Designing Markets for Reliability with Incomplete Information

Job Market Paper - Preliminary Draft

Leopold Monjoie

[1] Université Paris Dauphine, PSL Research University, LEDa, Place du Marechal de Lattre de Tassigny, 75016, Paris, France

[2] Chair European Electricity Markets

[3] leopold.monjoie@dauphine.psl.eu

Abstract.

This paper examines the challenges of allocating a good subject to capacity constraints when considering consumer preferences and investment decisions. A theoretical framework is developed where a market designer sequentially chooses a level of investment and proposes an allocation mechanism to consumers followed by a consumption stage. The market designer uses the allocation to maximize consumer surplus and finance the investment cost. He faces heterogeneous consumers who have private information about their demand level and belong to a publicly observed category. We show that the lack of complete information about consumer utility and constraints on the implementable mechanism leads to specific relations between the optimal allocation mechanism and the level of investment. Namely, we find that the optimal allocation implies discriminating consumers based on their types and that discrimination depends on the level of investment considered. This has significant welfare and distributive implications: an optimal pricing mechanism can minimize the investment cost and lead to a higher aggregate consumer surplus depending on the environment. However, it is not always a Pareto improvement for every consumer. We first study the benchmark case with complete information. We then analyze the current second-best situation, in which the market designer cannot obtain information about consumers and must choose fixed prices ex-ante. In the third step, we describe the optimal theoretical second-best allocation mechanism that considers the incentive and individual rationality constraints and the investment decisions.

1 Introduction

Economists have long advocated that pricing mechanisms should be carefully designed to allow the coverage of investment costs and promote efficient resource use. This is particularly true when providing *essential goods* characterized by the public-good nature of investment availability when supply is scarce. In those sectors, demand and supply fluctuate unpredictably, and if any demand exceeds the available capacity and cannot be efficiently rationed, it generates significant welfare losses.[1] For instance, without sufficient investment, the reliability of the electricity supply can be compromised, leading to frequent outages and power interruptions (IEA, 2020). The necessary increase of investment in clean but uncertain technology and the electrification of our end-use consumption implies that we must design the upcoming electricity sector carefully (IEA, 2021). However, electricity retail markets remain relatively simple, preventing consumers from reacting efficiently to the electricity system's scarcity. Similarly, transportation and distribution infrastructure are also central in current policymakers' debates in the energy sector. Network tariffs are usually designed to cover transmission lines' investment and operation costs. Still, the growing share of decentralized production and the intermittent nature of renewable production create new challenges(Eurelectric, 2021).[2] The COVID crisis has also shown that the lack of production capacity for medical goods, especially vaccines, has severe consequences. The absence of sufficient capacity to produce vaccines led to a worldwide lockdown and border closures, while also increasing contagions and hospital congestion. For instance, Kominers and Tabarrok (2022) and Athey et al. (2022) showed that the price incentives for providing new vaccines and expanding production capacities were largely sub-efficient compared to their social value. On the other hand, the crisis also highlights the issue of who should be allocated the vaccines, given the scarcity of available production capacity.[3] Finally, congestion in transportation systems continues to generate substan-

[1]This inefficient rationing usually stems from policymakers using price regulations, for instance, price caps, or because they are reluctant to implement complex pricing mechanisms due to technical, political, or equity reasons.

[2]The increase of volatility from both the supply and demand side implies that the sizing of networks must be rethought. Adding a new line to satisfy a high-magnitude but rare event is not necessarily optimal. In this case, the incentive to better size the network can also come through incentives through tariffs.

[3]In this context, (Akbarpour et al., 2023a) underlined that the classic opposition between prices and free, but random, allocation is not straightforward.

tial costs (Schrank and Lomax, 2021). It can also pose challenges and prevent the much-needed modal shift to low-carbon means essential for the energy transition (ITDP, 2021).

Previous discussions in the economic literature have recognized the public-good nature of available investment for such markets. They are mainly centered around the supply side of the problem. Namely, how to implement mechanisms to procure sufficient investment at the least cost and consider the private incentives producers face, which may differ from the optimum.[4]. However, most papers consider the demand as given without considering how consumers may value the available capacity.

The central contribution of this paper is to discuss the implications of considering the demand side when it comes to ensuring an efficient level of investment. Leaving aside any supply-side considerations, we provide a theoretical framework highlighting the inherent tensions that arise when implementing an allocation mechanism that (i) dictates how agents consume the goods and (ii) generates revenue to finance new investments in an incomplete information framework with heterogeneous consumers. We notably assume that the utility buyers derive from consuming the goods is uncertain and private information. The question of consumer heterogeneity, and by extension, that of redistribution generated by an allocation mechanism that is considered more efficient, has been recently studied in several empirical papers in the context of essential goods. For instance, in Cahana et al. (2022), the authors explore the redistributive effects of switching from a flat electricity price to real-time pricing. Depending on the design, low-income households may lose due to specific consumption patterns in the face of available supply.[5] The scarcity of vaccines creates a trade-off between protecting the most vulnerable (e.g., elderly), the likely spreader (e.g., students), or the individual bringing the highest economic benefits (e.g., front-line

[4]Those mechanisms can range from direct subsidies to the design of more complex competitive markets. An important stream of literature notable in electricity markets is focused on studying long-term markets in which producers offer either future production via long-term contracts (Ausubel and Cramton, 2010) or their future availability through, for instance, capacity remuneration mechanisms (Léautier, 2016; Holmberg and Ritz, 2020).

[5]Levinson and Silva (2022) have studied the rates implemented by utilities in the U.S. and how they take into account redistribution preferences in their design. Due to the rapid increase in residential rooftop solar photovoltaic, electricity network tariffs have also been studied, notably in the Californian markets. If the tariffs are mostly based on variable parts, then non-adopters tend to cross-subsidies adopters of such technologies. One central issue is that the latter are mostly high-income households (Burger et al., 2020).

health workers).[6] Finally, in the case of congestion pricing, Hall (2021) studies when the pricing of a lane portion leads to a Pareto improvement for all users. The author finds a fully efficient toll unnecessary for sufficient welfare and Pareto improvement. However, most of the recent works study the short-term effect of pricing issues without considering the long-term interactions with the level of investment. Therefore, in this paper, we provide a theoretical foundation to analyze the interaction between a set of heterogeneous consumers, the choice of the pricing mechanism, and the use of the revenue generated through this mechanism to increase available capacity.

The core results of the paper are the interaction between the level of investment and the proposed allocation mechanism under different sets of assumptions and constraints. We show that the lack of complete information about consumer utility and constraints on the implementable mechanism leads to specific relations between the optimal allocation mechanism and the level of investment. We find that the optimal allocation implies discriminating consumers based on their types and that discrimination depends on the level of investment considered. This means the most efficient mechanism is not always Pareto-improving for every consumer, even considering the increased available capacity. It has, therefore, significant welfare and distributive implications.

We study a market designer, which can be interpreted as a public authority or a regulated monopoly, that (i) determines the allocation in prices and quantities of a homogeneous good and (ii) chooses the level of investments that maximizes consumer surplus. The allocation mechanism defines the per-unit monetary transfer and the quantity for a set of consumers during the consumption stage subject to capacity constraint. Therefore, when the market designer chooses the investment level and the allocation of the good across its consumers, he needs to consider the potentially asymmetrical effects that a given allocation can have when the capacity is binding or not. We will show that depending on the ability to propose an allocation based on the different states of the world, the consideration of the capacity constraint can significantly impact the design of the efficient mechanism.[7] Hence, we find that the (potential) incompleteness of the mechanism pro-

[6](Sudarmawan et al., 2022) shows how countries choose who should receive the vaccine first. (Rahmandad, 2022) describes the tradeoff of allocating the vaccines between the most vulnerable and the high-transmission individuals. Finally, Persad et al. (2020) discusses the ethical consideration of allocating vaccines .

[7]In this paper, the capacity constraint is hard in the sense that we do not represent the costs associated with demand exceeding available supply. Therefore, the market designer can always reduce demand but at the cost of misallocation due to imperfect information. Several papers have described those costs in more detail, such as rolling

posed by the market designer due to implementation constraints can have significant impacts. As the market designer uses the allocation to maximize consumer surplus and finance the investment cost, he is also under a budget constraint.

The consumers are characterized by a linear utility function, which is uncertain when the market designer makes investment decisions and proposes an allocation mechanism. This uncertainty has two additive components: (i) a common shock that is identical across all consumers, and (ii) a private shock only observed by consumers before the consumption stage and the realization of the common shock. We also embedded each consumer with a category for which the market designer is publicly informed. The existence of private information with respect to their consumption implies that consumers' private incentives might also differ from the market designer's objective. Therefore, the allocation mechanism in our framework is used simultaneously to generate revenue to cover investment costs and screen for unobservable characteristics to ensure efficient consumption.

We analyze several market design cases. In section 3, we start with the first-best, in which the market designer perfectly observes the consumer type when choosing the investment level and the allocation. Then, we describe a *short-termist* market designer, which separates the investment decisions and the allocation mechanism proposal. Namely, the market designer sequentially makes the investment decisions under a budget constraint and chooses the allocation maximizing consumer surplus given capacity constraints. Section 4 analyzes the *current* second-best implemented across many markets. The market designer faces private information about the level of consumption and is constrained in the monetary transfer he can implement. Namely, the price is unique for every state of the world, and it can vary based on the category of consumers. Finally, in section 5, we look at the theoretical second-best case under incomplete information. We implement a mechanism design approach in the framework to study the relation between the allocation mechanism and the incentive and participation constraints. The remainder of this section discusses the related literature. Section 2 presents the environment.

Related Literature

blackouts in electricity (Fabra, 2018; Llobet and Padilla, 2018) or congestion costs in transports (Yoshida, 2008; de Palma et al., 2017).

We build our framework on several strands of literature. The dynamic interaction between investment decisions and the consumption stages stems from the *peak-load pricing theory* that originated from Boiteux (1949). It describes how capacity constraints interact with the provisions of a homogeneous good with time-varying uncertain stochastic demand. It has mainly been used in recent work to study the role of market power, as in Léautier (2016), where producers can increase the price on the spot market beyond marginal cost even though they are not capacity-constrained. The paper also introduces some long-term agreements with producers offering in capacity remuneration mechanisms and short-term markets. The effect of price regulation is also analyzed in Leautier (2018), where the author demonstrates that short-term inefficiencies can sometimes have long-term and counterintuitive effects. In this paper, the price cap changes the private incentives producers face, hence the final investment decisions. In line with the current paper, Holmberg and Ritz (2020) study the effect of having inefficient rationing (but due to inelastic demand, not private information). When demand exceeds capacity, an additional welfare loss exists. Consequently, electricity prices do not internalize this additional cost, and the market designer needs to implement an additional stream of revenue for the producers. Our work introduces two features in the model: (i) heterogeneous consumers with private information and (ii) inefficiencies due to the schedule commitment by the market designer before the uncertainty is resolved.[8]

A second stream of papers is also related to the electricity markets and is based on the seminal paper by Chao and Wilson (1987) on priority service. The central idea is to provide a mechanism design solution in the form of a contractual arrangement where consumers choose at the same time the allocation during the wholesale market, which is in the same vein as the allocation schedule of this paper, and the probability of being disconnected when demand exceed the level of capacity. This framework has been refined by a series of papers by the same authors, including the comparison with other market arrangements Chao et al. (2022) and the role of risk aversion Chao (2012). We also relate to a series of papers focusing on implementing the second-best pricing method for consumers with incomplete information in Spulber (1992a,b, 1993). The work in Spulber (1992b) focuses on an incomplete information framework without endogenous investment decisions. The

[8]Similarly, our work mirrors the literature from congestion pricing theory from Vickrey (1963, 1969). For a recent theoretical paper, see, for instance, de Palma et al. (2017), which also compares different allocation mechanisms, but without considering the demand side.

optimal allocation schedule is non-linear because consumers' type is private information. Therefore, the market designer faces some challenges when implementing such schedules. In Spulber (1992a), a regulated firm is introduced to consider its budget constraint. However, the focus of this paper remains circumscribed to the design of consumers' second-best tariffs. Finally, Spulber (1993) studies the case of a monopoly designing the rates under incomplete information. We depart from this literature by deepening the private incentives consumers might have by behaving strategically from the truthful reporting and by tightening the link with the investment decisions framework developed in the previous paragraph.

