DISCUSSION PAPER SERIES

DP13801

USE AND ABUSE OF REGULATED PRICES IN ELECTRICITY MARKETS: "HOW TO REGULATE REGULATED PRICES?"

David Martimort, Jerome Pouyet and Carine Staropoli

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

USE AND ABUSE OF REGULATED PRICES IN ELECTRICITY MARKETS: "HOW TO REGULATE REGULATED PRICES?"

David Martimort, Jerome Pouyet and Carine Staropoli

Discussion Paper DP13801
Published 14 June 2019
Submitted 06 June 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

USE AND ABUSE OF REGULATED PRICES IN ELECTRICITY MARKETS: "HOW TO REGULATE REGULATED PRICES?"

Abstract

We consider the regulation of the tariffs charged by a public utility in the electricity sector. Consumers differ in terms of their demands which are private information. When regulating the firm's tariffs, the government is concerned by redistribution across consumers classes. A conflict between redistribution and screening induces production distortions even when the firm is a monopoly. Introducing competition with an unregulated fringe may improve efficiency but jeopardizes redistribution. In response, the government may now want to manipulate information about the incumbent's cost so as to restrict entry and better promote its own redistributive objective. To prevent such obstacle to entry, the government's discretion in fixing regulated tariffs of the incumbent should be restricted. This can be done by imposing floors or caps on those tariffs and/or by controlling the market share left to the competitive fringe. We highlight the determinants of such limits on discretion and unveil to what extent they depend on the government's redistributive concerns.

JEL Classification: L51, L94, L98, Q41, Q48

Keywords: Electricity markets, regulated tariffs, government's redistributive concerns, optimal discretion

David Martimort - david.martimort@psemail.eu
Paris School of Economics and CEPR

Jerome Pouyet - jerome.pouyet@gmail.com
ESSEC Business School and CEPR

Carine Staropoli - carine.staropoli@univ-paris1.fr
Paris School of Economics

Acknowledgements
We thank Altermind and seminar participants at ESSEC for stimulating discussions. All remaining errors are ours.

Use and Abuse of Regulated Prices in Electricity Markets: "How to Regulate Regulated Prices?"*

David Martimort† Jérôme Pouyet‡ Carine Staropoli§

June 6, 2019

Abstract

We consider the regulation of the tariffs charged by a public utility in the electricity sector. Consumers differ in terms of their demands which are private information. When regulating the firm's tariffs, the government is concerned by redistribution across consumers classes. A conflict between redistribution and screening induces production distortions even when the firm is a monopoly. Introducing competition with an unregulated fringe may improve efficiency but jeopardizes redistribution. In response, the government may now want to manipulate information about the incumbent's cost so as to restrict entry and better promote its own redistributive objective. To prevent such obstacle to entry, the government's discretion in fixing regulated tariffs of the incumbent should be restricted. This can be done by imposing floors or caps on those tariffs and/or by controlling the market share left to the competitive fringe. We highlight the determinants of such limits on discretion and unveil to what extent they depend on the government's redistributive concerns.

Keywords: Electricity markets; regulated tariffs; government's redistributive concerns; optimal discretion.

JEL Code: L51, L94, L98, Q41, Q48.

*We thank Altermind and seminar participants at ESSEC for stimulating discussions. All remaining errors are ours.

†Paris School of Economics (EHESS). E-mail: david.martimort@psemail.eu. Address: PSE, 48 boulevard Jourdan, 75014 Paris, France.

‡THEMA-CNRS, ESSEC Business School, Université de Cergy-Pontoise. E-mail: pouyet@essec.edu. Address: ESSEC Business School, Avenue Bernard Hirsch, B.P. 50105, 95021 Cergy, France.

§Paris School of Economics-Université Paris 1. E-mail: carine.staropoli@univ-paris1.fr. Address: PSE, 48 boulevard Jourdan, 75014 Paris, France.

1. Introduction

In what is by now considered as a reference book *Market for Power* (1983), Paul Joskow and Richard Schmalensee presented various proposals to reform the electric power industry. Back then, the traditional model of monopolistic provision of electricity by publicly owned or heavily regulated utilities had already raised enough skepticism to consider that competition would be the best vehicle to improve industry performance and foster innovation. Joskow and Schmalensee did not provide a single *"one size fits all"* recipe for how to move from heavy government regulation to a more competitive field. Instead, they modestly acknowledged that no reform would represent a panacea to solve the genuine market failures of electricity markets. Echoing this prediction, the paths towards liberalization have greatly differed throughout the world even though they also have featured common elements. For instance, while generation, transmission, distribution and retail activities had always been bundled together within large vertically-integrated public utilities, most countries chose to unbundle those segments. Transmission and local distribution, which are viewed as being natural monopolies, have been separated from generation and retail activities, which were more prone to competition. At the same time, the standard *"cost-of-service regulation,"* that was designed to ensure that the price for electricity would cover the average production cost, was replaced by some forms of *"market-based pricing,"* at least on the more competitive segments of the supply chain. A rationale for introducing competition in generation was that economies of scale on this segment were modest and the potential benefits of mixing different sources of generation for the security of supply were supposedly substantial.

Competition in retail markets remains rather marginal in contrast to what happened at the generation level. To illustrate, in the E.U., the yearly switching rate of household customers from incumbent suppliers to new entrants remains quite small (around 6.5 percent on average) and retailers concentration is still significant even a few years after liberalization.[1] Several facts may explain such limited switching activities. First, competition has sometimes been slowly introduced.[2] Second, some countries still maintain some sort of price regulation that is found attractive by customers. Indeed, regulated and non-regulated retail tariffs (the so-called *"market offer"*) coexist in most cases. In other words, while the debate on the benefits of competition has focused on the technological features of the market, the repeated failures of competition to emerge suggest that scholars and practitioners should give more attention to the political and regulatory landscape that surrounds reforms in the electricity sector.

For decades, governments have indeed always expressed some sorts of redistributive concerns in the way they were regulating electricity tariffs. Advocates for such regulation would argue that high prices disproportionately impact poor households since energy expenditures represent a large fraction of their income. Governments have thus used different tricks to protect low-income households and reduce inequality across different classes of customers.[3] However, most governments also have other objectives on their

[1]The switching rate for non-household customers is at least twice as large as that of household customers in most countries.

[2]Even though the first directive for the common rules for the E.U. electricity market (Directive 96/92/EC) was already adopted in 1996, it is only in 2003 that Directive 2003/54/EC stipulated that this market would be fully opened by 07/01/2007. By that date, all electricity consumers (household and non-household) had the right to choose electricity suppliers.

[3]For instance, reduced electricity rates for low-income households, increasing block pricing, introduc-

agenda, their recent concerns towards environmental and climate change issues being two examples in order. These objectives call for improving energy efficiency and reducing electricity consumption and, as such, they might conflict with redistributive concerns. The optimal strategy for balancing those conflicting goals remains unsettled and subject to considerable political debate.

This complex picture of the economic and political landscapes behind electricity markets raises a number of important questions. What are the economic and political rationales behind the coexistence of regulated and non-regulated tariffs on those markets? How can we explain that only a small fraction of the overall demand ends up being served through *"market offers"* when efficiency considerations would call for a more open field? Is it possible that regulated tariffs might actually hinder the development of retail competition and somewhat erect inefficient barriers to entry on this segment of the supply chain? If regulated tariffs do harm competition, to what extent might those tariffs be actually manipulated to unduly favor the incumbent on the fallacious political grounds that regulated tariffs would shelter vulnerable customers from high prices?

Model and Results. To address these questions, we develop a model of retail price regulation for the electricity sector that intertwines economic and political considerations. The first key element is that customers belong to different groups that differ according to their specific demand, typically their willingness to pay for electricity. The supplier should thus rely on a nonlinear price to screen those different types of customers and better extract their surplus in a context where those types remain non-observable. To model regulation in a nutshell, we assume that the government can perfectly control the firm.[4] An immediate consequence of this assumption is that everything happens as if the firm's nonlinear price was *de facto* decided by the government itself. This assumption is meant to capture the long-lasting and closely knitted relationship between incumbent operators and public officials and the corresponding congruence of their interests.[5]

The second key element is that the government has redistributive concerns and maximizes a weighted sum of the utilities of the various types of customers. The distribution of the social weights associated to the different classes of customers captures the very nature of the redistributive concerns. This assumption may also be viewed as a shortcut for the fact that the government supports the sector by increasing production and consumption beyond efficiency. In any case, this assumption will also justify the existence of regulated tariffs and a wedge between prices and marginal costs even in a market environment where the incumbent faces the pressure of competitors. We will be *a priori* quite agnostic on the nature of the redistributive bias. The government may either favor low- or high-demand customers depending on the distribution of social weights that prevails.

Much in the spirit of the more traditional *"cost-of-service regulation,"* regulated tariffs must ensure that the firm's revenues cover its fixed cost.[6] When the retail segment

tion of electricity price-caps, social tariffs for vulnerable customers, energy vouchers that can be used to pay an energy bill or to cover the costs of home energy renovation, etc.

[4]In particular, we depart from the asymmetric information issue that has been the cornerstone of the whole modern theory of Regulatory Economics (Baron and Myerson, 1982; Laffont and Tirole, 1993; Laffont, 1994; Armstrong and Sappington, 2007).

[5]Marcel Boiteux, who was both the author of seminal papers on optimal electricity pricing (Boiteux, 1956) and the former CEO of the French public company EDF from 1979 to 1987, delivers his memories in a book (Boiteux, 1993) where he provides insightful feedback regarding public decision-making and the closely knitted relationship between the government and the management of a state-owned enterprise.

[6]Berg and Tschirhart (1988).

has been opened to competition, these tariffs must also be designed with an eye on the customers' incentives to switch to potential entrants. We characterize the optimal regulated tariffs, the corresponding consumption profile and the market segmentation between the incumbent and its competitors in those circumstances. A particular emphasis is given to the distortions induced by the government's political biases and how those distortions are modified by the threat of entry once competition is possible.

To understand the basic mechanisms at play, consider first the case without competition. This simplifying scenario is meant to depict the old institutional setting where the incumbent enjoyed unchallenged monopoly position. To illustrate, suppose there are two types of customers: low-valuation customers who can only afford a low price for electricity and high-valuation customers who are willing to pay a higher price for electricity.[7] If the government favors low-valuation customers, most of the fixed cost of production should be covered with the tariff charged to high-valuation customers. In response to this extra charge, high-valuation customers are willing to shade their demands to reduce their bill. When the customers' valuations are private information, incentive compatibility can be restored by reducing the consumption of low-valuation customers and making their allocation less attractive to high-valuation types; a standard result from the screening literature. Hence, a redistributive bias towards low-valuation customers calls for distorting downwards consumption and production with respect to efficiency.[8] When the government has instead a bias towards high-valuation customers, the reverse conclusion obtains: over-consumption arises. The key point is that regulated tariffs that promote a given redistributive objective must remain incentive-compatible and incentive-compatibility calls for allocative distortions.

The consequences on pricing of the government having redistributive concerns in such monopolistic environment are thus straightforward. To distort consumption downwards, as requested when the government favors low-valuation customers, marginal prices should be set below the firm's marginal cost. Redistributive concerns then justify below-cost pricing. This is a feature typically found in electricity markets that rely on increasing-block tariffs like, for instance, in South Africa, Australia, California, Japan or Ontario.[9] When the government is instead biased towards high-valuation customers, marginal prices remain above marginal cost.

Consider now the possibility of competition between the incumbent, which remains regulated, and a fringe of potential entrants. The fringe is competitive and prices at its marginal cost. The fringe's marginal cost is a random variable. This assumption allows us to link the probability of entry to various fundamental parameters of the market. Moreover, such randomness also echoes the fact that competitive generators supplying the fringe are more likely to be hit by supply shocks. Under those circumstances, efficiency dictates that the fringe should serve the whole demand as soon as it is more efficient than the incumbent. Given that the government has still some redistributive concerns, the regulated tariffs charged by the incumbent certainly impact on the probability of entry of potential competitors. Surprisingly, this impact does not depend on the nature of the government's redistributive concerns. Regulated tariffs are always chosen so as to favor the

[7] The model does not distinguish whether the difference between the customers' willingnesses to pay comes from the fact that high-valuation customers have higher revenues or higher needs for electricity.

[8] Because it cannot be stored at any reasonable cost, electricity must be produced at the same time it is consumed. Consumption is equal to production at all points in time (ignoring system's balancing services used to maintain the security and quality of electricity supply).

[9] On this issue, see Borenstein (2012).

incumbent and limit entry. Indeed, the cost of unregulated market-based pricing is that it never redistributes utility as the government would. For instance, under unregulated prices high-valuation customers may consume excessively, whereas low-valuations ones may not consume enough compared to the consumption profile that a government willing to redistribute towards the latter would induce through the convenient choice of regulated prices. There is thus a fundamental conflict between promoting redistributive concerns through the regulated tariffs of the incumbent and implementing fair competition on the retail market.[10]

The regulated tariffs and the allocation of market shares between the incumbent and its competitors both depend on the incumbent's cost. This cost is also known by the government, a consequence of our assumption of perfect regulatory control. In fact, knowledge of this cost structure is critical on efficiency grounds, not only to determine the optimal probability of entry but also the consumption level that should prevail when the incumbent serves the market. By the very nature of its long-lasting relationship with the incumbent, the government is ideally positioned to induce an efficient switching rule between the incumbent and the fringe. Yet, efficiency conflicts with the government's redistributive objectives. When information on the incumbent's cost structure is private, manipulating the incumbent's cost, and thereby limiting entry, is thus a way for an opportunistic government to promote its own redistributive concerns.

There is thus a fundamental trade-off when the government is granted full discretion in fixing regulated tariffs for the incumbent.[11] On the one hand, full discretion allows to use information on the incumbent's cost to better tailor production and entry decisions.[12] On the other hand, it also biases production and induces an inefficiently low probability of entry. In response to this fundamental trade-off, and if efficiency goals have to be maintained, regulated tariffs should also be somewhat regulated. Building on the literature on optimal delegation in organizations, we show that it is indeed optimal to limit the government's discretion.[13] When the government favors high-valuation customers, it might want to pretend that the incumbent's cost is lower than its true value so as to boost production and consumption. It does so with low marginal tariffs. Avoiding such manipulation requires to set a floor on regulated tariffs together with a minimal market share for entrants. When the government instead favors low-valuation customers, it might pretend that the incumbent's cost is greater than its true value so as to depress production and consumption. A cap on regulated prices might then be preferred and the share of the market left to the fringe should be limited.

Literature Review. Our paper touches, combines and contributes to several topics

[10]We focus on the retail market and ignore the strategic interactions between the retail and the wholesale markets, which is a salient issue when there remains some kind of vertical integration between those two segments as it is the case in many countries.

[11]In the case of countries where state-owned public utilities prevailed, liberalization further necessitated a clear separation of the roles of the state as a regulator and as an owner. This separation can be achieved either by privatizating the service or/and by establishing an independent regulatory authority. Regulatory independence avoids political interference in business decisions. However, energy can hardly be depolitized and some discretion for the government, especially in setting up regulated tariffs, generally remains.

[12]Strictly speaking, unless retailers are vertically integrated, they do not produce electricity themselves. Retailers rather purchase electricity on the wholesale market from producers and re-sell it to final customers. When we refer to their "production," we ignore this difference.

[13]See, among others, Holmström (1984), Melumad and Shibano (1991), Martimort and Semenov (2006), Alonso and Matouschek (2008), Amador and Bagwell (2013).

in economics. Each of those actually corresponds to a different building block of our analysis.

