics generally).

[138] See the bibliography cited *supra* in note 110.

[139] Simon (2004: 517) ('The central fining of coherence-based reasoning research is that the cognitive system *imposes* coherence on complex decision tasks' [emphasis in the original]).

[140] Ellickson (1991).

[141] Ostrom (1991).

[142] Sen (1977).

[143] Simon (1957).

On the other hand, there is an increasing body of literature, in different disciplines, showing that motivation may not be entirely self-serving. Individuals are not similarly motivated and selfish. On the contrary, their behaviour may reflect a wide range of attitudinal or intentional states. They often act cooperatively and respond kindly to others – either unconditionally (because of altruism, solidarity or lack of awareness of the consequences) or conditionally, punishing those who fail to cooperate (even at a cost to themselves!). As BENKLER summarises, the possibilities include altruism[144], committed mutualism[145], collective efficacy[146], heuristic reciprocity[147] and strategic mutualism[148] – only the latter being the kind of behaviour that game theoretical mechanism design seeks to elicit. The upshot is that people are diversely motivated, and on many occasions inclined to cooperate unselfishly.[149] In turn, this implies ways of exchanging communication that are much richer than the simple 'formal language' employed by P-A models.[150]

Indeed, it might be argued that predictions are more accurate and/or motivational factors are not so important in the context of utilities regulation. The firm is traditionally deemed to be an archetypical rational actor, pursuing the maximisation of its profit in a highly calculated manner. Informal ways of communication have, at most, weak importance. This assumption has been largely discredited.[151] Even if this were the case, however, P-A models could be criticised because of their narrow vision of the regulatory relation. Most models rely on a two-dimensional view of the players, meaning that interactions occur mainly (if not only) between regulators and firms. Literature has shown the weakness of this assumption. Other interest groups affect interactions and in most occasions participate in them. These groups may include not only the government as a separate entity from the regulator, but also non-governmental organisations, non-business corporations, consumer associations, and so forth.

[144] Benkler (2010: 311) (The 'action aimed to contribute to success of another, irrespective of success of self').

[145] The 'action aimed to contribute to success of another, consistent with success of self' (*Ibid.*).

[146] The 'action aimed toward the success of a common goal that transcends the agent's specifiable individual success' (*Ibid.*).

[147] The 'action aimed at benefiting one's own goal pursuit, guided by an implicit sense that "what goes around comes around"' (*Ibid.*).

[148] The 'action aimed at attaining one's own goal, which advances the goal of another as a by-product' (*Ibid.*).

[149] Benkler (2010: 302). More on cooperation in chapter II.

[150] More on this in chapter II.

[151] For example, Sperber & Wilson (1995), with their 'communicative principle of relevance', make a compelling argument that individuals are able to infer intentions and thinking even from the context.

It also, and importantly, includes courts. Their goals range in a wide spectrum, including economic, political and social objectives. But the key point is that whilst pursuing their own agendas and objectives (even though their motivations may not be entirely self-interested), each of these agents participates in, influences and moulds the regulatory relation.

Finally, P-A models normally assume a top-down, vertical process of regulation whereby regulatory rules are imposed exogenously on firms. In turn, rules are mostly designed with a single objective in mind: the achievement of competition through the work of markets is extolled as the most suitable way to achieve efficiency. All of these assumptions are nonetheless controversial. The main downside of the P-A literature is that institutional issues generally are not of interest, because institutions are taken as given. If they were considered, it would be recognised that it is at least debatable whether markets are the most suitable institution to be chosen – they are one among a number of institutional architectures capable of being applied.[152] However, even if institutions are taken as given, this literature uses normative considerations as a theoretical benchmark and not as a practical objective. In reality, regulation is a second best solution and therefore the objective of efficiency can be scrapped.

On the practical side, the P-A framework of analysis also produced some negative consequences. The first and foremost is that remedies based upon the insights of P-A models drastically increased the technical content and complexity of economic regulation. In some regimes, excessively detailed specifications were embedded in regulatory frameworks at the risk of producing over-inclusiveness. The Chilean Electricity Act, one of the first post-privatisation statutes, provides a particularly extreme example.[153] It is a very comprehensive, highly detailed and complex piece of utilities legislation. Different issues such as tariff-setting process, the awarding of concessions, security and safety issues, and quality standards

[152] See below, section II.C.4 of this chapter.

[153] The Chilean Electricity Act is contained in Ministry of Mining (1982): DFL No1, General Law for Electric Services referring to Electric Energy. Having been developed in late 70's, it was definitively introduced in 1982. Although it has been the subject of substantial amendments, it still remains the most important legislation governing the sector, alongside the Decree 327/98. Pollitt (2004) provides a detailed analysis.

are all embedded in the Act itself.[154] For historical and ideological reasons, it incorporates fine points that are usually and reasonably left to administrative discretion.[155] Also, it relies upon engineering principles more than economic ones – although it incorporates some blurred principles of incentive regulation. Conversely, for institutional reasons in the UK context complexity was reflected in the regulatory practice.[156] Whilst focusing on informational constraints, regulators acknowledged not a single, but many types of information which they need to deal with. This turned price controls into highly complex processes, with detailed information gathered, numerous consultation documents drawn up, various bilateral meetings and working groups convened, and so on.

Increased complexity may also impact negatively on the accountability (particularly judicial accountability) of regulators. Arguably, in the UK context not only did the personalised style of regulation increase regulatory discretion after privatisation, but the complexity of the framework also played an important part in this process.

D. New institutional economics

1. Conceptualisation

The last common framework of analysis of the regulatory relation develops under the umbrella of New Institutional Economics (NIE). The most salient features of NIE are now widely well-known within the field of law and economics, therefore, only its main aspects are repeated here.[157] NIE builds on, modifies, and extends neoclassical theory by incorporating a

[154] Note that in other presidential regimes the pattern is normally the opposite. Legislation is 'regulated' by secondary rules inserted by administrative decrees. Consequently, utilities regulation laws tend to be generic, leaving the details entrusted to the executive power.

[155] The act was dictated in the context of ideological change whose results were great distrust for administrative discretion and major institutional changes. As a result, there was a general reduction of the role of the administration in vast sectors of the economy. Institutional changes included a new Constitution with strong incentives for private initiative and heavy protection of property rights. A summary of the Chilean's privatisation process is presented in Fisher *et al.* (2003).

[156] See e.g. the adoption of 'menu' type of regulation (Laffont & Tirole, 1993) by Ofgem, the energy regulator, as explained by one official: Crouch (2006).

[157] For good overview and assessments, see Furubotn & Richter (1997); Williamson (2000: esp. 598-600) and (1998).

theory of institutions into economics, stressing the importance of institutions for social analysis. That is, it tries to explain what institutions are, how they arise, what purposes they serve, how they change and how they may be reformed.[158] Following NORTH, institutions are commonly defined as the humanly devised (formal and informal) constraints that structure political, economic and social interaction.[159] Nonetheless, the concept of institutions has many alternative and somewhat imprecise meanings, understood differently by different authors.[160] NIE is a broad field that includes works in transaction costs economics (TCE)[161], political economy[162], (positive) political theory[163], property rights[164], hierarchy and organisation[165], sociology and organisations[166], theory of the firm[167], and others.

The NIE framework is developed over two core pillars. On the one hand, the transaction costs economics perspective frames the discussion of both organizational form and legal rules, stressing the importance of the 'costs of running the economic system'.[168] On the other hand, from a decision theory viewpoint, NIE expressly takes into account individual cognitive limitations by assuming bounded rationality. Under bounded rationality, actors are '*intendedly* rational, but only *boundedly* so'.[169] This assumption tempered the traditional economic assumptions of self-interest and maximisation. Over these foundations, NIE analysis has been applied at four different levels – each level presenting a general description of the main institutions, their frequency of changes and the purposes of those changes.[170] From top to bottom, each level imposes constraints over the lower one.

---

[158] In this sense, it expressly built and expands on 'classic' institutional theory: see, e.g., Commons (1931).

[159] North (1991); see also North (1990a).

[160] *E.g.*, for a game theoretical understanding, see Schotter (1981).

[161] The TCE approach was initiated fundamentally after the work of Coase (1937) (stating that the main reason the firm emerges is because there is a cost of external exchange using the price mechanism of the market) and (1960) (stating that if property rights are well-assigned and transaction costs are zero, market transactions among the parties will lead to the most efficient outcomes). Both articles are reprinted in Coase (1988). For an early survey on TCE, see Williamson (1989: 135).

[162] See e.g., North (1990b); Weingast (1993).

[163] Spiller & Tommasi (2003) (developing a transaction theory to explain how political institutions affect transactions undertaken by political actors and the quality of policies that emerge).

[164] Barzel (1989).

[165] See e.g., Williamson (1985) and (1996).

[166] Powell & DiMaggio (1991).

[167] Klein *et al.* (1978).

[168] This is the straightforward definition of Arrow (1985).

[169] The famous quote belongs to Simon (1997: 88) (emphasis in the original). Simon claims that satisfaction rather than maximisation should be the goal sought in economics (Simon, 1959).

[170] This paragraph follows Williamson (2000).

The top level of analysis is *social embeddedness*[171], where the analysis is focused on informal institutions such as customs, traditions, norms, religion and other informal institutions that control and limit the next level. Most institutional literature takes the level of embeddedness as given, as does this work. The next two levels are the 'institutional environment' and the 'contractual governance institutions'. The institutional environment comprises political institutions and their implications for economic performance. Preferences at this second level imply 'first order economizing' choices – i.e., getting the formal rules of the game right.[172] The institutional environment constrains the governance institutions that form the third level of analysis. This level includes different contractual governance structures (i.e., markets, hierarchies and hybrids) that allow savings on transaction costs.[173] Commentators colloquially refer to this level as the 'play of the game'. The final level is where 'regulatory incentives' are displayed and affect firms acting in the marketplace – i.e., the level 'at which the neoclassical analysis works'.[174] This includes routine economic activities such as production and employment, financial aspects such as the ability to raise capital to implement large-scale investment, issues of risk management, pricing policies, government subsidies, competition and entry, etc.

The NIE framework has influenced regulatory literature at the second and third level: the institutional environment and institutional arrangements, in the manner described in the following two sections.

[171] The notion of embeddedness is commonly used in institutional analysis. The notion comes from institutional economic sociology, and was first introduced in the field by Granovetter (1985) (coining the term). For a critique, see Krippner (2001).

[172] This refers fundamentally to constitution, laws and the structure of property rights. The word 'laws' is taken here in a broader sense, including legal context. Indeed, rules vary not only according to their characteristics—i.e., extent, specificity, precision, intelligibility, legal status, etc.—but also according to their legal context. See Baldwin (1995).

[173] See Williamson (1991); Ménard (2004); Thorelli (1986). The assignment of least-cost governance structures to manage transactions is the so-called 'discriminating alignment hypothesis' (Williamson, 1985).

[174] Williamson (2000: 600).

2. Institutional environment and regulation

The application of the NIE framework to utilities comes largely from the oft-cited seminal work of LEVY and SPILLER.[175] Their analysis is situated at the institutional environment level, and departs from the *ex post* risk of expropriation of the firms' assets as a consequence of opportunistic behaviour on the part of the government. Because of the characteristics of utilities – i.e., large amounts of specific and non-redeployable assets that mostly constitute sunk costs, economies of scale and scope, and its broad range of domestic users – governments can make use of their administrative capabilities to expropriate the quasi-rents of the utilities firms administratively, either directly[176] or (more frequently) indirectly (i.e., by subtle means). An example of indirect expropriation is the adoption of changes in the regulatory framework financially inconvenient for investors or regulatory measures, such as setting prices under long-run average costs.[177] In this scenario, private investors would not be able to withdraw given the nature of their investment.[178] Hence from the outset they may not be willing to sink their money into the country. To be specific, opportunism reduces the amount of resources that private investors are willing to invest because the hazards are great.

Note that the source of concern changed with respect to P-A models. Underlying P-A models is the peril of firms exercising market power, whereas the central concern for NIE arises from governments exercising their own power of coercion. So to speak, 'public' expropriation instead of 'private' expropriation is the main concern.

---

[175] Levy & Spiller (1994). *See also* Levy & Spiller (1996).

[176] This usually takes the form of 'nationalisations', a method that does not require compensation in every case.

[177] If prices are set under LRAC, the monopoly is unable to survive in the market, since total revenue would be less than total cost. See, e.g., Posner (2007: 369).

[178] This is a manifestation of the so-called 'hold-up' problem—that is, the situation 'in which a party to a new or existing contract accedes to a very disadvantageous demand, owing to the party's being in a circumstance of substantial need'. See Shavell (2007: 325-6). Also, this is a topic related with policy risk (see *infra*, note 551 and accompanying text).

To overcome opportunism and attract long-term private investment, governments must refrain themselves from using their own capabilities against firms and assure investors that their assets will not be expropriated. In other words, they must show 'credible commitment'. The central question for governments (in a context where they are not willing to provide the required capital anymore) becomes how to make their commitments credible.[179] Utilities regulation becomes fundamentally a problem of design – i.e., how to devise *ex ante* safeguards against public expropriation and embed them within the regulatory framework.[180] The analysis moves away from the focus on incentives. Whilst these remain important in influencing performance, their impact crucially depends on the creation of a regulatory governance structure that provides such protections. The design of credible regulatory institutions is the key to successful regulatory outcomes.[181]

The design of the governance structure depends on the institutional 'endowment' of the country, which consists mainly of politics (including executive and legislative functions), judiciary (including the creation of mechanisms to resolve conflicts), administrative or bureaucratic capabilities, customs and other informal norms, and the country's social interests (including ideology). Depending on the country's endowment, the design of the regulatory framework needs to consider a number of checks-and-balances. However, LEVY and SPILLER argue that utility regulation is likely to be more credible if both executive and legislative discretion is curbed.[182] Since expropriation is more likely to occur if the government faces strong political pressures, the objective is to isolate the regulator from discretion. Therefore, the need to constrain unfettered discretion was seen as the fundamental problem for the design of regulatory governance structures.[183]

Later works stressed that both constraints of unfettered regulatory discretion and flexibility must be balanced in order to foster private investment and efficient operation of utilities. To some extent the emphasis was shifted away from discretion. One strand of the literature has

[179] E.g. Rose-Ackerman & Rossi (2000: 1436): 'How does a state attract foreign investment where there is some possibility that the commitments behind its current regulatory regime may change?'.

[180] On the concept of embeddedness, see *supra* note 171.

[181] In the categorical words of Holburn & Spiller (2002: 464): '…"having the institutions right" is more important than "having the structure right"' (double quotation marks in the original).

[182] Levy & Spiller (1996: 5).

[183] In fact, Levy & Spiller defined the governance structure as 'the mechanisms that societies use to constrain regulatory discretion and to resolve conflicts that arise in relation with these constraints' (1996: 4).

given independence central importance, considering the need to isolate the regulator from political intervention as much as possible. Most authors have proposed an arm's-length relationship with political authorities.[184] On the other hand, it has been argued that the most important role of the governance structure is not to constrain discretion, but to make regulators accountable for their decisions. In fulfilling that role, governance *procedures* must complement the regulatory institutional arrangements. Setting 'regulatory criteria for governance' is the primary task of regulatory design.[185] The criteria vary. Nonetheless, they usually comprise clarity of objectives and roles, coherence, independence or autonomy, accountability, transparency, predictability, and capacity.

3. Institutional arrangements and regulation, and the 'regulatory contract' approach

A second strand of NIE has matched other works on long-term contracting and contract incompleteness with the task of detailing specific contractual arrangements. From this perspective, regulation is seen as a contracting problem aimed at protecting parties from the regulatory relation and guaranteeing the respect of mutual commitments. As an exchange for the monopoly granted to a company, the State imposes constrains and/or requires the provision of certain obligations in favour of consumers – i.e., whereas consumers have a right to be served, firms have a right to serve. At issue is the trade-off between the possible extraction of monopolistic rents, on the one hand, and the encouragement of an efficient output level plus the provision of incentives for the firm to stay in the market, on the other.

Traditionally, such contractual relations were conceived as a discrete transaction comprising a well-defined object. Under the assumption of completeness, parties would enjoy the greatest degree of protection, because the contract could achieve first-best solutions with maximum efficiency. However, a complete contract can only be considered a theoretical construct. Since long ago both economic and legal doctrines have recognised the unfeasibility

[184] See Smith (1997a).
[185] On this, see Smith (1997a), (1997b) and (1997c); Stern & Holder (1999); Noll (2000).

of contractual completeness.[186] The economic theory focused upon 'contracting' – that is, the parties' choice of contractual form. Legal theory, by contrast, focused more upon 'contract law' – that is, the effect of legal rules on contractual behaviour, normally assuming that parties choose simple contracts.[187] Accordingly, incompleteness is understood differently in both fields.[188] Legal scholars use the term to refer to 'obligational' incompleteness – i.e., contracts in which the obligations are not fully specified for all future states of the world. Economists, in turn, use the term to refer to 'contingent' incompleteness – i.e., contracts that fail to fully realise the potential gains from trade in all states of the world. Notwithstanding these and other differences, the recognition of contractual incompleteness is universal.

Incompleteness is ubiquitous for a variety of reasons, such as the presence of transaction-costs[189] or bounded rationality. Because of these reasons, parties in a relationship can only specify contractual terms poorly – e.g., since investment is not verifiable by a court, the parties have no incentive to put the optimal level of investment in the contract. Also, parties cannot commit themselves either to keep the terms of the contract indefinitely or to renegotiate them in the future. Over this common ground, the incomplete contract theory matches NIE insights to emphasise comparative institutional aspects of governance decisions.[190]

Regulation is considered to be an 'administered contract' between consumers and utilities, with some degree of governmental oversight.[191] The administered contract approach recognises the importance of both reciprocity and the relational elements encountered in a

[186] In economics, the literature has its roots in Williamson (1975). Nevertheless, Hart & Moore (1988) is generally considered the seminal article in the field. See also Hart & Moore (1999) (defending the 'incomplete contracts' approach and providing theoretical foundations for it).

[187] For that reason, law-and-economics approach focuses upon the default rules that Courts should apply in a given case. See, e.g., Ayres & Gertner (1989). The contrast between the economic theory of incomplete contracts and the law-and-economics literature is shown by Posner (2003: 855-63) (although stating that differences are 'gradually disappearing').

[188] Ayres & Gertner (1992: 730) (arguing that 'by filling obligational gaps that are state contingent, courts can respond to both types of contractual incompleteness').

[189] However, contrast Maskin & Tirole (1999) (arguing that transaction costs do not always interfere with optimal—complete—contracting); and Posner (2003: 866-67) (arguing an ambiguity at the heart of the concept of transaction-costs).

[190] Note that a comparative institutional approach is not a necessary attribute or outcome of contract incompleteness; however, the latter allows placing a greater emphasis on the former.

[191] This view was initiated after the work of Goldberg (1976) (emphasising at 431 that: 'features which make long-term relationships between consumers and producers desirable and which further make it extremely difficult to determine at the outset the specific terms of the relationship'). See also Crocker & Masten (1996).

regulatory relation.[192] Furthermore, it stresses the incentives that all parties confront during their ongoing relationship. Regulation is seen as a long-term contracting problem not too different than contracting in private sector procurement.[193] Both situations differ primarily in degree rather than in kind. In both, relation-specific investments are seen as the main source of problems. Difficulties are however aggravated in the case of utilities, because technological reasons make relation-specific investments inevitable rather than optional. This means that whilst the risk of opportunistic behaviour may not be encountered in (public or private) procurement, it is always present in the regulatory relation. Nonetheless, since problems differ merely in magnitude but not in essence, solutions do not vary in one or other case. A variety of options similar to those that help to protect firms from each other against opportunism are also readily available to protect governments, firms and consumers from each other. As in the private sector, parties of the regulatory contract will adopt the mode of organisation that better fits with the attributes of the transaction at stake.[194]

There are a wide variety of possible regulatory arrangements that depend on both the level of political intervention and the source of investment. Regarding the level of political intervention, the range goes from private arrangements to forms of pure public ownership.[195] In the middle, a number of hybrid options provide more or less degrees of discretion to the official authority. Thus, the fiction of the regulatory contract defined the limits between which industries are left to the competitive forces of the market and which are more heavily regulated. The typical example is electricity versus telecommunications. In the latter, facilities-based competition would allow relying merely upon competition rules. Regulatory intervention would only be justified in the presence of significant market power, and only if such intervention were proportionate and necessary. By sharp contrast, it has been argued that wholesale electricity markets would be more prone to suffering from collective dominance, which makes it much more difficult to determine significant market power. Continuous

[192] See Macneil (1978) (providing a detailed discussion of the contractual alternative ways of organising transactions: classical, neoclassical and relational contract law).

[193] Laffont & Tirole (1993: 375) (stating that ideally the regulatory relationship should be governed by extensive long-term contracts); Gómez-Ibánez (2003: 9) (seeing the 'problem of an infrastructure monopoly' as a variant of private procurement); Priest (1992) (calling for a substantial reinterpretation of the regulatory relation based upon long-term contracts); and Joskow (1991) (applying an incomplete contracts approach to evaluate alternative organizational arrangements in utilities regulation and antitrust).

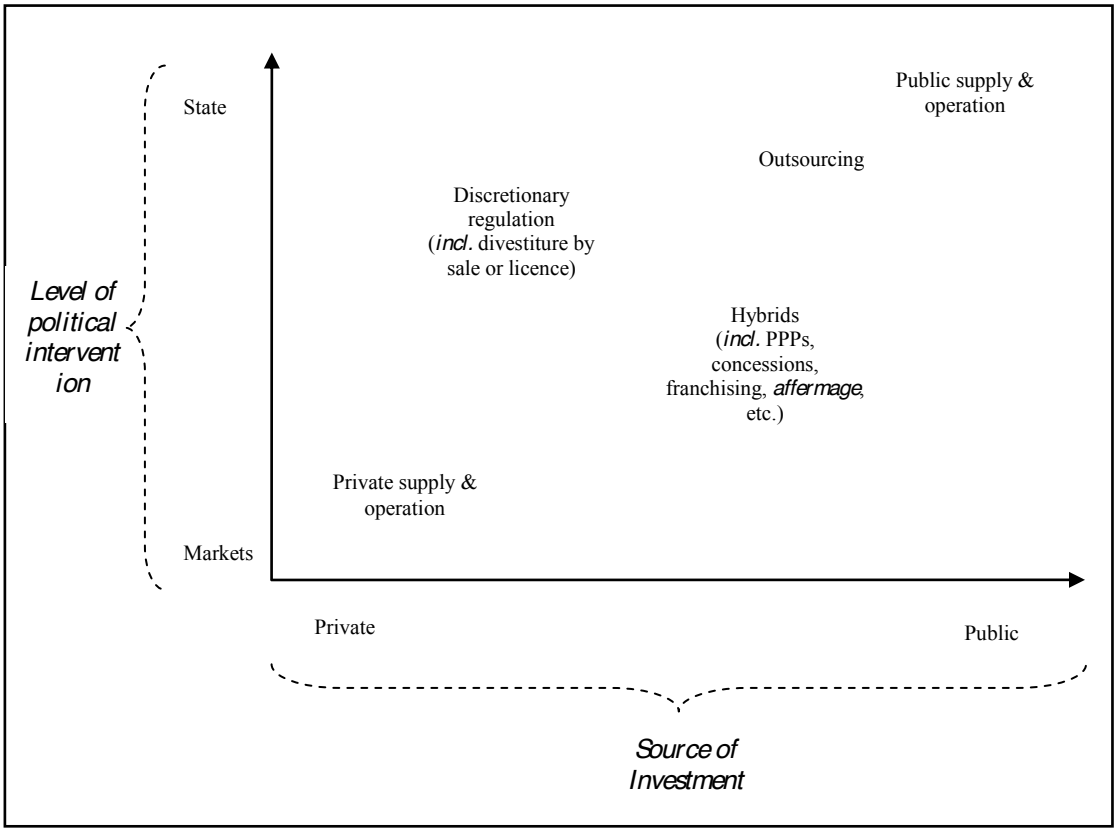
[194] See Ménard (2008: 286); Crocker & Masten (1996: 35) ('…depending on the specifics of the environment, some modes of governance will be preferred to others').

[195] Gómez-Ibanez (2006: 33) identifies eight options.

information gathering and permanent surveillance of the market would be desirable under poor conditions for competition.[196]

These regulatory arrangements may be considered alongside the options for private investment, which range from the total reliance on private supply and operation to total reliance on public provision and operation of services.[197] Alternative governance structures consider both dimensions, as illustrated in Figure 4.

Figure 4: *Range of options for regulatory design.*



The regulatory contract (or 'regulatory compact') approach has been the most common fictitious analytical framework for studying the relation between the State and regulated firms

[196] *See* Newbery, (1999) (arguing that a major role is still reserved to the regulator in the electricity sector, in contrast to the emphasis towards competition placed on the telecommunication industry).

[197] Guasch (2004), mentions up to 11 organisational forms with more or less private participation.

in US literature and practice.[198] Some commentators have even argued that the contract, although tacit and metaphorical, must be honoured by the State like any other legal contract.[199] However, that argument was short-lived.[200] Currently, it seems clear that more than the legal obligations that emanate from the contract (if there are any at all), the contract is better understood as a theoretical framework that accounts for the evolution of public utility regulation.[201]

4. Criticisms

Both the institutional environment and the institutional arrangements of the NIE approach have been challenged on a number of theoretical and practical grounds. One major drawback of the institutional environment perspective relates to the over-emphasis placed on restrictions to discretion and the level of inflexibility that this imposes on the system. This view reflects a lack of awareness that independent regulatory agencies have value not only when they have little or no discretion. If this were the case, most other features or institutions of good governance would become unimportant.[202] In practice, most regulators have some level of discretion to decide a course of action or to make administrative interpretation of the relevant statutes. As observed, the literature has already recognised that governments must not renounce to the necessary degree of flexibility while designing the regulatory framework, and thus must seek the right balance of good governance principles.

---

[198] E.g., *PacifiCorp v. Pub. Serv. Comm'n*, 103 P.3d 862, 871 (Wyo. 2004) ('The "regulatory compact" provides the fundamental basis for utility regulation. In general, the compact is a theoretical agreement between the utilities and the state in which, as a quid pro quo for being granted a monopoly in a geographical area for the provision of a particular good or service, the utility is subject to regulation by the state to ensure that it is prudently investing its revenues in order to provide the best and most efficient service possible to the consumer. In exchange, the utility is allowed to earn a fair rate of return on its rate base'); and *Office of Pub. Util. Counsel v. Pub. Util. Comm'n*, 104 S.W.3d 225, 227-28 (Tex. Ct. App. 2003) ('Under a fully regulated system, an electricity utility enters into a "regulatory compact" with the public: in return for a monopoly over electricity service in a given area; the utility agrees to provide service to all requesting customers and to charge only the retail rates set by the Commission.').

[199] Sidak & Spulber (1996), (1997a), (1997b) and (1998).

[200] A massive amount of literature has cumulated in response to Sidak & Spulber. A non-exhaustive list includes Rossi (2005: 95 *et seq.*) and (1998); Chen (1999); Baumol & Merrill (1998) and (1997); Williamson (1996); Williams (1996).

[201] Priest (1992).

[202] The exception being courts, which must be reliable enough to oversee the agency.

Another drawback is that governmental presence is taken for granted. The general view is somewhat 'optimistic': when governmental institutions are needed, they are readily available. Private investment perfectly matches with regulatory arrangements and institutions that could readily be established.

From an institutional arrangements perspective, a first ground of contention is the privileged status that NIE tends to assign to markets both as the starting point in the analysis of evolution of governance and as a sort of benchmark. Particularly in economic sociology, but also in other institutional subfields, the basic intuition is that markets are socially embedded. Critics of this approach remark that this intuition has led to the market being taken for granted.[203] As EISNER has rightly indicated, the search for non-market mechanisms is often framed as a response to market failures and not as part of a larger 'governance failure'.[204] Consequently, competition is the departure point that underlies most analysis. As will be shown further below, despite the strong case in favour of competition, that premise is not correct in all cases.

Also from a theoretical perspective, NIE's cognitive foundations have been criticised. Most studies are still built on the standard assumption of rationality, or perfect optimizing abilities of all agents, ignoring problems with bounds of human cognition. Even authors that expressly consider limits on rationality (i.e. assuming bounded rationality) still do not contemplate other important factors that may affect decision-making. Assumptions such as profit maximization and cost minimisation are not considered to be problematic. Likewise, studies of these bounds are mostly limited to individual problem-solving pursuing efficient results. All of this is controversial. Frequent conflicts of interest among parties (indeed, among actors of the regulatory relation) produce usual misalignment of interests. In resolving this, the role of power is crucial for decision-making.[205] The selection of governance structures entails a distribution of power among economic actors that help to shape the terms of the interactions.[206] Since the distribution implies strategic considerations on the part of the parties

[203] E.g. Krippner (2001).
[204] Eisner (2010).
[205] See Perrow (1986).
[206] Eisner (2010).

who seek to control the exchange, the result may be an asymmetrical distribution of power incompatible with notions of efficiency.

The theoretical criticisms also extend to motivational aspects underpinning some of NIE's analysis.[207] The focus of the criticism is on the central role of opportunism as the main factor affecting transactions. As in the case of the P-A model, the role of other factors such as reputation and trust are not considered. Transactions are treated as though they occur without any knowledge of previous transactions involving the parties concerned. Certainly, opportunism is a concern. However, the problem of opportunism is architecture-specific and therefore only secondary for regulatory design.[208]

The final question is whether the regulatory contract still a function to fulfil as a framework to analyse the regulatory relation or if it simply must be left behind. In the US, the contract analogy still has many advocates. They consider the contract to be a useful regulatory legal mechanism to attract investment. ROSSI, for example, has pointed out that the contract remains important as a framework for analysing regulation of public utilities since it still is able to explain and provide solutions for current concerns. However, he argues, the focus must change towards a bargaining approach – what he calls 'a political process theory of the regulatory contract'.[209]

To be sure, there are interesting features of the regulatory contract tradition that may and should be kept in any model of analysis. For instance, an interesting outcome of the economic theory of incomplete contracting is the prediction that contracts do not need to make reference to specific events in the world (in legal terms, they do not need to describe absolute or conditional obligations referring to 'physical' contingencies). On the contrary, what matters most is the specification in the contract of the bargaining procedures that parties must follow at the time of performance. Such procedures are designed to force parties to divulge

[207] E.g. Hodgson (2004); Love (2005).

[208] Aoki (2010). See Chapter V for a further development of this argument.

[209] See Rossi, (2005: 3) (arguing that contract remains relevant, but linked to a bargaining approach—a political process theory of the regulatory contract). *See also* Gómez-Ibánez (2003) (suggesting that legal regimes should move toward contractual arrangements); and Freeman (2000) (applying the contract analogy to other regulatory settings different from utilities).

their realised valuations and act efficiently on the basis of them.[210] The question is whether the fiction of the contract is still essential to keep these positive features and pursue the central regulatory goals. I consider that it is not.

Utilities regulation should move beyond the regulatory contract towards new frameworks of analysis. In fact, because the term 'contract' has been used with too little precision and caution in the regulatory context, it has given rise to many unfruitful and highly polarised debates – and that is in itself a strong reason to get rid of such a rough analogy. In fact, this framework has not been widely accepted in the UK. The reasons are mainly related to the essentially political nature of regulatory relations (which make it difficult to tie down regulatory discretion in ways which resemble contractual relations), the ambiguity as to who is principal and who is agent (with the attendant danger of adopting a single theoretical category for relationships which are radically different) and the growth of competition and social regulation (which fits best with other accounts of regulation, such as that of a network of stakeholders).[211]

CHAPTER CONCLUSION

In one way or another, the most important accounts of regulation were based on the sort of 'spirit of resistance' to government interference that MILL denounced more than a century and a half ago.[212] The models' assumptions normally have a cognitive bias towards markets as best adjudicators of competition – the prominent objective to pursue. Markets have a dominant role, so the government must act only to, and when it can, facilitate their functioning. This view is accurately captured in the words of FRIEDMAN, who envisaged the government as limited to doing 'something that the market cannot do for itself'.[213] As a consequence, the market has been raised as the benchmark institution. Likewise, most

[210] Posner (2003).

[211] Prosser (2005b). See also Prosser (1997) (stating that the regulatory contract is incompatible with a pluralistic view of regulation that includes a variety of goals and interests affected).

[212] See the quotation at the beginning of this chapter.

[213] Friedman (1962: 27).

traditional models focus single-mindedly on competition as the exclusive goal to be achieved. Alas, both assumptions are unrealistic.

There is no reason to assume that the market will be the institution of choice.[214] As social scientists have emphasised, the market is only one of several potential governance mechanisms available for institutional design. Market, in this sense, only represents a decentralised system of exchange. Even under the assumption of perfect governmental oversight, rational actors will only maintain safety levels to avoid penalties. Furthermore, as it has been shown, the assumption of rational actors is at most weak. All of this leads to the recognition that governments can and must play an important role.

Many years ago Karl POLANYI drew attention to the fact that the government is a central means for tempering and alleviating the deep dislocation associated with a dynamic economic system – a system that he understood as the societal way to meet material needs. Absent government discipline, he claimed, the entire market economy would have imploded.[215] In the far more modest context of this work, that claim is equally pertinent. High-risk, capital-intensive industries such as those subject to economic regulation, cannot survive without a great extent of regulation. Furthermore, in some cases private firms would not even enter these sectors without public financial support for research and development, underwriting risk, and guaranteed profits. One of the main lessons of the most recent account of regulation is that reforms to create competitive markets normally only redeploy regulation; they do not deregulate. The outcomes are a mix of successes and failures. They largely depend on the objective sought, among a wide variety of social and economic concerns that may underpin governmental action.

In this sense, it is not only necessary to appreciate that 'different mechanisms for allocation may be justified by different values, and whether we choose markets, a right-based approach, or one based on social solidarity will depend on which values we wish to promote in a

[214] Eisner *et al.* (2006: 7).

[215] As he eloquently once expressed at the beginning of one of his major works: '[T]he idea of a self-adjusting market implied a stark utopia. Such an institution could not exist for any length of time without annihilating the human and natural substance of society; it would have physically destroyed man and transformed his surroundings into wilderness'. See Polanyi (1944: 3).

particular context'.[216] It is also necessary to acknowledge the potential tensions between the objectives of the regulatory regime (at system-level) and the objectives of each of the multiple parties that participate in the regulatory process (at organisational-level). This is a problem largely left aside by traditional models, because they tend to focus on a two-tier hierarchy comprising regulators and firms. In practice, however, incentives and objectives need to be balanced across several parties. The balance in turn affects the outcomes and how these meet the expectations of the different parties.

[216] Prosser, (2005a: 38).

CHAPTER II

BROADENING THE ANALYTICAL FRAMEWORK FOR THE STUDY OF ECONOMIC REGULATION

> 'My strong opinion is that no abstract principle, and no absolute rule, can guide us in determining what kinds of industrial enterprise the State should undertake, and what it should not'.[217]
>
> - William S. Jevons

OVERVIEW

Regulatory changes are the result of path dependence – meaning that they are based on past experiences. Any method of analysis needs to acknowledge that the analytical frameworks studied in the previous chapter are not mutually exclusive. Examples abound. Notwithstanding their differences, both P-A models and NIE approaches recognise the importance of institutions and maintain that they are susceptible to comparative analysis. Also, as NIE and the regulatory contract traditions remark, incomplete contracting must be the base of the approach – law-making cannot be considered complete not only because there are enormous difficulties for lawmakers in predicting the future, but also because there also are numerous ambiguities stemming from the political process that play a major role. In many occasions, incompleteness might become the most desirable outcome. Actors will prefer to adopt a strategy of incompleteness that allows passing an imperfect law or regulation rather than having no legislation at all or risk over- or under-inclusions. These insights are complemented by the emphasis on information asymmetries, incentives and control that underlies the P-A perspective. All in all, traditional analytical frameworks of regulatory analysis are largely complementary and have left important lessons that cannot (and should not) be eluded.

[217] Jevons (1833: 267).

At the same time, however, the previous chapter has also stressed that traditional frameworks of analysis have important shortcomings. Chiefly, they tend to focus on a single objective (i.e. competition), show a strong preference for a single institution (i.e. the market) over others, and mostly ignore motivational and cognitive failures (i.e. are strictly based on rationality). Besides these theoretical concerns, the following chapters will highlight a number of practical shortcomings of the current approach to economic regulation – some of them long-standing concerns. So far, there is no model of regulation that addresses all these theoretical and practical factors explicitly.

The task remains how to develop a more comprehensive framework of analysis to study economic regulation – a framework that, building on the important insights of previous models, provides better guidance on when to make use of alternative regulatory tools and strategies. Ultimately, the relevant questions for regulatory design remain the same: to what extent is it essential for the State to fulfil certain functions (now specifically in economic regulation)? Which areas need stronger rules and which ones a lesser degree of intervention? In order to provide an answer, it is necessary to acknowledge that there are a number of different forms of intervention whose main use is a function of the nature and the context of the policy issue to be addressed. Recognising the limitations of different policy tools, it is possible to make a 'plea for pragmatism and pluralism' and acknowledge 'the value of designing complementary combination of instruments, compensating for the weaknesses of each, with the strengths of others, while avoiding combinations deemed to be counterproductive or at least duplicative'.[218] In other words, it is necessary to expand the framework of regulatory analysis.

To advance such a framework, this chapter is structured as follows. The first part proposes three avenues to broaden the theoretical framework of analysis of economic regulation: to acknowledge the multiplicity of regulatory objectives, the recognition of the plurality of actors that participate in the regulatory process, and to concede more importance to the role of collaboration within that process. In each case the main implications are elaborated. Part II applies the expanded framework, developing the implications for regulatory finance and risk

[218] Gunningham (2009) (identifying particular architectures for environmental problems).

analysis, and the role of courts. These two applications are further deal with in more detail in the remaining chapters of the thesis. Part III assesses two consequences of broadening the framework of analysis: (1) the application of the framework results in a different theoretical approach to regulation and (2) the application of the framework results in a more 'horizontal' view of regulation. The chapter ends with some concluding remarks.

I. BROADENING THE ANALYTICAL FRAMEWORK

Any analytical framework for analysing economic regulation should acknowledge that it is not sufficient to ensure that the system as a whole is optimal (which is the single focus of traditional frameworks). It is also important to consider the parts and the possible departures from optimality. In regulation, the pursuance of *multiple objectives* requires an important level of *collaboration* between *several parties* participating in the regulatory process. These three crucial aspects are considered in the next sections.

A. Plurality of objectives and the development of legal standards

Any country's regulatory regime likely will encompass, but not necessarily rank, a multiplicity of goals. Importantly, these goals are not only economic, but also non-economic. The design of any regulatory regime often revolves around other issues such as the vertical distribution of power among market participants, strategic decisions and others. This fact, acknowledged by general literature on regulation[219], is undoubtedly applicable to economic regulation. Alas, it has been largely ignored in the literature – except for notable exceptions. In the UK context, for example, PROSSER has highlighted how social objectives have been part of regulatory frameworks ever since privatisation took place, albeit at the beginning they were inserted in a non-systematic way.[220] He argues that since 2000, however, these objectives have been given a 'more coherent' position within the frameworks.[221] However,

[219] See generally Feintuck (2010: 39-63).
[220] Prosser (2005a: 67-70).
[221] *Ibid*, at 70-3.

besides these infrequent opinions, under traditional frameworks of analysis the discussion has remained largely centred on improvements of efficiency. Whilst it remains true that economics goals remain essential in economic regulation, there should be (and there is) an increased recognition that moral, social or political objectives may sometimes provide for more direct ways to allow consumers to be better off.

Multiple goals invariably will conflict, which means that the ascension of one of them will diminish the other goal's significance. For diverse reasons, conflict is a common state. For example, there may be different degrees to attain a goal; each objective may have shortcomings; objectives may be seldom expressible in concrete terms, so there may be vast space for interpretation; and so on. The resolution of the conflict involves normative trade-offs that should make possible to 'transplant' the multiple objectives into workable legal standards.

For reasons of accountability, in principle Parliament should be the institution in charge of resolving the trade-offs.[222] Nonetheless, most legislatures are quite reticent to accomplish such task.[223] In fact, even if objectives have been inserted 'coherently' in the regulatory framework, it is unlikely that they have been ranked or that some preference is explicitly mentioned. This means that it is more probable that courts and regulators will be the ones that in practice assess how to deal with several goals, ensuring an effective regulatory process. This is, to some extent, an undesirable situation for regulators (if not sometimes for courts) – a situation that is at least complex and at most dangerous. Thus, from the regulatory design perspective, the key question is not whether economic regulation should incorporate non-economic values. Rather, the issue is the degree of freedom that regulators and courts alike have in weighing the objectives in their respective analysis.

The activity of these bodies will consist of, primarily, in giving proper weight to all goals set up by the legislator that are relevant for their activity. In practice, 'to give proper weight' will mean either trading-off the multiple objectives every time they are in unsolvable conflict with

[222] Notwithstanding Parliament is subject to rent-seeking.

[223] Legislators normally limit themselves to narrow the range of possible goals, decreasing the regulators' range of options. However, the legislative election is subject to a valuation bias: only objectives akin to the government thinking will be embedded in the frameworks.

each other, or blend them, whenever possible, into legal standards.[224] Their task depends on the development of principles of priority (i.e., rules to rank or weigh interpretative rules) and harmonisation (i.e., rules to reconcile interpretative norms).[225] A crucial question is what markets participants can expect as outcomes of this process. Even if the right principles are developed (and, indeed, there is no certainty that this will be possible), multiple objectives do not allow much clarity on the *substance* of the standards. As a consequence, expectations cannot be substantive, but only procedural. The judicial and regulatory task is limited to the creation of standards that conform to the requirements of the due process and the 'rule of law' – i.e. standards that enhance transparency, accuracy, objectivity and predictability.[226]

Undoubtedly, the task of developing the 'right' standard still remains highly complex. Multiple objectives should be synthesised into rules that allow the plurality of market participants internalise them and comply with them in a way as easy as possible. If there were a narrow, single regulatory objective, an open-ended, fact-specific standard would suffice. As a single objective arguably circumscribe the outcome (possibly to one single option), discretion is – to a great extent – curbed by the act of choosing a single goal. However, in presence of multiple (and conflicting) policy goals, fact-specific legal standards seem incompatible with the due process and do not enable market participants to reasonably anticipate the costs of their actions. What the developers of legal standards should avoid in practice is that outcomes vary with the particular facts of each case. They should shift from case-by-case analysis to simpler and more administrable standards and rules[227]; and, whenever feasible, from directly regulating the activities of the market participants to maintaining the conditions of a competitive structure.

Further increasing the difficulties, the outcome of the weighing process will also depend on the extent that subsequent developments adjust to the legal standard.[228] This depends on two conditions. On the one hand, it depends on the market participants' selection of appropriate

[224] As said, ideally courts and regulators should stick to the second task (blending) as much as possible, because trading-off is a task of Parliament.

[225] Sunstein (1990a: 186-9). The situation in this case is similar for Parliament, courts or regulators.

[226] This is compatible with the ultimate objective of regulatory law under conditions of private provision, which is to curb discretion is order to boost investment. *See* Chapter I and Levy & Spiller (1994).

[227] These may include, for instance, presumptions of illegality, well-defined exceptions, etc.

[228] As Aoki (2001) states, statutory law and regulations *per se* are not institutions if they are not observed.

means to reach the ends desired by the regulator, from a variety of suitable means fit for achieving the regulatory objective(s). On the other hand, it depends on certain 'common belief' in the standard.[229] This not only means that both the regulator/courts and the other market participants should recognise the need for the standard. It also means that the standard should reflect the objective(s) in the manner desired by all the market participants. Somewhat counterintuitive, this requirement of a 'common belief' suggests certain 'commonality' of the underlying goal(s). Commonality is ubiquitous in settings such as intra-firm environments, but it may be at first glance contentious in economic regulation, traditionally characterised by the presence of opposite interests within the regulatory relation (particularly between the regulator and the firms, due to the influence of P-A models). In absence of a common goal (in this case, between the regulator or the court and the other market participants), commentators normally sustain that there may be no possibilities for cooperation. However, as it is explained further below, commonality is both desirable and possible within the economic regulatory regime.

Summarising the main ideas of this section: (i) there is a need to acknowledge the plurality of regulatory objectives; (ii) these objectives have to be translated into workable legal standards; (iii) normally, courts and regulators are the institutions that will trade-off or blend the objectives into legal standards; (iv) plurality makes hard to achieve certainty on the substance of the standard, so expectations can only be placed on procedural aspects; (v) the standard should not depend on a case-by-case analysis, but be simpler and administrable; and (vi) the applicability of the standard depends on both the selection of adequate means to implement it and the 'belief' that market participants have in the standard.

B. Plurality of actors and interactions between them

The second feature that needs to be explicitly taken into account by any framework of analysis is that, as seen in Chapter I, regulation is not a two-tier hierarchy limited to regulators and firms. This notion is, however, the basis of all traditional frameworks of analysis described in that chapter. Allegedly, the reason may be found in the view of the firm

[229] There is a third condition: the information available for each actor.

that underlies most of those frameworks. According to the traditional understanding, widely shared in the literature, businesses are a vehicle to maximise returns to the owners of capital – that is, shareholders. The firm has a binding fiduciary duty to put the needs of the owners first and increase value for them.[230] When applied to economic regulation, the task of the regulator becomes to 'emulate' the competitive environment and therefore maximise the returns of shareholders in order to boost private investment.[231]

However, the traditional view of the firm has been contended – if not surpassed – by the idea of the firm '...creating as much value as possible for stakeholders, without resorting to trade-offs'.[232] This view is the so-called 'stakeholder theory'. Managers are tasked with the duty of creating value for all stakeholders and, in case of conflict of interest, rethinking the problems so that divergent interests can go together in 'harmony' (thus creating even more value for each stakeholder). True, this approach is far from being undisputed on corporate and business law.[233] Nonetheless, beyond the controversy, the theory may provide a more comprehensive basis for understanding the regulatory process, as it allows recognising the plurality of actors interacting in that process.

The term stakeholder generally encompasses any group, entity or individual that exerts influence on, or is influenced by, the firm. However, for the purpose of the analysis, it is possible to ignore certain groups who will have little or no impact in it.[235] Therefore, suffice here is to deploy a sort of 'simplified' stakeholder approach and consider stakeholders only those actors that has an economic interest in the outcome of the regulatory process or whose support is necessary for the outcomes thereof. In regulation, the analysis may focus on five particular group of stakeholders: (1) the (private) regulated firm, (2) the regulator, which comprises not only regulatory agencies, but also the regulatory activities of certain other third parties that exert some form of direct control or direction over firms, within the meaning of

[230] This view underlies not only the frameworks of analysis described in Chapter I, but also most business laws.

[231] As will be seen, this is precisely what happened during the privatisation-cum-regulation of many utilities in the UK.

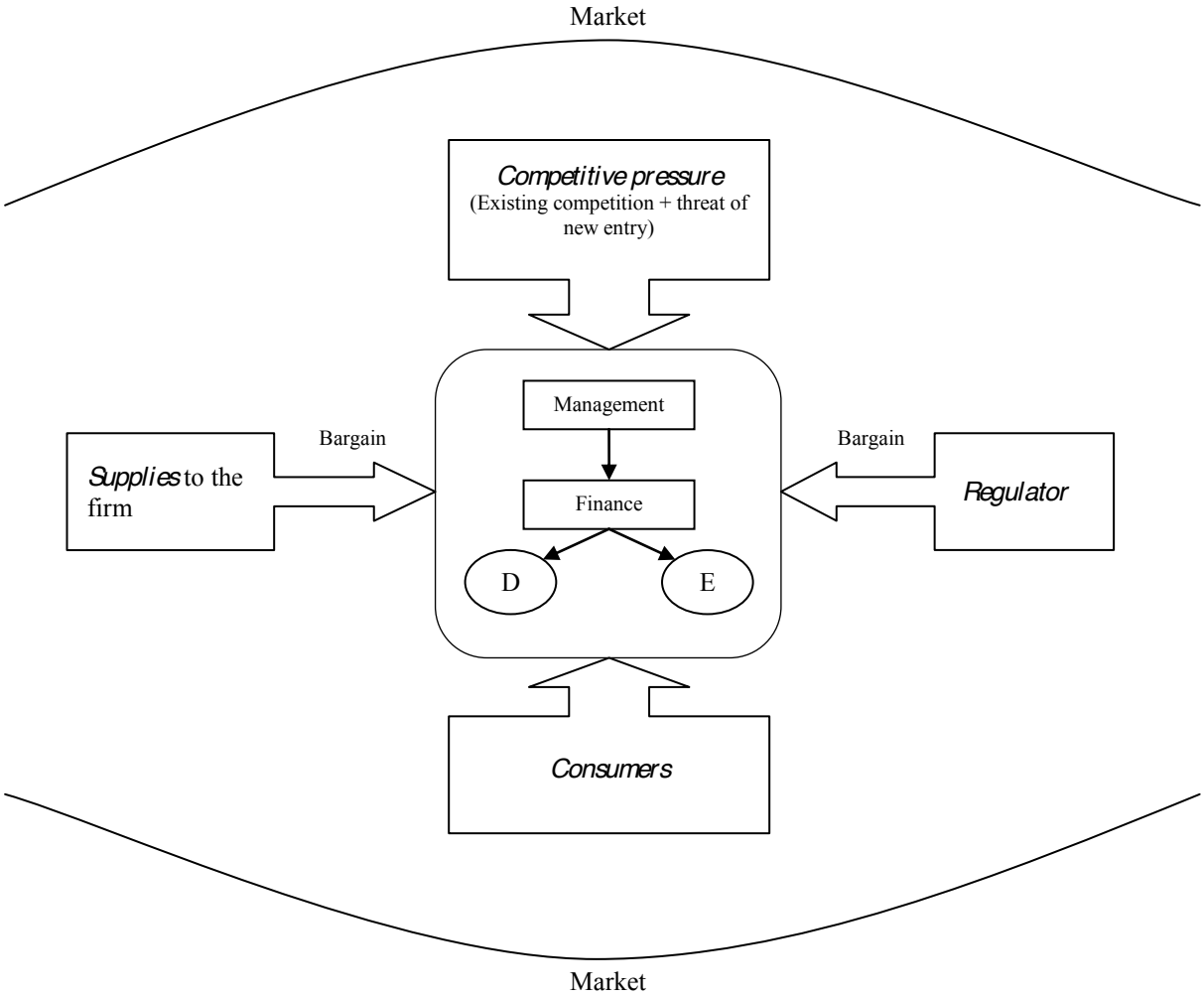
[232] Freeman *et al.* (2010: 28).

[233] For instance, confront Freeman (2010) with Jensen (2001).

[235] '[I]f the stakeholder concept is to have practical significance, it must be capable of yielding concrete actions with specific groups and individuals' (Freeman, 2010: 53).

political science models of fragmented regulation;[236] (3) consumers, which are represented both by the regulator (whose main duty normally is to protect consumers) and different associations that exert independent pressure (directly, or indirectly through the media) over the firm; (4) suppliers, which exert pressure by changing prices or quality; and (5) new entrants and existing competitors (see Figure 5).

Figure 5: *Five stakeholders acting in the regulatory regime*



[236] See Gunningham & Grabovsky (1998: esp. ch. 6).

Two important actors have been explicitly excluded from the list: courts and the government. The importance of courts is mentioned further below.[238] The government has been excluded because notwithstanding its essential role, this is only indirect: governments influence the actions of other market participants mainly by setting the formal rules of the game.[239] Beyond that, its strategy is deployed through the actions of the regulator. Indeed, the government can be considered as a separate actor from the regulator, not just because in the context of UK economic regulation regulators are mostly independent; also because beyond being a simple agent of the government, it is possible to regard the regulator as an arbitrator of the interest of most (if not all) participants in the regulatory process.[240] However, it seems more realistic to consider the government affecting regulation through the other stakeholders rather than to consider it as a constituent of the regulatory process in itself.

Not all of the stakeholders are present in each segment of the markets. Most importantly, their rights, functions and standing within the regulatory process differ. Therefore, the types of possible interactions are countless, and governance structures or architectures that arise from such interactions may adopt myriad forms or combinations.

Consider as an example the position and the distinct role that both the regulator and the (private) firms play a in the regulatory process.[242] On the one hand, since the regulator has a wide knowledge of both the systemic and the specific (local) environment, it is possible to consider its presence as essential or at least quasi-essential to the regulatory process. Essentiality means that the regulatory goal or goals cannot be achieved without some level of (quite strong) direction; quasi-essentiality implies that there are instances where such direction is 'less important' for the accomplishment of the goal(s) and hence can be softer. Regardless the degree of essentiality, however, both situations recognise that at least some level of regulatory oversight is always necessary. This conclusion is compatible with modern

[238] *See* section III.B in this chapter.

[239] *See generally*, North (1990a).

[240] In other settings, the approach has been acknowledged for a long time: e.g. Dodd (1932) (claiming that the board of directors act as trustee of all the stakeholders in a firm, including workers).

[242] Paraphrasing Alfred MARSHALL, '[t]he forces to be dealt with are so numerous that it is best to take a few at time…'.Marshall (1920: 20).

regulatory accounts described in the first chapter, according to which regulation has never receded despite the innumerable efforts to 'deregulate'.[244]

Asymmetrically, the presence of the (private) firm may or may not be strictly essential. In utilities markets firms can be deemed essential, quasi-essential or even non-essential to the provision of the service. To a great extent, firms can be replaced in their position by the State if it deems it necessary for the achievement of regulatory goals.[245] Whether the State actually decides to replace the firms largely depends on the relations between them – an issue further explored in the next section. For this purposes, it suffices to note that this standing of the firms within the regulatory process (ranging from essential to non-essential) is compatible with the analysis in Chapter I: in industries subject to economic regulation, competition is normally absent from at least some parts of the markets, either permanently or at some specific moment in time.

The position of the actors has implications for both the governance architecture that is advisable to adopt and the most suitable mode of provision of the service in different industries. On the one hand, if the regulator is essential, vertical hierarchy will probably characterise the governance architecture.[246] But if it is only quasi-essential, hierarchy is expendable and the governance architecture will be some hybrid form or the market, depending on whether the presence of the firms is more or less essential. Hybrid modes of governance (such as public-partnerships, concessions, contracting-out, etc.) serve well when neither the firm nor the regulator is strictly essential in the regulatory relation, meaning that none of them can fully attain the regulatory goal(s) in isolation.[247] When the presence of the firms is essential, but not that of the regulator, it is more advisable to use the market. On the other hand, if the need of private firms is deemed not necessary for the achievement of the goal(s), public provision of the service may be advisable.[248] This is unlikely to be the

[244] *See* Chapter I, section I.B.

[245] For example, if it deems it more efficient, although this is not a necessary condition, as it will be argued in the next section.

[246] Hierarchy 'can be defined as the asymmetric and incompletely defined authority of one actor to direct the activities of another within certain bounds'. See Miller (1992: 16).

[247] As will be seen, in this case the need to collaborate increases.

[248] Note that some weak degree of competition may arise, in which case communication between firms is likely to be very limited. However, competition is not essential, so the structure of the industry remains primarily monopolistic. An example may be the early stages of telecommunication privatisation in the UK.

common situation, though. That the firm is essential implies an immediate preference for private provision of the service, so in most settings, this would become the rule. These ideas are summarised in Table 1.[249]

Table 1: *Modes of governance*

Legal standing		Provision of the service	Governance architecture
Regulator	*Firms*		
E	QE or NE	Public	Hierarchy
E	E	Private	
QE	QE	Private	Hybrid
QE	E	Private	Market
QE	E	Private	

Summarising the main ideas of this section: (i) there is need to acknowledge the plurality of actors involved in the regulatory process – i.e., to identify a 'stakeholder map'; (ii) for the purposes of regulatory design, it suffices to deploy a 'simplified' stakeholder approach, which focuses on the main relevant actors; (iii) not all of the stakholders are present in each segment of the markets, and their rights, functions and standing within the regulatory process differ; (iv) as a consequence, interactions between them are countless, and governance structures may adopt myriad forms or combinations.

C. Recognising the important role of collaboration

Plurality of objectives and actors are not the only important features of a broadened analytical framework. The last feature of current regulatory regimes that needs to be better acknowledged is the increasing importance of collaboration in regulatory practice. According to the traditional view of economic regulation, the relation between the regulator and the firms affected by regulation is adversarial.[253] As seen in chapter I, the prevalence of the 'deregulation' view during the last three decades, driven at least partially by neoliberal thinking and strategies and individualistic premises, strengthened the notions of antagonistic

[249] As will be exposed in the next section, the standing of the parties and the mode of governance have implications for the regulatory strategy that will be chosen and applied.

[253] E.g. Breyer (1982: 6); Spulber (1989: 99).

players at the expense of solidarity. In regulation, such 'endemic conflict'[254] mainly reflected the US context of adversarial procedures and rather confrontational bargaining positions of business and consumers. However, even if the premise were true, adversarial positions do not imply the impossibility of collaborative outcomes. In fact, evidence does not always support the adversarial view. In practice, it is possible to find a fair degree of collaboration in key regulatory areas.

For instance, in the UK practice firms and regulators have numerous bilateral meetings during the price control period, as well as a number of interactions throughout the process (for example, by the issuance and answering of consultation documents[255]). Moreover, it is not uncommon for regulators and firms to sit together in working groups to draft licence conditions – it might even be the case that the drafting of heavy technical aspects is mostly left to the firms or private consultants, and only checked by regulators. In addition, in some sectors collaboration is embedded into the regime. In the UK energy industry, for example, a convoluted system of codes over-sighted by the regulator forms one of the core parts of the regulatory regime. Likewise, important parts of the telecommunications industry depend on the common agreement on technical standards.[256] In reality, the regulatory process might be too complex and intricate to be based purely on individual, selfish behaviour leading to confrontational positions.

The framework of analysis should recognise that collaboration is a key regulatory objective. Cooperative frameworks have been developed in a number of situations in law and economics – for instance, regulatory enforcement[257] and contract enforcement[258], among others. There is no fundamental reason why they should not work in economic regulation. Certainly, it is idealistic to assume that the whole utilities regulatory regime can be

[254] Balleisen (2010: 457).

[255] For example, during the last electricity distribution price control review (DPCR5, for the period 2010-15) the energy regulator issued five different consultation documents.

[256] In the water industry, the 'Cave review' (Cave, 2009) recently emphasised coordination when it recommended that the Government and regulators form a national water research and development (R&D) body and agree on a shared R&D vision for the industry.

[257] E.g. Scholtz (1984); Ayres & Braithwaite (1990).

[258] E.g. Rubinfeld & Cooter (1989) (enforcement 'in the shadow of the law'); Dixit (2004) (enforcement without any degree of governmental protection).

exclusively centred on this goal.[259] The claim here is more modest and realistic – a comprehensive view of regulation should *also* pay heed to the less confrontational exchanges taking place within the regulatory relation. Therefore, recognising the conflicting views, but also the multiple avenues for collaboration within the regime, the regulatory framework should lead stakeholders to work together in ways they would not do without institutional constraints.

To achieve collaboration, the framework must provide incentives for *cooperation*; create rules that allow for *coordination* of individual action when necessary; and create rules that direct *communications* amongst stakeholders. These three grounds are the fundamentals of a good regulatory design; moreover, as CALVERT rightly stated, they are essential for regulatory design:

> [T]he establishment and maintenance of endogenous incentives for *cooperation* among selfish individuals; the achievement of *coordination*, that is, standards, organization, or conventions, in complex settings; and the way in which *communication*, which turns out to be central to the first two phenomena, can take on a life on its own in the definition and functioning of institutions.[260]

Cooperation and coordination are not the same. They look at the individual and the group, respectively. The defining feature of cooperation is the engagement in some act in order to advance the goals of others and hence generate benefits for them. The action is taken regardless the attitude adopted by others at that specific moment.[261] Importantly, this does not mean that a selfish individual cannot take the action, only that it cannot be predicted that such a person would adopt the action. In fact, the immediate outcome of the action can be (and in most cases is) negative for the individual that carries it out.[262] This common situation is exemplified by the archetypical Prisoner's Dilemma game, whereby two players (i.e., prisoners) do not have an incentive to play the strategy that is in their better interest (i.e., to

[259] In fact, that would deny most of the valuable lesson learned from previous models of analysis.

[260] Calvert (1995: 218). On the assumption of selfish individuals, *see supra* Chapter I, especially section II.B.

[261] For example, see Ellickson (1991: 167) (stating that 'members of tight social groups will informally encourage each other to engage in cooperative behavior').

[262] One rather extreme example of cooperation with negative consequences is the case of suicide bombers, who sacrifice themselves for the efficacy of a collective goal.

remain silent in response to a prosecutor's questions) because each of their individual, most harmful strategies (i.e., confessing) strictly dominates in the game.[263] The main lesson is that without some direction or previously established practice or rule, cooperation is unlikely to occur spontaneously. Nonetheless, since regulation implies interactions between more than two individuals (stakeholders), problems normally go beyond the Prisoner's Dilemma situation, even if repeated. A broadened view of regulation resembles more closely more general standard collective action problems, of which the Prisoner's Dilemma is only the simplest variant. Within the regulatory relation, collaboration is closer to the notion of complementarity, which gives the idea of joint efforts to achieve the desired objective. Thus, coordination surpasses cooperation as key regulatory goal.

The defining feature of coordination is the aim to attain a common goal or goals that are difficult to reach.[264] Coordination can be achieved either voluntarily or in response to hierarchical rules. The option will depend on the barriers to coordinate. However, barriers are likely to take place, and overcoming them is fundamental (even in absence of cooperation). In other contexts, the agreement on the goal (or, most commonly a plurality of them – see next section) might in itself be problematic. But in regulatory settings the goal is to a large extent decided and imposed by the legislator, so obstacles to cooperation/coordination arise for different reasons. Either communication is imperfect and does not permit to choose amongst several conflicting ways to attain the goal (for example, because of conflicting views on the efficiency of the outcomes); or the actors have divergent preferences.[265] As a consequence, hierarchy works better to overcome barriers than reliance on the market.

Finally, communication plays a vital role in achieving both cooperation and coordination.[266] Communications is 'a process involving two information-processing devices. One device modifies the physical environment of the other. As a result, the second device constructs

[263] See generally, Baird *et al*. (1994).

[264] An archetypical example of the need for coordination, highly relevant in regulatory environments for utilities, is the setting of standards. If a firm has a pre-existing large investment in a potential standard (for example, due to its preference for the use of certain technology), it would prefer the adoption of such standard; but if another standard is preferred, the firm will still be better-off adhering to the chosen standard rather than having no standard at all. Note, however, that if one firm prefers the original production approach, the problem is not one of coordination anymore.

[265] These ideas have been captured by game theoretical literature in the simplest game of pure coordination: the 'battle of the sexes' game. See Luce & Raiffa (1957: 90-1).

[266] Sally (1995) (showing the effect of communication in experimental work on cooperation).

representations similar to representations already stored in the first device'.[267] Even if it is infeasible to enforce commitments, the ability of the market participants to transmit messages to the others and be understood increases the array of opportunities for interaction that are needed to reach collaborative outcomes. Allowing them to communicate more easily, each constituent of the regulatory relation will come to know the interest, belief and expectations of the others through communication rather than rational self-reflection. Communication can be either formal or informal. The expression of a message not only serves as a way to justify the creation of norms by the State or the regulator (i.e., to design the regulatory framework) or a given actual behaviour by any of the stakeholders. It may also signal the intention to follow certain patterns of future behaviour, creating expectations in other constituents. In this case, the message acts as a sort of minimal focal point '*a la* Schelling'[268], which helps to improve the prospects of collaboration clarifying what is expected from whom, what is considered defection and, eventually, what the punishment is. Communication fulfils both an explanatory and predictive role.

As game theoretical literature has shown, communication increases with repetitive interaction.[269] Repetition helps to give meaning to messages that may be given outside the setting of pre-existing conventions or social norms. Continuous interaction may also help to give a meaning to a variety of seemingly unconnected messages with apparently no inherent individual meaning. A collaborative model needs to include routinisation as one of its core principles, in order to enhance the possibilities of 'learning to cooperate'.[270] But, at first glance, the need for routinisation and the constant possibility of perfectibility that implies the search for stable collaboration seems to make long-term design unfeasible. Iteration involves behavioural variations from one stage to the next, complicating the attempts to maintain normatively defined patterns of interactions. It might be argued that this is a main drawback of relying on collaboration as base for an analytical framework. However, long-term design is still possible despite the continuous search for improvements. Regulatory rules can create

[267] Sperber & Wilson (1995: 1).
[268] Schelling (1960).
[269] For a non-technical approach, Dixit & Nalebuff (1991: 101) (indicating that cooperation may be sustainable if the game is repeated even a finite number of times, but that number is unknown).
[270] Calvert, (1995: 258).

stability and predictability even though they are subject to change.[271] In this sense, pressures for change arising from the needs of cooperation and coordination are not that different than pressures arising from exogenous forces, such as competitive foreign demands or global economic conditions, or endogenous forces, such as technological evolvement or new entry. Adaptation is an essential characteristic of any regulatory regimes that is not altered by repetitive interaction.