The issues related to distributive concerns are borrowed from a growing body of literature using mechanism design. In particular, consumers' characteristics with private information and publicly-observed categories are related to the papers from Akbarpour et al. (2023a) and Akbarpour et al. (2023b), which study the trade-off between allocating certain vaccines on a free but random basis or using prices to discriminate and extract information from consumers. In the two papers, the authors assumed that the market designer has distributive and exogenous revenue preferences. Therefore, the model exhibits a tension of allocating the good via prices, which generates some revenue, or via a random free allocation that minimizes distributive issues. Our paper allows us to endogenize the revenue preference by implementing investment decisions with a budget constraint. We also provide results when the policymaker can imperfectly implement prices. [9]

2 Environment

We describe in the section the idiosyncratic characteristics of an electricity system. Note that while the terminology is specific, the results can be applied to other essential goods as described in the introduction. (i) The demand side, which can be interpreted as households, industrial

[9] This paper fits the new literature on industrial organization using an incomplete information framework. Triple-IO (for Incomplete Information Industrial Organization) papers aim to underline traditional industrial organization issues and how they can be renewed when imperfect information exists. See, for instance, the literature review by Loertscher and Marx (2021). This paper deals with the effect of capacity-constrained systems where (inefficient) rationing must be implemented. It fits with some works by Loertscher and Muir (2020, 2021) and Gilbert and Klemperer (2000), which studies pricing and rationing decisions within imperfect information. We add to the existing literature by providing a similar framework but by including investment decisions and a different type of rationing mechanism.

consumers, or retailers participating in the electricity market; (ii) The allocation mechanism that defines how the market designer allocates (in terms of quantity and financial transfer) electricity to the demand side. (iii) The supply side concerns how investment and production decisions are made. This current version of the paper focuses solely on a market designer configuration where investment decisions are made to maximize consumer surplus. From an outcome perspective, this is similar to having either a monopolist subject to budget constraint or a set of perfect competitive producers without market failure or public interventions. (iv) The decisions' timing.

2.1 Consumers Preferences

There exists a unit mass of consumers for electricity. Each consumer is characterized by a type vector (i, θ, s). The first characteristic refers to the consumer category, such as, for instance, a consumer being a household or an industry. There is a finite set of categories such that $i \in \{1, 2\}$. It is publicly observed, and the size of each category, i.e., the number of agents, is denoted by $\mu_i > 0$ for each group. Each consumer is characterized by a demand level θ, which, under an incomplete information framework, is assumed to be privately observed by the consumer. Conditional on belonging to a category i, this value is drawn from a common-knowledge cumulative distribution function distribution G_i whose continuous density is $g_i > 0$ has full support and is strictly positive on $[\underline{\theta}_i; \bar{\theta}_i]$. With households, θ could represent the revenue shocks, the lowest type of consumer being the poor household and the highest type of consumer being the more prosperous household. Industrial consumers could also be modeled with this framework, where θ represents their buyers' orders (see Chao (2012) for a micro foundation). When we define a consumer category i as being of a higher type concerning a category j as follows:

Definition 1. *If the consumer category i is of a higher type than consumer category j, then $G_i(\theta)$ first-order stochastically dominates $G_j(\theta)$*

For instance, if we assume the same distribution for both categories, then it implies that we have $\bar{\theta}_i \geq \bar{\theta}_j$ and $\underline{\theta}_i \geq \underline{\theta}_j$. On the other hand, we suppose that consumers are also subject to an individual but identical shock represented by s. Every agent in the game knows this value. It can mean, for instance, weather shock or specific economic conditions (recession) observable by everyone. This shock follows a common-knowledge continuous distribution $F > 0$ whose density $f > 0$ has full support on $s \in [0, \bar{s}]$. In this framework, the demand shock is the same for all consumers, and the aggregate shock equals $2s$. For now, we also assume uniform distribution for the common shock and the private information.

We assume each consumer type is known before the demand shocks are realized in this initial environment. Therefore, this framework encompasses two interpretations of the demand shocks: (i) a static model, where a single shock is realized, and there is uncertainty concerning its realization. (ii) a repetition of multiple shocks over a given period (for example, one year), which are drawn from the distribution $F(.)$ Léautier (2016). In the last interpretation, we assume that the type of consumer does not change between different shocks and is determined before this given period. Expectation operator $\mathbb{E}_s$ will denote the expectations over every state of the world. All agents in the game are assumed risk-neutral.

We define a consumer's utility belonging to a category i of a type θ_i. The value for electricity consumption for each consumer is denoted: $U(q, \theta, s) = \int_0^q u(\tilde{q}, \theta, s) d\tilde{q}$, with q the quantity of electricity allocated to the consumer. u can be interpreted as the marginal willingness to pay for a given quantity of electricity. If a consumer receives a quantity q in exchange for a monetary transfer t, we define the indirect utility function, also referred to as the consumer surplus, as $V(q, \theta, s) = U(q, \theta, s) - tq$. If a consumer does not receive electricity, we assume its value is null. Finally, we assume that u is linear of the form: $u(q, \theta, s) = \theta + s - q$

2.2 Allocation design

Given a total quantity $Q(s)$ of electricity in state of the world s, a general allocation mechanism $\mathcal{M}$ can be described via a collection of functions $q_i : [\underline{\theta}_i, \bar{\theta}_i] \to \Delta(Q(s))$ where q_i is a function describing the quantity q of electricity allocation to a consumer with type θ in category i at a state s. The aggregate quantity allocated to a group i of consumers is $Q_i(s) = \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta)$. The total allocation is recover with $Q(s) = \sum_i Q_i(s) = \sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta)$.

We also define the function $t_i(\theta, s)$ as the monetary transfer assigned to a consumer with type $\theta \in [\underline{\theta}_i, \bar{\theta}_i]$ in category i at state s. To study the optimal second-best mechanism with incomplete information in section 5, we rely on the Revelation Principle. Given a direct mechanism $(q_i, t_i)_{i=\{1,2\}}$, for each category, consumers report their type θ, receive an allocation $q_i(\theta)$, and pays $t_i(\theta)$ to the market designer. From a pricing mechanism perspective, the mechanism design approach is similar to forward contracting, where the market designer fixed ex-ante both the allowable quantities at a given price for a given realization of s (Chao, 2011). In this paper, we also provide another pricing mechanism that we name *market allocation*, which is formally defined as follows:

Definition 2. *A market allocation for a consumer is defined such that:*

1. *The market designer offers the consumer an inverse supply function $O^{-1}(q)$ associating a quantity and a unit monetary transfer.*
2. *The consumer selects their quantity consumed given the supply function based on their demand function $d(t, \theta, s)$ with $d(t, \theta, s) = u^{-1}(t, \theta, s)$.*

When the market allocation is chosen and compared to the mechanism design approach, the market designer does not have to choose the quantity as the following relation defines it: $q_i(\theta, s) = d(t_i(\theta), \theta, s)$. For a category i of consumers, the aggregate electricity demand is $Q_i = d(t_i, s) = \mu_i \int_{\theta_i} d(t_i(\theta), \theta, s) dG_i(\theta))$. Finally, The inverse demand functions for each category as $p_i(Q_i, s) = D_i^{-1}(Q_i, s)$, and the aggregate function for all consumers is given by $p(Q(s), s) = \sum_i D_i^{-1}(Q_i, s)$. The comparison between the mechanism design and the market allocation will be used for different reasons. In particular, we will show that the market allocation can be seen as an *implementable mechanism* mimicking the mechanism design outcome, or on the contrary, as a constraint for the market designer. To present some current pricing mechanisms, we use market allocation as a source of inefficiencies.

2.3 Supply side

We assume the most straightforward form for the supply side. A direct interpretation is that the market designer collects total revenues $\sum_i \mu_i \int_{\theta_i} t_i(\theta) q_i(\theta_i) dG_i(\theta)$ and makes investment decisions in productive capacity. It encompasses the literature on the management of a public firm or the direct regulation of a private monopolist. The model could describe a market designer acting as an intermediary between consumers and producers, marking production and investment decisions. In that case, the mechanism between consumers and the market designer could be understood as a theoretical retail market, and the mechanism between producers and the market designer would be a wholesale market. The main idea is that we remain agnostic about the proper form of the allocation mechanism between producers and the market designer. However, we assume it is fully efficient in that the production and investment decisions are made in the same manner under optimal regulation (for instance, the market designer acts as a single buyer in a market with perfectly competitive producers). In the rest of the paper, we abstract from those details. We assume a market designer making both production and investment decisions.

We denote the level of investment k. The investment cost is linear with $I(k) = rk$. The production cost is unitary and normalized to 0. The capacity level k implies a capacity constraint such that for any total quantity allocation Q and any realization of s, we must have at $Q(s) \leq k$.

Following the description of the supply side, the market designer chooses the allocation and the investment decisions to maximize a long-term expected and aggregate welfare function:

$$\int_s \underbrace{\sum_i \mu_i \int_{\theta_i} \lambda_i(\theta) V_i(\theta, s) dG_i(\theta)}_{\text{Consumers surplus}} + \underbrace{\sum_i \mu_i \int_{\theta_i} t_i(\theta) q_i(\theta) dG_i(\theta)}_{\text{Revenue}} \; dF(s)$$

Where $\lambda_i(\theta)$ represents welfare weights on a given consumer. We will specify their role later in further research. For now, we assume that they are equal to $G_i(\theta)$. They are neutral for maximizing the welfare function, which boils down to the expected sum of the consumer's surplus function.

2.4 Timing

We assume a multi-period game where:

1. **Information stage.** The consumers (and the market designer under complete information) learn about consumer types.

2. **Investment Decision**. The market designer chooses the level of investment k

3. **Allocation Proposal**.

   (a) The market designer chooses an allocation schedule (which can be market or mechanism-based) offered to the consumers. The allocation can be fully complete if it depends on all the realization of s, or incomplete if some constraints limit the allocation to some realization of s.