Nonlinear pricing and redistribution. The design of the tariff structure by an incumbent firm selling a good to a population of heterogeneous customers has been extensively studied by the nonlinear pricing literature (Goldman, Leland and Sibley, 1984; Maskin and Riley, 1984; Wilson, 1993, among many others). When extraction of the customers' surplus is the firm's sole objective, this design is well-known. Consumption is always distorted downwards to facilitate rent extraction. The specificity hereafter is that the government (which perfectly controls the firm and thus sets the tariff itself) has some redistributive goals beyond efficiency. The literature on this particular issue is much more sparse. Related analyses from a technical viewpoint are found in Locay and Rodriguez (1992) and Vercammen, Fulton and Hyde (1996), although these articles address discriminatory pricing by firms or self-managed organizations in very different contexts. With redistributive concerns, the conflict between efficiency and redistribution is inherently different from the conflict between efficiency and rent extraction that prevails in the rest of the nonlinear pricing literature. As a result, consumption distortions are also somewhat different; over- or under-consumption may arise depending on the nature of the government's redistributive objectives. Prices are above (resp. below) marginal cost when low-valuation (resp. high-valuation) customers are targeted.

Pricing by public utilities and redistribution. The redistributive consequences of the pricing structure of public utilities is an old concern of Public Economics, going back at least to Atkinson and Stiglitz (1976). Putting aside incentive compatibility, Feldstein (1972a, 1972b) offers complete information models to analyze how to finance production fixed costs when customers may differ in terms of their willingnesses to pay and are all offered the same two-part tariff. One important issue that remains to understand is why price distortions should be preferred to more direct forms of redistributive transfers. An obvious response, often taken as granted by the existing literature, is that more traditional means of redistribution, like the income tax, are actually badly designed. One possibility investigated by Stiglitz (1982) and Naito (1996) is that asymmetric information prevents the use of first-best redistribution. Building on a two-dimensional screening problem (with both tastes and talent being private information of customers), Cremer and Gahvari (2002) argue that nonlinear pricing plays an effective redistributive role and is not a substitute for an ill-designed income tax policy. So doing, these authors provide a normative justification for universal service requirements and other similar regulatory policies found in practice.[14] Taking a more positive perspective, Posner (1971) argues that a regulator subject to political pressures may implement some form of cross-subsidization across different classes of customers, leading to what has been coined as some kind of "*taxation by regulation*."

Of course, it is often hard to disentangle the true rationales for price distortions and whether these distortions are due to genuine social preferences for equity and fairness, whether they are induced by some political pressures, or whether they are just a way to promote and/or harm competition. In this respect, our model is consistent with both a normative and a positive interpretation of the government's objectives. To the best of our knowledge, it is also the first model to explicitly introduce the costs and benefits of competition in such a setting.

[14] Similar issues have been addressed in related contexts by Cremer, Gahvari and Ladoux (1998), Bovenberg and Goulder (2002), and Russo (2015) among others.

Pricing by public utilities, contestability and entry. A long tradition in Regulatory Economics, especially following the seminal contribution by Baumol, Panzar and Willig (1982), addresses the consequences on pricing of the potential entry of competitors on contestable segments of a multi-product monopolist's activities. The sustainability of a given price structure is there entirely determined by the detailed properties of the monopolist's cost function (like its sub-additivity). Demand considerations play little role beyond the mere verification that Ramsey prices may (or may not) be sustainable.[15] One criticism of this line of research is that entry acts only as a threat. We depart from this literature by making entry an event of non-zero probability, thanks to some randomness in the cost of the competitive fringe. This assumption allows us to discuss the nature and determinants of the market share left to competitors, especially in view of the government's redistributive concerns.

Contracts as a barrier to entry. This latter approach is thus reminiscent of the I.O. literature on contracts as a barrier to entry. Following Aghion and Bolton (1987),[16] this literature looks at the welfare consequences of having manufacturers and retailers signing vertical contracts in a context where potential entrants, with random costs, may also provide the service. Inefficient entry may arise in response to imperfect extraction of the entrants' information rent. Formally, our contractual structure is quite similar. The incumbent's tariff is a contract passed with customers and that contract not only determines consumption with the incumbent but also the likelihood of entry. We show that inefficient entry always arises since the redistributive concerns of the government are never taken into account by the fringe.

Optimal delegation and institutional design. We append to this vertical structure between the government-incumbent pair and customers an extra layer aimed to model the possible *ex ante* constitutional constraints that may be put on regulated tariffs to avoid their manipulation by public officials. In other words, we determine the optimal degree of discretion that should be left to a government in regulating tariffs. The Political Science literature has long investigated the issue of the optimal mandate of a public official, stressing the trade-off between setting "*inflexible rules*" and leaving "*discretion*" to biased officials to rely on their better knowledge of market conditions (Epstein and O'Halloran, 1994, 1999; Huber and Shipan, 2002, 2006). This literature is rather abstract in that the policy being controlled lies in some abstract space and the conflict of interest with the constitutional level being fixed at the outset. Our approach is more prosaic, a policy being thereafter a set of regulated tariffs and an entry requirement. Relatedly, the nature of regulatory mandates in various structured environments has been investigated by Laffont (1996) for privatization, Boyer and Laffont (1999) for environmental regulation, Hiriart and Martimort (2012) for risk regulation and Iossa and Martimort (2016) for the design of procurement contracting and public-private partnerships.

Organization. Section 2 provides a brief overview of the restructuring of electricity markets, emphasizing reforms that have targeted the retail segment. This presentation serves as a background motivation for our analysis, which starts in Section 3 with the presentation of our model. Section 4 analyzes the optimal tariffs and consumption profiles without competition. Section 5 considers entry and discusses how our previous findings are modified. Section 6 studies why and how the government's discretion ought to be limited. Section 7 concludes. All proofs are relegated to an Appendix.

[15] See Faulhaber and Levinson (1981) or Spulber (1984) among others.

[16] See also Choné and Linnemer (2016) and Martimort, Pouyet and Stole (2019).

2. Retail Competition in Electricity Markets: A Quick Overview

This section provides a quick overview of various reforms that have impacted the electricity market. The emphasis is on retail distribution. We highlight that, in practice, retail competition is still impacted by regulatory intervention.

2.1. Price Regulation

The motivation for price regulation in the electricity market has historically been based on two ideas. First, a vertically integrated monopoly overcharges customers due to a lack of competitive pressure. Second, unlike other commodities, the physical characteristics and economic attributes of electricity imply that supply and demand may fluctuate a lot, threatening system security and price stability. The form of price control that prevailed was "*cost of service*" regulation. Such regulation aimed to replicate competition and protect consumers (especially vulnerable consumers) from the exercise of monopoly power. It didn't provide incentives for the containment of operating and investments costs but instead supported huge investments in essential facilities that were indeed needed at that time. It is by now acknowledged that this kind of regulation did not meet productive and allocative efficiency especially when the firm is able to keep its informational advantage.

The liberalization and restructuring of the electricity sector have somewhat changed the rationales for price regulation. Although the price of electricity has now some market-based components, some regulated components remain. Indeed, the bill that customers pay typically reflects the costs at the different stages of the supply chain. Those costs include the price for energy components (wholesale and retail prices) and the costs of transmission and distribution networks. They also include the taxes, fees and charges that are meant to cover policy-related costs such as environmental concerns. The higher those taxes are, the stronger the political bias might be.[17,18] The network and taxes components are by nature regulated. The wholesale generation cost component are determined by generators competing on the wholesale market, and the supply component by suppliers present in the retail market. The price of electricity is therefore both regulated and competitive.

2.2. Competition on Retail Markets

The opportunity to open electricity retail to competition and the way to proceed were not discussed before the controversy between Joskow (2000) and Littlechild (2000), even if retail activities were potentially competitive given the limited fixed and sunk costs. On one hand, advocates of a system of "*wholesale spot price pass-through*" considered

[17] Typically, investments in renewable energy sources and smart grid technology are a key driver of climate change policy and call for specific taxes and pricing schemes. Such environmental or redistribution-oriented taxes include, for instance, charges related to support schemes for renewable energy sources, costs of a country's energy efficiency measures, nuclear levies or territorial and social redistribution schemes.

[18] There are wide differences in the share and the type of taxation imposed on electricity consumers across countries. These differences reflect variations in national policy priorities and energy policies. In Denmark for instance, more than two thirds (67.7%) of the final price is made up of VAT, taxes and levies, whereas in Malta that share is only 4.78% (Eurostat, 2018).

that electricity retailers were not offering enough value-added services to consumers to justify competition among electricity retailers (Joskow, 2000). Joskow (2000)'s proposal was thus to bypass the retail market by allowing distributors to offer households and non-households the wholesale price plus a (regulated) mark-up designed to cover network and commercial costs.

On the other hand, Littlechild was the first to support competition at the retail level (Beesley and Littlechild, 1983), as well as the first regulator to implement such competition in practice.[19] Based on the separation of distribution and retail supply, the objective was to let consumers choose their retail supplier according to the type and terms of tariffs, services (such as demand response, energy efficiency or self-consumption) and equipments (like smart meters and remote control devices) offered.[20] As a matter of fact, the recent widespread diffusion of smart meters and the development of so-called smart grids have considerably increased the potential benefits of such competition among retailers (Gangale et al., 2017).[21] Retail competition was adopted first in Norway and in Britain in the late 1990s and finally applied to all consumers in the E.U. in 2007.[22]

2.3. An Incomplete Process Because of the Persistence of Price Regulation

As of 2019, the introduction of competition into electricity retail activities in the E.U. presents a mixed picture. Although all European countries should theoretically follow the same agenda, this is, however, not the case in practice.[23]

Indeed, in addition to high retail market concentration,[24] customers' switching rate is one of the key indicators for competitive development in energy retail markets. In the E.U., customers, especially households, mostly remain with their incumbent supplier.[25] Behavioral aspects are seen as important to explain such a situation.[26] Another explanation lies in the persistence of barriers to entry linked to regulation and policy-related intervention. To illustrate, some countries have still a dual retail market structure, whereby regulated and non-regulated markets exist in parallel. These countries have not

[19]Littlechild was at that time head of the British Office of Electricity Regulation.

[20]Innovative tariffs typically include "*dual-fuel tariffs*" (to allow customers to source both electricity and gas from a single retail supplier), "*green energy tariffs*" (electricity is sourced from renewable energy sources), or any "*dynamic pricing*" (to lessen the risk onto consumers associated to wholesale electricity price variations).

[21]Smart meters allow new bi-directional information flows between consumers and generation, thereby allowing risk and demand-side management.

[22]Directive 2003/54/EC concerning common rules for the internal market in electricity, followed by the Directive 2009/72/EC and the 2019 Commission's broader package of initiatives "*Clean Energy for All Europeans Package*" that aimed to promote a better regulation and protection of vulnerable consumers.

[23]For the E.U., see for instance the Market Monitoring Report, a yearly publication of the Agency for Cooperation of Energy Regulators (ACER) that assesses the development of retail competition and ranks countries according to their performances. Good performers include Finland, Sweden, Great Britain, Norway and the Netherlands, whereas Greece, Bulgaria and Cyprus are considered as bad performers.

[24]Between 2011 and 2016, the European average market share of the three largest suppliers in the retail electricity household segment fell from 84.3% to 80.6% (ACER, 2017). Many national differences remain among European countries

[25]The average switching rate in the European retail electricity markets for household equals 6.5% in 2017, the highest switching rate being reported by Norway (18.8%) and the lowest by Bulgaria (0.002%); see CEER (2018).

[26]Recent research empirically assesses the properties of the different type of tariffs, showing that consumers responsiveness to dynamic pricing depends on awareness, perception and potential cognitive biases (Faruqui and Sergici, 2010; Ito, 2014).

completely removed price controls and other regulatory oversight and supervision.[27] Contravening the 2003/54/EC Directive concerning common rules for the internal market in electricity, only half of E.U. countries have completely removed regulated prices. Other countries have kept some form of price regulation, either for the entire retail market or for households and small commercial and industrial customers only.[28] Some countries have not even announced any roadmap for their removal, while others simply seem to take their time.

2.4. Keeping Regulated Prices

Once competition in retail is introduced, the main justification for maintaining price control for incumbent retailers is to protect customers from significant increases in energy prices, especially in a context of limited retail competition. Price control can also help preventing price volatility, which is considered as socially unacceptable.

Price control may also be fixed so as to leave scope for new retailers to enter and operate more efficiently than the incumbent. Competition depends, indeed, on regulated prices: when, for instance, regulation leads to below-cost prices or generates insufficient margin to cover the risk of activity, it erects inefficient entry barriers and hinders the development of competition. Price control must therefore be transitional and removed as soon as sufficient competition has developed. Price calibration, that is, the determination of the appropriate level and structure of the regulated tariffs, is thus critical to avoid distorting competition among retailers.[29]

Texas is a good illustration of such choice to implement price control on incumbent suppliers to foster retail competition. The regulated price was calibrated as follows: the regulator (Public Utility Commission) set a "*Price to Beat*" that the incumbents had to offer to their customers, within their respective distribution service areas. This "*Price to Beat*" was transitional (anticipated to last 5 years from 2002 to 2007). After 3 years or until 40% of residential and small business customers are served by alternative providers, incumbent companies could start to offer a rate lower than this "*Price to Beat*". This rate was designed to give customers of incumbent companies a discount (a 6% rate reduction at the start of competition) and allow alternative suppliers and new entrants the opportunity to offer low rates to gain market shares. As a result, more than 70 firms have entered the retail market. The number of offers has been multiplied and switching rates reached almost 40% (Defeuilley, 2009). Retail competition in Texas has clearly been fostered by transitory regulation of incumbents' tariffs.

France provides another example of a regulatory reform that tried to reconcile strong regulation and increased competition through the coexistence of regulated prices and market offers. Two decades after the first European Directive on the electricity market,

[27] Other forms of price-setting interventions are often targeted at vulnerable customers but they may also have an impact on market competition. Examples include: the "*Acquirente Unico*" and standard offer prices in Italy, the new "*Safeguard Tariff*" in the U.K., "*Safety Net Regulation*" in Belgium and "*Tariff Surveillance*" in the Netherlands.

[28] Countries applying end-user price regulation in electricity are Bulgaria, Cyprus, Denmark, France, Hungary, Lithuania, Malta, Poland, Portugal, Romania, Slovakia and Spain.

[29] Price regulation may take different forms, such as setting or approving prices, standardization of prices or combinations of these tools using different mechanisms including rate-of-return, price-cap or other caps rule or other discretionary regulation.

French wholesale and retail electricity markets remains highly concentrated and households switching rate is low.[30] This is the main justification for continued regulation of the incumbent retailer EDF. Although France has taken a significant step towards more retail competition in 2010 with the NOME Law, regulated prices are still not explicitly removed and various measures have been taken to limit their impact on competition.[31,32] Interestingly, The French Energy Regulatory Commission (Commission de Régulation de l'Energie, CRE), is now responsible for setting regulated prices, whereas tariffs were previously set exclusively by the government (after advisory opinion from the CRE). However, the government always has the final say, which has already resulted in price disputes.[33] Ultimately, an unprecedented social crisis (the so-called "*yellow vests movement*"), which began as a protest against rising fuel taxes and has grown into a wider outpouring over inequality, has forced the government to postpone an increase of regulated tariff planned long in advance by the CRE. This emergency decision dictated by the social context illustrates the political bias that can affect electricity tariff regulation. At the end of the day, the government has proven to be able to turn its back on the decision of the independent regulator. This illustrates how far political bias can distort regulated tariffs.

3. MODEL

PREFERENCES. There is a continuum of customers of mass one. Customers have quasi-linear preferences defined over the quantity q of the good, say electricity, they buy at price p.[34] These preferences write as $\theta v(q) - p$, where θ is a preference parameter which is privately known by the customer. This parameter reflects the heterogeneity in the valuation of the good across customers. Although low levels of consumption may be correlated with income, and we might favor this interpretation in the sequel, this correlation might not be perfect. Some low users might have high incomes under some circumstances.[35]

[30] By the end of 2018, the 38 alternative suppliers accounted for just 21% of the household retail market (as a percentage of number of sites) and 34% as a percentage of annualized sales (in TWh) (CRE, 2018).