Recognising the importance of collaboration and its elements, it is possible to complement the conclusions of the previous section regarding the mode of governance and the type of provision of the service that better works in a give setting. The analysis of collaboration also allows the determination of the regulatory strategy that best suits the characteristics of a given industry.

As before, let's focus on the relation between regulators and firms. On the one hand, under conditions of private provision of the service, if possibilities for coordination are limited (for instance, because communication between firms is not possible or undesirable), command and control regulation will be recommendable.[272] The same is applicable when collaboration is necessary, but communication is difficult or imperfect because of its costs.[273] In this case, there is a need to force communication between firms. However, as regulatory rules can create a suitable environment for at least some level of collaboration and some degree of communication is feasible, this should not be considered the rule. This fact is compatible with the conclusion of most literature that argues that C&C regulation should be kept at minimum. So if collaboration is necessary, but communication is difficult or imperfect because of its costs, the need to overcome the costs should be achieved applying market mechanisms or a self-regulatory strategy. The election depends on the level of collaboration possible to implement in the market in order to comply with the legal standard. If fierce competition

[271] The definition of the patterns of interaction is, however, not impossible. The mere existence of the governance regime produces certain path-interdependence that constrains changes.

[272] Regulation in this case is strictly a 'sustained and focused control exercised by a public agency over activities that are valued by a community' (Selznick, 1985: 363).

[273] Possibilities are multiple. Archetypal costs that may undermine communication include familiar transaction costs such as the time and effort needed to establish communication and reach a consensus, and the cost of processing the information (including the cost of discarding potentially useful information). Other costs come from misunderstanding – either because the sender fails to communicate the desired message adequately or the receiver fails to decode the message. Also, there may be delays in the communication of the message; some information may be missing, and so on.

predominates, collaborative outcomes are unlikely to arise – they may even act in detriment of the regulatory goal(s). In such case, incentive-based regulation arises again as the most recommendable regulatory strategy. Conversely, if there is more possibility of collaborative outcomes, self-regulation should be implemented, placing the decision of how to comply with the regulatory goal(s) in the firms.[274] These ideas are summarised in Table 2, which extends the results of Table 1.

Table 2: *Extension of table 10 with role of collaboration and regulatory strategy*

Essentiality		Provision of the service	Governance architecture	Collaboration	Regulatory strategy
Regulator	*Firms*				
E	QE or NE	Public	Hierarchy	None	---
E	E	Private		Limited	Command & Control
QE	QE	Private	Hybrid	Displaced by competition	Incentive based
QE	E	Private	Market	Displaced by competition	
QE	E	Private		Predominant	Self-regulation

In sum: (i) one of the main tasks of the regulatory framework is to facilitate interactions in order to exploit the possibilities of collaborative outcomes; (ii) collaboration involves three aspects, cooperation, coordination and communication; (iii) cooperation, but mainly coordination, are the result of interactions that can only take place with the assistance of the framework, which means there is a need for some direction or previously established practice or rule to achieve a collaborative outcome; (iv) communication is vital for achieving collaboration and therefore should be considered when designing the rules; and (v) the analysis of collaboration allows the determination of the regulatory strategy that best suits the characteristics of a given industry.

II. SOME APPLICATIONS

This section explains in general terms how the broadened framework can be applied in two areas of regulation that are further detailed in the remaining chapters of the thesis. First, the implications for regulatory finance and the analysis of regulatory risks are stated. They are

[274] Self-regulation includes here some level of governmental intervention, sometimes referred as 'enforced self-regulation' (e.g., Ayres & Braithwaite, 1990: 305) or 'meta-regulation' (e.g., Parker, 2002), or 'mandated self-regulation' (e.g., Bardach & Kagan, 1982). On these concepts, see generally Coglianese & Mendelson (2010).

treated in chapters III and IV. Second, the broadened framework invites to consider an enhanced role of courts within the regulatory analysis – particularly specialised courts. This issue is treated in chapter V.

A. Regulatory finance and risk

As will be shown in Chapter III, regulatory finance is at the core of regulatory studies. Thus, any framework of analysis needs to consider how to address the complexities of regulatory finance. The situation is not different for the broadened approach advocated in this chapter. As in other areas, traditional financial theories have not embrace (at least in full) the idea of balancing competing interests of a broad group of stakeholders. However, several studies have advanced ways to apply such way of thinking into finance and have demonstrated that corporate financial policy also depends on the role of non-investor stakeholders. If this is true, and in regulation the regulator plays a key role in financial policy (as argued in Chapter III), it follows that the regulatory process must be widened in order to consider the interests of all actors in the regulatory relation.

Early literature started distinguishing between explicit and implicit claims of stakeholders. The latter come from expectations that result from vague promises, past actions or past experiences with the firm, such as tacit promises of timely deliver, product enhancement, security, and so on. By contrast, explicit claims come from legally binding contracts with stakeholders.[286] The important point for the purposes of this chapter is that both types of claims influence the financial policy of a firm. Consider first the situation when only explicit claims are considered. In such case, stakeholders other than investors do not play an important role in the financial policy of the firm, because their explicit claims are generally senior to those of shareholders and bondholders. If the risk of financial distress is low, the explicit claims of stakeholders are nearly risk-free and, consequently, cannot explain

[286] As Cornell & Shapiro (1987: 6) express: 'The distinguishing feature of implicit claims...is that they are too nebulous and state contingent to reduce to writing at a reasonable cost. For this reason implicit claims have little legal standing. Typically, the firm can default on its implicit promises without going bankrupt or liquidating'.

variation in the value of the firm.[287] Consider now the alternative (and more realistic) scenario. In practice, many of the claims issued by management to stakeholders other than investors take the form of mere implicit claims. Unlike explicit claims, payouts on implicit claims are not set (i.e., they are uncertain), so the price stakeholders pay for them depends on the overall condition of the firm, including its financial policy. Even when the risk of financial distress is low, the value of the implicit claims is will be sensitive to information about the firm's financial condition. As a consequence, the market value of the firm depends on the price at which both types of claims, explicit *and* implicit, can be sold.

On regulation, many implicit claims are dependent on actions or decisions made by regulators. This has at least two direct consequences. On the one hand, the costs of financial problems are likely to be larger than the direct cash drain indicates, so the information requested by regulators and the regulatory analysis should take into account this fact. On the other hand, regulators must take account of the manner in which they deliver information to the market, because it may affect the price of implicit claims. Most (if not all) non-investor stakeholders, including customers, employees and suppliers, make decisions contingent, at least in part, on that information. Exactly as information arising from the firm (revealed mainly by the price of stocks), regulatory announcements are likely to affect both investors and non-investors stakeholders, whose responses will affect the value of the firm and, ultimately, the regulatory outcomes. From this perspective, announcements that may imply rather large payoffs on implicit claims in the future should be supported by present actions (including the establishment of a suitable financial structure) that are supportive of the claim. In regulated environments, the risk of financial distress is a function of the ability of *both* the regulated firm and the regulator to honour implicit claims.

In addition, the broadened approach also makes possible to identify specific financial implications linked to each mode of governance architecture. Thus, hierarchy shall have some strong level of direction over the finance of the firms. Evidently, the strongest control implies that the State finance the firm. Softer forms of direction normally comprise some 'soft' degree of control of the capital structure of the regulated firms, as further detailed in Chapter III. Hybrid mechanisms will have contingent forms of control depending on the level of

[287] On the concept of risk of financial distress, see Chapter IV.

governmental direction. Generally, market governance should imply no control on the finances of the firms.

Table 3: *Summary of alternative modes of governance*

Mode of governance	Regulatory strategy	Financial implications
Hierarchy	Public provision	Control (strong or soft)
	Command and Control	
Hybrid	Incentive-based	Partial, contingent control
Market		No control
	Self-regulation	

B. An enhanced role of (specialised) courts

A second major consequence of the application of the broadened approach advocated in this chapter is a call for the enhancement of the role of courts in the regulatory process. As indicated, along with regulators, courts are crucial in the process of weighing the plurality of economic and non-economic objectives established by the legislator, either by trading-off or blending regulatory goals. In addition, there are a number of other aspects that commonly requires a third entity with sufficient capacity to scrutinise the decisions made by the stakeholders. Courts should be the final arbiters in the regulatory process.

This claim is contentious. Traditionally, courts have been deemed as unsuitable to deal with 'polycentric' questions affecting a large number of disparate interests.[288] As will be seen in Chapter V, a rather similar reasoning was taken to the extreme in the UK, were judges have traditionally been reluctant to get involved in regulatory affairs. But for the reasons indicated in that chapter, this claim seems hard to sustain under the present conditions. As the UK case shows, the criticisms against the judiciary may be largely overcome with the introduction of specialisation in the regulatory system. It is submitted that specialised courts are in the best position to scrutinise the regulatory decisions arising in a setting characterised by multiple objectives and actors. That is, even if specialisation does not fully resolve the problem of polycentrism, at least it largely decreases it. Specialisation provides judges with

[288] Fuller (1979: 397) (recognising that the polycentric problem is a matter of degree).

characteristics that more 'generalist' judges do not possess, which put them in a particular advantage to deal with issues arising from a governance view of regulation.

Specialisation is considered here as a comprehensive concept composed of three main parts. The first one is *expertise*: a greater level of knowledge in the specific substantive area to be subject to judgements – in this case, economic regulation. Note that it is not necessary for the court or tribunal to be more expert than the regulator.[289] It suffices that it possesses an important knowledge in the area so as to decide on the details of the question at issue. As long as judges' knowledge (and/or that of their supporting teams) improves, they may claim closer proficiency to regulatory authorities. The second part is *experience*: the accumulation of knowledge or skills. A court or tribunal also specializes when it must revise a large number of decisions coming from the same agency – particularly if those decisions are of similar kind or nature. This means that, contrary to common belief, expertise is not the defining feature of specialisation.[290] Finally, specialisation also comprises what I call *object-specificity*: the specific aim of introducing logical coherence to and protecting (at least one or some of) the objectives of one part of the legal system. The legislator considers that there are valuable public policy reasons to establish a specialised jurisdiction entrusted with a task which existing jurisdictions – it is deemed – are unable to accomplish. The three components of specialisation are equally relevant and arguably inseparable. A specialised tribunal must comprise all these attributes in order to match the regulator's capabilities.[291]

The three parts of the concept of specialisation are in line with the most recent developments in industries subject to economic regulation. At least in the area of economic regulation and its 'twin' of competition law, courts owning economic expertise have increased sharply, as well as that of their supporting teams. This has heavily contributed to the confidence in their

[289] Some commentators have wrongly made this assumption (e.g., Levin, 1985: 43).

[290] Indeed, day-to-day business does not transform the court in expert, but helps it to specialise. For this reason, it is a mistake that some legislation requires adequate *expertise* (for example, the article 4 of the 'Framework Directive', n 13, which indicates that the appeal body, which may be a court, 'shall have the appropriate expertise to enable it to carry out its functions effectively').

[291] Because, as Louis L. Laffe, pointed out, 'It is not merely the presence of expertness, but the wide-reaching and systematic character of an agency regulation which tends to choke out the normal jurisdiction of the courts' (Laffe, 1964: 1041).

own capacities to tackle complex concerns related to industries in this area.[292] Also, courts have been confronting a wider variety of regulatory matters and more cases, so they are more acquainted with the topics. Finally, the number of specialised courts in the area of economic regulation and competition law has recently increased.[293]

III. CONSEQUENCES

Broadening the framework of analysis in the manner developed in the previous sections has at least two direct consequences. First, it results in a different theoretical approach to study economic regulation. Second, the view of regulation as a top-down process is defeated by a more horizontal recognition of the relationship between firms and regulators.

A. A distinctive theoretical approach

The framework of analysis is dynamic in nature.[320] As stressed at the beginning of Chapter I, the foundations of any institutional framework of analysis need to consider a profound and sophisticated understanding of the structure of the industries at issue and how that structure evolves. Since institutional evolution is incremental[321], the modes of governance previously identified can also be sequentially changed for the consecution of the chosen objective.

From a theoretical perspective, the framework departs from previous analysis in several ways. First, collaboration makes the analytical framework move away from the individual constituents of the regulatory relation and their motivations fully formed before social

[292] An example is the application of what has come to be known as 'regulatory antitrust'. The mighty juggernaut of competition policy has been used with force to achieve new objectives emerged after deregulation. As a consequence courts dealing with antitrust issues have tended to intervene more in areas traditionally reserved almost exclusively for regulators, adopting an increasingly proactive role. See Monti (2008) (arguing in favour of a case-by-case assessment to determine whether the application of competition law would cut across regulatory policy choices; competition should not apply when it could harm regulatory goals).

[293] The last chapter presents an example of how specialised courts have operated in the UK practice.

[320] Other dynamic approaches have been recently advanced for particular utilities. See, e.g., the important 'ladder of investment' progression of Cave (2006) (see also Cave & Vogelsang, 2003).

[321] North (1990a: 83) (stressing that 'change typically consists of marginal adjustments to the complex of rules, norms, and enforcement that constitutes the institutional framework').

interactions. As seen in Chapter I, traditional P-A models put the individual at the centre of analysis – monitoring, rewarding, punishing and even manipulating the individual in order to achieve a desired outcome.[322] Conversely, collaboration emphasises the interactions between the stakeholders of the regulatory relation as the main object of the design. Likewise, cooperation and coordination rise as an alternative view to the markets and hierarchies dichotomy introduced by early strands of NIE.[323] Particularly, they allow a better recognition of the constitutive role of law and public policies in markets.[324] The task when designing an analytical framework is to find and learn the dynamics of cooperation and coordination in order to build successful collaboration when the processes and outcomes are normatively positive, and disrupt normatively harmful collaboration. The identification of those dynamics is, however, beyond the scope and needs of this work.[325] Suffice is to highlight here the need for collaboration and, as will be seen, its importance for the analysis of utilities regulation.

In addition, the analytical framework needs to move away from the focus on motivational factors.[326] Particularly, it needs to move away from the notion of 'opportunism' much emphasised by some strands of NIE literature (but severely criticised by other commentators) as the key factor in regulatory design.[327] Indeed, as long as some elements of incompleteness remain in place in the process of design, it is impossible to completely remove opportunistic behaviour from the regulatory relation. Likewise, other by-products of incompleteness such as moral hazard, misrepresentations, and so on, will also be ubiquitous. However, that does not mean the avoidance of opportunism constitutes the central task. As AOKI has rightly stressed (within the context of the firm), organisational modes of governance are '…not selected primarily in order to control people's opportunistic behavior, but in order to benefit

[322] *See supra* Chapter I, section II.B.

[323] E.g. Williamson (1973 and 1975). See also the conclusions of the previous chapter.

[324] *See* Eisner (2010: 513) ('Law and public policy are as foundational in a deregulated setting as they were under regulation, even if the effects are different').

[325] *See*, e.g., the important work of Benkler (2010) who identified a set of thirteen considerations ('design levers') for cooperation. According to Benkler, these factors may be intrinsic (such as solidarity, fairness, trust or efficacy), extrinsic (such as punishment and rewards, transparency and reputation, cost, crowding out, exit and entry conditions, leadership and asymmetric contribution) or mixed (norms).

[326] Similarly, the approach departs from most well-known economic literature which aims to uncover behavioural shortcomings.

[327] See *supra* note 178 and accompanying text.

from working together'.[328] The same applies to regulation. The problem of opportunism is architecture-specific and therefore only secondary for regulatory design.

B. A horizontal approach to regulation

The second consequence of the enhanced approach to regulation proposed in the previous sub-section is that the regulatory relation (i.e., the specific relation between regulators and firms) can be considered now primarily from a horizontal perspective. Up to now, most frameworks of analysis assumed at least some level of 'verticality' in the relation between the State/regulator and firms.[329] P-A models are explicit on this point, but the same assumption underlies the NIE approach and socio-legal accounts. Verticality is only dubious in the regulatory contract tradition, since the presence of the government is for the most part taken for granted. Although there is some governmental oversight, the State has no explicit objective function and acts more like an arbiter between the interested parties.

The horizontal approach implies that regulation is, in some sense, the result of some sort of 'bargain' between the regulator and firms. However, it is necessary to emphasise an important caveat. The allusion to bargaining helps mostly to shed light on the strategic positions of the parties within the regulatory relation. On the one hand, in some cases the special bargaining power of the firm places it in some equal position with the regulator. Such a bargaining power usually arises from the specific facts of a case and/or the characteristics of the regulated firm. These factual-specific circumstances are of relatively common occurrence and may be easily dealt with by the regulator. On the other hand, there may also be industry-specific situations where the regulator, in practice, is incapable of taking enforcement action against the firm. This occurs when the firm does not comply with its obligations and the outcomes have possible negative implications for the whole industry or even the wide economy – i.e., they become systemic.[330] Moreover, a number of actions taken

[328] Aoki (2010: 23) (within the context of intra-firm cognition).

[329] The assumption of verticality has its roots in the rationality of vertically integrated utilities common to pre-privatisation / 'deregulation' era.

[330] Industry-specific circumstances are more likely to be produced in industries where each firm plays an important role in the system.

by the regulator may even increase the negative effects. Both situations increase the bargaining power of the firm and improve its relative position in the regulatory relation. They give a key player some 'advantages' when confronting the regulator – advantages that may potentially be exploited in some specific circumstances. Conversely, they diminish the bargaining power of the regulator and reduce its relative position in the regulatory relation.

Beyond highlighting the strategic positions of the stakeholders, a horizontal approach to the regulatory relation, and the reference to bargaining between the regulator and the firms do not have further consequences. Particularly, the horizontal approach does not equate to an egalitarian bargaining between the regulator and firms – as implied, for instance, in some 'negotiated' approaches to regulation.[331] There are three interrelated reasons. First, in these approaches negotiations often revolve around single issues (e.g., price)[332] and predominate in one-time situations. Many important areas of regulation are left outside and must be dealt with in the traditional way – commonly, through command and control or incentive-based regulation. As a result, negotiation is at best a poor and incomplete account of the regulatory phenomena – it is more a regulatory tool and less a valid approach. Secondly, negotiated approaches normally disregard the importance of hierarchy in regulation. Command and control strategies are an important and necessary complement of negotiated settlements, particularly when the State has a diminished position in the bargaining relation. Finally, if regulation were the outcome of negotiation between two parties, regulatory law would lose much of its explanatory and practical role. In the extreme, courts would deviate from their role to dedicate their efforts to 'complete' the terms of the bargaining and decide outcomes – a task for which they are not prepared.

This is not to say that negotiation does not have any importance in regulation.[333] In fact, it has been argued that early post-privatisation regulation in the UK telecommunication sector

[331] E.g. Doucet & Littlechild (2009) and (2006); and Littlechild (2009).

[332] E.g. Scarpa (1994) (formalising the bargaining game in the specific price-regulation situation).

[333] Old regulatory literature stressed the relevance of negotiations in the determination of regulatory outcomes. E.g. Schmalensee (1979); Kahn (1988) (stating the regulatory relation resembles 'a bilateral monopoly'); Beesley & Littlechild (1989: 58-83) (stating that there is a greater scope for bargaining in price-cap regulation than in CoS regulation). Also Breyer (1982: 177-81), briefly analyses the bargaining situation between the regulator and the industry, and seems to consider the option economically unsatisfactory.

followed this approach.[334] However, the bargaining falls short to provide with an explanation for regulation and its consequences should not be over-emphasised. Likewise, a horizontal approach does not equate to a 'political bargaining' advocated by some authors.[335] Such an approach to regulation merely extends the regulatory contract approach and provides no guidance as to how stakeholders should engage in the regulatory process. Foremost, the horizontal approach stresses the need to go beyond vertical (hierarchical) approaches to regulation.

CHAPTER CONCLUSION

Unlike what JEVONS (quoted at the beginning of this chapter) believed, this chapter has argued that institutional analysis may provide some guidance on how and when the State should perform certain economic activities. Departing from the idea that at least some level of governmental activity is always necessary, it is still possible to specify 'the functions which are either inseparable from the idea of government, or are exercised habitually and without objection by all governments; as distinguished from those respecting which it has been considered questionable whether governments should exercise them or not'.[372] Previous frameworks of analysis have advanced on this task. However, their main shortcomings are the relatively poor efforts to fully understanding the diverse nature of economic regulatory goals and the excessive preference for the market as distinctive regulatory structure. The reason lies in the consideration of competition as the main objective to be pursued by regulators. There is a need to recognise the plurality of regulatory objectives and, consequentially, the multiple structures that may serve to attain any of them. A more comprehensive framework of analysis that studies economic regulation should depart from these facts.

Different market structures and various regulatory strategies can be employed. This approach put the regulator in a more horizontal position with respect to the firm, because it recognises

[334] Vickers & Yarrow (1988).
[335] Rossi (2005).
[372] Mill (2004: 728).

that hierarchy is not always the best alternative. Also, the approach calls for a more enhanced role of courts – particularly specialised courts – because they are in a better position to assess whether the right mode of governance has been chosen and whether such governance mode corresponds to the objective to be accomplished. To conduct this assessment, courts (and regulators alike) need to check the consistency of the chosen governance mode, considering the possible risks and unintended consequences that might arise.

CHAPTER III

THE REGULATION OF THE CAPITAL STRUCTURE OF UTILITIES FIRMS

> 'The greater part of the public works may easily be so managed, so as to afford a particular revenue sufficient for defraying their own expense, without bringing any burden upon the general revenue of society'.[373]
>
> - Adam Smith

OVERVIEW

Firms respond to the incentives of the regulatory regime in different ways. One such response is financial and consists of changes in their capital structure – i.e. the proportion of debt, equity or hybrid securities the firm uses to finance its assets. The understanding of the interactions between capital structure, the cost of capital and economic incentives within regulated environments is vital to critically assessing how the system has performed and providing avenues for improvements. Changes in the capital structures of regulated firms offer a suitable instrument to make these assessments.

Section I begins with a description of the late evolution of the capital structure of utilities firms in the UK since privatisation. The description draws on and specifies the general account of liberalisation and regulation made in Chapter I. The account is centred on the electricity and water sectors, but also applies to others expressly mentioned. Then, the second part of section I expounds upon some of the explanations that result from that relaxation. The theories are divided into two groups. First, there are general financial explanations. Given their relation with the regulatory system, they can be considered as 'exogenous' approaches: changes in gearing levels are associated with some action not necessarily correlated with the

[373] Smith (1991 [1776]: V.1.III.1).

regulatory framework. None of these theories, however, provides a complete explanation for what has been witnessed in UK utilities. Moreover, they are particularly incapable of providing a correct explanation of the changes of capital structures in regulated markets. A second group of theories incorporates features of the regulatory regime and views the changes as fundamental – i.e. associated with some 'endogenous' problem of the regulatory regime. In the presence of regulation, firms' decisions on capital structure are distorted, and therefore the regulatory regime in itself may affect the firm's choice of gearing. Features of the regime are hence a necessary complement of general financial explanations for the trend towards debt.

Section II critically assesses the main regulatory responses to the changes in capital structure, their underlying reasoning and the criticisms that have arisen. Mostly, regulators have responded by exerting some degree of control of the capital structure of the regulated firms. As explained below, control refers not only to traditional command and control (C&C) regulation, but also to softer forms of direction. This control is justified by the effect on both the cost of capital and the financial strength of the firms. However, it is shown that some concerns are overstated and provide no real reason for the control. Only a soft control may be justified to capture the benefits of low risk for consumers.

I. THE EVOLVING CAPITAL STRUCTURE OF UTILITIES FIRMS

A. The increase in the level of gearing

As in other parts of the world, liberalisation and regulation of utilities in the UK were part of a major privatisation process that led to a profound change in the ownership-pattern of the former state-owned monopolies and placed them in private hands. Although the process had no comprehensive list of goals, one of its principal underlying aims was the encouragement of wider employee share ownership as a means to alter the balance of political power: more

shareholders would offset the power of trade unions.[374] Political, rather than economic concerns were the main (but not exclusive) driver of the process.[375] To expand rapid share ownership, the government promoted an 'equity financial' model of the firms with concrete actions. Shares were deliberately underpriced, and bonuses were given as an incentive to small shareholders.[376] In addition, companies were privatised with very low levels of debt.[377] Examples abound. In the water sector virtually all the existing debt was written-off, so firms' balance sheets would appear un-geared before being sold to investors.[378] Likewise, gearing levels of electricity distribution companies were brought down to around 25%. Also, the allowed rate of return was generally high and investment programmes involving large sums were readily approved in order to facilitate the process and improve the level of services. Overall the policy was deemed successful.[379] Equity finance flourished and as a result a wide base of shareholders in the formerly state-owned utilities was created.

While it is undeniably true that political considerations played their part, it is also clear that the equity model was chosen (in those cases where it was) in order to give effect to the policy objective of transferring the firms to private ownership and control – that is to say, in order: to expose them fully to the managerial disciplines typically found in the private sector which were expected to drive efficiency gains; to enable them to access private sources of capital

[374] See Vickers & Yarrow (1988); Veljanovski, 1987). In addition, it would lead to a bigger base of supporters of capitalism sensitive to the then conservative government. Overall, the model aimed to change 'people's attitudes to business and trade' and create a 'more entrepreneurial society' (Veljanovski, 1987: 9).

[375] Indeed, during the first phase of privatisation (1979-84) there was an economic objective: the 'pursuit of quick cash' to improve the public finances (Parker, 2009: 88). From 1981 onwards, privatisation became more a matter of principle and 'entrenched as a central component of the Government's economic strategy' (*Ibid*, at 188).

[376] There were even cash injections. The so-called 'green dowry' (equivalent to £1 billion cash) was given to water and sewerage (but not water only) companies to fulfil the new enhanced environmental requirements promulgated by the European Commission.

[377] The only exception amongst utilities was the National Air Traffic Services. It was sold to a consortium that financed the purchase with 94% of debt (with a ratio of debt to regulatory capital value of 118%—see *infra* note 378 for an explanation of this concept).

[378] Although gearing is defined as 'debt to debt plus equity' (*D:D+E*), UK regulators commonly use the ratio 'debt to Regulatory Asset Base' (*debt:RAB*) as primary measure. The RAB is a regulatory concept, not a financial one. Generally speaking, it is a proxy value of the firm's regulated operating assets, upon which investors earn a return. It is formed by investors' stake at privatisation plus completed capital expenditures not paid out by customers. In monetary terms, it is also known as Regulatory Asset Value (RAV) or Regulatory Capital Value (RCV). The ratio 'debt to RAB' is known as 'regulatory gearing', and in general gearing measured on this basis is slightly lower than that using 'debt to debt plus equity'.

[379] Underpriced shares resulted in over subscriptions and immediate profits for shareholders. For that reason, as Vickers & Yarrow (1988: 160) point out, 'Politically, privatization was a winner, at least in the short term'.

and relieve them of public expenditure constraints; and to enable effective competition to be introduced into contestable markets. Wider share ownership was a relevant consideration in relation to those companies judged 'suitable' for retail investors (especially the public service utilities).[380] At that time, publicly listed companies in general exhibited rather low levels of gearing. Moreover, it was recognised that the introduction of competition, although progressive, would increase business risk; that the substantial capital investment expected to be needed to raise capacity and service quality in line with expected future demand (and, especially, to meet environmental quality obligations) implied large incremental financing requirements, and that, in the absence of any track record, regulatory and political risk would be perceived by investors to be high. All these factors led Government's advisers to advocate that utilities be offered for sale with low (in some cases, negative) gearing, notwithstanding that this was likely to reduce sale proceeds. In hindsight, it is clear that this approach was unduly cautious and that the consumer faced higher charges in the earlier years following privatisation than were, strictly speaking, necessary.[381]

During the mid-nineties the trend reversed. Infrastructure companies introduced major changes in their capital structures and there was a general trend to replace equity with debt.[382] With variations, the trend is still witnessed across all sectors.[383] For instance, since the last price control review, in 2004, the yearly overall industry weighted average gearing of electricity distribution companies has been between 45-50%. Recently, gearing levels have been in the range of 60% as an industry average in the gas distribution sector.[384] The situation

[380] But not all privatised entities fell in this category. *E.g.* Railtrack was initially thought unsalable: when eventually privatised, it was by way of a primarily institutional offering, and a number of other businesses were privatised by way of private sale.

[381] Some consider this as a symptom of an information deficit rather than a consequence of political choices.

[382] The period from mid-90s has been labelled the 'dash for debt', and seems to have started in electricity distribution (*see* Helm, 2003b: chapters 11 & 12). The trend coincides with a worldwide increase in financial gearing across privatised utilities in many developing countries (Correia da Silva *et al.*, 2004) and Europe (Bortolloti *et al.*, 2007). *See also* Alexander & Chia (2003), on evidence of the increased role of international bond-financing in utility and infrastructure companies.

[383] *E.g.* Wright *et al.* (2006) show ratios of debt to the market value of capital of nine utilities for the period 1995-2005. In most cases, but not all, there was considerable rise in gearing over the period. Indeed, there are variations. For instance, the optimal gearing ratio is likely to be higher in the water sector than in telecommunications – i.e. whilst the latter shows rapid technology change, changing market definition and -so far- growing demand, the former has both stable technology and demand. In fact, the trend towards debt has affected mainly the electricity, water (including both water and water and sewerage firms), gas and railways sectors.

[384] *See* Ofgem (2009a)

with regard to water is no different, with gearing levels rising sharply above 60%. Also, even a firm 100% reliant on debt (Network Rail) currently manages the rail infrastructure.[385] By sectors, overall, only airports still present relatively low levels of gearing. At the same time, however, there are considerable variations amongst firms within an industry. For example, whilst the most geared electricity distribution firm in 2008 had 61% of gearing (decreasing from 75% in 2005-06 and 68% in 2006-07), during the same year the least geared firm had only 17% (11% in 2006 and 23% in 2007). Likewise, in the water sector a study made by consultants identified four different corporate models, ranging from 100% debt to thin equity models and companies without 'structured finance' (Smith & Hannan, 2003). Notwithstanding the disparities, the tendency has been noticeable and remains fairly steady: up to now, UK utilities firms still present relatively highly geared capital structures.

B. What motivates the increase in gearing levels?

The explanations for the increasing levels of debt amongst utilities are not entirely clear, but this is hardly an issue exclusive to utilities. Since long ago the financial literature has tried to explain the so-called 'capital structure puzzle' (Myers, 1984).[386] The most traditional financial approach, due to Modigliani & Miller (1958) (the 'M-M model'), demonstrated that in perfect capital markets changes in the capital structure do not have any impact on the cost of capital – i.e. the value of the firm and its capital structure are independent.[387] Any restructuring which aims to substitute (cheaper) debt for (more expensive) equity is ineffective. The only consequence of such a policy would be to increase the equity risk premium demanded by equity investors. Any increase in the gearing would be matched by an

[385] Currently Network Rail's debts are explicitly underwritten and funded by the government. The government provides a financial guarantee through the so-called Financial Indemnity Mechanism (FIM). According to the regulator, the use of the FIM will be incrementally restricted and it is expected that by 2014 Network Rail raises its debt on an unsupported basis (ORR, 2008).

[386] The traditional departure point was normally an equity-financed firm, whereas a special rationale was sought for debt. As it will be seen, modern financial theories use a rather different approach.

[387] This is the well known 'Proposition I' of Modigliani & Miller (1958): capital structure does not affect the company's market value. In their words: 'the average cost of capital to any firm is completely independent of its capital structure and is equal to the capitalization rate of a pure equity stream of its class' (*ibid*: 268-9). This means that the value of the whole 'corporate pie' does not depend on how it is sliced (Brealey *et al*., 2008: 476; Tirole, 2006: 77-8; Myers, 2003: 219-20).

increase in the expected equity return.[388] In addition, this should be reflected in the risk faced by the firm. Holding constant all other factors, higher financial leverage increases the firm's risk profile.[389] Thus, any benefit from the greater use of (cheaper) debt should be offset by the change in the risk. Hence it can be said that, other things equal, higher gearing increases the variability of a firm's income.

The M-M model relies upon stringent assumptions. No taxation, no costs associated with financial distress, symmetric information (and the corresponding absence of agency costs), and perfect capital markets are all assumed.[390] In the model there is no 'optimal' capital structure. Nonetheless, the literature remarks that if debt policy were irrelevant equity-debt ratios would vary randomly from industry to industry and from firm to firm. That is not the case. In practice the capital structure does matter when capital market imperfections are taken into account. One example is Modigliani & Miller (1963), who added corporate income tax to their original proposition. This created a benefit for debt, because it serves to shield earnings from taxes. However, since the firm's objective function is linear and there are no offsetting costs of debt, the model would imply 100% debt financing. This is an extreme prediction that has two problems associated with it. First, there are just a few firms financed entirely with debt. Secondly, the model does not explain the presence of debt even before the introduction of corporate income taxes. Despite these criticisms, however, the point is that the assumptions of the traditional approach must be relaxed in order to understand the patterns of variations of the capital structure in the markets.

---

[388] This is equivalent to saying that the firm's cost of equity capital is an increasing, linear function of its debt-to-equity ratio. The increase in the value of equity is the 'Proposition II' of Modigliani & Miller (1958).

[389] UK regulators usually assess the cost of equity using the Capital Asset Pricing Model (CAPM) (in its standard version). The model assumes that the cost of equity for a given firm is equal to the expected return on investing in a single share in that firm. Appendix 1 to this chapter explains the technical details. For a more detailed analysis of the CAPM elements, *see* the bibliography cited *infra* in note 445. For alternative methods to calculate the return on equity in regulation, *see* Breyer (1982: 43-7). The most relevant element of the CAPM for this paper is the firm's equity beta, explained in Appendix 2. As stated in the text, within the CAPM an increase in the expected equity return means that β^e_i should increase linearly. However, one caveat is necessary. Changes in β^e_i can be attributed either to a change in leverage or a change in the firm's underlying asset beta (β^a_i). β^a_i reflects only the business risk in the market(s) where the company operates. A firm with no debt faces no financial risk, so $\beta^e_i = \beta^a_i$. Therefore, the proposition in the text holds only if β^a_i is held constant. Despite their differences, the overall pattern of both β^e_i and β^a_i over time should be similar. On the relation between β^e and β^a, *see* Appendix 2.

[390] 'Perfect' in this model means that capital markets are competitive, frictionless and complete (Myers, 2003: note 1).

1. Financial theories

Within financial economics, there are four leading theories of capital structure. The first one is the capital-structure irrelevance of Modigliani & Miller, already described. The remaining are: the trade-off theory, the pecking-order theory and the agency theory. Although these theories overlap (Myers, 2003) they are separately treated in this section, except for the agency theory. The latter is mainly focused upon the internal relations within the firm (i.e., the relations between managers and shareholders) and hence it does not appear in principle to apply to the external relation between regulators and the utilities, which is the focus of this thesis.[391] In addition, the transaction-cost economics theory has also come up with an explanation for the capital structure.

A first approach to the capital structure is known as the 'trade-off' theory.[392] It is an approach essentially 'strategic', meaning the firm determines its equity-debt ratio for reasons of mere convenience. The firm takes a decision balancing (i.e. trading off) the potential advantages of increasing gearing – primarily the tax shield – against the potential costs of financial distress. Debt finance would be more efficient for both the firm and shareholders than equity because corporate debt interest payments are (normally) tax deductible.[393] Thus, firms can reduce their tax liabilities by additional borrowing. Additional advantages, such as some regulatory measures, may also lead to increases in the level of debt. For the trade-off theory, the decision-making process is firm-specific. Debt-equity ratios vary from firm to firm and from industry to industry, with each firm picking a target (optimal) capital structure that maximises its own value.[394] Accordingly, in principle there is no security of 'last resort' suitable to all

[391] Agency theories have two main strands. The first one is the 'agency cost model' of gearing associated with the work of Jensen & Meckling (1976). It focuses upon the control of the managers by the owners in conditions of asymmetric information. Thus, debt limits the cash flow available to managers and therefore reduces their discretion. The second one is the 'control rights' model mainly associated with Hart (1995). It focuses upon the issuance of debt by the owner as a means of not ceding control rights to outside investors. Despite not being treated in the text here, these theories are highly relevant in other parts of this work (see *infra*, section IV.C in this chapter).

[392] The trade-off theory is in fact a family of related theories that evaluates the various costs and benefits of different gearing levels. The first formulation can be attributed to Kraus & Litzenberger (1973), who studied the balance between the dead-weight costs of bankruptcy and the tax saving benefits of debt.

[393] On tax deductibility of debt interest payments, see Graham (2000) and MacKie-Mason (1990).

[394] The targeted capital structure is discussed in the literature. Myers (1984) considers it as a key implication of the trade-off theory, and argues that deviations from the target are gradually eliminated. By

firms. The variations are explained in relation to assets – whilst firms with safe, tangible assets and enough taxable income to shield will tend to rely heavily upon debt, unprofitable firms with risky, intangible assets will prefer equity finance. However, as a rule, the model points to a positive relationship between gearing and profitability, given that the benefits of debt would tend to increase and the cost of financial distress decrease as profits rise.

The 'pecking-order' theory states that gearing changes in response to the financial need of the firms (their investment requirements) and the availability of internal resources.[395] Accordingly, firms would have a preference for -first- retained earnings over external finance and -secondly- if external finance is required, debt over hybrid securities and the latter over equity. The reason lies in the asymmetries of information that exist between lenders and borrowers. Anticipating that managers may exploit these asymmetries issuing new equity only when they deem the company's shares are overvalued, investors will discount any new and existing shares every time a new equity issue is announced.[396] To avoid the resulting risk, managers will seek to finance investments projects first internally and then externally. In sharp contrast with the trade-off theory, all firms adopt the same pecking order. This reflects a behavioural approach to finance: the use of retained earnings is favoured because they are the safest security. Unlike the trade-off theory, there is no optimal level of gearing. It does not advocate or target any debt-equity ratios *per se*. The debt-ratio only reflects the firm's cumulative requirements for external finance. The importance of financial slack is stressed. Therefore, equity is the financial instrument of last resort, to be issued only when the firm lacks the capacity to issue new debt, which in turn will normally occur only when there is a threat of financial distress. This suggests that there should be a negative relationship between gearing and profitability: with more profits the firm is capable of financing its projects with internal resources.[397]

---

contrast, Frank & Goyal (2007) hold that leverage exhibits target adjustment is neither necessary nor sufficient for a firm to be balancing tax savings against bankruptcy costs.

[395] *See* Myers & Majluf (1984) and Myers (1984).

[396] Note that an incentive alignment between managers and current stockholders is assumed.

[397] Consider also the case where levered firms finance new projects exclusively with retained earnings. In this case, more profitable firms will reduce their leverage (relative to less profitable firms) through retained earnings.

In the 'transaction-cost economics' (TCE) model the capital structure is seen as a pre-contractual problem. Debt and equity are seen as alternative modes of governance, whose selection critically depends upon the firm's assets – i.e. their redeployability outside the firm.[398] Thus, debt is the simplest governance structure, akin to the price system of the market, where agents are rewarded on the basis of output. In accordance, debt-holders will prefer low-specific investments – i.e. projects that present highly redeployable assets. By contrast, equity is a much more complex mode of governance that involves an intrusive involvement in the projects, has higher setup costs, and allows greater discretion. Accordingly, it is favoured as assets become highly specific. The TCE model arrives at the same conclusion as the pecking order theory: equity is the finance instrument of last resort. Nonetheless, it is based upon transactional rather than motivational assumptions. The option of debt or equity is seen as an economising problem: the 'optimal' capital structure will be chosen according to the transaction costs involved in the situation. As a consequence, every firm makes its choice based (as the trade-off theory) upon the characteristics of the assets. In contrast to the other two theories, the TCE model implies no specific relation with profitability: the firm chooses its capital structure regardless how large its profits are.

Table 4 summarises the main features of the three theories described in this section.

Table 4: *Financial theories on the capital structure.*

Theory	Purpose	Decision-making	Security of Last Resort	Relationship gearing-profitability
Trade off	Strategic	Firm-specific	None	Positive
Pecking order	Behavioural	Homogeneous	Equity	Negative
TCE	Economising	Firm-specific	Equity	No specific relation

2. Do financial theories provide a sufficient explanation?

Unfortunately, there is no clear-cut explanation for the trend towards debt. Within financial literature, it is far from undisputed which theory better elucidates the capital structure

[398] *See* Williamson (1988).

puzzle.[399] Numerous studies have been dedicated to this issue without reaching a final conclusion. However, some general observations can be advanced. First, amongst the three theories, arguably the TCE may be discarded as an explanation for the high debt/equity ratios witnessed in utilities. Its main downside is that its predictions move in the exact opposite direction of how reality has developed – utilities have non-redeployable assets, but despite this they have opted for highly geared capital structures.[400] Secondly, between the remaining two theories, the most recent empirical evidence tends to greater support the pecking order theory's predictions rather than the trade-off theory.[401]

Nonetheless, the trade-off theory – until recently the 'dominant theory' in corporate finance – seems to be preferred by 'official' sources to explain the trend towards debt amongst UK utilities. For instance, it underlies the report issued by the UK Government analysing the increased level of gearing in utilities (DTI & HM Treasury report, 2004). The report sees the trend as driven by four factors – taxation, risk redistribution, agency and informational effects and risk reduction – which combined would have helped firms to optimise their capital structure.[402] Also, the report identifies a number of sector-specific justifications. Amongst others, in the electricity and water sectors M&A activity (takeovers and leveraged buyouts) is mentioned as one of the main causes of the trend.[403] It is highly probable that upon this, the

[399] As Myers (2003: 217 & 244) expresses, 'There is no universal theory of capital structure, and no reason to expect one. […] [N]one of the theories gives a general explanation of the financing strategy. They are plausible as conditional theories, but we have only a partial understanding of the conditions under which each theory, or some combination of the theories, works'. Frank & Goyal (2007: 59), in their summary of the main features and evidence of the trade-off and the pecking order theories, find 'good reasons to question the standard versions of both theories'.

[400] Although the general support for this theory seems to lag behind the other two (pecking-order and trade-off), there are studies confirming the TCE approach. Amongst others, *see* Balakrishnan & Fox (1993), Bjuggren (1995), Kochhar (1996). None of these studies, however, refers to utilities.

[401] *E.g.*, Brierley & Bunn (2005); Fama & French (2002); Rajan & Zingales (1995). *See generally* Tirole (2006: 238): 'The pecking-order hypothesis has received substantial empirical support', although he recognises that '[a]s usual, things are more complicated than is suggested by this interesting hypothesis'. Certainly, the pecking order theory has also a number of downsides—foremost, that firms issue debt when they should not Fama & French, 2005). But the literature recognises that it is an adequate starting point. *See* Tirole (*Ibid.*: 238 *et seq.*) for an explanation of the downsides.

[402] *See also* Correia da Silva *et al.*(2003).

[403] For instance, in the electricity sector leveraged buy-outs followed the removal or expiration of the government 'golden shares' that the Government retained after privatisation, mainly during 1994-96. All the distribution companies were subject to at least one takeover by US companies (notably by Entergy and TXU) or multi-utilities. From 1998 onwards, US companies withdrew from the UK and the industry consolidated in the hands of few European energy companies. In the gas sector, in 2005 National Grid Gas sold four of the distribution companies to three new owners. As in the case of electricity distribution, this may explain the high levels of gearing of these companies. However, the remaining four networks are still operated by National Grid

report reflects the underlying thoughts on tax advantages derived from the acquisition of highly-leveraged companies.[404] If it is true that regulators are following the trade-off theory, there is a serious risk that they may be basing their policies on uncertain grounds.[405]

That risk increases when a second, more fundamental argument is considered. It is unclear that general financial theories are able to provide a complete explanation for the trend towards debt witnessed specifically in utilities markets. Indeed, in principle it may seem to be some grounds for the direct application of financial theories. There is evidence of a general sharp rise in corporate capital gearing in the UK economy since the late nineties until at least mid-2000 (Brierley & Bunn, 2005).[406] Favourable macroeconomic conditions may have exerted some pressure to gearing up – fundamentally easy credit for borrowing long-term debt and the monetary policy. The trend towards debt may simply indicate that utilities responded to the same general incentives that affect any other company in the market. This assertion seems to be corroborated by evidence of high levels of debt present in many capital-intensive industries, meaning that even if they are not affected by general macroeconomic trends, at least utilities are responding to the same incentives that affect industries with similar characteristics.[407]

Gas and present similar high levels of gearing. The M&A activity in the energy sector is described by HELM (2003b: chapter 12), who refers to the 'takeover mania'. Generally, the M&A activity allowed the creation of various multi-product utilities during those years. It raised concerns from the Competition Commission (CC) in 1996, when one water company, South West Water, was subjected to two competitive bids from Severn Trent and Wessex Water. The CC rejected the mergers—the first proposed takeovers of a water company by another—on the grounds that it would reduce the possibilities of the regulator to carry out yardstick competition. The CC adopted the view that 'no remedy, even in the shape of very significant price reductions aimed at forcing the merged enterprise beyond the current efficiency frontier, would be sufficient to compensate for the loss of [South West Water services] as a comparator' (*Severn Trent Plc and South West Water Plc: A report on the proposed merger* [25 October 1996], Cm 3429, para. 1.13, p.4. See also *Wessex Water Plc and South West Water Plc: A report on the proposed merger* [25 October 1996], Cm 3430). The M&A activity in the water sector has resulted in only 21 companies (11 water only firms and 10 water and sewerage firms) remaining from the original 39 statutory undertakings.

[404] Both the energy and water regulators have also applied the trade-off theory whilst choosing their policies. The 2004 price control reviews aimed to remove the tax advantages for highly geared companies and increase the returns to equity finance to affect the companies' gearing levels (Ofgem & Ofwat, 2006: 29 *et seq.*). Additionally, there have been attempts to claw-back tax benefits allegedly due to excess gearing (Ofgem 2004c: 111, para. 8.66 and 151, para. 11.53; and 2008a). On the tax advantages, see *infra* section III(C)(1).

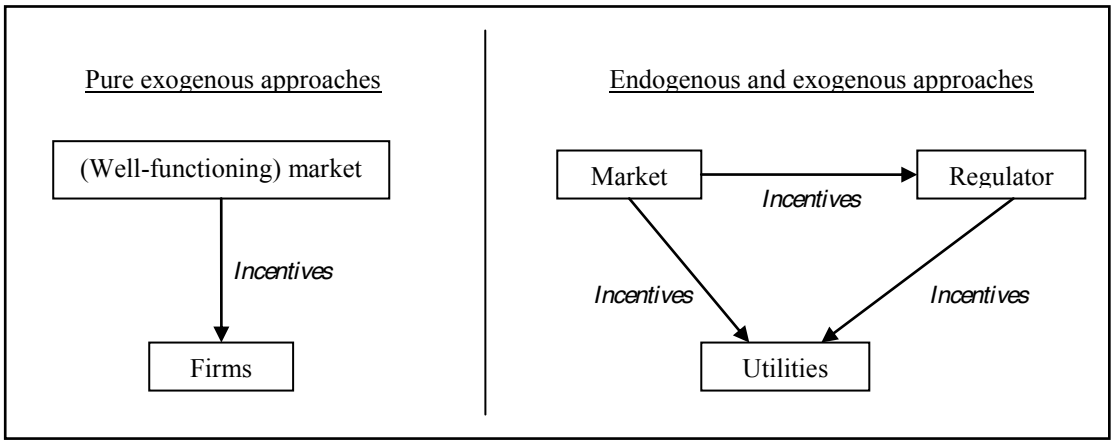
[405] In addition, the 'Achilles heel' of the trade-off theory is that tax benefits of debt are normally offset by the variety of costs that are due to capital market imperfections (Miller, 1989).

[406] Confirming that this is not a phenomenon exclusive to public utilities, the FSA issued a Discussion Paper assessing the risk associated with the general sharp growth of the private equity market in the UK (*see* FSA, 2006).

[407] '[A]lmost all airlines, utilities, banks, and real estate development companies rely heavily on debt. And so do many firms in capital-intensive industries like steel, aluminium, chemicals, petroleum, and mining'

But despite the apparent suitability of financial theories, a key finding of a number of studies is that capital structure decisions are not independent of the regulatory environment in which they operate.[408] Financial theories have been designed mainly to assess how *well-functioning* (but not necessarily perfect) capital markets respond to taxes, costs of financial distress and costs of bankruptcy. Hence particularities of the specific regulatory regime and the special characteristics of utilities markets are put aside. Nonetheless, in the presence of regulation the incentive structure becomes more complex (see Figure 5). Regulators, as intermediaries between owners and managers, severely affect property rights – crucially because they have the power to control revenues. On the one hand, shareholders see their residual returns modified.[409] On the other, under regulation managers are constrained in their choice of strategy (that is, managerial discretion decreases) and the cost of observing managerial performance is reduced.[410] By sharp contrast, financial theories presume managerial objectives as a given. However, the greater the regulatory intervention in the market, the stronger the constraints over management become.

Figure 6: *Incentives to gear up*.



(Brealey *et al.*, 2008: 496). From this perspective, there are some concerns related to general financial stability implications, because it arguably illustrates that companies may not be able to easily adjust their debt levels in response to shocks.

[408] *See e.g.* the studies cited *infra*, note 445.

[409] Jansson (2008) (stating that regulators control the firms' revenues).

[410] Kole & Lehn (1997) (*a contrario*, arguing that deregulation induces instability within the business environment that makes it harder to distinguish the effects of management decisions on firm performance from the effects of other factors).

To a great extent, regulation may help to alleviate agency problems.[411] Therefore, in regulated markets, regulators' behaviour and assumptions, the relation between regulators and firms, and the inner features of the regulatory regime may *all* be providing incentives to alter firms' capital structures.[412] All in all, then, in regulated markets there is a need to complement the explanations given by general financial theories before engaging in the design of public policies. Endogenous features of the regime should be incorporated in the analysis. They do not provide alternative but complementary explanations to the trend towards debt.[413] The next section explores in more detail how regulation may affect the decisions on the capital structure.

C. Incorporating endogenous features of the regulatory regime

Two main explanations for the trend towards debt have centred on the endogenous features of the regulatory regime. On the one hand, the increased gearing might be a means of influencing regulatory outcomes. Particularly, the trend to gear up might be used to curb opportunistic behaviour on the part of regulators. On the other hand, the increased gearing might be an unintended consequence arising from the weaknesses of the RPI-X system. Especially, there could be a mismatch between the time necessary to finance large capital expenditure projects and the five-year period that elapses between price settings. I turn now to the elaboration of these endogenous reasons for gearing up of regulated firms.

[411] Jiraporn & Gleason (2005) (founding a negative relationship between gearing and shareholder rights in unregulated industries, but not in regulated ones, and speculating that regulation mitigates the role of gearing in controlling agency costs).

[412] As Kole & Lehn (1997: 423) state: 'several governance mechanisms…are expected to change under deregulation, including ownership structure'.

[413] As will be noted, many of the flaws in the regulatory approach towards the high levels of debt are due to the regulators' reliance upon exogenous rather than endogenous theories.

1. Gearing as a 'shield' against regulators' opportunistic behaviour

The US long-established regulatory/financial literature, following the traditional approach to the capital structure, has implicitly assumed that the regulated industry was wholly financed by equity (SPIEGEL, 1994). That assumption was eventually relaxed, and analysts began to investigate how capital structure choice might be used to influence the decisions taken by regulators. New models showed how regulated firms use their capital structure as a means of influencing the regulated price. Firms *intentionally* opt for a higher debt-capital ratio because the regulator responds by setting a higher price to avoid the possibility of financial distress with negative effects on the services – i.e. the firms use debt financing as a 'shield' against opportunistic behaviour on the part of regulators.[414] In the European context, Bortolloti *et al.* (2007) have recently supported this argument. They find evidence showing that private European utilities increase their gearing after becoming regulated by independent regulatory agencies, and that behaviour has a positive and significant effect on regulated prices and investment levels.[415]

In the UK, Mayer (2005) has provided an analogous explanation for the trend towards debt observed in the water sector, which he sees as a part of the 'privatization-cycle'.[416] He explains that in the early stages of privatisation firms were less concerned about the level of commitment from the regulator. The need to boost investments protected the sunk costs to a large extent. Hence, gearing remained low. However, in later stages the focus of regulation changed: investment encouragement switched to higher efficiency requirements demanded by

[414] The assertion was supported with strong US evidence: Taggart (1981 & 1985) (introducing the 'price-influence effect': debt can provide incentives for regulators to choose higher prices in the output market); Dasgupta & Nanda (1993) (developing a model whereby firms use debt strategically to enhance their bargaining position); and Spiegel & Spulber (1994 & 1997) (developing a more general bargaining model whereby firms issue debt to increase the probability of bankruptcy, causing the regulator to raise the regulated price).

[415] In the UK, the CAA (2006) has applied this reasoning. The airports regulator issued a statement directed to a potential bidder for BAA, owner of three London airports, indicating that a highly geared takeover that would reduce the firm's credit rating would not induce an adjustment of prices.

[416] To some extent, DTI & HM Treasury report (2004: 13) also seems to partially support this explanation, when it states that the gearing up is a 'natural response to the fat balance sheets immediately after privatisation'. The roots of the explanation can be traced to the well-known 'life-cycle' theory that explained regulatory capture (*see* Bernstein, 1955).

regulators. In response, the capital structure of regulated firms changed, because debt provides greater 'rights to exit' from the regulatory contract.[417] On the one hand, shareholders may place the firm in the hands of creditors and receiver, if default threatens. On the other hand, the regulator may put the firm in special administration and/or revoke the licence. Accordingly, overall levels of gearing increased because of the reduction in the level of commitment, '[a]nd there was little that governments or regulators could do to avoid this' (*ibid*: 187).[418]

These 'shield-type' of explanations seem plausible in the context of privatised utilities. But to some extent they only present an 'enlarged' version of the trade-off theories, where the regulatory outcome is incorporated within the benefits the firm might obtain by rebalancing its capital structure. Yet there are insufficient studies empirically showing the validity of this approach – and most of the evidence is provided for the US case. In the US, however, regulators normally take the capital structure as given (De Fraja & Stones, 2004: 70). Hence, the task remains to determine whether this phenomenon exists in other jurisdictions. Also, 'cost of service' type of regulation was, and still is, largely dominant in many industries, particularly in the US. Therefore, the impact of specific features of PCR remains largely empirically unexplored.

2. Gearing as a response to the weaknesses of PCR

Another explanation for the trend towards debt may be found in the inbuilt incentives of PCR. It might be argued that, regardless of the financial structure of the firm, PCR may induce under-investment.[419] I will specify the reasons below.[420] But the option to under-invest

[417] On the regulatory contract, see *supra* note 198 and accompanying text.

[418] In Mayer's view, the tightening of the 1999 price control encouraged firms to adopt low-cost financing structures as a way to restrict potential regulatory adjustments further down.

[419] *See* Gómez-Ibánez (2003: 241-2), who provides reasons to the underinvestment claim and states that it is hard to say if this effect is better than the overinvestment effect; Cowan (2001: 52), who points out that 'whilst price-cap regulation seems to provide good incentives for operating cost efficiency, it is not clear that it provides optimal investment incentives for utilities'; and Newbery (1999: 51), stating that the advantages of price controls come at the expense of adequate capital investment. *See also* NAO (2002). Note that in sharp contrast with PCR, 'cost of service' ratemaking schemes provide incentives for overinvestment. This is the so-called 'Averch-Johnson effect' (Averch-Johnson, 1962). For empirical support, see e.g. Spann (1974).

[420] See *infra*, section IV.B in this chapter.

presents a problem for the firm: it must still comply with regulatory targets to generate revenues. That may be achieved by lowering financial costs. If debt is less costly than equity, increasing the proportion of debt would lower the total cost to the regulated firm, helping to achieve the target set by the regulator. In this sense, the regular cuts of costs that constitute the base of PCR may be providing incentives to alter the firms' capital structures in favour of debt. At the extreme, debt may totally replace equity within the capital structure.[421]

Incorporating financial terms, Helm has explained the trend towards debt as a result of an endogenous problem of PCR.[422] He argues that utilities have two radically different activities. On the one hand, there is the Regulatory Asset Base (RAB), against which the cost of capital is applied to determine the rate of return on assets.[423] On the other hand, they have an operational side, which is composed of capital expenditure (Capex) and operational expenditure (Opex). Whilst the RAB would be different in kind, Capex and Opex would only be different in degree. The RAB faces very low risk, because the rate of return is protected by the duty to finance the firm's activities.[424] In practice, this means the equity risk has almost been completely transferred to the customers. By contrast, with Capex and Opex firms face the equity risk in full. The reason is that the fixed-price period would not match the investment horizon of many projects, which are normally planned for periods of time longer than five years. As a consequence, investors would lack regulatory long-term commitments.[425]

The difference in risk and its interaction with the cost of capital would lead to gearing up of the capital structures. Helm argues that the method used by regulators does not provide

[421] Although this solution is not optimal (*see infra*, section IV.B in this chapter).

[422] His argument is long-standing and has been stated in several notes. A summary is presented in Helm (2009).

[423] PCR considers financial outcomes as a consequence of prices and not a determinant of them. By contrast, under a 'cost of service' scheme first the rate of return is set and then it is reflected in the price. On the concept of RAB, *see supra* note 378.

[424] This is a duty commonly replicated in statutes across sectors. For more details, see *infra* section IV in this chapter.

[425] Helm (2009) argues that unrewarded equity undermines the system because it produces a 'time-inconsistency' problem (i.e. a mismatch between long-term and short-term investment). Over-rewarded debt also undermines the system because customers pay a premium on the RAB. For more analysis on the commitment problem, *see infra*, section V of this chapter.

sufficient incentives for Opex and Capex.[426] These operating parts of the utilities' business require higher returns than the RAB. Since the cost of capital is calculated as an *average*, the regulator normally sets it in the middle range between value of the RAB (i.e. higher than the marginal cost of debt) and the more risky aspects of the business (i.e. lower than the marginal cost of equity). That is, the cost of capital is too high for the RAB and too low for the operating parts of the business. Hence investors do not earn full return on Capex and Opex, but earn returns in excess of the RAB. This creates incentives for the firms to gear-up, finance the RAB with debt, and securitise it in order to gain supra-normal profits. Therefore, arbitrage would be the rational response to a badly designed regulatory regime.[427]

* * * * * * *

Through any of these means, regulation (and the way it is applied) creates distortions that may undermine the regulatory aim to isolate firms, as much as possible, from the risk of financial distress. This is especially true under the PCR as it has been applied in the UK. As is shown in the next part, regulators have provided strong incentives to the firms to alter their capital structures. Unfortunately, the specific features of the regulatory regime have not always been taken into account in the analysis.

II. THE REGULATORS' RESPONSE

In theory, mimicking competitive markets would mean that regulated firms are free to opt for the ratio of debt and equity they consider better suits their needs. In principle, they should be able to pursue its goals as any other firm operating in an unregulated market, subject only to the public constraints of fulfilling their responsibilities for the maintenance of continuous quality of service and coverage (when it corresponds).[428] UK regulators, however, have

[426] For an explanation of the method to assess the cost of capital used by regulators (known as 'WACC'), *see infra* note 445 and accompanying text.

[427] As a solution, Helm proposes to 'split' the cost of capital. His diagnosis is compelling, insofar as it incorporates features of PCR within the analysis; his solution is not. I will not analyse here this specific solution, nor will I repeat the criticisms (*see e.g.* CC, 2008).

[428] This is the position, *e.g.*, of Jenkinson (2006) and Hillman and Braeutigam (1989: 91). Due to space considerations, the criticisms to this position are beyond the scope of this work.

generally applied some form of 'control' on the capital structure. They have reacted to the trend towards debt restricting in some form the freedom of the firms to set their own capital structures with leeway.[429] Although regulatory interventions present different degrees in each sector, the control affects the setting of the cost of capital and its different components, and has even given rise to concerns about the entire regulatory regime.[430]

A. Command and control and incentive-based approaches

The first form of control applies traditional command and control (C&C) regulatory strategies. There are two different approaches. In the 'strong' form, regulators may opt for an intrusive approach and compulsorily instruct the adoption of certain capital structure. With this, the firm is totally prevented from choosing the capital structure that it may deem optimal. By contrast, in the 'soft' form regulators establish minimum levels of equity or, conversely, prohibit gearing above a certain limit.[431] The firm is therefore not prevented from choosing the capital structure it deems preferable, but its freedom is circumscribed to the limits set by the regulator. This policy option is normally used in the banking sector, where systemic effects are highly probable and hence it is likely that the general public may end up bearing some or most of the losses of financial distress.[432] In utilities, the CAA (2010) has recently advocated this solution (despite recognising its problems of implementation). The mechanism would comprise two parts: (1) the one-year rolling average gearing level should be equal to or less than the gearing assumption in the cost of capital estimate used in the price cap determination; and (2) gearing at any time should be less than a predetermined level (set slightly higher than the targeted gearing assumption to provide with some flexibility for

[429] In theory the control may affect all the regulated firms in a given sector or apply only to some specific firm(s). An example of the latter would be to reduce the rate of return of high-leveraged firms to their lower cost of capital, whilst keeping the cost of capital of other firms unaffected. Generally, regulators in the UK have adopted a generalised control.

[430] *See e.g.* NAO (2002: 18): 'There are concerns that the standard form of price control regulation may no longer be appropriate for UK regulation'. Changes in utilities capital structures have even given rise to calls to reform the current regulatory approach (*e.g.*, amongst others, Barnard & Cooper, 2008; and Ofgem, 2009b). To some extent, there seems to be a general feeling that at least some British regulators have not completely understood or internalised the sort of 'paradigm shift' occurring in financing networks.

[431] Having a minimum threshold is an explicit transfer of risk to the private sector (see Erhardt & Irwin, 2004: 49-50).

[432] A common example is the imposition of minimum capital-adequacy ratios for banks.

trading purposes).[433]

The use of C&C mechanisms, however, is controversial. Arguably, C&C mechanisms seem more adequate to deal with risk of new infrastructure projects on an *ad-hoc* basis and not as a permanent regulatory regime. Aside from the familiar criticisms of inflexibility of rules and intrusion on managerial freedoms, limiting gearing is difficult because informational demands are severe and the applicable standard is likely to be highly discretionary and contentious – indeed, raising the level of exposure to judicial review. Also, the mechanisms would require close and constant monitoring of compliance. These criticisms apply to both the strong and the soft form of C&C mechanisms. Indeed, the stronger the level of intervention, the more accurate the criticisms are.

Alternative methods of control that use incentive-based techniques are often preferred – even in presence of something akin to systemic effects in some sectors.[434] For example, there is a promising but under-explored approach whereby the regulator in principle leaves the firm free to opt for the capital structure most convenient for its own interests, but include in the design of the regime some mechanism that tightens the price cap in the event that gearing increases above the level assumed in the price determination – e.g. a penalty that exceeds the amount of the benefit (tax benefit or other) arising from high gearing, or other less intrusive means. Subsequently the excesses would pass through to customers. Even though already considered by some regulators, this option has yet to be implemented in practice.

The most widespread incentive-based option, currently applied by most UK regulators, uses something akin to signalling.[435] Regulators implicitly indicate their preferences for a given capital structure providing incentives for companies to adopt it. This is commonly achieved using 'notional' rather than 'actual' levels of gearing on the calculation of the cost of capital. That is, regulators consider a level of gearing deemed 'appropriate' for the entire industry and

[433] The mechanism would be combined with a tax clawback.

[434] *See infra* note 511 and accompanying text.

[435] Before the last price controls, the CC (and its predecessor the MMC) used to set airports' cost of capital over the basis of actual gearing. However, it changed its position and decided to use notional levels of gearing for Heathrow and Gatwick (CC, 2007: Annex F, F7, para. 24) and Stansted (CC, 2008: 92, paras. 11.11 *et seq.*). The argument was linked to financeability (*see infra* section IV.A in this chapter). Regulators recognise that the reason behind the use of notional gearing is the increase in levels of debt by regulated firms. *E.g.* Ofgem (2009b). *See also* Helm (2009).

then use that figure as an input in the calculation of the firms' regulated returns. Underlying the method is an implicit assumption that certain capital structures are more efficient than others in a given regulated environment. However, firms are still adopting their own decisions on the capital structures at their own convenience.[436]

De Fraja & Stones (2004) indicate that in virtually all cases notional gearing has exceeded the actual level. However, it is difficult to make a definite assertion. Until fairly recently, not every regulator collected data on actual gearing. Ofgem, for instance, did not provide data before the 2004 electricity distribution price control. Also, gearing levels of gas distribution companies were published for the first time only in March 2009. In most sectors, the information is just recently being collected.[437] What is true is that levels of assumed gearing have been significantly increased over the time. Estimated gearing is now much more closely related to actual gearing – although it may be higher or lower depending on the sector-specific considerations taken into account during each price control. For instance, as shown in Table 5, in 2004-05 actual gearing was lower than notional gearing in the electricity sector, but it was higher in the water sector. Generally, however, it seems that regulators currently consider as efficient a level of debt of around 50-60%. This markedly contrasts with, for instance, the 12% estimated in the water first price control review in 1994.[438]

[436] Ofgem (2006a: 54, para. 8.17) has stated that '[Our] gearing assumption should not be interpreted as an endorsement of any particular capital structure. We believe that the companies and their financiers are best placed to decide on the most appropriate capital structure'.