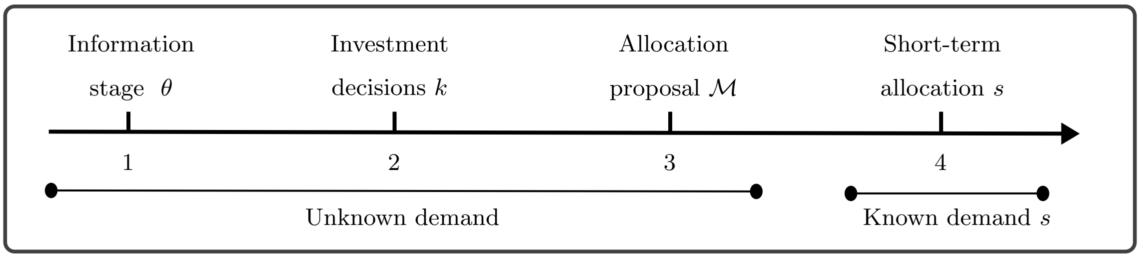
   (b) Consumers accept or reject the offer (in this case, the consumer does not participate in the third stage and receives no electricity.

4. **Short-term allocation**.

   (a) The realization of the common shock is known to every agent or the given period that occurs.

   (b) The allocations are realized following what has been proposed in the third stage.

We summarize the timeline of the game below:



3 Complete Information

3.1 Optimal allocation proposal

The first regime we study is the complete information case concerning consumer type. It can be understood as a nonstrategic regime with complete information in the sense that consumers reveal their type honestly. For each realization of the shock s, we define the allocation under complete information with $q_i^*(\theta, s)$ that maps the observed type of each consumer for each category to the quantity allocated. The monetary transfer $t_i^*(\theta, s)$ maps the observed type of each consumer to the per-unit payment made by the consumer to the market designer. This framework can be understood as the market designer offering a price/quantity allocation schedule that varies depending on the demand shock s. We derive its problem as follows:

$$\max_{\substack{t_i(\theta,s),\\ q_i(\theta,s),\\ k\geq 0}} \quad CS(k) = \sum_i \mu_i \int_s \int_{\theta_i} U(q_i(\theta, s), \theta, s) \; - \; t_i(\theta, s) q_i(\theta, s) dG_i(\theta) dF(s)$$

$$\text{s.t.} \quad I(k) \; \leq \; \sum_i \mu_i \int_s \int_{\theta_i} t_i(\theta, s) q_i(\theta, s) \; dF(s), \tag{R}$$

$$\sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) \leq k, \tag{K}$$

The first constraint follows the principle that the market designer should avoid any negative revenue at the optimum level of investment. It allows the rewrite of the objective function by replacing the payment part directly with the investment cost. For consistency with the rest of the analysis, we keep separated this constraint. In other words, under our supply-side assumption and given the absence of production cost, the entire income is allocated to financing the investment costs. The second constraint is the capacity constraint. We also include implicitly the conditions such that q_i and t_i are positive and that every consumer derives a null or positive surplus when participating in the mechanism. Finally, from a timing perspective, the constraints should be considered simultaneously. In the next section, we develop the implications of having sequential constraints.

We show in Proposition 1: (1) The optimal schedule in price (unit monetary transfer) and quantity. (2) The condition for the first-best investment level. (3) The market allocation implements the first-best schedule. Recall the market allocation is defined by a price t_i linked to the allocation schedule q_i such that $q_i(\theta, s) = d(t_i(\theta, s), s)$ with $d(t, \theta, s) = u^{-1}(t, \theta, s)$. Moreover, let defines the functions $p(q, s) = \sum \mu_i \int_s \int_{\theta_i} u(q, \theta, s) dG_i(\theta) dF(s)$ and $s_1(k) = \Big\{ s | \sum_i \mu_i \int_{\theta_i} d(0, \theta, s) dG_i(\theta) = k \Big\}$.

Proposition 1. *(1) The optimal unity monetary transfer and quantity schedule is defined for each realization of s as follows:*

$$t^*(k,s) = \begin{cases} 0 \\ p(k,s) \end{cases} \quad \textit{and} \quad q_i^*(k,\theta,s) = \begin{cases} d(0,\theta,s) & \textit{if} \quad s \in [0, s_1(k)) \\ d(p(k,s),\theta,s) & \textit{if} \quad s \in [s_1(k), \bar{s}] \end{cases}$$

With $s_1(k)$ is defined as the first state of the world when the capacity is binding.

(2) The optimal level of investment is given by the equality between the marginal investment cost and the expected aggregate marginal utility when the capacity is binding:

$$k^* = \left\{ \quad k \quad | \quad r = \sum_i \mu_i \int_{s_1(k)}^{\bar{s}} \int_{\theta_i} u(k,\theta,s) dG_i(\theta) dF(s) \right\}$$

(3) If the market designer implements a market mechanism with a supply function given by the monetary transfer schedule $t^*(k,s)$ and when consumers offers truthfully their demand functions, the market outcome is the first-best allocation.

Solving for the Lagrangean shows that when the capacity is not biding, the optimal allocation is characterized by an expected marginal utility null: $\int_{\theta_i} u(q_i^*(\theta,s),s)dG_i(\theta) = 0$. On the other hand, when the capacity is binding, the optimal allocation should be equal to the marginal investment cost: $\int_{\theta_i} u(q_i^*(\theta,s),s)dG_i(\theta) = r$. It implies that the optimal allocation is such that the marginal utility should be equal in every state of the world. The equivalence between the first-best and market allocation can be understood by adding a new constraint to the maximization problem called (M) and equal to $q_i(\theta,s) = d(t_i(\theta,s),\theta,s)$. In that case, the two maximization problems lead to the same outcomes.

The results of this proposition are at the core of how markets in the electricity system should work. Whenever the capacity is not constraining, prices equal the short-term marginal cost, i.e. the marginal production cost, which is null in our framework. When the capacity is binding, prices should be raised above the long-term marginal cost such that the expected prices during those periods equal the marginal investment cost. Given the maximization objective, the optimal transfer between consumers and the market designer for each s is identical to implementing the single price given by the aggregate inverse demand function at the capacity level.

With the linear model, we can express the expected consumer surplus in three intermediate cases depending on the level of investment k and the realization of the demand shock s: (i) the capacity always binds for any s, that is for any low $k \in [0, k^-]$ with $s_1(k^-) = 0$, (ii) the capacity never binds for any s, that is for any high $k \in [k^+, +\inf)$ with $s_1(k^+) = \bar{s}$ (iii) the capacity binds for some s, that is for any $k \in [k^-, k^+]]$. For the last case, we can express, for instance, the expected consumer utility under the optimal single-price allocation as follows[10]:

$$\sum_i \mu_i \left[\int_0^{s_1(k)} \underbrace{\int_{\theta_i} U(d(0,\theta,s),\theta,s) dG_i(\theta)}_{\text{off-peak utility}} dF(s) + \int_{s_1(k)}^{\bar{s}} \underbrace{\int_{\theta_i} U(d(p^k,\theta,s),\theta,s) dG_i(\theta)}_{\text{on-peak utility}} dF(s) \right]$$

$p^k = p(k, s)$ is defined for notation clarity as the aggregate demand function at the investment level $p(k, s)$.

3.2 Long-term vs. Short-term consumer surplus

The previous section showed that the optimal mechanism allocation is identical to a market under complete information when the market designer seeks to optimize the expected consumer surplus. In practice, the market designer might also pay attention to consumer surplus on a short-term horizon. In this section, we analyze the consequence of choosing first the investment level and second the consumer surplus under a market allocation. [11] We define the new objective for the market designer as follows. For clarity, we note 0, the period under which the investment decision

[10] In the rest of the paper, we do not always study all the possible cases depending on the value of k, as they do not change the theoretical results. We keep the last one as our main reference.

[11] Recent events in the European electricity markets showed that a market designer might adopt this short-term-oriented policy. Some interventions focused on reducing short-term prices via diverse interventions without considering long-term investment decisions. Therefore, this modeling approach could mirror those interventions.

is made, and period 1, the period under which the choice of mechanism is made.

$$\max_{k\geq 0} \max_{\substack{t_i(\theta,s),\\ q_i(\theta,s)}} CS(k) = \sum_i \mu_i \int_s \int_{\theta_i} U(q_i(\theta,s),\theta,s) \; - \; t_i(\theta,s)q_i(\theta,s)dG_i(\theta) \quad dF(s)$$

s.t.

$$\text{period} \quad 0: \qquad I(k) \; \leq \; \sum_i \mu_i \int_s \int_{\theta_i} t_i(\theta,s)q_i(\theta,s) \; dF(s), \tag{R}$$

$$\text{period} \quad 1: \qquad \sum_i \mu_i \int_{\theta_i} q_i(\theta,s)dG_i(\theta) \leq k, \tag{K}$$

$$q_i(\theta,s) = d(t_i(\theta,s),\theta,s), \tag{M}$$

Compared to the long-term case, we dissociate the maximization problem into two sub-problems, which are solved using backward induction. First, the market designer maximizes the consumer surplus by choosing prices and quantity such that the quantity cannot be above investment level (K), and the market designer implements a market allocation (M).[12] Then, it selects the investment level given the revenue made in the second period. The market allocation with individualized prices leads to the consumer surplus maximizing allocation under this objective. We describe the mechanism in the following proposition assuming w.l.o.g. that consumers of category 1 are of a higher type than category 2. For clarity we assume that $\bar{\theta}_1 \geq \bar{\theta}_2 \geq \underline{\theta}_1 \geq \underline{\theta}_2$. We define $s_j(k)$ with $j \in \{1,2,3\}$ such that the total quantity of consumers at a null price equals the investment level: $\sum \mu_i \int_{\theta_i} d(0,\theta,s_1(k))dG_i(\theta) = k \quad , \quad \mu_1 \int_{\underline{\theta}_1}^{\bar{\theta}_1} d(0,\theta,s_2(k))dG_1(\theta) + \mu_2 \int_{\underline{\theta}_1}^{\bar{\theta}_2} d(0,\theta,s_2(k))dG_2(\theta) = k \quad , \quad \mu_1 \int_{\underline{\theta}_2}^{\bar{\theta}_1} d(0,\theta,s_3(k))dG_1(\theta) = k$

Proposition 2. *Under short-term consumer surplus maximization, the best allocation schedule is an individualized price system. Assuming w.l.o.g. that consumer category* 1 *is of a higher type than category* 2*, then the price and quantity schedule is defined for each realization of* s *as follows.*

- *If* $s \in [0, s_1(k))$ *then* $t_1^{st}(\theta,s) = t_2^{st}(\theta,s) = 0$
- *If* $s \in [s_1(k), s_2(k))$ *then* $t_1^{st}(\theta,s) = 0$ *for all consumers* 1 *and* $t_2^{st}(\theta,s) = 0$ *for consumers* 2 *with* $\theta \in [\underline{\theta}_1, \bar{\theta}_2]$*. Define* $Q_0^1(s)$ *the total quantity for consumers having* $t_i^{st}(\theta,s) = 0$*. Then, for consumers 2 with* $\theta \in [\underline{\theta}_2, \underline{\theta}_1]$, $t_2^{st}(\theta,s)$ *is defined such that*

[12]Without the constraint, the market designer would choose a null price.

$$\int_{\underline{\theta}_2}^{\underline{\theta}_1} d(t_2^{st}(\theta, s), \theta, s)dG_2(\theta) + Q_0^1(s) = k$$