[31] NOME stands for New Organization of Electricity Market.

[32] First, to avoid risk and loss aversion biases, any customer choosing to switch away from the regulated tariffs can switch back at any time free of charge. Second, price calibration explicitly aims to give new entrants room to develop and, at the same time, support investment in peak electricity generation power plants and renewable energy sources (eventually through vertical integration by new retailers). The NOME Law stipulates that, until 2025, alternative suppliers will have access to part of the incumbent's historical nuclear production capacity in base-load (up to a maximum of 100TWh/year, equivalent to 25% of production capacity) at a regulated price called ARENH (Regulated Access to Historical Nuclear Electricity). The objective is to partly neutralize the incumbent's historical advantage associated to nuclear power plants and vertical integration. Third, the NOME Law has put a progressive end (up to the 31st of December 2015) to the regulated tariffs for industrial customers and revised the structure of regulated prices for household customers.

[33] Since 2012, the government has systematically set retail prices to a different level than the one recommended by the CRE, leading to tariff deficits, court rulings and consumer confusion. In 2013, while the CRE recommended an increase by 11.3% of the regulated tariff for households, the government decided an increase by only 5%. In 2014, the CRE recommended a lower increase (1.6%) than what the government finally set (2.5%), while, in 2015, the CRE recommended a decrease by 0.9% while the government set an increase by 2.5% (ACER, 2017).

[34] The main motivation for our analysis is of course the electricity sector though we keep our presentation more general to broaden the scope of our analysis to other sectors like gas, telecommunication or transportation where regulated prices may remain despite the competitive context.

[35] To illustrate that such correlation may be less than perfect, Borenstein and Davis (2012), in their analysis of the gas sector in the U.S., argue that income and consumption are only weakly correlated. Poor households with multiple children leaving in less energy efficient houses consume a disproportionate

The function $v(q)$ is strictly increasing and strictly concave in the quantity q ($v' > 0$ and $v'' < 0$) over a domain $\mathcal{Q} = [0, Q_{max}]$ where Q_{max} is large enough to ensure interior solutions under all circumstances below. The conditions $v(0) = 0$ and $v'(0) = +\infty$ ensure that all types of customers are always served by the firm.[36] The preference parameter θ is drawn from a common knowledge (and atomless) distribution function F on the support $\Theta = [\underline{\theta}, \overline{\theta}]$ (with $\underline{\theta} \geq 0$). The corresponding positive density is denoted by f with $f = F'$.

TECHNOLOGY. The public utility's marginal cost of production $c \geq 0$ is constant and common knowledge until Section 6 below. Production requires also to incur a fixed cost denoted by $K \geq 0$ whose value remains common knowledge throughout.[37]

EFFICIENCY. The efficient level of consumption $q^{\varnothing}(\theta, c)$ maximizes the overall surplus $\theta v(q) - cq - K$ and is thus obtained as

$$\theta v'(q^{\varnothing}(\theta, c)) = c,$$

provided that $\theta v(q^{\varnothing}(\theta, c)) - cq^{\varnothing}(\theta, c) - K \geq 0$, an assumption that is made throughout the whole analysis. The superscript '$\varnothing$' stands for the fact that there are no redistributive concerns in this benchmark, only efficiency considerations matter.

NONLINEAR PRICING AND INCENTIVE-FEASIBLE ALLOCATIONS. Following an approach that is now well know from the existing nonlinear pricing literature, the public utility offers a nonlinear price $P(q)$ to screen customers according to their consumption.[38]

To ensure the existence of a maximizer to the customers' problem, we assume that a feasible nonlinear price $P(\cdot)$ is upper semi-continuous. Denoting by $q(\theta)$ a selection in the best response correspondence for the customer with type θ, we accordingly define

$$q(\theta) \in \arg\max_{q \in \mathcal{Q}} \theta v(q) - P(q).$$

The equilibrium payoff $U(\theta)$ of a customer with type θ when facing a nonlinear price $P(\cdot)$ is thus defined as

$$U(\theta) = \max_{q \in \mathcal{Q}} \theta v(q) - P(q).$$

The next result is the standard characterization of incentive compatible allocations $(U(\theta), q(\theta))_{\theta \in \Theta}$, i.e., allocations that can be implemented by means of a nonlinear price $P(\cdot)$.[39]

LEMMA 1. *An allocation $(U(\theta), q(\theta))_{\theta \in \Theta}$ is incentive compatible if and only if:*

- *$U(\theta)$ is absolutely continuous and thus almost everywhere differentiable with the following condition at each point of differentiability*

$$\dot{U}(\theta) = v(q(\theta)); \tag{3.1}$$

amount of their wealth in energy.

[36] Shutting off some low-valuation customers is thus not an issue in this context. This reflects the fact that, in practice, customers are often protected against a company shutting off or terminating their service, even if they cannot afford to pay their bills.

[37] Production costs encompass the variable cost of producing electricity (buying energy sources and managing power plants if the firm is vertically integrated) or of buying electricity on the wholesale market (if the firm has no generating plants).

[38] See Goldman, Leland and Sibley (1984), Maskin and Riley (1984) and Wilson (1993).

[39] See Rochet (1987) and Laffont and Martimort (2002, Chapter 3).

- *$U(\theta)$ is convex, or equivalently*

$$q(\theta) \text{ is non-decreasing.} \tag{3.2}$$

A customer with type θ accepts the nonlinear price $P(q)$ whenever $U(\theta) \geq 0$. A direct consequence of (3.1) is that $U(\theta)$ is non-decreasing. Therefore, participation of all types is ensured whenever the customer who has the lowest valuation for the good finds it attractive to consume

$$U(\underline{\theta}) \geq 0. \tag{3.3}$$

An allocation that satisfies the incentive compatibility conditions (3.1) and (3.2) as well as the participation constraint (3.3) is said to be *incentive-feasible.*

Observe that a rent profile $U(\theta)$ that satisfies the incentive compatibility condition (3.1) can be written as

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v(q(\tilde{\theta}))d\tilde{\theta}. \tag{3.4}$$

REMARK. In our analysis, we shall rely on a dual approach that focuses on the allocation $(U(\theta), q(\theta))_{\theta\in\Theta}$ as the object of prime interest instead of the nonlinear price $P(q)$ that implements such allocation. A simple duality argument from convex analysis nevertheless allows to recover that nonlinear price from the payoff profile as[40]

$$P(q) = \max_{\theta\in\Theta} \theta v(q) - U(\theta).$$

At any $q(\theta)$ that is chosen by a consumer with type θ, the price $P(q(\theta))$ can finally be expressed as

$$P(q(\theta)) = \theta v(q(\theta)) - \int_{\underline{\theta}}^{\theta} v(q(\tilde{\theta}))d\tilde{\theta} - U(\underline{\theta}).$$

□

GOVERNMENT'S OBJECTIVE. The government exerts a complete control of the public utility. This assumption amounts to saying that the nonlinear price used by the incumbent is actually chosen by the government itself.[41] Supposing full control is, of course, a simplifying assumption that nevertheless captures the idea that the long-term relationship between the incumbent (often a vertically integrated monopoly which has been publicly-owned for decades) and the public sphere has reduced the informational asymmetries that are the tenets of the modern theory of Regulator Economics.[42] Alternatively, our model is consistent with the fact that the public utility may be publicly-owned and, in that case, ownership gives access to information.[43]

[40] See Rochet (1985) and Basov (2005).

[41] Typically, when an independent regulatory agency is in charge of regulating price, it might have no or limited information to target low-income households compared to what is available to direct tax authorities (governments or public administrations). Our implicit assumption here is that the regulator and the tax authorities are a single entity (the government) which has some redistribute objectives.

[42] See Baron and Myerson (1982), Laffont and Tirole (1993), Laffont (1994), Armstrong and Sappington (2007).

[43] See Arrow (1975).

The government has some redistributive concerns. This objective should nevertheless bring the allocation induced by the price structure on the Pareto-frontier of the set of incentive-feasible allocations. Following Holmström and Myerson (1983) and Ledyard and Palfrey (1999), who characterize such *interim efficient* allocations in related contexts, we assume that the government maximizes a weighted sum of the consumers' gains

$$\int_{\underline{\theta}}^{\overline{\theta}} \alpha(\theta)U(\theta)f(\theta)d\theta, \tag{3.5}$$

where $\alpha(\theta)$ is the positive social weight given to the customer with type θ.

For future reference, we define the *downward cumulated social weight* as

$$\Lambda(\theta) = \int_{\theta}^{\overline{\theta}} \alpha(\tilde{\theta})f(\tilde{\theta})d\tilde{\theta},$$

and we normalize the distribution of weights so that $\Lambda(\underline{\theta}) = 1$.

The government's objective is consistent with various redistributive scenarios. When the distribution of social weights gives few importance to low-valuation types, i.e., $\Lambda(\theta) < 1 - F(\theta)$ for θ small enough, redistribution goes towards the high-valuation customers. When instead $\Lambda(\theta) > 1 - F(\theta)$ for θ small enough, redistribution goes towards the low-valuation customers.

Taking a normative perspective, equity or fairness considerations might then justify the choice of a given distribution for those social weights. Yet, the above objective is also consistent with a couple of political economy interpretations.[44] First, the government may want to favor specific interest groups for electoral purposes. As far as electricity market reforms and electricity consumption are concerned, there is indeed a wide diversity of interest groups and stakeholders that may oppose the level and structure of electricity prices (industrial vs. agriculture or households customers, urban vs. rural customers, pro-environmental vs. pro-industry parties, market-oriented vs public intervention, etc.). The distribution of social weights may thus reflect the relative socio-economic importance of these various groups and their political strengths. Second, and in the spirit of Posner (1971)'s so-called "*Taxation by Regulation,*" cross-subsidization between groups of users is a form of taxation with the regulator redistributing its proceeds in response to the political pressures of those groups.

Implicit in our above modeling is also the fact that, by considering a single entity in charge of price regulation (the so-called *government*), we do not distinguish between elected politicians and regulators in charge of controlling the firm's price structure. Most often, regulators are mostly concerned with efficiency and much less with the redistributive consequences of pricing which is by nature an important concern for politicians.[45]

[44]See Erdogdu (2014) for a political economy analysis of electricity market liberalization.

[45]Referring to the telecommunication sector in the U.K., Burns, Crawford and Dilnot (1995) report that the Director General of Telecommunications (DGT) clearly sets out his role in his first annual report back in 1985: "*I should make it clear that I do not think that it would be appropriate for me to seek to impose a balance of prices in a way that was motivated primarily by a desire to achieve some particular redistribution of income amongst members of the community, nor do I think my powers would permit me to do this [....] I do not believe, for example, that I could properly put forward a proposal for a rule that all people on low incomes should be given telephones free of rentals; such a proposal would involve arbitrary judgments about matters of income redistribution and my making it would involve the usurping*

BREAK-EVEN CONSTRAINT. Of course, the public utility has to operate under budget constraint.[46] The revenues raised from the service must cover its cost

$$\int_{\underline{\theta}}^{\overline{\theta}} \Big(P(q(\theta)) - cq(\theta)\Big) f(\theta)d\theta \geq K.$$

For future references, it is useful to rewrite this budget constraint in terms of the profile of consumer's payoffs and consumption levels $(U(\theta), q(\theta))_{\theta\in\Theta}$ as follows

$$\int_{\underline{\theta}}^{\overline{\theta}} \Big(\theta v(q(\theta)) - cq(\theta) - U(\theta)\Big) f(\theta)d\theta \geq K. \tag{3.6}$$

4. PRICING STRUCTURE WITH A REDISTRIBUTIVE OBJECTIVE

The nonlinear price chosen by the government, or equivalently the allocation $(U(\theta), q(\theta))_{\theta\in\Theta}$ that it induces, must maximize (3.5) subject to the incentive compatibility constraint (3.1), the participation constraint (3.3) and the break-even condition (3.6).[47] In a first step, we also neglect the participation constraint (3.3) and check *ex post* that it is verified.

The next Proposition characterizes the allocation induced by a government with some redistributive concerns. Distortions away from the first-best level of consumption are there completely characterized by the new term

$$\gamma(\theta) \equiv \frac{1 - F(\theta) - \Lambda(\theta)}{f(\theta)},$$

whose sign depends on the magnitude of the bias towards low-valuation customers.

For future reference, we impose the following monotonicity condition familiar from screening models

$$\theta - \gamma(\theta) \text{ non-decreasing.} \tag{4.1}$$

Less familiar is the following condition that also ensures that the solutions to the optimization problems found below are always regular (i.e., those solutions do not entail bunching)

$$\min_q -\frac{qv''(q)}{v'(q)} \geq \frac{\gamma(\theta)}{\theta} \quad \forall \theta \in \Theta. \tag{4.2}$$

This condition simply says that social weights should be small compared to inverse price elasticity for electricity, a rather weak requirement given the small value of such elasticity for household customers, especially in the short run.

---

of the proper role of government." Yet, Burns, Crawford and Dilnot (1995) also recognize that "*[...] in practice, all the regulators, including the DGT, have taken into account the social impact of relative price movements, a point recently acknowledged by the electricity regulator, Stephen Littlechild, who said at a recent conference that regulators do have a social concern, which influences regulatory policy.*"

[46] See Boîteux (1956).

[47] For the time being, we neglect the monotonicity condition (3.2). We will impose later on a condition on the joint distribution of types and weights that ensures that (3.2) is indeed satisfied.

PROPOSITION 1. *The optimal allocation induced by the government has the following properties:*[48]

- *A consumption profile* $q_r(\theta, c)$ *such that*

$$(\theta - \gamma(\theta))\, v'(q_r(\theta, c)) = c; \tag{4.3}$$

- $q_r(\theta, c)$ *is non-decreasing in* θ *when (4.1) holds.*[49]
- *The participation constraint* (3.3) *holds provided that*

$$\int_{\underline{\theta}}^{\overline{\theta}} \left(\left(\theta - \frac{1 - F(\theta)}{f(\theta)} \right) v(q_r(\theta, c)) - c q_r(\theta, c) \right) f(\theta) d\theta \geq K.^{50} \tag{4.4}$$

Much as in the spirit of the nonlinear pricing literature, the consumption level $q_r(\theta, c)$ now maximizes a "*virtual surplus*" $(\theta - \gamma(\theta))v(q) - cq$, where the preference parameter θ is replaced by a virtual type $\theta - \gamma(\theta)$.[51] This virtual type differs from the true preference parameter to reflect how incentive compatibility considerations interact with the government's biased objective. This interaction is novel since, in the nonlinear pricing literature, the monopoly's objective is simply to extract consumer surplus.

Of course, when the government has no redistributive concerns, every customer types receive the same social weight, i.e., $\alpha(\theta) = 1$. It follows that $\Lambda(\theta) = 1 - F(\theta)$ and the first-best consumption is implemented, i.e., $q_r(\theta, c) = q^{\varnothing}(\theta, c)$ for all (θ, c).

The interaction between incentive compatibility and the government's redistributive concerns can be best understood by signing $\gamma(\theta)$. Suppose first $\alpha(\theta)$ is small for types θ close to the highest valuation $\overline{\theta}$. Thus $\Lambda(\theta) < 1 - F(\theta)$ and $\gamma(\theta)$ is positive in that neighborhood. The government is thus biased towards low-valuation customers and there is a downward distortion of production below the first-best level. Intuitively, the government

[48] We make the dependence on the incumbent's cost c explicit for future reference.

[49] When Condition (4.1) does not hold, the consumption profile $q_r(\cdot)$ that solves the relaxed problem where the monotonicity constraint (3.2) is omitted is no longer a solution to the maximization problem $(\mathcal{P})$. Ironing procedures, following Myerson (1981) or Guesnerie and Laffont (1984), are necessary to characterize the solution and, in particular, bunching areas. Taking into account these technicalities would complicate the analysis without changing neither the nature of our analysis nor our results.