[437] Since privatisation is a relatively recent process, information about firms is relatively scarce. Hence the use of real debt/equity ratios (i.e. measured from the firm's book values) seems useless, forcing regulators to still rely heavily on an average cost of capital under notional debt-equity ratios.

[438] The gearing data in Table 2 needs to be interpreted with care. For the period the table covers, most of the companies to which the data relate were part of wider groups, many of which had adopted holding-company based financing structures such that the levels of debt dependent on the assets and cash flows of the regulated businesses for continuing service were significantly higher than the borrowings of the legal entity housing the regulated business alone. Thus the data tend to understate the true levels of gearing.

Table 5: *Levels of gearing officially reported for some UK utilities' firms*.[439]

Gearing	Elect. Dist.		Gas Dist.		Water		London Airports	
Year	*Actual*	*Notional*	*Actual*	*Notional*	*Actual*	*Notional*	*Actual*	*Notional*
1994-95						12		---
1995-96		---				---		---
1996-97		---		---	28.3	---	25	30
1997-98		---		---	36.3	---	32.9	---
1998-99		---		---	42.7	---	30.3	---
1999-00		50		---	44.2	45-55	25.1	---
2000-01		---		---	49.7 [d]	---	24.6	---
2001-02		---		62.5 [c]	55.6	---		---
2002-03		---		---	59.4	---		25
2003-04		---		---	59.3	---		---
2004-05	50 [a]	57.5		---	61.2	55		---
2005-06	51 [b]	---		---	58.5	---		---
2006-07	45	---	61	---	62.2	---		---
2007-08	44	---	63	62.5	66	---		60 [e]
2008-09		---		---		---		50 [f]
2009-10		65				57.5		

Source: various regulatory reports.
[a] Incorrectly reported as 44% in OFGEM (2008b). [b] 50% in OFGEM (2007c); corrected in (2007b)
[c] In 2001 Transco still operated as a monopoly in gas transport and distribution.
[d] Reported as 48.3% in OFWAT (2001b); 49.7% in later reports.
[e] Heathrow and Gatwick airports. [f] Stansted airport.

B. The regulatory concerns

Why are regulators concerned about high gearing? The increase in gearing levels has two (potentially conflicting) effects (Bucks, 2005). The first is the impact on the cost of capital. Regulators make a decision about the cost of debt and equity that would be incurred by an efficiently financed regulated firm at the notional gearing level. Since the increase in gearing may decrease the cost of capital, the regulatory approach has the potential to 'minimise' the latter (when notional gearing is higher than actual gearing) and hence lower the price paid by consumers.[440] However – and this is the second effect – high gearing also introduces price

[439] In other sectors the percentages are similar. In 2000, the ORR assumed a gearing level of 50% for Railtrack, the company that used to manage the railways infrastructure. Railtrack's successor, Network Rail, is a company financed by debt and limited by guarantee. In 2001, Ofgem assumed 60% for the electricity transmission companies. However, that year Oftel (Ofcom's predecessor) only assumed 10-30% of gearing for mobile phone operators (the reasons may be connected to the characteristic of the industry – *see supra*, note 383).

[440] Needless to say, the aim of the regulatory approach must be focused upon consumers: if gearing is efficiency enhancing, those benefits should be captured for them. The CC has recently confirmed this point: 'It might be in consumers' interest that a positive adjustment to [the regulated price] be made, for example if this were required to avoid an increase to [the firm]'s cost of capital' (CC: *Sutton and East Surrey Water plc:*

volatility that might lead to a fragile financial position for the firms with (allegedly) negative general consequences for the regulatory regime. In addition, high gearing can drive the cost of capital upwards because of the increase in prices variability.[441] Therefore, preventing firms to increase debt over a certain threshold (or incentivising them not to do so) would help to *ex ante* prevent financial distress and thus avoid unplanned *ex post* cash-injections. The regulatory task is to trade-off the costs of price uncertainty against the benefit of price reductions associated with more debt.[442] As Stones (2007: 147) explains,

> [B]y manipulating consumer prices, the regulator can change the distribution of risks between consumers and shareholders and arrange for shareholders' returns to be positively correlated, uncorrelated, or even negatively correlated with the market return. The last possibility would provide shareholders with a form of insurance against market risk, and thus the cost of equity would be lower than the cost of debt.

III. CAPITAL STRUCTURE AND THE COST OF CAPITAL

The first effect linked to the trend towards debt is the impact of the capital structure on the cost of capital. In practice, price controls are set by a financial methodology. The methodology integrates future cash flows (economic modelling) – i.e. opex and capex forecasts); accumulated balances from the past; and balances at the end of the review period that are attributable to subsequent periods (accounting numbers) – i.e. the RAB. As illustrated

---

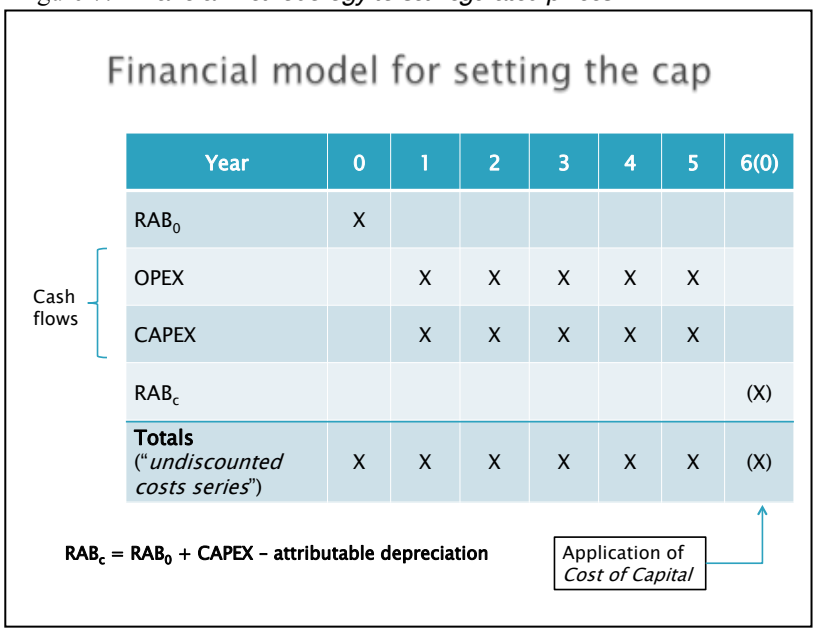
Interim Price Determination [17 June 2009], at 78, para. 4.94). The protection of the interests of the consumers is a central statutory duty established for all regulators. Its importance has been confirmed by case-law. *E.g.*, Competition Appeal Tribunal: *T-Mobile (UK) Limited, British Telecommunications Plc, Hutchison 3G UK Limited, Cable & Wireless UK & Ors, Vodafone Limited, Orange Personal Communications Services Limited v Office of Communications* [2008] CAT 12, at 98: "In any event, the lack of pass through is relevant only to the question of whether the proposed prices had an adverse effect on end-users…".

[441] Therefore, debt-only firms are not optimal, and the socially optimal capital structure always leaves some price uncertainty. *See* De Fraja & Stones (2004).

[442] There are different models to address the trade-off. For example, it has been proposed that if consumers are not price-risk averse, then prices could be subject to more variations. The outcome would depend upon the prevailing economic conditions. Specifically, prices might increase in presence of adverse economic conditions and negative shocks (Cowan, 2003; Stones, 2007). Therefore, both firms and consumers would absorb the price variability, with the attitudes of both groups to risk determining how the price should move (STONES (2007) shows that the less price variability should not necessarily lower the return for investors. A rather different approach is the model proposed by Ofgem (2010), whereby the price control is focused on outputs instead of inputs. The analysis of these proposals exceeds the scope of this work.

in Figure 6, the cost of capital is applied over this figure. Therefore, income is provided in part to finance current expenditure and in part to ensure the recovery of past investment.

Figure 7: *Financial methodology to set regulated prices*



The cost of capital has become a central component for the determination of the regulated price during periodic price controls even against expectations of those who originally introduced PCR into the UK. Under PCR the cost of capital should simply be one more element to calculate the cap. The expectation was that the X-factor would be mainly technologically driven, with the return for investors affected, but not determined, by the regulatory framework (Littlechild, 1983). Nonetheless, the importance of the rate of return was soon recognised by regulators.[443] Even so, in practice, regulators gave strong weight to the cost of capital. As Foster (1992: 213 *et seq.*) explained, 'it looked as if [the then director of Oftel] was undermining the spirit of RPI-X and reintroducing rate-of-return regulation by the back door'.[444] Even Littlechild (largely considered the architect of PCR in the UK) conceded later that the rate of return was implied in the calculation of the X-factor (Littlechild, 1986).

[443] First in telecommunications: *see e.g.* Oftel (1985, 1986a, 1987a, 1987b, 1988, 1991, 1992a and 1992b).
[444] *See also* Ofgas (1991) and Ofwat (1991).

Because the cost of capital is a central element of the regulated price, regulators have incentives to exert endogenous control over the capital structure – exactly as investors may do when they attempt to maximise the return on their assets (De Fraja & Stones, 2004). Generally, UK regulators assess the cost of capital using the so-called 'Weighted Average Cost of Capital' (WACC).[445] If certain conditions hold, the capital structure may directly affect the overall cost of capital computed through this method. For instance, when the assumed levels of gearing are lower than the actual gearing of some firms (and consequently the WACC is higher than their real cost of capital), the outcome of the price revision represents an advantage for them.[446]

Two main aspects influence the relation between the capital structure and the cost of capital. The first is the approach to the WACC adopted by the regulator in relation with taxes (pre-tax, post-tax or the so-called 'vanilla WACC'), which has implications for the tax advantages the firm might obtain from having more debt. When the WACC is calculated on a post-tax basis, there is less incentive to increase gearing because the advantage is already incorporated into the model.[447] The approach, however, largely depends on the preferences of the regulator and the characteristics of the industry, so I will not deal with this aspect in detail.[448] The second aspect is related to the incorporation of risk perception (which largely depends on the capital structure) within the calculation of the cost of capital.

---

[445] The WACC is calculated over the basis of the expected costs of the debt and equity. The result gives the average cost of capital for the regulated firm as a whole. Both the cost of debt and the cost of equity are computed over the basis of a number of individual components, some estimated by the regulator and others market-determined. Given the individual component estimates, the result is normally a range. Regulators then need to select a single point estimate from within that range, trading-off the benefits for consumers against the cost implications. A more technical explanation of the WACC method may be found in many standard textbooks of Corporate Finance (*e.g.* Brealey *et al.*, 2008; Copeland *et al.*, 2005) and in Wright *et al.* (2003).

[446] As the MMC stated: 'A number of regulated utilities have increased their level of gearing in the period since privatization and since debt appears relatively cheaper than equity in CAPM calculations this might be expected to have reduced their WACC' (MMC, 1997). On the CAPM, *see supra* note 389.

[447] *E.g.* Ofgem (2004c: 109, para. 8.59): one of the objectives to opt for a post-tax approach to the cost of capital was 'reduce the incentives to increase gearing'.

[448] In the UK, some regulators have traditionally followed the pre-tax approach. This is the case of the CC (*e.g.* 2008 & 2007) and its predecessor (MMC, 1993a, 1993b). The same is the case for Oftel and Ofcom, Postcomm and the CAA. By contrast, others have applied a post-tax approach or the vanilla version. Ofwat has consistently applied a post-tax approach to the cost of capital. Recently, both ORR and Ofgem have been using a post-tax approach to the cost of capital in its 'vanilla' version. Ofgem changed its position in the electricity distribution price control review for the period 2005-2010 and has maintained it for 2010-15 (*compare* Ofgem [2004c: para. 8.32; 2004b: para. 7.18; and 2004a] with Ofgem [1999b: para. 5.24]).

A. Cost of capital and Risk

Risk perception is reflected in the WACC through the so-called CAPM betas: asset, equity and debt betas.[449] Whilst equity betas include the combined effects of ***both*** business and financial risk, asset betas reflect only the former.[450] Traditionally, the relation between both types of betas has led to the widespread conclusion that higher gearing leads to higher equity betas, and that the increase will be greater the smaller the value of the corresponding debt beta. This should be followed by higher expected returns on equity and hence a higher cost of capital.[451] For this to occur, the perception of risk should be reflected in the firm's betas as accurately as possible. Yet that has not been the regulatory policy. In practice, regulators have sought to 'protect' firms from the risk of default (section 4 below). Hence they have not only looked to influence the financial behaviour of firms by exerting control over gearing levels, but also 'manipulate' equity betas. The tendency since the late 90s has been to set the equity betas coefficients close or equal to one (see Table 6), implying that the risk of the regulated firm is equal or very similar to an average firm in the market.[452]

Table 6: *Equity Betas used by regulators in the UK*.

Regulator	Company	Year	Equity Beta
Ofwat	Water companies	1994	0.67 – 0.75
		1999	0.7 – 0.8
		2004	1.0
		2009	0.9
ORR	Railtrack	2000	1.1 – 1.3 [a]
Ofgem	DPCR	1999	1.0
		2004	1.0
		2009	> 1.0
	GDPCR	2001	1.0 [b]
		2007	1.0
	TPCR	2001	1.0 [b]

[449] On this concept, *see* appendix and the bibliography cited *supra* in note 445.

[450] For an explanation of the concepts of business and financial risk, *see* chapter III, section I.B.

[451] This is the traditional logic of the CAPM, according to the M-M model (*see supra*, section I.B in this chapter). In jurisdictions such as the UK, where debt and equity are taxed differentially and there is a tax benefit from debt compared to equity funding, relaxation of the M-M conditions may result in the cost of capital actually falling as gearing rises, up to the point at which the incremental expected costs of failure offset the increased value of the tax shield. However, this is not to undermine the substantive point made in the text, relating to the seemingly consistent over-estimation (in relation to the gearing assumption) of equity beta coefficients made by UK regulators.

[452] There are no theoretical boundaries for the value of equity betas (*see* appendix for details).

		2006	1.0
MMC / CC	London Airports	1996	0.7 – 0.9
		2002	0.8 – 1.0
		2007	0.9 – 1.3 [c]
		2008	1.0 – 1.24

[a] ORR (2000) argued that Railtrack's betas reflected a 'premium' over other regulated industries due to its high level of operational gearing.
[b] In 2001 Transco operated as a monopoly in gas transport and distribution.
[c] The table expresses the minimum and maximum point-estimated for *both* airports included in the report. The range for Heathrow was 0.9 – 1.15; and for Gatwick 1.0 – 1.3. *See* CC (2007).

UK regulators sustain their estimation of equity beta at unity for different reasons. The main one is based upon inaccurate information. It is argued that there is risk inherent to large capital investment programmes, whose long-term horizon would mean that they are constantly subject to uncertainties that would not be captured by historical measures of risk. Thus, either the lack of reliable evidence or instabilities in the estimates over time is blamed for the greater weight given to unconditional expectations, which in turn leads to consider that utilities are of average risk. Additionally, certainty must also be taken into account. There are just a few publicly-quoted utilities on the London Stock Exchange with publicly available information.[453] Therefore, regulators should estimate a proxy of the risk faced by all of the firms based only upon the information available. This may be a difficult and demanding task, and the results are unlikely to be accurate. Finally, regulators may not want to change their estimations looking for a sustainable long-term approach. If the long-term estimation was that firms behave like an average firm in the market, it may be unreasonable to reduce the estimation of betas while the risk is increasing for other firms.

As a result of the regulatory policy, the assumption on gearing levels has exerted a major influence in the calculation of the cost of capital *because* the increased level of gearing has *not* been matched with the estimated betas. However, public utilities should not be 'assumed' to be equal to an average firm. Their risk should be modelled according to the specific characteristics of the industry.[454] Also, the different nature of the risks should be factored into

[453] For instance, this claim has been made by Ofgem (2008: 93, para. 5.7). Amongst the companies listed are Severn Trent, United Utilities, Northumbrian, BT, NG, Pennon and Scottish & Southern.
[454] For instance, electricity transmission should be a lower risk activity than distribution. Nonetheless, Ofgem (2006) has claimed that the evidence is not sufficiently robust to quantify differences between sub-sectors with accuracy.

the models.[455] Overall, the regulatory argument for equity betas coefficients equal to one is 'distinctly generous' (Wright *et al.*, 2006: 14).

First, the different characteristics of the regulatory regime must be taken into account. The evidence indicates that, in principle, betas should be higher under PCR than under a 'cost of service' (CoS) ratemaking scheme – i.e. investors bear the greatest non-diversifiable risk in PCR (Alexander *et al.*, 2000; 1999; 1996), whilst CoS regulation immunizes shareholders from shocks to long-term cash flows (Guthrie, 2006). However, the literature also points out that in practice CoS schemes and PCR tend to be similar (*e.g.* Kessides, 2004; Grout, 1995; Baumol, 1967). The key differences lie mostly in the level of cost pass-through and how fast changes are reflected in the regulated price (which in turn is normally a function of the length of time that elapses between price reviews).[456] In the US, under a CoS scheme, equity betas are very low.[457] Considering the practical similarities between both regulatory regimes, it is at least hard to justify the doubling of betas made by the UK regulators.

In addition, there is mounting evidence that betas should decrease. Old estimations made by Grout (1995) indicate that the level of betas should be in the bottom half of the distribution, ranging between 0 and 1, with the lowest betas appearing in the water sector. Likewise, Alexander *et al.* (1999) estimated average betas of 0.6 for electricity, 0.84 for gas and 0.67 for water. It is unlikely that these estimations have severely changed in recent times. Indeed, recent evidence by Wright *et al.* (2006) confirms that an estimation of 0.5 is a better measure

[455] Armstrong *et al.* (1994) have convincingly argued that the correct way to model risk is to see it as a factor that lowers expected future cash flows, thus reducing expected profitability without altering the cost of capital *per se*. They indicate that regulatory risk is firm-specific and should be diversifiable, but recognise that if regulators applied this model they would be recognising that there is a possibility of future expropriation. In this line, it can also be argued that risk should be confined only to political risk, not regulatory risk. If the cost of capital represents the *expected* rate of return, then regulatory risks should not have any effect on it.

[456] Under 'cost of service' regulation, investors only bear the consequences of shocks *until* the next regulatory revision of prices. By contrast, under PCR investors only bear the consequences of shocks *after* the regulatory control. Hence the frequency of the price reviews (or the possibility to pass on the costs to consumers) is crucial. More frequent revisions under a 'cost of service' scheme will lead to less risk and to a decrease in the cost of capital (Brennan & Schwartz, 1982). More frequent revisions or more pass-through (*e.g.*, via 'claw-backs' or 're-openers') under PCR will lead to more risk and an increase in the cost of capital (Evans & Guthrie, 2006).

[457] Old evidence by Grout (1995) and Alexander *et al.* (1999) showed average betas of 0.2 in the case of electricity, gas and water utilities, and 0.5 in telecoms. The values have probably slightly increased recently.

of equity betas for long-term utilities.[458] Similarly, Jenkinson has argued that equity betas above 0.4 are hard to justify (Jenkinson, 2006). Given their captive base of customers and regulated tariffs, the revenues of utilities hardly vary with the general state of the economy.[459] In addition, most of the risk associated with large capital programmes would be diversifiable. Jenkinson's estimation seems to be supported by recent calculation made by the ORR (2008) and the CC (2008).[460] The latter found that utilities' asset betas are low (0.37) compared to airports (between 0.50 and 0.61) and well below the risk an investor faces in the UK stock market (0.72).

Third, there is also strong evidence pointing in the opposite direction to the traditional relation between asset and equity betas assumed by regulators. That is, there would be a *negative* relationship between betas and gearing: in utilities, higher levels of gearing would be associated with *lower* equity betas and therefore lower asset betas.[461] The explanation may lie in the own regulators' behaviour: their control over the capital structure may be producing a decrease in the underlying asset betas.[462] That is, regulation itself may have reduced the volatility and uncertainty inherent to high levels of gearing.

[458] The estimation, however, presents a wide confidence interval. Recently Ofwat has recognised that equity betas in the water sector are very low compared with market and historical trends (Ofwat, 2007: 46).

[459] This is, indeed, a generalisation. In the case of those utilities subject to revenue cap regulation (e.g. energy networks), revenues do not vary with the level of demand (at least not in the short term). By contrast, in the case of those subject to price cap regulation (e.g. water/wastewater, telecoms), revenues do vary with the level of demand. In the recent economic downturn, revenues of some water supply companies (especially those with a high proportion of industrial demand) have been significantly reduced. However, some effects have been moderated *ex-post* by regulatory protections.

[460] Importantly, the ORR (2008: 230, para. 14.17) has recognised that there is 'strong evidence that Network Rails' risk profile is below that of the airports and is similar to the energy and water sectors. We are providing Network Rail with some very significant protections against risk, particularly related to its capital investment programme. It also faces very little volatility in revenues. The majority of its income is fixed for the five-year control period'.

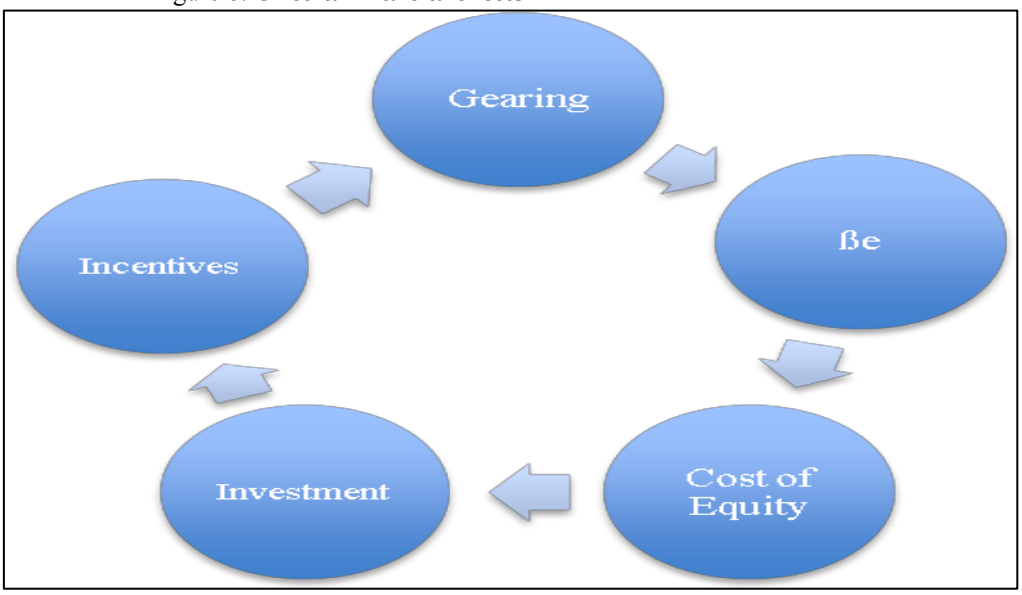
[461] The evidence is presented by Wright *et al.* (2006). Their evidence is compatible with that of Marston & Perry (1996) for US firms. By the same token, the CC has stated that it does not accept the argument that higher levels of gearing produce a higher cost of capital (CC, 2007: F23, para. 90).

[462] Similarly, Wright *et al.* (2006: 63): 'Paradoxically, the attempt by the regulator to limit the "dash for debt" may actually have sped up the process'. This fact provides an additional support for the stated view that, under regulation, changes in capital structures are influenced by regulators' behaviour: by itself the regulatory control has influenced the changes in the capital structures of regulated firms.

B. Cost of capital, gearing and the incentives of the regime

The relation between the capital structure and the cost of capital provides a good example of the need to consider endogenous features of the regulatory regime. Regulators cannot set the notional gearing value for the WACC calculations definitively until the incentives are calibrated, so that uncertainty of the price control has been reduced. The effect of other investment in the notional gearing (for example, changes in capex programmes) also needs to be taken into account. The cost of equity range will influence the way the incentives are calibrated. In turn, the range of potential returns under the incentives will influence how we set the notional gearing. Lastly, the notional gearing will feed back into the cost of equity estimates by affecting the equity beta. This is illustrated in Figure 7.

Figure 8: *Circular financial effects*



IV. CAPITAL STRUCTURE AND THE FINANCIAL POSITION OF THE FIRMS

Let's focus now on insolvency risk. Allegedly, the increase in gearing might lead to a fragile financial position of the firms, generally affecting the regulatory regime. This idea is commonly deployed through one or more of three concerns that I shall call the 'financeability concern', the 'under-investment concern' and the 'control concern'.

A. The 'financeability concern'

According to the traditional economic view on the duty to finance networks, which finds its roots in the classic theory of the natural monopoly, regulators must set the price so as to allow the monopoly simply to breakeven. If the price is set 'correctly', a firm that raises capital in the capital markets should not discriminate between long- and short-term when making investment decisions, provided that the projects generate revenues higher than the cost of capital in net present value (NPV) terms. In practice, this means that the firm needs to earn a return at least equal to the cost of capital on the RAB.[463]

In addition, most regulators now interpret the duty to finance as a duty to maintain the so-called 'financeability' of the firms.[464] This idea refers to the possibility that firm's revenues,

[463] Originally, it was thought the RPI-X did not require the measurement of capital and rate of returns (Littlechild, 1983). Nonetheless, it was later recognised that 'RPI-X and rate-of-return regulation have certain common features. Both accept the need to secure an adequate return for the company's shareholders in order to induce them to continue to finance the business, without conceding unnecessarily high prices at the expense of customers' (Beesley & Littlechild, 1989: 460). Currently, financial models used to determine the X-factor already incorporate a mechanism to ensure that new investment earns the cost of capital. Regulatory concerns for allocative efficiency are the ones that have led regulators to consider measurements of (and control) capital and rate of return (Armstrong *et al.*, 1994: 174). On the concept of RAB, *see supra* note 378.

[464] In this sense, *e.g.*, CC (2007: 73, para. 5.11): 'The CAA has no express statutory obligation to ensure that regulated airports, including BAA, are adequately financed. The CAA is, however, required to promote the efficient, economic and profitable operation of such airports, and also to encourage investment in new facilities at airports in a timely fashion. In this context, we consider appropriate for us to establish whether our proposals would enable Heathrow and Gatwick to finance their [businesses] on reasonable terms in the banking and capital markets through some combination of debt and equity'; Ofwat (2004: 217): 'We have a duty to secure that companies are able to finance the proper carrying out of their functions as licensed undertakers ('finance functions'). We look at this as having two strands. One is to secure that, if a company is efficiently managed and

profits and cash flows allow it to raise finance in the capital markets – on terms considered 'reasonable' by the regulator. Thus, the regulated price should provide the firm with sufficient returns to (1) keep certain credit ratings issued by credit rating agencies (with some 'headroom' in the rating to avoid an immediate slip in the event of a negative shock) and (2) meet certain 'financeability targets' that underlie the credit ratings and are linked to the firm's achievements (*e.g.*, CC, 2007). With this aim, most regulators now apply a 'financeability test' during price controls as follows: First, a decision is made regarding the appropriate credit rating that the firm should target (typically 'Investment Grade').[465] Moreover, benchmark provisions have been embedded into some utilities' licences compelling them to maintain minimum financial ratios. Secondly, it is decided which financial indicators and thresholds are going to be used. Financial indicators are mostly cash-based.[466] Finally, each company is modelled and the financial indicators are computed and analysed. Within the financeability test, it is decided what level of gearing would be consistent with the targets.

The financeability concern stems from the mismatch between the calculation of the cost of capital made by regulators and means of financing projects by the utility (Oxera: 2006 & 2010; Bucks, 2005). Whilst the regulatory model provides for a real rate of return, most firms pay out returns to creditors in nominal terms.[467] The result is that, due to the difference between real inflows and nominal outflows, firms may be left with negative cash flows for a certain period of time (possibly longer than one single price control period). If – and only if – a firm invests at a rate above the level required to replace its existing assets, this 'cash-flow gap' increases to an extent that may lead to a deterioration of the financial indicators.[468] In

---

financed, it is able to earn a return at least equal to the cost of capital. The second is that its revenues, profits and cash flows must allow it to raise finance on reasonable terms in the capital markets. We refer to this second strand as financeability'; and ORR (2008: 273, para. 17.1): 'We have a duty to act in a manner that it will not render it unduly difficult for Network Rail to finance its activities'.

[465] This means at least BBB-, Baa3 or any equivalent rating, depending on the issuer. However, at least the CC has increased the level, stating that in the current financial 'turbulence…an efficiently financed company might reasonably seek to target an A3/A– credit rating' (CC, 2008: 93, para. 11.13).

[466] E.g. funds from operations/interest cover; net debt/RAV; and retained cash flow/net debt.

[467] When applying a real WACC, the effects of inflation are compensated through indexation of the RAB. Alternatively, the cost of capital might be calculated using a nominal WACC without RAB indexation. Both approaches are neutral in NPV terms, and the election depends on the regulator. In the UK, only Ofcom uses a nominal approach.

[468] That is a common situation in large infrastructure investments. Note, however, that the indexation of the RAB may reverse this effect (i.e. it may lead to positive cash balances), but only in the long run. Note also that differing credit rating agencies use different methods when reflecting this level of capital expenditure in their ratings – situation that, considering the regulatory context, may be regarded as unavoidable. Some agencies

turn, this may result in a decrease in credit ratings; drive the rate of return up (increasing the cost of finance projects); and allegedly render it more difficult for the firm to access the financial markets.

If a firm fails the financeability test, a number of adjustments can be applied – typically to 'bring revenues forward'. In the past regulators have used NPV positive mechanisms, such as the allowance of additional revenues ('revenue uplifts') whilst warning the firms not to distribute them as dividends.[469] Also, ORR has recently provided Network Rail with a 'risk buffer' of £1bn. over 5 years, with the explicit purpose of managing risk within the regulatory settlement (ORR, 2008). The company has discretion over the use of this fund.[470] The financial effect of payments '…is to set allowed returns at a level above the cost of capital' (Jenkinson, 2006: 13); lead to re-valuations of firms' shares and premiums on purchases; and make consumers to pay more for services than they would otherwise have done.[471] The tendency, however, is to use NPV-neutral mechanisms. For instance, Ofgem have re-profiled revenues to advance cash flows from future periods by accelerating depreciation and shortening assets lives. Likewise, the CAA (2002) and the CC (2007) have used a point estimate towards the higher end of the WACC.[472] The effect of neutral-NPV mechanisms is

---

reflect actual maintenance and renewal expenditure; others use accounting or regulatory depreciation as a proxy (data on this being generally more readily available). As the two are not necessarily coincident, at least in the short run, this could result in deteriorating financial indicators even where investment is at or below the level required to renew existing assets.

[469] Yes, firms get extra revenues to spend on themselves (!). At least Ofgem and Ofwat have included vast revenue uplifts in their price controls amounting to large sums in NPV terms. In the price control 2004 Ofgem allowed one company an extra payment of £1.6m precisely to 'provide a small cushion against downside risks and improve the projected financial ratios' and to 'maintain a credit rating conformably within investment grade' (Ofgem, 2004c: 114-5). Regulators considered that 'the materiality of the adjustment was small' (Ofgem & Ofwat, 2006: 46, para. 136). In turn, Ofwat allowed payments for approximately £400m for the period 2005-10 (Ofwat, 2004).

[470] ORR (2008: 235, para. 14.43) has even argued that although it increased the protections against risk, it is 'taking a cautious approach by not reducing the risk buffer to take account of those further protections'.

[471] Fortunately, after taking stock of the weaknesses, regulators have recently signalled some future changes. For instance, Ofwat has indicated that for the next price control review (due at the end of 2009) it will try to ease the financing constraint through several market mechanisms, as it considers the revenue uplift is not appropriate. Should this tool be adopted, it would be applied in a NPV-neutral manner (Ofwat, 2007). Ofgem (2008a) has also stated that if the cost of capital is set at the right level, revenue uplifts are unnecessary.

[472] *E.g.*, the CC recommended two uplifts during the price control review of BAA in 2002, in order to reflect the financial constraint associated with the new Terminal 5 in Heathrow, along with the use of the mid-point of the WACC range. Nonetheless, they were not adopted by the CAA (which has the 'last word' in setting the prices for the airports) and instead a point estimate towards the higher end of the WACC range recommended by the CC was used. The latter approach was also applied by the CC (2007).

allegedly less detrimental for consumers: prices increase in the short-term, but in the long run they should decrease. Whatever the solution, though, consumers are affected.[473]

The solution, however, is not clear. Suitable proposals include the conversion of non-regulated assets into cash and the renegotiation of creditor protection arrangements. Their implementation, however, might be highly contentious and their reach out of the regulators' means. Further, Jenkinson (2006) has proposed the adjustment of the profile on prices within a control period to match the cash-flow profile and keep financial ratios at appropriate levels – in practice a solution closer to the options for risk-sharing analysed in section V below.[474] Equity injections may also be appropriate – with the regulator providing guidance as to whether and when it estimates new equity would be needed during the price control period.[475] Nonetheless, too much reliance on this mechanism may be imprudent. If the objective of the financeability test becomes chiefly to attract equity, the test happens to be tautological and to some extent whimsical – regulators boost equity because they want to avoid the 'flight from equity'; and the way to avoid the flight from equity is by boosting equity. This is nothing more than a convenient loophole. In the end, the call should be for a holistic approach whereby the potential actions to address the financeability concern are ideally discussed with the firm, consulted with consumers' groups where possible and implemented on an *ad-hoc* basis.

Notwithstanding the solution, it is clear that the financeability test as it is currently applied must be revised. Importantly, the over-reliance on credit rating agencies (CRAs) is unjustified. First, there is an issue of circularity. Regulators rely heavily upon the information provided by CRAs. They even meet agencies' representatives and learn about the factors considered by the latter whilst grading a firm (CC, 2008: L6, para. 25). At the same time, however, CRAs rely heavily on regulatory measures to ascribe ratings. Secondly, – and

[473] As the CC (2010b: 99, para. 10.16) has stressed, 'It is not consistent with good regulatory practice or the consumer objective to determine that customers should pay higher prices to rectify a financeability problem resulting from a company's own decision about financial structure'. A similar statement is done in CC (2010a: 71, para. 10.24).

[474] Jenkinson recognises that such an approach would trade-off stability of financial ratios against stability of price paths for consumers. But he argues that if cash flows cannot be stabilised, the regulator might alter the depreciation profile of the firm and spread the effect over the years.

[475] In fact, Cowan (2003) notes that one of the ways whereby consumers and firms can share risks without regulatory intervention is the buying of utilities' shares by consumers. However, he also points out that for this to be applicable shareholding should be implausibly large.

related – appropriate ratings in infrastructure and utilities sectors depend largely on a good understanding of the regulatory regime. It is at least doubtful that CRAs, more used to assessing standard competitive sectors, have the expertise and specialised knowledge to evaluate all the relevant implications.[476] Finally, some actions that CRAs may deem positive for the purposes of achieving a certain rating may in fact have negative consequences for consumers. Considering their general tasks, CRAs tend to prioritise actions that benefit debt-holders. Conversely, regulators need to act for the consumers' benefit.

It does not follow from the above that CRAs should not play a part in the assessment of financeability. As CEPA (2010) has correctly stressed, the alternatives are more difficult to implement and demand expertise on credit analysis that regulators do not (and should not) have. However, the recent financial crisis at least invites to adopt a cautionary approach. High profile episodes (e.g. the collapse of big financial institutions) have undermined the alleged ability of CRAs to act as 'vehicles' that monitor and spread information. The financeability test should rely less on the fact that a firm may not achieve the targeted credit rating – its achievement should be considered only as *minimum* condition. CRAs should fulfil an informational role and not be the central part of the assessment.

Finally, at a more fundamental level, it can be questioned whether the financeability concern exists at all – that is, whether any advance of cash flows that goes beyond the price set during the control is necessary. Contrary to some studies that have argued that at least in some sectors the access to equity markets is more difficult or costly (*e.g.* Oxera, 2006, regarding water), it seems improbable that utilities reach a point where they cannot access financial markets by themselves. Theoretically, as in any other business, an adequate rate of return should be enough to provide incentives to attract capital to utilities. In fact, by definition the regulated price is set higher than incremental cost, so the firm has incentives to invest and produce. Regulation only prevents *significant* departures from incremental cost (Breyer, 1982: 18-9). This fact alone should suffice to exclude any type of expectations of advancement of cash flows (i.e., short-term cash flow gaps should be allowed for by

[476] In contrast, CEPA (2010, at 9: 'It is difficult to conclude that on the basis of the methodology the credit agencies do not understand the industry').

investors); to compensate any potential increase in the rate of return demanded by investors; and certainly to avoid concerns about financeability.

Nonetheless, for this to occur it is required that the regulated price is set correctly by the regulator. This is not an oxymoron – it is a complex, but not impossible task. If the regulator calculates the WACC using notional gearing, as they do, it is possible to base that assumption taking into account long-term considerations – the more so if utilities are supposed to 'mimic' competitive markets. In turn, this requires that investors have confidence that the regulator will not expropriate long-term returns. This leads to the problem of 'regulatory commitment', revised in section 5 below. As will be shown, political risk and regulatory risk are widely seen as relatively low in the UK.[477] To a large extent, low risk reassures the availability of financial resources – either from equity or debt. In fact, so far there is no evidence showing that firms are or have been unable to raise finance.[478]

In sum, there are at most only weak theoretical and empirical grounds supporting the thinking that equity investors will not be willing to invest in utilities when debt investors are willing to do so. The financeability concern may well be overstated.

B. The 'under-investment concern'

The financeability concern is closely related with the 'under-investment concern'. According to this argument, stockholders might be willing to accept some positive NPV projects only if the firm is financed by equity, but not if it is largely debt financed – in fear that the pay-offs might be large enough to be profitable, but not sufficient to repay the debt-holders. Therefore, some 'good' projects may be inefficiently rejected.[479] The situation is aggravated if there is a possibility of financial distress, because the higher the gearing of the firms, the lower the possibility of absorbing financial shocks. In utilities, it is argued, the firm will postpone long-

[477] See *infra* chapter IV.

[478] For instance, the CC recently disallowed an increase in prices requested by one water company, on the grounds that it was 'able to finance the proper carrying out of its functions with prices at their current level' (CC: *Sutton and East Surrey Water plc: Interim Price Determination* [17 June 2009], at 78, para. 4.96).

[479] This problem is akin to the 'asset substitution problem' analysed *infra* in note 488 and accompanying text.

term investments demanded by regulators in favour of short-term decisions.[480] In the extreme, all of this may lead to under-investment.

The under-investment concern is unsound. First, it is based on wrong managerial assumptions. On the one hand, it is assumed that managers always act in the interest of shareholders or at least that shareholders can control managers' incentives and align them with theirs. As seen, that is a weak assumption. On the other hand, if the under-investment problem exists at all, it is erroneous to base the explanation only on the capital structure. If, in general, firms with less debt invest more, as financial evidence shows, there is no evidence showing *why* investment is positively correlated with less debt.[481] Indeed, agency theories have shown that even if the capital structure is 100% reliant on equity, certain managerial attitudes may lead to under-investment.

Second, and perhaps more essential, important incentives created by the own regulatory regime must be taken into account. Whilst generally PCR leaves regulators a great margin of discretion, the regime only encourages firms to make improvements that are repaid within a single review period. On the one hand, this may create incentives to cut back long-term programmes already agreed with the regulator.[482] Since many types of infrastructure are durable enough to make that underinvestment not result in a decline of quality of service or performance, the firm may be tempted to cut back on an agreed investment at the beginning of the price-cap period, increasing its profits.[483] On the other hand, PCR does not encourage efficiency improvements that payback beyond the next price control (*e.g.* Gómez-Ibánez, 2003: 241 *et seq.* and Newbery, 1999: 50-2). That is, PCR in itself provides firms with incentives to underinvest in Capex and quality of service.[484] The reason is the lack of

[480] Unlike non-regulated companies, regulators demand from regulated firms quality improvements and capital investments to be made in certain periods of time. The expectation is that they should not be reallocated across time.

[481] As also firms with more cash do. For a survey of the literature, *see* Stein (2003), who points out that the evidence does not clearly reveal *why* firms with more cash and less debt invest more.

[482] Two related effects may happen. The firm may prefer to delay improvements with short-term paybacks and make them after the price control, hoping to capture the savings as profits. Also, the firm may increase its investments as the price review approaches, hoping to get a higher cap that allows greater future savings.

[483] For this reason, with PCR the regulatory burden increases, since the agency must monitor the firm more closely to discover whether any saving is due to efficiencies rather than deviations from previous commitments.

[484] *See* Armstrong *et al.* (1994), who formally show that by itself, PCR does nothing to encourage improvements in the quality of service.

regulatory commitment. Firms will make long-term investments only if the regulator commits to recognise them and incorporate them into future price controls. However, the essence of the system is that every n-years period the price is reset. Intrinsically, a problem of time-inconsistency between short-term and long-term projects exists.[485]

In sum, the incentives to under-invest are a distinctive problem associated with PCR *regardless* of the capital structure of the firm. The under-investment problem is not as much associated with the firms' financial stability as with the regulatory regime in itself. The problem largely derives from the alleged inability of regulators to make future commitments (see section V). Therefore, it is at least dubious that controlling the capital structure is the right way to address what is in essence a design problem.[486]

C. The 'control concern'

The debt model is also disputed on the grounds of managerial corporate control.[487] It is argued that one of the key advantages of the equity-based model is that the structure of corporate control is familiar and clear. In the event of troubles, whatever the reason may be, shareholders have a strong incentive to replace the incumbent management. By contrast, when a company operates under a debt-based model, incentives become less clear and heavily depend on the deeds of covenant that are in place to protect debt-holders' rights. The ambiguous and poor control increases the risk: weaknesses in decision-making are severely exposed and result in lack of incentives for efficiency. In other words, principal-agent problems would tend to increase with debt, and hence the overall economic incentives of PCR might be undermined.[488]

---

[485] The under-investment question might be seen as the 'the flip side' of the regulatory commitment problem. The latter is related with the investors' perception of risk (*see infra* section V).

[486] As Newbery (1999: 50) points out: 'Rate-of-return regulation *evolved* through a series of landmark court cases in the United States … Price regulation was *designed* in the United Kingdom to create an efficient system of regulation…' (emphasis in the original).

[487] In this sense, Currie (2003) and DTI & HM Treasury (2004).

[488] Regarding principal-agent problems, *see generally* Milgrom & Roberts (1992). Note that even though the 'control question' is centred on the relation between managers and principals, there is also an agency problem arising between stockholders and bondholders in firms with hybrid financial structures (i.e. structures with presence of both debt and equity) – a problem known as the 'asset substitution problem' in the financial literature (Myers, 1977). The potential conflict arises from the type of projects that will be taken on by

Underlying the argument is the (incorrect) assumption that the adoption of PCR *in itself* implies an equity financial model of the firms – especially if its 'high-powered' characteristics were to be preserved.[489] Since the system works on the basis of the cost reductions set by the regulator, it requires agents responsive to those incentives. According to this argument, shareholders would be the only actors capable of responding correctly. Whilst they have incentives to outperform expected productivity of the firm and seek managerial improvements, the only incentive for debt holders would be not to engage in activities that may undermine their rather secure repayments.[490] Therefore, shareholders would be more suitable than debt-holders to fulfil the regulatory targets. The equity model of private ownership becomes not only necessary, but crucial to improve efficiency under PCR.

The 'control concern' completely misses the target – mainly because it pays no attention to the insights of agency theories. First and foremost, an equity-model of the firm does not necessarily create managerial incentives. The 'control concern' implicitly assumes that the incentives of the firm's managers and shareholders are perfectly aligned. Moreover, it is assumed that the alignment is in favour of the latter – i.e. that managers adopt their decisions

---

stockholders. Once debt holders have advanced capital to the stockholders, the latter have an incentive to undertake projects riskier than those the former would prefer. Recognising this incentive, debt holders will charge a higher price for debt capital (Jensen & Meckling, 1976). The agency costs may be attenuated by the use of contracts – covenants, call provisions, conversion rights, etc. – but they can never be eliminated as long as the debt holders cannot perfectly discern if the outcome is the result of uncertainty or is due to the actions of the manager. In turn, in a regulated environment, this agency problem will have an impact in the outcome of the WACC.

[489] For instance, the DTI & HM Treasury (2004: 35, para. 99) states that 'Shareholders play an important role in driving managers to respond to the efficiency incentives of the RPI-X regime. The corporate governance model of shareholder ownership provides a clear structure for decision-making. However, these incentives can be reduced or removed in highly geared structures, through a loss of shareholder pressure'. Likewise, a consultants' report argues that 'The introduction of equity into the industry at the time of privatisation is seen as the agent of the substantial improvements in performance and efficiency in the sector. Indeed, the RPI-X regime is seen as predicated on the active participation of equity, with the profit maximising interests of shareholders spurring management on continually to outperform regulatory settlements' (Smith & Hannan, 2003: 48, para. 6.1.2). With similar reasoning, ORR (2006: 45, para. 5.1) has stated that 'Although there are clear justifications for Network Rail's current financial structure [100% reliant on debt], it does mean that the incentive-based regulatory framework provides weak financial incentives on the company to strive for continuous improvements in performance and efficiency'.

[490] Shareholders are concerned about the upper part of the probability distribution of all the possible performance outcomes – i.e., above the outcome required to pay the debts. By contrast, claimholders receive nothing of the cash flows above the amount specified in their debt contracts. In accordance, riskier projects reduce their expected pay-offs.

for the benefit of shareholders.[491] This is not always true. Managers have an incentive to over-expand the size and scope of the firm to satisfy their own ends at the expense of the shareholders (Jensen & Meckling, 1976). Particularly, managers in profitable companies are likely to waste money on inefficient investment. One of the reasons is the personal benefit associated with the so-called 'empire-building': a larger firm (allegedly) creates more opportunities for career enhancement and promotions, higher rewards and status. In addition, managers may be more concerned about their own reputation and career. If that is the case, they have incentives to take short-term actions at the expense of the long-run shareholder value; exhibit an excessive tendency to 'herd' in their investment decisions ignoring their own private information; or simply under-invest to not reveal their managerial ability.[492]

Indeed, shareholders may intend to exert control over managers in an attempt to avoid the misalignment of incentives. On this, however, the 'control concern' underestimates collective action problems resulting from dispersed ownership. Primarily, shareholders have an incentive to free-ride. The costs of monitoring managers are high for an individual shareholder, who must invest a vast amount of time and effort to check the firm performance. This problem is especially acute the smaller the relative wealth of each shareholder is to the size of the firm. Hence each shareholder would prefer others to carry out that function. If free-riding is prevalent, the result is a low and inefficient level of monitoring.[493]

But even if managers act in the interest of shareholders (or the monitoring is optimal), the 'control question' still assumes that equity is associated with proactive managerial attitudes that would not be present in debt-models. That, I submit, is wrong. Indeed, at the extreme some contracts may *de facto* reproduce the property rights structure normally seen in an

[491] As mentioned in Chapter II, this 'shareholder value maximization', according to which shareholders are legally entitled to demand managerial deference to their interest (even if this is limited by *de facto* opportunism of managers), is a traditional assumption of the Anglo-American corporate environment. It is normally understood as a sort of 'duty' to generate (high) profits. Williamson (1984) argues in favour of this assumption, stating that shareholders are relatively less well protected than other parties such as workers or creditors. Since shareholders need protection the most, rules should primarily be designed to protect their interests.

[492] See Stein (2003) for a revision of the literature on these issues.

[493] There are mechanisms to solve the collective action problem. Two examples are the executive compensation models and the blockholder models. Systems of compensation for performance, commonly seen in firms, aim to align *ex ante* managers' interests with those of shareholders. A semi-concentrated ownership structure with at least one large shareholder may also help to improve monitoring.

equity-based model, so the incentives to managerial performance may remain in place.[494] However, with or without protections the argument is still flawed. There is no evidence whatsoever supporting some sort of proactive managerial attitude. On the contrary, the literature points out that managers may prefer a 'quiet life' and be prone to excessive inertia (Aggarwal & Samwick, 1999; Bertrand & Muullainathan, 2003).[495] This resistance to change may eventually lead to the continuation of negative-NPV projects and/or underinvestment – precisely one of the main regulatory concerns.

On the other hand, managers may have incentives to outperform even within firms heavily reliant on debt – mainly because they face high personal costs associated with financial distress.[496] Amongst these costs are reputational and career concerns, costs associated to finding another job, etc. Only by exerting effort managers can meet the interest payments on debt and avoid the pernicious consequences of bankruptcy.[497] In the end, however, the (personal) costs associated with bankruptcy may also lead to over-expand the size and scope of the firm, because a larger firm decreases the risk of bankruptcy if the company is well-diversified. In this sense, Jensen has argued that increasing gearing may solve the overinvestment problem (Jensen, 1986).[498] Increasing debt decreases the scope for inefficient managerial behaviour because managers are committed to return cash to claimholders. 'Free cash-flow' (i.e. residual cash flow after all positive NPV investment needs have been met)

[494] *See* Smith & Warner (1979) for one of the first treatments about the use of covenants in bond contract to prevent opportunistic behaviour of firms in financial distress. On covenants, *see generally* Tirole (2006: 85 *et seq.*).

[495] This finding is not strange in utilities regulation: remember the 'quiet life' of the monopoly (HICKS, 1935) and the so-called 'X-inefficiency' (Leibenstein, 1966). Although the latter is more associated with overinvestment and managerial 'empire building', when adapted to PCR it might well be the result of underinvestment (*see supra* section IV.B).

[496] Titman (1984) argues that managers personally bear a substantial portion of the bankruptcy costs. *See also* Fama (1980) for one of the first discussions on career concerns; and Holmström (1999) for a formal model of how managers' actions affect their perceived value in the labour market.

[497] Note that this threat is not applicable to Network Rail, a company whose debt is guaranteed by the State. The same reason undermines the positive effect of the more efficient monitoring: bondholders have their payments guaranteed regardless of the firm's performance.

[498] Others have also stressed the 'agency benefits' to debt if it reduces the scope of managerial discretion: *e.g.* Grossman & Hart, 1982). As Jensen (1986: 323) states: 'The problem is how to motivate managers to disgorge the cash rather than investing it below the cost of capital or wasting it on organizational inefficiencies'. Note that if managers are unwilling to increase debt, there are incentives for other groups to take over the firm and restructure it. Indeed, as Jensen recognises, debt does not always have positive control effects. Rapidly growing firms with large and profitable investment projects but no free cash flow will need to go regularly to financial markets to obtain capital. In this case investment bankers fulfil the role of monitoring. On the contrary, the control function is more important in organisations with large cash flows and low growth prospects and in organisations that must shrink.

will be used to pay off lenders. Therefore, the higher the level of debt, the more powerful the incentives created for managers to work harder to meet the debt-holders payments.[499]

In sum, agency theories demonstrate two crucial findings that undermine the 'control concern'. First, that the effect of managerial incentives is unclear: they may lead either to underinvestment or overinvestment *regardless* the capital structure of the regulated firm. The final effect will depend largely upon managers' attitudes and other incentives created by the regulatory regime.[500] Secondly (and maybe more fundamental), PCR in itself has no specific financial model attached to it. The contrary assumption is merely a misconception.[501] Shareholders are not a necessary condition for PCR to operate, nor are they a necessary condition for the efficiency of the regulatory regime.[502]

---

[499] For a critique, Myers (2003: 243) (warning that the free cash-flow theory does not provide a model for the incentives and actions of managers and does not indicate how they should arrange financing).

[500] This may be observed from the attempts to apply PCR to non-stockholder owned firms, most notably Network Rail, Glas Cymru (Welsh Water) and Royal Mail. The key question is how the financial incentives inherent in PCR (e.g. to reduce unit costs) are transmitted. 'Received wisdom' holds that the scrutiny of financial markets in which regulated firms must compete for resources, whether debt or equity, are an important element in the transmission of regulatory incentives. A priori, there is no particular capital structure that performs better in this sense than others. However, it may be argued that only firms with a sufficient financial 'cushion' to absorb risk can compete successfully in debt capital markets. (Indeed, the extent of the cushion required is, of course, a function of the perceived risk and also of the level of risk appetite in the market at any time. In other words, it is a dynamic function. This may in part explain the equity 'heavy' model adopted in the UK at time of privatisation). Whether the financial 'cushion' takes the form of stockholder equity, accumulated reserves (as in a mutual) or external support (e.g. the Government indemnity of Network Rail's bonds) has implications for managerial incentives and, thus, for the design of regulatory incentives. There is literature around this: for example, Bolt (2010). *See supra* section IV.B, explaining the incentives to under-invest created by PCR.

[501] In its origins the regulatory regime was never associated with any specific financial model. The financial structure is not even mentioned in the report that was used as a base for the introduction of RPI-X in the UK (Littlechild, 1983). As seen in the introduction, the equity-model was a feature purely associated with the politics of privatisation. This aim, which underlay the first wave of privatisations, was intact when utilities started to be privatised in 1984. As Foster (1992: 124) points out, '[i]n turning British Telecom [the first utility being privatised] into private company whose shares were sold to the public, the creators of this structure were following precedent'. The regulatory regime complemented privatisation. The new regulatory regime was instrumental to privatisation (*e.g.* Beesley & Littlechild, 1989: 457: 'The initial level of X is set by the government at the time of privatization, *as part of the privatization process* … the initial level of X is set as part of a whole package of measures…' [emphasis in the original]; Newbery, 1999: 50: 'Price regulation was *designed*…to enable publicly owned utilities to be transferred to private ownership' [emphasis on the original]). Protecting the firms' profits would encourage new shareholders to continuously inject cash into the firms. By contrast, according to Vickers & Yarrow (1988), before privatisation the regulatory regime was no object of concern because the industries were reasonably competitive. Despite being complements, however, the concerns of one and the other regime may be clearly differentiated.

[502] Note that even when the plans to privatise British Telecom started, the improvements in efficiency were thought to arise from a source different than the RPI-X. The government emphasised the idea of allowing BT to borrow freely from the capital markets, without the constraints faced in the public sector (DTI, 1982) – i.e., 'privatization would facilitate more efficient capital allocation' (Vickers & Yarrow, 1988: 158).

V. REGULATORY COMMITMENT

Most of the regulatory concerns converge on the issue of regulatory commitment (or sometimes called the 'time-inconsistency problem') – how does one encourage long-term investment, considering the impossibility of binding future regulators and/or politicians? A number of solutions to deal with this problem have been proposed (e.g. Helm, 2009; CEPA, 2010), and there is a large amount of discussion around this certainly important issue. However, for a number of reasons, I doubt regulatory commitment constitutes an issue as important in UK regulation as it appears to be.

From a 'macro' level, regulatory commitment is largely a reputational problem – and on that regard it is undeniable that UK utilities face very low political and regulatory risk. The fact that utilities were a low-risk business had already been acknowledged shortly after BT's privatisation – the first utility privatised in the UK. As Foster (1992: 213) recounts, when in 1986 the then director of the telecoms' regulator (Oftel, Ofcom's predecessor) approved BT's proposed price increases, he warned that when he renegotiated the X-factor in 1989 (as he was required to do so by the licence) he bore in mind what rate of return BT was earning. 'He said he felt that he had no choice, given its low level of risk'.[503] Indeed, some actions and decisions adopted by regulators in the past may have slightly increased the risk perceived by investors. But generally there is a strong precedent of 'responsible behaviour' sustained on powerful institutions that constrain opportunistic behaviour. Despite the lack of constitutional protections, the steady evolution of the regulatory frameworks (with enough check and balances embedded on them) and the strong protection of property rights by an independent judiciary, reassure commitment to a large extent. True, institutions are perfectible and responsible behaviour may be insufficient to build credible commitment – thus there is always some level of remaining risk. However, overall the literature is right in agreeing that commitment concerns remain low.[504]

[503] *See also* Oftel (1986b: 9).
[504] In this sense, *e.g.*, Cowan (2001), Newbery (1999) and Levy & Spiller (1996).

From a more 'micro' perspective, that view is sustained in the large number of measures to prevent potential situations of financial distress (either individual or systemic) already embedded in the system.[505] First, a wide range of information on the financial health of regulated firms is available to regulators. On the one hand, there are a number of market indicators of financial distress they monitor on a regular basis – *e.g.*, share prices and credit ratings. On the other hand, regulators collect information about the firms from a number of different sources, including cost and revenue reporting arrangements and annual visits to the companies' premises in order to follow up on the reporting submissions. Regulators normally use that information to observe firms' performance against their price controls. They also collect information on the network companies' Capex and monitor changes in that expenditure to ensure the interests of consumers are safeguarded against under-investment.

Second, it has been argued that regulators should adopt *ad-hoc* measures in case of deterioration of the financial conditions of the firm (section IV.1). However, there are already a number of measures specially designed to address a case of financial distress. A good example is the so-called 'ring-fencing licence conditions'. These are several provisions whose purpose is primarily to ensure that 'assets, cash-flows and other financial resources of regulated energy network operators are applied to meet the needs of the regulated companies. The conditions seek to ensure that those resources are not diverted to any other purposes nor exposed to any unrelated risks'.[506] Likewise, they allow the regulator to obtain access to the information in order to monitor the financial position of the firms, have early warning of severe deterioration and being able to take effective action in the event of a breach. Certainly, as it has been recognised by regulators, the financial health of a firm may deteriorate too quickly for the ring-fence conditions to be effective. But in addition regulators count with general enforcement powers to deal with a broad range of circumstances.

[505] For a summary of these measures in the electricity sector, *see* Ofgem (2008a). Additionally, as CEPA (2009: 23) points out, other mechanisms within the price control that reduce uncertainty are not incorporated in the assessment of risk, which normally remains unchanged between the initial and final regulatory proposals for the cost of capital.

[506] Ofgem (2008c: 7).

Ultimately, the provisions for 'special administration' that replace insolvency proceedings in some sectors are particularly important.[507] Under a special regime, an administrator is appointed primarily to fulfil the regulatory obligations of the insolvent firm – i.e. the administrator is tasked with continuing to develop and maintain the network.[508] With that aim, a number of special measures may be implemented. For instance, licence conditions may be modified to secure funding, and the Secretary of State may even provide loans and grants, guarantees or indemnities to the failed firm. Also, if there are outstanding costs from the administration process, they are recovered from other licensees and passed on to consumers. The special administration ends only when either the licensee is 'rescued as a going concern' or its assets are transferred to another party.[509] In this process, the interests of shareholders and creditors are secondary: the principal focus is on rescuing the firm in distress.[510]

When combined, all these measures largely mitigate the potential impact of financial failures. Certainly, they do not dissipate the risk of distress completely. Some actions can only be adopted after a failure, and therefore provide little guidance on what could happen when the finance becomes an issue between price controls. Also, the measures do not allow the anticipation of possible situations of something akin to systemic risks faced by some utilities.[511] But when considered together, the measures reassure investors that the risk of losing their investment has been kept to minimum levels.

---

[507] Special administration regimes are set out in Part 3, Chapter 3, sections 154-159 of the Energy Act (2004), for gas and electricity companies; in Part II, chapter III, sections 23-25, and schedule 3 of the Water Industry Act (1991), for water companies; and in sections 59-65 of the Railways Act (1993). Administrative complementary rules and some sections of the Insolvency Act 1996 may apply, depending on the sector. There is no such special administration regime for airports (only for air traffic service licence companies), but the common company law provisions apply.

[508] This is the essential difference with the function of a liquidator or a receiver, whose principal responsibility is to sell enough assets to pay off the firm's debts. In the special administration regime, the administrator manages the assets. At issue is not only the assurance of continuity of supply. There is also an economic underlying reason: many of the firm's assets are specific and hence cannot be redeployed in alternative uses.

[509] *E.g.* Section 155(2) of the Energy Act 2004.

[510] *E.g.* Section 158(3) of the Energy Act 2004 states that the administrator must exercise and perform his powers and duties in the manner that best protects the interests of the creditors of the company, but only so far as they are consistent with the objective of the energy administration.

[511] As Hood *et al.* (2001: 178) indicate, 'Government central *raison d'etre* is indeed often held to include 'system risk' roles such as risk-taker of last resort and the regulation of collective or public risk […]'. However, so far regulators have not been keen to include systemic failure in their price-cap calculations and offset it. This is to some extent comprehensible, since that position would imply the recognition that most of the risk of financial distress is borne by customers.

On top of this, there is no clear commitment not to bailout firms – in fact, the situation seems to point to the opposite.[512] Although bailouts are in principle discarded for inefficient companies, regulators have been unable to fully commit against State financial aid.[513] Each case is likely to be considered on its own merits. This opens the door to firm-specific bargaining and implies that the government will certainly intervene bailing out utilities – at least in extreme cases. Furthermore, the own regulatory framework openly provides a sort of 'institutionalised' alternatives to bail out utilities. There is an option for firms to apply for the (unilateral) 'disapplication' of their revenue restrictions or a 'reopener' of the settlement. The latter possibility arises when some exceptional circumstances in which the revenue stream set in the price control ceases to provide sufficient funds for the regulated firm. In such cases, the 'reopener' allows to re-set revenue allowances or the parameters that give rise to those allowances. In some cases, however, the re-opener provisions are unlikely to be of use in *relieving* financial distress, because they tend to be related to the occurrence of specific events. A disapplication request enables control parameters to be reset for a broader range of reasons than those determined by specific reopener provisions and hence is more likely to be applied.

It might be argued that the regulatory framework protects against misuse of these provisions. In reviewing price controls, regulators must follow their statutory duties. Thus, to grant regulatory relief which would alleviate financial distress, the regulator must be satisfied that the cause of the distress is: (i) unforeseen when the price control was set; (ii) outside the firm's scope to avoid; and (iii) mitigated to the greatest extent practicable by management action. In principle, this should exclude distress caused by the impacts of a capital structure choice, which would be avoided with a different feasible structure. However, there is no *specific* statutory duty to follow these principles and, in fact, these are measures that has been used by regulators more often than expected. For instance, Ofgem has made provision for

[512] Calls for bailouts increase in times of general financial crises, mainly due to three effects. Financial crises raise the cost of funding projects; induce pressures to cut prices; and scaling back expectations of future demand.

[513] Regulators normally stress that the risk of financial failure must be borne by the firms, not consumers – especially when it arises from firm's action or inaction. Nonetheless, the message is contradictory. *E.g.* Ofwat & Ofgem (2006: 27, para. 80) paradoxically have noted that '[…] it is not possible to rule out the possibility that a regulator may be asked by the Special Administrator to consider a case for re-opening price limits', although 'in making any changes to price limits, regulators will want to ensure that it is investors not consumers that would be expected to bear the costs arising from inefficiency'. Ofgem (e.g., 2008c: 13, para. 2.31) and the CAA have expressed in similar sense.

price control re-openers in each of its current price controls that relate to specific events.[515] In water, there is the so-called 'shipwreck clause' or 'substantial effects' provision, whereby firms may ask for a revision of their K factors when facing adverse and unavoidable circumstances. Recently, ORR (2008) has also issued provisions regarding the events that may trigger a re-opener. In brief, the government not only has the ability to absorb insolvency risk; it also seems willing to do so.

VI. CHAPTER CONCLUSIONS

All of the above leads to the conclusion that, in reality, utilities face very low levels of risk in the UK. If the State acts as a guarantor of risk, it should be recognised within the regulatory models – for the benefit of consumers. Regulation may cost a great deal for the latter and firms alike when it is based on wrong assumptions and false concerns. Unfortunately, the assessment of the main regulatory responses to the changes in the capital structure of UK utilities firms highlights that many of the evils of high gearing are exaggerated. Likewise, the effectiveness of various aspects of the regulation of the capital structure and the setting of the cost of capital are highly questionable. Some risks are overstated and some approaches lack adequate justification. Most problems or failures of regulation, however, are not systemic, but accidental. The current search for improvements may be enhanced not only with regulators' fine-tuning in their approach to price controls, but also with future research addressing at least some of the downsides here detected.

[515] *See e.g.* Ofgem (2004c & 2007a). A similar measure is the so-called 'Capital Expenditure Safety Net' in the current transmission price for the four gas and electricity transmission companies: if a firm's investment falls below 20% of its Capex allowance at any stage in the five-year control period it will trigger an automatic review of that company's allowance (Ofgem: 2006, paras. 7.13 *et seq.*).

APPENDIX 1: MEASURING THE COST OF CAPITAL

A. The WACC

UK regulators usually assess the cost of capital using the so-called 'Weighted Average Cost of Capital' (WACC). Although the WACC is not the only method to assess the cost of capital, it has been widely applied in most utilities sectors. The WACC calculates an average (not marginal) cost of capital over the basis of two components: the expected cost of the debt and the cost of the equity. Formally:

$$(1) \qquad WACC_i = g_i R^D_i + (1 - g_i) R^E_i$$

where,

'i' represents any firm;
g_i is the proportion of debt finance or 'net gearing';
R^D_i is the required rate of return on debt;
$(1 - g_i)$ is the net proportion of equity; and
R^E_i is the required rate of return on equity.