- *If* $s \in [s_2(k), s_3(k))$, *then* $t_1^{st}(\theta, s) = 0$ *for consumer* 1 *with* $\theta \in [\bar{\theta}_2, \bar{\theta}_1]$. *Define* $Q_0^2(s)$ *the total quantity for consumers having* $t^{st}(\theta, s) = 0$. *Then, for all consumer with* $\theta \in [\underline{\theta}_1, \bar{\theta}_2]$, *a unique price for same types of both category* $t^{st}(\theta, s)$ *is defined such that*

$$\sum \mu_i \int_{\underline{\theta}_1}^{\bar{\theta}_2} d(t^{st}(\theta, s), \theta, s)dG_i(\theta) + Q_0^2(s) = k$$

- *For* $s \in [s_3(k), \bar{s}]$ *then define* $t_1^{st}(\theta, s)$ *for consumers* 1 *such that*

$$\int_{\bar{\theta}_2}^{\bar{\theta}_1} d(t_1^{st}(\theta, s), \theta, s)dG_1(\theta) = k$$

The prices are set to $\theta + s$ *for all other consumers, so their demand is null. Moreover, the demand schedule is determined by the demand function* $d(t, \theta, s)$ *at the defined price schedule.*

Given the capacity constraint, the individualized price system can be understood as a rationing mechanism. Hence, the price schedule is constructed to ration consumers from the lowest type to the highest one. Given the ordering between the categories, the second price (between s_1 and s_2) consists in reducing the consumers belonging to category 2 whose type is comprised between $\underline{\theta}_2$ and $\underline{\theta}_1$.[13]. For any states between s_2 and s_3, the market designer is indifferent between rationing consumers from both categories (as soon as it does in increasing order). Finally, the last prices are defined to exclude category 2 from the market while continuing rationing the lowest consumers type from category 1 whose type is between $\bar{\theta}_2$ and $\bar{\theta}_1$. For other consumers whose prices are not defined, we assume that the market designer excludes them such that the price implies a null consumption.

Illustrative Example The short-term mechanism can be understood in a setting with a discrete set of consumers. Assuming that only two consumers with type θ_1 and θ_2 drown from the corresponding distribution, such as $\theta_1 > \theta_2$. In that case, the allocation under a short-term mechanism can be described as follows:

[13]When $\underline{\theta}_2 = \underline{\theta}_1$ this schedule is not needed

$$t_1^{st}(k, \theta_1, s) = \{0, 0, u(k, \theta_1, s)\} \quad | \quad t_2^{st}(k, \theta_1, s) = \{0, u(k - d(0, \theta_1, s), \theta_2, s), \theta_2 + s\}$$

$$q_1^{st}(k, \theta_1, s) = \{\theta_1 + s, \theta_1 + s, k\} \quad | \quad q_2^{st}(k, \theta_1, s) = \{\theta_2 + s, k - (\theta_1 + s), 0\}$$

As consumer surplus always decreases with prices, the optimal situation when there is no capacity constraint ($s \in [0, s_1(k)]$) is when $t_1^{st} = t_2^{st} = 0$. It implies a quantity equal to $d(0, \theta_i, s) = \theta_i + s$ and corresponds to the first terms in the set of prices and quantities. When the capacity starts to bind, ($s \in [s_1(k), s_3(k))$), prices must rise to ration consumers. Again, as consumer surplus always decreases with prices, this implies that the constraint will always bind and that it is never optimal to set prices such that quantity is below capacity. Note that the capacity constraint implies: $\sum_i d(t_i^{st}, \theta_i, s) = k$, under linear marginal utility this is similar as having $t_2^{st}(t_1^{st}) = 2s + k - \sum_i \theta_i - t_1^{st}$. It allows us to express consumer surplus at the capacity constraint only concerning t_1^{st}. We find that the consumer surplus exhibits a U shape. We formally prove this result in the Appendix.

The key difference between the two allocation mechanisms can be understood as follows. The market designer chooses prices and quantities given the investment level in the short-term consumer mechanism allocation. So prices, when understood from a market perspective, are set only to ration and reduce the quantity. In that case, maximizing the consumer surplus always implies a form of discrimination against the lowest consumers as soon as consumers have heterogeneity and a capacity constraint. The revenue from prices acts only as (residual) transfers to cover fixed costs and choose the investment level. Given the investment level, the best allocation follows this personalized price system.

On the other hand, under the first-best consumer allocation, the market designer cares about the long-term decisions of choosing the optimal level of investment. In that perspective, prices are chosen to generate revenues and efficiently ration consumers. When increasing the (marginal) level of investment, the market designer internalizes the opposite effect of sustaining a marginal investment cost and increasing the available quantity for consumers. This increase in quantity allows both an increase in consumer surplus and prices to cover the fixed costs. We compare the quantity and price schedule in Figure 1 for the discrete case.[14] The red curves represent the higher consumer 1, and the blue curves represent the lower consumer 2. For the (expected) quantity schedule, the black line represents the total quantity, which is by definition equal between the two mechanisms (the plateau is equal to k). They only differ with respect to the allocation between consumer. The first-best mechanism is shown in the first plot with dashed curves. Whenever the capacity is constraining, the central idea of the short-term mechanism is to exclude the lowest type of consumers gradually.[15]. The second plot shows this via individualized (expected) prices. We have excluded prices used to exclude consumers for clarity. Each increasing price is assigned to a specific group of consumers. The first blue line is assigned to consumers of category two between with $\theta \in [\underline{\theta}_2, \underline{\theta}_1]$. The second mixed-colored curve is applied to consumers of both categories with $\theta \in [\underline{\theta}_1, \bar{\theta}_2]$. Finally, when the demand is too high, it is always optimal to exclude all consumers with types lower than $\bar{\theta}_1$ and set a price, given by the red curve, for consumers above such that capacity is binding.

To compare this allocation schedule with the first best solution, we provide in Corollary 1 a description of the optimal quantity (rationing) that maximizes the first-best surplus. Instead of setting prices and letting the quantities adjust, the market designer could select quantities and impose a transfer on each consumer. Such a policy can be implemented when prices do not emerge due to price regulation, such as in a price cap case (see, for example, Leautier (2018); Zöttl (2011)). When the capacity is sufficiently binding such that the price cap constrains the price on the market, the market designer needs to set a quantity rationing policy. The corollary describes this policy under the linear model for ease of exposition. We denote θ_i^{av} as the average type of a category i.

Corollary 1. *Under linear marginal utility assumption and uniform distribution.*

(i) When the capacity starts to bind, then the optimal rationing is independent of s:

$$\alpha_i^*(s,k) = \mu_i + \frac{\mu_i \mu_j (\theta_i^{av} - \theta_j^{av})}{k} \qquad \forall s \in (s_1(k), \bar{s}]$$

[14]The shape of the graphs would be the same when plotting the aggregate quantity for each category.

[15]Note that even if the market designer does not differentiate between the category for consumers of type $\theta \in [\underline{\theta}_1, \bar{\theta}_2]$, the existence of consumers for category 1 with higher type and priced at 0 implies the increasing red curve between s_2 and s_3.

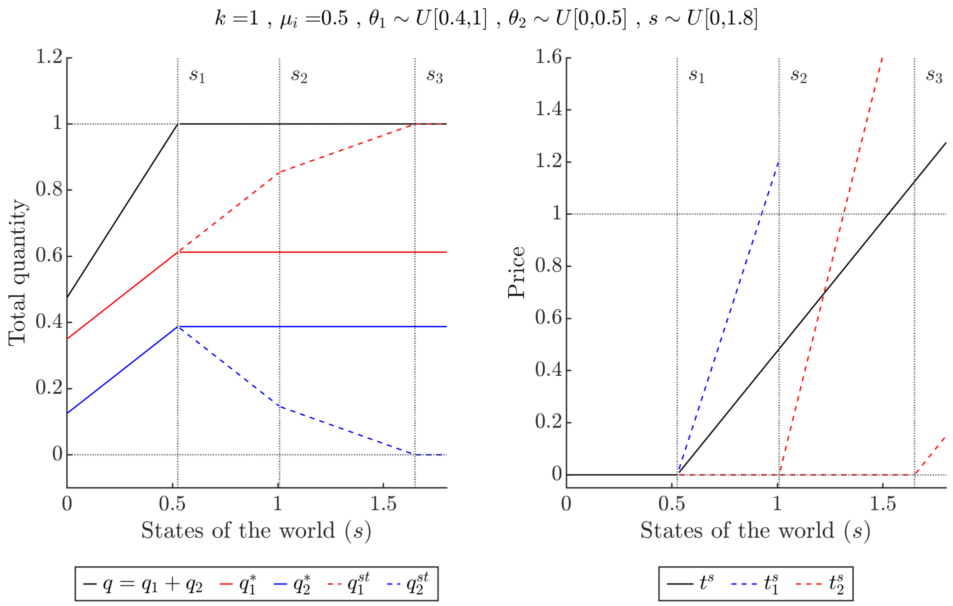


Figure 1: Quantity and price schedules under first-best and short-term consumer surplus mechanism. (The price schedules on the plot represent only prices associated with a positive quantity.)

(ii) Assuming that consumer 1 is the higher type ($\theta_1 > \theta_2$) Under the individualized price system, the rationing strategy consists of rationing first the lowest type and then the highest type. If α_i^{st} is the rationing policy under the individualized price system, then we have:

$$\alpha_1^{st}(s,k) = \mu_1 \frac{s + \theta_1^{av}}{k} \quad \textit{and} \quad \alpha_2^{st}(s,k) = 1 - \alpha_1^{st}(s,k) \quad \forall s \in (s_1(k), s_2(k)]$$
$$\alpha_0^{st}(s,k) = \frac{1}{k}\left(\frac{\bar{\theta}_1 - \bar{\theta}_2}{\bar{\theta}_1 - \underline{\theta}_1}\right)\left(s + \frac{\bar{\theta}_1 + \bar{\theta}_2}{2}\right) \quad \textit{and} \quad \alpha_{-0}^{st}(s,k) = 1 - \alpha_0^{st}(s,k) \quad \forall s \in (s_2(k), s_3(k)]$$

With α_0^{st} the ratio for consumers from category 1 with $\theta > \bar{\theta}_2$ receiving a null price, and $\alpha_{-0}^{st}(s,k)$ the ratio for consumer belonging to both category receiving positive price.

We turn now to the choice of k. Under this framework, we dissociate two notions: The short-term consumer surplus without lump-sum transfers and the long-term consumer surplus with lump-sum transfers. The first notion describes the consumer surplus at the level of investment without taking into account the revenue constraint of period 0. Recall that the backward induction implies that we find prices and quantities given a value of k. Still, the sequence of decisions does not mean that the revenue generated in the second period covers the corresponding investment cost of the first period. Hence, we define this short-term consumer surplus as follows: $CS^{cs}(k) = U(k) - R(k)$, with U the expected aggregate utility, which depends solely on the quantity schedule, and $R(k)$ the expected aggregate revenue, which depends on the quantity and price schedule. On the other hand, we can also define a feasible long-term consumer surplus that considers only the utility and the investment costs, which can be expressed as $CS^{lt}(k) = U(k) - I(k)$. In that case, the market designer can always implement a non-distortive lump-sum transfer T (in any period) such that investment costs are covered: $T = \max(0, I(k) - R(k))$. In the proof relative to the first-best allocation k^*, we show that at the first-best investment level, we always have $CS^{lt}(k^*) = CS^{st}(k^*)$. Lemma 1 describes the level of investment that maximizes the long-term consumer surplus given the short-term allocation schedule.