[50] Suppose that Condition (4.4) (the participation constraint of the lowest valuation consumer) does not hold for $q_r(\cdot)$. This means that the fixed cost of production K is so large that more revenues have to be raised to cover that outlay. This creates a tension with the participation constraint of the lowest valuation customer. Let us denote by λ the Lagrange multiplier of this constraint. Proceeding as in the proof of Proposition 1 to show that (3.3) can be rewritten as $\int_{\underline{\theta}}^{\overline{\theta}} \left(\left(\theta - \frac{1-F(\theta)}{f(\theta)} \right) v(q(\theta)) - cq(\theta) \right) f(\theta) d\theta \geq K$, it is then straightforward to derive the optimal consumption as

$$\left(\theta - \frac{1 - F(\theta) - \frac{\Lambda(\theta)}{1+\lambda}}{f(\theta)} \right) v'(q_r(\theta, c)) = c.$$

Provided now that $\theta - \left(1 - F(\theta) - \frac{\Lambda(\theta)}{1+\lambda}\right)/f(\theta)$ is non-decreasing, $q_r(\theta, c)$ so defined is non-decreasing and the monotonicity requirement (3.2) again holds. The role of the multiplier λ is to reduce the impact of the government's biased objective in the design of the optimal policy. Modulo this minor change, results are qualitatively similar. The same holds true throughout the paper.

[51] Myerson (1981) coined this expression.

would like that most of the burden of financing the fixed cost of production be borne by high-valuation customers. As a result, those types are willing to mimic the low-valuation ones, and this strategy is made less attractive when consumption is distorted downwards.

Suppose instead that $\alpha(\theta)$ is large for θ close to $\overline{\theta}$. The government is now biased towards high-valuation customers. $\gamma(\theta)$ is then negative and there is over-production in comparison with the first-best level. The government would like that most of the burden of paying for the firm's fixed cost be borne by low-valuation customers. Those types are thus willing to mimic high-valuation ones to minimize their contribution. To make such a strategy less attractive, the consumption level is distorted upwards.

More generally, incentive compatibility shapes the structure of the distortions implemented to favor some types of customers relative to others. Favoring consumers who have higher (resp. lower) valuations requires upwards (resp. downwards) distortion.

Moreover, we observe that

$$q_r(\theta, c) = q^{\varnothing}(\theta, c) \text{ for } \theta \in \{\underline{\theta}, \overline{\theta}\}.$$

In other words, there are no distortions in consumption levels at both extreme points of the set of types. Distortions are concentrated in the middle of the interval. The next two figures represent consumption distortions under different redistributive concerns, either towards low-valuation customers or towards high-valuation ones.

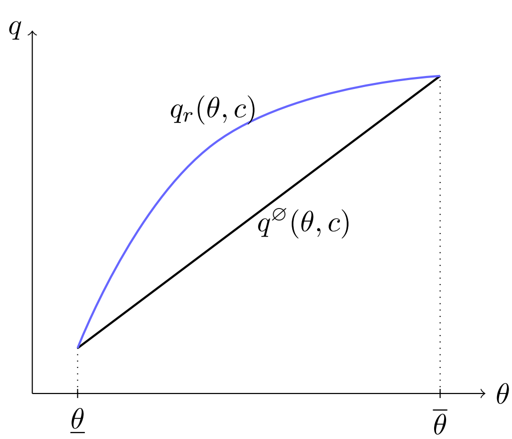

Figure 1: Over-production for $\gamma(\theta) > 0$.

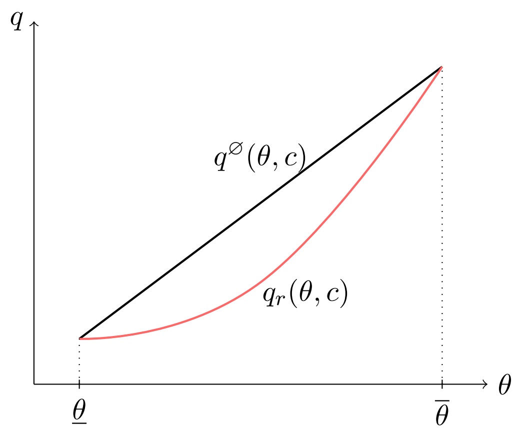

Figure 2: Under-production for $\gamma(\theta) < 0$.

Below-Cost Pricing. Denote by $p(q) \equiv P'(q)$ the marginal price for q units of consumption. From the consumers' optimality condition, we have

$$p(q_r(\theta, c)) = \theta v'(q_r(\theta, c)).$$

Simplifying using the characterization of the optimal consumption profile (4.3), we obtain

$$p(q_r(\theta, c)) = c + \gamma(\theta) v'(q_r(\theta, c)).$$

From this, it immediately follows that

$$p(q_r(\theta, c)) \gtreqless c \Leftrightarrow \gamma(\theta) \gtreqless 0 \Leftrightarrow q_r(\theta, c) \gtreqless q^*(\theta, c).$$

In words, below-cost pricing arises for low-valuation customers in contexts where the government has a redistributive bias towards high-valuation customers. We will keep stock of this insight for the rest of our analysis.

5. The Costs and Benefits of Entry

We now consider the possibility for consumers to also buy the good from a competitive fringe on a market. The price on this market is a random variable, denoted by c_e, that corresponds to the marginal cost of the fringe. Its common knowledge probability distribution is denoted by G, with the corresponding density $g = G'$ and with support $[0, \bar{c}_e]$.

Surplus and consumption levels when purchasing from the fringe write respectively as

$$U_e(\theta, c_e) = \max_{q \in \mathcal{Q}} \theta v(q) - c_e q \quad \text{ and } \quad q_e(\theta, c_e) = \arg\max_{q \in \mathcal{Q}} \theta v(q) - c_e q.$$

Of course, $q_e(\theta, c_e)$ is defined through the following first-order condition

$$\theta v'(q_e(\theta, c_e)) = c_e.$$

Similarly, and much as in the analysis of the previous sections, the regulated tariff $P_i(q, c)$ of the incumbent induces surplus and consumption profiles defined respectively by

$$U_i(\theta, c) = \max_{q \in \mathcal{Q}} \theta v(q) - P_i(q, c) \quad \text{ and } \quad q_i(\theta, c) \in \arg\max_{q \in \mathcal{Q}} \theta v(q) - P_i(q, c).$$

When entry is possible, the contracting game unfolds as follows. First, the government sets a regulated tariff $P_i(q, c)$, knowing neither the demand parameter θ, nor the cost of the fringe c_e. Yet, the government knows the incumbent's cost structure at this stage. Second, the fringe's cost parameter c_e is realized. Third, consumers decide whether to buy from the incumbent at the regulated tariff or from the fringe at price c_e.[52]

Incentive-Feasible Market Allocations. We now define a *market allocation* as the payoff and production profile induced by a regulated tariff $P_i(q, c)$ together with a "*cut-off rule*" that determines when consumers should switch to purchase from the fringe. From the *Revelation Principle,* there is no loss of generality in viewing this cut-off rule $b(\hat{\theta}, c)$ as dependent of the customer's announcement on his type and the (common knowledge) cost of the incumbent c. The expected surplus $\mathcal{U}(\theta, c)$ of a consumer with type θ is thus defined as

$$\mathcal{U}(\theta, c) = \max_{\hat{\theta} \in \Theta} \left(1 - G(b(\hat{\theta}, c))\right) U_i(\theta, c) + \int_{\underline{c}}^{b(\hat{\theta}, c)} U_e(\theta, c_e) g(c_e) dc_e.$$

Mimicking our characterization of incentive compatibility found under the incumbent's monopoly, we obtain

Lemma 2. *An incentive compatible market allocation $(\mathcal{U}(\theta, c), q_i(\theta, c), b(\theta, c))_{\theta \in \Theta}$ has the following properties:*

[52] Hortacsu, Madanizadeh and Puller (2017) argue that consumers exhibit significant inertia as illustrated by the low switching activity. Our model is consistent with this finding if we interpret the entrant's cost as its true production cost plus any switching costs that consumers may have.

- *$\mathcal{U}(\theta, c)$ is absolutely continuous and thus almost everywhere differentiable with the following derivative at each point of differentiability, where $b(\theta, c)$ and $q_i(\theta, c)$ are both continuous*

$$\frac{\partial \mathcal{U}}{\partial \theta}(\theta, c) = (1 - G(b(\theta, c)))\, v(q_i(\theta, c)) + \int_{\underline{c}}^{b(\theta,c)} v\left(q_e(\theta, c_e)\right) g(c_e) dc_e. \tag{5.1}$$

- *The following monotonicity condition together with (5.1) are sufficient for incentive compatibility*

$$\Psi(\theta, \hat{\theta}) \geq 0 \quad \forall (\theta, \hat{\theta}) \in \Theta^2, \tag{5.2}$$ [53]

where

$$\Psi(\theta, \hat{\theta}) = \left(1 - G(b(\hat{\theta}, c))\right) v(q_i(\hat{\theta}, c)) + \int_0^{b(\hat{\theta},c)} v\left(q_e(\theta, c_e)\right) g(c_e) dc_e.$$

Participation constraints require that the market allocation gives to all types of customer a non-negative payoff. As above, a sufficient condition is that the customer with the lowest valuation chooses to accept such mechanism

$$\mathcal{U}(\underline{\theta}, c) \geq 0. \tag{5.3}$$

The set of incentive-feasible market allocations is thus fully determined by (5.1), (5.2) and (5.3).

BREAK-EVEN CONDITION. This condition writes now as

$$\int_{\underline{\theta}}^{\overline{\theta}} (1 - G(b(\theta, c)))\left(P_i(q_i(\theta, c)) - cq_i(\theta, c)\right) f(\theta) d\theta \geq K. \tag{5.4}$$

That the consumer may switch to the competitive fringe *a priori* hardens the firm's break-even condition since the incumbent raises less revenues from the service with entry. Yet, and much as in the spirit of the I.O. literature on vertical contracts (Aghion and Bolton, 1987; Choné and Linnemeur, 2015; Martimort, Pouyet and Stole, 2019) the incumbent may also structure the tariff so as to make switching costly for the customers and thereby extract some of the gains from entry. Those gains are then *in fine* used to cover the incumbent's fixed cost.

OPTIMAL MARKET ALLOCATION. Assuming, as we did in the previous section, that (5.3) and the monotonicity condition (5.2) both hold, the optimal market allocation should thus maximize

$$\int_{\underline{\theta}}^{\overline{\theta}} \alpha(\theta) \mathcal{U}(\theta, c) f(\theta) d\theta$$

[53] Suppose that G puts all mass on very high values of c_e. Then, switching never occurs and the model boils down to the monopoly setting we have been studying so far. It amounts to set $G \equiv 0$ in the above formulation. Lemma 2 then boils down to Lemma 1. The monotonicity condition (5.2) holds when $q_i(\theta, c)$ is non-decreasing in θ as expected. When the distribution G puts positive mass on low values of c_e switching can be used as a screening device. The screening problem has thus multiple instruments and the characterization of incentive compatible allocations is notoriously more difficult under those circumstances. See Laffont and Martimort (2002, Chapter 3).

subject to the break-even constraint (5.4) and the incentive compatibility condition (5.1).

Proceeding as in the Proof of Proposition 1, it is straightforward to show that the government's problem actually consists in finding the allocation $(q_i(\theta,c), b(\theta,c))$ that maximizes

$$(5.5)\quad -K + \int_{\underline{\theta}}^{\overline{\theta}} \Bigg(\left(1 - G(b(\theta,c))\right)\left(\left(\theta - \gamma(\theta)\right) v(q_i(\theta,c)) - cq_i(\theta,c)\right) \\ + \int_0^{b(\theta,c)} \Big(\left(\theta - \gamma(\theta)\right) v(q_e(\theta,c_e)) - c_e q_e(\theta,c_e)\Big) g(c_e) dc_e \Bigg) f(\theta) d\theta.$$

This expression is nothing else than an expected virtual surplus, computed with the relevant social weights, but also taking into account that the government may find it optimal to induce consumers to switch to the fringe for efficiency reasons. This expression also makes it clear that the government is not hostile to entrants *per se* and if there is any inefficient entry, it is not due to the government willing to erect inefficient barriers to entry. Instead, by playing on the regulated tariffs, the government can induce an optimal probability of switching and grasp the virtual surplus that entrants may generate.

The solution to this maximization problem is summarized in the next proposition.

PROPOSITION 2. *The optimal market allocation has the following properties:*

- *The optimal cut-off rule $b^*(\theta,c)$ satisfies*

$$(5.6)\quad \left(\theta - \gamma(\theta)\right) v(q_r(\theta,c)) - cq_r(\theta,c) \\ = \left(\theta - \gamma(\theta)\right) v(q_e(\theta, b^*(\theta,c))) - b^*(\theta,c) q_e(\theta, b^*(\theta,c)).$$

- *For $c_e \geq b^*(\theta,c)$, a consumer with type θ purchases from the incumbent a quantity $q_r(\theta,c)$ at the regulated tariff. For $c_e \leq b^*(\theta,c)$, a consumer with type θ purchases instead from the competitive fringe a quantity $q_e(\theta,c_e)$.*

- *Condition (4.1) is always sufficient to ensure that (5.2) holds when $\overline{\theta} - \underline{\theta}$ is not too large and (4.2) holds.*

When efficiency is the sole goal of a government giving an equal social weight to all customers, i.e., $\gamma(\theta) = 0$ for all θ, the optimal cut-off rule is efficient: $b^*(\theta,c) = c$. The fringe then supplies the market as soon as it is more efficient than the incumbent.

When the government has some redistributive concerns, the switching rule is determined so as to equalize consumer's virtual surpluses with the incumbent and with the fringe. The benefit of switching to the fringe is that it produces at a lower cost than the incumbent. The cost is that production on the market by the competitive fringe is only driven by efficiency considerations.

The allocation described in Proposition 2 implies a strong form of dichotomy between, on the one hand, how regulated prices are fixed and, on the other hand, how demand is split between the incumbent (operating under regulated tariffs) and competitive entrants. More precisely, the quantity purchased from the incumbent remains identical to that found in the absence of entry, and it is independent of the probability that customers switch to the market. The quantity bought from the incumbent always maximizes the virtual surplus $(\theta-\gamma(\theta))v(q)-cq$. In other words, only the absolute level of regulated prices impacts market shares, not their margins. When the government wants to reduce entry, it decreases uniformly the level of regulated prices.

Another important result can be derived from Proposition 2.

PROPOSITION 3. *Whatever the nature of the bias in the government's objective the allocation of market shares is biased against entry, i.e., whatever the sign of* $\gamma(\theta)$

$$c \geq b^*(\theta, c) \quad \forall(\theta, c). \tag{5.7}$$

With redistributive concerns towards low-valuation (resp. high-valuation) customers, consumption from the market is greater (resp. lower) than from the incumbent, i.e., if $\gamma(\theta) \geq 0$ *(resp.* ≤ 0) $\forall\theta \in \Theta$ *then*

$$q_e(\theta, b^*(\theta, c)) \geq q_r(\theta, c) \ \textit{(resp. } \leq 0\textit{)} \quad \forall(\theta, c). \tag{5.8}$$

Differently put, there is less entry than what efficiency requires and this result does not depend on the nature of the redistributive concerns of the government. To understand such a systematic bias towards the incumbent, let us rewrite (5.6) as follows[54]

$$\left[(\theta-\gamma(\theta))\,v(q)-cq\right]_{q_e(\theta,b^*(\theta,c))}^{q_r(\theta,c)} = (c-b^*(\theta,c))\,q_e(\theta,b^*(\theta,c)). \tag{5.9}$$

The left-hand side represents the loss in virtual surplus when switching from the incumbent, which offers a quantity $q_r(\theta,c)$ that precisely maximizes this virtual surplus, to the fringe, which produces a quantity $q_e(\theta,c)$ that clearly does not (except of course when there is no redistributive concerns across consumers types, i.e., $\alpha(\theta)=1$ for all θ). It is therefore positive. The right-hand side represents how much cost is saved by having the production $q_e(\theta,c)$ being supplied by the competitive fringe at a cost $b^*(\theta,c)$ rather than by the incumbent at cost c. The government therefore distorts the incumbent's tariff in a way that restricts entry, i.e., $c \geq b^*(\theta,c)$. Intuitively, if entrants were just as efficient as the incumbent, the government would always prefer the production profile implemented by the incumbent because it fulfills its biased objective, whereas that of the entrants only promotes efficiency. Competition has to come with non-marginal efficiency gains to make it attractive to give up the redistributive role that the incumbent's tariffs play.