The result of adding the cost of debt and the cost of equity gives the cost of capital for the regulated firm as a whole. Both types of costs are computed over the basis of a number of individual components, some estimated by the regulator and others market-determined. Given the individual component estimates, the result is normally a range. Regulators need to select a single point estimate from within that range. That is a matter of judgement involving a trade-off between benefits for consumers against the cost implications.

B. The Cost of Debt

The cost of debt (R^D_i) refers to the expected return on debt. For instance, if a bondholder is paid a fix coupon rate of interest, the regulated firm should be allowed to earn enough profit to pay that interest, exactly as an unregulated firm treat their own bondholders.[516] Formally,

$$(2) \qquad R^D_i = g_i(r_f + \rho)$$

where,

g_i is level of net gearing;
r_f is the risk free rate; and
ρ is the 'debt premium' over the risk free rate

- g_i is defined as:

$$(3) \qquad g_i = \frac{ND_i}{ND_i + VE_i}$$

where,

ND is the net debt of the firm; and
VE_i is the market value of the equity.

ND is equal to the interest bearing debt without considering the cash and short term securities, and it is estimated from the book value information from the accounts of the firm.

- r_f is the return on any safe investment for which there is no repayment risk (or that risk is the minimum). This is the benchmark against which all other investments in an economy should be assessed. Hence, the risk free rate is normally assumed to be fixed. The most common gauges used as base value are the Treasury (Government)

[516] Note that, in the regulatory context, normally preferred stock is treated like debt if it has a fixed dividend. Nevertheless, differences in the taxation regime remain.

Bonds, given the ability of governments to raise finance through taxation.[517] Alternatively, it can also be measured by the return on indexed-linked long-term bonds.

- ρ reflects the investor subjective reaction to a number of uncertainties that its investment will confront during the investment horizon. In practice, it is either measured directly from the observable yield of the firm's bond (Y^D_i) or through comparator information.[518]

Alternatively, the cost of debt can be (also easily) calculated taking into account the probability of default of the firm (π) and the observable yield on its debt. The probability of default is usually estimated from default rates on debt of a similar credit rating published by credit rating agencies. Formally:

$$(4) \qquad R^D_i = (1 - \pi)\, Y^D_i$$

where,

$(1 - \pi)$ is the probability the firm does not default on its debt; and
Y^D_i is the observable yield of the firm's debt.

C. The Cost of Equity

The determination of the cost of debt is relatively uncontroversial. By sharp contrast, the source of many regulatory disagreements has historically been the cost of equity (R^E_i) – i.e., the amount of profits required by the firm over the amount needed to pay bondholders.

[517] With some exceptions, in developing countries the US or UK interest rates on Government Bonds are taken as base value, rather than local rates. The risk in the case of Treasury Bonds is normally limited to inflation risk.

[518] Both components are easily observable. Yields on new bonds are readily available from commercial information sources on a daily basis (e.g. the Financial Times).

There are several financial methods to calculate the proper return on equity.[519] Regulators have applied a number of them with more or less success. The 'comparable earnings method' was the traditional system used by the public utilities commissions in the US. Under this method the regulator first chooses some comparable industry and then assesses its RoR, which is applied to the regulated firms.[520] A second method employed is the 'Discounted Cash Flow' (DCF). It was first used during the 60's in the US, but it has also been studied in the UK. In the simplest version, the regulator assesses the cost of equity equalising it to the sum of the next period's dividend yield (i.e. the dividend divided by the share price) plus an expected rate of growth of dividends.[521] Note, however, that whilst the divided yield is observable, an assumption must be made regarding the growth rate, which in addition must be assumed constant. Hence, the level of regulatory discretion in this method is relatively high.

Lately, the so-called Capital Asset Pricing Model (CAPM) has been the main method applied by regulators.[522] The CAPM model has different versions, but the 'standard' or 'single factor'

[519] Breyer (1982: 43-7) presents a summary of the different methods in regulation. For a general financial discussion, see Patterson (1995).

[520] Such method was upheld by two old decisions by the US Supreme Court. In *Hope* the Court stated that '…the fixing of "just and reasonable" rates, involves a balancing of the investor and the consumer interest…The investor interest has a legitimate concern with the financial integrity of the company whose rates are being regulated…By that standard the return to the equity owner should be commensurate with returns on investments in other enterprises having corresponding risks. That return, moreover, should be sufficient to assure confidence in the financial integrity of the enterprise, so as to maintain its credit and to attract capital.' (*Federal Power Commission v. Hope Natural Gas Co.*, 320 U.S. 591, 603 [1944]). Likewise, in *Bluefield* the Court said that 'A public utility is entitled to such rates as will permit it to earn a return on the value of the property which it employs for the convenience of the public equal to that generally being made at the same time and in the same general part of the country on investments in other business undertakings which are attended by corresponding risks and uncertainties; but it has no constitutional right to profits such as are realized or anticipated in highly profitable enterprises or speculative ventures…A rate of return may be reasonable at one time and become too high or too low by changes affecting opportunities for investment, the money market and business conditions generally.' (*Bluefield Waterworks and Improvements Co. v. Public Service Commission*, 262 U.S. 679, 692-93 [1923]). The main problem of the system is the election, by the regulator, of an industry 'comparable' to the regulated one.

[521] *See* SIEGEL (1985) for a description of this methodology. Fundamentally, the DGM reverses the hypothesis according to which if the dividend growth is constant, the price of the share is the initial dividend divided by the difference between the cost of equity and the rate of divided growth.

[522] Both the DCF and the CAPM model are complementary. *See* Patterson (1995: 21): 'In practice, because one is implicit [DCF] and the other explicit [CAPM], the two approaches complement each other as largely independent assessments of the same unobservable quantity, the equity investors' required return'. In practice, DCF is normally used as a check on CAPM-based results.

CAPM is the most widely used.[523] The model assumes that the cost of equity for a given firm is equal to the expected return on investing in a single share in that firm. Formally:

$$(5) \qquad E(R^E_i) = r_f + \beta_{Ei} (E(R_m) - r_f)$$

where,

r_f is the risk free rate;
β_{Ei} is the firm's equity beta; and
$E(R_m) - r_f$ is the 'equity risk premium'.

- r_f was analysed in the previous section. The concept of Beta is analysed in Appendix 2.

- $E(R_m) - r_f$ is composed by the difference between the cost of equity for an 'average' firm (also known as 'market return') ($E(R_m)$) and the risk free rate r_f. The result shows the level of additional return required by an investor to hold equities in a given firm in preference to the risk free instrument.

  $E(R_m)$ is market-estimated. The estimation can be based on an arithmetic mean or a geometric mean. The former is an average of the percentage returns in each year covered by the sample. The latter is the compound rate of return over *t* years.

The result is generally not the same – the arithmetic mean will normally be higher than the geometric mean. As one author explains,

> [A] high return appears to be lower when it is seen as part of a compound average growth rate (i.e. the geometric mean), because the investor will be measuring the rate of growth by reference to a measure of capital employed during the period that includes part of the

[523] The single factor CAPM was first developed by Sharpe (1964) and Lintner (1965), and it is essentially a linear model, because the expected excess of return on a given asset *vis-à-vis* the risk free rate is in fixed proportion to the expected excess return on the market. The asset's beta determines the degree of proportionality. Alternative non-linear approaches (i.e. where this proportionality does not hold) have been developed, but they are not much applied in practice. A 'conditional CAPM', where the parameters of the model vary over time, have also been advanced without much success in practice. See Wright *et al.* (2003: esp. chapter 3).

> growth itself. // By contrast, if the return is seen as a risk or insurance premium, i.e. part of income or consumption — as is the case in setting price controls — then the only relevant measure of capital is the capital employed at the beginning of the period.[524]

There is considerable debate about each of the terms in the CAPM equation. The debate has affected the regulatory determination of the cost of capital and particularly the attitude of regulators towards the assumption of some components. Yet it remains the most uncontested model to assess the cost of capital and, accordingly, the most widely used.[525]

D. The Effects of Taxation

Replacing equations (2) and (5) in (1), the WACC is given by:

$$(6) \qquad WACC = [g_i(r_f + \rho)] + [(1 - g_i)\, r_f + \beta_{Ei}\, (E(R_m) - r_f)]$$

This equation does not contain any tax adjustment, neither in the cost of debt nor in the cost of equity. This form is the so-called 'vanilla' form of the WACC. According to the 'vanilla WACC', the cost of capital is calculated over the basis of a pre-tax cost of debt and a post-tax cost of equity, as a cash flow item, and added to the operating costs of the business.

Nonetheless, the regulated tariff must provide firms with enough revenues to meet their (corporation) tax liabilities. Considering the effects of taxation, the cost of capital can also be calculated using two other approaches: a pre-tax approach and a post-tax approach.

In the pre-tax approach (also known as 'tax wedge'), the *cost of equity* is 'grossed up' by the tax shield, i.e. it is multiplied by a 'wedge'. Thus, the post- tax cost of equity is converted in a pre-tax cost of equity. Formally:

---

[524] Reckon Open, *'Arithmetic and geometric mean returns | viewpoint: Franck'*, 2008-05-28, available at: <www.reckon.co.uk/open/Arithmetic_and_geometric_mean_returns_%7C_viewpoint:_Franck> (last visited: 16 April 2011).

[525] Arguably the most serious contender to the CAPM is the 'Arbitrage Pricing Theory' (APT) of Ross (1976). As CAPM, it is a linear model. But in contrast to the latter, it assumes that there is more than one factor influencing asset returns – i.e. there can be more than one source of systematic risk. Another multifactor model is the French and Fama multifactor CAPM (Fama & French, 1996). Nonetheless, difficulties to satisfactorily identify the relevant factors remain a key problem of these multifactor models.

$$(7) \qquad \text{pre-tax } WACC_i = [g_i(r_f + \rho)] + \left[(1 - g_i)\, \frac{r_f + \beta_{Ei}\,(E(R_m) - r_f)}{(1 - t_c)}\right]$$

where,

t_c is the corporation tax rate.

The corporate tax rate used in the calculation is normally the standard marginal tax rate of the corporation tax. At the UK statutory corporate rate of 28%, the factor $1/(1 - t_c)$ is approximately equal to 1.38.[526] Despite the fact that in the short term the effective rates of firms deviate quite substantially from this assumption, in the long run the average effective rate should be very similar to the standard rate. Therefore, when the resulting cost of capital is applied it provides enough revenues for investors to meet their tax liabilities.

Under a post-tax approach, the *cost of debt* is adjusted for the tax shield:

$$(8) \qquad \text{post-tax } WACC_i = [g_i(r_f + \rho)\,(1 - t_c)] + [(1 - g_i)\, r_f + \beta_{Ei}\,(E(R_m) - r_f)]$$

Since interest is deducted before tax is calculated, this formula captures the tax benefits associated with gearing up. Nevertheless, two downsides make this approach inconvenient for use in practice. First, if tax payments are being allowed for separately as a cost (i.e. the tax liabilities are being considered as part of the efficient costs of the regulated business), in parallel, regulators need to establish policy on the extent of pass-through of tax – especially in presence of high levels of debt. The increase in use of debt reduces the tax levels of the firm, and hence required revenues. This critique also applies to the 'vanilla' version of the WACC. Secondly, the interest payable on debt has already taken into account taxable profit.

Theoretically both approaches can be equalised. When the same number for corporation tax is adopted, *ceteris paribus*, the pre-tax cost of capital translates into the post-tax cost of capital (and vice-versa) according to the following formula:

[526] The current rate and factor were used by the CC (2007 & 2008). At the former UK statutory corporate rate of 30%, the factor $1/(1 - t_c)$ is approximately equal to 1.42. This was mostly used in past calculations.

(9) post-tax $WACC_i$ = pre-tax $WACC_i$ $(1 - t_c)$

Therefore, the final election depends on the preferences of the regulator and the characteristic of the industry.

E. Summary

Table A.1: *Summary of main elements of the WACC reported by some UK regulators in the latest price controls reviews.*

Regulator:	**OFGEM**			**OFWAT**	**CC**			**ORR**
WACC	*DCPR5 2009*	*TPCR4 2006*	*GDPCR 2007*	*PR09 2004*	*Bristol Water*	*Stansted 2008*	*Heathrow & Gatwick 2007*	*CP4*
Risk free rate (pre tax) (%)	2.00	2.50	2.50	2.0	2.0	2.0	2.5	
Debt premium (pre tax) (%)	1.35	1.25	1.05					
Cost of Debt real (pre tax) (%)	3.6	3.75	3.55			3.4 – 3.7	3.55	
Equity Premium real (post tax) (%)	5.25	4.75	4.50	5.4	5.0	3.0 – 5.0	2.5 – 4.5	
Equity Beta	0.90	1.00	1.00	0.94	0.92	1.00 – 1.24	0.90 – 1.30*	
Cost of Equity real (post tax) (%)	6.73	7.00	7.25	7.08	6.60	5.00 – 8.20	4.75 – 8.35**	6.50 – 7.00
Debt gearing (%)	65	60.0	62.5			50	60	
Vanilla WACC (%)	4.7	5.05	4.84					
WACC pre tax (%)	---	6.25	5.97			7.1	6.2 (H) – 6.5 (G)	
WACC post tax (%)	---	4.40	4.18	---				

Source: Various regulatory reports.

* The table expresses the minimum and maximum point-estimated for *both* airports included in the report. The range for Heathrow was 0.9 – 1.15; and for Gatwick 1.0 – 1.3. *See* CC (2007).

** The table expresses the minimum and maximum point-estimated for *both* airports included in the report. The range for Heathrow was 4.75 – 7.68; and for Gatwick 5.00 – 8.35. *See* CC (2007).

APPENDIX 2: ESTIMATING BETAS

The firm's equity beta is a concept that captures the sensitivity of the firm's equity to 'systematic' risk – i.e., the risk that is common to an entire class of firms in the market. Thus, this coefficient allows assessing the relative 'riskiness' of its equity when compared with the market as a whole.[527] It is the only element in the model that is specific to the firm. There are no theoretical boundaries for the value of betas. A positive beta will imply that the price varies in correlation with the market price. A beta close to the unity implies that the firm behaves similarly to an average firm in the market (and therefore the expected return on equity $E(R^E_i)$) will be equal to the expected return on the market. A beta hypothetically zero would mean that the price is not correlated with the market. Finally, a negative beta implies that the value of the firm's equity generally decreases when the market goes up. Normally the beta coefficient is calculated using between three and five years worth of daily or weekly share prices. In regulation, however, it has been argued that different frequencies can be used for estimates of different companies (Wright *et al.*, 2003: 104).

A. Asset Betas

The systematic risk represented by equity betas (β_{Ei}) in the single-factor CAPM includes the combined effects of ***both*** business and financial risk. Business risk is associated with the unique circumstances of a particular firm and the overall economic environment. It arises from the operating characteristics of an investment in real assets. Therefore, it basically represents the possibility that the firm will not have the cash flow necessary to meet its operating expenses.[528] Financial risk arises from the means by which the assets of the firm are financed. It represents the possibility that the firm will not have the cash flow necessary to

[527] Technically, the beta factor is the covariance of the returns of an asset and the market dividend (i.e. the return on a well-diversified portfolio, usually the stock market as a whole) divided by the variance of the market (i.e. the variance of the return on the diversified portfolio). Under the common financial assumption that 'the future resembles the past', the estimation is rather straightforward: it is the slope of the regression line of the firm's rate of return against the market return.

[528] Some elements that are included in the business risk are revenue uncertainty, labour cost uncertainty, regulatory and political uncertainty, and risk of technological obsolescence, amongst others. The distinctive feature of all of these elements is that they are independent of the means by which the assets are financed.

pay its equity holders when the assets are partially funded by borrowing. For this reason, a firm with no debt has no financial risk.[529]

However, the first and foremost factor affecting the cost of capital is simply the underlying business risk. Therefore, in order to assess the appropriate cost of capital the regulated firm will be allowed to earn, it would be necessary to remove the effect of the financial risk or gearing from equity betas – i.e., the equity beta has to be adjusted to derive a so-called 'asset beta' (β_{Ai}). The asset beta reflects only the business risk in the markets where the company operates. If a company has no gearing (it is an all equity firm), and hence no financial risk, its equity beta and its asset beta are identical. By contrast, both concepts are dissimilar in a geared firm, because the corresponding level of gearing must be taken into account, along with the corresponding risk (β_{Di}).[530] Accordingly:

$$(10) \quad \beta_{Ai} = \beta_{Ei} (1 - g_i) + \beta_{Di} g_i \qquad \text{if } g_i \neq 0$$

In the UK, the calculation of asset betas has been considered an alternative method to the WACC and has not being widely used by regulators.[531] As seen, the practice is normally to assess equity betas and use them in the context of the CAPM. Yet during the price control review of the London Airports in 2007 the CC used the asset betas to determine the corresponding equity betas (CC, 2007).

[529] Amongst the elements included in the financial risk are the risk of bankruptcy; the risk of restructuring in event of default; and the risk of suboptimal operating and inefficient investment decisions when the firm is in condition of financial distress.

[530] In order to simplify the calculation, a general assumption of $\beta_{Di} = 0$ is normally applied. The assumption is questionable, though, especially when debt accounts for a large proportion of the capital structure of a firm.

[531] Despite the differences, the overall pattern of both equity betas and asset betas over time should be similar.

CHAPTER IV

REGULATORY DESIGN, RISK ALLOCATION AND STATE GUARANTEES

> 'There are things we know that we know. There are known unknowns; that is to say, there are things that we now know we don't know. But there are also unknown unknowns; there are things we do not know we don't know'.[532]
>
> - Donald Rumsfeld

OVERVIEW

The discussion on the cost of capital and the capital structure of utilities firms, developed in the previous chapter, leads to broader arguments about what the role of the State in infrastructure industries is and, crucially, the significant role that risk plays in economic regulation. Risk is, however, a fuzzy notion that can be studied from several perspectives. One of them, accurately captured in the words quoted at the beginning of this chapter, is related to actors' attitudes facing risk, which affect the outcome of their actions. Beyond the controversial context in which those words were expressed, they condense a quite complex truth: knowledge has different states that affect behaviour, some of which are not even possible to grasp. A second, different perspective is the consideration of risk as justification for regulation. Indeed, even economic regulation may be subsumed in risk terms 'if the object of regulation is seen as market risk and the purpose of regulation is simply framed in terms of managing risk …'.[533] This chapter does not aim, however, to add to the large body of literature on human behaviour and risk. Nor does it dwell on risk as basis for regulation—

[532] Donald Rumsfeld, US Secretary of Defense (2001-06), press briefing, February 12, 2002, available at <http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636>.

[533] Black (2010: 305) (referring specifically to the 'risk of market failure'. The usage of the term in this chapter is, however, broader).

either from a broad perspective or a narrow one that does not include economic regulation.[534] Its purposes, needless to say, centre on the role of risk in regulatory design.

What is the role of regulatory design when facing the variety of risks (including known-knowns, known-unknowns and unknown-unknowns)? This chapter briefly explores these questions, building on the recognition that if, generally speaking, the ultimate goal of economic regulation may be considered to be the maximisation and protection of investment[535], there is a need to consider the alternatives that lower the risk created by public actions that may undermine the willingness to provide funds. On this, institutional design has a key role to play. A key regulatory task is to determine the cheapest bearers of risk and allocate the latter accordingly.

Consequently, the analytical framework should also provide tools to figure out what the relevant trade-offs are. As analysed in the first section of the chapter, these tools (which comprise the decision-making principles and procedures devised by the State to allocate risk) depend on both the type of risk and the specific regulatory regime. On the one hand, a number of uncertainties in economic regulation, such as unpredictability in demand, legislative or regulatory changes, sunk costs, etc., make the treatment of investment in infrastructure industries more complicated (perhaps more risky) than ordinary business. On the other hand, there are a variety of approaches that can be adopted to deal with investment decisions[537] —most of them implying stricter scrutiny than what is normally applied in non-regulated businesses. Different possibilities for risk allocation will emerge from the interplay between both variables (i.e., risks and regulatory regimes), altering the incentives for different actors.

Amongst the many types of risks, however, financial risk is arguably the most important type of risk (or one of the most) from the economic regulation standpoint. Financial risk underlies many policies adopted by regulators and legislators alike. Hence the second section focuses on that risk and, specifically, on a crucial aspect of regulatory design: State guarantees. These

[534] On the latter, see e.g. Hood *et al.* (2001).
[535] See *supra*, introduction.
[537] See, e.g., the analysis and case studies in Alexander and Harris (2005).

can be defined as the specific assurances given by the State that a firm will receive financial support in case of financial distress, accompanied by the expectation on the part of the firm that it will be rescued. As will be shown, the inclusion of State guarantees into the regime (in any of the varieties describe below) severely alters the incentives for investors, managers and the State.

I. ALLOCATING RISKS

The classic distinction introduced by Frank KNIGHT separates the notion of risk from that of uncertainty.[538] The dividing criterion is measurability: under situations of risk, the probabilities of different outcomes are known (i.e., they are measurable); under uncertainty they are not (i.e., they are unmeasurable).[539] Most real life cases refer therefore to uncertainty rather than risk. Modern literature, however, has criticised the distinction mainly for its lack of practicality or incompleteness. On the one hand, some have suggested that the distinction is unnecessary due to the possibility of assigning subjective probabilities to the outcomes.[540] On the other hand, commentators have pointed out that the states of knowledge are in fact not two, but three radically different states of knowledge, colloquially referred to as 'known-knowns' (risk), 'known-unknowns' (uncertainty), and 'unknown-unknowns' (radical ignorance).[541]Various alternative classifications have been advanced. Nonetheless, for infrastructure industries it suffices to characterise risk generally as an 'unpredictable variation in value'— either the total value of the project or the value of the stakeholder's interest in the project.[542] The former is an aggregation of the latter, that is, the sum of the values accruing of each stakeholder.[543]

[538] Knight (1921).
[539] *Ibid*, at 233.
[540] *E.g.*, Jeffreys (2004).
[541] E.g. Black (2010: 310). See also the quotations at the beginning of this chapter.
[542] Irwin (2007: 5).
[543] For these purposes, the concept of stakeholder is broad: it may include third parties such as workers, insurers, subcontractors, and so forth. The concept of stakeholder is further contextualised and specified in Chapter V.

The decision-making principles and procedures devised by the State to allocate risk depend on two aspects: the specific regulatory regime and the specific type of risk at hand. Both aspects are analysed in this section.

A. Risks and the regulatory regime

Risks are normally shared. It is a function of the regulatory framework to establish in advance, as much as possible, the rules that determine which stakeholder bears what risk (i.e., all or part thereof) and to what extent, taking into account the distribution of rights arising from the framework. The task is subject to the constraints imposed by the presence of behavioural and cognitive obstacles[544], transaction costs[545] and eventually the particular position of the Government within the regulatory relation. For example, consider three common regimes: *pure price cap*, whereby the regulator set a maximum price per unit of volume; a *pure revenue cap*, whereby regulator controls the total amount of revenue that the firm can collect from its customers; and a *hybrid regime*, whereby the firm is entitled to a fixed amount of starting revenue, and a further per unit allowance. As Table 7 indicates, these three regimes are roughly the ones that different sectoral regulators have been applying to UK utilities so far.

[544] See *supra* Chapter I.
[545] Arrow (1971); Klein *et al.* (1996).

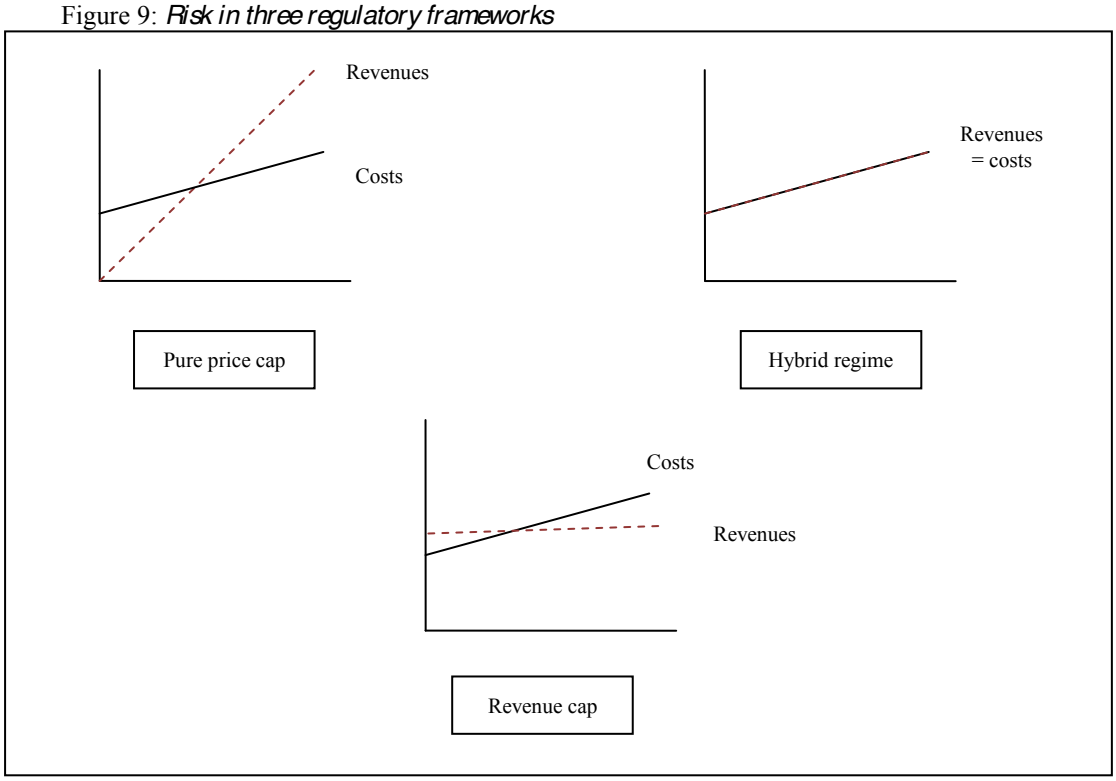
Table 7: *UK regulatory regimes (up to 2010).*

Sector	Regulatory Regime
Airports	(Pure) price cap: amount of income proportional to the number of passengers
Elect. Distribution	50:50 hybrid revenue/price cap: number of units distributed as the output measure in the price cap component
Elect. Transmission	Mostly revenue cap (but with entitlement to additional income when firms connect additional customers to the network)
Gas Distribution	(Pure) revenue cap
Gas Transmission	(Pure) revenue cap
Railways	Hybrid regime: fixed and variable component of revenue calibrated to match the mix of fix and variable costs
Water and Sewerage	(Pure) price cap: amount of income proportional to number of properties (unmetered households) or volume of water delivered (metered consumers)

As shown in Figure 9 below, a pure price cap transfers the highest amount of risk to the firm. This is so because under a pure price cap the regulator adjusts the prices at the end of the period, passing on the benefits of the efficiency gains to consumers through lower prices for the next period. Revenue cap transfers less amount of risk to the firm, whereas the hybrid regime transfers the least risk to the firm and more risk to consumers.[546]

[546] For a complete assessment of the regulation of investment, see Alexander and Harris (2005).

Figure 9: *Risk in three regulatory frameworks*



Source: First Economics (2008)

Beyond the overall regime, however, firms care about risks arising in *each* specific infrastructure project. From an economic/financial perspective, then, the allocation rule should aim to maximise the total value of each project—which means that total welfare considerations take precedence over distributional concerns. To achieve this objective, governments should isolate the election of the allocation rule from political considerations and follow the rule according to which the risk should be borne by the agents who can better manage/diversify it. Or, as IRWIN more accurately states, each risk should be allocated taking into account each party's ability to 'influence the corresponding risk factor' (i.e., minimise moral hazard), 'influence the sensitivity of total project value to the corresponding risk factor' (i.e., avoid adverse selection) and 'absorb the risk'.[547] This, in turn, depends on the specific type of risk.

[547] Irwin (2007: 5 and 56-7). He recognises the need to make trade-off whilst following the three-part principle.

B. Specific risks and the rules for their allocation

The characterisation of risk as an unpredictable variation of value is comprehensive of both good and bad outcomes, and includes not only Knightian risk and uncertainty, but also radical ignorance. This is correct for infrastructure industries, chiefly because investment is commonly made in the face of uncertainty.[548] However, it is still possible to deconstruct the general definition. This one comprises many types of particular risks depending on a number of 'risk factors' or 'variables whose outcome affects total project value and whose value is uncertain'.[549] These variables are normally random, but they can also be partly deterministic and partly stochastic. Considering the risk factors, risks may be subject to almost infinite divisions. However, four types of risks are particularly important in the context of industries subject to economic regulation and deserve more attention (especially in the UK context): policy risk; demand and commercial risk; business risk; and insolvency risk.[550]

A great part of a good regulatory design depends on the possibilities of isolating firms from *policy risk*—that is, the risk arising from government's actions and those of its agents. This definition of policy risk includes both political and regulatory risks.[551] As such, the definition is general and inaccurate, as it does not recognise different levels of responsibility. However, it is useful for the purpose of highlighting the wide variety of heterogeneous events with high levels of uncertainty that may undermine any intent to attract private investment to industries subject to economic regulation.[552] As highlighted by NIE literature[553], the most basic policy risk is the fear for (public) expropriation. Nonetheless, it is usual that provisions embedded in countries' regulatory frameworks and international treaties outlaw this risk to some extent. Also, some large investors, such as multinationals, are well placed to diversify their business and limit their assets at risk.[554] In fact, in many countries financial markets already

[548] And also because, as Posner (2010: 290) says, 'Calculable risk and uncertainty should not be thought of as dichotomous. Often one can say that a decision is more likely, or even much more likely, to be preferable to the alternatives, without being able to quantify the probability'.

[549] Irwin (2007: 51).

[550] This chapter does not deal with other important types of risk that affect all firms more generally, such as exchange risk.

[551] There are a large number of conflicting definitions of political and regulatory risk (see Smith, 1997d).

[552] Hill (1998: 288).

[553] See *supra*, chapter I.

[554] Stansbury (1990: 677-88).

incorporate a risk premia that reflect this policy risk.[555] For this reason, policy risks that fall short of expropriation may imply more damage to investors. Among these risks are the introduction of competition in a formerly monopolistic market (with the consequential risk of stranded assets)[556], opportunistic behaviour in joint ventures (and their rather similar 'public-private partnerships') involving a State party, and the renegotiation of franchising contracts[557]. In all these cases, the level of risk premium will largely depend on the overall protection of property rights within the specific country.

In the UK, regulatory risk is enrooted in the institutional regime, due to the characteristics of the parliamentary system. On the one hand, utilities face higher regulatory risk because constitutional protection against administrative decisions is deemed low.[558] Due to party control over legislature and the executive power, legislation can be amended to 'accommodate' even new judicial trends. Thus, parliamentary sovereignty leads courts to seldom challenge decisions by regulators. This makes the statutory duties which safeguard the regulated firms' returns insufficient. On the other hand, as seen, the judiciary played no role in the creation of the regulatory regime (including the landmark introduction of price-caps). This contrasts, e.g., with the situation in the US, whose regulatory regime *evolved* over a long period of time largely thanks to judicial decisions.[559] The system of regulation in Britain was purposely *designed* as a suitable (as thought then) form of regulation for concentrated markets that (it was expected) would become competitive in the short- or medium-term.[560] As seen, the design included statutes with largely vague duties for

[555] See Kolbe & Tye (1991: 115) (for the US experience); Diamonte *et al.* (1996) (for emerging countries).

[556] As a rule of thumb, if stranded assets—that is, the impossibility of earning an economic return on previously sunk investment as a consequence of changes in the market conditions—are a consequence of regulatory action, the firm should be compensated only when the changes could not have been reasonably anticipated. In the UK, there has been no compensation for stranded assets. That possibility was even ruled out in a recent case (*National Grid*, analysed *infra* in chapter IV).

[557] Guasch (2004) (studying renegotiation of contracts in Latin America and the Caribbean).

[558] In this sense, Spiller (1995: 65) (stating that 'The lack of constitutional protection, then, makes utilities particularly at risk in U.K.-like systems'). *See also* Levy & Spiller (1996: esp. chapters 1 and 3).

[559] In the same sense, Newbery (2006: 2): 'Cost-of-service (or rate-of-return) regulation evolved in the United States over nearly two centuries in response to legal challenges to clarify the nature of property rights and place restraints on regulatory activism'. The US legal evolution is summarised in Breyer *et al.* (2002: 260-7).

[560] Littlechild (1983), famously stated that 'Regulation is essentially a means of preventing the worst excesses of monopoly; it is not a substitute for competition. It is a means of "holding the fort" until competition arrives' (para. 4.11).

regulators, which left considerable room for administrative interpretation.[561] Conversely, vagueness left little space for judicial control.[562] Judges, in turn, have been generally reluctant to get involved in regulatory disputes—although the situation is fast changing.[563] The consequence of both the lack of constitutional protection (which implies more leeway for the legislative power) and the scarce judicial protection (which implies more leeway for regulators) was that political risk has traditionally been deemed intrinsically high in the UK.[564]

A second particular risk arises when *demand* estimates are unreliable. This is a common situation especially in infrastructure projects, due to qualitative and quantitative informational constraints and the number of uncertainties underlying the forecasts. As a consequence, firms are unable to diversify this type of risk, so they commonly seek to insert some cost-sharing mechanism within the framework—either some pass-through arrangements or a direct adjustment of the price in advance or on an *ad-hoc* basis. The plausibility of such demands depends on the regulatory design and regulatory attitudes.

An example where the regime was expressly designed to avoid the firm bearing demand risk may be found in one of the sub-sectors of the rail industry: the underground.[565] In the case of the London underground, in March 1998 a major programme to modernise and restructure London underground system was announced. The regulatory framework was designed as public-private partnerships (PPP) agreements. The details of the regime and its eventual failure are of no concern here.[566] Suffice is to stress, generally, that the government opted for a contractual scheme whereby two consortia—Metronet and Tube Lines—won the temporal

[561] Foster (1992: 212) explains: 'A regulator [in the UK] can use its discretion to revise RPI-X for the sake of efficiency rather than fairness, though as it is his discretion he does not have to use it so'. Given the judicial (adversarial) nature of cost-of-service regulation in the US, conclusions may give more weight to fairness than to efficiency.

[562] Armstrong *et al.* (1994: 170): 'The vaguely defined duties of the [regulator] mean that his decisions are unlikely to be struck down by the courts…'. Indeed, this was an *intended* feature of the original design of the regime after privatisation; see Prosser (1997) and Rawlings (2010).

[563] As will be seen in Chapter IV, there has been a trend towards less deference that has crucially changed the role of the UK judiciary in economic regulation.

[564] This, however, seems to have not excessively affected the risk premia, mainly due to the stability of the regime.

[565] The UK rail industry if formed by four sub-sectors: the main network, the underground, light rails and tramways, and the so-called minor and heritage railways.

[566] The system collapse gradually, and the underground was brought fully into public hands again in 2009. For more antecedents, *see* NAO (2009b and 2004).

(although long-term) right to repair and upgrade the underground's infrastructure in exchange for certain remuneration.

The system of remuneration was based on 'availability payments', whereby the private contracting party was entitled to receive fixed periodic fees contingent on a service quality standard being met. The contracts established a performance-based incentive scheme, along with penalties in case of underperformance vis-à-vis some given baseline levels.[567] In one of the contracts (the one signed with Metronet), the private firm was first entitled to receive four-weekly payments financed by the government through different public bodies.[568] Broadly speaking, Metronet was assigned £600 million per annum under this scheme. In addition, the PPP agreements allowed for £360 million of contingency to cover extra costs. These costs were expected, as the condition of many assets transferred to Metronet was unknown. As a result of the regime, then, Metronet's remuneration essentially paid for the upfront investment. The revenue stream was not connected to consumer fees and therefore there was little or no risk for the firm. This one was therefore able to make a normal profit on investment *regardless* of the demand. In practice, the risk was transferred to customers or taxpayers.

However, in this case there was at most a very weak case for effecting such risk allocation. Since the majority of costs in the rail industry are fixed, irrespective of volumes (and the underground is no exception to this rule), demand risk was negligible. Had the government checked for consistency of its chosen regulatory design, it should have found that it did not need to compensate private firms for bearing such risk—as in practice it did, due to the remuneration regime inserted into the PPP contracts.

[567] The companies were paid £3 for every 'passenger hour of benefit' they created. That term was measured by improvements in journey times compared to the historic base line combined with the number of passengers who experienced the savings.

[568] These payments (so-called 'Infrastructure Service Charge') varied according to four criteria of performance: (1) *availability*, which measured the reliability of the network under Metronet's control, i.e. the delay attributable to Metronet. This was the main indicator of performance. (2) *Capability*, which measured the capacity of that network, based on the journey times for passengers for a given line or part of a line. (3) *Ambience*, which measured customer experience of the trains, platforms and station facilities, i.e. a reflection of their cleanliness and general condition. And (4) *service points*, which measured delivery against a number of varied obligations (e.g. speed with which service faults were rectified).

The third common risk faced by investors is *business risk*, sometimes known as *construction and maintenance risk*. This kind of risk arises when the costs of building and/or maintaining infrastructure differ from projections. Amongst the costs included are operating expenditure (opex), capital maintenance/renewals expenditure (capex) and enhancement expenditure.[569] Business risk is associated with the unique circumstances of a particular firm and the overall economic environment. It arises from the operating characteristics of an investment in real assets. Therefore, it mostly represents the possibility that the firm will not have the cash flow necessary to meet its operating expenses. Some elements that are included in the business risk are revenue uncertainty, labour costs uncertainty, regulatory and political uncertainty, and risk of technological obsolescence, amongst others. The distinctive feature of all these elements is that they are independent of the means by which the assets are financed.

In principle, the firm normally controls business risks.[570] This is certainly true in the case of renewals and enhancement expenditures.[571] However, there are two potential exceptions. First, in utilities there are a number of factors that make the risk asymmetric – i.e. the possibility of affecting capex costs is bigger than in other sectors (even other regulated sectors), affecting the risk profile of the firm. This exception arises, for example, when a significant part of the network must be shut down in order to carry out works. If there are delays and disruption extended for longer than expected, the firm might face penalties and pay significant amounts in compensation.[572] The second exception arises in situations in which there is a severe constraint in the forecast of the costs due to lack of information about the quantity and/or quality of the assets. Both exceptions may give rise to demands for embedding risk-sharing mechanisms within the regulatory framework.

---

[569] Normally, in the rail industry a distinction is made between capital maintenance and renewals. In other regulated sectors both types are included in 'capex'.

[570] In terms of the CAPM (see Appendix I to Chapter III), equity beta should not increase with this type of risks and the normal calculation of the (already risk-adjusted) cost of capital should not be adjusted. Least cost-sharing mechanisms should be allowed. There are exceptions, though, whereby consumers, not the firm, bear the risk. For instance, in the rail industry there may be pass-through arrangements for electricity purchase costs, given the proportionally superior amount of energy consumed and the volatility of energy prices (First Economics, 2008).

[571] For instance, in the case of Railtrack (and certainly now its successor Network Rail) under-spend in renewals and efficient over-spending in renewals was incorporated to the RAB after some time.

[572] Note that the cause of the disruption for longer than expected is not important. It might be a consequence of poor planning or simply unexpected circumstances. What is important is that even short overruns may have severe impact on costs. This is a feature not usually encountered in other sectors.

The final type of risk relevant in the context of industries subject to economic regulation is *insolvency (or financial) risk*. It represents the possibility that the firm will not have the cash flow necessary to pay its equity holders when the assets are partially funded by borrowing. That is, it is the risk that arises from the inability of the firm to pay its debt—a situation generally known as financial distress. Financial risk, hence, depends on the means by which the assets of the firm are financed. Amongst the elements included in the financial risk are: the risk of bankruptcy, the risk of restructuring in event of default, and the risk of suboptimal operating and inefficient investment decisions when the firm is in a condition of financial distress. Normally, regulatory frameworks allocate the risk of financial distress to investors rather than consumers and taxpayers. However, on occasions such risk-allocation is changed by the State, which assures a particular outcome; i.e., that the firm will receive financial support to cover its deficit and will not be left to fail, regardless of efficiency considerations.[573] This situation is further considered in the next section.

II. AVOIDING GUARANTEES AND UNINTENDED EFFECTS

To a great extent, the allocation of risk between the constituents of the regulatory relation is influenced by the actions actually taken by governments to deal with firms in situations of financial distress. As seen, normally regulatory frameworks allocate the risk of financial distress to investors rather than consumers and taxpayers. However, such initial risk-allocation may be changed by the State by giving *assurance* that the firm will receive financial support to cover its deficit and will not be left to fail, regardless of efficiency considerations. When the assurance is accompanied by an *expectation* on the part of the firm that it will be rescued from financial distress, the State is providing a financial guarantee that is not normally allowed for within regulatory frameworks. As a consequence, the government (and therefore taxpayers or customers) assumes risks that would otherwise be faced by shareholders and/or creditors. As a consequence of the flawed risk-allocation, incentives provided to the firms and other interested parties are severely distorted. Both moral hazard and adverse selection give rise to suboptimal consequences. Therefore, regulators should

[573] The next part analyses in detail the consequences of this situation.

always control for consistency of their chosen actions by checking that no guarantees have been (nor shall be) provided to firms.

In this work, State guarantees are considered from a wide perspective—'wide' with respect to both their theoretical basis and their scope of application. Regarding the theoretical basis, there are at least three strands in literature that refer, either directly or indirectly, to State guarantees. First, the late 90s and early 2000s saw a growing number of studies, mostly linked to international donor organisations, which analysed how governments explicitly assumed risks well in excess of their capacities of control with the aim of attracting private investment.[574] Second, the so-called 'soft budget constraint (SBC) syndrome', first acknowledged by KORNAI, began to be used to analyse a number of diverse economic problems.[575] Despite the fact that the idea was originally applied to the study of socialist economies, the same theoretical framework was later extended to capitalist economies as well and used for analysing a wide variety of topics—including aspects related to economic regulation.[576] Finally, a number of studies, particularly from the socio-legal perspective, analysed 'counterproductive regulation' and 'paradoxes' of the so-called regulatory state.[577] It is submitted that, when applied to economic regulation, these three bodies of literature refer to a similar phenomenon and hence can be analysed within the common framework of State guarantees, in the manner explained in this section.

Regarding the scope of application, the concept of State guarantee is widely applicable to both developing and developed countries. Heretofore, most studies (at least those that analyse SBC and explicit guarantees) have focused on emerging economies. Since most industrial countries (but also others) have been able to provide credible levels of commitments by the creation of adequate institutions, it has been assumed that guarantees for them are not needed. However, even though the phenomenon is especially pervasive in developing economies,

[574] *See, e.g.*, the compilation of articles in Irwin *et al*. (1997).

[575] Kornai (1980) first used the concept as an explanation of economic behaviour in socialist economies marked by shortage and in transit to capitalism. Dewatripont & Maskin (1995) further developed the model. Since Shaffer (1989), the SBC problem is normally modelled as a dynamic commitment problem.

[576] For an application of the SBC problem in regulation and privatisation see, *e.g.*, Martimort and Straub (2006); for an application to monopoly, see Segal (1998).

[577] *See, e.g.*, Sunstein (1990b) (defining paradoxes as self-defeating regulatory strategies— strategies that achieve an end opposite to the one intended, or to the only public-regarding justification that can be brought toward their support); and Grabosky (1995b) (referring to 'ways by which regulatory initiatives may defeat themselves or may otherwise inflict collateral damage).

there seems to be no strong reason to think that it cannot appear in different economic settings. Particularly when a broad definition is used (as in this chapter), guarantees are far more common than expected in a number of regulated industries in developed economies. The UK, for example, has been traditionally considered as an example of a country that does not bear risks except when they relate specifically to a project.[578] Nevertheless, the evidence discussed further below provides some indication of the use of State Guarantees.[579]

State guarantees can be analysed from an *ex ante* perspective (i.e., before the financial distress of the firm has occurred) and from an *ex post* perspective (i.e., after the situation of financial distress). From an *ex ante* perspective, they consist of the promise of financial aid for distressed firms. More precisely, they consist of the explicit or implicit assurance of financial aid matched with the investors' expectation that the aid will be provided if it is required. From an *ex post* perspective, the guarantee consists of the actual provision of financial aid through any action that averts financial failure of the firm in distress. The remain of this section first explains both perspectives; then proposes a simple typology of guarantees in order to facilitate the process of control of consistency; and finally analyses the reasons for providing guarantees and possible ways to avoid them.

A. Promising financial aid: the ex ante perspective

From an *ex ante* perspective, a State guarantee has two core elements. The first is the governmental *assurance* of rescue from financial distress. The second is the reaction to the assurance: for the State guarantee to exist, it is necessary that investors and/or managers form an *expectation* that the firm will be rescued from financial distress.

---

[578] *E.g.* Irwin *et al.*(1997: 6).

[579] Perhaps the increasing adoption of guarantees is due to the increase in the fear for distress during the last decade. *See, e.g.,* OFGEM (2003: 1): 'In the past, the possibility of a monopoly network operator becoming insolvent has been regarded as remote. It is much less clear that this remains the case today, in spite of the introduction of financial ring-fencing conditions. The growing trends of combining network businesses with other, more risky activities in the same corporate group, and of highly leveraged financing structures, increase the risk of financial failure. Moreover, the consequences of failure would be severe if the company was not able immediately to continue to trade. Security of supply and public safety would be immediately and progressively threatened'.

The assurance may be given in two different ways. First, the State may guarantee the revenues of the firms by explicitly reassuring the payments from the outset—that is, promising financial aid before the event of financial distress manifests. Second, the State may implicitly assure the protection against situations of financial distress. This implicit assurance will occur when, despite its initial explicit promise not to provide financial aid to firms in distress or its silence on the issue, the government is incapable of make credible commitments given its own previous behaviour.[580] In a nutshell, there are two possible *ex ante* attitudes: either the government manifests its interest in providing the guarantee (explicit assurance) or it makes no affirmative manifestation (implicit assurance). In both cases, investors' expectations match the governmental assurance.

The motivations of the government for providing assurance of future financial aid vary. There are a number of structural conditions affecting those motivations, which can be political, social or economic. These conditions create a distinctive incentive structure to which governments normally respond when making decisions for regulated industries. To appreciate this, consider first the incentives that arise in well-functioned markets. In these markets, firms are normally permitted to fail. As a consequence, they incorporate such an event as a variable to make decisions in their ordinary business. The answer to the risk of financial distress becomes how to prevent the inefficient liquidation of cash-poor firms.[581] In contrast, in markets subject to higher levels of governmental intervention, two effects may happen. On the one hand, different types of efficiency must be traded off; on the other, efficiency may simply be rendered as secondary.

An example of the trade off between different types of efficiencies can be found in the Competition Commission (CC) decision on the merger between South East Water Limited and Mid Kent Water Limited in 2007.[582] In that case the CC assessed, for the first time, a merger between two water companies under the provisions that give it powers to decide whether a merger prejudices, or may be expected to prejudice, the ability of the regulator to

[580] Note that remaining silent is equivalent to making an *ex ante* explicit commitment not to provide a guarantee.

[581] See Aghion *et al.* (1992).

[582] Competition Commission, *South East Water Limited and Mid Kent Water Limited – A report on the completed water merger*, 1 May 2007.

make comparisons between different water enterprises—that is, the ability to apply yardstick competition. This system, which aims to emulate rivalry in a competitive market, is crucial in the regulation of the UK water sector.[583] Despite concluding that the prejudice was likely (although limited) in the actual case, the CC allowed the merger subject to the imposition of a remedy. The remedy was a price reduction on the terms set in the own judgement. Importantly, the CC considered the benefits for the customer to be substantially more important than the prejudice. Efficiency considerations of the regulatory system were set aside and in favour of consumer welfare—*i.e.*, allocative efficiency took priority over dynamic considerations.[584]

On the other hand, it is not difficult to envisage cases where efficiency is rendered (at least in a particular case) as a secondary goal. For example, applying the same example, the CC opted for the maintenance of the number of competitors in order to avoid the decrease in the possibilities of applying yardstick competition. On balance, a number of considerations can be deemed more relevant than efficiency and take precedence over it.[585] The reason lies in the multiplicity of (sometimes incompatible) objectives governments have. This plurality of goals is a feature that manifests both within a certain period of time (*i.e.*, a number of objectives are pursued by the government in a given sector) and within an inter-temporal horizon (*i.e.*, new governments have or emphasise different objectives than previous ones).[586] Even under 'normal' conditions of the market some objectives may take precedence over efficiency as a result of the necessary trade-offs. Under conditions of financial distress, however, the need to set efficiency aside becomes stronger. For a number of reasons that vary

[583] As the Competition Commission has explained, comparative competition enables the regulator 'to take account of the objective differences in the operating environments of the companies before making comparisons between them. These comparisons enable [the regulator] to come to an informed assessment of how each company's performance compares with that of the most efficient companies. On this basis, [the regulator] sets a scoring system for levels of customer service and efficiency assumptions that are incorporated into price controls specific to each company which require the company to become more efficient. Companies have the incentive, once these price limits have been set, to achieve additional cost savings by operating even more efficiently. These savings are then reflected in efficiency assumptions set in subsequent periodic reviews. The benefits achieved are thus both local and national, since a well-performing company will help to set future targets for the others. The process of comparative competition therefore maintains downward pressure on prices for all customers' (*Ibid.*para. 2.17).

[584] As mentioned in the Introduction, in the water sector the Cave report (Cave, 2009) recently gave precedence to dynamic considerations, proposing a number of changes to the merger regime.

[585] Amongst them: the maintenance of the number of firms in order to avoid the decrease in the possibilities of applying yardstick competition (e.g. water sector); avoiding systemic effects in the market (e.g. financial sector); or keeping security and/or continuity of supply (e.g. energy and rail sectors).

[586] On plurality of regulatory objectives, see Chapter V.

according to the sector, governments might opt to allow some level of inefficiency in the market. The most well-known consideration is the avoidance of systemic effects that commonly underpin bailouts in financial markets.[587] Nonetheless, even in absence of spillover effects, there are still a quite large number of considerations that may potentially justify a guarantee. Under the threat of default, these concerns may lead governments to save a firm in first place, regardless of efficiency.[588]

In infrastructure markets, many (if not most) of these 'extra-efficiency' considerations are linked to reputational effects.[589] These effects arise from the hierarchical side of the regulatory relation, whereby failures by the firm may be publicly perceived as government failures. Under the traditional P-A approach[590], the government may attempt to overcome the problems of hierarchy by designing a suitable contract. An ideal incentive system would overcome information asymmetries and other market failures bringing the relation closer to an 'invisible hand'. Individual self-interest behaviour of the firm would be harnessed and the welfare goals of the regulator would be achieved. However, it can be shown that mechanism design has limits and regulation is by definition inefficient. If inefficiencies translate into failures perceived by the public, such as disruptions in the service or others, the outcome may be regarded as a failure to exercise proper planning or control. In turn, this loss of reputation may rapidly become social/political unrest. In principle, the guarantee allows governments to avoid the problems generated by informational constraints. It partially avoids reputational losses by deviating public perception from qualitative to quantitative considerations (*i.e.*, from failure to the monetary value of the rescue package).

Note that when providing explicit assurances, the State demonstrates an overt interest in providing financial aid in the future. Its motivation may or may not be clear and unambiguous at present, but it is clear that there is a reason underpinning the promise of financial aid. That is not necessarily the case under implicit assurances. However, that does not mean they are unmotivated. The government may remain silent, but still act under a non-declared

[587] Systemic risks are those affecting the entire market as a consequence of the cascading failure produced by a single unit due to the interdependence of all the entities in the system.

[588] Similarly, in well-functioned markets there might still be incentives to ask and provide bailouts. The presence of sunk (or fixed) costs and asset specificity are normal motivations to bailout a firm.

[589] Other motivations are the response to political pressures, capture, and the extreme case of bribery.

[590] For more details on this explanation, see the discussion in Chapter I.

motivation. Conversely, the government may even declare that it has no intention whatsoever to provide financial aid, but still creates expectations for investors. As will be explained, the distinction between explicit and implicit assurances mainly illustrates that governmental motivations are certainly important, but they are neither necessary nor sufficient for State guarantees. From an *ex ante* perspective, there are only two elements: the explicit or implicit governmental assurance of financial aid, and the creation of some expectation amongst investors that aid will be received if it is required.

B. Providing financial aid: the ex post perspective

Regardless its *ex ante* attitude, *ex post* the government may or may not provide financial aid. When provided, the aid manifests in any action that averts financial failure. The most common response is financial. There is a wide range of methods that can be used to ensure the firm continues existing. Nonetheless, bailouts are the most archetypical manifestation of the guarantee. A bailout always places the risk of default on either customers or taxpayers. When the bailout is made through an (unscheduled) adjustment in the price of the service provided by the firm, customers bear the risk. When the bailout is made providing the firm with a (unscheduled) payment (generally cash), or reducing the amount of the liabilities that the Government is entitled to obtain from the firm, the risk is borne by taxpayers.[591] The main consequence of this risk allocation is that bailouts always involve a transfer of wealth from the less to more financially-constrained agents. This is an *ex ante* unintended effect of any regulatory policy. It is at least dubious that customers and taxpayers are the agents better prepared to diversify the risk of default and/or distress.

Considered from a purely *ex post* perspective, conferring financial aid has numerous intrinsic inconveniences. Moreover, they are mostly irresolvable. One main difficulty is the need to define when it should be considered that a firm is in the situation of financial distress. Unlike the event of default (that triggers bankruptcy), distress is not a univocal concept. Informational constraints make the definition highly infeasible—cash flows cannot be

[591] The difference between both scenarios depends on the identification of consumers and taxpayers, which is both country- and sector-specific.

verified with certainty, least by regulators. This certainly occurs when both parties of the contract are private, but the situation is arguably enlarged by asymmetries of information when one of the parties is a public authority, such as a regulator, in a public-private partnership contract. Thus, it is necessary to rely upon verifiable 'signs of financial distress'. Theoretically, this might be done by directly identifying the event(s) that triggers a bailout, or through the creation of a mechanism that allows to objectively determine *ex ante* whether a firm is in distress. However, the same informational constraint that would make a definition of distress necessary, would turn this definition into inexact and to some extent arbitrary.

In absence of an objective *ex ante* definition of financial distress, there are a number of options. Having no definition at all should certainly be discarded. Such a policy would render even more inefficient solutions and increase the level of legal uncertainty. One option is to allow the decision of declaring the situation of distress up to the own firm. In this case, however, the incentive is perverse. Due to informational constraints, the incentive is for all firms to declare themselves in distress as soon as the minimum probability of economic problems is presented. As an alternative option, it has been proposed to pay for the debt of a fraction ψ of all companies, chosen at random.[592] Nonetheless, the inconveniences of this alternative are multiple. On the one hand, in many utilities the possibilities of distress are arguably less likely to occur for several firms at the same time. In fact, given the characteristic of natural monopoly present at least in some segments of these markets, distress can be expected to be an isolated event. In that case, the practicability of these options largely diminishes. On the other hand, even if the solution were feasible, the bailout would allocate a significant amount of the budget to firms that are not cash-constrained, hence turning the bailout into a less efficient option. Significantly, by definition there is no possibility to assess *ex ante* the likelihood and magnitude of the inefficient outcome.[593]

[592] Hotchkiss *et al.* (2008).

[593] The uncertainty surrounding the definition of a bailout may be contrasted with the situation under a normal bankruptcy regime. Bankruptcy is far more predictable than bailouts. It depends upon specific parameters possible to be measured on an *ex ante* basis. The probability of bankruptcy can be (relatively easily) calculated as a function of the capital structure (specifically, the level of gearing) and the uncertainty in the value of the firm. Insolvency risk is mainly project specific. Whilst an unpredictable variation in the value of the firm has many causes, both project-specific and economy-wide; the choice of gearing is project-specific. See Irwin (2007).

A further difficulty is that the extension of the bailout is somewhat arbitrary unless there is a clear *ex ante* definition; meaning that the extension of financial help is largely unpredictable. Indeed, to be viable it must cover at least a large fraction of the debt of the firm in distress. Therefore, its extension is a function of the level of gearing of the firm—which in infrastructure sectors tends to be rather high. For this reason, in absence of clear commitment not to bailout firms in distress (i.e., there is no possibility for customers/taxpayers to avoid insolvency risk), some rationality for controlling the capital structure of regulated firms may arise. The extent of the control is, however, debatable.[594]

C. A typology of State guarantees

Combining both *ex ante* and *ex post* perspectives, it is possible to advance a typology of State guarantees that may help designers to carry the check of the solution that resulted from the application of the synthesised approach.

As seen, there are two possible *ex post* outcomes of the explicit assurance. First, the government may keep its promise and provide the guarantee; secondly, it may break its commitment and provide no guarantee at all. If the guarantee exists, I will call it 'intended'. When the State Guarantee is intended, the government agrees in advance to make a payment if there is a deterioration of the firm's financial conditions, firms expect that payment and the payment is eventually made after the situation financial distress. This means there is a perfect alignment of *ex ante* and *ex post* governmental objectives and investors' expectations. One clear example of the intended State Guarantee in the UK is the explicit assurance given to the creditors of Metronet (as seen, one of the firms vested with the task of modernising a great part of London underground's infrastructure) and the subsequent payment of the debt by the Government.[595] Suffice is to say here that from the outset of the contract, the Government guaranteed 95% of the senior debt acquired by Metronet to cover the event of distress and ended up paying approximately £1,600 million after Metronet's collapse made the guarantee

[594] See Irwin (2007: 83-4), and *supra* chapter III.
[595] See *supra* note 566.

enforceable.[596] Other common examples of intended guarantees includes subsidies, minimum revenues, debt guarantees, State insurances and others.

Intended guarantees mainly arise because of the awareness of governments of the impossibility of writing complete contracts. There are at least two (compatible) views of this phenomenon. On the one hand, from a contractual perspective, incompleteness is due to the lack of adequate 'presentiation' (*i.e.*, the perception of future foreseen effects into the present) first identified by MACNEIL.[597] Thus, 'anything preventing either binding of the future or recognition of that binding frustrates presentation'.[598] In regulated settings, flexibility is the foremost factor affecting presentiation. For multiple reasons, regulatory regimes are frequently open to choices and the development of new outcomes. Flexibility, in turn, prevents predictability and hence contract completeness. On the other hand, from a transaction-cost point of view, presentiation may be too costly. In any event, the more market-oriented the contract is, the higher the lack of response to future consequences of the parties' acts. Awareness of this limitation may at least partially explain the need for the State to give reassurances on future outcomes to private firms.[599]

Early literature on State guarantees focused exclusively on the intended meaning. Whilst advocating more suitable ways to allocate risks, most commentators restricted the use of the term to explicit reassurances. For example, SMITH defines guarantee in the following terms:

> [A] contractual arrangement under which a third party (the guarantor) agrees to fulfil the financial or other obligations of the guaranteed party (the principal obligor) to another party

[596] Whilst designing the contracts that would be the basis for the programme of modernisation and restructuring of the London underground system, the government decided to guarantee 95% of the senior debt acquired by Metronet to cover the event of distress. Considering that Metronet's debt at the outset represented an 88% of its capital structure, it was likely that the intended guarantee would be of a considerable amount. In 2007 the firm went into administration due to their inability to meet their financial obligations. After Metronet's collapse made the guarantee enforceable, the government paid approximately £1,600 million to settle the debt obligations net of calculated benefits (NAO, 2009). See more details in chapter IV below.

[597] Macneil (1974a). As the same author explains in a different work (1978:863), presentiation 'it is a recognition that the course of the future is so unalterably bound to present conditions that the future has been brought effectively into the present so that it may be dealt with just as if it were in fact the present. Thus the presentiation of a transaction involves restricting its expected future effects to those defined in the present, *i.e.*, at the inception of the transaction'. Enhancement presentiation in order to limit freedom for exercise of choice is a norm of all types of contracts. *See also* Macneil (1974b).

[598] Macneil (1974a: 591).

[599] This explanation is different from the immediate (or actual) explanation for the guarantee.

> (the beneficiary) in the case of default by the principal obligor. In private infrastructure arrangements, the principal obligor will typically be a government entity that has given undertakings to the private investor.[600]

Note that the agreement to provide financial support reveals the intended nature of the guarantor, who gives the assurance to a third party. According to this definition, then, the government cannot merely commit to follow some course of action, but only to guarantee the financial performance of a legally independent entity. SMITH argues that the use of a wider approach 'obscures important differences between cases in which the government is expressing a primary obligation…and when it is doing so in support of the primary obligations of another'.[601] However, as it is discussed below, government self-commitment to bear risk must be considered a guarantee. Some authors agree. They interpret the term guarantee in a broader sense, as the government assumption of a risk, including the self-commitment to bear that risk.[602] That is, they assume that governments make a conscious decision to bear risks that should in principle be borne by private investors. Nevertheless, the definition is still restricted to the explicit meaning of the State Guarantee.

The excusive focus on intended State guarantees is too formalistic. It does not cover all the possible situations of mistaken risk-allocation between the government and investors.[603] Recall that the government may also provide implicit *ex ante* assurances. Unlike the previous situation, the State is in principle unwilling to provide financial support.[604] As before, implicit guarantees have the same two possible *ex post* outcomes. Either the government renege on its promise and explicitly provide financial aid, or give no financial aid at all. In the first case the government shows its intention to reassure the payments once the financial conditions have

[600] Smith (1997d: 71). He goes on to argue that 'to be enforceable as a guarantee, the government's undertaking must evidence an *unambiguous* commitment to meet the obligations of the principal obligor in the case of default. … In some cases the compromise is a comfort letter, which expresses the government's support for the venture, entity, and perhaps contract in question, but falls short of an unequivocal guarantee' (*Ibid.*, emphasis added).

[601] *Ibid.*

[602] Irwin *et al.* (1997).

[603] Indeed, dealing with explicit guarantees with adequate policies should be a priority. A first and foremost measure to avoid intended guarantees (and also other types) is to provide an explicit reassurance that distressed firms will not be rescued (Irwin, 2007). As will be shown, however, such a declaration is neither commonly found in regulatory frameworks nor normally supported by other institutional measures to enhance its credibility.

[604] Recall that the lack of *ex ante* assurance may have consisted of an initial promise not to provide a guarantee, but this is not essential: the key point is the lack of affirmative manifestation.

changed and the firm has become unprofitable.[605] There is a State guarantee that I will call 'unintended'. The unintended guarantee reflects a mismatch between the *ex ante* governmental promise (implicit assurance) and its own *ex post* behaviour. The literature on soft-budget constraints has made this type of guarantees apparent. The reason is that the government's behaviour produces an anticipation of a given outcome that changes managers' and investors' perception of risk, discouraging financial discipline. From the regulatory design point of view, the *ex ante* perspective is more important than the *ex post*: if the assurance creates expectations in investors, there is a guarantee.

In the UK, an example of the unintended guarantee is provided by the case of Railtrack—the company that used to manage the British rail infrastructure until its collapse in 2001.[606] After the firm's collapse, the government made an explicit promise not to bailout shareholders. However, shareholders pressured for compensation. Primarily, they threatened to take legal actions, claiming illegalities in the procedure that put Railtrack under administration. Crucially, pressures resulted in the government stepping back from its original intent. Along with taking over Railtrack's debt, the firm was sold to a 'not for dividend' company named Network Rail. In the sale and purchase agreement, Network Rail agreed the acquisition of the entire issued shares of Railtrack at £500 million. From that amount, the Government provided £300 million in the form of a grant, with the remaining £200 million funded by the issuance of new debt by the purchaser (debt that in turn was guaranteed by the Government). The governmental payment was justified in the need 'to reflect economic benefits accruing in the public interest from an earlier end to administration than would otherwise have been the case'. The payment ultimately constituted a surreptitious, but effective bailout of shareholders.

The last variant of the State guarantees is arguably the most contentious. I will call it the 'hidden' guarantee. As with the unintended guarantee, in this situation the government makes no *ex ante* declaration of its intention to provide future financial aid—that is, there is only an implicit assurance. Likewise, there is no explicit *ex post* provision of financial aid.

[605] The ownership of the firm is not relevant for these purposes. It can be either private or State-owned.

[606] A further example analysed in chapter II is the decision adopted by the UK government to privatise some utilities firms with underpriced shares.

Nevertheless, a number of financial provisions embedded in the regulatory regime (perhaps deemed optimal at the outset) distort the incentives of investors and managers and result in a misallocation of risk, producing inefficient outcomes. Unintentionally, the government is affecting the allocation of risks between investors and itself.

Indeed, it is highly controversial to sustain that this situation constitutes a guarantee. Its inclusion crucially depends on the objective of the classification and the relevant characteristics of the guarantees. On the one hand, recall that the *ex ante* perspective is far more relevant from the design point of view. The *ex post* analysis is indeed important, but consequential. When designing a regulatory regime, regulators should control for consistency of their options assessing even the future outcomes of the distribution of risks between all stakeholders. This is done from an *ex ante* perspective (i.e. before designing the regime), but crucially taking into account the possible *ex post* consequences. On the other hand, even though subjectively there is no explicit governmental motivation (indeed, there is no intention to help the firm in distress), nor even actual financial aid, it is submitted that the motivational element is secondary for the purposes of regulatory design. The designer only needs to check for the *consequences* of the policy, not for governmental intentions. If a policy produces an assurance unintentionally and that assurance meets the investors' expectations, there is a guarantee. The checking the designer must conduct needs to cover unintended effects because they may configure State guarantees.[607] Indeed, this situation makes the check unavoidably inefficient. But, once again, first best solutions are unachievable under conditions of contractual incompleteness.