Lemma 1. *The long-term consumer surplus under a short-term maximizing regime can exhibit non-concavity. There is at least one local maximum and at most two local maxima.*

The value of each investment derives from the first-order condition of the consumer surplus under the price and quantity schedule described in Proposition 2. It significantly differs from the first-order condition of Proposition 1. We express the condition as follows:

$$r = \int_{s_1(k)}^{s_2(k)} \overbrace{\mu_2 \int_{\underline{\theta}_2}^{\underline{\theta}_1} t_2^{st}(\theta, s)\, \frac{\partial t_2^{st}(\theta, s)}{\partial k} dG_2(\theta)}^{\text{marg. utility from category 2}} \, dF(s) + \int_{s_2(k)}^{s_3(k)} \overbrace{\sum \mu_i \int_{\underline{\theta}_1}^{\bar{\theta}_2} t_i^{st}(\theta, s)\, \frac{\partial t_i^{st}(\theta, s)}{\partial k} dG_i(\theta)\, dF(s)}^{\text{marg. utility from cat. 1 and 2}}$$
$$\int_{s_3(k)}^{\bar{s}} \underbrace{\mu_1 \int_{\bar{\theta}_2}^{\bar{\theta}_1} t_1^{st}(\theta, s)\, \frac{\partial t_1^{st}(\theta, s)}{\partial k} dG_1(\theta)\, dF(s)}_{\text{marg. utility from category 1}}$$

When choosing the investment level, the market designer needs to weigh the effect of a change of k on the positive prices that generate positive quantities for consumers (i.e., It excludes consumers that are not rationed (null prices) or fully rationed (null quantity). The first-order conditions capture those effects. The existence of two possible maxima stems from the boundary conditions due to the capacity constraints. Indeed, there must be a coherence between a maximum and the values of all the $s_j(k)$. For instance, if k solve the previous equation, it must be that $s_1(k) > 0$ and $s_3(k) < \bar{s}$. We therefore study the (six) possible maxima for the different values of the s_j. We find that there are only four possible values.

The first one corresponds to the case where the (maximizing) capacity level is very high, such as it never leads to rationing of the highest type of consumers of both categories (above $\underline{\theta}_1$), that is $s_2(k) = \bar{s}$, implying the second and third term in the first order cancels out. The second one corresponds to the case where the capacity level is moderately high, such as it never leads to rationing of the highest type of consumers from category 1 (above $\bar{\theta}_2$), that is $s_3(k) = \bar{s}$, implying the third term in the first order cancels out. Finally, the two last equilibria are mutually exclusive. They can coexist with the other maxima but never with each other. They respectively imply a lower level of capacity such that (i) the lowest type consumers of category 1 are always excluded ($s_2(k) = 0$, first term cancels out), or (ii) the lowest consumers of both categories are always excluded ($s_3(k) = 0$, first and second term cancel out)). On the contrary, the value is always unique for the first-best maximizing investment level, such that the capacity is either always binding or binding with positive probability. The result in Corrolary 2 shows the short-term consumer surplus is higher under the short-term allocation, and the long-term consumer surplus is always higher under the first-best allocation.

Corollary 2. *The individualized price system that maximizes the short-term consumer surplus always leads to a lower long-term consumer surplus.*

We illustrate the results in Figure 2. The first panel shows the short-term consumer surplus $U(k)-R(k)$, and the second shows the long-term surplus, $U(k)-I(k)$, for the respective short-term and first-best allocations. The solid curves represent the first-best allocation, while the dashed curves represent the short-term allocation. As we can see, the associated short-term consumer surplus is always higher under short-term allocation. However, when considering the necessary lump-sum transfer to fully cover the fixed costs, the long-term consumer surplus is consistently lower. Note that the consumer surplus under the single-price system is concave, while under the

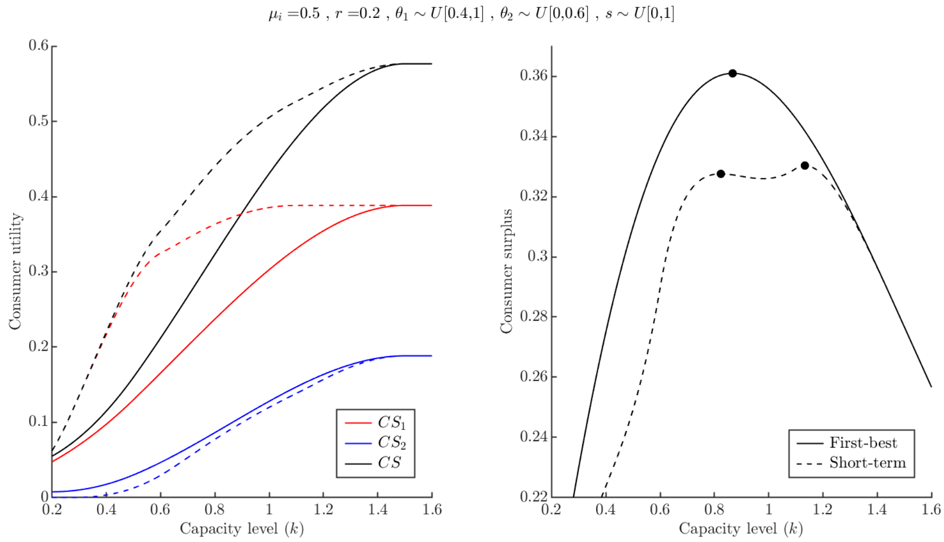


Figure 2: Expected consumer utility and surplus under the short-term (ST) and first-best (LT) allocation.

multi-price system, it exhibits some non-concavity. The dots represents the investment level that maximizes the consumer surplus.

4 Incomplete Information - Fixed price

In this section, we study the second regime under which the market designer has to choose the best allocation, given the following assumptions:

A4.1 The consumer's type is unknown to the market designer.

A4.2 The market designer cannot extract any information from consumers.

A4.3 The price schedule offered by the market designers is constrained to a unique price for every state of the world.

A4.4 Given a set of prices p_i^r

- The market designer implements a market allocation until the capacity is not binding.

- The market designer implements a rationing policy when the capacity is binding.

The first assumption implies an incomplete information framework. The second assumption means that when offering the best allocation, the market designer is not subject to incentive compatibility and participation constraints. The third assumption provides a more realistic approach between the complete first-best allocation and the incomplete information case with a mechanism design setup described in the next section. Indeed, it approximates the actual management of essential goods such as retail electricity or public transport, where a market designer is constrained in its short-term allocation while having imperfect information on its consumers' type. To capture the effect of incomplete information, the market designer must be constrained when implementing the mechanism. It can come (i) from the quantity allocation - that is, consumers do not maximize their utility[16] - (ii) or from the proposed monetary transfer. In the last case, the price schedule is incomplete because it is not optimal for every value of s. It distorts the quantity consumers demand, even though it maximizes their utility.[17] In this paper, we take the second interpretation: We assume that the market designer can only choose a single price for every state s. From a policy perspective, this is similar to a market designer offering a fixed-price contract to consumers.[18] In this section, we study two cases: (i) the market designer does not discriminate between different categories, and the offered price is unique for every consumer; (ii) the market designer can discriminate between different categories, and he offers a price for each category.[19] The first case allows us to focus our analysis on highlighting the trade-off the market designer faces when collecting revenue for the investment cost. In contrast, the second case highlights the distributive effect between consumers of different groups. This modeling approach underlines a market designer's trade-off concerning the uncertainty of the consumers' types, even without strategic inefficiencies.

[16]See for instance Martimort and Stole (2020) which studies the optimal monopoly nonlinear pricing in an incomplete information setting where consumers wrongly equal marginal benefit with average price. For empirical evidence, see Ito (2014).

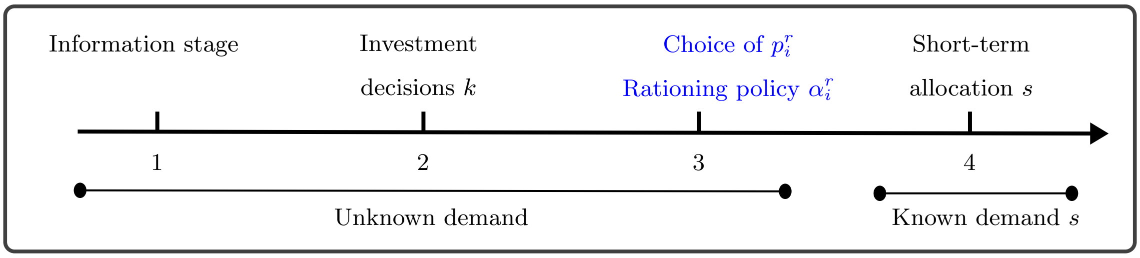
[17]In current practice, political and technological constraints imply that the market designer (or any retail agent) can only propose a finite number of schedules. See, for instance, Astier (2021) for theoretical and empirical implications for consumer surplus of allocation incompleteness.

[18]An important caveat is that when considering the case with a discrete number of consumers, a first-best allocation does not replicate the optimal allocation in incomplete information. That is, even with a complete set of prices, private information leads to inefficiencies. We provide in the Appendix the results with the discrete framework.

[19]Political or technical reasons can prevent the market designer from distinguishing between different categories. It will be apparent in the section that even though the outcome of discrimination is welfare-enhancing, it is not always Pareto-improving for every category.

The core idea of the model is that, without any information, a market designer has to choose a price p_i^r independent of the world's states s. However, one issue remains unanswered, which is how quantities are allocated within the framework. While the choice of a unique price forever consumer does not need any information concerning consumer type, the market designer still has to choose how to allocate the goods between consumers. In this section, we make the following assumption.

Finally, the last assumption can be understood as follows. When the capacity is not binding, quantity should be adjusted at a price p_i^r. When the capacity is binding, a random allocation is implemented because of incomplete information.[20] Following the framework description, we modify the actions taken by the market designer. The allocation proposal comprises (i) choosing p_i^r and (ii) defining the rationing policy α_i^r described below (that is, the share of capacity each consumer is receiving when capacity is binding). We provide below an updated figure considering the decisions the market designer has to make within this framework.