COMPARATIVE STATICS. Even though the regulated tariffs are uniformly shifted downward to induce entry, the probability of entry is not uniform across customers. Since a customer's payoffs on the market (namely $U_e(\theta,c_e)$) and with the incumbent (namely $U_i(\theta,c)$) are both type-dependent, the probabilities that different types switch to the entrant are also so. Indeed, we have:

$$\frac{\partial b^*}{\partial\theta}(\theta,c) = \frac{1-\dot{\gamma}(\theta)}{q_e(\theta,b^*(\theta,c)) + \frac{\gamma(\theta)v'(q_e(\theta,b^*(\theta,c))}{\theta v''(q_e(\theta,b^*(\theta,c))}}\left(v(q_e(\theta,b^*(\theta,c)) - v(q_r(\theta,c)))\right). \tag{5.10}$$

Since the denominator is positive when (4.2) and (5.8) hold, and $\theta-\gamma(\theta)$ is non-decreasing by assumption, (5.10) leads to

$$\frac{\partial b^*}{\partial\theta}(\theta,c) \geq 0 \text{ (resp. } \leq 0) \text{ if } \gamma(\theta) \geq 0 \text{ (resp. } \leq) \ \forall\theta \in \Theta. \tag{5.11}$$

[54]We use the compact notation $[f(x)]_{x_1}^{x_2} = f(x_2) - f(x_1)$.

Last, but not least, observe that the cut-off rule $b^*(\theta, c)$ depends explicitly on the cost parameter c. More precisely, we have

$$\frac{\partial b^*}{\partial c}(\theta, c) = \frac{q_r(\theta, c)}{q_e(\theta, b^*(\theta, c)) + \frac{\gamma(\theta)v'(q_e(\theta, b^*(\theta, c)))}{\theta v''(q_e(\theta, b^*(\theta, c))}}, \tag{5.12}$$

and, therefore,

$$\frac{\partial b^*}{\partial c}(\theta, c) > 0 \tag{5.13}$$

when (4.2) holds. In other words, as the incumbent's cost increases, the switching rule recommends to shift to the fringe more often. This effect opens the possibility that, in a context of asymmetric information on the value of the incumbent's cost, the government may want to strategically manipulate information on this cost so as to make entry less likely than what would be optimal had such information been common knowledge. This scenario will be studied in the next section.

PARTICIPATION CONSTRAINT. The possibility that entry may arise affects the participation constraint of the consumer with the lowest valuation, which we remind was assumed to be satisfied. Simple manipulations show that $\mathcal{U}(\underline{\theta}, c) \geq 0$ is equivalent to

$$\begin{aligned} &\int_{\underline{\theta}}^{\overline{\theta}} \left((1 - G(b^*(\theta, c))) \left(\left(\theta - \frac{1 - F(\theta)}{f(\theta)} \right) v(q_r(\theta, c)) - cq_r(\theta, c) \right) \right. \\ &+ \left. \int_0^{b^*(\theta, c)} \left(\left(\theta - \frac{1 - F(\theta)}{f(\theta)} \right) v(q_e(\theta, b^*(\theta, c))) - c_e q_e(\theta, b^*(\theta, c)) \right) g(c_e) dc_e \right) f(\theta) d\theta \geq K. \end{aligned} \tag{5.14}$$

It is interesting to compare the consumer's participation constraints with entry, namely (5.14), and without entry, namely (4.4). Intuitively, (5.14) is more easily satisfied than (4.4). This highlights that the government benefits from the possibility of entry by competitive firms that are more efficient than the incumbent. Consumers indeed have to accept or reject the incumbent's regulated tariff before knowing the fringe's efficiency parameter. Henceforth, the government can play on the regulated tariff to indirectly extract the expected surplus that consumers obtain from entry. This surplus helps financing the fixed cost of production of the incumbent.

6. How to Regulate Regulated Prices?

We now assume that the incumbent's cost parameter c is only known from the incumbent and the government. This cost parameter is drawn from an atomless distribution function $H(\cdot)$ with positive density $h(\cdot) = H'(\cdot)$ on the same support $\mathcal{C} = [0, \overline{c}]$ as the entrant's own cost. Private information on c allows the government to manipulate the tariffs structure to strategically impact entry.

6.1. Benchmarks

Suppose, as a first benchmark, that the decision to allow entry is chosen so as to maximize an *ex ante* efficiency criterion. In the absence of any information on the incumbent's cost, entry is thus favored if and only if[55]

$$\mathbb{E}_c(c) \geq c_e. \tag{6.1}$$

This cut-off rule induces either excessive or insufficient entry depending on whether c is below or above the mean.

A possible response to the inadequacy of this rule would be to delegate the choice of this cut-off to the government. The benefit of granting such discretion to the government is that it shares cost information with the firm. The cost of such delegation is that the government is not only concerned with efficiency but also with redistribution. Indeed, if the government is given full discretion, it implements the market allocation described in Proposition 2, and that allocation is always biased against entry even though it is cost-dependent. The flip side of such discretion is therefore that the government can manipulate costs so as to limit entry and favor its own redistributive concerns.

6.2. Strategic Manipulations of the Incumbent's Cost

We now study what sorts of institutional constraints can be imposed on the government to prevent such strategic manipulations and limit its discretion. To find the optimal mix between *rules and discretion*, we adapt the approach of the delegation literature (Holmström, 1984; Melumad and Shibano, 1991; Martimort and Semenov, 2006; Alonso and Matouscheck, 2008), especially when it addresses the constraints imposed on public bodies (Epstein and O'Halloran, 1994; Hiriart and Martimort, 2012; Iossa and Martimort, 2016) and characterize *delegation mechanisms*.

Such mechanisms take a very simple form in our context. The most natural instruments consist in setting a floor or a cap on regulated prices and imposing that competitors always get access to a minimal or maximal market share. The government, informed on the incumbent's cost parameter, is thus bound to choose consumption levels under regulated prices and market shares within a predetermined set, which we call the *delegation set*. From the Revelation Principle,[56] this set of options can be fully described by means of a direct revelation mechanism of the form $(q_i(\theta, \hat{c}), b(\theta, \hat{c}))_{\hat{c} \in \mathcal{C}}$, that determines both the consumption under the regulated tariff and the switching rule as a function of the report $\hat{c}$ on the incumbent's cost parameter.

Taking stock of the expression found in (5.5), the government's payoff when announcing a cost $\hat{c}$ and implementing the market allocation $(q_i(\theta, \hat{c}), b(\theta, \hat{c}))$ may be written as follows

$$\Phi(c, \hat{c}) = -K + \int_{\underline{\theta}}^{\overline{\theta}} \Bigg((1 - G(b(\theta, \hat{c})))\left((\theta - \gamma(\theta))v(q_i(\theta, \hat{c})) - cq_i(\theta, \hat{c})\right)$$
$$+ \int_{\underline{c}}^{b(\theta, \hat{c})} \big((\theta - \gamma(\theta))v(q_e(\theta, c_e)) - c_e q_e(\theta, c_e)\big) g(c_e) dc_e \Bigg) f(\theta) d\theta.$$

The following incentive compatibility constraints ensure that the government does not manipulate the incumbent's cost

$$c \in \arg\max_{\hat{c} \in \mathcal{C}} \Phi(c, \hat{c}).$$

[55] $\mathbb{E}_c(\cdot)$ denotes the expectation operator with respect to the distribution of c.
[56] Myerson (1982).

Differentiating with respect to $\hat{c}$, we obtain

$$(6.2)\quad \frac{\partial \Phi}{\partial \hat{c}}(c,\hat{c}) = \int_{\underline{\theta}}^{\overline{\theta}} \Bigg[(1-G(b(\theta,c)))\,((\theta-\gamma(\theta))v'(q_i(\theta,c))-c)\,\frac{\partial q_i}{\partial c}(\theta,c)$$

$$+g(b(\theta,c))\left(-\left[(\theta-\gamma(\theta))v(q)-cq\right]_{q_e(\theta,b(\theta,c))}^{q(\theta,c)}+(c-b(\theta,c))q_e(\theta,b(\theta,c)))\right)\Bigg]\frac{\partial b}{\partial c}(\theta,c)f(\theta)d\theta.$$

To understand the incentives of the government to manipulate cost information, suppose that the mechanism aims to implement the *ex ante* efficient market allocation, i.e., a threshold level $b^*(\theta,c)\equiv c$ and a consumption level $q^{\varnothing}(\theta,\min(c,c_e))$. For such mechanism, we can use (6.2) to obtain

$$\left.\frac{\partial \Phi}{\partial \hat{c}}(c,\hat{c})\right|_{\hat{c}=c} = -\int_{\underline{\theta}}^{\overline{\theta}}(1-G(c))\gamma(\theta)v'(q^{\varnothing}(\theta,c))\frac{\partial q^{\varnothing}}{\partial c}(\theta,c)f(\theta)d\theta.$$

It immediately follows that

$$\left.\frac{\partial \Phi}{\partial \hat{c}}(c,\hat{c})\right|_{\hat{c}=c}\begin{cases}\leq 0 & \text{if } \gamma(\theta)<0 \quad \forall\theta,\\ \geq 0 & \text{if } \gamma(\theta)>0 \quad \forall\theta.\end{cases}$$

The mechanism consisting in asking the most efficient supplier between the incumbent and the fringe to produce the first-best quantity is thus not incentive compatible. The intuition is clear. When the government wants to induce over-production with respect to the first best (i.e., when $\gamma(\theta)<0$ for all θ), it just claims that the incumbent's cost is lower than its true value. When under-production is preferred (i.e., when $\gamma(\theta)>0$ for all θ), the government exaggerates the incumbent's cost.

6.3. *The Optimal Level of discretion*

REDISTRIBUTION TOWARDS HIGH-VALUATION CUSTOMERS. To understand how the government's incentives to manipulate costs can be controlled, let us first consider the case $\gamma(\theta)<0$ for all θ. The government would then like to claim that the incumbent has a lower cost than its true value. This scenario is also interesting because higher production levels are likely to yield higher profit levels for the incumbent and might thus also reflect some kind of regulatory capture from the incumbent. To be precise, the impact on profits is *stricto sensu* absent in our model since the incumbent's break-even constraint always binds at the optimum. However, had the incumbent earned strictly positive profits, higher production levels would imply higher profit levels. Such extra profits will appear as soon as the firm's control by the government is imperfect and several scenarios are consistent with such an assumption. First, the government might not be fully aware of the incumbent's cost and the incumbent may thus enjoy some informational rents. Second, incentivizing the incumbent to exert non-verifiable specific investments may require to give up some liability rents. Third, and more generally, the firm's profit may not be fully appropriated by the government when bargaining power is more evenly split between the government and the incumbent.

Consider thus the following simple delegation mechanism:

$$q_i(\theta,c)=\begin{cases} q_r(\theta,c) & \text{for } c\geq c^* \\ q_r(\theta,c^*) & \text{for } c\leq c^* \end{cases} \quad \text{and} \quad b(\theta,c)=\begin{cases} b^*(\theta,c) & \text{for } c\geq c^* \\ b^*(\theta,c^*) & \text{for } c\leq c^* \end{cases} \tag{6.3}$$

for some $c^* \in \mathcal{C}$.[57] With such mechanism, the government is free to choose among a restricted set of options corresponding to the optimal allocations it would implement had the incumbent's costs been common knowledge.

It can be readily checked that this mechanism is incentive compatible, as (6.2) is satisfied. This mechanism has also a very simple interpretation. It consists in leaving full discretion to the government provided that it announces a cost for the incumbent above the threshold c^*. For cost announcements that fall below that threshold, the government is bound to offer profiles of type-dependent quantities and switching rules that are independent of the cost announcement. Therefore, since $q_r(\theta,c)$ is a decreasing function of the incumbent's cost c, a first requirement is that consumption from the incumbent cannot be too high, or, in over words, the marginal values of regulated tariffs cannot be too low. Since $b^*(\theta,c)$ is an increasing function of c, a second requirement imposes a minimal market share for the fringe, or regulated prices cannot be too low in absolute levels as well.

As a benchmark, suppose the government has no redistributive concerns and efficiency is its sole concern, i.e., $\gamma(\theta)=0$ for all θ. Clearly, it should be that $c^*=\underline{c}$. Full discretion is granted to the government in the setting the incumbent's regulated tariffs. More generally, next proposition characterizes key properties of the optimal threshold c^* that would maximize an *ex ante* efficiency criterion when the government has redistributive concerns towards high-valuation customers.

Proposition 4. *Suppose that $\gamma(\theta)<0$ for all $\theta\in\Theta$. Then, it is always optimal to limit the government's discretion in setting regulated prices*

$$c^* > \underline{c}.$$

An interior optimum c^*, when it exists, is given by

$$\begin{aligned} -\underline{\mathcal{H}}(c^*)\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right)f(\theta)d\theta = \\ \int_{\underline{\theta}}^{\overline{\theta}}\Big(g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\left(v(q_e(\theta,b^*(\theta,c^*)))-v(q_r(\theta,\underline{c}))\right) \\ +(1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)v'(q_r(\theta,c^*))\Big)\gamma(\theta)f(\theta)d\theta, \end{aligned} \tag{6.4}$$

[57] In comparison with most of the delegation literature, the delegation mechanism we consider here, although it keeps much of the tractability found in the one-dimensional case that prevails in this literature, is bi-dimensional. Both the profile of consumption and the switching rule are determined by the mechanism. In the case of one-dimensional mechanisms, the extant literature has stressed that bunching is frequently found at the optimum as a result of the conflict of interests between the uninformed principal and his informed agent. Instead, with bi-dimensional mechanisms, new screening possibilities might be feasible. This issue is studied in Koessler and Martimort (2012) in much details.

where $\underline{\mathcal{H}}(c^*) \equiv \int_{\underline{c}}^{c^*} H(c)dc$.

When the government manipulates the information about the incumbent's cost, it becomes optimal to limit its discretion in setting the incumbent's regulated tariffs. In this first scenario, the government is tempted to understate the incumbent's cost as a way to implement an implicit redistribution across customers through an excessively high production level. Limiting the government's discretion amounts to preventing the government from unduly claiming a low cost for the incumbent operator. A cap on the incumbent's market share and a floor on its marginal prices are two instruments that achieve this goal.

REDISTRIBUTION TOWARDS LOW-VALUATION CUSTOMERS. For completeness, consider now the case $\gamma(\theta) > 0$ for all θ. The government would like to exaggerate the incumbent's cost in that scenario. Mirroring the above analysis, consider thus the following delegation mechanism

$$(6.5) \qquad q_i(\theta, c) = \begin{cases} q_r(\theta, c) & \text{for } c \le c^* \\ q_r(\theta, c^*) & \text{for } c \ge c^* \end{cases} \quad \text{and} \quad b(\theta, c) = \begin{cases} b^*(\theta, c) & \text{for } c \le c^* \\ b^*(\theta, c^*) & \text{for } c \ge c^* \end{cases}$$

for some $c^* \in \mathcal{C}$. With such a mechanism, both the marginal prices charged by the incumbent and the market share of the fringe cannot be too large.