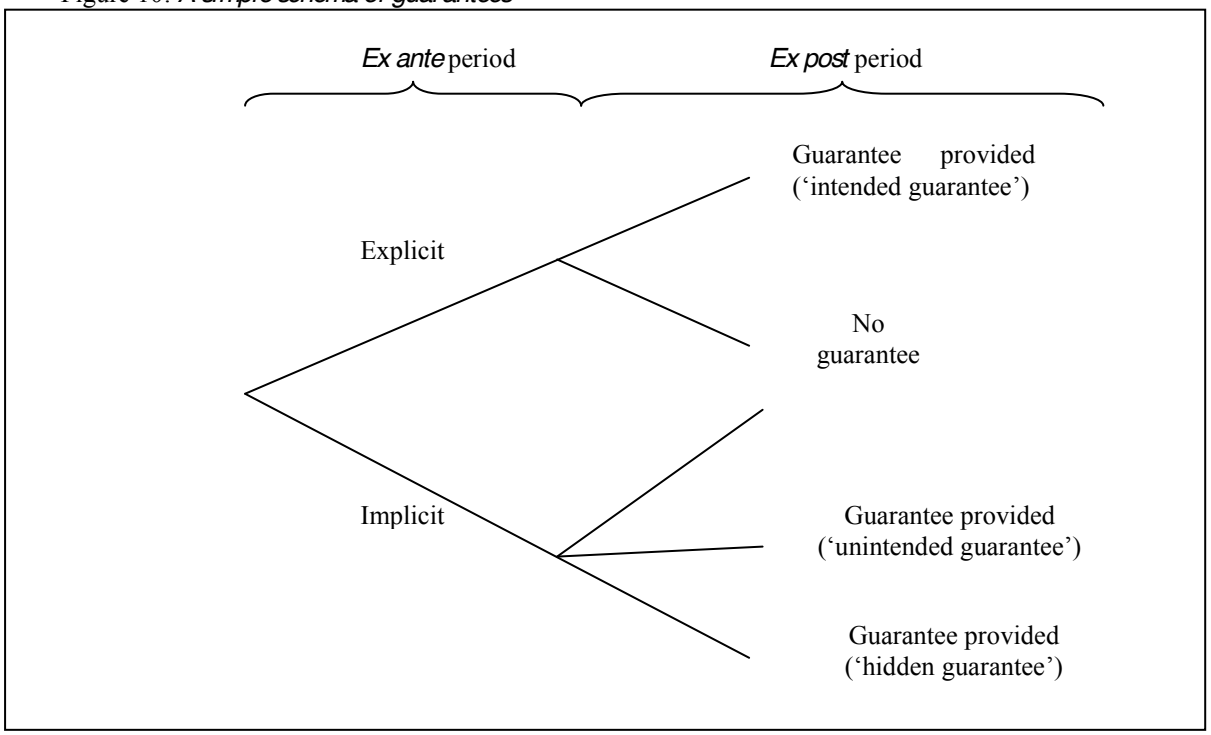
The hidden State guarantee reflects a different source of contractual incompleteness. In this situation, there is no awareness of the consequences of present actions and an impossibility to define *ex ante* future contingencies. Lack of presentiation becomes almost absolute lack of representation—but attributable to faults in the design. Thus, the guarantee is related to the unintended effects of a given regulatory policy. This signifies that, unlike the other two types of State guarantees, the outcome of the hidden guarantee extends beyond mere financial aid

[607] This claim follows only partially from the literature (and will be qualified below), but as will be seen, it is fully supported by the regulatory practice.

and therefore presents more challenging concerns for regulatory design: how to minimise some effects that are not even foreseen?

Figure 10 summarises the previous ideas presenting a simple schema of State guarantees.

Figure 10: *A simple schema of guarantees*



D. Consequences

The outcomes of State guarantees vary. Importantly, they are not necessarily negative from a welfare perspective. Somewhat counter-intuitively, State guarantees are not intrinsically inefficient or a self-defeating strategy.[608] On occasion, governments may engage in these kinds of policies and produce positive welfare effects. However, negative effects seem to be most common and thus it is important to balance positive and negatives effects when checking for the consistency of the regime. Certainly, it is difficult to value costs and especially benefits, and indeed is not easy to do this *ceteris paribus*. Despite this

[608] Sunstein, (1990a: 412): 'whether a strategy is self-defeating depends on how its purposes are described'.

methodological problem, the point here is that the designer must take into account both welfare-enhancing and damaging effects. If negative effects dominate, the central task is to determine *ex ante* what kind of policies may constitute a guarantee, so as to avoid embedding them in the regulatory regime in first place or remove them from the framework and avoid them in the future.

The assessment depends on the benchmark. An important limitation is that (Pareto) efficiency cannot be used as the standard to assess the existence of a State guarantee. If regulation is by definition suboptimal, second-best results are the only possible outcome of all types of regulation. That is, the focus on efficiency would mean that regulation would always constitute a hidden guarantee by itself. If that were the case, there would be no reason for defining a hidden guarantee in order to explain the misallocation of risk. On the contrary, when a benchmark different from efficiency is used, the hidden guarantee does not equate regulation, because expectations and incentives arising in one case and another are essentially different. In this sense, most other goals are fit for purpose.

Arguably a good benchmark (one at least suitable for the purposes of checking the consistency of the selected mode of governance in infrastructure industries), is to consider the impact of the State guarantee on investment. Provided that the attraction of (mostly private) investment is one of the key goals of economic regulation, this goal is an apt yardstick. From this perspective, guarantees are mostly negative, because of two interrelated effects. First, allegedly most industries subject to economic regulation are (or deemed themselves) 'too big to fail'. This means that governments seem far too keen to bail them out, hence given an implicit guarantee to investment. The effect has lately been witnessed mainly in financial markets, but its conditions seem easily replicable in infrastructure.[609] The fear of service disruption provides the regulator with an incentive to renege on its promise not to rescue the firm. Second, there is what may be called the 'systemic effect' of the guarantee. When this guarantee is given for a single sector, an expectation may be created that all the others will receive similar treatment. Note that the negativity of these two effects for investment attraction does not follow automatically. It may seem that providing a guarantee should

[609] The catchphrase came into fashion in the financial sector during the 2008-09 crisis. However, it also applies to many utilities.

attract more investment (because investors perceive a less risky environment), so the guarantee would be positive according to this standard. However, State guarantees do produce an important detrimental effect that becomes apparent when the incentives of the actors of the regulatory process are analysed.

The guarantee severely alters the incentives of the actors. Rent-seeking behaviour is arguably the most well-known consequence of a guarantee.[610] In particular, rent-seeking behaviour produces a number of moral hazard and adverse selection consequences suboptimal for both investors and managers. In its most extreme version, rent-seeking might lead to some form of regulatory capture. In practice, however, the real possibility of the market participants to influence the initial risk allocation is a function of the regulatory regime in place. Generally speaking, the more market-oriented the governance structure in place, the fewer the possibilities to try to renegotiate the initial conditions and obtain a financial guarantee.

On the one hand, the lack of commitment not to bailout firms in distress produces an unintended risk allocation that influences private investors' *ex ante* incentives. A bailout always puts the risk of default on either customers or taxpayers. When the bailout is made through an (unscheduled) adjustment in the price of the service provided by the firm, customers bear the risk. When the bailout is made providing the firm with a (unscheduled) payment (generally cash), or reducing the amount of the liabilities that the Government is entitled to obtain from the firm, the risk is borne by taxpayers. The difference between both scenarios depends on the identification of consumers and taxpayers, which is a feature both country- and sector-specific. The main consequence of this risk allocation is that, unlike bankruptcy procedures, bailouts involve a transfer of wealth from the less financially-constrained agents to those who are more so. This is an unintended effect of any regulatory policy. It is at least dubious that customers and taxpayers are the agents better prepared to diversify the risk of default and/or distress.[611] That is one of the main reasons why regulatory frameworks allocate from the outset these kinds of risk to investors rather than the public.

[610] *See* the literature cited *supra* in note 94.

[611] As seen, isolating the election of the alternative from political considerations, governments should follow the rule according to which the risk should be borne by the agents who can better diversify it. Considering all the *ex-post* options, the threat of bankruptcy should stand out—exactly as it occurs in non-regulated markets.

Since in a bailout the true allocation of risk differs from the anticipated allocation, investors have an *ex ante* incentive to concentrate their efforts in trying to alter the initial allocation and increase the risk borne by the public.[612]

It might be thought that between the different categories of investors, creditors have the greatest incentive to advocate a guarantee—particularly a bailout. The reason is that bankruptcy procedures tend to treat debt-holders more benignly than equity-holders (i.e. their possibilities to be expropriated are limited to a greater extent), an effect that does not seem to repeat itself in a bailout—despite any sort of formal declaration on the contrary arising from governments. Whilst this is true, there is an opposite incentive arising from the guarantee. Equity is artificially made far too expensive, providing firms with incentives to gear-up—a variation of the adverse selection problem but specifically applied to different sorts of regulated markets. Exactly as it has been shown for the financial sector, in utilities markets the implicit guarantee also means that the cost of debt is kept artificially low.[613] For this reason, somewhat counter-intuitively, creditors have an incentive to prevent guarantees and exert managerial control. The outcome of the balance of incentives is difficult to predict.

On the one hand, the guarantee also alters managerial behaviour. Anticipating the bailout, managers will also have few incentives to avoid adverse shocks—i.e., they will not incur in any private effort cost to prevent cost overruns. For a number of reasons, in infrastructure industries moral hazard reasserts itself. The problem becomes clearer if the opposite situation (i.e. the government gives no guarantee) is considered in first place. Assume that the regulator hardens the budget constraint of the firm. This situation may force the management to adopt short-term, risky decisions that may render substantive benefits at the expense of long-term objectives.[614] This is especially true in the case of short-period price reviews, because the investment may not pay beyond the current period. In turn, the focus on immediate performance may decrease the quality of investments. At first glance, providing the firm with no guarantee seems to produce pernicious consequences for the long-term

[612] There is evidence of such transference of risk. The most well-known study is Guasch (2004), who studies renegotiation of concessions in Latin America.

[613] In addition, there are indications that the 'soft' control of the capital structure is also incentivising firms to gear up. Hence the evil is greater.

[614] Von Thadden (1995).

economic incentives of the regulatory regime. However, the situation is worse without the hardening. If the firm knows with greater certainty that there is a State guarantee, the perverse managerial incentives to adopt short-term decisions will remain. That is, by anticipating the bailout the firm will have little incentive to generate revenue, least of all to reduce its costs.[615]

There are two complementary reasons for this. First, the guarantee provides perverse incentives to managers to deviate their efforts to other activities (especially risky ventures). Due to adverse selection, unprofitable projects may nevertheless be financed and continue to be financed even though their quality is low[616]—particularly if quality is not contractible. In a nutshell, the anticipation of the guarantee provides managerial incentives to overinvest. In addition (or alternatively), the manager may simply 'goldplate' the project by considerably increasing their demand for unnecessary inputs.[617] In markets subject to government control, the most likely secondary effect of either of these two situations is that the firm will not meet the targets required by the regulator.

E. Avoiding State guarantees

It remains to provide some possible solutions to avoid embedding negative State guarantees within regulatory frameworks. The task for the government is how to make its commitment not to bailout firms in financial distress stronger – i.e. how to harden the budget constraint of the firm. Providing an explicit reassurance that distressed firms will not be rescued should be the starting point.[618] In some cases the declaration is embedded in the regime, and it is especially strong when it has been incorporated in the country's constitution. Somewhat surprisingly, however, this kind of reassurance is difficult to find within regulatory frameworks—indeed, it has never been provided in the UK context. Informal declarations are

[615] Traditional literature studied the problem considering State-owned enterprises. See *e.g.* Kornai (1980), who analyses soft budget constraint in centrally planned economies.

[616] Dewatripont & Maskin (1995).

[617] The literature has shown that the option for a given regime directly affects managerial incentives. From this perspective, the more efficient option to deal with distress is the one that severely penalises managers for not paying the debt (Aghion *et al.* 1992). That is, the system should keep the bonding role of debt and at the same time penalise managers.

[618] Irwin (2007).

more or less frequent, but their credibility is severely reduced by the absence of institutional support.[619] Beyond these attempts, however, there is simply no possibility of preventing bailouts in all cases or, the same, making the commitment credible and optimal. Furthermore, when the guarantee is embedded in the regulatory contract (as in the case of a hidden guarantee) the problems cannot be overcome by regulatory design.

However, there are specific methods of governmental risk management possible to be applied. At the most basic level is the use of market-based solutions. As seen, in well-functioned markets the most common policy to avoid inefficient liquidation is the creation of a bankruptcy regime that permits and facilitates both liquidation and reforming of firms in distress. Allowing bankruptcy to operate normally may be one possible way to avoid guarantees. Nevertheless, the range of options is wider, and includes more the intrusive alternatives. For example, there is the alternative of exerting *ex ante* control of the capital structure and the imposition of ring-fencing conditions. However, as it was argued in the previous chapter, for a number of reasons it is at least unclear that the control of the capital structure is an effective method to deal with distress at all.[620]Some other measures may be implemented so as to force the firm to internalise liquidity shocks—for example, the imposition over the firms of the obligation to set aside assets to cover unexpected losses.[621] As a result of risk management, the level of risk should be kept to reasonable levels.

When the possibility of use of any of these options—from simple bankruptcy regimes to alternatives closer to command and control—is neglected, or their 'toughness' is decreased with parallel measures that creates opposite incentives, governments undermine the credibility of all their line of actions as a threat. Regrettably, however, that is precisely what has happened in practice. Particularly, bankruptcy has been increasingly abandoned as a mechanism to deal with financial distress. In many cases, the main reason not to allow firms to go bankrupt is the quite strong but erroneous tendency to identify bankruptcy with

[619] Recall that from an *ex post* perspective the guarantee is always in the interest of the State.

[620] Some reasons are related to managerial incentives; others with economic incentives arising from the regulatory regime. At least a 'strong' form of control (i.e. the requirement of certain minimum debt/equity ratio) may not be the most recommendable option for industries subject to economic regulation. That leaves us with the option for *laissez faire*, or for adopting only a 'soft' form of control (i.e. that the regulator only signalises its preference for a given debt/equity ratio), which is the most supported alternative amongst UK regulators.

[621] See Irwin *et al.*(1997: 17).

liquidation. Also, bankruptcy has been undermined by the introduction of less draconian insolvency regimes for some infrastructure sectors.

In the UK, most regulated sectors currently contain specific provisions that vary the applicable insolvency law.[622] These regimes are known as Special Administration regimes (SAR). Whilst in some industries the Insolvency Act may still apply and a receiver or administrator be appointed under it, the SAR allows the relevant regulator and the Secretary of State to alternatively ask the Court for appointment of a Special Administrator. Unlike a receiver or administrator appointed under the traditional insolvency law regime, who would simply be obliged to obtain the best possible price for the assets of the licensee, the Special Administrator is tasked with making a scheme for the transfer of the licensee's assets and undertaking a new licensee and, in the interim, ensuring that the licensee continues to trade.[623] In any case, the result is that the degree to which firms are permitted to fail is low in these markets.

CHAPTER CONCLUSION

Risk allocation and the avoidance of State guarantees are a decisive task of governments. Different rules for risk allocation correspond to different types of risks, and these in turn correspond to different institutional regulatory regimes. Likewise, State guarantees (which arise from the specific risk of financial distress) are also influenced by the regulatory regime. Most guarantees have pernicious consequences, but some of them may occasionally be beneficial. Furthermore, some of them are unavoidable. However, the provision of most State guarantees depends considerably on the regulator's attitudes—mainly, the provision of bailouts for firms in distress and the measures of risk governance. This chapter has shown that regulators should consider carefully their options in order to discriminate between actions that render positive or negative consequences. They have a duty to secure the finance

[622] This is the case of the water, railways and energy sector, the National Air Traffic Control (NATS), and the PPP arrangements for London Underground.

[623] The SAR normally empowers the relevant Secretary of State to make available finance from funds for the purpose of enabling a licensee in special administration to continue to trade.

of private firms' activities. The task is therefore to balance such duty with the protection of consumers, without distorting the incentives provided by the framework.

CHAPTER V

JUDICIAL SCRUTINY OF REGULATORY DECISIONS

> 'When *I* use a word,' Humpty Dumpty said in a rather scornful tone, 'it means just what I choose it to mean – neither more nor less'.[624]
>
> - Lewis Carroll

OVERVIEW

An effective mechanism for scrutiny of decisions made by regulators is an essential part of the regulatory system – there can be little doubt about that. But institutional design in this area is hardly a straightforward task. The designer faces what is arguably one of the central problems of administrative law – namely, to define the proper relation of the body that scrutinises to the regulatory agencies. When considering which mechanism is more suitable for dealing with that problem in a particular regulatory regime, the designer needs to take into account two sets of fundamental questions. The first one is related to the form of the administrative system: is there a need for a special body (i.e. not the common judicature) applying the law? If affirmative, what would be the institutional characteristics of that body? The second question is related to the substance of the law: If there is a need for a special body, does the specific part of the regime subject to this body (e.g., in the context of this chapter, competition law or economic regulation) require the application of different standards of scrutiny? Both questions involve trade-offs. The answer is normally related to a question of expertise and (arguably more important) to the specificities of the part of the regulatory regime at issue.

[624] Carroll (1992 [1871]: 254).

More importantly, the answer to both questions is largely related. The standard of scrutiny needs to match the institutional characteristics of the appeal body. If judges are generalists, non-experts, it is highly likely that they will have incentive not to differ with the expert administration when dealing with special matters – at least in questions of policy, but perhaps also in questions of law. This means they will only review factual questions. In principle, the standard of review should not instruct judges to revise more than these questions; the outcome could be harmful for the objectives the regulatory system pursues. Conversely, a body that is specialist in the area has incentives to differ with the expert administration, because its own knowledge (either accumulated in time or acquired by learning) allows it to confidently examine questions of policy (and law).[625] If the specialist body is somewhat restrained, there seems to be no reason to create such body in the first place. The most common reason for restricting the standard of scrutiny the specialist body must apply is that specialisation may act to the detriment of the regulator, which would only become a fact-finder for the body. For the reasons explained below, that fear and others are largely overstated. The institutional characteristics of the regime should match substantive aspects.

The chapter studies the case of judicial scrutiny in the UK. For many years the study of judicial scrutiny of regulatory decisions and its implications in this area was absent from specialised literature in regulation. A number of the major studies of the UK regulatory regime devoted no more than a few lines – if any – to the topic of judicial control. When it was considered, its treatment was normally included within the broader subject of accountability.[626] Indeed, the opportunity to check decisions of regulators through appeal or judicial review is an important element of accountability.[627] But the major reason for the lack of a detailed treatment was the belief that judges do not (and should not) affect economic regulatory policy.[628] As mentioned in Chapter II, this was the orthodox view of judicial review to economic/commercial matters. The judiciary tended to restrained itself from intervening unless the regulator has exceeded its mandate (the *ultra vires* rule). Conversely,

[625] As it was discussed in Chapter II, specialisation comprises three interrelated notions: *expertise* (i.e. the greater level of judicial knowledge in a specific substantive area), *experience* (i.e., the accumulation of knowledge or skills), and *object-specificity* (i.e., the specific aim of introducing logical coherence to and protecting the objectives of one part of the legal system).

[626] E.g., Baldwin and Cave, 1999; and Ogus (1994).

[627] Prosser (2007: 352).

[628] E.g., Cosmo Graham (2000: 75) predicted a minor role for judges in regulation: 'it is doubtful whether judicial review would have a major impact on the work of regulators'.

the last decades have seen an increasing movement toward, first, the protection of individuals and control of power – a movement that represents a leap forward from the strict *ultra vires* approach, inspired in concepts linked to the principle of the rule of law, such as 'legality', 'fairness', 'reasonableness' and, lately, 'proportionality'.[629]

Indeed, these questions delve into a more general debate regarding the scope of judicial review beyond commercial regulation and the effects on constitutional values – issues that exceed the limits of this thesis and therefore are only tangentially touched upon here.[630] However, the deferential approach has been particularly strong in cases involving regulatory decisions and, as this chapter shows, is incompatible with the current regulatory regime, characterised by the embeddedness of more specialisation within the institutional structure. This feature is particularly notorious after the creation of the Competition Appeal Tribunal (CAT) – a tribunal whose characteristics of expertise, experience and object-specificity place it in a better position to match regulators' expertise. This has resulted in judicial scrutiny having arguably more influence on substantive regulatory policy outcomes. Somewhat contradictorily, however, the CAT provides a case where institutional characteristics do not seem to work in harmony with the substantive aspects of the area. Whilst specialisation has generated incentives for the application of a less deferential approach, the legislator has chosen to restrict the scope of the revision, pushing toward more deference. The result is a rather flawed regulatory design that may impact negatively on regulatory policy, possibly inducing procedural formality to the detriment of substance and increasing regulatory risk.

It is submitted that there is scarce risk in having an approach closer to non-deference, but carefully applied so as to avoid interfere in regulatory policy.[631] The chapter advocates a more intensive form of review, based on the principle of proportionality, and argues that the gains of such an approach could be vast.[632] A well-designed system of regulation needs the courts to play a more interventionist role in order to protect individuals from capricious or

[629] Oliver (1987). That these concepts are part of the rule of law is uncontroversial. *See, e.g.*, Dicey (1885), Jowell (1994), Craig (1997), Allan (2000), and Bingham (2007).

[630] For a recent account, *see* Arancibia (2011). *See also* Engelman (2001: 5): 'the trends developing in commercial Judicial Review are simply illustrative of the developing jurisprudence in Judicial Review generally'.

[631] As it will be explained, the 'no deference' approach has an important impact on the overall dynamic of the regulatory process (*see* section III below).

[632] Paraphrasing Roscoe POUND, this will better permit regulatory law to be stable, but do not stand still.

arbitrary regulatory actions. Such broader approach to judicial review is especially important in a regulatory context of plurality of objectives and actors.

Before continuing, two caveats are important. First, this chapter does not intend to argue that the CAT is necessarily the most suitable tribunal to deal with all regulatory cases (especially if there is a need to recognise plurality of objectives). Its more modest purpose is merely to show how specialisation has worked in practice and draw some lessons from the assessment. Secondly, this chapter does not include decisions of the OFT (the competition agency), although the structure of scrutiny is similar to that of sectoral regulators.[633]

The rest of the chapter is organized as follows. The first part provides a background of the most salient institutional features of the UK regime since privatisation. This part concludes that the main characteristics of the regime are specialisation and the existence of various forms of appeals and standards of review. They are revised in the following two sections. Subsequently, a way forward to tackle the main problems of the design is proposed, analysing its pros and cons. Concluding remarks follow.

I. THE INSTITUTIONAL BACKGROUND

Unlike other systems, both courts and tribunals in the UK are understood to be external to the regulators whose decisions they have power to scrutinise. In economic regulation, there is now a complex web of institutions with different (and sometimes overlapping) competences. Since the privatisation process began during the mid-1980s, the introduction and enhancement of competition became one of the main objectives economic regulators must pursue. In such context, it seemed reasonable to place them under the same institutional framework of competition authorities, whose decisions were mainly reviewed by regular courts. Nonetheless, the designers decided to keep courts away from regulatory developments as much as possible. That would only change with the introduction of specialisation.

[633] For an assessment of the CAT's decisions brought by the OFT, see Whish (2007) and Rodger (2009a, b) and (2006).

A. The original design (1984 – 2002) and its legacy

For some years after privatisation began, the institutional structure of utilities sectors was complex, with a number of different bodies applying and adjudicating discrete aspects of the law. Foremost, the privatisation model was characterised by the creation of independent regulators in charge of the task of overseeing, directing and controlling the newly-privatised firms. To achieve that objective, they were given broad, discretionary powers that could be applied with substantial leeway. These powers included the concurrent application of competition law (other than for mergers) in their respective sectors.[634] Regulatory decisions could only be challenged before the High Court of Justice – most commonly the Administrative Court within the Queen's Bench Division.[635] There was only one exception to the regulatory scrutiny by courts: in case of licence modifications by the regulator (including, indeed, the outcome of the price control), firms were entitled to refer their disagreements with the regulatory decision to the then existent Monopolies and Merger Commission (MMC), which had to conduct a full investigation and report on the issue. The MCC was in essence a regulatory agency, but both its structure and powers gave it a 'quasi-judicial' character.[636] Its rulings were generally amenable to judicial review. Beyond this exception, the original legislation establishing economic regulators only allowed reviews by courts.

The Administrative Court is commonly referred to as 'generalist', because its members deal with disputes arising throughout a wider part of the legal spectrum. Indeed, administrative judges have experience and a high degree of expertise in general public law. However, that characteristic does not amount to specialisation in the sense used in this thesis.[637] They do not have specific cross-disciplinary expertise in economics, business and accountancy (as the members of the CAT must have) and perhaps in some other aspects of public policy that eventually may weigh in economic regulation.[638] Indeed, this feature is entirely anticipated and logic given the one of the key components of the specialisation concept. Courts are

[634] Some regulators held limited concurrent powers in relation to the pre-Competition Act 1998 regime.
[635] In turn, the Court of Appeal may eventually review judgements of the Administrative Court.
[636] Wilks (1999: 51). More details on these characteristics, *infra* section II.C.
[637] See *supra* note 625.
[638] 'Generalist' is by no means a pejorative expression; it only highlights this fact.

expected to pursue justice, not to protect a specific objective. They are not result-oriented – i.e. they lack object-specificity. Their judicial action only aims to protect the legal coherence within the legal system *as a whole*.

The design, however, did not show any special preference for (not even a consideration of) the advantages of having generalist judges. The inclination was instead to keep the *status quo* (so to speak). The Administrative Court has inherent jurisdiction governed by the rules of precedents.[639] It revises decisions presented mainly through judicial review claims. Judicial review is available to the parties whenever no other special right of appeal is available. Broadly speaking, its content comprises the revision of illegality, unreasonableness and procedural unfairness, *plus* the control of proportionality.[640] This is fundamentally different from merit revisions, which extend to the widest variety of questions. At time of privatisation there was a very limited involvement in the substance of regulation – that is, merit reviews were scarce. Judicial scrutiny was mainly formalistic, confining its role to procedural issues rather than its substance.[641] Judges, it was commonly held, are debarred from substituting their own decision for that of the primary decision-maker.

Designers of the regime acknowledged the nature of the scrutiny courts tended to carry out at time of privatisation. They deemed courts were 'ill-suited' to deal with economic regulatory affairs, so they made a 'conscious decision…to keep the law and the courts out of the new system'.[642] The designers tried to avoid the high levels of judicialisation of regulation, as it had occurred in the US.[643] With that purpose, a discretionary, co-operative and closed style of government and regulation was largely favoured. Some authors argue that '[t]he vaguely defined duties of the [regulator] mean that his decisions are unlikely to be struck down by the

[639] Supreme Court Act 1981, s. 31; and Civil Procedure Rules, part 54 (judicial review and statutory review).

[640] The content of each one of these areas is further considered below.

[641] See *infra* III.B.1. This approach rested in wider views about judicial scrutiny. See Le Sueur *et al.* (2007).

[642] Foster (1992: 267).

[643] See Prosser (1997: 57) ('An important reason for not adopting older or US models was a desire to avoid judicialization of regulatory procedures and to prevent the courts from having a major role as a means of challenge of regulatory decisions'). By the same token, Black *et al.* (1998). However, according to Foster (1992: 267), most important may have been the influence of a new piece of labour legislation that aimed to introduce regulations to get labour relations out of the courts as far as possible. Foster (1992) recognises that this was mere coincidence, but 'the reasons that seemed compelling in that case flowed over into the other'.

courts…'.[644] However, vagueness of duties does not imply that the role for courts fades – as the case of antitrust rules so clearly illustrates.[645] Most accurate is to acknowledge that statutes were drafted 'in a very permissible language hitting at determinations of unlawfulness'.[646] Also, it was estimated that the remedies under English administrative law did not provide the adequate safeguards. Overall, the design resulted in regulators with ample powers to act the way they deem fit (and, as will be shown, in very few challenges made against the use of those discretionary powers).

The situation remains relatively stable until 1998. That year the government passed the Competition Act, which brought the English regime of competition in line with the European developments. The Act, which came into force in March 2000, introduced a change that would eventually become vital for the judicial institutional regime: the Competition Act replaced the MMC with the Competition Commission (CC).[647] The CC was *internally* divided into two parts, one of which was the 'Competition Commission Appeal Tribunals' (CCAT).[648] The reason for its creation seems to lie in the desire to closely align UK competition law with European law.[649] The need for an appeal body was evident. Locating the CCAT within the CC remained the preferred scenario, because it would give the former access to the staff and expertise of the latter. In fact, following a long-standing UK tradition of mixing lawyers and lay members to deal with competition law matters[650], it was first envisaged that the 10 members of the CCAT were to be drawn from an expanded MCC. This proposal did not prosper, but the mixture was conserved: judicial specialisation was inserted for the first time in the UK regime.

---

[644] Armstrong, Cowan & Vickers (1994: 170).

[645] Antitrust statutes are purposely vague so as not to discourage business. Despite this fact, courts have developed a refined jurisprudence through common law. See, among others, Baxter (1982) and Easterbrook (1983).

[646] Rawlings (2010). On unlawfulness and other concepts, see *infra* III.B.1.

[647] The CC was originally introduced by s. 48 of the Competition Act 1998.

[648] The CCAT was first envisaged in a White Paper of 1989, which concerned cartels and restrictive agreements (DTI, 1989). It followed the earlier Green Paper (DTI, 1988). For a valuable account of the historical background to the establishment of the CCAT, see Grinyer & Pryor (2007).

[649] In one of the preparatory documents, the Court of First Instance (currently Court General) was mentioned as an example of a body whose decisions would be based on both judicial and economic elements (see DTI, 1996).

[650] In the UK, there is a long tradition to appoint lay members in the assessment of competition law issues. This goes back at least to the Restrictive Practice Court created by the Restrictive Practice Act 1956. See Green (2007) (indicating that a fundamental part of the character of the CAT was 'the continuation of a long tradition of appointing non-lawyers to panels to sit in conjunction with the legally qualified chairman')

Notably, the CCAT was given extensive powers and enjoyed full jurisdiction on all appeals from decisions by the Director General of Fair Trading (the OFT predecessor) and the sector regulators. During the passage of the Competition Bill in 1997, the minister responsible for the Bill highlighted the importance that appeals against the substance of the regulatory decisions should be dealt with by 'more court-like arrangements', and indicated that 'the tribunal will have many of the characteristics of a court… it will be recognized as having the same importance and status as the High Court, as appeals from the tribunal will be to the Court of Appeal'.[651]

Indeed, the CCAT 'operate[d] much like a court'.[652] But in its beginnings it had to surmount certain historical legacies, mainly due to its 'internal' location within the CC and the different modus operandi of the latter. Somewhat conflictingly, the nature of procedures was fundamentally different in both bodies. Whilst the CC's procedures were fundamentally administrative in nature, the CCAT would hear 'proper appeals'. Also, the possibility of sharing staff (a characteristic regarded as an advantage by the designers) divided the views and eventually was put into practice. All problems were overcome in the end and the tribunal managed to create its own reputation – some have argued that personal leadership may have played a crucial role in this.[653] However, as will be seen, the caseload developed more slowly than predicted.

With the creation of the CCAT, the Competition Act 1998 introduced two crucial changes to the judicial scrutiny of the regulatory regime. On the one hand, it created – perhaps unintentionally – an 'administrative model' somewhat akin to the system of administrative judges that operates inside the US Federal Trade Commission.[654] On the other, it introduced specialisation. The administrative model was short-lived; specialisation survived and became a notable feature of the UK regulatory regime.

---

[651] Quoted in Competition Appeal Tribunal, *Guide to Proceedings* (October 2005), at 4.

[652] See DTI (2001: para. 4.4). More details on the CC are given *infra* II.C.

[653] See Grinyer & Pryor (2007) and Green (2007).

[654] By referring to an administrative model, I intend to highlight only that the adjudicator is placed outside the judicial branch of the government—i.e. either being part of the Executive or, as in the case of the CCAT, being part of an independent administrative agency.

B. Developments after 2002

In 2002 the model changed with the passing of the Enterprise Act 2002. The model maintained the feature of specialisation, but instead of an administrative model of review, it adopted a judicial model. In this sense, the UK model diverges expressly from the US and Australian model, for example, which rely on generalist courts. Also, it established a double jurisdiction for the newly created tribunal.

1. Specialised justice: the CAT

Along with granting more powers and independence to the CC[655], the Enterprise Act 2002 changed the institutional landscape in a crucial way. The administrative model that the Competition Act 1998 had introduced was abandoned in favour of a judicial model of review. Despite the fact that the White paper that presaged the passing of the Enterprise Act indicated that 'the Government proposes to widen the remit of the Competition Commission Appeal Tribunals enabling them to hear claims for damages brought by harmed parties', the Act in the end transformed the administrative tribunal into a new judicial tribunal: the CAT.[656] Since the tribunal acquired new functions as appeal body against the CC's decisions, it became independent.

Importantly, specialisation was kept. First, the objective was given by the nature of the specific topic where the judgements would be made – namely, competition law and economic regulation.[657] Secondly, the CAT's design matched object-specificity with increased expertise. By statute the President must appear to have appropriate experience and knowledge of competition law and practice; and the chairmen must appear to have appropriate experience and knowledge of any relevant law and practice. Generally the CAT's wing

[655] See *infra* II.C.

[656] Enterprise Act 2002, s. 12 *et seq.* and Sched. 2.

[657] See *Napp Pharmaceutical Holding Ltd v The Director General of Fair Trading* [2002] EWCA Civ 796, [2002] UKCLR 726, at para 34 (per Buxton LJ): '[The CAT is] an expert and specialist tribunal, specifically constituted by Parliament to make judgements in an area in which judges have no expertise'.

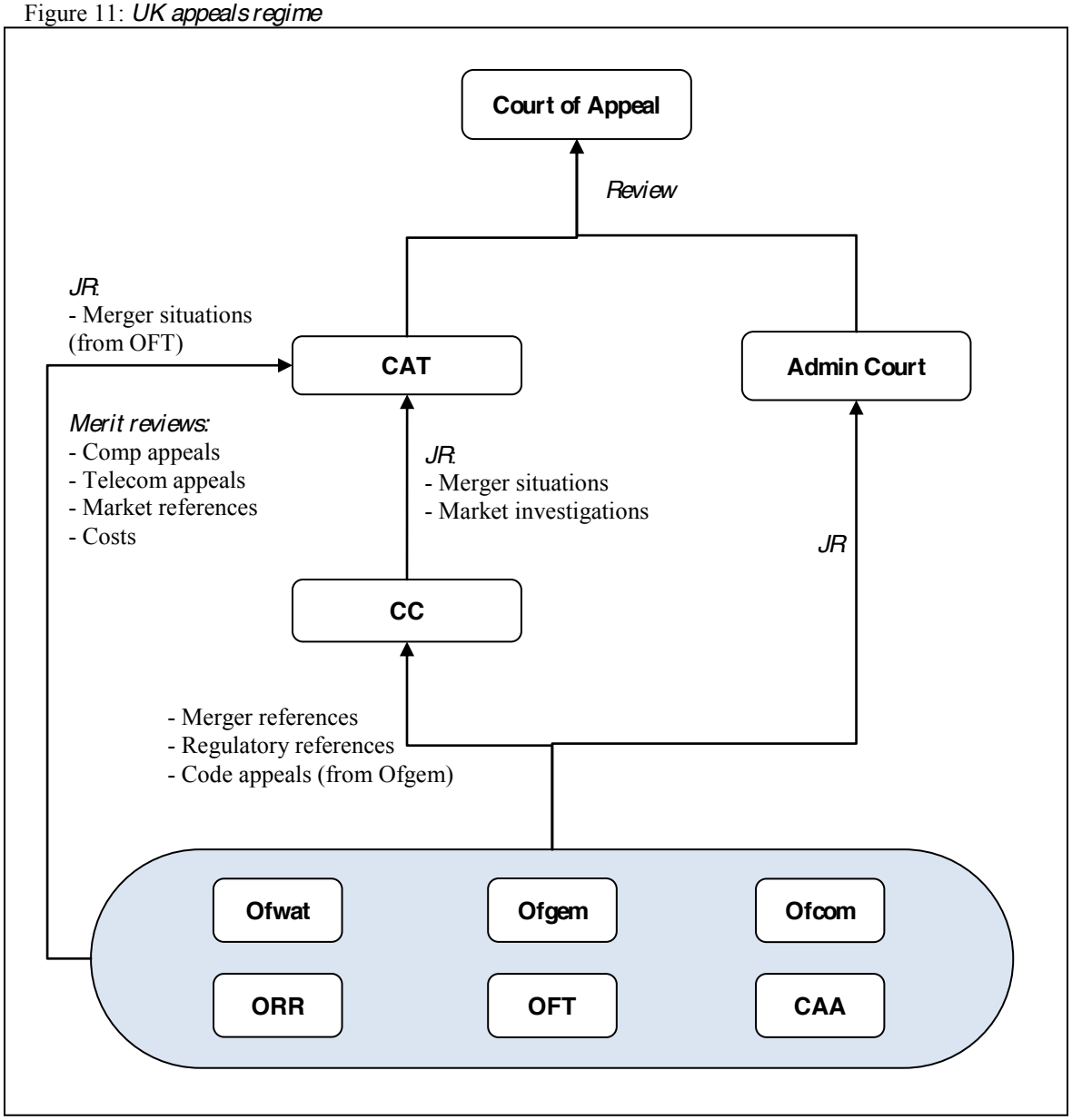
members have cross-disciplinary training in economics, business, accountancy, and perhaps in some other areas of public policy eventually important for carrying out their functions. This is uncommon in generalist courts.[658] Nonetheless, most of the chairmen and the President are judges of the Chancery Division of the High Court.[659] Finally, since the creation of the CAT, judicial interventions in economic regulation have had a multiplier effect (see below). The CAT has become the front-line of the justice system. This has increased the experience of judges, therefore completing all the requirements to be considered a specialised body: the CAT is a tribunal whose members may claim a deep level of expertise and experience in a specific domain.

Depending on the territorial jurisdiction affected, judgements of the CAT (either on a point of law or in penalty cases as to the amount of the penalty) are reviewable by the Court of Appeal in England and Wales; the Court of Session in Scotland; or the Court of Appeal in Northern Ireland. Any of these courts or the CAT can give permission to further appeal. When reviewing a decision of the CAT, courts should recognise the special feature of this tribunal. As the House of Lords has ruled:

> [T]he ordinary courts should approach appeals from [specialized tribunals] from an appropriate degree of caution; it is probable that in understanding and applying the law in their specialised field the tribunal will have got it right … They and they alone are the judges of the facts. It is not enough that their decision on those facts may seem harsh to people who have not heard and read the evidence and arguments which they have heard and read. Their decisions should be respected unless it is quite clear that they have misdirected themselves in law. Appellate courts should not rush to find such misdirections simply because they might have reached a different conclusion on the facts or expressed themselves differently…[660]

[658] But see *supra* note 637 and accompanying text.
[659] On the importance of this 'mixed' composition, see *infra* note 820 and accompanying text.
[660] *AH (Sudan) v Secretary of State for the Home Department*, [2007] UKHL 49, [2008] 1 AC 678, at para.30 (per Baroness Hale).

Figure 11: *UK appeals regime*



2. The wide variety of avenues to challenge regulatory decisions

The CAT was given an ample, double jurisdiction that encompasses both revisions of the merits of the case and judicial review.[661] Different statutory provisions mandate what kind of

[661] Four main statutes confer jurisdiction on the CAT: the Competition Act 1998; the Enterprise Act 2002; the Communications Act 2003; and the Energy Act 2008. See also Rayment (2005). Note that it is not a strange feature that the CAT's jurisdiction is created by statute. The whole competition law regime in the UK is generally a creation of statute. Competition law plays only a minor role in the common law through the doctrine

review it must apply in each case. The CAT is bounded by the option already made by Parliament as to which standard of review applies. Hence, it can only make a discrete choice regarding the type of scrutiny. However, for the parties the distinction creates an important trade-off between the speed of the scrutiny *vis-a-vis* its scope. Whereas the scope of judicial review may be more limited, judicial review at least delivers quicker answers.[662]

The double jurisdiction covers four areas:

- *Judicial review in merger situations* and *market investigations*: the CAT deals with applications for judicial review of decisions of the OFT (the competition authority), the Secretary of State or the CC in relation to these matters.[663] This includes modification of licenses to exclude alterations proposed by the regulator at the end of the price review, which were already possible to be 'appealed' to the CC.

  In cases of judicial review, statutory provisions normally mandate the CAT to apply 'the same principles as would be applied by a court on an application for judicial review'. Section III of this chapter further explains this standard in more detail.

- *Merit review (generally) on telecommunications appeals*: some provisions order appeal on the merits in issues such as price control and others.[664] The base is on European legislation, which requires a revision of the merits of the case in the field of telecommunications.[665] In issues related to price controls, the CAT must make a reference to the CC before disposing of the case.

---

of restraint of trade. See *Esso Petroleum Ltd v Harper's Garage (Stoutport) Ltd* [1968] 1 AC 269. The doctrine has been recognised as part of the domestic competition law, e.g. in *Days Medical Aids Ltd v Pihsiang Machinery Manufacturing Co* [2004] EWCH 44; [2004] 1 All ER (Comm) 991, at para. 265.

[662] Eventually, the length of the process can even be considered as part of the due process.

[663] Enterprise Act 2002, s. 120(4) (regarding merger situations) and s. 179(4) and Sch. 9, pt. 2 (regarding market investigations).

[664] Communications Act 2003, s. 192 (providing merit appeals on price controls, with exceptions established in schedule 8); and s. 316 and s. 317 (providing merit appeals on broadcasting). Judicial review is the other route of appeal for Ofcom's decisions under Part 2 of the Communications Act 2003 and Parts 1 to 3 of the Wireless Telegraphy Act 2006.

[665] See article 1(4) of Directive 2009/140/EC of the European Parliament and of the Council of 25 November 2009 [2009] OJ L 337/37, that amends Article 4 of Directive 2002/21/EC on a common regulatory framework for electronic communications networks and services (hereinafter, the 'Framework Directive'): 'Member States shall ensure that the merits of the case are duly taken into account' in an appeal.

- *Merit review (generally) on competition appeals*: the CAT hears appeals against all sector regulators and the OFT for infringements and non-infringements of the Competition Act 1998.[666]There are a few exceptions where the appeal is not on the merits.[667] When the statute creates such exceptions, the issue may still be challengeable by way of judicial review in the Administrative Court.

- *Merit review on costs*: some penalties imposed by the CC can also be appealed to the CAT, which therefore revises the merit of the sanction.

The CAT has indicated that when applying revision on the merits, its standard is one of 'profound and rigorous scrutiny'.[668] Any reason underpinning the regulatory decision, including the existence or lack thereof, may be scrutinized in order to determine the correctness of the decision. The CAT may revise not only the procedural aspects of a decision, but also the underlying policy and factual considerations. As explained in *Freeserve*, 'the Tribunal has, in principle, jurisdiction…to decide whether the [regulator] has made an error of fact or law, or an error of appraisal or of procedure, or whether the matter has been sufficiently investigated'.[669]

Strikingly enough, the possibilities for regulatory scrutiny do not exhaust in the ones mentioned above ones. First, note that under the Enterprise Act 2002 sectoral regulators also have powers to undertake market investigations, although this power does not carry any right

---

[666] Competition Act 1998, s. 46 and s. 47, and Sched 8.

[667] Competition Act 1998,s. 46(g) and 46(h), and s. 47 (b) and 47(c).

[668] *Vodafone& others v Ofcom*, [2008] CAT 22, at para. 46: '…[I]t is still incumbent on [regulators]… to conduct their assessment with appropriate care, attention and accuracy so that their results are soundly based and can withstand the profound and rigorous scrutiny that the Tribunal will apply on the appeal on the merits…'.

[669] *Freeserve v Director General of Telecommunications* [2003] CAT 5, at paras. 110-1. See also *Hutchison 3G v Ofcom* [2008] CAT 11, at para. 164: 'this is an appeal on the merits and the Tribunal is not concerned solely with whether the [appealed decision] is adequately reasoned but also with whether those reasons are correct'. These statements are compatible with the intentions of the drafters of the Act. As the Minister for Competition and Consumer Affairs during the passage of the Competition Bill (pre-Competition Act 1998) in committee on 18 June 1998 said:

> It is our intention that the tribunal should be primarily concerned with the correctness or otherwise of the conclusions contained in the appealed decision and not with how the decision was reached or the reasoning expressed in it. That will apply unless defects in how the decision was reached or the reasoning make it impracticable for the tribunal fairly to determine the correctness or otherwise of the conclusions or of any directions contained in the decision. Wherever possible we want the tribunal to decide a case on the facts before it, even where there has been a procedural error, and to avoid remitting the case to the [regulator]. We intend to reflect that policy in the tribunal rules.

to appeal. Secondly, notice that current legislation gives some regulators (namely, in the energy, rail and postal sectors) powers to levy penalties on regulated firms.[670] In these cases, affected parties can 'appeal' before the High Court, not the CAT. However, the court does not carry a full revision on the merits. The appeal looks more like judicial review, although the court has the power to amend the amount of the penalty. Finally, the Energy Act 2004 introduced a right of appeal to the CC against energy code modification decisions of the regulator.[671] This appeal is available to anyone materially affected by the regulatory decision or a body that represents such persons.

* * * * * * *

In sum, after privatisation a number of different bodies administered or adjudicated discrete aspects of the law. Substantial changes to the model were introduced in 2002. Two core characteristics lie at the heart of the current institutional design. The first one is judicial specialisation; the second is the division between judicial review and appeals on the merits. The next sections analyse these two characteristics.

II. SPECIALISATION

One of the key differences between the different bodies in charge of reviewing economic regulatory decisions lays in their dissimilar degree of specialisation.[672] This section explains the main empirical and theoretical consequences arising from the introduction of specialisation in the UK regime and why these results justify the increase in specialisation.

[670] Utilities Act 2000, ss. 59 and 95; Transport Act, s. 225; and Postal Services Act, ss. 30-7.
[671] Energy Act 2004, ss. 173-4.
[672] Sir Robert Carnath has argued that the 'principal distinguishing features [of the tribunals], as compared to the courts, are flexibility, specialisation, and accessibility' (Carnath, 2009: 69).

A. How does specialisation affect regulatory outcomes?

It remains to assess why specialisation matters, from both the theoretical and the empirical perspective.

1. The theoretical underpinning

Courts delimit the extent of sector specific economic regulators' decision-making powers (that is, their discretion) by exerting different levels of intervention. At least theoretically, judicial revision can be deployed throughout a continuum that ranges from minimum scrutiny of the primary decision to total substitution of judgement. These two extremes represent the two administrative doctrines that set limits to the judicial control of regulatory decisions – the 'deference' and 'non-deference' approaches.[673] Both deference and non-deference are useful heuristic devices used simply as a sort of 'shortcut' to stress agency discretion or judicial discretion, respectively. For that reason, notwithstanding the fact that judicial control is a question of degree, the dichotomy is still useful as an analytical tool. In this sense, the study of judicial scrutiny may be rightly framed in this language.

The deference view indicates that whenever the meaning of a statute is ambiguous with respect to the specific question at issue, courts should refuse to consider the legal merit of the regulator's decisions. That is, the regulator is given a wider margin of discretion in its pronouncements. On the other hand, non-deference can simply be defined negatively as lack of deference – i.e., the judicial assessment (or reassessment) and the decision of part or all the substantive issues already resolved by the regulator. Only in its most extreme version (which hereinafter will be called 'interference') does it mean the entire substitution of the administrative judgement made by the primary decision-maker by what a court believes to be the right interpretation of the issue in question. Therefore, the closer non-deference is to

[673] The reference to 'degrees of deference' is common in the literature. *E.g.*, Craig (2008: 615); Young (2009). US authors also use the terms regularly. See *e.g.*, Tiller (1998: 117); Breyer (1986: 366); Levin (1985).

interference, the more reduced the level of judicial acceptance of the regulator's substantive interpretation.

The level of specialisation crucially affects the scope of the judicial revision and the extent of the regulators' and judges' discretion. Discretion may generally refer to questions of law, fact and/or policy.[674] Because of their lack of expert knowledge or their lack of experience, it is expected that generalist courts do not revise all the aspects of the regulatory decision. Commonly, assessment will be directed to procedural questions and to review whether there are no noticeable departures from the law. But questions of law and policy will normally be left to policymakers. That is, courts will tend to maintain themselves bound to *discrete* choices, with the expected result that they defer more to agency interpretation. In fact, this is exactly what has occurred in English law. Whilst all questions of law are reviewable by courts since the House of Lords' decision in *Page*[675], generalist courts have traditionally declined to become involve in questions of fact and policy.[676] In fact, under the traditional model of scrutiny there was an important recognition of the regulator's expertise: the CC's predecessor, the MMC, did not lose any judicial review case despite its decisions being ordinarily challenged.[677] Arguably, the lack of judicial specialisation resulted in the embeddedness of deference in the core of the newly created regimes for economic sectors.

As it will be revised later, arguments of economic complexities and technicalities are normally stressed to justify this sort of 'light touch' approach. Such arguments, however, are not particularly sound in presence of specialist tribunals. Specialisation creates incentives to increase judicial discretion, because the tribunal has incentives to revise *all* aspects of the regulatory decision – that is, a specialist tribunal makes *continuous* choices. Greater

[674] Amongst many others, Dworkin (1977: 22); Wade & Forsyth (2004: 250 *et seq.*). The distinction is also widely used in American literature. See e.g., Breyer *et al.* (2002: 227), and Spiller & Tiller (1996).

[675] *R v Hull University Visitor ex parte Page* [1993] AC 682.

[676] E.g., Harlow & Rawlings (2009: 311). Confirming that English Courts will not normally decide factual issues: *R v DG Telecommunications ex parte Cellcom*, [1999] COD 105, at para. 26 (per Lightman J): '[T]he court should be very slow to impugn the decisions of fact made by an expert and experienced decision-maker, it must surely be slower to impugn his educated prophesies and predictions for the future'. On regulatory policy: *Holder v Law Society* [2003] EWCA Civ 39, [2003] 1 WLR 1059. Similarly, *E v SSHD* [2004] EWCA Civ 49, [2004] QB 1044, and *Unichem v OFT* [2005] CAT 8, [2005] 2 All ER 440, at para. 177.

[677] As usual, there are exceptions: *South Yorkshire Transport v Monopolies and Merger Commission* [1993] 1 WLR 23 (considering the decision 'aberrant'). The point has also been noted by Harlow & Rawlings (2009: 312).

specialisation allows tribunals to feel more confident of their own capabilities when dealing with topics that in principle would appear to fall almost exclusively under the remit of regulators – and hence to judge even technically complex matters.[678] Specialist tribunals have a deeper understanding of relevant policy objectives, reduce the probability of simple inadvertence, may pursue coherence and minimize dependence on the views and adversarial skills of opposing counsel.[679] Indeed, all of this does not imply an argument in favour or disfavour of specialisation. The argument, at least at this stage, remains positive: if a tribunal is specialized, it will arguably be less inclined to be deferential to the regulator, no matter how expert that regulator is.

To appreciate in more detail why specialisation leads to less deference, it is necessary to understand first how regulators adopt their decisions and the content thereof. The regulatory decision is bound by legislation, which means that it must obey the principles and conform to the duties already established within the regulatory framework (mainly in statutes). For example, assume that the regulator confronts only discrete choices (say, X_1 and X_2) and must opt for one of them. Whilst X_1 maximizes competition law, one of the objectives mandated by the statute, X_2 maximizes any objective other than competition (e.g., security of supply in the electricity sector, or reliability of the network, universal service, and so on). Indeed, regulators cannot choose or change the boundaries of their decisions, so for the purposes of this chapter it is assumed that they act within these boundaries. The content of the regulatory decision may vary. For simplicity, assume that it has only two central contents; let's say facts and law. That is, each decision (X_1 or X_2), whatever the objective it maximizes, comprises a sub-decision within the factual space and a certain rule (or principle), each being individually contributory to the process of maximization. Generally, regulators have substantial leeway to choose the main content of each choice. The variety of relevant sources of law (EU instruments, national statutes, etc.) and even the precise wording and categorisation of the objective (e.g., the use of words such as *shall* or *may*, whether there is a principal or secondary duty, specific indications, etc.) all affect the regulators' degree of freedom.

[678] E.g. Lianos (2010) (focusing on the admissibility and assessment of economic expertise in the related field of EC competition law).

[679] Legomsky (1990: 22).

The *ex post* type of review that a court or tribunal adopts has an impact on the decision the regulator will finally adopt *ex ante*. The reason is that, depending on the specific institutional characteristics, the regulator can to some extent anticipate the possible outcome of the revision.[680]

Consider first the case of the non-specialised court. Since a non-specialised court will only examine whether one part of the regulatory decision (the law) complies with the objective that must be pursued according to the statute, the outcome of the revision is relatively clear – to a great extent. Either the court will affirm the regulator's option if the chosen rule maximizes the objective, or it will reverse the option if the rule does not pursue the objective's maximisation. The regulator may easily anticipate this result. It knows in advance that following certain rules (commonly, procedural rules) and/or using some mechanisms (including guidelines, agency adjudication, and rulemaking, among other tools[681]) the probabilities of a court reversing its decisions decrease. By sharp contrast, the outcome is less clear when the court is specialised. Since a specialized court focuses on every aspect of the regulatory option, the regulator does not know whether the court will agree that considering certain facts in the decision maximises the same objective. The reason is straightforward. Whilst assessing the regulatory rule is relatively simple on an *ex ante* basis, the assessment of facts (or policies) comprises much more discretion, which leaves greater space for judicial revision of regulatory decisions. A specialised court can reach the conclusion that the facts, the law, or both, in the simplified model, do not maximize the objective.[682] As one commentator has accurately said, for regulators this reflects a 'crude equation: review of experts by generalists – wide margin of appreciation; review of experts by other experts (potentially even "more expert experts") – narrow margin'.[683]

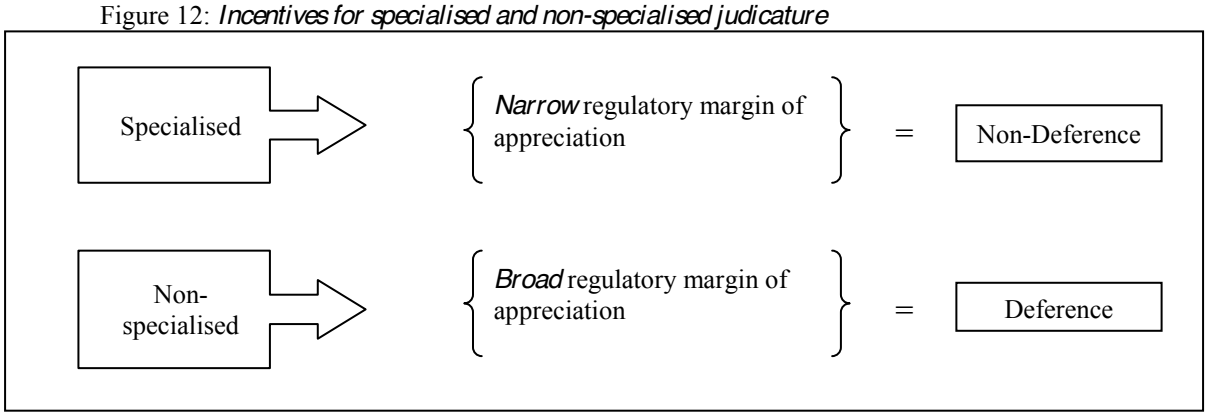
[680] By contrast, Graham (2000: 75) has argued that judicial review does not have an impact on the work of regulators because 'it is difficult to predict the outcome of a judicial review case'.

[681] In the US context, see Hanssen (2000: 539).

[682] Indeed, the same analysis may in principle apply to questions of policy.

[683] de la Mare (2007: para. 14).

Figure 12: *Incentives for specialised and non-specialised judicature*



2. Empirical consequences.[684]

The years after privatisation may be characterised as a period where parties made few legal challenges to regulatory decisions. The role of MMC (which was in practice, as seen, a 'quasi-judicial body') remained largely confined to those areas where its intervention was required by statute (i.e., airports), but it did not play a major role where its intervention depended on parties' referrals.[685] As late as 1992 it was still considered 'an unfired gun'.[686] By 1993 it had dealt with one inquiry.[687] In the 14 years up to 1999, there were two gas, two electricity, four various aspects of telecommunications and two water licence adjudications – out of a total of 35 inquiries.[688] Moreover, even some overtly erroneous decisions were not referred to either the MMC or its successor the CC. For instance, in the water sector the calculation of the X factor in the well-known formula RPI-X applied by the regulator was at the centre of a conflict during the price control in 1999, due to the over-reliance on technical analysis.[689] Likewise, the miscalculations made by the regulator during the electricity price control for 1995-2000 are considered an example of reneging on own commitments caused by an excess of expectations on regulatory outcomes. These or similar cases did not give rise to referrals. True, there were some important decisions during the years that followed

[684] This part does not assess judicial review, only scrutiny by specialised bodies. See *R v Director General of Gas Supply, ex parte Smith*, CRO/1398/88, QBD, 31 July 1989, for an early example of judicial review.

[685] Kay (1993: 59) (discussing the account of the MMC given by its chairman).

[686] Wilks (1999: 261).

[687] *Chatline and Message Services: a report on the provision of Chatline and Message Services by means of the British Telecommunications public switched telephone network*, MMC No. 238, 21 February 1989.

[688] Wilks (1999: 258-9).

[689] For a detailed account of this case, Gómez-Ibánez (2003: 223-40). The formula was actually RPI + K.

privatisation.[690] However, the apparent use regulators did of referrals, using them as 'sanctions', may have lead firms to back off in the face of the prospect of a reference, decreasing the importance of the MMC and even the CC in its early days.[691]

After 2001 the situation changed. Particularly after the commencement of the CAT's functions in April 2003, the CAT has become the most notorious body that carries out judicial scrutiny of regulatory agency decisions. Table 8 shows the number of judgements issued during the period 2001-10 related to economic regulation. Note that, in accordance with the theoretical explanation set out in the next section, the table includes only judgements that arguably affect regulators' decisions – that is, final judgments, judgments on admissibility of appeals, and judgments disposing appeals. Conversely, the table does not include judgments on jurisdiction, preliminary issues, commitments, remedies, costs, amendments on notice of appeal, admissibility of evidence, rates in dispute, and so on; nor any ruling or decision.[692] The total number shown in Table 8 is 26. It is apparent that the major impact has been in the telecommunications sector. The number of cases ruled per year is relatively stable, although it has slightly increased since 2006.

Table 8: *Number of judgements of the CAT in Economic Regulation 2001-10*

Year	Ofcom	Ofgem	Ofwat	CC	Total
2001	---	---	---	---	0
2002	1	---	---	---	1
2003	1	---	1	---	2
2004	3	---	---	---	3
2005	2	---	---	---	2
2006	2	---	3	---	5
2007	1	---	2	---	3
2008	4	---	1	---	5
2009	3	1	---	1	5
2010	---	---	---	---	0
Total	17 (65%)	1 (4%)	7 (27%)	1 (4%)	**26 (100%)**

Notes: (1) Table only includes judgements that are final, dispose appeals or rule on admissibility
(2) ORR & NIAUR have no cases so far.
Source: CAT website (www.catribunal.org.uk).

[690] For example, Monopolies and Merger Commission: *British Gas plc* (1993), brought under the powers of the MMC to deal with licence modifications, which initiated the progressive liberalisation of the gas sector.

[691] Various commentators have remarked this point: Prosser (2007: 352), Wilks (1999: 256), Kay (1993: 59).

[692] 'In practice, a number of different expressions is used in relation to the Tribunal decisions: they are referred to as "judgements" where the decision deals in detail with the substantive issues in the case; and at some other times as "rulings" or "decisions" where predominantly procedural or ancillary issues such as costs are involved' (CAT *Guide to Proceedings*, *supra* note 651, at 56 [note 18]).

However, there are different biases in the type of regulatory cases the CAT scrutinises in practice. First, there are just a few competition cases.[693] This may be linked to attitudes of regulators. Various sources have found an apparent tendency for regulators, where they have a choice, to rely on their sector-specific powers rather than to pursue a case under the Competition Act.[694] Furthermore, most competition decisions are non-infringement decisions. This outcome is shown in Table 9:

Table 9: *Competition enforcement decisions by regulators since 2002*

Regulator:	***Ofcom***	***Ofgem***	***ORR***	***Ofwat***	***Total***
Type of decision	23	4	7	2	36
Infringement	0	1	1	0	2
Non-infringement	23	3	6	2	34

Source: Adapted from NAO (2010, fig. 2 & 4).

The reasons mentioned are allegedly related with sectoral powers. They would be clearer, more appropriate for imposing structural remedies and removing barriers, and they would lead to speedier results, because they do not enable third party complainants to appeal to the CAT (hence scarce resources would not be devoted to the appealed cases rather than other, more urgent, priorities). However, the reason may well lie in specialisation. According to the theoretical underpinnings shown in the previous section, arguably the increased deference that comes with specialisation is pushing regulators to avoid being subject to review. In fact, the biases toward telecommunications' cases make this point stronger: the choice does not exist in that sector, because parties can raise a dispute in such a way that the regulator has to use its regulatory powers rather than its competition powers. Therefore, the incentives arising from the institutional design are affecting regulatory choices.

[693] For example, the first successful judicial review of any CC remedy since the Enterprise Act 2002 came into force was *Tesco* (*Tesco plc v. Competition Commission ad others*, [2008] CAT 20), handed down by the CAT on 4 March 2009, which followed a CC market investigation into the supply of groceries in the UK (*The supply of groceries in the UK, market investigation*, CC, 30 April 2008). As mentioned before, competition cases are not included in this assessment. But this fact stresses the relatively few cases the CAT has confronted in competition matters.

[694] *E.g.* NAO (2010), DTI (2006), and Pimlott (2010).

In fact, the NAO has recently found that there are strong incentives in the system to appeal against a regulatory decision.[695] However, the explanations it provides are weak: the relative immaturity of the system; the importance of the decision; the nature of the firms (i.e., powerful and well resourced); and the complexity of the subject matter.[696] A better explanation lies in the institutional design: parties want to appeal because they trust that the specialised nature of the body that scrutinises will provide it with enough capacities to enable it to analyse the regulatory decision in detail.

The second bias in the type of regulatory cases the CAT scrutinises is related to the previous one: unsurprisingly, when decisions are actually scrutinised, telecommunications stands out as the most reviewed area. Because of the institutional design, the regulator in this case cannot avoid merit review. The total of judgments showed in Table 8 above (26 judgements) correspond only to a handful of cases: 17 cases, of which 11 are appeals brought under the Telecommunications Act 2003. Seven judgments in the water sector correspond to only four cases, with four of those judgments being issued in a single case. Overall, appeals on the merits represent approx. 95% of the total number of cases.[697]

Table 10: *Regulatory cases rules by the CAT in each sector (2001-10)*

Regulator	Type of Review & Statutory Provision	Case	Judgements by case	Neutral citation	Content of Decision
Ofgem	Merit review / Competition Appeal	National Grid	1	[2009] CAT 14	Judgement
Ofwat	Merit review / Competition Appeal	Albion (DwrCymru)	4	[2008] CAT 31	Judgement (unfair pricing)
				[2007] CAT 8	Refusal of permission to appeal
				[2006] CAT 36	Judgement (dominance and other issues)
				[2006] CAT 23	Judgement
		Thames Water	1	[2006] CAT 7	Judgement
		Independent Water Company	1	[2007] CAT 6	Admissibility
		Aquavitae	1	[2003] CAT 17	Admissibility of appeal
Ofcom	Merit Review /	Freeserve	2	[2002] CAT 8	Admissibility

[695] NAO (2010: para. 3.5).
[696] *Ibid.*, at para. 3.7.
[697] So far, the analysis has included only some judgments of the CAT. To assess the real workload of the tribunal, however, it would be necessary to include all other judgments and rulings. In that case the total number increases sharply.