4.1 Single price policy

We start by assuming that the market designer is constrained by setting a unique price for each category, so we drop the index and assume that p^r is the price chosen by the market designer. The incomplete information set-up in this section has an important implication regarding quantity

[20]We show in the Appendix that this is not the only mechanism the market designer can implement. Indeed, we could assume a mechanism design approach where the market designer can also choose $q_i(\theta, s)$ when the capacity is not binding. In that case, the market designer produces the total quantity as in the first best and allocates it randomly to the consumers. We show in the Appendix that there is no clear ranking between the two mechanisms regarding consumer surplus, which depends on the model parameters. The market allocation minimizes the cross-subsidies between consumers, but when prices determine the quantity when the capacity is not binding, it induces a negative price effect. This section aims to illustrate current practice, and the two mechanisms do not fundamentally differ. Therefore, we choose to represent the market allocation approach.

allocation. Indeed, combining a single-price policy and imperfect knowledge implies that some inefficient rationing should be expected in the market. To see this, recall that $d(p^r, \theta, s)$ is the quantity a consumer asks given the price p^r. Let's define $s_1^r(k)$ the first states of the world when the capacity is binding when the price is p^r that is:

$$s_1^r(k, p^r) = \left\{ s : \sum \mu_i \int_{\theta_i} d(p^r, \theta, s) dG_i(\theta) \right\}$$

For any $\leq s_1^r(k, p^r)$, the price is such that capacity is not binding. That is, the quantity asked by each consumer is short-term. In that case, there is no need for rationing. Note, however, that when $p^r > 0$, the model does imply an inefficiency similar to the effect of market power. Due to the price being higher to marginal, it prevents some Pareto-improving trade from happening. For any $s \geq s_1^r(k)$, capacity is binding, and the total quantity each consumer asked is above the available capacity. To avoid market failure, the market designer needs to reallocate quantity between consumers. However, we assumed that he does not observe consumer type. Without any possibility of extracting information, the only option for the market designer is to allocate equally across consumers a quantity equal to the investment level. Therefore, the individual quantity k and the expected quantity for each category is $\mu_i k$. We illustrate the implications by defining the expected utility under the single-price policy with incomplete information.

$$\sum \mu_i \int_0^{s_1^r(k)} \overbrace{\int_{\theta_i} U(d(p^r, \theta, s), \theta, s) dG_i(\theta)}^{\text{off-peak utility}} dF(s) + \sum \mu_i \int_{s_1^r(k)}^{\bar{s}} \overbrace{\int_{\theta_i} U(k, \theta, s) dG_i(\theta)}^{\text{on-peak utility}} dF(s)$$

We turn now to determining the best single-price policy given the framework. Compared to the previous analysis, the optimal price p^r depends not only on the first-best condition but on the budget constraint. The optimal price p^r is given by the following lemma.

Lemma 2. *If it exists, the optimal value $p^r(k, p^r)$ satisfies the net revenue condition $R^k(k, p^r) = 0$ with:*

$$R^k(k,p^r) := \underbrace{p^r\left(\sum \mu_i \int_0^{s_1^r(k,p^r)} \int_{\theta_i} d(p^r,\theta,s)dG_i(\theta)dF(s) + \int_{s_1^r(k,p^r)}^{\bar{s}} kdF(s)\right)}_{Expected\ revenue} - I(k)$$

This result is close to what can be found in the literature on peak pricing with price-inelastic consumers. In that case, the optimal price is simply the average cost. Under our framework, the optimal single price is different due to the price response of the consumers during off-peak periods and to the inefficient rationing occurring in the on-peak periods. Next, we provide in Proposition 3 the relation between the investment level and the optimal single-price

Proposition 3. *If an optimal single-price $p^r(k)$ exists, it increases in k.*

The intuitions of the lemma and proposition are closely related. When choosing the price p^r, the market designer must trade off opposite effects. Indeed, increasing p^r lowers quantity during off-peak. Hence, the revenue effect during off-peak is ambiguous. For on-peak periods, the revenue effect is always positive as the expected quantity is k and is not affected by a change of p^r. Note that the revenue is concave in p^r, meaning that the second-order effects are negative, limiting the market designer's ability to extract revenue from consumers.[21] Those effects can be shown by expressing the first derivative of the expected net revenue:

$$\frac{\partial R^r(k)}{\partial p^r} = \underbrace{\int_0^{s_1^r(k)} \overbrace{d_p p^r}^{-} + \sum_i \mu_i \int_\theta \overbrace{d(p^r,\theta,s)}^{+} dG_i(\theta)dF(s)}_{\text{off-peak marg. revenue}} + \underbrace{\int_{s_1^r(k)}^{\bar{s}} \overbrace{k}^{+}\ dF(s)}_{\text{on-peak marg. revenue}}$$

With $d_p = \frac{\partial d}{\partial t}$ the derivative of the demand function with respect to prices. Calculation shows that $\frac{\partial s_1^r(k)}{\partial p^r} > 0$, as a higher price, means consumers decrease their consumption, and the capacity is binding less often. Next, we show how k modifies the marginal effect of p^r. Again, the market designer faces trade-offs. Increasing k increases the revenue that can be collected from on-peak

[21]Increasing p^r lowers the occurrence of on-peak periods, and the revenue during off-peak is concave due to the linearity assumption of the marginal utility.

periods but also decreases the occurrence of a binding capacity. Therefore, at the margin, it reinforces the negative effect of the price on the quantity during off-periods and increases the marginal revenue during on-peak periods. Those direct effects are represented in the following expression of the cross-derivative of the expected revenue.

$$\frac{\partial^2 R^r(k)}{\partial p^r \partial k} = \overbrace{d_p p^r f(s_1^r(k)) \frac{\partial s_1^r(k)}{\partial k}}^{-} + \overbrace{\int_{s_1^r(k)}^{\bar{s}} 1 dF(s)}^{+}$$

So if $\frac{\partial R^r}{\partial k} > 0$, then at least one quantity is negative, which cannot happen.

Then we use the fact that $\frac{\partial^2 R}{\partial p^r 0^2} < 0$ is concave. At $p_0^r = r \mid k = \tilde{k}$: $s_1 = 0$. We have $\frac{\partial R}{\partial p^r 0} > 0$, implying that the revenue is increasing at the limit in p_0^r. If the function is concave, there could be at least two potential values for the optimal value of p_0^r. However, note that consumer surplus is always decreasing in prices; therefore, a lower price is always optimal compared to a higher price. So, the optimal value corresponds to the first increasing part.

Proposition 3 shows that expanding the capacity level always leads to the positive (revenue) effect dominating the adverse (price) effects. That is, the effect of the increase in the revenue collected during on-peak periods offsets the compound negative impact of a price increase that (may) lower the revenue during off-peaks and reduces the occurrence of on-peak periods.[22]

4.2 Category-price policy

We extend the previous findings by assuming that the market designer knows the category each consumer belongs to (but his type remains private information). Therefore, the market designer can imperfectly discriminate between consumers and implement a category-based price to finance its investment cost. We start by defining the new rationing policy under this framework. This stage boils down to allocating the capacity k in the first step between the two categories and the second step, randomly for each consumer within each category. This procedure implies that the market

[22] From a policy perspective, the market designer never wants to lower price so as the increase the consumption during off-peak.

designer allocates the same expected quantity to each category under the first best allocation (even though the within-category allocation remains inefficient). The problem is solved as follows. Let the quantity for category i be q_i; then, when the capacity is constraining, we must have for every state of the world: $\sum_i \mu_i q_i = k$, implying that the relation between the quantity is equal to $q_i(q_j) = \frac{k-\mu_j q_j}{\mu_i}$. Then, we maximize the short-term expected utility: $\sum_i \mu_i \int_{\theta_i} U(q_i, \theta, s) dG_i(\theta) dF(s)$, given the previous relation. Solving using the first-order condition leads to a capacity share for each consumer belonging to a group i of $1 + \frac{\mu_j(\theta_i^{av} - \theta_j^{av})}{k}$, which in for all consumers implies the same allocation as in Corollary 1.

Given this optimal rationing policy, we now define the new problem the market designer faces:

$$\max_{\substack{k, \\ p_i^r \to \mathbb{R}^+}} \quad CS^r(k, p_1^r, p_2^r) = \sum_i \mu_i \int_s \int_{\theta_i} U(q_i^r(\theta, s), \theta, s) - p_i^r q_i^r(\theta, s) dG_i(\theta) \quad dF(s)$$

$$\text{s.t.} \quad I(k) \leq \sum_i \mu_i \int_s \int_{\theta_i} p_i^r q_i^r(\theta, s) \; dF(s), \tag{R}$$

We drop the capacity constraint as the rationing policy defines it implicitly. To see this, we can redefine the state of the world when the capacity starts binding:

$$s_1^r(k, p_1^r, p_2^r) = \left\{ s : \sum \mu_i \int_{\theta_i} d(p_i^r, \theta, s) dG_i(\theta) \right\}$$

Then, the quantity $q_i^r(\theta, s)$ allocated in the market for each consumer of catagegory i is equal to $d(p_i^r, \theta, s)$ when $s < s_1^r(k, p_1^r, p_2^r)$ and $\alpha_i^r k = k + \mu_j(\theta_i^{av} - \theta_j^{av})$ when $s \geq s_1^r(k, p_1^r, p_2^r)$. Note that while the total quantity for each category is identical under the first-best and this framework, the total utility does differ. It implies a very similar delta in terms of utility as the equation in the previous section with a single price policy.

First, let's note $\mathcal{L}^r$, the Lagrangian associated with the market designer program such that

$$\mathcal{L}^r(k, p_1^r, p_2^r, \gamma^r) = CS^r(k, p_1^r, p_2^r) + \gamma^r R^r(k, p_1^r, p_2^r)$$

With CS^r the aggregate expected consumer surplus defined as the sum of consumers' utility net of monetary transfers, γ^r the lagrangian multiplier associated with the budget constraint, and $R^r(k)$ the budget constraint (expected revenue net of investment costs). Then, using the Envelop Theorem, we can express the derivative of an optimal price with respect to k as follows:

$$[] \frac{\partial p_i^r(k)}{\partial k} = \overbrace{\frac{-\mathcal{L}_{jj}}{H^r}}^{\leq 0} \overbrace{\Bigg(CS_{ik} + \rho CS_{jk} + (R_i - \rho R_j) \sum_i (CS_{ik} - \rho CS_{jk}) \frac{R_i}{L_{jj} bH^r}}^{\text{effect of } k \text{ on CS with holding R fixed}} \tag{1}$$

$$\underbrace{+ \gamma^r \Bigg(R_{ik} - \rho R_{jk} + (R_i - \rho R_j) \sum_i (R_{ik} - \rho R_{jk}) \frac{R_i}{L_{jj} bH^r} \Bigg) - R_k \frac{H^r}{bH^r} \Bigg)}_{\text{effect of } k \text{ on } R} \tag{2}$$

With $\rho_c = \frac{CS_{ij}}{L_{jj}}$, $\rho_r = \gamma^r \frac{R_{ij}}{L_{jj}}$ and $\rho = \rho_r + \rho_c$. $H^r = \mathcal{L}_{11}\mathcal{L}_{22} - \mathcal{L}_{12}\mathcal{L}_{21}$ being the determinant of the Hessian matrix of the Lagrangian. $bH^r = \mathcal{L}_{ij}R_iR_j - \mathcal{L}_{jj}R_i^2 - \mathcal{L}_{ii}R_j^2 + \mathcal{L}_{ji}R_iR_j$ being the determinant of the bordered Hessian matrix of the lagrangian. Each variable's index is associated with the corresponding derivative. For instance, CS_{ik} reads as the cross derivative between the price of category i with respect to the investment level. It measures the (marginal) change of the marginal effect of price p_i^r on the consumer surplus. We summarize the findings in the following proposition.

Proposition 4. *Suppose that category 1 consumers are of higher types than category 2 consumers and that the average private demand for each category is sufficiently different, then:*

- $p_1^r(k)$ *is increasing with* k
- $p_2^r(k)$ *is first decreasing, then increasing with* k.

Morevover, given a threshold value of k^r*, any value below* k^r *implies:* $p_2^r(k) > p_1^r(k)$*, and any value above* k^r *implies:* $p_1^r(k) > p_2^r(k)$*. At* k^r*, optimal prices equal the optimal single price, such that* $p^r(k) = p_i^r(k)$*.*

Figure 3 illustrates the results. The red curve shows $p_1^r(k)$, the blue curve shows $p_2^r(k)$, and the black dashed curve shows the optimal single price $p^r(k)$ found in the previous section. Following the proposition, we observe that the blue curves corresponding to the group with a lower expected demand exhibit a non-monotonic relationship with the level of investment such that it decreases for low values of k and then increases again following a similar behavior to the optimal price for the higher category of consumers represented in the red curve.