PROPOSITION 5. *Suppose that $\gamma(\theta) > 0$ for all $\theta \in \Theta$ and that the following condition*

$$(6.6) \quad g(b^*(\theta, \bar{c}))\frac{\partial b^*}{\partial c^*}(\theta, \bar{c})\left(v(q_e(\theta, b^*(\theta, \bar{c}))) - v(q_r(\theta, \bar{c}))\right) + (1 - G(b^*(\theta, \bar{c})))\frac{\partial q_r}{\partial c^*}(\theta, \bar{c})v'(q_r(\theta, \bar{c})) > 0$$

holds. Then, it is optimal to limit the government's discretion in setting regulated prices

$$c^* < \bar{c}.$$

An interior optimum c^*, when it exists, is given by

$$(6.7) \quad \overline{\mathcal{H}}(c^*) \int_{\underline{\theta}}^{\bar{\theta}} \frac{d}{dc^*}\left((1 - G(b^*(\theta, c^*)))q_r(\theta, c^*)\right) f(\theta)d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \Big(g(b^*(\theta, c^*))\frac{\partial b^*}{\partial c^*}(\theta, c^*)\left(v(q_e(\theta, b^*(\theta, c^*))) - v(q_r(\theta, \underline{c}))\right) + (1 - G(b^*(\theta, c^*)))\frac{\partial q_r}{\partial c^*}(\theta, c^*)v'(q_r(\theta, c^*))\Big)\gamma(\theta)f(\theta)d\theta$$

where $\overline{\mathcal{H}}(c^*) \equiv \int_{c^*}^{\bar{c}}(1 - H(c))dc$.

In contrast with the case of redistribution towards high-valuation consumers, the condition for restricting the discretion of the government is now ambiguous. Indeed, notice, first, that $\frac{\partial b^*}{\partial c^*}(\theta, \underline{c}) > 0$ from (5.13). Second, (5.8) ensures that $v(q_e(\theta, b^*(\theta, \underline{c}))) - v(q_r(\theta, \underline{c})) > 0$ for all θ when $\gamma(\theta) < 0$. Hence, the first term in the integrand is always positive. The second term in the integrand is, however, negative.[58] Condition (6.6) holds

[58]This comes from the fact that $\frac{\partial q_r}{\partial c^*}(\theta, c^*) < 0$ as shown in the Appendix (see (A.13)).

when $1 - G(b^*(\theta, \bar{c}))$ is small enough, i.e., when it is very likely that the competitive fringe can produce at a much lower cost than the incumbent. In that scenario, it is again valuable to restrict the government's discretion.

7. Conclusion

Redistributive concerns imply that the production of the regulated incumbent departs from efficiency. Such concerns also justify the existence of regulated tariffs, even when the market is competitive. Our analysis shows, first, that regulated tariffs are always biased against the incumbent's competitors, and, second, that information about the incumbent may be strategically used so as to promote the government's objective. Limiting the discretion of the government in terms of price regulation is warranted. Yet, the precise manner of doing so depends on the bias in the government's objective. With a government biased towards low-valuation customers, it is best done by putting a cap on regulated tariff and imposing a minimal market share for the incumbent's competitors. The reverse holds for a government which is biased towards high-valuation customers.

Our analysis could be extended in several ways. First, we could relax the assumption that the government has full control of the incumbent's tariff and explicitly model how the government regulates this tariff even when, for instance, it has only imperfect information on the firm's cost. This imperfect knowledge would be the source of frictions.[59] We are confident that our findings would carry over to such a more complex scenario. The key ingredient to preserve a rationale for limiting the discretion of the government is to suppose that the government remains more informed than the rest of society on the incumbent's technology and it also pursues redistributive objectives that depart from efficiency.

Second, we have also viewed the government as a black-box and made the simplifying assumption that the regulator and the government are indeed a single entity. In practice, although regulators are subject to implicit political pressures to implement redistributive policies, they also are given mandates which, at least on surface, look like being only driven by efficiency considerations. Again, we are rather confident that a more detailed analysis of the political process would not change the basic thrust of our analysis.

Third, we have kept aside an important aspect of pricing in the electricity retail market, namely that it can also be used as a tool to induce customers to adopt new technologies (solar panels or new electrical appliances with remote control systems for demand side management, for instance) or change their behavior to reduce energy consumption.[60] In other words, the government's objectives might go beyond redistribution and also account for other environmental considerations. Distortions on pricing will be different than the ones our analysis has highlighted. Yet, these distortions will depart from efficiency and the very same trade-off between rules and discretion will carry over although details will of course differ.

Lastly, we have modeled competition on the retail market, ignoring the upstream segments of the supply chain. A more detailed analysis of the relationship between the

[59]The literature on the regulation of nonlinear pricing is almost inexistant. An exception is the complete information model of Sappington and Sibley (1992). Of course, the literature on price cap regulation analyzes how a firm chooses its prices for a bundle of services when subject to a global constraint. But, again, it is in complete information scenario where customers have known demand.

[60]This issue is studied in Feger, Pavanini and Radulescu (2017).

incumbent and entrants could be developed by taking into account the upstream segments of the supply chain and considering with more details the issue of vertical integration. To simplify, we have assumed that the incumbent faces a fixed cost and must raise enough revenue to break even. Some countries have, however, chosen a least stringent unbundling regime in which the incumbent remains somehow vertically integrated.[61] It justifies the fact that the incumbent remains regulated on the natural monopoly segment but also heavily monitored on the competitive segments (wholesale and retail markets) to avoid foreclosure and other restraints to competition. Taking those considerations into account could give a more realistic view of the market without changing the main insights.

References

ACER (2017). "Retail Markets Monitoring Report - CEER Report." Agency for Cooperation of Energy Regulators, November 2017, Ref: C17-MMR-83-02.

Aghion, P. and P. Bolton (1987). "Contracts as a Barrier to Entry." *The American Economic Review,* 77: 388-401.

Alonso, R. and N. Matouschek (2008). "Optimal Delegation." *The Review of Economic Studies*, 75: 259-293.

Amador, M. and K. Bagwell (2013). "The Theory of Optimal Delegation with an Application to Tariff Caps." *Econometrica,* 81: 1541-1599.

Armstrong, M. and D. Sappington (2007). "Recent Developments in the Theory of Regulation." In *The Handbook of Industrial Organization*, M. Armstrong and R. Porter eds., Vol 3: 1557-1700. North-Holland.

Arrow, K. (1975). "Vertical Integration and Communication." *The Bell Journal of Economics,* 173-183.

Atkinson, A. and J. Stiglitz (1976). "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics,* 6: 55-75.

Baron, D. and R. Myerson (1982). "Regulating a Monopolist with Unknown Costs." *Econometrica*, 50: 911-930.

Basov, S. (2005). *Multi-Dimensional Screening.* Springer.

Baumol, W., J. Panzar and R. Willig (1982). *Contestable Markets and the Theory of Industry Structure.* Harcourt Brace Javanovich.

Beesley, M. and S. Littlechild (1983). "Privatization: Principles, Problems and Priorities." *Lloyds Bank Review*, 149: 1-20.

Berg, S. and J. Tschirhart (1988). *Natural Monopoly Regulation: Principles and Practice.* Cambridge University Press.

Boiteux, M. (1956). "Sur la Gestion des Monopoles Publics Astreints à l'Equilibre Budgétaire." *Econometrica,* 24: 22-40.

Boiteux, M. (1993). *Haute tension.* Editions Odile Jacob.

Borenstein, S. (2012). "The Redistributional Impact of Nonlinear Electricity Pricing." *American Economic Journal: Economic Policy*, 4: 56-90.

[61]The least stringent form entails "*accounting unbundling*" whereas "*ownership unbundling*" is the most extreme one.

Borenstein, S. and L. Davis (2012). “The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets.” *The Journal of Law and Economics,* 55: 75-128.

Bovenberg, A. and L. Goulder (2002). “Environmental Taxation and Regulation.” In A. Auerbach and M. Feldstein eds., *Handbook of Public Economics,* Vol. 3, Chapter 23: 1471-1545.

Boyer, M. and J.J. Laffont (1999). “Towards a Political Theory of the Emergence of Environmental Incentive Regulation.” *The RAND Journal of Economics,* 40: 137-157.

Burns, P., I. Crawford and A. Dilnot (1995). “Regulation and Redistribution in Utilities.” *Fiscal Studies,* 16: 1-22.

Choné, P. and L. Linnemer (2016). “Nonlinear Pricing and Exclusion: I. Buyer Opportunism.” *The RAND Journal of Economics,* 46: 217-240.

CRE (2018). “Observatoire des Marchés de Détail.” 14 Décembre 2018, *Commission de Régulation de l’Energie.*

Cremer, H. and F. Gahvari (2002). “Nonlinear Pricing, Redistribution, and Optimal Tax Policy.” *Journal of Public Economic Theory,* 4: 139-161.

Cremer, H., F. Gahvari and N. Ladoux (1998). “Externalities and Optimal Taxation.” *Journal of Public Economics,* 70: 343-364.

Defeuilley, C. (2009). “Retail Competition in Electricity Markets,” *Energy Policy*, 37: 377-386.

Epstein, D. and S. O’Halloran (1994). “Administrative Procedures, Information and Agency Discretion.” *American Journal of Political Science*, 38: 697-722.

Epstein, D. and S. O’Halloran (1999). *Delegating Powers*. Cambridge University Press.

Erdogdu, E. (2014). “The Political Economy of Electricity Market Liberalization: A Cross-country Approach.” *Energy Journal*, 35: 91-128.

Faruqui, A. and S. Sergici (2010). “Household Response to Dynamic Pricing of Electricity: A Survey of 15 Experiments.” *Journal of Regulatory Economics*, 38: 193-225.

Faulhaber, G. and S. Levinson (1981). “Subsidy-Free Prices and Anonymous Equity.” *The American Economic Review,* 71: 1083-1091.

Feger, F., N. Pavanini and D. Radulescu (2017). “Welfare and Redistribution in Residential Electricity Markets with Solar Power.” CEPR discussion paper No 12517.

Feldstein, M. (1972a). “Equity and Efficiency in Public Sector Pricing: The Optimal Two Part Tariff.” *The Quarterly Journal of Economics,* 86: 175-187.

Feldstein, M. (1972b). “Distributional Equity and the Optimal Structure of Public Prices.” *The American Economic Review,* 62: 32-36.

Gangale F., J. Vasiljevska, C. Covrig, A. Mengolini and G. Fulli (2017). "E.U. Smart Grid Project Outlook (2017)." JRC Science for Policy Report. *Publications Office of the European Union*, ISBN 978-92-79-68898-0 (online).

Goldman, B., H. Leland and D. Sibley (1984). "Optimal Nonuniform Prices." *The Review of Economic Studies,* 51: 305-319.

Guesnerie, R. and J.J. Laffont (1984). "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm." *Journal of Public Economics,* 25: 329-369.

Hiriart, Y. and D. Martimort (2012). "How Much Discretion for Risk Regulators?" *The RAND Journal of Economics*, 43: 283-314.

Holmström, B. (1984). "On the Theory of Delegation." In *Bayesian Models in Economic Theory*, M. Boyer and R. Khilstrom eds., Elsevier Science Publishers.

Holmström, B. and R. Myerson (1983). "Efficient and Durable Decision Rules with Incomplete Information." *Econometrica*, 51: 1799-1819.

Hortacsu, A., S. Madanizadeh and S. Puller (2017). "Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market." *The American Economic Journal: Economic Policy,* 4: 192-226.

Huber, J. and C. Shipan (2002). *Deliberative Discretion: The Institutional Foundations of Bureaucratic Autonomy.* Cambridge University Press.

Huber, J. and C. Shipan (2006). "Politics, Delegation, and Bureaucracy." In B. Weingast and D. Wittman eds., *The Oxford Handbook of Political Economy*, Oxford University Press.

Iossa, E. and D. Martimort (2016). "Corruption in PPPs, Incentives and Contract Incompleteness." *International Journal of Industrial Organization,* 44: 85-100.

Ito, K. (2014). "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." *American Economic Review*, 104: 537-563.

Joskow P. and R. Schmalensee (1983). *Markets for Power - An Analysis of Electrical Utility Deregulation* . MIT Press.

Joskow P. (2000). "Why Do We Need Electricity Retailers? Or Can You Get It Cheaper Wholesale?" *Mimeo MIT*.

Koessler, F. and D. Martimort (2012). "Optimal Delegation with Multi-Dimensional Decisions." *Journal of Economic Theory*, 147: 1850-1881.

Laffont, J. J. (1994). "The New Economics of Regulation Ten Years After." *Econometrica*, 62: 507-537.

Laffont, J. J. (1996). "Industrial Policy and Politics." *International Journal of Industrial Organization,* 14: 1-27.

Laffont, J.J. and D. Martimort (2002). *The Theory of Incentives.* Princeton University Press.

Laffont, J.J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation.* MIT Press.

Ledyard, J. and T. Palfrey (1999). "A Characterization of Interim Efficiency with Public Goods." *Econometrica*, 67: 435-448.

Littlechild, S. (2000). "Why We Need Electricity Retailers: A Reply to Joskow on Wholesale Spot Pass-Through." Mimeo Faculty of Economics, University of Cambridge. Cambridge.

Locay, L. and A. Rodriguez (1992). "Price Discrimination in Competitive Markets." *Journal of Political Economy,* 100: 954-965.

Martimort, D. and A. Semenov (2006). "Continuity in Mechanism Design without Transfers." *Economic Letters*, 93: 182-189.

Martimort, D., J. Pouyet and L. Stole (2019). "Screening Contracts as a Barrier to Entry." Mimeo PSE.

Maskin, E. and J. Riley (1984). "Monopoly with Incomplete Information." *The RAND Journal of Economics,* 15: 171-196.

Melumad, N. and T. Shibano (1991). "Communication in Settings with No Transfers." *The RAND Journal of Economics*, 22: 173-198.

Milgrom, P. and I. Segal (2002). "Envelope Theorems for Arbitrary Choice Sets." *Econometrica*, 70: 583-601.

Myerson, R. (1981). "Optimal Auction Design." *Mathematical Operations Research*, 6: 58-73.

Myerson, R. (1982). "Optimal Coordination Mechanisms in Generalized Principal-Agent Models." *Journal of Mathematical Economics,* 10: 67-81.

Naito, H. (1996). "Tariffs as a Device to Relax the Incentive Problem of Progressive Income Tax System." Research Seminar in International Economics, Working Paper No 391, University of Michigan.

Posner, R. (1971). "Taxation by Regulation." *The Bell Journal of Economics and Management Science,* 22-50.

Rochet, J.C. (1985). "The Taxation Principle and Multi-Time Hamilton-Jacobi Equation." *Journal of Mathematical Economics*, 14: 113-128.

Rochet, J.C. (1987). "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context." *Journal of Mathematical Economics*, 16: 191-200.

Russo, A. (2015). "Pricing of Transport Networks, Redistribution, and Optimal Taxation." *Journal of Public Economic Theory,* 17: 605-640.

Sappington, D. and D. Sibley (1992). "Strategic Nonlinear Pricing under Price-Cap Regulation." *The RAND Journal of Economics*, 23: 1-19.

Spulber, D. (1984). "Scale Economies and Existence of Sustainable Monopoly Prices." *Journal of Economic Theory,* 34: 149-163.

Stiglitz, J. (1982), "Self-Selection and Pareto Optimal Taxation." *Journal of Public Economics,* 17: 213-240.

Vercammen, J., M. Fulton and C. Hyde. (1996). "Nonlinear Pricing Schemes for Agricultural Cooperatives." *American Journal of Agricultural Economics,* 78: 572-584.

Wilson, R. (1993). *Nonlinear Pricing.* Oxford University Press.