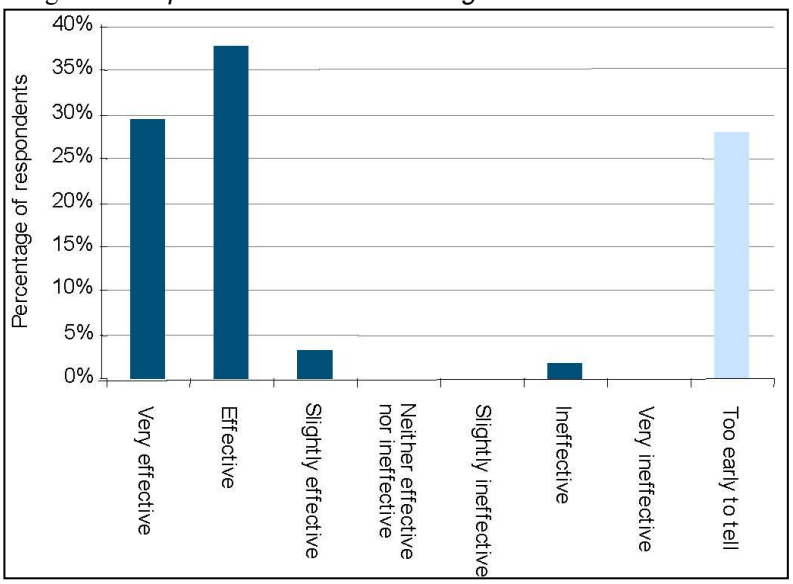
	Telecoms Appeal	(1007/2/3/02)		[2003] CAT 5	Final judgement
		BT (CPS) (1025/3/3/04)	1	[2004] CAT 23	Final judgement
		Floe (1024/2/3/04)	3	[2004] CAT 18	Final judgement
				[2005] CAT 28	Reasons for refusing permission to appeal
				[2006] CAT 17	Judgement
		BT (1018/3/3/03)	1	[2004] CAT 8	Judgement
		Hutchison 3G (1047/3/3/04)	2	[2005] CAT 39	Judgement
				[2009] CAT 11	Disposal of the appeals
		Mobile Call Termination (1083/3/3/07) (1085/3/3/07)	2	[2009] CAT 1	Tribunal's power on disposal
				[2008] CAT 11	Judgement (non price control matters)
		The Number (UK) ltd. (1100/3/3/08)	2	[2009] CAT 4	Permission to appeal
				[2008] CAT 33	Judgement
		Media Marketing Promotions (1053/3/3/05)	1	[2006] CAT 12	Final judgement
		Vodafone (1094/3/3/08)	1	[2008] CAT 22	Judgement
		VIP communications (1027/2/3/04)	1	[2007] CAT 3	Admissibility
		Termination Rate Disputes	1	[2008] CAT 12	Judgement (core issues)
CC	Judicial Review / s. 179 Enterprise Act 2002	BAA	1	[2009] CAT 35	Final judgement

Source: CAT website (www.catribunal.org.uk).

Notwithstanding the number of cases, it now seems more accepted that judicial pronouncements affect economic regulatory policy and may influence substantive outcomes. In fact, already in 2004 – just a year after the tribunal began functioning – a study made by consultants found that the CAT produced an immediate impact among stakeholders (although a substantial minority felt it was too early to comment on the effectiveness of the CAT), as shown in Figure 13.[698] Reportedly, independence of operation, clarity of action and strength of leadership were the key factors valued by respondents. The view among practitioners and academics alike seems to be that such a positive opinion has not varied.

[698] KPMG (2004).

Figure 13: *Impact of the CAT in reviewing cases at 2004*



Source: KPMG (2004)

However, the influence of judicial pronouncements in regulatory outcomes depends on whether judgments are biased to one of the parties – i.e. whether the decisions are pro-respondent/defendant or pro-appellant. Table 11 includes the same cases mentioned above. In each case it is detailed whether the regulator's decision has been upheld (therefore the decision is pro-respondent) or quashed (therefore the decision is pro-appellant). Only final judgment in each case has been considered. As the table shows, the CAT's decisions have been slightly balanced in favour of appellants, although the numbers do not show any tendency. Judicial influence in outcomes seems, so far, rather unpredictable.

Table 11: *Nature of CAT's decisions 2001-10*

Regulator	Case	Nature of the Decision (*)
Ofgem	National Grid	Pro-respondent
Ofwat	Albion (*DwrCymru*)	Pro-respondent
	Thames Water	Pro-appellant
	Independent Water Company	Pro-respondent
	Aquavitae	Pro-respondent
Ofcom	Freeserve	Pro-appellant
	BT (CPS)	Pro-respondent
	Floe	Pro-appellant
	BT	Pro-appellant
	Hutchison 3G	Pro-appellant
	Mobile Call Termination	Pro-respondent
	The Number (UK) ltd.	Pro-appellant
	Media Marketing Promotions	Pro-respondent
	Vodafone	Pro-appellant

	VIP communications	Pro-appellant
	Termination Rate Disputes	Pro-appellants
CC	BAA	Pro-appellant

(*) The decision is 'pro-appellant' if the challenge has succeeded on at least one ground.

Finally, it is also important to assess the number of judgments that are reviewed and overturned by superior courts. Specialisation can hardly deploy its advantages if decisions in practice only move from a lower judicial level to the upper level, where non-specialised courts constantly overturn them with careless substantive analysis. Constant overruling may incentivise the specialised court to change the way it decides cases (perhaps even at the expense of economic reasoning). Nonetheless, the reality might also be the other way round. Outside the remit of economic regulators, WHISH has noted that for the period 2000-06 the OFT had adopted 19 infringement decisions, of which 12 were appealed to the CAT and all but one were upheld, or substantially upheld (although in some cases the level of the penalty was varied).[699] In contrast, as Table 12 shows, in regulatory domains it is not possible to show a clear tendency. At least in principle, therefore, specialisation can (and do) work well in this area.

Table 12: *Court of Appeal judgements (2001-10)*

Regulator	Case	Citation	Decision (**)
Ofgem	National Grid	[2010] EWCA Civ 114	Confirmed, but penalty was varied
Ofwat	Albion (DwrCymru)	[2008] EWCA Civ 536	Dismissing the appeal
Ofcom	T-Mobile (*)	[2008] EWCA Civ 1373	Dismissing the appeal
	Floe	[2006] EWCA Civ 768	Ruling against the CAT's decision
		[2009] EWCA Civ 47	Upholding the appeal
	Mobile Call Termination	[2010] EWCA Civ 391	Allowing the appeal
	The Number (UK) ltd.	[2009] EWCA Civ 1360	Referring the matter to the ECJ

(*) Not mentioned in Table 7.
(**) The decision is referred here only in general terms.

B. Is specialisation justified?

Generally speaking, specialisation is the answer to heightened complexity. In public policy, the more basic reason to establish a specialised jurisdiction is that the legislator considers that there are valuable reasons which existing generalist jurisdictions are unable to deal with in an

[699] Whish (2007: 401).

effective way. In other words, object-specificity constitutes the main reason to embed specialisation in the regulatory regime. Indeed, there are many others. Uniformity, predictability, expertise and better decision-making are among the most commonly mentioned. In the specific case of the CAT, the designers saw particular advantages in the flexibility to tailor the rules under which the tribunal operates to suit the nature of the matters it deals with and having a body with wide competences to deal with issues relating to the whole of the UK.[700]

A main reason that supports specialisation in the current context is the multiplicity of objectives regulators deal with. This contrasts sharply with the early years of privatisation. At that time, regulatory objectives were supposed to be centred on the introduction of competition in newly open markets – more precisely, one specific view of competition.[701] Although 'political compromises had to be made, and the legislation…contained a number of duties mixing economic and social rationales'[702], the objectives were generally economic oriented – at least at the beginning of the process. If there were other objectives, these were introduced in a somewhat incoherent manner.[703] Conversely, the twin goals of efficiency and competition were embraced across sectors in a context largely dominated for an ideology of free markets. As the House of Lords would later recognise, '… there is no clear distinction between sectoral regulation and the promotion of competition in those sectors'.[704] The effect of this approach was double-fold. On the one hand, regulatory design was 'downgraded', meaning that little attention was given to regulators and the processes.[705] On the other, important substantive features of regulation were neglected – foremost, the plurality of objectives.

Indeed, regulation is not only about introduction of competition. As seen in Chapter I, such a view stems from the over-reliance of markets as governance mode. It makes markets seem as the only alternative to State intervention. However, the dichotomy is fictitious. As Chapter II showed, in reality there is a continuum of various modes of governance. Regulation plays a

[700] CAT, *Guide to Proceedings*, *supra* note 651, at 4.
[701] See *supra* chapter I.
[702] Prosser (2005a: 68).
[703] *Ibid*, at 77.
[704] Lord Simon of Highbury, *Hansard*, HL, vol. 583, col. 914 (25 November 1997).
[705] Wilks (1999: 253).

role in all of them, and its objective is far from a single one. This fact was reflected in the UK regulatory practice. Over time, as many features of the regulatory regimes changed, a number of other objectives were increasingly added to the regulatory frameworks. New social and environmental concerns signalled that competition was no longer the almost exclusive concern.[706] In many occasions, these new objectives are misaligned with competition, which has been somewhat 'downplayed' by the legislator. There are a few examples in utilities legislation. In the case of energy regulation, for example, legislation now indicates that the main objective of the regulator is to respect competition 'whenever it is possible', whilst protecting the interests of 'existing *and future* customers' – which includes the protection of security of supply.[707] The statute also indicates that when carrying out its functions the energy regulator must *have regard* to the need to contribute to the achievement of sustainable development.[708] Similarly, the Postal Services Act mandates the regulator to exercise its functions in the manner which it considers is best calculated to ensure the provision of a universal postal service, and then indicates that the regulator must promote competition 'whenever appropriate'.[709] In these cases, the statute leaves regulators the option to either give less weight to competition or to pursue solely the other objective when implementing a particular policy.

Indeed, competition remains important. It should always be present in the analysis, so as to diminish restrictions to it as much as possible. As the early privatisation model, the current approach also stresses some form of competition as an underlying core idea. However, in principle it is up to the regulator to choose how to deal with the incompatibility of an objective in an actual case. Regulators, as policymakers, are tasked with choosing the objective they deem worthy of protection or advance in each case, and have the burden to demonstrate why a specific objective was chosen. In economic regulation, most of the reasons are not principled, but adopted for technical motives or any other reason generally linked to expertise. The question is what are the desirable characteristics of the institution that

[706] See Prosser (2005a) for a complete appraisal of the changes.

[707] E.g., Electricity Act 1989, s. 3A(1): The principal objective of the [regulator]... is to protect the interests of existing and future customers in relation to electricity conveyed by distribution systems or transmission systems, *wherever appropriate by promoting effective competition* between persons engaged in, or in commercial activities connected with, the generation, transmission, distribution or supply of electricity...' (emphasis added). Similarly, s. 4AA(1) of the Gas Act 1986.

[708] E.g., Electricity Act 1989, s. 3A(2)(c) and Gas Act 1986, s. 4AA(2)(c).

[709] Postal Services Act 2000, s. 3(1) and 5(1), respectively.

is entrusted with the task of scrutinising regulatory decisions made in this context. In order to design the most suitable institution, some considerations must be taken into account. First, despite the fact that statutes may give some guidance on the 'ranking' of the objectives, often such an order does not exist. As mentioned, it is up to the policymaker to give precedence to one or another in each case, and this needs to be acknowledged by the institution in charge of the scrutiny of the decision. Secondly, the objectives are normally open, with a variety of possible contents equally suitable to achieve them. The regulator provides the actual content, and during this process the underlying reasoning normally relies on questions of expertise.

The enlargement of objectives decreases the capabilities of generalist courts to scrutinise regulatory decisions. The incentive to maintain themselves bound to *discrete* choices makes the courts' reasoning incompatible with the need to occasionally make deep revisions of the content of a decision to assess whether the objective has been achieved. At most, they can check whether the means have been appropriate – something courts commonly do, as will be seen in the next section. Moreover, discrete choices and generalist knowledge is incompatible with an assessment of the specific content of the objective considered by the regulator. Conversely, specialisation decreases these shortcomings to a great extent. There is a strong case to be made that specialised tribunals are a better place to deal with complex regulatory decisions that respond to a variety of objectives.

The own Enterprise Act 2002 acknowledged this argument indicating that when the CC considers remedies in relation to regulated markets it needs to take into account various objectives regulators possess (which, once again, may go beyond the prevention of adverse effects on competition).[710]

C. Are there two specialist appeal bodies?

As seen, in the UK specialisation in judicial bodies (or quasi-judicial bodies) was introduced in 1998 with the creation of the CCAT, the predecessor of the CAT. In order to fully understand the reasons for the change, it is necessary to appreciate the role and position of the

[710] Enterprise Act 2002, s. 168.

CC inside the institutional design. This role was a consequence of historical developments in the area of competition law.

The CC's predecessor, the MMC, has its origins in 1948, with the passing of the first statutory regulation of competition policy: the Monopolies and Restrictive Practices (Inquiry and Control) Act.[711] As CRAIG has explained, 'the title of the 1948 Act accurately conveys the intention of the framers of the legislation'.[712] Despite the fact that 'most of the early reports of the Commission concerned cartels'[713], its jurisdiction was restricted to actions of dominant firms, leaving anti-cartel enforcement to other bodies.[714] Subsequent legislation enlarged the jurisdiction of the MMC to the area of merger control[715], resale price maintenance[716], monopoly references[717], and 'anticompetitive practices'.[718] Hence, it formed the regulatory frameworks of the privatised utilities.[719] Broadly speaking, regulators were given powers to make references to the MMC to ask the latter to report on whether the referred matters operated, or were expected to operate, against the public interest and whether modification of licence conditions could remedy or prevent those adverse effects. In its report the MMC could specify modifications, but it was up to the regulator or the government to carry them out.[720]

[711] The original body was the Monopolies and Restrictive Practice Commission (MRPC), split into two in 1956 (the Monopolies Commission and the Restrictive Practices Commission), and finally renamed MMC in 1965.

[712] For an account of the institutions before the 1990s, see Craig (1987: 205). See also Scott (2009).

[713] Craig (1987: 205).

[714] Fundamentally, the Restrictive Practice Court, created by the Restrictive Trade Practice Act 1956 (which was then replaced by the Restrictive Trade Practices Act 1976).

[715] First by the Monopolies and Mergers Act 1965 and then by the Fair Trading Act 1973. According to the latter, the Director of Fair Trading or a government minister could refer mergers to the MMC.

[716] Resale Prices Act 1964 (later replaced by the Resale Prices Act 1976).

[717] Besides mergers, the Fair Trading Act 1973 also contemplated referrals to the MMC in three cases: fact-finding investigations, which objective was to determine only whether a statutory monopoly exists and whether it is being exploited; investigations to specific practices of dominant firms; and references to determine whether a monopoly is likely to operate against the public interest.

[718] Further enlargements were made with the passing of the Competition Act 1980, which introduced the concept of 'anticompetitive practice' in UK law. The concept, however, was vague and did not match the current categories commonly know in UK or EC competition law.

[719] S. 54 and Sch. 10 of the Competition Act 1998 gave regulators powers in competition matters in concurrence with the OFT. See also The Competition Act 1998 (Concurrency) Regulations 2004, SI 2004/1077.

[720] See, e.g., sections 13 and 15 of the Telecommunications Act 1984 (original text). The Secretary of State retained some veto powers over the regulator's decision.

This was referred by the legislation as 'appeals'. However, this appeal was different to the common appeal in two senses. On the one hand, the MMC was unconstrained by the actual dispute, but could investigate whether any matter referred to it operated against the public interest. The public interest test contained an ample variety of objectives, including not only the maintenance of 'effective competition', but also the maintenance and promotion of the 'balanced distribution of industry and employment' and the promotion of the interest of consumers, among others.[721] On the other hand, only the regulated firm allegedly affected by the decision could make a referral. However, notwithstanding the explicit powers given to the MMC, it lacked the possibility of directly applying remedies –at least in theory its role remained largely one of an adviser. The MMC constituted a sort of 'enhanced-regulator' that guarded against damages to competition, but which lacked the necessary tools to fully comply with its mission. That did not seem to be an impediment for regulators to make use of the threat to refer companies to the MMC as a tool to enhance their bargaining position – i.e. regulators used referrals as 'sanctions' to pressure companies to agree to modifications of licence conditions.[722] Nonetheless, as noted in the empirical analysis, the MMC in itself remained for many years 'the dog that has not barked'.[723]

The CC replaced the MMC in 1999, following the passing of the Competition Act 1998. As its predecessor, specialisation remains a feature of the CC. Its members are experts, experienced people whose task is specifically defined. For each inquiry there is a decision-making body of at least three independent experts (drawn from a wider panel of around 50 appointed members), which are supported by a specialist staff team. Soon after, the Enterprise Act gave the CC wider powers and greater independence. It introduced a new regime for the assessment of mergers and markets in the UK and turned the CC's role in this area clearly focused on competition issues, replacing the wider public interest test in the previous regime.[724] Most importantly for present purposes, though, were the 'regulatory' powers given to the CC. The CC carries out revisions of some regulatory decisions, including certain aspects related to price controls, in areas dictated by the relevant sector-specific

[721] Fair Trading Act 1973, s. 84.
[722] See *supra* note 691 and accompanying text.
[723] Kay (1993: 58).
[724] Although the public-interest test remains in place in some parts of the legislation: media plurality, national security and financial stability.

statute. In a nutshell, if a regulated company does not agree to a modification of its licence proposed by the regulator during the price review, the regulator must refer the question to the CC, which will uphold the regulator's decision or order the modification of the licence.

The Competition Act 1998 has kept for the CC the same role that the MMC formerly had – that is, simply to make recommendations to government.[725] By contrast, the Enterprise Act gave the CC *remedial* powers to direct companies to take certain actions to improve competition. In utilities regulation, these powers include modifying licenses or not to include alterations proposed by the regulator at the end of the price review. That is, the Enterprise Act not only centred the role of the CC on competition assessment; it also provided it with the adequate tools to protect competition. Nonetheless, the test that the CC must apply in utilities cases remains a public-interest test. Generally, the CC must consider whether any matter referred to it may be expected to operate against the public interest and, if so, whether this could be remedied by modifications to the licence. The public-interest test is broader than the competition test, and involves more discretion. Along with the enlargement of the powers, it significantly increased the stance of the CC.

As a result of the mix of specialisation and remedial powers, in practice the CC became a truly 'appeal body' in all but name. That feature remained even after the Enterprise Act 2002 split up the judicial function. Therefore, in parallel to judicial review claims made before the Administrative Court, parties that disagreed with certain regulatory decisions were given the option to challenge the decision to the CC, which has the power to revise the full 'merits' of the decision – its decision in turn being subject to judicial review before the CAT. Allegedly, a noticeable proportion of the CC's resources are now devoted to conducting regulatory inquiries, often in the form of appeals of proposed licence modifications.[726] The transformation took a further step with the changes introduced by the Energy Act 2004. With the introduction of the new right to appeal against energy code modification decisions of the regulator, the act expressly transformed the CC into a 'proper' appeal body.[727]

[725] See *supra* note 717.

[726] Geroski (2006: 331-40).

[727] The CC has so far determined one such appeal: that brought by E.ON against GEMA's decision to mandate certain changes to the Uniform Network Code (UNC) governing GB's gas network. The outcome disfavoured the regulator. See CC 02/07, *E.ON UK plc v Gas and Electricity Markets Authority* (10 July 2007).

Notwithstanding the true nature of the CC's function (i.e. whether quasi-appellate or proper appellate functions), there is a question to be made on whether it is acceptable for an appellate body to be located outside the judicial branch of government – particularly in a jurisdiction where there is no single formal written constitution. Despite the fact that, as seen, the administrative model was short-lived in the UK, the CC still remains an odd organisation that does not seem to fit within the current institutional design – it seems to be more an incongruous reminiscence of the past with a jurisdiction incoherently created. In practice, it means that there is a double specialised scrutiny of regulatory decisions in some areas of regulatory law. Within the current design, this is unjustified.

* * * * * * *

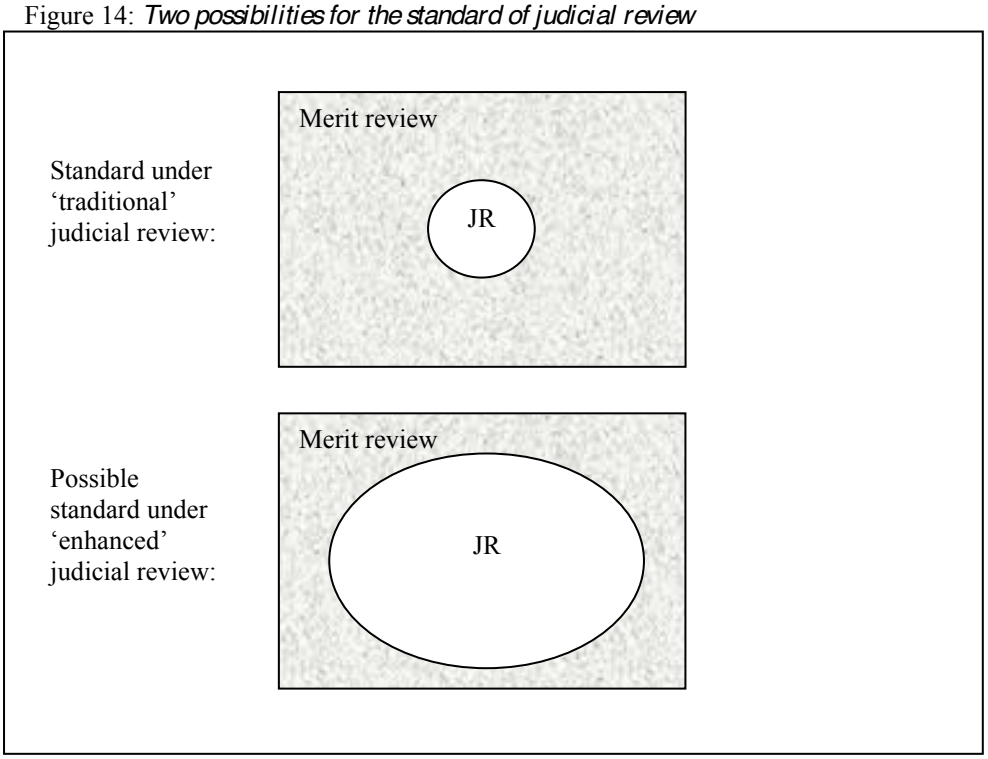
In sum, the empirical evidence indicates no strong support for specialisation. From a theoretical standpoint, however, specialisation creates an incentive to act with less deference to the regulator. Hence the regulator has fewer chances to try to circumvent judicial review in the first place making use of various mechanisms. For this to occur, the standard of review that the specialised body must apply should arguably match the incentives for more deference – i.e. the standard should allow the court to act with substantial leeway. The next section analyses the standard in the UK context.

III. THE EVOLUTION OF STANDARDS

The second characteristic of the institutional design is the distinction between judicial review and appeals on the merits. This section further explains the distinction and its current application to economic regulation. As seen, whilst some statutory provisions normally mandate to apply 'the same principles as would be applied by a court on an application for judicial review', others allow merit revisions, which extend to the widest variety of questions in order to determine the correctness of the decision. This is a traditional differentiation that lies at the heart of English public law. An important caveat, however, that is not always

sufficiently stressed, is that all judicial review claims are, to some extent, review on the merits – as Figure 14 illustrates. In this sense, the aforementioned nomenclature of deference and non-deference seems more appropriate. Although neither of these concepts represents a standard in itself (e.g., there is no 'standard of deference'), they rightly emphasise that the scrutiny is a question of degree. What matters is to indicate where the frontier lies within the continuum – i.e. to find out what the standard of review is, what the normative arguments that support the chosen standard are, and what the legal test used to determine the compliance with the standard is.

Figure 14: *Two possibilities for the standard of judicial review*



These questions lie at the core at the core of a more general discussion regarding the constitutional relationship between the judiciary and the executive. A broader, more substantive standard of review implies a reduction in regulatory agency autonomy which is inconsistent with the courts' perception of their own position within the fundamental order. In the orthodox view, the agency, not the judge, determines how the balance between private rights and regulatory needs ought to be struck. A different standard reverses such view.

Therefore, it is necessary to acknowledge that any movement toward a more intrusive judicial approach will express a rather different conception of the separation of powers – a conception that must be based on solid principles. As it will be argued, such principle is proportionality.

This section begins by exemplifying how a broad standard of review (i.e. one that encompasses merit review, but without incurring in substitution of judgment) works in practice. The narrow standard is then explained, showing how it applies to economic regulation. As previously mentioned, at the time of privatisation courts were reluctant to apply a broad standard of review. They did not revise substance, but rather confined themselves merely to procedural issues when scrutinising regulatory decisions by way of judicial review. That system disabled courts from protecting investors and hence protected regulatory discretion.

A. Review of the merits

Whilst reviewing the merits of a decision, the CAT has the amplest jurisdiction. Therefore, it has not hesitated to engage in the discussion of complex economic issues and challenge even the regulatory premises of action of the regulatory authorities. Only an example should suffice to demonstrate this approach, and the clearest one is probably *Albion* – a case related to disputes about access to facilities operated by water undertakings.[728] The CAT upheld an appeal against the decision of the water regulator (Ofwat) and found that the undertaking in question had engaged in margin squeeze practices.[729] The main facts of the case are explained in Box 1.

[728] The procedural history of the case is highly intricate. I mainly refer to the following judgements: [2008] CAT 31 (hereinafter, '*Albion*, CAT unfair pricing judgement'); [2006] CAT 23 (hereinafter, '*Albion*, CAT main judgement'); [2006] CAT 36 (hereinafter, '*Albion*, CAT further judgement'); [2008] EWCA Civ 536, [2008] WL 2033584, (hereinafter, '*Albion*, CoA judgement); and [2009] CAT 12 (hereinafter, '*Albion*, CAT remedies judgement').

[729] Generally speaking, a form of abuse of dominance in competition law.

Box 1: *The Albion case*

In 1999 *Albion* became the first new statutory water firm since privatisation of the water industry in England and Wales. It replaced *DŵrCymruCyfyngedig*, a water firm, as the statutory water undertaker in respect of *Shotton Paper Mill*, in the area of Deeside. *Albion* continued purchasing non-potable water in bulk from *Dŵr* at the point of supply to *Shotton* and then resold the water to the latter. The bulk supply agreement expired in May 2003, but the parties continued to operate it as if it had not expired. The price paid by *Albion* to *Dŵr* under the agreement was the same that *Shotton* agreed to pay to *Albion* under a separate supply agreement.

In September 2000 *Albion* manifested its wish for a different arrangement: a 'common carriage arrangement'. Under the new deal, Albion would purchase water directly from *United Utilities*, the current supplier to *Dŵr*, at the point where the water was abstracted from the River Dee, and would pay *Dŵr* for using its water distribution network system and treatment works while transporting the water to *Shotton*.

Albion complained about the price applied by *Dŵr* for the common carriage, arguing that it was excessive, gave rise to a 'margin squeeze' and was discriminatory. The complaint related to an alleged breach of a prohibition stated in the Competition Act 1998, of conduct amounting to an abuse of dominant position. It was made to the Director General of Water Services, whose functions were later transferred to the Water Services Regulation Authority. When the Director rejected the complaint, *Albion* appealed to the CAT, which in turn found, *inter alia*, that the Director's conclusion on the issue of margin squeeze was 'erroneous in law and incorrect, or at least, insufficient, from the point of view of the reasons given, the facts and analysis relied on and the investigation undertaken'. Nonetheless, the CAT did not mention which course of action had to be taken. It also did not resolve the excessive pricing issue, and expressed concerns with respect to the matter of dominance (although the latter was not the focus of the appeal). This led to a new judgment.

The CAT issued a new decision in November 2006. It found that at all material times *Dŵr* had a dominant position in the relevant market. Hence, a possible abuse was analysed. The CAT considered that further investigation and new hearings were needed before adopting a decision on the excessive pricing issue. However, the CAT established that *Dŵr* had abused a dominant position by imposing a margin squeeze, and decided to confirm an existing order for interim relief. Later, the CAT refused to grant permission to appeal. *Dŵr* then made an application for permission to appeal to the Civil Division of the Court of Appeal. It was refused on the papers, but permission was granted on two specific issues: the correct legal test for finding a margin squeeze, and the jurisdiction of the CAT to decide the issue of dominance, rather than remitting the matter for decision by the Authority. The Court issued a decision on 22 May 2008, rejecting *Dŵr*'s appeal on the issue of margin squeeze, and confirming the CAT's jurisdiction to issue a decision on the matter of dominant position.

The CAT analysed in great detail several aspects of the regulatory decision – particularly those related to the identification of the anticompetitive margin squeeze. First of all, the CAT chose the relevant benchmark. As it is well-known in competition law, there are two main 'imputation tests': the 'as efficient competitor' (AEC) test and the 'reasonable efficient competitor' (REC) test.[730] Both of them have been endorsed by European legislation, so there

[730] Whilst the AEC test focused upon the costs of the dominant undertaking's own downstream operation, the REC paid more attention to the costs of an actual or potential competitor – even one less efficient than the incumbent – in the downstream market. As the *Notice on the Application of the Competition Rules to Access*

is no definitive test and directives leave space for regulatory determination. However, the competition case-law has continuously followed the AEC test, and so did the CAT.[731] It might be argued that this is not the best outcome. Whilst the election of the relevant test is certainly a judicial task, an allegedly better approach should be that courts establish the 'appropriate circumstances' for the application of a given test, and only then should they choose the most suitable test for the specific case. Nonetheless, even if courts are correct in their approach of relying on a test from the outset, they still need to consider the regulatory implication arising from the preference for the AEC test.

In *Albion*, three main regulatory aspects following from that election were the objects of in-depth scrutiny. First, the CAT said that the application of the AEC test leads to the consideration of a 'notional business' – i.e., a hypothetical retail arm of the incumbent. According to this view, to establish whether there is a squeeze, the regulator must create a theoretical downstream business of the integrated monopoly.[732] Then, the costs must be allocated to that business, including an appropriate amount for profits. If this retail arm can trade profitably at the upstream price charged to the competitors, there is no squeeze. The CAT unambiguously noted that this is *the* right course of action to determine whether there is a squeeze of the competitor's margin. Regulators are not free to decide other methods they deem more suitable for their purposes.

Agreements in the Telecommunications Sector ('Access Notice'), OJ [1998] C 265/2, at paras. 117 & 118, explains: the AEC implies 'that the dominant company's own downstream operations could not trade profitably on the basis of the upstream price charged to its competitors by the operating arm of the company'. Regarding the REC, it indicates: 'In appropriate circumstances, a prize squeeze could also be demonstrated by showing that the margin between the price charged to competitors on the downstream market (including the dominant company's own downstream operations, if any) for access and the price which the network operator charges in the downstream market is insufficient to allow a reasonably efficient service provider to obtain a normal profit (unless the dominant company can show that its downstream operation is exceptionally efficient)'.

[731] The Commission clearly applies the AEC test since its 1988 Commission decision in *Napier*. In 2000, the Court of First Instance (nowadays General Court) reaffirmed the AEC test in *Industries de Poudres*, and did it again in *Deutsche Telekom* in 2007. According to Geradin & O'Donoghue (2005), the Commission apparently applied the REC test in *National Carbonising*, but the decision does not elaborate further on the margin squeeze issue.

[732] *Albion*, CAT main judgement, n 728, para 900. The failure to consider the costs of a notional retail arm of the incumbent was in the CAT's view a 'central weakness' of the regulator's decision (*ibid*, at para. 906). On this approach, the CAT relies explicitly on the Commission's decision in *Deutsche Telekom AG* [2003] OJ L 263/9, at para. 140 (upheld by the General Court in case T-271/03, *Deutsche Telekom v Commission* [2008] ECR II 000, at para. 188).

Secondly, and related to the previous point, the application of the AEC test also led the CAT to reject the notion of 'avoided costs' as a 'satisfactory basis' for the margin squeeze test.[733] According to this argument, the application of such a notion would imply that the competitor would need to be 'more efficient' than (as opposed to 'equally efficient' to) the incumbent in order to be able to compete in the market. The reason is that the avoided costs method takes no account of the incumbent's fixed costs or the entrant's total costs. Nonetheless, the Court of Appeal accepted an alternative, more deferential use for the notion of avoided costs. Despite the fact the notion cannot be used as part of the test for squeeze, the court held that it might still be used as an objective justification (alongside displacement, amongst others).[734] Needless to say, one approach or the other has important consequences for the burden of proof and the potential substantial findings.

Finally, and arguably more relevant, the CAT ruled on the suitability of the so-called 'Efficient Component-Pricing Rule' (ECPR) to determine margins and any alleged squeeze.[735] In a nutshell, the ECPR is a rule of marginal-cost pricing according to which it is optimal to set the access price to a bottleneck equal to the direct cost of providing access plus the opportunity cost of providing access to the interested provider, which is equivalent to the reduction of the incumbent profit caused by the provision of access – i.e., the price minus the direct cost and the marginal cost. As such, the ECPR is a second-best access rule for cases where the user-level price has already been fixed (ensuring absence of monopoly rents) and – crucially – the regulator is concentrated *solely* on productive efficiency.[736] In practice, neither the ECPR nor its modifications has received much application. On the contrary, it has been by and large looked at with general mistrust by regulators. Nonetheless, for many years the access pricing debate has been centred on this rule and its potential applicability.[737]

---

[733] *Albion*, CAT main judgement, n 728, para 910 (ruling out the 'avoided costs' principle as the basis of reasoning of some European decisions); and *Albion*, CAT further judgement, n 728, at para. 305 (the 'avoided costs' argument is open to the same objections of principle as the ECPR approach). The Court of Appeal upheld: *Albion*, CoA judgement, n 728, at para. 103.

[734] *Albion* CoA judgement, n 728, at para. 106: 'In our view [the objective justification for a zero or negative margin] was the appropriate context within which to consider such matters'.

[735] For more details on the ECPR, see Annex 1 to this chapter.

[736] The rule was first, but separately, developed by Willig (1979) and Baumol (1983). See also Baumol & Sidak (1994); Sidak & Spulber (1997: 283 *et seq.*); Vickers (1997).

[737] Despite its apparent simplicity, the ECPR is highly controversial, being the focus of strong academic debates regarding its usefulness as a practical rule. See, amongst others, Armstrong *et al.* (1996), Economides & White (1995) and (1998). *Contrast* Baumol *et al.* (1997), Larson (1998) and Baumol & Sidak (1994).

In *Albion*, the regulator had decided to apply the ECPR. Indeed, there is no legal reason that prevents it from using that approach; whether it induces more or less competition is a different (and again, highly debatable) matter. However, in a decision that borders interference, the CAT felt confident to decide whether the application of such debatable rule is meritorious or not, or whether the rule is reasonable enough to be applicable. It clearly rejected the use of the ECPR when it stated that 'it cannot be assumed that [the incumbent's] upstream price is reasonable…[t]he margin squeeze in question cannot be justified on the basis of an ECPR approach which is itself unsound'.[738] Beyond the rightness of the decision and the value of the criticisms that can be made to the ECPR, the ruling clearly illustrates the non-deferential approach. The CAT put competition (one view thereof) in front of any other objective that the regulator could have considered when deciding to apply the ECPR. Also, its ruling has an important practical downside: with the rejection of ECPR as a practical access pricing rule, the CAT may have left regulators with tied hands, because arguably the two main theoretical solutions to tackle the access pricing issue (the other being the impractical Ramsey pricing rule) were discarded as workable rules.[739] This may be an important unintended effect of non-deference whose consequences still cannot be assessed. In fact, when asked to specify a minimum retail margin, the CAT considered it was 'not appropriate' to issue an order on this which would bring the infringement to an end, alleging that 'courts normally avoid direct price administration… [t]his was not an appropriate role for this Tribunal in the present case'.[740] One may at least wonder whether the practical aspects also played some part in denying the calculation.

---

[738] *Albion*, CAT main judgement, n 728, at para. 873. The main reasons are stated at para. 835. There was no permission to appeal against the 'ECPR part' of the CAT's reasoning. Note that some commentators have argued that the General Court took exactly the opposite direction in *Deutsche Telekom* (*supra* note 732), where it adopted 'something broadly akin to the ECPR'. See Vickers (2009: 8) (providing no further reasons for his statement). The assertion is based in para. 237 of the judgment. Although it is not entirely persuasive that the General Court did embrace the ECPR, the controversy surrounding the substantial aspects of the rule is beyond the scope of this work.

[739] The use of Ramsey prices was discarded, e.g., in 2003 by the CC as base for the calculation of the costs of call termination because (i) it would breach the cost-causation principle; (ii) it is impractical setting Ramsey prices correctly; (iii) there would be a demand for the own CC to set prices, which is erroneous; and (iv) there are 'formidable problems' associated with the correct computation (*see* 'termination charges' case, at para. 1.6). See generally Laffont & Tirole (2000).

[740] *Albion*, CAT remedies judgement, n 728, at para. 55. At para. 54 indicated that setting the minimum retail margin would require 'a yet further fact-finding investigation'.

B. The scope of judicial review

1. The traditional approach

Unlike substantive revisions, where judges are entitled to revise the 'correctness' of the decision as such (i.e., whether it is right or wrong), procedural revisions are understood as a challenge to the way in which a decision has been made. The concern is not the conclusion of the process and its rightness, but whether the correct procedure has been followed.[741] The assessment of the procedure must follow the principles laid down in the landmark cases *Wednesbury* and *GCHQ*, whereby courts analyse the 'illegality' (unlawfulness), 'irrationality' (unreasonableness) and the 'procedural impropriety' (unfairness) of the administrative proceedings.[742] Arguably, the more relevant (and controversial) basis for challenging a regulatory decision is unreasonableness, so this section will mainly focus upon this particular ground.[743] Generally, irrationality means rejection of 'absolute or unfettered discretion'.[744] That is, a court may intervene and quash a decision where an authority: acted outside its powers; took into account factors that ought not to have been taken into account; or failed to take into account factors that ought to have been; or no reasonable authority would have ever considered imposing the decision.

For a long time courts have tried to draw the frontiers of judicial review. This is somewhat a 'battle for the soul' of judicial scrutiny between supporters of further judicial involvement in

---

[741] See <http://www.judiciary.gov.uk/you-and-the-judiciary/judicial-review> (accessed 02.08.10).

[742] *Associated Provincial Picture Houses v Wednesbury Corporation* ([1947] 2 All ER 680) (per Lord Greene MR) set down the standard of judicial review of public body decisions in a tautology: the test asks whether a decision is 'so unreasonable that no reasonable authority could ever have come to it'. *Wednesbury* synthesised many years of UK courts decisions (in this sense Craig, 2008: 645, n. 137). In *GCHQ* ([1984] 3 All ER 935), which constitutes the basis of the modern doctrine of judicial review, Lord Diplock indicated that 'one can conveniently classify under three heads the grounds on which administrative action is subject to control by judicial review. The first ground I would call "illegality" [unlawfulness], the second "irrationality" [unreasonableness] and the third "procedural impropriety" [unfairness]'. Recently, the approach has been confirmed, for example, in *R (Corner House Research and another) v Director of Serious Fraud Office* [2009] 1 AC 756.

[743] For examples of challenges on grounds of illegality, see *e.g. BAA v Competition Commission* [2009] CAT 35, and *The Queen v The Gas and Electricity Markets Authority* [2008] EWHC 1415 (Admin). An example on grounds of procedural unfairness is *Interbrew*, the first case lost by the CC (*Interbrew S.A. v Competition Commission* [2001] EWHC Admin 367).

[744] Wade and Forsyth (2004: 354).

regulating the regulators, and those who hold less expansionary views.[745] The most supported view simply indicates that supervisory jurisdiction translates into reviews, not appeals.[746] Therefore, the intensity of the review does not encompass a decision on the merits of the case. As CRAIG has indicated, courts are generally limited in their possibilities of interventions over administrative discretion: 'it is *not* for the courts to *substitute their choice* as to how discretion ought to have been exercised for that of the administrative authority'.[747] Judicial control has limits, either demanded by 'basic conceptions of political theory and the allocation of governmental functions'[748] or based upon practical considerations.[749] However, these statements do not say anything about the actual content of judicial review. If merit reviews allow the widest scrutiny, defining judicial review as *not* a review of the merits does not provide any criteria as to where to draw the frontier of judicial review.

The 'classic' version of the judicial review test applied a very restrictive view of scrutiny.[750] This is especially apparent in the case of irrationality. Despite the fact that the concept is essentially the same as applied in American law, British judges traditionally did not use it to the same extent or with as broad a scope. It merely 'tend[ed] to be defined only in relation to logic, not in relation to the evidence deployed or the conclusions that might reasonably be drawn from the evidence'.[751] Judicial review was essentially a control of procedures.

Indeed, the same criteria were applied, and still apply, to economic regulatory decisions. Courts have systematically affirmed that, provided the decision-making process has been

[745] Poole (2009) (explaining the 'structural and fundamental' reconfiguration of judicial review and English administrative law). This is also one of the main topics in US administrative law. See generally Breyer *et al.* (2002).

[746] E.g., *R (on the application of exoteric Gas Solutions Ltd) v Gas & Electricity Market Authority (OFGEM)* [2003] EWHC 2072, at para. 83: 'that these are judicial review proceedings. This is not a Court of Appeal from OFGEM's decision on the merits. It is an expert body, upon which Parliament has deliberately conferred a broad discretion. Its judgment that investigation rather than immediate enforcement was appropriate was well within the bounds of that broad discretion'.

[747] Craig (2008: 613) (emphasis in the original). Similarly, Wade and Forsyth (2004: 362): 'the court must not usurp the discretion of the public authority which Parliament appointed to take the decision'.

[748] Craig, (2008: 613).

[749] Ogus (1994: 117) argues that a cautious approach to judicial review might not be inappropriate, because 'the administrative costs of regulation may escalate and private interests will have an incentive to exploit the process for tactical purposes'. Economists have gone even farther, criticising the 'excess of legalism' of judicial control (!). E.g., Newbery (1999: 401) (stating that 'judicial review is slow [and]… it can produce rather perverse and legalistic findings').

[750] Harlow & Rawlings (2009: 99) (stating that the test was born in a context where 'judges were concerned to avoid accusations of meddling in policy').

[751] Foster (1992: 273).

robust enough, they are 'slow to interfere' with the specialised judgement of the regulator.[752] The test mainly limits the inquiry to whether the conclusions reached by the agency are adequately supported by evidence.[753] In general, judges have allowed regulators a relatively wide margin of discretion, on the basis of experience and expertise.[754] They have declared, for example, that it is entirely for the regulatory agency to attribute to the relevant considerations such weight as it thinks fit.[755] Furthermore, the regulator may not necessarily require giving *weight* to one factor if its only statutory obligation is to *have regard* to it – i.e., to consider it.[756] Overall, judges' own opinions cannot replace those of the administrative decision-maker unless the decision is truly outrageous.

Therefore, even though privatisation judicial institutions have always exerted some control on the rationality of the sector specific economic regulators' decisions, the approach has *prima facie* been one of deference. There are two main consequences.[757] First, regulatory

[752] There is now a well-established case-law: *Hutchison 3G UK Ltd v The Office of Communications* [2009] EWCA Civ 683, at para. 107 (per Lord Justice Etherton); *Royal Mail Group v Postal Services Commission* (CA) [2008] EWCA Civ 33; *Nichols v Security Industry Authority* [2006] EWHC 1792, at para. 87 (per Kenneth Parker QC); *R (Great North Eastern Railway Ltd) v Rail Regulator* [2006] EWHC 1942; *Fisher v English Nature* [2003] EWHC 1599 Admin [2004] 1 WLR 503 Civ 663, at para. 46 (per Lightman J); *R v Director General of Telecommunications ex parte Cellcom* [1999] ECC 314.

[753] That is, UK courts effectuate a control of the rationality of the regulator's decisions which in broad terms is similar to that performed by the Court of Justice in *Tetra Laval*: case C-12/03, *P Commission v Tetra Laval*, [2005] ECR I-987. See Graham (2009).

[754] E.g. *R (on the application of) Centro v The Secretary of State for Transport* [2007] EWHC 2729 (Admin); [2008] ACD 19, at 36: 'If this… is analysed as based on irrationality the claimant has to overcome a high threshold. This is because the issues for decision concerned the application of complex economic concepts in particular the elasticities applied to price increases to be used as part of the reimbursement rate paid to transport operators providing travel concessions. It is clear that, when considering decisions of this nature in the context of judicial review, the court is particularly cautious and reluctant to intervene…'. Similar reluctance is seen in the context of the review of the decisions of economic regulators. In *R (London and Continental Stations and Property Ltd) v The Rail Regulator* ([2003] EWHC 2607 Admin, at 32) Moses J stated that 'it must be borne in mind that the regulator was concerned with issues of economic policy and of economic theory and practice'. In that case the regulator was concerned with determining a method of compensating the operator for its loss of business in the future. Moses J stated that there was no way in which such damage could be measured with any exact precision, even after the event. He also stated that in considering the various challenges to the regulator's directions, the court must 'bear in mind that he was reaching his conclusions in a field in which he was both expert and experienced. He was advised by experts'. Moses J stated that 'these factors demonstrate that the constraining role of the Courts is modest' (*Ibid*, at 36).

Similarly, lack of expertise is also primarily brought into play in explaining why courts are willing to accept fresh expert evidence when dealing with complex matters. See Lidbetter & Roberts (2004).

[755] *Tesco v Secretary of State for the Environment*, [1995] 1 WLR 759, at 764 G-H (per Lord Keith).

[756] *Ibid*, at 784 (per Lord Hoffman): '[I]f the decision to give that consideration no weight is based on rational planning grounds, then the planning authority is entitled to ignore it'.

[757] These specific arguments are in addition to other criticisms made to deference in more general terms. Allan, for example (2006: 675), considers that deference is 'either empty or pernicious', because it does not contribute to implement the already secured separation of powers between courts and other branches of

consistency may be affected due to the *ad-hoc* nature of the decisions. In fact, that was arguably the case during the first years following privatisation: consistency decreased and the system possibly deteriorated. Each regulator was free to pursue the objectives in the manner it deemed suitable and challenges to the alleged reasons of decisions were largely unproductive due to courts' inaction. As FOSTER explains,

> [C]ourts did not use their review powers to interpret the laws passed by Parliament; they exposed the inconsistencies between successive pieces of legislation, but left to Parliament resolve them. Because Parliament did not do so, and given its practical powers could not have been expected to do so, this inaction of the courts was a powerful reason why that system of regulation deteriorated rather than improved.[758]

Most importantly, the 'judicial threat' may lead to over-emphasise procedural formality in the promulgation of regulatory policies. Indeed, it could be argued that the focus on process is not so much a peril (perhaps even an advantage) if it is considered on a stand-alone basis. If that were true, the deferential approach would bear no danger. However, the preference for procedural formality may well come at the expense of substance if both are assumed substitutes.[759] A procedural, formal standard of review invites regulators not to give reasons for their decisions, but merely *state* them in order to prevent challenges to their decisions.[760] There is no incentive (least a duty) to explain the decision in detail or provide clear criteria. So the focus is on regulatory procedures. The more robust the process to reach the decision is, the stronger the position of the regulator when facing judicial scrutiny. Also, regardless the substantive context, the emphasis on procedures may be a source of inefficient delays.[761]

---

government, or it permits the abdication of judicial responsibility in favour of reliance on the alleged expertise, good faith or good sense of public officials.

[758] *Ibid.*, at 281.

[759] See Stephenson (2006: 530 note 2) ('…increasing agency procedural formality decreases the marginal benefit to the court of a more stringent standard of substantive review, while an agency's decision to decrease its level of procedural formality increases the marginal benefit to the court of a more stringent substantive standard'). Similarly, see Balwin & McCrudden (1987: 60-3) (on potential adverse consequences of judicial review).

[760] See Veljanovski (1991: 17) (stating that 'English administrative law acts as a positive inducement for regulatory agencies *not* to state clear criteria nor to give reasons for their decisions, thus preventing the courts from reviewing their actions'; emphasis in the original).

[761] It is not hard to find examples in regulatory practice. A notable one is the increased recourse to 'minded to' consultations, whereby regulators 'announce' their most likely position previous to the issuance of the final decision.

Furthermore, it must be considered that a central procedural drawback of the UK regulatory regime (still characterized by the use of price controls) is that it does not reduce the burden of regulatory proceedings.[762] The frequency of regulatory controls (normally every five years) demands intense and highly intrusive regulatory processes, given that the stakes are much higher. Price controls usually start two or three years before the next period and they are normally open and consultative. Also, price controls are usually complemented by controls on quality and the introduction of incentive schemes for regulatory targets. All of this results in highly complex and discretionary administrative procedures, with enormous coverage, which require cumbersome informational factual-specific details needed by the regulator to fulfil its tasks. Therefore, the UK system seems to demand as many judicial checks on substance as (or arguably more than) any other jurisdiction – a task that a low standard of review barely helps to carry out.

2. Judicial Review in economic regulation

The CAT somewhat 'inherited' all of the aforementioned principles when it received its statutory mandate to apply 'the same principles as would be applied by a court'.[763] The same restrictive standard traditionally used in judicial review was supposed to be applied to regulatory disputes. This presents two questions.

The first question is to what extent advances made in a context different than economic regulation should exert decisive influence in this subject. There seems to be an internal contradiction within the design. On the one hand, specialisation leads to a rather high degree of judicial discretion. On the other, the 'extension' of the traditional standard to regulatory matters runs in the opposite direction towards more discretion: knowing the well-established content of the standard, it seems plausible to infer that the parliamentary option for it had the implicit assumption that there is no greater need for specialisation in the matters subject to

[762] Still price-cap regulation is simpler to operate than other schemes (particularly cost-of-service). Whilst there is an intensive period of administrative work before the new price-cap period comes into effect, once the level of the cap is determined regulation is limited to the relatively simple task of checking that price changes prepared by the firm satisfy the regulatory standard. However, as it is noted in the main text, the regulatory burden may increase.

[763] See generally Kennelly (2006).

judicial review by the CAT. If that was the case, it is an odd way to design the regime – the suit does not seem to fit the institutional body. Indeed, this may have been simply an unintended effect. Another possibility is that it was expected the frontier would be sufficiently ample to include something more than the traditional standard. In fact, by 2002 there already were important advances in other areas that extended the remit of judicial review.[764] However, in this case the question remains the same: if the scope of judicial review was enlarged because of the substantive developments in other areas of law, why should those advances dictate the approach to economic regulation (and competition law)? Ideally, the standard of scrutiny should be in accordance with the substantive law that it is supposed to control. In the case of economic regulation, the introduction of specialisation should have been matched with an appropriate wide standard of review (as in the case of telecoms); otherwise, traditional courts provided were institutionally sufficient.

This is not a *per se* case in favour of a broad standard of review in economic regulation. It may well have been the case that the designers of the regime opted for a narrow standard of review – as the case seems to have been. But if they believed that economic regulation was a matter for which generalist judges were not fully prepared to deal with, they should have avoided to restrain specialists by forcing them to consider how a generalist court would have judged. This claim is justified on grounds of consistency of the regulatory design. The essential distinction between generalisation and specialisation may be blurred if the specialised tribunal is compelled to follow the same standard of review that the non-specialised court applies. In both cases the regulator will anticipate the result, avoid giving reasons for its decision, and instead focus its efforts on having more robust procedures. Ultimately, as seen in chapter IV, the consequence is an increase in regulatory risk.[765]

The second question (which is different than, but underlies the previous one) is whether economic regulation deserves a deeper scrutiny. So it seems the CAT has thought. Early after its creation the CAT made a bold attempt to broaden the scope of its own powers applying judicial review in a more flexible way – that is, giving it a form slightly similar to a merit

[764] See *infra* section IV.B.2.
[765] See the reference regulatory objectives, *supra* chapter II, section I.A.

review.[766] Its arguments were precisely that its own expertise, flexibility of procedures and capacities did not match the restrictive traditional approach to judicial review. The Court of Appeal, however, readily rejected the endeavour, despite recognising that the CAT 'was right to observe that its approach should reflect the "specific context" in which it had been created as specialized tribunal'.[767] In later cases, then, the CAT has acknowledged that it is bound to deference. In *BSkyB* it denied the claims made by the applicant that the CAT's own 'degree of familiarity with the statutory regime, the relevant case-law and some of the legal and economic concepts which arise', all of which would allow it to 'offer a "faster and less expensive route to justice"', meant that it must behave less deferentially.[768] In the CAT's view:

> [N]one of this means that the Tribunal is applying judicial review principles in a different way or is exercising a higher intensity of review than would be the case if the matter were before the Administrative Court. Further, by no means all of the findings which may be the subject of a section 120 challenge are such as would necessarily call for expertise in competition law and practice.[769]

The Court of Appeal confirmed the judgement in this part and also rejected the expanded approach. In the court's view, the specialisation feature does not alter the nature of the tribunal's task.[770] Judicial review was proudly reconciled with the traditional view. Actually, it is not too fanciful to think that the CAT's wording on *BSkyB* was carefully drafted merely to avoid being overturned by the Court of Appeal's on the same matter.

---

[766] See *IBA Health Ltd v OFT* [2003] CAT 27, where the CAT suggested to 'readover' the standard of judicial review (para. 220).

[767] *IBA Health* [2004] EWCA Civ 142, [2004] 4 All ER 1103, at para. 90. The court also held that 'notwithstanding the [CAT]'s specialised composition, the [judicial] review was not to take the form of an appeal on the merits, but was limited to the ordinary principles applied in the Administrative Court' (*Ibid*, at para. 88-9, per Carnwath LJ).

[768] *British Sky Broadcasting Group plc v Competition Commission and the Secretary of State for Business, Enterprise and Regulatory Reform* [2008] CAT 25, at para. 61.

[769] *Ibid*, at para. 62.

[770] *British Sky Broadcasting Group plc v Competition Commission and another*, [2010] EWCA Civ 2, [2010] WLR (D) 5, at para. 37 ('… the Tribunal is to apply the normal principles of judicial review, in dealing with a question which is not different from that which would face a court dealing with the same subject-matter. It will apply its own specialised knowledge and experience, which enables it to perform its task with a better understanding, and more efficiently. The possession of that knowledge and experience does not in any way alter the nature of the task.').

The question for the level of the standard has also arisen in the context of proportionality – a second notion related to judicial review. The CAT has explicitly applied the proportionality test to cases dealing with discretion. A recent illustration is *Tesco*. The CAT agreed with the Competition Commission that the adoption or recommendation of any remedy must satisfy the 'proportionality principles', which cannot be divorced from the statutory context and the framework under which the remedy is imposed.[771] The CAT even recommended to the Competition Commission the application of a 'double proportionality' approach: 'for example, the more important a particular factor seems likely to be in the overall proportionality assessment, or the more intrusive, uncertain in its effect, or wide-reaching a proposed remedy is likely to prove, the more detailed or deeper the investigation of the factor in question may need to be'.[772] This goes well beyond simple statements such as 'a penalty must be proportional to the offence' used in some traditional cases.[773]

Nevertheless, reflecting the inner struggle, the own CAT seems to have downplayed the significance of the double proportionality test.[774] In *Barclays* it clarified that double proportionality is not a new legal principle but 'simply a convenient label for the common sense proposition that, within a wide margin of appreciation, the depth and sophistication of analysis called for in relation to any particular relevant aspect of the inquiry needs to be tailored to the importance or gravity of the issue within the general context of the [Competition] Commission's task'. Nonetheless, the statement still goes some way beyond deference. For example, references to the depth and sophistication of the analysis imply recognition of the particular nature of the issues at stake. This is in line with the nature of the CAT. Specialist tribunals are likely to accept that certain factors or principles embedded in statutes may be given *particular* weight. And even if these tribunals are not assessing the 'correctness' of the weight given to certain principles by the regulators, the appraisal will probably have a wide margin of appreciation. Proportionality applied by specialist tribunals might be gunpowder ready to blow the frontiers of the traditional deferential approach to judicial review.

[771] *Tesco Plc v Competition Commission*, [2009] CAT 6, at 135.
[772] *Ibid.*, at 139.
[773] E.g., *R. v Barnsley MBC Ex p. Hook* [1976] 1 WLR 1052, 1057 (per Lord Denning M.R.).
[774] *Barclays Bank Plc v Competition Commission* [2009] CAT 27, at para. 21.

* * * * * * *

In sum, the typical judicial answer to a challenge on economic regulatory matters declined to assess questions of facts and/or policy. Traditional principles of judicial review call for a standard bias towards deference, unless the decision is truly outrageous. The broader 'conceptual matrix' of the English system of judicial review and its core formal concerns relating to the examination of power and procedures, which allowed little room for substantive review[775], was extended and applied to economic regulation. There is little room for de- and re-constructing the decision-making process – even though the judicial scrutiny is now carried out by a specialist tribunal, which incentive leads it precisely to apply a less deferential approach.

IV. THE WAY FORWARD

A. Taking stock

At this stage, it is sensible to take stock and summarise the main insights presented heretofore, along with key questions they present.

The first insight is that current institutional design is rather messy. The current design has been the product of historical developments that do not follow a unique philosophy – if they follow one at all. At most, the driver of changes has been the need to accommodate institutions on an *ad-hoc* basis to the transformation of UK and European competition law. This has resulted in a tangled web of bodies with occasionally overlapped competences. But it seems clear that the design of the judicial institutional structure for economic regulation has developed without having the regulatory perspective in mind. Indeed, initially the design was influenced by the strong confidence in the 'Austrian' view of competition that dominated the early days of the privatisation process. Provided that competition was at the core of regulatory policy, regulators were created to be the image of the competition authority (the OFT and its predecessor) and it seemed reasonable to place them under a similar institutional

[775] Poole (2009: 143).

structure. Allegedly poor judicial understanding of regulatory matters led the designers to try to keep courts as far as possible from the regulatory system, boosting regulatory discretion. The only 'supervision' was eventually given by the CC – under a rather odd relationship with sectoral regulators similar to that it had with the OFT. Even though the Austrian vision was increasingly abandoned as the process advanced, the objective of pursuing competition as main task still dominated the institutional design. When the system of competition was brought into line with the European regime, regulators were placed again under similar oversight to the OFT. This time, a more specialised tribunal was thought to increase the possibilities of advance competition. Only the need to comply with other European legislation would create a slight differentiation for telecommunications (on an ad-hoc basis) that would eventually expose the limits of the institutional model.

There are two interrelated problems arising from the first insight. The first is whether the current design is the most suitable to pursue the ultimate goal of economic regulation: to attract and protect foreign investment. On the other hand, a distinctive feature of the current design is the presence of a specialised tribunal that has jurisdictions in parallel or alternatively to non-specialist courts. How does this feature affect the scrutiny? As seen, the proposed answer is that, generally speaking, specialisation incentivises a 'hard look' approach and generalisation leads to 'light touch' approach.

The second insight is that there has been an enlargement in the number of regulatory objectives. Competition, which was the first and foremost objective to pursue at the time of privatisation, has been coupled with a number of other goals, the importance of which equals that of competition and on occasions surpasses it. The question that arises is now normative: what are the most desirable characteristics of the institution that will deal with regulatory decisions in the actual context of plurality? Arguably, the enlargement of objectives has decreased the capabilities of generalist courts to scrutinise regulatory decisions. Regulators, as policymakers, are tasked with choosing the objective they deem worthy of protection or advance in each case, and have the burden to demonstrate why a specific objective was chosen. In economic regulation, most of the reasons are not principled, but adopted for technical motives or any other reason generally linked to expertise – which can also vary in degree. There is a strong case to make that specialised courts are a better place to deal with

complex regulatory decisions that respond to a variety of objectives. However, such a task cannot properly be done if the specialised body cannot determine the standard of review, but this is imposed externally – a problem that leads to the last claim.

The third insight is related to the standard of review. This insight has two normative claims. The first one is related to the incoherence of the diverse arrangements to challenge regulatory decisions embedded in the system. The main consequence of the actual design is the dissimilar treatment of similar issues, with important differences both in the scope of the scrutiny and the availability of some specific arrangements.[776] In practice, for example, price controls are subject to two different regimes. On the one hand, most regulatory decisions in utilities sectors are subject to an in-depth revision by the CC, the outcome of which can be subject to judicial review by the CAT. On the other hand, in the specific case of telecommunications, the decision is subject to merit review by the CAT, which must refer to the CC before disposing of the case. Why should a revision of a market investigation demand less profound scrutiny than, say, a price control inquiry? Undoubtedly, the first one may even be more burdensome.

The second claim related with the standard of review is related to the body that set the standard in practice. Both standards of review apply to different specific situations by express mandate of the respective statute. Whilst the CAT determines the content of merit reviews, it has to apply judicial review in the exact manner as generalist courts would apply it. This means that in practice the courts set the content of the standard applicable to economic regulatory decisions and, oddly enough, that content depends on the developments of other areas of law different from regulation. Unlike Humpty Dumpty, quoted at the beginning of this chapter, when the CAT uses a word, it does not necessarily mean just what it chooses it to mean.

The general point is whether it is justifiable that the design makes two regimes comparable even though they respond to different underlying objectives. There seems to be a tendency in

[776] See Prosser (2006) (acknowledging the 'mish-mash' produced by the 'maze of appeals'). Similarly, Rose & Richards (2010) (concluding that the various different appeal and review mechanisms created by statute are unnecessarily complicated, and lead to increasing costs and delay).

the UK regulatory design to subject institutions underpinned by unequal goals alike. Indeed, in some cases uniformity may be justified on grounds of consistency. For example, when the Competition Act 1998 introduced concurrent powers, the rationale was regulators would apply competition prohibitions in the same way as the competition agency (the OFT). Concurrency avoids 'turf wars' at the boundary of each jurisdiction and avoids potential gaps in the application of competition law arising from partial definitions applicable only to one specific field.[777] Although arguably justified, indisputably concurrent powers 'do not reflect a major acceptance that utilities industries are different from others'.[778]

The same applies in particular to the CAT and the standard of review. But unlike the previous case, uniformity seems a bit exaggerated. It is debatable whether a general judicial institution should impose a standard of review to another judicial institution – especially in a context that requires a high level of specialisation. As will be explained, there have been crucial changes underpinning the transformation of the traditional approach to judicial review in areas such as human rights and EU law. These important changes, that have broadened the standard of review, are all *substantive* in nature. They reflect particular advances in a given area of law and those advances are in turn reflected in the judicial approach. So far, however, the standard of review remains narrower than what specialisation incentivises and regulatory law requires. In practice, the (legal) differentiation of standards may well become a distorting factor that prevents the CAT from carrying out an accurate revision of the case, distorting the outcomes and thus affecting substantive policy. If this argument were true, there are strong grounds to get rid of the frontier and find a more suitable alternative.

There is a crucial tension embedded in the own institutional design of the CAT. Specialisation, on the one hand, presses against more deferential forces that run parallel to it, on the other. The dominance of the latter, though, threatens to leave judicial scrutiny isolated from the substantial developments of regulatory law (or even competition law, again, one of the main targets of judicial protection in this area). To repeat: the CAT's institutional design has two forces that exert pressure in opposite directions. On the one hand, specialisation demands a highly intense scrutiny of all parts of the regulatory decision. On the other, the

[777] Pimlott (2010: 163).
[778] Prosser (2005a: 54) (justifying the difference on utilities' public service role).

statutory boundaries and the limitations imposed by the courts on judicial review proceedings push towards the other side. The result is an inner-struggle reflected in the judgments whereby non-deference comes and goes – like the high and low tides produced by the waves of the sea. The current institutional design does not recognize that 'the degree of deference which the courts should show will, of course, depend on and vary *with the context*'.[779] The standard of review should ideally reflect mainly changes in the substantive approaches to competition and economic regulation, not in other areas of law. Currently, there is an erroneous mismatch between substance and judicial revision.

B. Looking ahead: a unified approach?

Looking ahead, there is a first straightforward step to take to make the system more coherent: to abolish the CC. As seen, in practice two different bodies alternatively carry out the scrutiny of regulatory decisions. In telecommunications, the CAT has jurisdiction to assess the merits of some issues related to price controls; in other sectors, the CC is entitled to make the assessment. This differentiation is not justified. If the model of scrutiny expressly opted for a judicial one, the role of the CC in economic regulation (and arguably also in competition law) is superfluous. Recent proposals of the government that aim to change the institutional design and give it a more thorough structure have auspiciously acknowledged this need.[780]

The second step may be more controversial. It is submitted that there is a strong case for a change in the standard of review applicable to economic regulation and competition law. The change is nonetheless simple: to expand the frontiers of judicial review. On either side of the spectrum there is an area of scrutiny – namely, judicial review and merit review, respectively – whose content has been increasingly slackened and, in reality, enlarged. At midpoint there

[779] Lord Steyn (2005: 352) (emphasis added). See also King (2008: 440) (stating that 'The key is that any reference to principles of restraint should remain contextual'); and Rivers (2006: 177) (arguing that 'the correct intensity of review should be set by the seriousness of right-infringement in the case at hand').

[780] The UK Government announced the plan on 14 October 2010. The aim is to create a single agency called 'Competition and Markets Authority' (CMA). At March 2011 the consultation has not be unveiled. The plan also intends to make other important changes in economic regulation sectors. For example, Postcomm (the postal services regulator) will be merged with Ofcom (the telecoms regulator).

is a grey frontier that, as a consequence of the broadening of what were once rather clear opposite standards, has become thin and blurred. The question is whether the frontier remains important in a context of flexible standards. As it is now, that frontier can be easily trespassed. If this is the case, either the mere existence of the frontier, as it is now, is superfluous; or there is a case to make it stronger. Since the latter solution implies to step back in the development of standards, the former is the most logical consequence.