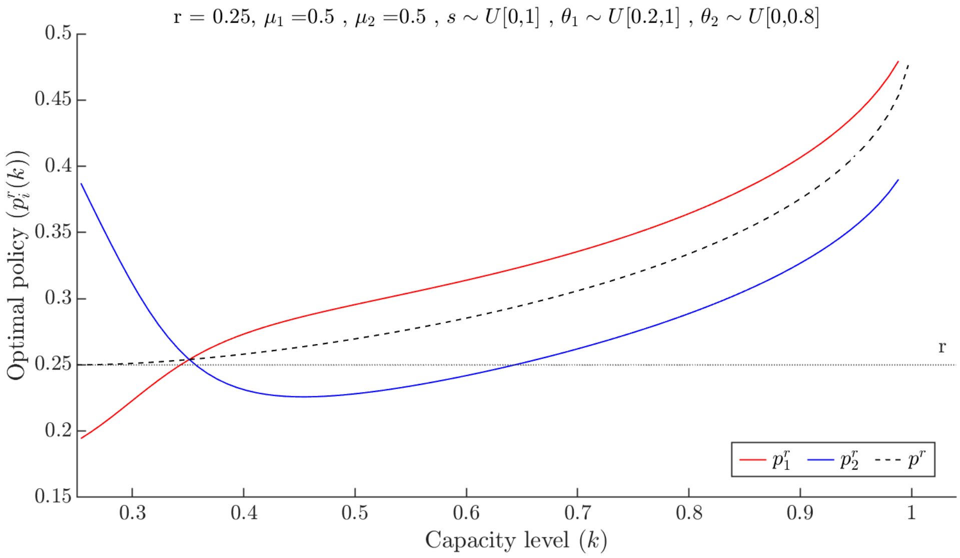


Figure 3: Optimal prices under the category-price policy with respect to the investment

The proof of such behavior of the optimal prices can be understood by distinguishing the first-order and the second-order effects of prices and level of investment on (a) the aggregate consumer surplus and (b) the budget constraint as illustrated with Equation 1. The non-monotonicities of prices are the result of a consumer surplus effect dominating first a revenue effect for low values of k, then the revenue effect dominating the consumer for higher values of k.

From a consumer surplus perspective, and as illustrated in Proposition 2, the market designer prefers (i) discrimination and (ii) favoring the consumers from the category of the highest type. Preferring discrimination implies that consumer surplus exhibits a U shape form. Due to the asymmetry between the consumers, the function is skewed towards lower types. This is because higher types bring relatively more surplus than the smaller types. On the other hand, when studying the variation of the expected revenue with respect to prices, the market designer faces the opposite effect. Increasing p_1^r generates more revenue as they consume, on average, more.

To see the fundamental tension between revenue and consumer surplus, we study the difference between the derivative of the consumer surplus with respect to p_1^r and p_2^r. It is negative and the opposite of the difference between the derivative of the expected revenue. In absolute terms, they are both equal to $\frac{1}{2}(\theta_1^{av} - \theta_2^{av}) > 0$. Note that we also have the opposite effect between consumer surplus and revenue from the price level. Namely, the consumer surplus is higher when prices are low, and revenue is increasing (although not everywhere) with prices. Therefore, the net effect on the optimal values of p_1^r and p_2^r ultimately depends on which effect dominates.

The next step for understanding the results lies in how those opposite effects change with the level of investment. As shown in the previous section, our framework implies that a change of k does affect both revenue and the consumer surplus, which is captured via the direct effect on prices needed for financing this investment and the change of occurrences between off-peak and on-peak periods. First, the level of investment induces a positive first derivative of the consumer surplus and a negative second derivative. That is, increasing k always increases the surplus, but for a higher level of investment, the positive impact is relatively smaller. On the other hand, an analysis of how the budget constraint behaves shows a convex effect with respect to k. It implies that an increase of k leads to the iso-revenue associated with the constraint shifting at an increasing rate.[23]. The switch between the decreasing and increasing parts is associated with the consumer surplus effect dominating the budget effect first. As k increases, the respective concavity and convexity of the functions lead to the budget effect dominating the surplus effect: this explains the increasing parts and the ranking between the two prices on the right part of Figure 3.

Now, let us turn towards the left part of Figure 3. For low values of k, the budget effect is small. Now, for the sake of clarity, let's assume the budget is fixed. Note that the iso-revenue curve is convex and skewed towards the prices of consumers 1 due to the preference for symmetry of the budget constraint. That is, when choosing between two pairs of prices, the market designer needs to balance two opposite effects. The market designer faces a high degree of discrimination for relatively different prices.

[23]The pure revenue effect of increasing prices is found in the previous section.

On the other hand, for prices with closer value, the market designer meets, on average, lower prices. As consumers' surplus benefits from a higher degree of discrimination and lower prices, choosing the optimal pair of prices depends on the net opposite effects. As k increases, the discrimination effect is lower, which induces the market designer to choose relatively less different prices. For the final increasing parts of the optimal prices, the budget constraint and the preference for lower prices imply that both prices are increasing, and the price for the higher type is above the price for smaller types. We illustrate this tension in Figure 4. We represent the contour map of the aggregate consumer surplus with respect to p_1^r and p_2^r for two values of k. The convex curve represents a fixed revenue constraint that we assume is independent of k. As k increases, we show that the indifference curves tend to be more convex. It is this shift in the shape of the consumer surplus that implies a decrease in the optimal price, meaning that the gain in lower prices is higher than the gain from discrimination.

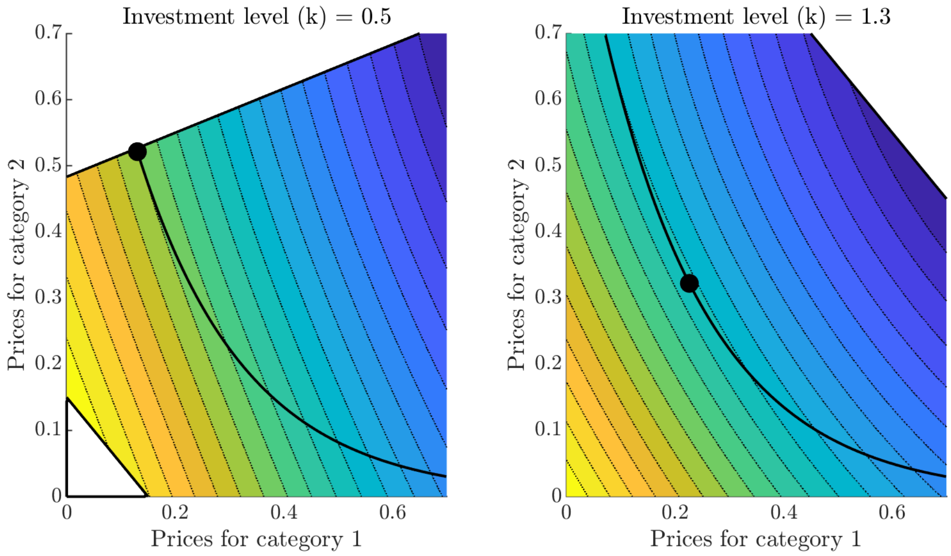


Figure 4: Illustration of the decreasing part of p_2^r with respect to k

We conclude this section by analyzing the implication of the optimal policy when choosing the level of investment to maximize the consumer surplus. We define the first-order conditions using the Envelop Theorem for constrained optimization. That is, it is sufficient to derive the derivative of the Lagrangian with respect to k: $\frac{\partial \mathcal{L}^r}{\partial k} = \frac{\partial CS^r}{\partial k} + \gamma^r \frac{\partial R^r}{\partial k}$. We start with the consumer surplus:

$$\frac{\partial s_1^r(k)}{\partial k}\overbrace{\sum \mu_i \int_{\theta_i}\int_{\alpha_i^r k}^{d(p_i^r,\theta,s)} u(q,\theta,s)dqdG_i(\theta)}^{\Delta \text{ in quantity btw. off-peak and on-peak}}|_{s=s_1^r(k)} + \int_{s_1^r(k)}^{\bar{s}} \overbrace{\sum \mu_i \int_{\theta_i} u(\alpha_i^r k,\theta,s) - p_i^r \; dG_i(\theta)}^{+ \text{ in on-peak cons. surplus}} dF(s)$$

The second term is similar to the complete information benchmark. It represents the gain in consumer surplus during on-peak as capacity expands. Note that the gain in surplus does not depend on the price as the quantity is randomly assigned to each consumer in each category due to imperfect information. The first term stands for the change at the margin of quantities for each consumer. Under complete information, the quantities allocation is continuous in s. However, due to incomplete knowledge, the market designer creates a discontinuity in the allocation when capacity starts binding, which implies that the value at $s_1^r(k)$ does not cancel out. For the revenue, the derivative can be expressed as follow:

$$\frac{\partial s_1^r(k)}{\partial k}\overbrace{\sum \mu_i \int_{\theta_i} p_i^r\Big(d(p_i^r,\theta,s) - \alpha_i^r k\Big) dG_i(\theta)}^{\Delta \text{ in quantity btw. off-peak and on-peak}}|_{s=s_1^r(k)} + \int_{s_1^r(k)}^{\bar{s}} \overbrace{\sum \mu_i p_i^r}^{+ \text{ in on-peak rev.}} \quad dF(s) - r$$

The first term is similar and originates from discontinuity. The second term comes from the increase in available quantity during on-peak. As expected, and similarly to the first-best investment level, the sign of the first-order condition is ambiguous as it has positive and negative effects of an increase in k. For instance, it raises investment costs, but it also raises available revenue. Calculations shows that there exists a level of investment that maximizes the expected aggregate consumer surplus, as the consumer surplus and the revenue are concave in k. We now compare the outcomes in terms of welfare given the optimal policy for the single-price and category-based prices at the aggregate surplus level and also from the surplus for each category. Figure **??** illustrates our findings when we vary either the investment costs or the degree of heterogeneity between the consumers. Namely, we consider that category 1 is of higher types than category 2, and we increase the lower type from category 1: $\underline{\theta}_1$. In the first panel, we show the relation between the optimal investment level given the (i) first-best complete information case (ii) the second-best with single price policy (iii) the second-best with category-based policy for different levels of r. The second panel shows the aggregate consumer surplus at the corresponding investment level. The third panel is the consumer surplus for category 1, and the fourth is the consumer surplus for category 2. The last three figures are shown with different values of $\underline{\theta}_1$

The figure suggests multiple implications for our comparison. First, we find that the imperfect information and the price constraints could imply a lower or higher level of investment at the second best, depending on the model parameters. This result can be compared to other analyses for the cause of underinvestment, such that a price cap implies a lower investment level, and the public good nature of the investment leads to a higher second-best. Then, we find that, on an aggregate basis, allowing the market designer to discriminate between categories of consumers increases the aggregate consumer surplus. Indeed, as we have shown earlier, maximizing consumer surplus or revenue implies a form of preference for discrimination. Therefore, our model provides a basis for the positive effect of discriminating consumers based on their category due to the absence of information and the constraints on prices. This has to be compared to the first-best mechanism under which the market designer does not discriminate.