Appendix

Proof of Lemma 1. The proof is standard and follows from Rochet (1987) or Milgrom and Segal (2002). It is thus omitted. □

Proof of Proposition 1. We consider the relaxed problem obtained by omitting the constraints (3.2) and (3.3). We provide later on a condition that ensures the solution of the problem so relaxed satisfies those constraints.

Condition (3.6) must obviously hold as an equality at the optimum, which using (3.4) allows to obtain the following rewriting

$$U(\underline{\theta}) = -K + \int_{\underline{\theta}}^{\overline{\theta}} \left(\theta v(q(\theta)) - cq(\theta) - \int_{\underline{\theta}}^{\theta} v(q(\tilde{\theta}))d\tilde{\theta} \right) f(\theta)d\theta. \tag{A.1}$$

Integrating by parts the right-hand side of (A.1) yields

$$U(\underline{\theta}) = -K + \int_{\underline{\theta}}^{\overline{\theta}} \left(\left(\theta - \frac{1-F(\theta)}{f(\theta)} \right) v(q(\theta)) - cq(\theta) \right) f(\theta)d\theta. \tag{A.2}$$

We may now rewrite the government's objective, using (3.4) again, as follows

$$\int_{\underline{\theta}}^{\overline{\theta}} \alpha(\theta)f(\theta) \left(U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v(q(\tilde{\theta}))d\tilde{\theta} \right) d\theta,$$

or, using the normalization of social weights $\int_{\underline{\theta}}^{\overline{\theta}} \alpha(\theta)f(\theta)d\theta = 1$,

$$U(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} \alpha(\theta)f(\theta) \left(\int_{\underline{\theta}}^{\theta} v(q(\tilde{\theta}))d\tilde{\theta} \right) d\theta. \tag{A.3}$$

Denote now $\Lambda(\theta) = \int_{\theta}^{\overline{\theta}} \alpha(\tilde{\theta})f(\tilde{\theta})d\tilde{\theta}$ and observe that $\Lambda(\underline{\theta}) = 1$ and $\Lambda(\overline{\theta}) = 0$. Integrating by parts (A.3) yields a new definition of the government's objective as

$$U(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} \Lambda(\theta)v(q(\theta))d\theta. \tag{A.4}$$

Gathering finally (A.2) and (A.4) yields the new expression of the government's objective

$$\int_{\underline{\theta}}^{\overline{\theta}} \left(\left(\theta - \frac{1-F(\theta)-\Lambda(\theta)}{f(\theta)} \right) v(q(\theta)) - cq(\theta) \right) f(\theta)d\theta - K. \tag{A.5}$$

Maximizing pointwise this expression with respect to $q(\theta)$ yields (4.3).

Condition (4.1) then ensures that $q_r(\cdot)$ is non-decreasing. Condition (4.4) is obtained by imposing that $U(\underline{\theta})$ as defined in (A.2) is positive. □

Proof of Lemma 2. To simplify notations, denote $\varphi(\theta, c_e) = \theta v(q_e(\theta, c_e)) - c_e q_e(\theta, c_e)$ and observe that $\frac{\partial \varphi}{\partial \theta}(\theta, c_e) = v(q_e(\theta, c_e))$.

Necessity. Incentive compatibility constraints can be written as

$$\mathcal{U}(\theta, c) - \mathcal{U}(\hat{\theta}, c) \geq (\theta - \hat{\theta})(1 - G(b(\hat{\theta}, c)))v(q_i(\hat{\theta}, c)) + \int_{\underline{c}}^{b(\hat{\theta}, c)} (\varphi(\theta, c_e) - \varphi(\hat{\theta}, c_e))dG(c_e) \quad \forall(\theta, \hat{\theta}) \in \Theta^2$$

or

$$\text{(A.6)}\quad \mathcal{U}(\theta,c)-\mathcal{U}(\hat{\theta},c)\geq \int_{\hat{\theta}}^{\theta}\left((1-G(b(\hat{\theta},c)))v(q_i(\hat{\theta},c))+\int_{\underline{c}}^{b(\hat{\theta},c)}v(q_e(\tilde{\theta},c_e))dG(c_e)\right)d\tilde{\theta}\quad \forall(\theta,\hat{\theta})\in\Theta^2.$$

From this, we deduce that

$$|\mathcal{U}(\theta,c)-\mathcal{U}(\hat{\theta},c)|\leq |v(Q_{max})||\hat{\theta}-\theta|,$$

and thus $\mathcal{U}(\theta,c)$ is Lipschitz continuous. It is therefore absolutely continuous and in particular, differentiable almost everywhere.

It also follows from (A.6) that

$$\frac{\mathcal{U}(\theta,c)-\mathcal{U}(\hat{\theta},c)}{\theta-\hat{\theta}}\geq (1-G(b(\hat{\theta},c)))v(q_i(\hat{\theta},c))+\int_{\underline{c}}^{b(\hat{\theta},c)}\left(\frac{\varphi(\theta,c_e)-\varphi(\hat{\theta},c_e)}{\theta-\hat{\theta}}\right)dG(c_e),$$

and, permuting the roles of θ and $\hat{\theta}$,

$$\frac{\mathcal{U}(\theta,c)-\mathcal{U}(\hat{\theta},c)}{\theta-\hat{\theta}}\leq (1-G(b(\theta,c)))v(q_i(\theta,c))+\int_{\underline{c}}^{b(\theta,c)}\left(\frac{\varphi(\hat{\theta},c_e)-\varphi(\theta,c_e)}{\hat{\theta}-\theta}\right)dG(c_e).$$

Passing to the limit as $\hat{\theta}$ converges towards θ yields (5.1) which holds at any point of differentiability θ where $b(\theta,c)$ and $q_i(\theta,c)$ are both continuous.

Sufficiency. Consider a market allocation that satisfies (5.1). Then, we compute

$$\mathcal{U}(\theta,c)-\mathcal{U}(\hat{\theta},c)=\int_{\hat{\theta}}^{\theta}\left((1-G(b(\tilde{\theta},c)))v(q_i(\tilde{\theta},c))+\int_{\underline{c}}^{b(\tilde{\theta},c)}v(q_e(\tilde{\theta},c_e))dG(c_e)\right)d\tilde{\theta}\quad\forall(\theta,\hat{\theta})\in\Theta^2.$$

Observe that (A.6) holds when

$$\Gamma(\theta,\hat{\theta})=\int_{\hat{\theta}}^{\theta}\Bigg(\left((1-G(b(\tilde{\theta},c)))v(q_i(\tilde{\theta},c))+\int_{\underline{c}}^{b(\tilde{\theta},c)}v(q_e(\tilde{\theta},c_e))dG(c_e)\right)$$
$$-\left((1-G(b(\hat{\theta},c)))v(q_i(\hat{\theta},c))+\int_{\underline{c}}^{b(\hat{\theta},c)}v(q_e(\tilde{\theta},c_e))dG(c_e)\right)\Bigg)d\tilde{\theta}\geq 0\quad\forall(\theta,\hat{\theta})\in\Theta^2.$$

Observe that $\Gamma(\theta,\theta)=0$. Hence, we have

$$\Gamma(\theta,\hat{\theta})=\int_{\theta}^{\hat{\theta}}\frac{\partial\Gamma}{\partial\hat{\theta}}(\theta,\tilde{\theta})d\tilde{\theta}.$$

Moreover, we can check that

$$\Gamma(\theta,\hat{\theta})=-\int_{\hat{\theta}}^{\theta}\left(\frac{d}{d\hat{\theta}}\Big((1-G(b(\hat{\theta},c)))v(q_i(\hat{\theta},c))\Big)+\frac{\partial b}{\partial\hat{\theta}}(\hat{\theta},c)g(b(\hat{\theta},c))v(q_e(\tilde{\theta},b(\hat{\theta},c)))\right)d\tilde{\theta}$$

and thus $\frac{\partial\Gamma}{\partial\hat{\theta}}(\hat{\theta},\hat{\theta})=0$. Hence,

$$\Gamma(\theta,\hat{\theta})=-\int_{\theta}^{\hat{\theta}}\int_{\theta}^{\tilde{\theta}}\frac{\partial^2\Gamma}{\partial\theta\partial\hat{\theta}}(\theta_0,\tilde{\theta})d\theta_0 d\tilde{\theta}.$$

Finally, observe that $-\frac{\partial^2 \Gamma}{\partial\theta\partial\hat{\theta}}(\theta,\hat{\theta}) = \Psi(\theta,\hat{\theta})$ which ends the proof when (5.2) holds. □

Proof of Proposition 2. Using the incentive compatibility constraint (5.1), the budget constraint may be expressed as follows

$$K + \mathcal{U}(\underline{\theta}, c) \leq \int_{\underline{\theta}}^{\overline{\theta}} \Big((1 - G(b(\theta,c)))(\theta v(q_i(\theta,c)) - c q_i(\theta,c)) - \int_{\underline{\theta}}^{\theta} (1 - G(b(\tilde{\theta},c))) v(q_i(\tilde{\theta},c)) d\tilde{\theta} \\ + \int_{\underline{c}}^{b(\theta,c)} (\theta v(q_e(\theta,c_e)) - c_e q_e(\theta,c_e)) g(c_e) dc_e - \int_{\underline{\theta}}^{\theta} \int_{\underline{c}}^{b(\tilde{\theta},c)} v(q_e(\tilde{\theta},c)) dG(c_e) d\tilde{\theta} \Big) f(\theta) d\theta.$$

Integrating by parts to simplify the right-hand side, we obtain

$$\text{(A.7)} \quad K + \mathcal{U}(\underline{\theta}, c) \leq \int_{\underline{\theta}}^{\overline{\theta}} \left((1 - G(b(\theta,c))) \left(\left(\theta - \frac{1 - F(\theta)}{f(\theta)} \right) v(q_i(\theta,c)) - c q_i(\theta,c) \right) \right. \\ \left. + \int_{\underline{c}}^{b(\theta,c)} \left(\left(\theta - \frac{1 - F(\theta)}{f(\theta)} \right) v(q_e(\theta,c_e)) - c_e q_e(\theta,c_e) \right) g(c_e) dc_e \right) f(\theta) d\theta.$$

The government's objective writes now as follows

$$\int_{\underline{\theta}}^{\overline{\theta}} \alpha(\theta) f(\theta) \mathcal{U}(\theta,c) d\theta = \int_{\underline{\theta}}^{\overline{\theta}} \alpha(\theta) f(\theta) \Big(\mathcal{U}(\underline{\theta}, c) \\ + \int_{\underline{\theta}}^{\theta} \Big((1 - G(b(\tilde{\theta},c))) v(q_i(\tilde{\theta},c)) + \int_{\underline{c}}^{b(\tilde{\theta},c)} v(q_e(\tilde{\theta},c_e)) g(c_e) dc_e \Big) d\tilde{\theta} \Big) d\theta \\ = \mathcal{U}(\underline{\theta}, c) + \int_{\underline{\theta}}^{\overline{\theta}} \Lambda(\theta) \Big((1 - G(b(\theta,c))) v(q_i(\theta,c)) + \int_{\underline{c}}^{b(\theta,c)} v(q_e(\theta,c_e)) g(c_e) dc_e \Big) d\theta,$$

where the first line uses (5.1) and the second line relies on an integration by parts.

Solution to the relaxed problem. We first consider the maximization problem neglecting the incentive compatibility condition (5.2). Of course, the break-even condition (A.7) is binding at the optimum, i.e, the regulated tariff is chosen to exactly cover the firm's overall cost. Inserting the expression of $\mathcal{U}(\underline{\theta}, c)$ so obtained into the government's objective as expressed in the text, the optimal allocation solves

$$\max_{(q_i(\theta,c), b(\theta,c))_{\theta\in\Theta}} -K + \int_{\underline{\theta}}^{\overline{\theta}} \Big((1 - G(b(\theta,c))) \left((\theta - \gamma(\theta)) v(q_i(\theta,c)) - c q_i(\theta,c) \right) \\ + \int_{\underline{c}}^{b(\theta,c)} \left((\theta - \gamma(\theta)) v(q_e(\theta,c_e)) - c_e q_e(\theta,c_e) \right) g(c_e) dc_e \Big) f(\theta) d\theta.$$

Pointwise optimization with respect to $q_i(\theta,c)$ leads to $q_i(\theta,c) = q_r(\theta,c)$. Pointwise optimization with respect to the cut-off rule $b(\theta,c)$ leads to (5.6).

Sufficient conditions for (5.2) to hold. Observe that (5.2) is implied by

$$\text{(A.8)} \qquad \frac{\partial \Psi}{\partial \tilde{\theta}}(\theta,\tilde{\theta}) \geq 0 \quad \forall (\theta,\tilde{\theta}) \in \Theta^2$$

which amounts to

$$\text{(A.9)}\quad g(b^*(\tilde\theta,c))\left(v(q_e(\theta,b^*(\tilde\theta,c))-v(q_r(\tilde\theta,c))\right)\frac{\partial b^*}{\partial\theta}(\tilde\theta,c) + \left(1-G(b^*(\tilde\theta,c))\right)v'(q_r(\tilde\theta,c))\frac{\partial q_r}{\partial\theta}(\tilde\theta,c)\geq 0.$$

First, observe that Condition (4.1) is sufficient to ensure that $\frac{\partial q_r}{\partial\theta}(\theta,c)\geq 0$. Second, differentiating (5.6) with respect to θ yields (5.10). Inserting this equality into (A.9) taken for $\tilde\theta=\theta$ yields the condition

$$\frac{\partial\Psi}{\partial\tilde\theta}(\theta,\theta)\geq 0\quad\forall\theta\in\Theta$$

which can be rewritten as

$$\text{(A.10)}\quad \frac{(1-\dot\gamma(\theta))g(b^*(\theta,c))}{q_e(\theta,b^*(\theta,c))+\frac{\gamma(\theta)v'(q_e(\theta,b^*(\theta,c))}{\theta v''(q_e(\theta,b^*(\theta,c))}}\left(v(q_e(\theta,b^*(\theta,c)))-v(q_r(\theta,c))\right)^2 + \left(1-G(b^*(\theta,c))\right)v'(q_r(\theta,c))\frac{\partial q_r}{\partial\theta}(\theta,c)\geq 0.$$

Hence, this condition holds provided that (4.2) is satisfied. Finally, (A.8) holds provided that $\overline{\theta}-\underline{\theta}$ is not too large. □

PROOF OF PROPOSITION 3. Observe that (5.7) immediately follows from (5.6). Now, we may also rewrite (5.6) as

$$\text{(A.11)}\qquad \left[(\theta-\gamma(\theta))\left(v(q)-qv'q\right)\right]_{q_e(\theta,b^*(\theta,c))}^{q_r(\theta,c)}=-\gamma(\theta)q_e(\theta,b^*(\theta,c))v(q_e(\theta,b^*(\theta,c))).$$

The result then follows from observing that $v(q)-qv'(q)$ is increasing in q (since its derivative is worth $-qv''(q)>0$). □

PROOF OF PROPOSITION 4. The optimal threshold c^* maximizes an *ex ante* efficiency criterion and is thus solution to the following problem

$$\max_{c^*\in\mathcal{C}}\underline{\mathcal{W}}(c^*)$$

where

$$\begin{aligned}\underline{\mathcal{W}}(c^*)=&\int_{\underline{c}}^{c^*}\left(\int_{\underline{\theta}}^{\overline{\theta}}\left(\left(1-G(b^*(\theta,c^*))\right)\left(\theta v(q_r(\theta,c^*))-cq_r(\theta,c^*)\right)\right.\right.\\&\left.\left.+\int_{\underline{c}}^{b^*(\theta,c^*)}\left(\theta v(q_e(\theta,c_e))-c_eq_e(\theta,c_e)\right)g(c_e)dc_e\right)f(\theta)d\theta\right)h(c)dc\\&+\int_{c^*}^{\overline{c}}\left(\int_{\underline{\theta}}^{\overline{\theta}}\left(\left(1-G(b^*(\theta,c))\right)\left(\theta v(q_r(\theta,c))-cq_r(\theta,c)\right)\right.\right.\\&\left.\left.+\int_{\underline{c}}^{b^*(\theta,c)}\left(\theta v(q_e(\theta,c_e))-c_eq_e(\theta,c_e)\right)g(c_e)dc_e\right)f(\theta)d\theta\right)h(c)dc.\end{aligned}$$

Of course, such maximum exists since $\underline{\mathcal{W}}(c^*)$ is continuous and $\mathcal{C}=[0,\overline{c}]$ is compact.