Ideally, the standard of review should be integrated – that is, it should be a unified (not uniform) but flexible standard whose intensity is determined by the own CAT according to the development of the regulatory objectives it must protect. Flexibility calls for a variable intensity of review that would allow assessing with more accuracy different levels of regulatory expertise applied by the regulator. It would also accommodate better with the revision of the reasons alleged by the regulator to choose one regulatory objective over others. As a result, the adoption of an integrated approach would benefit the overall quality of the scrutiny. Whilst developing their function, the CAT may help to surmount legislative inertia and confer a necessary degree of dynamism to the regulatory process.[781]

What matters most is to delimit the boundary between being a primary decision-maker and being a body that scrutinises regulators' decisions. It is important to recall that a broader standard does not mean judgment substitution. 'Management' of the appeal mechanism therefore becomes the important issue.[782] There *is* a difference between oversight and administrative decision-making. But it is uncertain whether the legislator is capable enough to find that limit and express it correctly in a legal rule. Judges are much better placed to do so – particularly in areas that require specialisation. As long as specialist judges themselves delimit the boundaries of their own competences (and, as stressed below, they do), or they remain subject to effective oversight (and, as also stressed below, they are), the standard of scrutiny can be enlarged as much as to allow the reviewer to consider the merits of a case – again, without incurring in judgment substitution.

[781] See Sunstein (1990a: 111): 'Indeed, legislative reform must overcome an enormous burden of inertia. It is through interpretation, in the courts and the executive branch, that regulatory improvements, interstitial to be sure, can be brought about most easily'.

[782] House of Lords (2003-4a: para. 225).

Hence what should be clarified is that the principles of judicial review the CAT will apply are not necessarily the *same* as those a court would apply. The CAT needs to be free to apply the principles it deems appropriate within the remit of its own jurisdiction. With that caveat, the case for a unified, broad standard of review is strong. There are two ways to implement the change: one that requires a statutory change and one that does not require statutory change.

1. The dramatic way: broadening the standard through statutory change

The first one is through a statutory modification that eliminates the need for the CAT to apply 'the same principles as would be applied by a court on an application for judicial review'. Leaving aside the intrinsic difficulty to carry out legislative changes, there are two possible objections to this argument. First, it can be argued that parties would be left with no guidance on what the applicable legal test would be. Certainty would be undermined. Moreover, specialised judicial discretion would arguably be as dangerous as the regulatory discretion is supposed to be controlling. The second argument, related to the first one, is that in case the CAT could give some guidance, it would be – so to speak – 'reinventing the wheel'. The CAT would have more freedom to apply whatever test it deems suitable to the actual case, on an *ad-hoc* basis, even paying no attention to the historical developments on judicial review. However, these concerns are simply speculative. First, it has been shown that the CAT has given guidance on how its approach would be. Secondly, the composition of the CAT, which mixes judicial and non-judicial members, creates a strong incentive not to deviate from judicial tendencies more than what eventually is necessary for the resolution of the conflict at issue.

In fact, the UK government recently seems to be advancing toward supporting a statutory change – albeit only a partial one.[783] Responding to proposed changes to European Legislation and the possible increase in regulatory uncertainty caused by these changes, the government is proposing modifications to the Communications Act 2003. In particular, it is proposing that an 'enhanced' judicial review would be capable of complying with the European requirement of having a proper 'right to appeal'. This would be a narrower form of

[783] See DBIS (2010).

review than the actual review of the merits applicable in some telecommunications cases. Note that the government is approaching a similar solution as the one proposed in this chapter from the other end. Its concern is that merit reviews have become too burdensome and unmanageable for the CAT.

Importantly, however, the government recognises that judicial review may have the sufficient flexibility *to take the merits of the case duly into account*.[784] In this sense, it has said:

> [W]e believe an effective appeal should, as a minimum, consider whether the Regulator acted lawfully, and followed the correct procedures, took relevant issues and evidence duly into account and generally acted in accordance with their statutory duties. In considering these issues, it should take the merits of the case into account.[785]

The intention to make the change is positive and goes in line with the developments of judicial review in economic regulation and other areas. Indeed, the devil is in the details, so it seems prudent to wait until the idea actually becomes a law. However, as a project it already falls short of providing a full answer to the standard of review.[786] On the one hand, it leaves other areas untouched, so the existing 'maze of appeals' would remain in place.[787]

On the other hand, the proposal does not delimit what an enhanced judicial review would be. An enhanced judicial review could consider a number of relevant aspects that the narrow approach does not currently consider. Let's consider some hypothetical examples. Enhanced judicial review could include the ability of courts (or tribunals) to order *access* to certain documentation. However, this seems unlikely in the actual context of judicial review, since so far courts have not accepted a full right of discovery. Likewise, it seems improbable that courts would allow enhanced judicial review to cover an assessment of the *evidence* used by the regulator. Perhaps an assessment of the *reasons* of the regulator might be included. So far, this is an aspect that has only occasionally been considered, but a specialised judiciary is certainly well-positioned to make this requirement. An eventual obligation to present

[784] *Ibid.*, at 18, para. 59.
[785] *Ibid.*, at 17, para. 55.
[786] In fact, it is the second time the government rejects a general appeal right on the merits. In 2004 the House of Lords Constitution Committee had already unsuccessfully proposed such an approach. See HL (2003-4a: paras. 219-32), and (2003-4b: paras. 60-73).
[787] Prosser (2006: 196).

evidence is, however, a stronger step than the need to give reasons, because it forces the regulator to support its reasoning with substantive evidence. Courts will surely be reluctant to extend the review considering that there are no open hearings involving the participation of affected interests in regulatory processes and that such a requirement risks to produce 'overjudicialisation'.[788]

In addition, enhanced judicial review might extend not only to the scrutiny of the rationality of the regulatory decision, as it usually does, but also reach the *relevancy* of the decision for the objective sought – which includes an evaluation of both the considerations taken into account and their purpose.[789] These aspects seem to have more prospects of success than previous ones, as they are already included in traditional judicial review. Nonetheless, relevancy has not been greatly used in practice. The 'extension' is appropriate and straightforward.

Finally, enhanced judicial review could go beyond the consideration of questions of law and include *questions of facts* as well. Indeed, would this distinction be inexistent, there be no question in applying judicial review to these categories. For a long time now, some commentators have argued in favour of this position, claiming that the distinction between questions of law and fact is open to manipulation in practice.[790] Judges would characterize an issue as one of law when they sought to exercise control. Conversely, the same issue may be characterised as one of fact when exerting no control is the preferred option. But (the argument goes) the terms may also be conceptually inseparable, so the distinction would be misleading. There would be a number of 'mixed questions' of law and fact.[791] Courts would normally accept revision of facts in recognition that regulators' factual decisions may well be reflecting a judgement on policy. Having no revision at all on whether the evidence is right or

[788] Certainly, there have been great advances in openness, including the occasional use of public hearings and the issuance of extensive consultations, but they have been more sectoral developments rather than a thorough mandatory and structured regulatory procedure. However, there seems to be a far more coherent procedural regime in place now.

[789] Recall that *Wednesbury* about three different limbs, two of them being closely related to relevancy: the body failed to take into account factors that ought to have been taken into account or, in making the decision, took into account factors that ought not to have been taken into account.

[790] See e.g. Wilson (1969: 376) (equating questions of law to questions of 'degree' and affirming that 'it may be that questions of degree are classified as questions of fact merely as a matter of procedural convenience'). In American law, this was the (recurrent) argument of some adherents of legal realism.

[791] On the problems of classifying such mixed questions, see generally Levin (1985).

the inferences drawn from it comply with the legal standard may result in an alteration of the legal rules. Therefore, the aim of applying a separate, different standard of review to questions of law and questions of facts seems, according to the supporters of this position, particularly difficult and illusory – especially (although not exclusively) if the regulator has not divided its decision in both parts. The argument, however, does not seem entirely convincing. True, there are a number of 'grey areas' that may be difficult to categorise. However, to jump into the conclusion that the distinction is useless seems exaggerated. It is indeed valuable as an analytical tool, and in practice different standards apply to each category. More than getting rid of the distinction, the point here is that enhanced judicial review could cover questions of law and also questions of fact.

2. Broadening the standard without statutory change

Perhaps surprisingly, a statutory change is not compulsory to broaden the frontier of judicial review. There are two main reasons. First, in practice the standard of judicial review has become 'mouldable' into other areas of law.[792] Secondly, it is not clear whether in economic regulation the frontier between judicial review and merit review remains in place.

The first argument has two interrelated strands. First, as it is largely acknowledged, there have been utterly important advances over time regarding the nature of the *Wednesbury* test to control unreasonableness of regulatory decisions. The test has been 'modified', 'loosened' and increasingly broadened.[793] It now applies to cases where the court estimates that the logic was 'flawed' or the decision-maker misdirected itself in law. There are several cases, particularly in the area of human rights, where English Courts have applied 'anxious scrutiny' in judicial review cases – acting with deference to social, political and economic judgments made by the authorities, but closely scrutinising the facts.[794] In other areas, English courts have not only shown an ever-increasing willingness to permit challenges against

[792] The same reasoning, in a slightly different but close context, was long ago applied by Black (1998: 156) (stating that 'A preferable approach…would be to adopt the principle of rationality in reviewing all interpretations by regulators of their own rules and general statutory provisions').

[793] Le Sueur *et al.*, (2007). The expression 'heightened *Wednesbury*' is commonly used to differentiate irrationality challenges with special treatment.

[794] In the human rights domain confront Young (2009); with Allan (2006) and Edwards (2002).

administrative agencies, but also some willingness to broaden the scope of the review. For instance, the Administrative Court has indicated that although the evaluation of the facts may correspond to another entity (e.g., the regulator), it 'must inquire whether those facts exist, and have been taken into account, whether the judgement has been made upon a proper self-direction as to those facts, whether the judgement has not been made upon other facts which ought not to have been taken into account'.[795] That is, even though factual disputes should not be decided via judicial review, the soundness of the factual foundations can be revised when the court deems that the evidence is not adequate.

A second strand of the general movement of courts towards less deferential approaches has been accurately labelled as 'cross-fertilisation'.[796] Due to the influence of external sources of law, particularly European Law, courts have limited the scope of the regulatory decisions. For instance, in *T-Mobile* the Court of Appeal not only declared that judicial review 'can be suitable adapted when necessary', but that it 'must be adapted' to comply with European Law.[797] This strand pressures for non-deference in two ways. On the one hand, courts have held that the principles of judicial review laid down in European Law are 'stricter' than the traditional unreasonableness test used in English Law.[798] On the other hand, European Law opened judicial control to the potential use of a wide-range of administrative law principles previously of not extensive use in the UK.[799] Foremost amongst these is proportionality. As YOUNG explains, under this test '[t]he court does not merely determine whether the decision of the public authority was "reasonably" or "rational" but must also "assess the balance which the decision-maker has struck", paying attention to "the relevant weight accorded to

[795] *Secretary of State for Education and Science v Tameside MBC* [1977] AC 1014, at 1047 (per Lord Wilberforce).

[796] Rawlings (2010).

[797] *T-Mobile v Office of Communications* [2008] EWCA Civ 1373, [2009] 1 WLR 1565 (per Jacob LJ), at paras. 23 and 29, respectively. Indeed, the court also recognised that the Framework Directive (the European legislation applicable to the case) does not require the court to transform into a 'a fully equipped duplicate regulatory body waiting in the wings just for appeals' (*Ibid.*, at para. 31).

[798] *R (Mabanaft) V. Secretary of State for Energy and Climate Change* [2009] EWCA Civ 224 (per Arden LJ), at para. 30.

[799] The duty of national courts to construe national rules in accordance with relevant directives was first referred by the ECJ in cases 14/83, *von Colson* [1984] ECR 1891, and 79/83 *Dorit Harz* [1984] ECR 1921. See also case 106/89, *Marleasing* [1990] ECR 4135, para. 8, stressing that courts are only obliged to interpret domestic legislation in conformity with Community directives *in so far as it is possible to do so.*

interests and considerations"'.[800] That is, courts might eventually need to go beyond the traditional 'rationality' test to further enquire into the weight attached to the relevant consideration. A compelling argument has been made that any decision must be justified by conformity with the principle of proportionality.[801]

Proportionality entails a greater judicial attenuation of administrative autonomy than does the traditional approach to substantive review.[802] According to this principle, in order to determine whether a decision is justified in law, the courts should be satisfied that the qualification placed on the applicant's right is proportionate to the competing policy objective being pursued. The mere existence of public policy considerations is insufficient and therefore the range of options available to the decision-maker at the substantive level is narrowed. As a result, courts become the main arbitrators of public and private interests. Moreover, they are tasked with ensuring that the regulatory action does not go beyond what is strictly necessary to secure the regulatory objective.[803]

However, the acknowledgment of European influence remains generally reduced; and the recourse to proportionality in economic regulation is limited[804] – despite the fact that, precisely, 'the proportionality principle is most useful for cases of economic regulation'.[805] Notwithstanding the application of this principle may set important limits to the discretion of regulatory agencies, courts have been reluctant to use it over and beyond *Wednesbury* in purely domestic cases (i.e., cases without a European dimension).[806] There are some noteworthy exceptions where courts have reasoned in an analogous manner to, or explicitly applying, proportionality. Regrettably, however, beyond few exceptions the principle generally does not stand as independent standard of review – including to the exercise of

[800] Young (2009: 554-555), quoting *R (Daly) v Secretary of State for the Home Department* [2001] UKHL 26; [2001] 2 WLR 1622.
[801] Alexy (2002).
[802] Elliott (2001).
[803] For an in-deep review of these questions, *see* Arancibia (2011).
[804] See Wilsher (2008).
[805] Sunstein (1990a: 182).
[806] Craig (2008: 622 *et seq.*) See also Craig (1999: 99) (stating that there is no reason why courts should not be able to apply proportionality in pure domestic cases). Obviously, because proportionality is a general principle of European Law, national courts are obliged to apply it in cases with European dimension. The leading authority on the status of proportionality is the House of Lords' decision in *R v Secretary of State for the Home Department Ex p. Brind* [1991] 1 A.C. 696.

discretion by regulatory agencies.[807] The general reasoning is that the overturn of a decision because the court balanced the conflicting interests differently than the agency amounts to substitution of judgement.

Increased judicial control should not undermine, but make stronger the positive incentives of the regulatory regime by settling the rule of law. True, there are compelling reasons to limit the role of the courts and tribunals, some of which are examined in the next part. However, in other domains, courts constantly tread the tightrope of deference / non-deference searching for the right balance. There is no reason why the CAT cannot carry out the same exercise. If the CAT is allowed to broaden the standard of judicial review in the area of economic regulation, the frontier is removed with no need of further reform. This is exactly what has happened, for example, in the area of human rights. What is important, though, is that the CAT is the organism that should determine the standard within the constitutional system. The different nature of the subject should be recognized and not shielded due to the threat of interference. Whatever words courts may use to clothe their fears, the current fabric cannot protect judges from them. A change of suit may result in positive benefits for the entire economic regulatory regime.

The second argument to support an enhanced judicial review without statutory change is that the frontier between judicial review and merit review has, to a large extent, been already blurred in economic regulation. Notwithstanding the reason to give the CAT double jurisdiction, in practice there is no clear-cut division. Indeed, the content of the revision might in some circumstances be the same; regardless the standard the court or tribunal is applying. For instance, on many occasions procedure equates substance. An easy, straightforward example is given by some allegedly procedural rule that prevents the enforcement of a substantive right or makes its exercise excessively difficult.[808]

---

[807] There are cases in which some revisions (especially made by the Competition Commission, CC) are based upon a standard of 'public interest'. For an example, *see* CC, *Vodafone, O2, Orange & T-Mobile: Report on references under section 13 of the Telecommunications Act 1984 on the charges made by Vodafone, O2, Orange and T-Mobile for terminating calls from fixed and mobile networks* (hereinafter, 'termination charges') (2003). The CC has indicated that the duty to decide whether existing licences operate or may be expected to operate contrary to the public interest is 'closely related' to proportionality (*Ibid*, at para. 2.570). However, I will not deal here with that kind of argument.

[808] In European Law there is a long-established case-law recognising this situation. E.g., case C-312/93, *Peterbroeck* [1995] ECR I-4599.

In fact, a number of conducts of the firms can be framed as competition-related infringements (subject to merit review) or as breach of a licence condition (subject to judicial review).[809] The same happens with many challenges to regulatory decisions. The difficulty normally does not go beyond the imagination of the parties. Moreover, it is highly likely that the CAT will look at the rationality test in judicial review cases through the lenses of competition law – in particular considering that the majority of cases it revises are appeals (as seen, telecommunications appeals). Competition infringements, which are normally subject to merit reviews, are much more structured, complex and impose a harder scrutiny.[810] There are a number of requirements that must be complied with, such as the completion of a statement of objections, the disclosure of evidence, the type of evidence that can be presented, etc. If judicial review is looking to have those considerations in mind, the distinction disappears to a large extent, as the starting point will expectedly be a heightened standard.

In practice, more than being close, somewhat antagonistic categories, both types of scrutiny often mingle. A foremost example is the review of the use of the proportionality test by regulators. There are a number of cases in which hierarchy between principles may simply not exist; or cases in which the lack of hierarchy is transitory and applicable only to the specific case.[811] In all these cases proportionality becomes crucial. Even specialist tribunals are less prone to interfere when the regulator has correctly applied the proportionality test. In *T-Mobile*, for example, (a decision where the statute allowed for a review of the merits) the CAT indicated that:

> [T]here may, in relation to any particular dispute, be a number of different approaches which [the regulator] could reasonably adopt in arriving as its determination. There may

[809] Recently the National Audit Office has confirmed this: 'Regulators can usually choose to use either their regulatory powers or their competition powers to achieve the desire outcome' (NAO, 2010: para. 10).

[810] Particularly since the landmark decision of the Court of Justice in *Tetra Laval*: Case C-12/03 P, *Commission v Tetra Laval*, judgment of 15 February 2005.

[811] For instance, the Competition Commission indicated that 'Initially, we were concerned that we should understand whether any hierarchy existed between these various obligations on [the regulator]. However, [the regulator] submitted that it would not be helpful, at least for the purposes of the present appeal, to attempt to set out its duties and objectives in a hierarchical way. [The regulator] submitted that it had to pursue all of its objectives and duties simultaneously, in so far as it was possible for it to so do, and that in relation to the present appeal, there was no conflict between any of [the regulator's] duties and objectives… we accept [the regulator's] submission…' (*E.ON UK plc and GEMA and British Gas Trading Limited*, CC 02/07, 10 July 2007).

> well be no single "right answer" to the dispute. To that extent, the Tribunal may, whilst conducting a merits review of the decision, be slow to overturn a decision which is arrived at by an appropriate methodology even if the dissatisfied party can suggest other ways of approaching the case which would also have been reasonable and which might have resulted in a resolution more favourable to its cause.[812]

The CAT adopted a very similar approach in *Albion* (another case where the statute allowed for a review of the merits). The CAT reasserted its entitlement to form its own view of what it considers is the appropriate decision, but it recognised the range of options that the regulator may have:

> [70] The Tribunal has jurisdiction… to reach its own decision in respect of a matter forming part of the decision under appeal… // [72] We are conscious, however, that in determining the lawfulness of an access price, there may be a number of different approaches which a regulator, exercising its concurrent powers with the OFT, could reasonably adopt in arriving at its decision. There may well be no single "right price"… To that extent, the Tribunal will, whilst still carrying out an assessment of the merits of the case, give due weight to a finding which is arrived at by an appropriate and reliable methodology, even if a dissatisfied party could suggest other ways of approaching the issue which would also have been reasonable and which might have resulted in a resolution more favourable to its case.[813]

Note that in both cases the mere label of merit review does not produce any substantive difference in the assessment. There is no reason why the same kind of appraisal cannot be done in judicial review procedures. In fact, the analysis is not too different than the approach taken in *Barclays* – a judicial review case. The CAT held that

> [S]o far as concerns evidence, the important distinction is between a decision based upon no evidence, with which the Tribunal may interfere, and one based upon the weight given to particular evidence, which is a matter for the Commission, and with which the Tribunal should not interfere, in the absence of irrationality.[814]

---

[812] *T-Mobile (UK) Ltd. and others v Ofcom*, [2008] CAT 12, at para. 82. The need to be slow to interfere when some balance between competing views is necessary has been reasserted in other decisions. In *Vodafone*, for example, the CAT stated that it 'may, depending on particular circumstances, be slower to overturn certain decisions where, as here, there may be a number of different approaches which [the regulator] could reasonably adopt' (*Vodafone*, n 68, para. 46).

[813] *Albion*, CAT unfair pricing judgement, n 71, at paras.70 & 72.

[814] *Barclays*, n 67, at para. 23.

At first glance, in *Barclays* the CAT aimed to limit the scope and arguably the intensity of the review.[815] It seems to look for a 'balanced' test of no substitution that, at the same time, allows it to express greater concern for the provision of reasons and information.[816] In this sense, it considered that the 'fashioning of an effective, reasonable and practicable remedy requires not merely fact-finding about the market as it is, but analysis as to the *probable effect of alternative remedies* upon that market in the future', which 'calls for (*inter alia*) quantification, evaluation and the analysis of causation, sensitivity and risk'.[817]

However, despite the efforts to limit itself, in the same judgement the CAT indicated that '[t]he relevant failing must satisfy a materiality test', which 'will require the finding or decision to be quashed *unless the Tribunal is satisfied* that a reasonable decision-maker in the position of the [Competition] Commission would still have reached the same finding or decision'.[818] In fact, the CAT struck down part of the Competition Commission's analysis of adequate remedies for lack of proportionality and examined in detail the economic methodology employed by the Competition Commission, finding that this was defective and that in conjunction with the other failings of the decision should lead to its quashing.[819] That is, it was not satisfied that a third party would have reached the same conclusion – which is the same as saying that the CAT itself was not satisfied with the path followed by the Competition Commission.

In sum, despite the fact that the principle seems to be that the reassessment of primary facts is *prima facie* not accepted and that the regulatory policy decision is likely to stand unless there is an error in the basis of the judgement, the CAT seems to analyse the way a regulator applies proportionality exactly in the same way either under judicial review or under merit reviews.

---

[815] On this, it was following precedent, perhaps to avoid being overturned by the Court of Appeal: see *supra* note 767 and accompanying text.

[816] This resembles closely the 'hard look' approach of the US courts. See Breyer *et al.*, (2002).

[817] *Barclays*, n 67, at para.26 (emphasis added).

[818] *Ibid*, at para. 28 (emphasis added). In the previous paragraph the CAT bases its argument in *Tetra Laval*, indicating that 'these tasks call for the application of particular care, beyond that necessitated by mere fact-finding' (*Ibid*, at para. 27).

[819] *Ibid*, at paras. 142-175.

All in all, it is hard to think that the possibility the CAT has to make continuous choices (because of its specialisation) will decline for the mere fact that the statute mandates a specific type of revision. Judicial review is already flexible enough to encompass broader scrutiny and in practice, notwithstanding the type of revision, in most occasions there is enough room for the CAT to avoid the unhelpful categorisation as much as possible without engaging in an interventionist approach. In light of the revision's specific object (i.e. economic regulation) and the characteristics of the tribunal (i.e. specialised), a wider, unified but variable standard of review is both necessary and possible.

C. Rebuttals and rejoinders

1. Potential bias towards non-deference, separation of power and accountability

A first counter-argument that might be argued is that, given the CAT's specialisation, the unified approach would almost certainly be highly biased towards non-deference. Conversely, a deferential approach would purport to implement separation of powers. In turn, non-deference would increase judicial discretion in a context of scarce judicial accountability: judges are unelected officials and hence their powers must be restrained.

However, this is but a minor objection to the integrated approach. The argument is cause of no concern if the CAT remains itself adequately restrained, avoiding the peril of unleashed interference. Several features of the institutional design prevent this risk. First, recall that most of the CAT's chairmen and the President are judges of the Chancery Division of the High Court. At least one of them must always be one of the three members that hear the cases.[820] This should act as deterrent to avoid great departures from the common practice of the generalist courts. Secondly, there are principled substantive demands that cannot be overlooked. Proportionality, for instance, must always be respected. Thirdly, there is the simple argument that the risk of open excesses is not exclusive of tribunals, so it does not undermine the need for an integrated approach – at most, it only stresses the need for some

[820] Moreover: interim measures are dealt with either by the President or, in case of urgency, by one of the chairmen sitting alone.

form of hierarchical control. That control in fact exists (and this constitutes a separate argument in itself). The CAT is not the last instance: its judgements are revised by the Court of Appeal or other by way of judicial review. With just reasons, these courts are vigilant and readily available to control open unrestrained behaviour on the part of the tribunal.[821]

Moreover, even if deference is still preferred, ALLAN has convincingly argued that the separation of power is already secure 'by the proper application of legal principles defining the scope of individual rights or the limits of public powers'.[822] The standard of scrutiny does not contribute to the task. In fact, since separation of powers suggests independence of authority, the idea does not reflect the proper nature of the UK administrative system.[823] Furthermore, nor does it reflect the true nature of government activity – least under a governance approach as the one advocated in this thesis. The administrative system is better described as one of 'check and balances' – an idea that suggests intermingled authority.[824] In this sense, a unified approach would serve to better scrutinise regulatory decisions, secure compliance with the law and counterbalance discretion.

2. Decreased role for regulators

Arguably the strongest counter-argument against the application of a unified approach is that the role and importance of regulators may be largely reduced. There is an enormous risk of regulatory inefficiency arising from courts systematically second-guessing decisions made by regulators within the remit of their own competences. If courts constantly quash regulatory decisions, many of the advantages of agency administration may be lost – knowledge, experience (e.g., working with highly detailed statutes), speedy and cheap resolution of problems, availability of wider consultation, and even the possibility that the agency

[821] E.g., *Ofcom v Floe Telecom Ltd* [2006] EWHC Civ 768 (limiting the CAT's powers to impose a timetable on the regulator for re-investigation).

[822] Allan (2006: 675).

[823] Jowell & Oliver (2007: xiii) (stating that 'separation of powers…has never been wholeheartedly embraced in the UK'). Clear separation of the judiciary exists only since 2004, with the passing of the Constitutional Reform Act.

[824] Breyer *et al.* (2002: 37).

implements governmental sector-specific programmes and policies. Regulators risk being transformed into mere fact-finders, gathering evidence for the upstream real decision-maker.

The argument is unsound for two reasons. First, judges – it cannot be forgotten – remain bound to their institutional competences and are often aware of their own institutional limitations.[825] This is the case of the CAT. As Green has stated, 'although the tribunal does possess all the powers of the decision maker…in practice many merit cases turn out to be decided upon quasi-judicial review grounds', because of the need to save resources in fact-finding investigation.[826] As seen, the tribunal is also willing to enhance judicial review wherever necessary. The awareness of its position, nature and boundaries impedes that the CAT pretend to be a covered regulator.

Secondly, in economic regulation, the regulatory framework normally provides in itself for solutions of conflicts of norms – in several forms. First, it is likely that the regulator will have a duty to have regard to principles, or aim to meet objectives, which might be in conflict. Secondly, some 'target duties' may be established within the framework – i.e., duties that are not necessarily required to be achieved immediately. In both cases, the mandate to the regulator is to 'do its best' to achieve the principle.[827] Thirdly, the framework may contain some discretionary area. This is normally reflected in phrases such as 'so far as possible…', 'have regard to the need to / interests of…', 'in the manner best calculated to achieve…' and so on.[828] In all these cases there is a need to attach some precise weight to the principle sought – weight that depends on the specific facts of the case. Therefore, in some cases

[825] Jowell (2003: 598) (stating that 'it is quite appropriate for courts modestly to acknowledge a practical appreciation of their own institutional limitations').

[826] Green (2007: 255).

[827] For instance, both section 3A of the UK Electricity Act (1989) and section 4AA of the Gas Act (1986) indicate that the principal objective of the regulator is to protect the interests of consumers, 'wherever appropriate by promoting effective competition'.

[828] See *e.g.*, article 1(2) of Directive 2009/73/EC of the European Parliament and of the Council, of 13 July 2009, concerning common rules for the internal market in natural gas, [2009] OJ L 211/94 (stating that the rules established by the directive for natural gas shall also apply in a non-discriminatory way to biogas and gas from biomass or other types of gas 'in so far as' such gases can technically and safely be injected into, and transported through, the natural gas system); UK Competition Act 1998, s 38(8) (stating that when setting the amount of a penalty, the OFT must only 'have regard' to its own guidance issued according to section 38(1)), and *Aberdeen Journals Ltd v The Office of Fair Trading* [2003] CAT 11, at para. 472 (confirming that interpretation); UK Water Industry Act 1991, section 2(2) (indicating that the regulator shall exercise and perform its powers and duties 'in the manner that he considers is best calculated' to achieve certain objectives described in the Act), and *Marcic v Thames Water Utilities Ltd* [2003] UKHL 66, at para. 65 (confirming the interpretation).

regulatory inaction may indicate a conscious choice for a 'regulatory holiday' in order to provide the regulated firm with resources that are necessary to fulfil other aims.[829] In other words, the statute itself may be deferring to regulatory interpretation.[830]

3. Unintended effects

Somewhat counter-intuitively, the main perils of non-deference do not arise from the overt excesses of specialist tribunals. Unintended effects – hidden excesses hard to be uncovered – are the most challenging risks for the common approach. When judges impose higher standards of scrutiny there is always a risk related to unforeseen effects that are beyond their control. The result of the unintended effects is regulatory uncertainty, which in the economic context may affect the financing of the industry and put up costs to consumers – exactly the opposite effect of what should be the regime's aim. It has been argued that the risk of unintended consequences may increase particularly when specialist tribunals tend to provide 'general advice' that goes beyond the actual dispute.[831] In addition, it could be added that the risk may also arise from the object-specificity nature of the CAT's task. Competition law is nowadays only one amongst a number of objectives that regulators must pursue. This plurality not only compels regulators to adopt a more holistic approach, but also binds the CAT to act with interpretive flexibility.

Nonetheless, it can be stressed that tribunals and courts alike must avoid unintended consequences and regulatory uncertainty. A recent case that illustrates this is *National*

[829] See also Monti (2008: 132) (considering the 'regulatory holiday' as a means to solve tensions between allocative and dynamic efficiency).

[830] See *e.g.*, *R. v Ministry of Agriculture, Fisheries and Food Ex p. First City Trading* [1997] 1 CMLR 250, 278 (indicating that decisions as to political and social choice are made by the legislature, or by a person assigned the task by the legislature). Note that the argument is valid in the UK context because regulatory institutions (including statutes) were *not* specifically designed to curb discretion.

[831] The Court of Appeal highlighted these points recently in *Floe*: 'Specialist tribunals seem to be more prone than ordinary courts to yield to the temptation of generous general advice and guidance…They should take care not to be, or feel, pressured…to do things which they are not intended, qualified or equipped to do' (*Ofcom and T-Mobile (UK) Limited v Floe Telecom Limited (in liquidation)* [2009] EWCA Civ 47, at para. 21). For that reason, the court indicated that it 'should take care to confine its judgement to those points on which there are very good reasons, in the interests of the parties and in the public interest, for departing from the normal prudent course of deciding only what it is necessary for the adjudication of the actual dispute between the parties' (*ibid*, at para. 24).

Grid.[832] The case is in the competition law domain, so the CAT had jurisdiction to revise the merits. In *National Grid* the energy regulator (Ofgem) fined a firm for abuse of its dominant position in the market for domestic-sized gas meters. The CAT dismissed the appeal, but significantly reduced the fine imposed on the firm.[833] The CAT considered that the regulator's close involvement in the facts that led to the infraction was an important mitigating factor. Moreover, it castigated the regulator for 'not consider[ing] that it was part of its role either as an industry regulator or as a competition law enforcement agency to steer the industry participants away from making private arrangements which risked jeopardising the competitive process to a serious degree'.[834] In practice, the CAT forced the regulator to carry out an assessment of their own past actions when assessing possible infringements. This is rather odd. It is unlikely that the agency's involvement in a given case may create legitimate expectations that the regulator will not take action against a firm. However, after *National Grid* the same regulator must consider its involvement as a mitigating factor whilst setting a fine. That is, the regulator must use two different sets of 'robes and wigs' – the regulatory gown, when dealing with day-to-day matters, and the competition gown, whilst assessing infringements.

This is a clear example of unintended effects of judicial control. Concurrent powers do not equal separation of functions inside the same agency. Neither the statutes nor the policy guidance of competition authorities provide for a double role. Both roles are parts of a whole that allows the agency to fulfil its multiple duties. With the burden imposed by the CAT, the pursuance of the goal of promoting competition (one of the primary duties of regulators) may be severely weakened. It is not hard to imagine practical problems arising from the requirement to have regard to their own past behaviour. Regulators may at least face delays whilst fulfilling their competition role, because the fear of 'shooting their own foot' may lead them to be overzealous in analysing their own past 'regulatory behaviour' within the context of an investigation. At most, they may decide not to take competition actions that would otherwise have been taken, due to their deep past regulatory involvement in the issue.

---

[832] *National Grid plc v Gas and Electricity Markets Authority*, [2009] CAT 14.

[833] The fine was slashed from £41.6 million to £30 million.

[834] *Ibid*, at para. 208. The CAT added that the regulator was 'the architect and main driver of the process' that led to the infringement.

However, unintended effects are more common when a non-expert court revises the same kind of issues. The approach adopted by the CAT risks to 'force' traditional courts to come out of their feuds and decide on substantial regulatory matters. Consider the decision in *National Grid* again. The CAT recognized that both the burden imposed upon the regulator and the significant mitigation in the fine, were imposed considering 'the history of the discussions in the particular circumstances of this case'.[835] That should have served to avoid extensive interpretations. However, the Court of Appeal reaffirmed and increased the burden. The court criticized the CAT for not giving the regulator's 'involvement in the process' that led to the condemned actions '*sufficient weight*' as a mitigating factor.[836] As a result, the court halved the fine imposed by the CAT.[837] The consequence is not trivial. Not only must the regulator now pay regard to its own past behaviour as a mitigating factor; it must give it *considerable* weight.

More importantly, however, the reasoning of the Court of Appeal sounds at least unconvincing and at most insufficient to sustain the conclusion that the CAT did not take the infringement seriously enough. The court reached its conclusion accepting that 'there was [no] legal error in the Tribunal's approach to the issue of penalty' (no unlawfulness); that the conclusion was not 'an unreasoned one' (no unreasonableness); and that it has 'little familiarity with penalties in the field of competition law' itself (lack of specialisation).[838] The court only relied upon an irrelevant argument already considered by the CAT – the 'history' of the case.[839] In addition, the court gave weight to other factors that the CAT had dismissed. One of them was that the case raised a novel point. For example, the court indicated that 'the basis on which the finding of abuse was established did involve a substantial change of position from that originally adopted by the [regulator], which suggests an element of

[835] *Ibid*, at para. 218. As the CAT put it: 'This is not to say that sectoral regulators are in all cases required to step in and sound some warning bells on competition grounds if they see market developments taking a worrying turn. Neither are we saying that if a company sends a draft contract out of the blue to an official within the regulator it can then claim to have tacit approval if the regulator does not take action' (*Id.*).

[836] *Gas and Electricity Markets Authority and others v National Grid plc*, [2010] EWCA Civ 114, [2010] All ER (D) 296 (Feb) (per Lord Justice Richards), at para.108 (emphasis added).

[837] The Court set the fine at £15 million.

[838] *Ibid*, at paras.102, 103 and 115, respectively. It is fair to presume that there was no unfairness on the CAT assessment.

[839] Cf. para. 209 of the CAT's judgment with para.108 of the court's judgment.

uncertainty about the correct analysis'.[840] Hence it recommended reflecting that issue in the penalty. But the statement seems to refer to the normal development of a regulatory activity. There is no reason why a regulator cannot change its mind (even substantially and/or on *ex post* basis in the course of an investigation) without producing uncertainty. Even if it is correct that the involvement of the regulator must be taken into account in the context of a penalty, it seems to be an exaggerated manifestation of interference to give that consideration such a weight as the court did (one half of the fine imposed by the CAT).

In sum, the sole fact that the CAT may involuntarily err does not undermine the claim for its 'right' to set its own standard of review. Actually, it is highly likely that in such cases the margin of error is lower. Given its specialisation the CAT probably has a better understanding of the regulator's reasoning and may better appreciate aspects that a non-specialized court is either not in a position to consider or it must not assess because of the general nature of its task. Therefore, a unified and specific standard of review may help to both decrease judicial fallibility and enhance interpretive flexibility.[841]

CHAPTER CONCLUSION

Designing an effective mechanism for scrutiny of decisions made by regulators is not an easy task. On the one hand, there is a question related to the form of the administrative system – whether specialisation is necessary at all and what the institutional characteristics of the specialised body would be. On the other hand, there is a question related to the substance of the law – whether regulatory law requires the application of different standards of scrutiny. As this chapter has shown, the answer to both questions involves trade-offs and is largely related. The standard of scrutiny needs to match the institutional characteristics of the appeal body.

[840] *Ibid,* at para. 115. The decision also mentions a few other factors, mostly technical details of the controversy at issue.

[841] See Sunstein & Vermeule (2003: 949) (denouncing the 'cognitive trap': the demand of specialists that generalist judges should decide cases like specialists, overlooking the possibilities of far worse outcomes due to the emulation).

The UK experience gives important lesson for regulatory design. For many years the study of judicial scrutiny of regulatory decisions and its implications in this area was absent from specialised literature in regulation. Fortunately, the last decade has seen an increasing importance of the topic – particularly after the creation of the CAT. Its characteristics of expertise, experience and object-specificity have resulted in judicial scrutiny having arguably more influence on substantive regulatory policy outcomes. The question that remains is how to integrate judicial review within the more complex task of designing a regulatory regime.

APPENDIX: THE EFFICIENT COMPONENT PRICING RULE (ECPR)

The ECPR is a form of marginal-cost pricing according to which it is optimal to set the access price to a bottleneck (a) equal to the direct cost of providing access (C_0) plus the opportunity cost of providing access to the interested provider, which is equivalent to the reduction of the incumbent profit caused by the provision of access – i.e., the price (P) minus the direct cost (C_0) and the marginal cost (C_1):

$$(1) \qquad a = C_0 + (P - C_0 - C_1)$$

The key message of the ECPR is extremely simple. It supports the revenue neutrality for the incumbent, since the access charge should be set equal to the difference between the final price (P) and the marginal cost (C_1) on the competitive segment:

$$(2) \qquad a = P - C_1$$

However, to achieve its objective, the rule has strong assumptions. First, final products must be homogeneous ($p = p_1 = p_2$).[842] Secondly, there is no possibility of bypassing the incumbent. Finally, the market must be contestable –i.e., a market where firms can 'hit-and-run' due to the absence of sunk costs and free entry.[843] Only when these conditions hold, the ECPR becomes a margin rule: the margin between the incumbent's final price and the access charge (P – a) will be equal to the marginal cost in the competitive segment (C_1). Alternatively, the ECPR is also widely known as 'retail minus'.[844] That is, the access price is the price at which the incumbent sells a given service to a given end-user in the downstream market minus the costs avoided by the incumbent when the new entrant bears some of the costs of providing the service to the end-user.

[842] Note that if the incumbent's product is different from the entrant's product, the access charge should be lowered to reflect the opportunity cost of entry.

[843] Baumol *et al.* (1982).

[844] OECD (2004), at 215-16, defines retail minus as '[a]n approach to setting access prices under which the access prices are explicitly set on the basis of the end-user or retail prices of the corresponding final services. The discount off retail prices is usually set as a fixed percentage of the retail price'.

BY WAY OF CONCLUSION…

Generally speaking, the issue of how social progress and adaptation occur is one of those 'big' questions open for debate. In argumentation theory, for example, the dominant thinking considers conceptual change as evolutionary, not revolutionary.[845] It involves innovation and selection, rather than a competition of paradigms to supersede one another. At risk of over-simplification, this theory holds that knowledge evolves when a new concept focuses a problem from a different angle and, after being subject to a process of comparison, it replaces or revises a more traditional concept. The underlying aim is primarily to reject absolutisms. In regulation, by contrast, arguments somewhat resemble the idea of a Hegelian dialectic model of change based on the interaction of opposing forces: that the regulatory 'pendulum' swings from one extreme to the other (so the metaphor goes) seems to be an image deeply enrooted amongst many commentators and practitioners alike. Every time the mood fluctuates one way, there seems to be a rediscovery and update of rather old ideas that were, up to that moment, forgotten in the past, but that suddenly prove to be useful and plainly applicable to new concerns. Recent examples abound – from the now discredited initiatives towards 'light touch' regulatory techniques and their resemblance (to a certain extent) to *laissez faire*, to the 'renaissance' of Keynesian ideas for financial markets and a nearly universal call for more regulation. These and other examples demonstrate that in regulation there seems to be an uncontainable impetus for declaring paradigm shifts.

One major indication of the predominance of absolute ideas in regulatory thinking was the bogus feeling that the 'deregulation' movement that began during the last decades of the 20th century was unstoppable; and even more generally, that some sort of 'end of the history' had finally arrived with the undeniable triumph of capitalism. However, important theoretical insights have correctly insisted that the impression that capitalism equates the retreat of the State is utterly erroneous. In fact, historically, regulation never receded – in the UK, for example, from the highly interventionist regimes of the Tudor and Stuart periods, passing through the misguided so-called period of *laissez-faire*, to the 20th century nationalisations

[845] See, e.g., Toulmin (1972).

and then privatisations and deregulations[846], what has been witnessed is no more than a reshuffle of the boundaries between the State and markets with the consequent rearrangement of the involvement of the government in the economy. Furthermore, it is not just the idea that, from time to time, the State retreats in favour of markets and vice-versa that has proved wrong. As recent events have shown, the predicted outcomes of this mutual advance and retreat have been shown to be incorrect. Deregulation proved to be merely re-regulation. Likewise, capitalism is – it always has been – 'regulatory capitalism'.[847]

Indeed, capitalism is no more than a denomination for a tangled web of evolving economic and legal institutions.[848] More than being akin to the swing of a pendulum, institutional evolutions are like spiral, incremental movements.[849] Notwithstanding the nature of their character (either endogenous or exogenous)[850], past, present and future institutions are interlinked. Therefore, along with recognising past contributions to the study of regulation, it is also necessary to pay attention to the number of insights from different disciplines (such as political sciences, law and economics, but also others) which stresses the need to move forward, particularly in four directions: (1) beyond the State and firms to a plurality of players; (2) beyond the 'economic man' to a broader theoretical underpinning that also takes into account behavioural and cognitive factors; (3) beyond the dichotomy of State and markets to a wide variety of regulatory instruments and modes of governance; and (4) beyond the focus on a single regulatory objective to the recognition of the plurality of goals that underlies policy-making. Overall, these ideas summarise an undeniable existing truth: underneath the complex regulatory system there is a growing diversity in constant flux. Arguably, within the current state of affairs, the theoretical underpinnings of regulation should be bound to recognise various interrelated dimensions and domains.

So, where does the relatively narrow sub-field of economic regulation stand within this broad picture? Unfortunately, for too long now, economic regulation has lagged behind other major developments of regulatory theory (see Chapter I). It seems to be trapped in its self-imposed

[846] Ogus (1994: 6-12).
[847] Levy-Faur (2005).
[848] *E.g.*, Posner (2010) (analysing the financial crisis).
[849] By the same token, North (1990a), Aoki (2007).
[850] *See*, e.g., Aoki (2007).

boundaries, so to speak. These were fundamentally laid during the early 1990s, when the field was 'reshaped' coinciding with the privatisation-cum-deregulation trend that became widespread across jurisdictions. Both the 'new economics of regulation', centred on asymmetric informational concerns, and the 'new institutional economics' approach, with its focus on transaction-cost and modes of governance, gained momentum, building on, or positioning against more traditional rational choice approaches. Since then, theories of economic regulation have either wrongly believed in the possibility of finding single first-best mechanisms that solve discrete regulatory problems, or focused excessively on high-level regulatory institutions. These are all important substantive areas – unquestionably, all worth being considered by policymakers. However, contributions in the sub-field have remained, for too long, almost exclusively circumscribed to these topics, without embracing new, wider progresses. Even now, economic regulation fails to take full account of the evolving varieties underlying not only regulation generally, but also the specific industries subject to it (i.e., generally speaking, those traditionally considered natural monopolies). The sub-field has neither changed paradigm (if this were at all possible), nor recognised incremental advances in general regulatory theory. This lack of advances has had pernicious consequences. Not only has it been just a pervasive theoretical nuisance; the practical consequences of such a line of thinking have also been undermined.

The scarcity of fresh insights and enhancements in the theory has added to the view that, perhaps, economic regulation is not fit for purpose anymore. Moreover, the very purposes of regulation are currently an issue open to debate. Depending on the particular circumstances, and because of the interdependencies existent between social, economic and political factors, these ideas may end up putting into question the foundations of institutional design. However, given that institutional arrangements are for the most part contingent upon country-specific factual conditions and history, the analysis of the validity of such questioning needs to be carried out in a particular context. On this, economic regulation in the UK presents a privileged, almost unique case study. Since privatisation, the scope of both competition law and economic regulation has played a major role in defining powers and competences of private and public actors in different sectors. Above all, the study of the evolution of regulation in the UK shows that competition can be effective, but it needs to be carefully complemented with a wide range of other tools and policies. Recently, though, the

questioning of several areas of regulatory policy and its ultimate goals has been paralleled by calls for discrete institutional adaptations.

There can be little doubt that regulatory design in industries subject to economic regulation is challenging. Difficult trade-offs are likely to arise. This work has claimed that only by undertaking comparative institutional analysis of the governance architectures and the processes and formal rules that channel political and economy activity, is it possible to properly structure the institutional design of regulation. For regulatory rule-making needs to recognize diversity and plurality; the analysis requires a profound understanding of both the structure of the industry at issue and its evolution. Undoubtedly, this thesis has not intended to cover all structural facets of the industries subject to economic regulation in the UK – arguably a haughtier enterprise and a rather unachievable task. In order to show the difficulties and how to overcome them, the focus has been instead on a key part of such a structure: the assessment of financial implications of price controls, particularly those related to the capital structure and the cost of capital.

Changes in the capital structure of UK utilities firms since privatisation reflect with clarity the interactions between regulatory policy and private incentives (see Chapter III). The early recognition of the importance of the cost of capital (even against the theoretical basis of the regulatory regime) influenced, at least partially, subsequent changes in private equity and debt. It also gave rise to a number of regulatory responses aiming both to exert some form of control over the capital structures and to develop new ways to better compute the cost of capital. Alas, many regulatory approaches have been adopted as a reaction to what are at least partial, or at most wrong, explanations for private attitudes. On the contrary, when regulators have taken into account the specific institutional features of the regulatory regime, both the justifications for interventions and the measures adopted have seriously improved. However, the importance of studying the regulation of the capital structure goes beyond the particular assessment of regulators' behaviour regarding financial aspects of pricing policies. As the analysis has exposed, interactions between the capital structure and the cost of capital are essentially an exceptional example of how firms respond to the regime's incentives in industries where revenues are controlled and to a large extent guaranteed.

In turn, the discussion on the cost of capital and the capital structure of utilities firms leads to broader arguments about who are the cheapest bearers of risk and what the role of the State in infrastructure industries is. Risk allocation and the avoidance of State guarantees are a decisive task of governments (see Chapter IV). Different rules for risk allocation correspond to different types of risks, and these in turn correspond to different institutional regulatory regimes. Likewise, State guarantees (which arise from the specific risk of financial distress) are also moulded by the regulatory regime, which means that some of them are totally unavoidable – unintended effects are ubiquitous in economic regulation. At the same time, however, the provision of some guarantees depends considerably on the regulator's attitudes – typically, bailouts. Most guarantees have pernicious consequences for the regime, but some of them may occasionally be beneficial. Considering this, regulators must consider carefully their options in order to discriminate between actions that render positive or negative consequences. Regulators have a duty to secure the finance of private firms' activities. Their task is therefore to balance such duty with the protection of consumers without distorting the incentives provided by the framework.

Diversity and plurality, specific regulatory tasks (such as financial aspects of regulation), and the allocation of risks and avoidance of unintended effects – all of these lead us to consider which agents are better placed to assess not only whether regulators are fulfilling their task properly within the remit of their mandate and competences, but also to consider more broadly the possible shortcomings of institutional design and point them out to the corresponding institutional bodies. Judges might naturally fulfil these tasks. However, they have been traditionally reluctant to be involved in regulatory matters. Maybe due to this fact, judicial scrutiny of regulatory decisions has been a topic rather neglected by a great part of the UK literature focused on regulation and left almost entirely to more general public law studies. Despite this fact, what matters is recognising the importance of judicial scrutiny beyond its role as institutions that hold regulators accountable for their acts. Courts should be placed within an institutional context where their importance is stressed along with that of regulators, the State and other third parties.

Most importantly, it should be acknowledged that, in the current state of affairs, judges do not fulfil a passive role: they influence regulatory outcomes significantly. Indeed, there is a need

to recognise the natural limitations of judicial scrutiny – a humble judicature fits better with the emphasis on competition and other competing regulatory goals than judicial activism. However, within an institutional setting dominated by plurality and diversity, it has become undeniable that blind deference to other bodies does not contribute to the success of the regime. Judges are capable of adopting an appropriate and balanced approach to regulatory law. Such an approach will advise in which occasions a deferential standard is more desirable or whether the creation of judicial safeguards constitutes a more apt alternative to restrain negative incentives for private behaviour and public discourse and practice. The enhancement of the legitimacy of judicial review within the regulatory process crucially depends on judges' ability to keep this fine balance. These insights take on more fundamental and general questions regarding the current role of public law in UK regulatory judicial review. I have not attempted to provide answers for these questions. Instead, in Chapter V it has been argued that as judges' specialisation increases, they are better placed to fulfil this task. Thus, as long as there is an increased flexibility in the applicable standards and any mismatch between the legal rules and institutional incentives is avoided, confidence in specialised justice may grow. This, in turn, should make healthy room for a certain degree of 'experimentation' that, along with carefully circumventing unintended effects and errors, may allow courts to play a very positive role in industries subject to economic regulation.

Considering all of the above, the key question that arises is how to face the future: How should economic regulation adapt to new challenges and advances? This query, it is submitted, demands an answer that goes not only beyond particularities of certain regulatory domain, but also extends outside traditional boundaries of single, closed disciplines. Institutions, finance and public law have been used here. Combined, they have served to attempt to provide a broader avenue to address the main areas where economic regulatory theory and practice lag new developments (see Chapter II). Unfortunately, most works in economic regulation still tend to focus on isolated perspectives, being either concerned with sector-specific subject areas or drawn from the insights of a single theory. Thus, for example, most traditional economic theories look at regulatory law as a way to correct market failures, frequently ignoring important issues of political processes and their failure highlighted by political science literature. Also, whilst psychological and modern behavioural economic works have consistently stressed that there are a number of cognitive and behavioural

shortcomings of the actors involved in the regulatory process that should be taken into account in the analysis, they are habitually set aside – stubbornly – in favour of more traditional assumptions of rationality. As a consequence of this fragmented approach, many valuable lessons learnt in other stances are at least undervalued or at most completely lost. An important (albeit implicit) claim of this thesis is that the study of regulation can gain considerably from an approach that embraces and combines different perspectives, traditional and novel. Accordingly, one of the main aims of this work has been to offer a modest contribution towards a more coherent alternative for regulatory policymaking.

Beyond isolated disciplines and particular industries, regulation is a broad subject capable of embracing an ample variety of perspectives. Possibilities for mixed research abound. The study of economic regulation in particular should reach beyond the remit of those viewpoints that are most commonly employed to its study. Moreover, interdisciplinary scholarship should be strongly encouraged if the lessons of recent crises and theoretical advances are to be taken seriously. For such an approach does make a difference. An interdisciplinary approach may update the thinking of regulation to present challenges; provide suitable solutions; overcome current weaknesses of regulatory processes; and strengthen the institutional responses to vulnerability, improving ways in which the system reacts and adjusts to both internal and external shocks. This leads us back to the initial point that all institutional change is incremental. Interdisciplinarity should not be taken as a call to shake the foundations of economic regulation and exchange them for a new paradigm – a comprehensive structural solution to what are, in essence, the same old concerns. It is more a call to be aware of the need for urgency: advances should not await calmer days.

REFERENCES*

AGGARWAL, Rajesh K. and Andrew A. SAMWICK (1999): 'Empire-Builders and Shirkers: Investment, Firm Performance, and Managerial Incentives', *NBER Working Paper 7335*, available at <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=196029>.

AGHION, Phillip, Oliver HART and John MOORE (1992): 'The Economics of Bankruptcy Reform', *Journal of Law, Economics and Organization,* 8(3): 523-46.

ALCHIAN, Armen A. (1950): 'Uncertainty, Evolution, and Economic Theory', *Journal of Political Economy,* 58: 211-21

ALEXANDER, Ian and Clive Harris (2005): *The Regulation of Investment in Utilities*. Washington, DC: The World Bank.

_____ and Sie-Chien CHIA (2003): 'Bond Finance: A Growing Source of Funds for Utility and Infrastructure Companies?, *International Journal of Regulation and Governance*, 2(1): 1-25.

_____, Antonio ESTACHE and Adele OLIVERI (2000): 'A few things transport regulators should know about risk and the cost of capital', *Utilities Policy*, 9: 1-13.

_____, Colin MAYER and Helen WEEDS (1999): 'Regulatory Structure and Risk: An international Comparison', *Policy Research Working Paper* 1698. Washington: The World Bank.

_____ and Timothy IRWIN (1996): 'Price Caps, Rate-of-Return Regulation, and the Cost of Capital', *Public Policy for the Private Sector*, note 87. Washington: The World Bank.

ALEXIADIS, Peter and Martin CAVE (2010): 'Regulation and Competition Law in Telecommunications and Other Network Industries', in *The Oxford Handbook of Regulation* (R. Baldwin, M. Cave and M. Lodge, eds.), chapter 20: 500-22. Oxford, UK: Oxford University Press.

ALEXY, Robert (2002): *A Theory of Constitutional Rights* (J. RIVERS, tr.). Oxford, UK: Oxford University Press.

ALLAN, Trevor R.S. (2006): 'Human Rights and Judicial Review: A Critique of "Due Deference"', *Cambridge Law Journal,* 65: 671-95

_____ (2000): 'The Rule of Law as the Foundation of Judicial Review', in C. FORSYTH (ed.), *Judicial Review and the Constitution* (Oxford: Hart Publishing), 419

AOKI, Masahiko (2010): *Corporations in Evolving Diversity – Cognition, Governance, and Institutions*. Oxford, UK: Oxford University Press

_____ (2001): *Toward a Comparative Institutional Analysis*. Cambridge, MA and London, UK: The MIT Press.

ARANCIBIA, Jaime (2011): *Judicial Review of Commercial Regulation*. Oxford, UK: Oxford University Press

ARMSTRONG, Mark, and David SAPPINGTON (2006): 'Regulation, Competition and Liberalization', *Journal of Economic Literature*, 44(2): 325-66

_____, Chris DOYLE and John VICKERS (1996): 'The Access Pricing Problem: A Synthesis', *Journal of Industrial Economics*, 44: 131-50

* All the documents issued by UK regulators can be found in their respective websites.

_____, Simon COWAN and John VICKERS (1994): *Regulatory Reform: Economic Analysis and British Experience*. Cambridge, MA & London, UK: The MIT Press.

ARROW, Kenneth (1985): 'The Economics of Agency', in, *Principals and Agents: The Structure of Business* (John W. Pratt and Richard J. Zeckhauser), chapter 2: 37-51. Boston, MA: Harvard Business School Press.

_____ (1971): *Essays in the Theory of Risk-Bearing.* Amsterdam: North-Holland.

ARTERBURN, Norman F (1927): 'The Origin and First Test of Public Callings', *University of Pennsylvania Law Review*, 75: 411-28

AVERCH, Harvey and Leland L. JOHNSON (1962): 'Behavior of the Firm Under Regulatory Constraint', *American Economic Review,* 52: 1052-69

AYRES, Ian and Robert GERTNER (1992): 'Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules', *Yale Law Journal,* 101(4): 729-73

_____ and _____ (1989): 'Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules', *Yale Law Journal,* 99: 87-130

_____ and John BRAITHWAITE (1990): *Responsive Regulation: Transcending the Deregulation Debate.* New York, NY and Oxford, UK: Oxford University Press.

BAIRD, Douglas G., Robert H. GERTNER and Randal C. PICKER (1994): *Game Theory and the Law.* Cambridge, MA and London, UK: Harvard University Press.

BALAKRISHNAN, Srinivasan and Isaac FOX (1993): 'Asset Specificity, Firm Heterogeneity and Capital Structure', *Strategic Management Journal*, 14(1): 3-16.

BALDWIN, Robert, (1995) *Rules and Government.* Oxford: Clarendon Press.

_____, Martin CAVE and Martin LODGE (eds.) (2010): *The Oxford Handbook of Regulation.* Oxford, UK: Oxford University Press.

_____ and _____ (1999): *Understanding Regulation.* Oxford, UK: Oxford University Press.

_____ and Christopher MCCRUDDEN (1987): *Regulation and Public Law.* London, UK: Weidenfeld and Nicolson.

BALLEISEN, Edward J. and David A. MOSS (eds.) (2010): *Government and Markets – Toward a New Theory of Regulation.* Cambridge University Press.

_____ (2010): 'The prospects for effective coregulation in the United States: a historian's view from the early twenty-first century', in *Government and Markets* (E. Balleisen and D. Moss, eds.), chapter 14: 443-81. Cambridge University Press.

BARDACH, Eugene and Robert KAGAN (1982): *Going by the Book: The Problem of Regulatory Unreasonableness*. Philadelphia, Penn.: Temple University Press.

BARNARD, Roger and John COOPER (2008): 'Is Ofgem still fit for Purpose?', *International Energy Law Review*, 5: 158-61.

BARON, David P. and David BESANKO (1986) 'Regulation, Asymmetric Information, and Auditing', *Rand Journal of Economics*, 15(4): 447-70

_____ and Roger B. MYERSON (1982): 'Regulating a Monopolist with Unknown Costs', *Econometrica,* 50(4): 911-30

BARZEL, Yoram (1989): *Economic Analysis of Property Rights*. Cambridge University Press.

BAUMOL, William J. (1967): 'Reasonable Rules for Rate Regulation: Plausible Policies for an Imperfect World', in: *Prices: Issues in Theory, Practice, and Public Policy* (A. Phillips & O. E. Williamson, eds.). Philadelphia: University of Pennsylvania Press.

_____ and Thomas W. MERRILL (1998): 'Does the Constitution Require that we Kill the Competitive Goose? Pricing Local Phone Services to Rivals', *New York University Law Review,* 73: 1122.

_____ and _____ (1997): 'Deregulatory takings, Breach of the Regulatory Contract, and the Telecommunications Act of 1996', *New York University Law Review,* 72: 1037.

_____ Janus A. ORDOVER and Robert D. WILLIG (1997): 'Parity Pricing and Its Critics: A Necessary Condition for Efficiency in the Provision of Bottleneck Services to Competitors', *Yale Journal on Regulation,* 14: 145.

_____ and J. Gregory SIDAK (1994): 'The Pricing of Inputs Sold to Competitors', *Yale Journal on Regulation,* 11: 171-202

_____ (1983): 'Some Subtle Issues in Railroad Regulation', *International Journal of Transport Regulation,* 10: 341-55

_____ John C. PANZAR and Robert D. WILLIG (1982): *Contestable Markets and the Theory of Industry Structure* (Harcourt College Pub).

BAXTER, William F. (1982): 'Separation of Powers, Prosecutorial Discretion, and the "Common Law" Nature of Antitrust Law', *Texas Law Review,* 60: 661.

BECKER, Gary (1983): 'A Theory of Competition among Pressure Groups for Political Influence', *Quarterly Journal of Economics*, 98(3): 371-400

BEESLEY, Michael E. and Stephen LITTLECHILD (1997 [1983]): 'Privatization: Principles, Problems and Priorities', in *Privatization, Regulation and Deregulation* (M. E. BEESLEY), chapter 2: 26-42 [original: *Lloyds Bank Review,* July 1983]. London, UK: Routledge, 2nd ed.

_____ and _____ (1989): 'The Regulation of Privatized Monopolies in the United Kingdom', *Rand Journal of Economics*, 20(3): 454; reprinted in Michael Beesley (1997): *Privatization, Regulation and Deregulation*, 2nd edition, chapter 4: 58-83. London: Routledge, 2nd ed.

BENKLER, Yochai (2010): 'Law, Policy, and Cooperation', in *Government and Markets* (E. Balleisen and D. Moss, eds.), chapter 9: 299-332. Cambridge University Press.

BERNSTEIN, Marver (1955): *Regulating Business by Independent Commission.* Princeton, NJ: Princeton University Press.

BERTRAND, Marianne and Sendhil MULLAINATHAN (2003): 'Enjoying the Quiet Life? Corporate Governance and Managerial Preferences', *Journal of Political Economy*, 111(5): 1043-75.

BIGLAISER, Gary and Michael H. RIORDAN (2000): 'Dynamics of Price Regulation', *Rand Journal of Economics*, 31(4): 744-67

BINGHAM, Sir T. (2007): 'The Rule of Law', *Cambridge Law Journal*, 66(1): 67-85

BISHOP, Matthew, John KAY & Colin MAYER (eds.) (1994): *Privatization & Economic Performance.* Oxford University Press.

BJUGGREN, Per-Olof (1995): 'A Transaction Cost Perspective on Financial Distress and Capital Structure', *International Review of Law and Economics*, 15: 395-404.

BLACK, Julia (2010) 'The Role of Risk in Regulatory Processes', in *The Oxford Handbook of Regulation* (R. Baldwin, M. Cave and M. Lodge, eds.), chapter 14: 302-48. Oxford, UK: Oxford University Press.

_____ (2001): 'Decentring Regulation: The Role of Regulation and Self-Regulation in a "Post-Regulatory" World', [2001] *Current Legal Problems* 103.

_____, Peter MUCHLINSKI and Paul WALKER (eds.) (1998): *Commercial Regulation and Judicial Review.* Oxford, UK: Hart Publishing.

_____ (1998): 'Reviewing Regulatory Rules: Responsibility to Hybridisation', in *Commercial Regulation* (J. Black, P. Muchlinsky and P. Walker, eds.). Oxford, UK: Hart Publishing.

BOLT, Chris (2010): 'Regulating a state owned company: can we create the right incentives?', *Beesley Lecture*, 28 October 2010, available at <http://www.chrisbolt.me.uk/resources/Regulating+a+state+owned+company.pdf>

BONBRIGHT, James C., Albert L. DANIELSEN and David R. KAMERSCHEN (1988): *Principles of Public Utility Rates*. Arlington, Virginia: Public Utilities Reports, Inc.

BORTOLLOTI, Bernardo, Carlo CAMBINI, Laura RONDI and Yossi SPIEGEL (2007): 'Capital Structure and Regulation: Does Ownership Matter?', *Fondazione Eni Enrico Mattei, Nota di Lavoro 94.2007*, available at <http://ssrn.com/abstract=1022342>

BÖS, Dieter (1991): *Privatization: A Theoretical Treatment*. Oxford University Press.

BRAITHWAITE, John (2008): *Regulatory Capitalism: How it works, ideas for making it work better.* Cheltenham, UK and Northampton, MA: Edward Elgar

BREALEY, Richard A., Steward C. MYERS and Franklin ALLEN (2008): *Principles of Corporate Finance*. McGraw-Hill, 9th edition (international edition)

BRENNAN, Michael J. and Eduardo S. SCHWARTZ (1982): 'Consistent Regulatory Policy and the Uncertainty', *Bell Journal of Economics*, 13(2): 506-21

BREYER, Stephen (1982): *Regulation and Its Reform.* Cambridge, MA and London, UK: Harvard University Press

_____, Richard B. STEWARD, Cass R. SUNSTEIN and Matthew L. SPITZER (2002): *Administrative Law and Regulatory Policy.* New York: Aspen Publishers, 5th ed

_____ (1986): 'Judicial Review of Questions of Law and Policy', *Administrative Law Review,* 38: 363.

BRIERLEY, Peter and Philip BUNN (2005): 'The Determination of UK Corporate Capital Gearing', *Bank of England Quarterly Bulletin*, Autumn: 356-66

BROWN, Ashley C., Jon STERN, and Bernard TANENBAUM with Defne GENCER (2006): *Hardbook for Evaluating Infrastructure Regulatory Systems*. Washington, DC: The World Bank.

BUCHANAN, James and Gordon TULLOCK (1962): *The Calculus of Consent.* Michigan: University of Michigan Press

BUCKS, Peter (2005): 'The Financeability Gap', *The Utilities Journal*, Oxera, March

BURTON, John (1997): 'The competitive order or ordered competition?: The "UK model" of utility regulation in theory and practice', *Public Administration*, 75: 157-88

CAA (Civil Aviation Authority) (2010): 'NATS (En Route) plc CAA price control proposals (2011-2014)', May 2010

_____ (2006): 'Possible offer for BAA plc: Statement by the UK Civil Aviation Authority', February 2006

CABRAL, Luis M.B. and Michael H. RIORDAN (1989): 'Incentives for Cost-Reduction under Price Cap Regulation', *Journal of Regulatory Economics*, 1(2): 93-102

CALVERT, Randall L. (1995): 'The Rational Choice Theory of Social Institutions: Cooperation, Coordination, and Communication', in *Modern Political Economy* (J. S. Banks and E. A. Hanushek, eds.), chapter 8: 216-67. Cambridge University Press

CARNATH, Robert (2009): 'Tribunal Justice – a New Start', [2009] *Public Law* 48-69

CARROLL, Lewis (1992 [1871]): *Through the Looking-Glass – and What Alice Found There*. London, UK: Vintage Books.

CAVE, Martin (2009): *Independent Review of Competition and Innovation in Water Markets: Final Report*, April 2009.

_____ (2006): 'Encouraging Infrastructure Competition Via the Ladder of Investment', *Telecommunications Policy*, 30: 223-37

_____ & Ingo VOGELSANG (2003): 'How Access Pricing and Entry Interact', *Telecommunications Policy*, 27: 717-27

CC (Competition Commission) (2010a): ''Bristol Water plc, A reference under section 12(3)(a) of the Water Industry Act 1991 – Report', August 2010.

_____ (2010b): 'Bristol Water plc, A reference under section 12(3)(a) of the Water Industry Act 1991 – provisional findings report', June 2010.

_____ (2008): 'Stansted Airport Ltd. – Q5 Price Control Review', October 2008.

_____ (2007): 'BAA Ltd. – A Report on the Economic Regulation of the London Airports Companies (Heathrow Airport Ltd. and Gatwick Airport Ltd.)', September 2007.

_____ (2002): 'BAA plc – A Report on the Economic Regulation of the London Airports Companies (Heathrow Airport Ltd., Gatwick Airport Ltd. and Stansted Airport Ltd.)', October 2002.

_____ (1996): 'BAA plc – A Report on the Economic Regulation of the London Airports Companies (Heathrow Airport Ltd., Gatwick Airport Ltd. and Stansted Airport Ltd.)', June 1996

CEPA (Cambridge Economic Policy Associates Ltd.) (2010): 'RPI-X@20: providing financeability in a future regulatory framework', May 2010

_____ (2009): 'The Use of RPI-X by Other Network Industry Regulators, Final Report, February 2009.

_____ (2007): 'Indexing the Allowed Rate of Return, ORR / OFWAT. Final Report', September 2007

CHEN, Jim (2004): 'The Nature of the Public Utility: Infrastructure, the Market, and the Law', *Northwestern University law Review*, 98: 1617

_____ (1999): 'The Second Coming of *Smyth v. Ames*', *Texas Law Review*, 77: 1535

COASE, Ronald (1937): 'The Nature of the Firm', *Economica*, 4: 386

_____ (1960): 'The Problem of Social Cost', *Journal of Law and Economics*, 3: 1

_____ (1988): *The Firm, the Market, and the Law*. Chicago and London: The University of Chicago Press

COGLIANESE, Cary and Evan MENDELSON (2010): 'Meta-Regulation and Self-Regulation', in *The Oxford Handbook of Regulation* (R. Baldwin, M. Cave and M. Lodge, eds.), chapter 8: 146-68. Oxford, UK: Oxford University Press.

COMMONS, John R. (1931): 'Institutional Economics', *American Economic Review*, 21(4): 648

COPELAND, Thomas E., J. Fred WESTON and Kuldeep SHASTRI (2005): *Financial Theory and Corporate Policy*, 4th edition. Pearson

CORNELL, Bradford and Allan C. SHAPIRO (1987): 'Corporate stakeholders and corporate finance', *Financial Management*, 16(1): 5-14

CORREIA DA SILVA, Luis, Antonio ESTACHE and Sakari JARVELA (2004): 'Is Debt Replacing Equity in Regulated Privatized infrastructure in Developing Countries?', *Policy research Working Paper* 3374. Washington: The World Bank

_____, Tim JENKINSON and Colin MAYER (2003): 'The Capital Structure of Water Companies', in: *Water Sustainability and Regulation* (D. Helm ed.), chapter 10: 155-66. Oxford: Oxera

COWAN, Simon (2006): 'Network Regulation', *Oxford Review of Economic Policy*, 22(2): 248-59

_____ (2003), 'Optimal Risk Allocation for Regulated Monopolies and Consumers', *Journal of Public Economics*, 88: 285-303

_____ (2001): "Developments in Regulatory Principles: The UK Experience", in: *Regulation of Network Utilities: The European Experience* (Henry, C., M. Matheu and A. Jeunemaître, eds.), chapter 2: 39-53. Oxford: OUP

CRAIG, Paul (2008): *Administrative Law* (London, UK: Sweet & Maxwell)

_____ (1999): 'Regulation and Judicial Review', in C. MCCRUDDEN (ed.) *Regulation and Deregulation*. Oxford, UK: OUP

_____ (1997): 'Formal and Substantive Conceptions of the Rule of Law: An Analytical Framework', *Public Law*: 467-87

_____ (1991): 'Constitutions, Property and Regulation', [1991] *Public Law* (winter): 538-54

_____ (1987): 'The Monopolies and Merger Commission: Competition and Administrative Rationality', in Robert BALDWIN and Christopher MCCRUDDEN, *Regulation and Public Law* (London, UK: Weidenfeld and Nicolson), 201-26

CROUCH, Martin (2006): 'Investment under RPI-X: Practical Experience with an Incentive Compatible Approach in the GB Electricity Distribution Sector', *Utilities Policy*, 14: 240

CROCKER, Keith J. and Scott E. MASTEN (1996): 'Regulation and Administered Contract Revisited: Lesson from Transaction-Cost Economics for Public Utilities Regulation', *Journal of Regulatory Economics*, 9: 5

CST (COUNCIL FOR SCIENCE AND TECHNOLOGY) (2009): *Improving Innovation in the Water Industry: 21st Century Challenges and Opportunities*, March

CUSHMAN, Robert E. (1941): *The Independent Regulatory Commissions*. New York: Oxford University Press

CURRIE, David (2003): 'Mutualization and Debt-Only Vehicles: Which Way for RPI-X Regulation?, *Competition and Regulation in Utility Markets* (C. Robinson, ed.), chapter 7: 187-98. Cheltenham, UK & Northampton, MA, USA: Edward Elgar

DASGUPTA, Sudipto and Vikram NANDA (1993): 'Bargaining and Brinkmanship: Capital Structure Choice by Regulated Firms', *International Journal of Industrial Organization*, 11: 475-97

DBIS (Department of Business, Innovation and Skills, UK) (2010): *Implementing the Revised EU Electronic Communication Framework: Overall Approach and Consultation on Specific Issues* (13 September)

DE LA MARE, Tom (2007): 'Regulatory Judicial Review: The Impact of Competition Law' (paper presented at the 2007 ALBA conference), available at <http://www.adminlaw.org.uk/docs/Thomas%20de%20la%20Mare%20-%20June%202007.pdf> (accessed: 30 June 2010)

DE FRAJA, Gianni and Clive STONES (2004): 'Risk and Capital Structure in the Regulated Firm', *Journal of Regulatory Economics*, 26(1): 69-84

DEMSETZ, Harold (1968): 'Why Regulate Utilities?', *Journal of Law & Economics*, 11(1): 55

DERTHICK, Martha and Paul J. QUIRK (1985): *The Politics of Deregulation*. Washington D.C.: The Brooking Institution

DEWATRIPONT, Mathias and Eric MASKIN (1995): 'Credit Efficiency in Centralized and Decentralized Economies', *Review of Economic Studies*, 62: 541

DIAMONTE, Robin L., John M. LIEW and Ross L. STEVENS (1996): 'Political Risk in Emerging and Developed Markets', *Financial Analyst Journal* 71 (May/June)

DICEY, Albert V. (1885): *An Introduction to the Study of the Law of the Constitution*. London: Macmillan

DIXIT, Avinash K. (2004): *Lawlessness and Economics – Alternative Modes of Governance*. Princeton, NJ and Oxford, UK: Princeton University Press

_____ and Barry J. NALEBUFF (1991): *Thinking Strategically.* New York and London: W.W. Norton & Company

DODD, E. Merrick (1932): 'For whom are corporate managers trustees?', *Harvard Law Review*, 45: 1145-63

DOUCET, Joseph and Stephen LITTLECHILD (2009): 'Negotiated settlements and the National Energy Board in Canada', *Energy Policy*, 31(11): 4633

_____ and _____ (2006): 'Negotiated Settlements: The Development of Legal and Economic Thinking', *Utilities Policy,* 14: 266

DOWNS, Anthony (1957): *An Economic Theory of Democracy*. New York: Harper & Brothers

DTI (Department of Trade and Industry, UK) (1982): *The Future of Telecommunications in Britain* Cmnd 8610 (London: HSMO)

_____ (1989): *Opening Markets: New Policy on Restrictive Trade Practices*, Cm 727 (London: HMSO)

_____ (1988): *Green Paper Review of Restrictive Trade Practices Policy: A Consultative Document*, Cm 331 (HMSO)

_____ (1986): *Tackling Cartels and Abuse of Market Power: Implementing the Government's Policy for Competition Law Reform*, March

_____ (2001): *A World Class Competition Regime* (white paper), Cm 5233 (30 July)

_____ and HM TREASURY (2006): *Concurrent Competition Powers in Sectoral Regulation – A Report by the Department of Trade and HM Tresury*, URN 06/1244 (May)

_____ and _____ (2004): *The Drivers and Public Policy Consequences of Increased Gearing*, report (October)

DWORKIN, Ronald (1977): *Taking Rights Seriously*. London, UK: Duckworth

EASTERBROOK, Frank H. (1983): 'Statutes' Domains', *University of Chicago Law Review*, 50: 533

ECONOMIDES, Nicholas and Lawrence WHITE (1995): 'Access and interconnection Pricing: How Efficient is the Efficient Component Pricing Rule?', *Antitrust Bulletin*, 40: 557

_____ and _____ (1998): 'The Inefficiency of the ECPR Yet Again: A Reply to Larson', *Antitrust Bulletin*, 43: 429

EDWARDS, Richard A. (2002): 'Judicial Deference under the Human Rights Act', *Modern Law Review*, 65: 859-82

EISNER, Marc Allen (2010): 'Markets in the Shadow of the State: An Appraisal of Deregulation and Implications for Future Research', in *Government and Markets* (E. Balleisen and D. Moss, eds.), chapter 16: 512-37. Cambridge University Press.

_____, Jeff WORSHAM and Evan J. RINGQUIST (2006): *Contemporary Regulatory Policy*, 2nd ed. London: Lynne Rienner Publishers, Inc.

ELLICKSON, Robert (1991): *Order Without Law: How Neighbors Settle Disputes*. Cambridge, MA: Harvard University Press

ELLIOTT, Mark (2001): 'The Human Rights Act 1998 and the Standard of Substantive Review', *Cambridge Law Journal*, 60(2): 301-36

EMMONS, Willis (2000): *The Evolving Bargain: Strategic Implications of Deregulation and Privatization*. Boston, MA: Harvard University Press

ENGELMAN, Philip (2001): *Commercial Judicial Review*. London: Sweet & Maxwell

EVANS, Lewis T. and Graeme A. GUTHRIE (2006): 'Inventive Regulation of Prices when Costs are Sunk', *Journal of regulatory Economics*, 29(3): 239-64

FAMA, Eugene F. (1980): 'Agency Problems and the Theory of the Firm', *Journal of Political Economy*, 88: 288-307

_____ and _____ (1996): 'Multifactor Explanation of Asset Pricing Anomalies', *Journal of Finance*, 51: 55-84

_____ and _____ (2002): 'Testing Trade-Off and Pecking Order Predictions about Dividends and Debt', *Review of Financial Studies*, 15: 1-33

_____ and Kenneth R. FRENCH (2005): 'Financing Decisions: Who Issues Stock?', *Journal of Financial Economics*, 76: 549-82

FEIGENBAUM, Harvey B. and Jeffrey R. Henig (1994): 'The Political Underpinnings of Privatization: A Typology', *World Politics*, 46(2): 185

FEINTUCK, Mike (2010): 'Regulatory Rationales Beyond the Economic: In Search of the Public Interest', in *The Oxford Handbook of Regulation* (R. Baldwin, M. Cave and M. Lodge, eds.), chapter 3: 39-63. Oxford, UK: Oxford University Press.

FISHER, Ronald, Rodrigo GUTIÉRREZ and Pablo SERRA (2003): 'The Effects on Privatization on Firms and on Social Welfare: the Chilean Case', *Research Network Working Paper R-456* (Inter-American Development Bank)

FOSTER, C.D. (1992): *Privatization, Public Ownership and the regulation of Natural Monopoly*. Blackwell, Oxford

FRANK, Murray Z. and Vidhan K. GOYAL (2007): 'Trade-off and Pecking Order Theories of Debt', available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=670543 (accessed: 18 April 2011)

FREEMAN, R. Edward (2010 [1984]): *Strategic Management: A Stakeholder Approach*. Cambridge University Press

_____, Jeffrey S. HARRISON, Andrew C. WICKS, Bidhan L. PARMAR and Simone DE COLLE (2010): *Stakeholder Theory: The State of the Art*. Cambridge University Press

FREEMAN, Jody (2000): 'The Contracting State', *Florida University Law Review*, 28: 155

FRIEDMAN, Milton (1953): *Essays in Positive Economics* (Chicago: University of Chicago Press)

_____ (1962): *Capitalism and Freedom*. Chicago, IL: University of Chicago Press

FSA (Financial Service Authority, UK) (2006): 'Private Equity: A Discussion of Risk and Regulatory Engagement', *discussion paper* 06/6, November 2006

FUKUYAMA, Francis (1992): *The End of History and the Last Man*. New York, NY: Avon Books, Inc.

FULLER, Lon L. (1979): 'The forms and limits of adjudication', *Harvard Law Review*, 92: 353-409

FURUBOTN, Eirik G. and Rudolf RICHTER (1997): *Institutions and economic theory: The contributions of the New Institutional Economics*. Ann Arbour, University of Michigan Press

GERARDIN, Damien and Robert O'DONOGHUE (2005): 'The Concurrent Application of Competition Law and Regulation: the Case of Margin Squeeze Abuses in the Telecommunications Sector', *GCLC Working Paper* 04/05

GEROSKI, Paul (2006): 'Competition policy and practice under the Enterprise Act', in Centre for the Study of Regulated Industries, *Regulatory Review 2004/05*. University of Bath, School of Management

GERWIG, Robert (1962): 'Natural Gas Production: A Study of Costs of Regulation', *Journal of Law & Economics*, 5: 69-92

GILOVICH, Thomas, Dale GRIFFIN and Daniel KAHNEMAN (eds.) (2002): *Heuristics and Biases: The Psychology of Intuitive Judgement*. Cambridge University Press

GÓMEZ-IBÁNEZ, José A. (2003): *Regulating Infrastructure – Monopoly, Contracts, and Discretion*. Cambridge, MA, and London, England: Harvard University Press

GOLDBERG, Victor P. (1976): 'Regulation and Administered Contracts', *Bell Journal of Economics*, 7(2): 426-48

GRABOSKY, Peter N. (1995a): 'Regulation by Reward: On the Use of Incentives as Regulatory Instruments', *Law & Policy*, 17(2): 257

_____ (1995b): 'Counterproductive Regulation', *International Journal Soc. and Law*, 23: 347

GRAHAM, John R. (2000): 'How Big Are the Tax Benefits of Debt?', *Journal of Finance*, 55(5): 1901-41

GRAHAM, Cosmo (2000): *Regulating Public Utilities: A Constitutional Approach*. Oxford, UK and Portland Oregon: Hart Publishing.

_____ (2009): 'Judicial Review of the Decisions of the Competition Authorities and the Economic Regulators in the UK' in O. ESSENS, A. GERBANDY and S. LAVRIJSSEN (eds), *National Courts and the Standard of Review in Competition Law and Economic Regulation* (Europa Law Publishing, Groningen 2009) 258-61.

_____ and Tony PROSSER (1991): *Privatizing Public Enterprises*. Oxford: Clarendon Press.

GRANOVETTER, Mark S. (1985): 'Economic Action and Social Structure: The Problem of Embeddedness', *American Journal of Sociology*, 91: 481

GRANT, Wyn (2000): *Pressure Groups in British Politics*'. MacMillan Press Ltd.

GREEN, Donald and Ian SHAPIRO (1994): *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven, Conn.: Yale University Press

GREEN, Nicholas (2007): 'The early years: reflections and a view from the bar', *Competition Law Journal*, 6(4): 252

GRINYER, Peter and Arthur PRYOR (2007): 'The origins of the Competition Appeal Tribunal and the decisive role played by sir Christopher Bellamy in its development', *Competition Law Journal*, 6(4): 241

GROUT, Paul A. (1995): 'The Cost of Capital in Regulated Industries', in *The Regulatory Challenge* (M. Bishop, J. Kay & C. Mayer, eds.), chapter 16: 386-407. Oxford: OUP.

_____ and Anna ZALEWSKA (2006): 'The Impact of Regulation on Market Risk', *Journal of Financial Economics*, 80: 149-84

GROSSMAN, Sanford and Oliver HART (1982): 'Corporate Financial Structure and Managerial Incentives', in *The Economic of Information and Uncertainty* (John McCall ed.) chapter 4: 107-40. Chicago: University of Chicago Press

GUASCH, J. Luis (2004): *Granting and Renegotiating Infrastructure Concessions – Doing it right*. Washington D.C.: The World Bank

GUNNINGHAM, Neil (2009): 'Environmental Law, Regulation and Governance: Shifting Architectures', *Journal of Environmental Law*, 21(2): 179

_____ and Peter GRABOVSKY (1998): *Smart Regulation: Designing Environmental Policy*. Oxford, UK: Clarendon Press

GUTHRIE, Graeme (2006): 'Regulating Infrastructure: The Impact on Risk and Investment', *Journal of Economic Literature*, 44: 925-72

HANSSEN, F. Andrew (2000): 'Independent Courts and Administrative Agencies: An Empirical Analysis of the States', *Journal of Law, Economics and Organization*, 16: 534

HARLOW, Carol and Richard RAWLINGS (2009): *Law and Administration*. Cambridge University Press

HART, Oliver (1995): *Firms, Contracts, and Financial Structure*. Oxford: Oxford University Press

_____ and John MOORE (1999): 'Foundations of Incomplete Contracts', *Review of Economic Studies*, 66: 115

_____ and _____ (1988): 'Incomplete Contracts and Renegotiation', *Econometrica*, 56: 755

HAYEK, Friedrich A. (2002): 'Competition as discovery procedure' (trans. Marcellus S. Snow), *The Quarterly Journal of Austrian Economics*, 5(3): 9

HELM, Dieter (2009): 'Utility Regulation, the RAB and the Cost of Capital', paper presented at the Competition Commission Spring Lecture, March 6th 2009, available at: <www.dieterhelm.co.uk>

_____ (2003b): *Energy, the State, and the Market: British Energy Policy since 1979* (revised edition). Oxford: OUP.

_____ and Tim JENKINSON (1998): 'Introducing Competition into Regulated Industries', in D. HELM and T. JENKINSON (eds.), *Competition in Regulated Industries*. Oxford, UK: OUP

HICKS, John R. (1935): 'Annual Survey of Economic Theory: The Theory of Monopoly', *Econometrica*, 3(1): 1-20

HL (House of Lords, UK) (2004a): 'The Regulatory State: Ensuring its Accountability', Select Committee on the Constitution, HL 68

_____ (2004b): 'The Regulatory State: Ensuring its Accountability: The Government's Response', HL 150

HILL, Claire A. (1998): 'How Investors React to Political Risk', *Duke Journal Comp. & International Law,* 8: 283

HILLMAN, Jordan Jay and Ronald BRAEUTIGAM (1989): *Price Level Regulation for Diversified Public Utilities*. Kluwer Academic Publishers.

HODGSON, Geoffrey M. (2004): 'Opportunism is not the only reason why firms exist: Why an explanatory emphasis on opportunism may mislead management strategy', *Industrial and Corporate Change,* 13: 401-18

HOLBURN, Guy L.F. and Pablo T. SPILLER (2002): 'Institutional or Structural: Lessons from International Electricity Sector Reforms', in *The Economic of Contracts: Theories and Applications* (E. Brousseau and J-M. Glachant, eds.), 463-502. Cambridge University Press

HOLMSTRÖM, Bengt R. (1999): 'Managerial Incentive Problems: A Dynamic Perspective', *Review of Economic Studies*, 66: 169-82

HOOD, Christopher, Henry ROTHSTEIN and Robert BALDWIN (2001): *The Government of Risk*. Oxford: OUP

HOTCHKISS, Edith S., Kose JOHN, Robert M. MOORADIAN and Karin S. THORBURN (2008): 'Bankruptcy and the Resolution of Financial Distress', in B. ESPEN ECKBO, *Handbook of Corporate Finance: Empirical Corporate Finance*, vol. 2. Amsterdam, The Netherlands: Elsevier/North-Holland

HOVENKAMP, Herbert (1991): *Enterprise and American Law 1836-1937*. Cambridge, Mass. and London, England: Harvard University Press

_____ (1984): 'Technology, Politics, and Regulated Monopoly: An American Historical Perspective', *Texas Law Review*, 62: 1263

HURWICZ, Leonid (1973): 'The design of mechanisms for resource allocation', *American Economic Review*, 63: 1

_____ and Stanley REITER (2006): *Designing Economic Mechanisms*. Cambridge University Press

IRWIN, Timothy C. (2007): *Government guarantees – Allocating and valuing risk in privately financed infrastructure projects*. Washington, D.C.: The World Bank

_____ Michael KLEIN, Guillermo E. PERRY and Mateen THOBANI (eds.) (1997): *Dealing with Public Risk in Private Infrastructure*. Washington, DC: The World Bank

JACOBS, Alan M. and Steven TELES (2007): 'The Perils of Market Making: The Case of British Pension Reform', in *Creating Competitive Markets – The Politics of Regulatory Reform* (M.K. Landy, M.A. Levin and M. Shapiro, eds.), chapter 8, 157-83. Washington D.C.: Brooking Institution Press

JANSSON, Eva (2008): 'Deregulation, Property Rights, and Legal System', *European Journal of Law and Economics*, 26: 187-211

JEFFREYS, Richard (2004): *Subjective probability: The real thing*. New York: CUP

JENKINSON, Tim (2006): 'Regulation and the Cost of Capital', *International Handbook on Economic Regulation* (M. Crew & D. Parker, eds.), chapter 7. Cheltenham, UK & Northampton, MA, USA: Edward Elgar

JENSEN, Michael C. (2001): 'Value Maximization, Stakeholder theory, and the Corporate Objective Function', *European Financial Management*, 7(3): 297-317

_____ (1986): 'Agency Costs of Free Cash Flow, Corporate Finance and Takeover', *American Economic Review*, 76(2): 323-29

_____ and William H. MECKLING (1976): 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure', *Journal of Financial Economics*, 3: 305-60

JEVONS, William S. (1883): 'On the Analogy between the Post Office, Telegraphs, and Other Systems of Conveyance of the United Kingdom, as Regards Government Control', in *Methods of Social Reform*. London, Macmillan

JIRAPORN, Pornsit and Kimberly C. GLEASON (2005): 'Capital Structure, Shareholders Rights, and Corporate Governance', available at <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=792604> (accessed: 17 April 2011)

JOLLS, Christine, Cass R. SUNSTEIN and Richard THALER (1998): 'A Behavioural Approach to Law and Economics', *Stanford Law Review*, 50: 1551

JONES, Leroy P., Pankaj TANDON and Ingo VOGELSANG (1990): *Selling Public Enterprises: A Cost-Benefit Methodology*. Boston, Massachusetts: The MIT Press

JORDANA, Jacint (2005): 'Globalizing Regulatory Capitalism', *Annals of American Academy of Political and Social Science*, 598: 184

JOSKOW, Paul (2008): 'Lessons Learned from the Electricity Market Liberalization', The Energy Journal (David Newbery, special edition)

_____ (1991): 'The Role of Transaction Cost Economics in Antitrust and Public Utilities Regulatory Policies', *Journal of Law, Economics, & Organization*, 7(special issue): 53-83

_____ and Nancy L. ROSE (1989): 'The Effects of Economic Regulation', in R. SCHMALENSEE and R.D. WILLIG (eds.), *Handbook of Industrial Organization*, vol. II (Amsterdam: Elsevier), 125-208

_____ and Roger Noll (1981): 'Regulation in Theory and Practice: An Overview', in *Studies in Public Regulation* (G. Fromm, ed.), MIT Press Series on the Regulation of Economic Activity. Cambridge, MA: The MIT Press

JOWELL, Jeffrey and Dawn OLIVER (eds.) (2007): *The Changing Constitution*, 6th ed. Oxford: OUP

_____ (2003): 'Judicial deference: servility, civility or institutional capacity?' [2003] *Public Law* 592

_____ (1994): 'The Rule of Law Today', in J. JOWELL and D. OLIVER (eds.), *The Changing Constitution* (Oxford: OUP), 17-23

KAHN, Alfred E. (1988): *The Economics of Regulation: Principles and Institutions*. Cambridge, MA and London, UK: The MIT Press

KAHNEMAN, Daniel, Paul SLOVIC and Amos TVERSKY (eds.) (1982): *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge University Press

KAY, John (1993): 'Utility Regulation and the monopolies and Merger Commission: Retrospect and Prospect', in *Major Issues in Regulation*. London: Institute of Economic Affairs and London Business School

KEARNEY, Joseph D. and Thomas W. MERRILL (1998): 'The Great Transformation of Regulated Industries Law', *Columbia Law Review*, 98(6): 1323-1409

KENNELLY, Bryan (2006): 'Judicial Review and the Competition Appeal Tribunal', *Judicial Review*, 11: 160-170

KESSIDES, Ioannis (2004): *Reforming Infrastructure. Privatization, Regulation, and Competition*. The World Bank and OUP

KING, Jeff A. (2008): 'Institutional Approaches to Judicial Restraint', *Oxford Journal of Legal Studies*, 28: 409-41

KLEIN, Benjamin, Robert G. CRAWFORD and Armen A. ALCHIAN (1978): 'Vertical Integration, Appropriable Rents, and the Competitive Contracting Process', *Journal of Law & Economics*, 21(2): 297

KLEIN, Michael, Jae SO and Ben SHIN (1996): 'Transaction Costs in Private Infrastructure Projects – Are They Too high?', *Public policy for the private sector*, note 95. Washington D.C.: World Bank

KNIGHT, Frank H. (1921): *Risk, Uncertainty and Profit*. Boston, MA and New York, NY: Houghton Mifflin Co.

KOCHNAR, Rahul (1996): 'Explaining Firm Capital Structure: The Role of Agency Theory vs. Transaction Cost Economics', *Strategic Management Journal*, 17(9): 713-28

KOLBE, A. Lawrence and William B. TYE (1991): 'The *Duquesne* Opinion: How Much "Hope" Is There for Investors in Regulated Firms?', *Yale Journal on Regulation*, 8: 113

KOLE, Stacey R. and Kenneth LEHN (1997): 'Deregulation, the Evolution of Corporate Governance Structure, and Survival', *American Economic Review*, 87(2): 421-25

KOLKO, Gabriel (1965): *Railroads and Regulators, 1877-1916*. New York: Norton

KORNAI, Janos (1980) 'The Dilemmas of a Socialist Economy: The Hungarian Experience', *Cambridge Journal of Economics*, 4(2): 147

KOROBKIN, Russel B. and Thomas S. ULEN (2000): 'Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics', *California Law Review,* 88: 1051

KPMG (2004): *Peer Review of Competition Policy* (report made for the Department of Trade and Industry) (17 May), available at <http://bis.ecgroup.net/Publications/BusinessLaw/CompetitionMatters.aspx> (accessed: 02 November 2010)

KRAUS, Alan and Robert H. LITZENBERGER (1973), 'A State-Preference Model of Optimal Financial Leverage', *Journal of Finance*, 28(4): 911-22

KRIPPNER, Greta R. (2001): 'The Elusive Market: Embeddedness and the Paradigm of Economic Sociology', *Theory and Society,* 30(6): 775

KRUEGER, Anne O. (1974): 'The Political Economy of a Rent-Seeking Society', *American Economic Review,* 64: 291

LAFFE, Louis L. (1964): 'Primary Jurisdiction', *Harvard Law Review*, 77: 1037

LAFFONT, Jean-Jacques (2003): *The Principal Agent Model: The Economic Theory of Incentives*. Cheltenham, UK and Northampton, MA: Edward Elgar Publishing Ltd.

_____ (1994) 'The New Economics of Regulation Ten Year After', *Econometrica*, 62(3): 507

_____ and Jean TIROLE (2000): *Competition in Telecommunications*. Cambridge, MA and London, UK: The MIT Press

_____ and _____ (1993): *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA and London, UK: The MIT Press

_____ and _____ (1986): 'Using Cost Observation to Regulate Firms', *Journal of Political Economy* 94: 614

_____ and David MARTIMORT (2002): *The Theory of Incentives – The Principal Agent Model*. Princeton and Oxford: Princeton University Press

_____ and _____ (1999): 'Separation of Regulators Against Collusive Behaviour', *Rand Journal of Economics*, 30(2): 232

LANDIS, James M. (1938): *The Administrative Process*. New Haven: Yale University Press

LANDY, Marc K., Martin A. LEVIN and Martin SHAPIRO (eds.) (2007): *Creating Competitive Markets – The Politics of Regulatory Reform*. Washington D.C.: Brooking Institution Press

LARSON, Alexander C. (1998): 'The Efficiency of the Efficient Component-Pricing Rule: A Comment', *Antitrust Bulletin*, 43: 403

LE SUEUR, Andrew, S.A. DE SMITH, R. Hon Lord WOOLF and Jeffrey L. JOWELL (2007): *Judicial Review of Administrative Action*, 6th ed. Sweet & Maxwell.

LEGOMSKY, Stephen H. (1990): *Specialized Justice*. Oxford, UK: Clarendon Press

LEIBENSTEIN, Harvey (1966): 'Allocative Efficiency vs. "X-Efficiency", *American Economic Review*, 56(3): 392-415.

LEIGHT, Jessica (2010): 'Public Choice: A Critical Reassessment', in *Government and Markets* (E. Balleisen and D. Moss, eds.), chapter 7: 213-55. Cambridge University Press

LEVIN, Ronald M. (1985): 'Identifying Questions of Law in Administrative Law', *Georgetown Law Journal*, 74: 1-63

LEVINE, Michael E. and Jennifer L. FORRENCE (1990): 'Regulatory capture, public interest, and the public agenda. Toward a synthesis', *Journal of Law Economics, and* Organization, 6: 167-198

LEVY, Brian and Pablo T. SPILLER (1994): 'The Institutional Foundations of Regulatory Commitment: A Comparative Analysis of Five Country Studies of Telecommunications Regulation', *Journal of Law, Economics, and Organization*, 10: 201

_____ and _____ (1996): 'A framework for Resolving the Regulatory Problem', in *Regulations, Institutions, and Commitment: Comparative Studies of Telecommunications* (B. Levy and P. Spiller, eds.), 1-35. Cambridge University Press

LEVY-FAUR, David (2005): 'The Global Diffusion of Regulatory Capitalism', *Annals of American Academy of Political and Social Science*, 598: 12

LIANOS, Ioannis (2010): 'Judging' Economists: Economic Expertise in Competition Law Litigation - A European View', in *New Challenges in EC Competition Law Enforcement* (I. Lianos and I. Kokkoris (eds.). The Hague, Netherlands: Kluwer

LIDBETTER, Andrew and Veronica ROBERTS (2004): 'Expert evidence in judicial review and Competition Appeal Tribunal cases', *Competition Law Journal*, 3(2): 113

LINTNER, John (1965): 'The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets', *The Review of Economics and Statistics*, 47(1): 13-37

LITTLECHILD, Stephen (2009): 'Stipulated Settlements, The Consumer Advocate and Utility Regulation in Florida', *Journal of Regulatory Economics*, 35(1): 96

_____ (1986): *Economic Regulation of Privatised Water Authorities* (HMSO)

_____ (1983): *Regulation of British Telecommunications' Profitability* (HMSO)

_____ and G. OWEN (1980): 'An Austrian Model of the Entrepreneurial Market Process', *Journal of Economic Theory*, 23: 361-79

LOEB, Martin and Wesley A. MAGAT (1979): 'A Decentralized Method of Utility Regulation', *Journal of Law & Economics*, 22: 399

LOVE, James H. (2005): 'On the Opportunism-Independent Theory of the Firm', *Cambridge Journal of Economics*, 29: 381

LUCE, Duncan R. and Howard RAIFFA (1957): *Games and Decisions: Introduction and Critical Survey*. New York: Wiley

MACAVOY, Paul W. (1971): 'Regulation-Induced Shortage of Gas natural', *Journal of Law & Economics*, 41(1): 167

_____, W. T. STANBURY, George YARROW and Richard J. ZECKHAUSER (eds.) (1989), *Privatization and State-Owned Enterprises*. Springer

MACKIE-MASON, Jeffrey (1990): 'Do Taxes Affect Corporate Financing Decisions?', *Journal of Finance*, 45(5): 1471-93

MACNEIL, Ian (1978): 'Contracts: Adjustment of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law', *Northwestern University Law Review*, 72(6): 854

_____ (1974a): 'Restatement (Second) of Contracts and Presentiation', *Virginia Law Review*, 60: 589

_____ (1974b): 'The Many Futures of Contracts', *California Law Review*, 47: 691-816

MAJONE, Giandomenico (1994): 'The Rise of the Regulatory State in Europe', *West European Politics*, 17: 77

MARSHALL, Alfred (1920 [1890]): *Principles of Economics*. London, UK: Macmillan and Co., Ltd., 8th ed.

MARTIMORT, David (1999): 'The Life Cycle of Regulatory Agencies: Dynamic Capture and Transaction Costs', *Review of Economic Studies*, 66(4): 929

_____ (2006): 'An Agency Perspective on the Costs and Benefits of Privatization', *Journal of Regulatory Economics*, 30: 5

MASKIN, Eric and Jean TIROLE (1999): 'Unforeseen Contingencies and Incomplete Contracts', *Review of Economic Studies*, 66: 83

MAYER, Colin (2005): 'Commitment and Control in Regulation: The Future of Regulation in Water', *Governments, Competition and Utility Regulation* (C. Robinson, ed.), chapter 8: 183-98. Cheltenham, UK & Northampton, MA, USA: Edward Elgar

MCCRAW, Thomas K (1984): *Prophets of Regulation*. Cambridge, MA and London, UK: Harvard University Press.

MCCUBBINS, Matthew and Thomas SCHWARTZ (1984): 'Congressional Oversight Overlloked: Police Patrols versus Fire Alarms', *American Journal of Political Sciences*, 28(1): 165

MCNOLLGAST (MCCUBBINS, Matthew, Roger NOLL and Barry R. WEINGAST) (1987): 'Administrative Procedures as Instruments of Political Control', *Journal of Law, Economics and Organization*, 3(2): 243

_____ (1989): 'Structure and Process, Politics and Policy: Administrative Arrangements and the Political Control of Agencies', *Virginia Law Review*, 75(2): 431

MÉNARD, Claude (2008): 'A New Institutional Approach to Organization', in *Handbook of New Institutional Economics* (C. Ménard and M. Shirley, eds.), 281-318. Springer

_____ (2004): 'The Economics of Hybrid Organizations', *Journal of Institutional and Theoretical Economics*, 160: 345

MILGROM, Paul and John ROBERTS (1992): *Economics, Organization and Management* (international edition). Pearson Education

MILL, John Stuart (2004 [1848]): *Principles of Political Economy*. Amherst, NY: Prometheus Books

MILLER, Gary (1992): *Managerial Dilemmas – The Political Economy of Hierarchy*. Cambridge University Press

MILLER, Merton H. (1989): 'The Modigliani-Miller Propositions after Thirty Years', *Journal of Applied Corporate Finance*, 2: 6-18

MITNICK, Barry M. (1980): *The Political Economy of Regulation. Creating, Designing, and Removing Regulatory Forms*. New York: Columbia University Press

MODIGLIANI, Franco and Merton H. MILLER (1963): 'Corporate income Taxes and the Cost of Capital: A Correction', *American Economic Review*, 53: 443-53.

_____ and _____ (1958): 'The Cost of Capital, Corporation Finance and the Theory of Investment', *American Economic Review*, 48: 261-97

MONTI, Georgio (2008): 'Managing the Intersection of Utilities Regulation and EC Competition Law', *Competition Law Review*, 4(2): 123

MORAN, Michael (2003): *The British Regulatory State*. Oxford, UK: OUP

_____ and Tony PROSSER (eds.) (1994): *Privatization and Regulatory Change in Europe* . Philadelphia

MORRISON, Rt. Hon. Herbert (1933): *Socialisation and Transport*. London, UK: Constable & Co. Ltd.

MYERS, Steward C. (2003): 'Financing Corporations', *Handbook of the Economics of Finance, Volume 1A: Corporate Finance* (G.M. Constantinides, M. Harris and R.M. Stulz, eds.), chapter 4: 215-53. Elsevier, North Holland

_____ (1984): 'The Capital Structure Puzzle', *Journal of Finance*, 39: 575–92

_____ (1977): 'Determinants of Corporate Borrowing', *Journal of Financial Economics*, 5: 147-75

_____ and Nicolás MAJLUF (1984): 'Corporate financing and investment decisions when firms have information that investors do not have', *Journal of Financial Economics*, 13: 187–221

MYERSON, Roger (1982): 'Optimal Coordination Mechanisms in Generalized Principal-Agent Problems', *Journal of Mathematical Economics*, 11: 67

_____ (1979): 'Incentive compatibility and the Bargaining Problem', *Econometrica*, 47: 61

NAO (National Audit Office, UK) (2002): 'Pipes and Wires', report by the Comptroller and Auditor General, April 2002

_____ (2009): *The Department of Transport: The Failure of Metronet*, HC 512 Session 2008-2009 (5 June 2009), available at: <http://www.nao.org.uk/publications/0809/the_failure_of_metronet.aspx> (accessed: 17 April 2011)

_____ (2010): *Review of the UK's competition landscape* (22 March), available at <www.nao.org.uk/publications/0910/competition_landscape.aspx> (accessed: 12 April 2011)

_____ and ONS (Office for National Statistics) (2002): 'Network Rail: NAO ONS Joint Statement', news release, 24 October 2002

NEWBERY, David M. (2006) 'The Relationship Between Regulation and Competition Policy for Network Industries', CWPE 0631 and EPRG 0611

_____ (1999): *Privatization, Restructuring, and Regulation of Network Utilities*. Cambridge, MA and London, UK; The MIT Press

NOBEL PRIZE (PRIZE COMMITTEE OF THE ROYAL SWEDISH ACADEMY OF SCIENCES) (comp.) (2007): 'Mechanism Design Theory: Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2007', available at <http://nobelprize.org/nobel_prizes/economics/laureates/2007/sci.html> (accessed: 17 November 2010)

NOLL, Roger G. (2000): 'Telecommunications Reform in Developing Countries', in Anne O. Krueger (ed.), *Economic Policy Reform: The Second Stage*. Chicago, Il and London, UK: University of Chicago Press

NORTH, Douglas (1991): 'Institutions', *Journal of Economic Perspectives*, 5(1): 97

_____ (1990a): *Institutions, Institutional Change and Economic Performance*. Cambridge University Press

_____ (1990b): 'A Transaction Cost Theory of Politics', *Journal of Theoretical Politics*, 2(4): 355

OFGAS (Office of Gas) (1991): 'New Gas Tariff Formula: Economic Aspects'.

OFGEM (Office of Gas and Electricity Markets): (2010a): 'RIIO: A New Way to Regulate Energy Networks', final decision, October 2010.

_____ (2010b): 'Regulating energy networks for the future: RPI-X@20 Recommendations', July 2010.
_____ (2009a): 'Gas Distribution Annual Report for 2007-08', March 2009, document 27/09.
_____ (2009b): 'Regulating Energy Networks for the Future: RPI-X@20. Principles, Process and Issues', February 2009, document 13/09.
_____ (2009c): 'Electricity Distribution Price Control Review – Final Proposals', December 2009, document 144/09.
_____ (2008a): 'Distribution Price Control Review: Policy Document', December 2008, document 159/08.
_____ (2008b): 'Electricity Distribution Cost Review 2007-2008', December 2008, document 165/08.
_____ (2008c): 'Arrangements for Responding in the Event that a Network Company Experiences Deteriorating Financial Health', December 2008, document 158/08.
_____ (2007a): 'Gas Distribution Price Control Review: Final Proposals', December 2007, document 285/07.
_____ (2007b): 'Electricity Distribution Cost Review 2006-2007', December 2007, document 289/07.
_____ (2007c): 'Electricity Distribution Cost Review 2005-2006', January 2007, document 18/07.
_____ (2006): 'Transmission Price Control Review: Final Proposals', December 2006, document 206/06.
_____ (2004a): 'Electricity Distribution Price Control Review: Background information on the cost of capital', March 2004.
_____ (2004b): 'Electricity Distribution Price Control Review: Policy Document', March 2004, document 62/04.
_____ (2004c): 'Electricity Distribution Price Control Review: Final Proposals', November 2004, document 265/04.
_____ (2001): 'Review of Transco's Price Control from 2002: Final Proposals', September 2001.
_____ (1999a): 'Review of Public Electricity Suppliers 1998 to 2000 - Distribution Price Control Review: Update', October 1999.
_____ (1999b): 'Review of Public Electricity Suppliers 1998 to 2000 - Distribution Price Control Review: Final Proposals', December 1999.
_____ and OFWAT (2006): 'Financing Networks: A Discussion Paper', February 2006.
OFTEL (Office of Telecommunications) (1992a): 'Regulation of BT's Prices', January 1992.
_____ (1992b): 'Cost of Capital', January 1992.
_____ (1991): 'Annual Report 1990', May 1991, HMSO.
_____ (1988): 'Control of British Telecom's Prices Statement', July 1988.
_____ (1987a): 'Annual Report 1986', July 1987 HMSO.
_____ (1987b): 'Telephone Service and Price Statement', November 1987.
_____ (1986a): 'Review of British Telecom's Tariff Changes: Report by the Director General of Telecommunications', November 1986.
_____ (1986b): 'Annual Report 1985', June 1986, HMSO.
_____ (1985): 'British Telecom's Price Increases November 1985: A Statement by the Director General of Telecommunications', December 1985.

OFWAT (Office of Water Services) (2009): 'Future water and sewerage charges 2010-15: Final determinations', November 2009.

_____ (2008): 'Financial Performance and Expenditure of the Water Companies in England and Wales 2007-2008'.

_____ (2007): 'Setting price limits for 2010-15: Framework and approach –a consultation paper', October 2007.

_____ (2004): 'Future Water and Sewerage Charges 2005-2010: Final Determinations', December 2004.

_____ (2001): 'Financial Performance and Expenditure of the Water Companies in England and Wales 2000-2001'

_____ (1999): 'Future Water and Sewerage Charges 2000-05: Final Determinations', November 1999.

_____ (1994): 'Future Charges for Water and Sewerage Companies: The Outcome of the Periodic Review', July 1994.

_____ (1991): 'The Cost of Capital: A Consultation Paper', July 1991.

OGUS, Anthony (1994): *Regulation: Legal Form and Economic Theory*. Oxford, UK and Portland Oregon: Hart Publishing

OLIVER, Dawn (1987): 'Is the Ultra Vires Rule the Basis of Judicial Review?', *Public Law*, 543-5.

ORR (Office of Rail Regulation) (2008): 'Periodic Review 2008: Determination of Network Rail's Outputs & Funding for 2009-14', October 2008.

_____ (2006): 'Periodic Review 2008: Enhancing Incentives for Continuous Improvements in Performance: A Consultation Paper', July 2006.

_____ (2000): 'The Periodic review of Railtrack's Access Charges: Final Conclusions' (2 volumes)

OSTROM, Elinor (1991): *Governing the Commons*. Cambridge University Press

OXERA (2010): 'What is the impact of financeability on the cost of capital and gearing capacity?' (July 2010)

_____ (2006): 'Testing for Financeability: an assessment', report prepared for Water UK (March 2006)

PARKER, Christine (2002): *The Open Corporation: Effective Self-Regulation and Democracy*. Cambridge, UK: Cambridge University Press

PARKER, David (2009): *The Official History of Privatisation, volume 1: The Formative Years 1970-1987*. London and New York: Routledge

PATTERSON, Cleveland S. (1995): *The Cost of Capital: Theory and Estimation*. Greenwood Publishing Group

PELZTMAN, Sam (1989): 'The Economic Theory of Regulation after a Decade of Deregulation', *Brooking Papers on Economic Activity: Microeconomics*, 1-59

_____ (1976): 'Toward a more General Theory of Regulation', *Journal of Law and Economics*, 19(2): 211.

PERROW, Charles (1996): *Complex Organizations: a Critical Essay*, 3rd ed. New York: McGraw-Hill.

PIMLOTT, Nick (2010): 'Concurrency and the Role of the Competition Appeal Tribunal as Supervisor of the Sectoral Regulators', *Competition Law Journal*, 9(2): 162

POLANYI, Karl (1944): *The Great Transformation*. Boston, MA: Beacon Press

POLINSKY, A. Mitchell (1983): *An Introduction to Law and Economics*. Boston and Toronto: Little, Brown and Co.

POLLITT, Michael (2004): 'Electricity Reform in Chile: Lessons for Developing Countries', *Journal of Network Industries*, 5(3-4): 221

POOLE, Thomas (2009): 'The Reformation of English Administrative Law', *Cambridge Law Journal*, 68: 142

PORTER, Michael (1980): *Competitive Strategy – Technique for Analyzing Industries and Competitors*. New York: The Free Press

POSNER, Richard (2010): *The Crisis of Capitalist Democracy*. Cambridge, MA and London, UK: Harvard University Press.

_____ (2007): *Economic Analysis of* Law, 7th ed. Wolters Kluwer,

_____ (1999 [1969]): *Natural Monopoly and Its Regulation*. Washington, D.C.: The CATO Institute

_____ (1974) 'Theories of Economic Regulation', *The Bell Journal of Economics & Management Science*, 5(2): 335-58.

_____ (1971): 'Taxation by Regulation', *Bell Journal of Economics & Management Science*, 2(1): 22-50

POSNER, Eric (2003): 'Economic Analysis of Contract Law After Three Decades: Success or Failure?', *Yale Law Journal*, 112(4): 829-80

POWELL, Walter and Paul DIMAGGIO (eds.) (1991): *The New Institutionalism in Organisational Analysis*. Chicago and London: University of Chicago Press

PRIEST, George (1992): 'The Origins of Utility Regulation and the "Theories of Regulation" Debate', *Journal of Law & Economics*, 36: 289

PROSSER, Tony (2007): 'Regulation, Markets, and Legitimacy', in *The Changing Constitution* (J. Jowell and D. Oliver, eds.), 339-358. Oxford: OUP

_____ (2006) 'The Place of Appeals in Regulation – Continuity and Change', in *Regulatory Review 2004/05*, 195-211

_____ (2005a): *The Limits of Competition Law*. Oxford, UK: OUP

_____ (2005b): 'Regulatory Contracts and Stakeholder Regulation', *Annals of Public & Cooperative Economics*, 76(1): 35

_____ (1997): *Law and the Regulators*. Oxford: Clarendon Press

QUIRK, Paul J. (1981): *Industry Influence in Federal Regulatory Agencies*. Princeton, NJ: Princeton University Press

RAJAN, Raghuram G. and Luigi ZINGALES (1995): 'What do We Know about Capital Structure? Some Evidence from International Data', *Journal of Finance*, 50(5): 1421-60

RAMANADHAM, Venkata Vemuri (ed.) (1993): *Privatization: A Global Perspective*. London: Routledge

RAYMENT, Ben (2005): 'Powers of the Competition Appeal Tribunal', *Competition Law Journal* 4: 273

RAWLINGS, Richard (2010): 'Changed Conditions, Old Truths: Judicial Review in a Regulatory Laboratory', in, *The Regulatory State* (D. Oliver, T. Prosser and R. Rawlings, eds.), chapter 14. Oxford, UK: OUP

RIVERS, Julian (2006): 'Proportionality and variable intensity of review', *Cambridge Law Journal*, 65: 174

RODGER, Barry J. (2009a): 'Competition law litigation in the UK courts: a study of all cases 2005-2008: Part 1', *Global Competition Litigation Review*, 2(2): 93

_____ (2009b): 'Competition law litigation in the UK courts: a study of all cases 2005-2008: Part 2', *Global Competition Litigation Review*, 2(3): 136

_____ (2006): 'Competition law litigation in the UK courts: a study of all cases to 2004: Part 1', *European Competition Law Review,* 27(5): 241

ROQUES, Fabien A. and Niccos SAVVA (2009): 'Investment and Uncertainty with Price Ceilings in Oligopolies', *Journal of Economic Dynamics & Control*, 33(2): 507

ROSE-ACKERMAN, Susan and Jim ROSSI (2000): 'Disentangling Deregulatory Takings', *Virginia Law Review*, 86(7): 1435.

ROSE QC, Dinah and Tom RICHARDS (2010): 'Appeal and Review in the Competition Appeal Tribunal and High Court', *Judicial Review*, 15(3): 201

ROSS, Stephen A. (1973): 'The Economic Theory of Agency: The Principals' Problem', *American Economic Review, Papers and Proceedings*, 63: 134-9

_____ (1976): 'The Arbitrage Theory of Capital Asset Pricing', *Journal of Economic Theory*, 13: 341-60

ROSSI, Jim (2005): *Regulatory Bargaining & Public Law*. Cambridge University Press

_____ (1998): 'The Irony of Deregulatory Takings', *Texas Law Review*, 77: 297

RUBINFELD, Daniel L. and Robert D. COOTER (1989): 'Economic Analysis of Legal Disputes and Their Resolution', *Journal of Economic Perspectives*, 27: 1067

SALLY, David (1995): 'Conversation and Cooperation in Social Dilemmas', VII *Rationality and Society*, 57, 57-92

SAPPINGTON, David E. M. (1991): 'Incentives in Principal-Agent Relationships', *Journal of Economic Perspectives*, 5(2): 45

SAVAS, E. S. (2000): *Privatization and Public-Private Partnerships*. New York: Seven Bridges Press, LLC

SCARPA, Carlos (1994): 'Regulation as a Bargaining Process: Negotiation over Price and Cost-Reducing Investments', *Oxford Economic Papers*, 46(3): 357

SCHELLING, Thomas C. (1960): *The Strategy of Conflict*. Cambridge, MA: Harvard University Press

SCHLEIFER, Andrei and Robert W. VISHNY (1997): 'A Survey of Corporate Governance', *Journal of Finance*, 52(2): 737

SCHMALENSEE, Richard (1979): *The Control of Natural Monopolies*. Lexington, MA.: Lexington Books

SCHOLTZ, John T. (1984): 'Deterrence, Cooperation and the Ecology of Regulatory Enforcement', *Law and Society Review*, 18: 179

SCHOTTER, Andrew (1981): *The Economic Theory of Social Institutions*. Cambridge University Press

SCOTT, Andrew (2009): 'The Evolution of Competition Law and Policy', *LSE Law, Society and Economy Working Papers* 9/2009

SCOTT, Colin (2004): 'Regulation in Age of Governance: The Rise of the Post-Regulatory State', in *The Politics of Regulation: Institutions and Regulatory Reform for the Age of Governance* (J. Jordana and D. Levy-Faur, eds.), 145-74. Cheltenham, UK: Edward Elgar.

SELZNICK, Phillip (1985): 'Focusing Organizational Research on Regulation', in *Regulatory Policy and the Social Science* (R. Noll, ed.). Berkeley: University of California Press

SEN, Amartya (1977): 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory', *Phil. & Pub. Aff.*

SHAFFER, Mark E. (1989): 'The Credible-Commitment Problem in the Center-Enterprise Relationship', *Journal of Comparative Economics*, 13: 359

SHARPE, William F. (1964): 'Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk', *The Journal of Finance*, 19(3): 425-42

SHAVELL, Steven (2007): 'Contractual Holdup and Legal Intervention', *Journal of Legal Studies*, 26: 325

SHEARING, Clifford and Jennifer WOOD (2003): 'Nodal Governance, Democracy and the New "Denizens"', *Journal of Law and Society*, 30: 400

SIDAK, J. Gregory and Daniel F. SPULBER (1998): 'Deregulation and managed Competition in Network Industries', *Yale Journal on Regulation*, 15: 117

_____ and _____ (1997a): *Deregulatory Takings and the Regulatory Contract: The Competitive Transformation of Network Industries in the United States*. Cambridge University Press

_____ and _____ (1997b): 'Givings, Takings, and the Fallacy of Forward-Looking Costs', *New York University Law Review*, 72: 1068

_____ and _____ (1997c): 'The Tragedy of the Telecommons: Government Pricing of Unbundled Network Elements Under the telecommunications Act of 1996', *Columbia Law Review*, 97: 1081

_____ and _____ (1996): 'Deregulatory Takings and Breach of the Regulatory Contract', *New York University Law Review*, 71: 851

SIEGEL, Jeremy J. (1985): 'The Application of the DCF Methodology for Determining the Cost of Equity Capital', *Financial Management*, 14(1): 46-53

SIMON, Dan (2004): 'A Third View of the Black Box: Cognitive Coherence in Legal Decision Making', *University of Chicago Law Review*, 71: 511

SIMON, Herbert (1997): *Administrative Behavior – A Study of Decision-Making Processes in Administrative Organizations*. New York, NY: The Free Press, 4th ed.

_____ (1959): 'Theories of Decision-Making in Economics and Behavioral Science', *American Economic Review*, 49: 253-83

_____ (1957): 'A Behavioral Model of Rational Choice', in *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behaviour in a Social Setting*. New York: John Wiley and Sons, Inc.

SMITH, Adam (1991 [1776]): *The Wealth of Nations*. Amherst, NY: Prometheus Books

SMITH, John and Duncan HANNAN (2003): 'Structure of the Water Industry in England: Does it Remain Fit for Purpose?', report for Department for Environment, Food and Rural Affairs (DEFRA) and the Office of Water Services (OFWAT), November

SMITH, Warrick (1997a): 'Utility Regulators – The Independence Debate', 127 *Viewpoint*

_____ (1997b): 'Utility Regulators – Roles and Responsibilities', 128 *Viewpoint*

_____ (1997c): 'Utility Regulators – Decisionmaking Structures, Resources, and Start-up Strategy', 129 *Viewpoint*

_____ (1997d): 'Covering Political and Regulatory Risks: Issues and Options for Private Infrastructure Arrangements', in *Dealing with Public Risk in Private Infrastructure* (T. Irwin, M. Klein, G. E. Perry and M. Thobani, eds.), 45-87. Washington, DC: The World Bank

SMITH JR., Clifford W. and Jerold B. WARNER (1979): 'On Financial Contracting: An Analysis of Bond Covenants', *Journal of Financial Economics*, 7(2): 117-61

SPANN, Robert M. (1974): 'Rate of Return Regulation and Efficiency in Production: An Empirical Test of the Averch-Johnson Thesis', *Bell Journal of Economics and Management Science*, 5: 38

SPERBER, Dan and Deirdre WILSON (1995): *Relevance: Communication and Cognition*, 2nd ed. Blackwell Publishing,

SPIEGEL, Yossef (1994): 'The Capital Structure and Investment of Regulated Firms Under Alternative Regulatory Regimes', *Journal of Regulatory Economics*, 6: 297-320

_____ and Daniel F. SPULBER (1997), 'Capital Structure with Countervailing Incentives', *The RAND journal of Economics*, 28(1): 1-24.

_____ and _____ (1994): 'The Capital Structure of a Regulated Firm', *The RAND journal of Economics*, 25(3): 424-40

SPILLER, Pablo T. (1995): 'Regulatory Commitment and Utilities' Privatization: Implications for Future Comparative Research', in *Modern Political Economy* (J. S. Banks and E. A. Hanushek, eds.), chapter 2: 63-79. Cambridge University Press.

_____ and Emerson H. TILLER (1996): 'Invitations to Override: Congressional Reversals of Supreme Court Decisions', *International Review of Law & Economics*, 16: 503-21

_____ and Mariano TOMMASI (2003): 'The Institutional Foundations of Public Policy: A Transactions Approach with Application to Argentina', *Journal of Law, Economics and Organization*, 19(2): 281-306

SPULBER, Daniel F. (1989): *Regulation and Markets*. Cambridge, MA and London, UK: The MIT Press

STANSBURY, Philip R. (1990): 'Planning Against Expropriation', *International Law*, 24: 677

STEIN, Jeremy C. (2003): 'Agency, Information and Corporate Investment', in *Handbook of the Economics of Finance, Volume 1A: Corporate Finance* (G.M. Constantinides, M. Harris and R.M. Stulz, eds.), chapter 2: 111-65. Elsevier, North Holland

STEPHENSON, Matthew C. (2006): 'The Strategic Substitution Effect: Textual Plausibility, Procedural Formality, and Judicial Review of Agency Statutory Interpretations', *Harvard Law Review*, 120: 528

STERN, Jon and Stuart HOLDER (1999): 'Regulatory Governance: Criteria for Assessing the Performance of Regulatory System. An Application to Infrastructure Industries in the Developing Countries of Asia', *Utilities Policy*, 8(1): 33

STEYN, Lord (2005): 'Deference: a tangled story', [2005] *Public Law* 346

STIGLER, George (ed.) (1988): *Chicago Studies in Political Economy*. Chicago, IL and London, UK: The University of Chicago Press

_____ (1971): 'The Theory of Economic Regulation', *Bell Journal of Economics & Management Science*, 2(1): 3

STONE-SWEET, Alec (1999): 'Judicialization and the Construction of Governance', *Comparative Political Studies*, 31: 147.

STONES, Clive (2007): 'Risk Sharing, the Cost of Equity and the Optimal Capital Structure of the Regulated Firm', *Review of Industrial Organization*, 30: 139-59

SUNSTEIN, Cass R. (ed.) (2000): *Behavioral Law and Economics*. Cambridge University Press

_____ (1990a): *After the Rights Revolution – Reconceiving the Regulatory State*. Cambridge, MA, and London, England: Harvard University Press

_____ (1990b): 'Paradoxes of the Regulatory State', *University of Chicago Law Review*, 57: 407

_____ and Adrian VERMEULE (2003): 'Interpretation and Institutions', *Michigan Law Review,* 101: 885

TAGGART JR., Robert A. (1985): 'Effects of Regulation on Utility Financing: Theory and Evidence', *Journal of Industrial Economics*, 33(3): 257-76.

_____ (1981): 'Rate-of-Return Regulation and Utility Capital Structure Decisions', *Journal of Finance*, 36: 383-93

TARGETTI, Ferdinando (ed.) (1991): *Privatization in Europe: West and East Experiences*. Dartmouth, Aldershot.

THATCHER, Mark and Alec STONE SWEET (2002): 'Theory and Practice of Delegation to Non-Majoritarian Institutions', *West European Politics*, 25(1): 1-22.

THORELLI, Hans B. (1986): 'Networks: Between Markets and Hierarchies', 7 *Strategic Management Journal* 37

TILLER, Emerson H. (1998): 'Controlling Policy by Controlling Process: Judicial influence on regulatory Decision Making' 14 *Journal of Law, Economics and Organization* 114

TIROLE, Jean (2006): *The Theory of Corporate Finance*. Princeton and Oxford: Princeton University Press

TITMAN, Sheridan (1984): 'The Effect of Capital Structure on a Firm's Liquidation Decision', *Journal of Financial Economics*, 13(1): 137-51

TOULMIN, Stephen E. (1972): *Human Understanding: The Collective Use and Evolution of Concepts*. Princeton: Princeton University Press.

TROXEL, Emery (1947): *Economics of Public Utilities*. New York, NY: Rinehart & Company, Inc.

TULLOCK, Gordon (1967): 'The Welfare Cost of Tariffs, Monopoly, and Theft', 5 *West. Econ. J.* 224

VELJANOVSKI, Cento (2010): 'Strategic use of regulation', in *The Oxford Handbook of Regulation* (R. Baldwin, M. Cave and M. Lodge, eds.), chapter 5: 85-103. Oxford, UK: Oxford University Press.

_____ (1991): 'The Regulation Game', in C. VELJANOVSKI (ed.), *Regulators and the Market* (London: The Institute of Economic Affairs)

_____ (1987): *Selling the State* (London, England: Weidenfeld and Nicholson)

VICKERS, John (2009): 'Competition Policy and Property Rights', *University of Oxford Department of Economics Discussion Paper Series* No. 436

_____ (1997): 'Regulation, Competition, and the Structure of Prices', 13 *OXREP* 15

_____ and George YARROW (1988): *Privatization: An economic Analysis*.

VISCUSI, W. Kip, John M. VERNON and Joseph E. HARRINGTON JR. (2005): *Economics of Regulation and Antitrust* (Cambridge, MA and London, UK: The MIT Press, 4th ed.)

VOGEL, Steven K. (1996): *Freer Markets, More Rules – Regulatory Reform in Advanced Industrial Countries* (Ithaca and London: Cornell University Press)

VON THADDEN, Ernst-Ludwig (1995): 'Long-Term Contracts, Short-Term Investment and Monitoring', 62 *Rev. Econ. Studies* 557

WADE, H.W.R. and C.F. FORSYTH (2004): *Administrative Law* (Oxford, UK: OUP, 9th ed.)

WEAVER, R. Kent and Bert A. ROCKMAN (1993): *Do Institutions Matter? Government Capabilities in the United States and Abroad* (Washington: Brooking Institution Press)

WEINGAST, Barry R. (1993): 'Constitutions as Governance Structures: The Political Foundations of Secure Markets', 149(1) *Journal of Institutional and Theoretical Economics* 286

WHISH, Richard (2007): 'Appeals to the Competition Appeal Tribunal under the Competition Act 1998 and the Enterprise Act 2002', in Peter Vass (ed.), *Regulatory Review 2006/7* (Bath, UK: Centre for the Study of Regulated Industries)

WILKS, Stephen (1999): *In the Public Interest – Competition Policy and monopolies and Merger Commission* (Manchester, UK: Manchester University Press)

WILLIAMS, Stephen F. (1996): 'Deregulatory Takings and Breach of the Regulatory Contract: A Comment', 71 *New York University Law Review* 1000

WILLIAMSON, Oliver (2000): 'The New Institutional Economics: Taking Stock, Looking Ahead', 38(3) *J. Econ. Lit.* 595

_____ (1998): 'Transaction Cost Economics: How it Works; Where it is Headed', 146(1) *De Economist* 23

_____ (1996a): 'Deregulatory takings and breach of the Regulatory Contract: Some Precautions', 71 *New York University Law Review* 1007

_____ (1996b): *The Mechanisms of Governance* (New York and Oxford: OUP)

_____ (1989): 'Transaction Cost Economics', in *Handbook Industrial Organization*, vol. I, (Amsterdam: Elsevier).

_____ (1988): 'Corporate Finance and Corporate Governance', *The Journal of Finance*, 43(3): 567-91

_____ (1985): *The Economic Institutions of Capitalism* (New York: The Free Press)

_____ (1984): 'Corporate Governance', *The Yale Law Journal*, 93: 1197-230

_____ (1975): *Markets and Hierarchies: Analysis and Antitrust Implications* (Free Press)

_____ (1973): 'Markets and Hierarchies: Some Elementary Considerations', *American Economic Review*, 63(2): 316-25.

WILLIG, Robert D. (1979): 'The Theory of Network Access Pricing', in H.M. Trebing (ed), *Issues in Public Utility Regulation* (The Institute of Public Utilities)

WILSHER, Daniel (2008): 'Judicial Review of Economic Regulators in the UK by the High Court: Is Ultra Vires Enough to Satisfy the Demands of the European Economic Constitution?', 7(1) *Comp. L. J.* 83

WILSON, James Q. (1991): 'Comparative Economic Organization: The Analysis of Discrete Structural Alternatives', 36 *Adm. Sci. Quarterly* 269

_____ (1980): *The Politics of Regulation* (New York: Basic Books, Inc.)

WILSON, W.A. (1969): 'Questions of Degree', 32 *Modern Law Review* 361-76

WRIGHT, Stephen, Robin MASON and David MILES (2003): *A Study into Certain Aspects of the Cost of Capital for Regulated Utilities in the UK*. Smithers & Co. Ltd.

_____, _____, Steve SATCHELL, Kenjiro HORI and Meltem BASKAYA (2006): *Report on the Cost of Capital*, Smithers & Co. Ltd.

YOUNG, Alison L. (2009): 'In Defence of Due Deference', 72 *Modern Law Review* 554-80