Since $\underline{\mathcal{W}}(c^*)$ is in fact differentiable, we now also compute

$$\underline{\mathcal{W}}'(c^*) = \int_{\underline{c}}^{c^*}\int_{\underline{\theta}}^{\overline{\theta}} \Bigg((1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)(\theta v'(q_r(\theta,c^*))-c)+$$
$$g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\Big(\theta v(q_e(\theta,b^*(\theta,c^*)))-b^*(\theta,c^*)q_e(\theta,b^*(\theta,c^*))-(\theta v(q_r(\theta,c^*))-cq_r(\theta,c^*))\Big)\Bigg) f(\theta)d\theta h(c)dc.$$

Taking into account the definitions of $q_r(\theta,c^*)$ and $b^*(\theta,c^*)$ from (4.3) and (5.6), we obtain

$$\underline{\mathcal{W}}'(c^*) = \int_{\underline{c}}^{c^*}\int_{\underline{\theta}}^{\overline{\theta}} \Bigg((1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)\Big(\gamma(\theta)v'(q_r(\theta,c^*))+c^*-c\Big)+$$
$$g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\Bigg(\gamma(\theta)\Big(v(q_e(\theta,b^*(\theta,c^*)))-v(q_r(\theta,c^*))\Big)+(c-c^*)q_r(\theta,c^*)\Bigg)\Bigg) f(\theta)d\theta h(c)dc,$$

which we rewrite as

$$\begin{aligned}\underline{\mathcal{W}}'(c^*) = &\left(\int_{\underline{c}}^{c^*}(c^*-c)h(c)dc\right)\left(\int_{\underline{\theta}}^{\overline{\theta}}\frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right)f(\theta)d\theta\right)\\
&+\int_{\underline{c}}^{c^*}\left(\int_{\underline{\theta}}^{\overline{\theta}} g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\gamma(\theta)\left(v(q_e(\theta,b^*(\theta,c^*)))-v(q_r(\theta,c^*))\right)f(\theta)d\theta\right)h(c)dc\\
&+\int_{\underline{c}}^{c^*}\left(\int_{\underline{\theta}}^{\overline{\theta}}(1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)\gamma(\theta)v'(q_r(\theta,c^*))f(\theta)d\theta\right)h(c)dc.\end{aligned}$$

Integrating by parts, we find

$$\int_{\underline{c}}^{c^*}(c^*-c)h(c)dc = \int_{\underline{c}}^{c^*}H(c)dc \equiv \underline{\mathcal{H}}(c^*).$$

Inserting above, we obtain

$$\begin{aligned}\underline{\mathcal{W}}'(c^*) = &\underline{\mathcal{H}}(c^*)\int_{\underline{\theta}}^{\overline{\theta}}\frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right)f(\theta)d\theta\\
&+\int_{\underline{c}}^{c^*}\left(\int_{\underline{\theta}}^{\overline{\theta}} g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\gamma(\theta)\left(v(q_e(\theta,b^*(\theta,c^*)))-v(q_r(\theta,c^*))\right)f(\theta)d\theta\right)h(c)dc\\
&+\int_{\underline{c}}^{c^*}\left(\int_{\underline{\theta}}^{\overline{\theta}}(1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)\gamma(\theta)v'(q_r(\theta,c^*))f(\theta)d\theta\right)h(c)dc.\end{aligned}$$

We now compute

$$
\begin{aligned}
\underline{\mathcal{W}}''(c^*) =& \underline{\mathcal{H}}'(c^*)\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right) f(\theta)d\theta \\
&+ \underline{\mathcal{H}}(c^*)\int_{\underline{\theta}}^{\overline{\theta}} \frac{d^2}{dc^{*2}}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right) f(\theta)d\theta \\
&+ h(c^*)\int_{\underline{\theta}}^{\overline{\theta}} g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\gamma(\theta)\left(v(q_e(\theta,b^*(\theta,c^*)))-v(q_r(\theta,c^*))\right) f(\theta)d\theta \\
&+ h(c^*)\int_{\underline{\theta}}^{\overline{\theta}} (1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)\gamma(\theta)v'(q_r(\theta,c^*))f(\theta)d\theta \\
&+ \int_{\underline{c}}^{c^*}\left(\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left(g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\gamma(\theta)\left(v(q_e(\theta,b^*(\theta,c^*)))-v(q_r(\theta,c^*))\right)\right) f(\theta)d\theta\right) h(c)dc \\
&+ \int_{\underline{c}}^{c^*}\left(\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)\gamma(\theta)v'(q_r(\theta,c^*))\right) f(\theta)d\theta\right) h(c)dc.
\end{aligned}
$$

Observe that $\underline{\mathcal{H}}(\underline{c}) = \underline{\mathcal{H}}'(\underline{c}) = 0$ and, by assumption, $h(\underline{c}) > 0$. Hence, we get $\underline{\mathcal{W}}'(\underline{c}) = 0$ and

$$
\begin{aligned}
\underline{\mathcal{W}}''(\underline{c}) = h(\underline{c})\int_{\underline{\theta}}^{\overline{\theta}} \Big(& g(b^*(\theta,\underline{c}))\frac{\partial b^*}{\partial c^*}(\theta,\underline{c})\left(v(q_e(\theta,b^*(\theta,\underline{c})))-v(q_r(\theta,\underline{c}))\right) \\
&+ (1-G(b^*(\theta,\underline{c})))\frac{\partial q_r}{\partial c^*}(\theta,\underline{c})v'(q_r(\theta,\underline{c}))\Big)\gamma(\theta)f(\theta)d\theta.
\end{aligned}
$$

Observe now that when

$$
\begin{aligned}
\text{(A.12)}\quad & g(b^*(\theta,\underline{c}))\frac{\partial b^*}{\partial c^*}(\theta,\underline{c})\left(v(q_e(\theta,b^*(\theta,\underline{c})))-v(q_r(\theta,\underline{c}))\right) \\
&+ (1-G(b^*(\theta,\underline{c})))\frac{\partial q_r}{\partial c^*}(\theta,\underline{c})v'(q_r(\theta,\underline{c})) < 0
\end{aligned}
$$

holds, and since $\gamma(\theta) > 0$ for all $\theta \in \Theta$, we have $\underline{\mathcal{W}}''(\underline{c}) > 0$ and thus $\underline{\mathcal{W}}(c^*) > \underline{\mathcal{W}}(\underline{c})$ for c^* in a right-neighborhood of $\underline{c}$. Hence, restricting the government's discretion is optimal.

Now, observe, first, that $\frac{\partial b^*}{\partial c^*}(\theta,\underline{c}) > 0$ from (5.13). Second, Condition (5.8) ensures that $v(q_e(\theta,b^*(\theta,\underline{c}))) - v(q_r(\theta,\underline{c})) < 0$ for all θ when $\gamma(\theta) < 0$. Finally, using the definition of $q_r(\theta,c)$ and differentiating with respect to c, we obtain

$$
\text{(A.13)}\qquad \frac{\partial q_r}{\partial c^*}(\theta,c^*) = \frac{1}{(\theta-\gamma(\theta))v''(q_r(\theta,c^*))} < 0.
$$

Gathering those three facts, Condition (A.12) is shown to hold.

When interior, the optimal value of c^* is given by $\underline{\mathcal{W}}'(c^*) = 0$, which writes as (6.4).

Observe now that

$$
\begin{aligned}
\text{(A.14)}\quad & \frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right) = \\
& (1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*) - g(b^*(\theta,c^*))q_r(\theta,c^*)\frac{\partial b^*}{\partial c^*}(\theta,c^*).
\end{aligned}
$$

Taking into account (5.13) and (A.13), the expression in (A.14) is found to be negative.

□

Proof of Proposition 5. The optimal threshold c^* again maximizes an *ex ante* efficiency criterion which now writes as

$$\max_{c^*\in\mathcal{C}} \overline{\mathcal{W}}(c^*)$$

where

$$\overline{\mathcal{W}}(c^*) = \int_{c^*}^{\overline{c}} \Bigg(\int_{\underline{\theta}}^{\overline{\theta}} ((1-G(b^*(\theta,c^*)))\,(\theta v(q_r(\theta,c^*)) - cq_r(\theta,c^*)) \\ + \int_{b^*(\theta,c^*)}^{\overline{c}} (\theta v(q_e(\theta,c_e)) - c_e q_e(\theta,c_e))\, g(c_e)dc_e\big) f(\theta)d\theta \Bigg) h(c)dc \\ + \int_{\underline{c}}^{c^*} \Bigg(\int_{\underline{\theta}}^{\overline{\theta}} ((1-G(b^*(\theta,c)))\,(\theta v(q_r(\theta,c)) - cq_r(\theta,c)) \\ + \int_0^{b^*(\theta,c)} (\theta v(q_e(\theta,c_e)) - c_e q_e(\theta,c_e))\, g(c_e)dc_e\big) f(\theta)d\theta \Bigg) h(c)dc.$$

Of course, such maximum exists since $\overline{\mathcal{W}}(c^*)$ is continuous and $\mathcal{C}$ is compact.

Since $\overline{\mathcal{W}}(c^*)$ is in fact differentiable, we now also compute

$$\overline{\mathcal{W}}'(c^*) = \int_{c^*}^{\overline{c}} \int_{\underline{\theta}}^{\overline{\theta}} \Bigg((1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)(\theta v'(q_r(\theta,c^*)) - c) + \\ g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\Big(\theta v(q_e(\theta,b^*(\theta,c^*))) - b^*(\theta,c^*)q_e(\theta,b^*(\theta,c^*)) - (\theta v(q_r(\theta,c^*)) - cq_r(\theta,c^*))\Big)\Bigg) f(\theta)d\theta h(c)dc.$$

Taking into account the definitions of $q_r(\theta,c^*)$ and $b^*(\theta,c^*)$ from (4.3) and (5.6), we obtain

$$\overline{\mathcal{W}}'(c^*) = \left(\int_{c^*}^{\overline{c}} (c^*-c)h(c)dc\right)\left(\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right) f(\theta)d\theta\right) \\ + \int_{c^*}^{\overline{c}} \left(\int_{\underline{\theta}}^{\overline{\theta}} g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\gamma(\theta)\left(v(q_e(\theta,b^*(\theta,c^*))) - v(q_r(\theta,c^*))\right) f(\theta)d\theta\right) h(c)dc \\ + \int_{c^*}^{\overline{c}} \left(\int_{\underline{\theta}}^{\overline{\theta}} (1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)\gamma(\theta)v'(q_r(\theta,c^*))f(\theta)d\theta\right) h(c)dc.$$

Integrating by parts, we find

$$\int_{c^*}^{\overline{c}} (c-c^*)h(c)dc = \int_{c^*}^{\overline{c}} (1-H(c))dc \equiv \overline{\mathcal{H}}(c^*).$$

Inserting above, we obtain

$$\overline{\mathcal{W}}'(c^*) = -\overline{\mathcal{H}}(c^*)\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left((1-G(b^*(\theta,c^*)))q_r(\theta,c^*)\right) f(\theta)d\theta \\ + \int_{c^*}^{\overline{c}} \left(\int_{\underline{\theta}}^{\overline{\theta}} g(b^*(\theta,c^*))\frac{\partial b^*}{\partial c^*}(\theta,c^*)\gamma(\theta)\left(v(q_e(\theta,b^*(\theta,c^*))) - v(q_r(\theta,c^*))\right) f(\theta)d\theta\right) h(c)dc \\ + \int_{c^*}^{\overline{c}} \left(\int_{\underline{\theta}}^{\overline{\theta}} (1-G(b^*(\theta,c^*)))\frac{\partial q_r}{\partial c^*}(\theta,c^*)\gamma(\theta)v'(q_r(\theta,c^*))f(\theta)d\theta\right) h(c)dc.$$

We now compute

$$\begin{aligned}
\overline{\mathcal{W}}''(c^*) = &- \overline{\mathcal{H}}'(c^*) \int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left((1 - G(b^*(\theta, c^*)))q_r(\theta, c^*)\right) f(\theta)d\theta \\
&- \overline{\mathcal{H}}(c^*) \int_{\underline{\theta}}^{\overline{\theta}} \frac{d^2}{dc^{*2}}\left((1 - G(b^*(\theta, c^*)))q_r(\theta, c^*)\right) f(\theta)d\theta \\
&- h(c^*) \int_{\underline{\theta}}^{\overline{\theta}} g(b^*(\theta, c^*))\frac{\partial b^*}{\partial c^*}(\theta, c^*)\gamma(\theta)\left(v(q_e(\theta, b^*(\theta, c^*))) - v(q_r(\theta, c^*))\right) f(\theta)d\theta \\
&- h(c^*) \int_{\underline{\theta}}^{\overline{\theta}} (1 - G(b^*(\theta, c^*)))\frac{\partial q_r}{\partial c^*}(\theta, c^*)\gamma(\theta)v'(q_r(\theta, c^*))f(\theta)d\theta \\
&+ \int_{c^*}^{\overline{c}} \left(\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left(g(b^*(\theta, c^*))\frac{\partial b^*}{\partial c^*}(\theta, c^*)\gamma(\theta)\left(v(q_e(\theta, b^*(\theta, c^*))) - v(q_r(\theta, c^*))\right)\right) f(\theta)d\theta\right) h(c)dc \\
&+ \int_{c^*}^{\overline{c}} \left(\int_{\underline{\theta}}^{\overline{\theta}} \frac{d}{dc^*}\left((1 - G(b^*(\theta, c^*)))\frac{\partial q_r}{\partial c^*}(\theta, c^*)\gamma(\theta)v'(q_r(\theta, c^*))\right) f(\theta)d\theta\right) h(c)dc.
\end{aligned}$$

Observe that $\overline{\mathcal{H}}(\overline{c}) = \overline{\mathcal{H}}'(\overline{c}) = 0$ and, by assumption, $h(\overline{c}) > 0$. Hence, we get

$$\overline{\mathcal{W}}'(\overline{c}) = 0$$

and

$$\begin{aligned}
\overline{\mathcal{W}}''(\overline{c}) = -h(\overline{c}) \int_{\underline{\theta}}^{\overline{\theta}} \Big(g(b^*(\theta, \overline{c}))\frac{\partial b^*}{\partial c^*}(\theta, \overline{c})\left(v(q_e(\theta, b^*(\theta, \overline{c}))) - v(q_r(\theta, \overline{c}))\right) \\
+ (1 - G(b^*(\theta, \overline{c})))\frac{\partial q_r}{\partial c^*}(\theta, \overline{c})v'(q_r(\theta, \overline{c}))\Big)\gamma(\theta)f(\theta)d\theta.
\end{aligned}$$

Since $\gamma(\theta) < 0$ for all $\theta \in \Theta$, $\overline{\mathcal{W}}''(\overline{c}) > 0$ when (6.6) holds. Thus $\overline{\mathcal{W}}(c^*) > \overline{\mathcal{W}}(\overline{c})$ for c^* in a left-neighborhood of $\overline{c}$. Hence, restricting the government's discretion is here also optimal.

When interior, the optimal value of c^* is given by

$$\overline{\mathcal{W}}'(c^*) = 0,$$

which writes as (6.7). □