5 Incomplete Information - Mechanism Design

We extend our framework by allowing the market designer to choose an allocation mechanism such that (i) consumers behave truthfully and (ii) the market designer is not constrained in its choice of prices given the realization of s. The two assumptions combined allow him to bypass the spot market allocation because truthful behavior implies that he can also set quantities for each consumer. In other words, the market designer can now offer a complete set of prices and quantities such that the schedule depends on each consumer, for every state of the world s, and every type θ. The following figure summarizes the new action set for the market designer. As we will show, the incentive compatibility constraint pins down the optimal monetary transfer t_i^*, leaving the market designer only with the quantity choice.

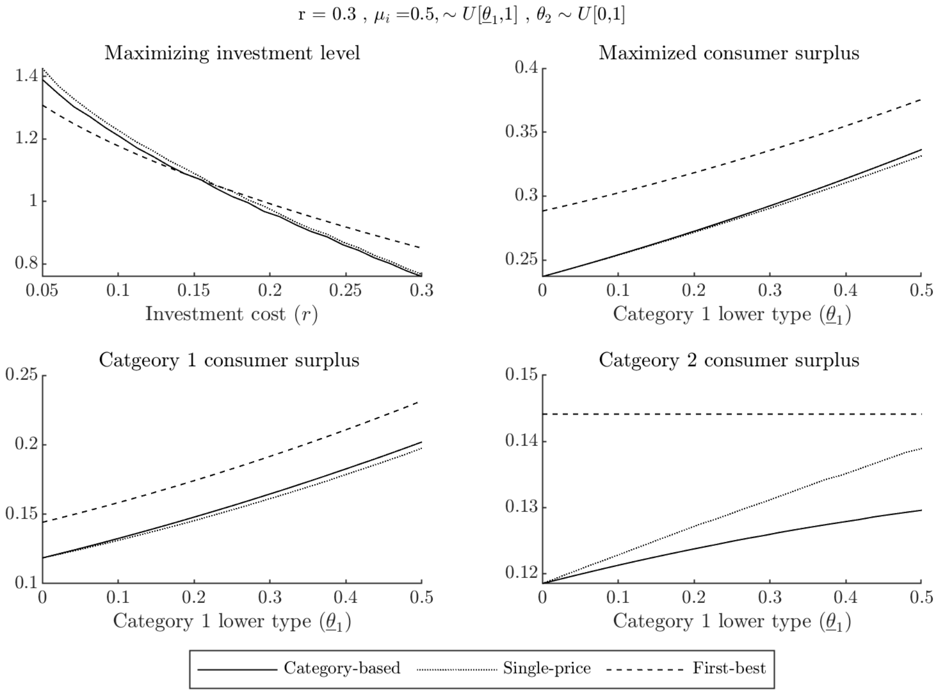
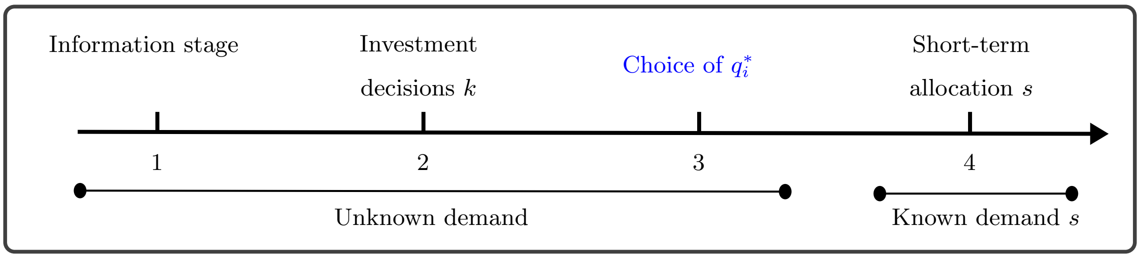


Figure 5: First panel: Optimal investment level. Second panel: Optimal aggregate consumer surplus. Third panel: Optimal category 1 consumer surplus. Fourth panel: Optimal category 2 consumer surplus.



To induce true reporting from consumers, the market designer needs to require the following:

$$\theta_i = \arg\max_{\tilde{\theta}} \int_s U(q_i(\hat{\theta}, s), \theta, s) - t_i(\hat{\theta}, s)q_i(\hat{\theta}, s)\, dF(s) \qquad \text{(IR)}$$

While the participation of every consumer implies that:

$$0 \leq \int_s U(q_i(\theta,s),\theta,s) \ - \ t_i(\theta,s)q_i(\theta,s)\ dF(s) \tag{IC}$$

Putting all of this together, the mechanism design problem faced by the market designer is given by:

$$\max_{k\geq 0} \max_{\substack{t_i(\theta,s)\to\mathbb{R}^+,\\ q_i(\theta,s)\to\mathbb{R}^+}} \quad CS(k) = \sum_i \ \mu_i \int_s\int_{\theta_i} U(q_i^*(\theta,s),\theta,s) \ - \ t_i^*(\theta,s)q_i^*(\theta,s)dG_i(\theta)dF(s)$$

$$\text{s.t.} \quad I(k) \leq \sum_i \mu_i \int_s\int_{\theta_i} t_i(\theta,s)q_i(\theta,s)dG_i(\theta)dF(s), \tag{R}$$

$$\sum_i \mu_i \int_s\int_{\theta_i} q_i(\theta,s)dG_i(\theta)dF(s) \leq k, \tag{K}$$

$$0 \leq \int_s U(q_i(\theta,s),\theta,s) \ - \ t_i(\theta,s)q_i(\theta,s)\ dF(s), \tag{IR}$$

$$\theta_i = \arg\max_{\hat{\theta}} \int_s U(q_i(\hat{\theta},s),\theta,s) - t_i(\hat{\theta},s)q_i(\hat{\theta},s)\ dF(s), \tag{IC}$$

We start describing the optimal allocation schedule given the new constraints and for a given level of investment k. We analyze the implications regarding investment decisions in the Appendix available upon request.

Using the method developed similarly to Guesnerie and Laffont (1984) and the incentive compatibility constraint, we characterize the monetary transfer $t_i(\theta,s)$ in terms of quantity $q_i(\theta,s)$. As our problem is well-defined, the incentive compatibility constraint is satisfied as soon as the optimal allocation $q_i(\theta,s)$ increases with respect to the type θ. Lemma 3 describes the revenue from a consumer given its type, and a realization of s is

Lemma 3. *The payoff equivalence implies the following relation between optimal transfer and quantity allocated to a consumer of type θ, from category i and given a realization of s:*

$$t_i(\theta,s)q_i(\theta,s) = U(q_i(\theta,s),\theta,s) - \int_{\bar{\theta}}^{\theta}\int_s q_i(\tilde{\theta},s)dF(s)d\tilde{\theta} + Cst$$

Where Cst is an arbitrary constant.

The proof uses the canonical approach of the Envelope Theorem (see Milgrom and Segal (2002)) that we modify to consider the capacity constraints. Next, we use the approach from Spulber (1992a) to characterize a feasibility constraint that associates the budget and individual rationality constraints. The core idea is that if one of the constraints is satisfied but not the other, a feasible lump-sum transfer from the non-binding constraint could exist that allows for relaxing the binding constraints. To say it differently, when there is, for instance, some excess revenue but the individual rationality is constraining, it is possible to transfer a lump-sum positive amount of money to the lowest types of consumers, which allows for less constraint optimal allocation. We describe in the following equation the corresponding new constraint, noted $R - IR$:

$$\sum_i \mu_i \int_s \int_\theta U(q_i(\theta, s), \theta, s) - \Gamma_i(\theta) \int_s q_i(\theta, s) dF(s) dG_i(\theta) dF(s) - I(k) \geq 0$$

With $\Gamma_i(\theta) = \frac{1-G_i(\theta)}{g_i(\theta)}$ the inverse hazard rate. Under our uniform distribution assumption considering the distribution of θ, the inverse hazard rate is decreasing with θ. We then solve for the Lagrangian. The following lemma shows the first-order condition to find the optimal allocation $q_i^*(\theta, s)$.

Lemma 4. *Given IC and IR constraints, the optimal allocation for a consumer of type θ from category i and for a given realization of s $q_{i,l}^*$ satisfies the following condition.*

$$u(q_{i,l}^*, \theta, s)(1 + \zeta) - \zeta \Gamma_i(\theta) - \varepsilon = 0$$

With ζ and ε, the Lagrangian multipliers for, respectively, the $R-IR$ condition and the capacity constraint. We denote $l = \{1, 2, 3, 4\}$ the index variable such that when $l = \{3, 4\}$ implies that $R - IR$ is binding while $l = \{1, 2\}$ means it does not, and $l = \{2, 4\}$ implies that the capacity is binding while $l = \{1, 3\}$ means it does not.

Given the lemma, we can prove that the optimal allocation increases with the type. For instance, the following equation shows the derivative of the optimal allocation when both capacity and the $R - IR$ constraint are binding.

$$\frac{\partial q^*_{i,4}}{\partial \theta} = \left(\frac{1}{2} - \frac{\partial u(q^*_{i,4}, \theta_i, s)}{\partial \theta}\right)\left[1 - \frac{\zeta}{1+\zeta}\Gamma'_i(\theta)\right] / \frac{\partial u(q^*_{i,4}, \theta, s)}{\partial q}$$

Following our model specification, $\frac{\partial u}{\partial \theta} = 1$ and $\frac{\partial u}{\partial q} < 0$, the assumption concerning $\Gamma'_i(\theta) < 0$ implies that $\frac{\partial q^*_{i,4}}{\partial \theta} > 0$. Lemma 4 provides four solutions to the problem faced by the market designer depending on which constraints are binding or not. We can regroup them in pairs such that $\{q^*_1(\theta, s), q^*_2(\theta, s)\}$ is the set of quantities when the optimal allocation is not constraint by $R - IR$. That is, the revenue generated by the mechanism is sufficient to cover the fixed costs and provide enough incentive for every consumer to consume electricity. On the other hand, $\{q^*_3(\theta, s), q^*_4(\theta, s)\}$ is the set of quantities such that the constraint is binding, implying that the optimal allocation needs to be distorted to covers both fixed costs and participation. In the following corollary, we characterize the first set of quantities.

Corollary 3. *The optimal allocation under the mechanism design approach when the $R - IR$ constraint is not binding is identical to the first-best allocation.*

When $\zeta = 0$, the condition in Lemma 4 is identical to the conditions described in the complete information section. Moreover, it can also be shown that the spot market schedule in prices and quantities is also incentive-compatible. We next analyze the threshold between the two sets of quantities. That is, we described under which value of k the market designer faces a binding $R - IR$. We summarize our findings in the following proposition.

Proposition 5. *There exists a unique value of k such that the $R - IR$ is null. Moreover, for any value of k below this threshold, the constraint $R - IR$ is not binding, while any value above the constraint is binding.*

Proposition 5 that it is possible to cover both fixed costs and participation constraints without distorting the allocation only when the level of investment is low. The intuition for this result can be understood as follows. First, we denote the marginal virtual utility: $J_{i,k} = u(q^*_i(\theta, s), \theta, s) - \Gamma_i(\theta)$, which is the marginal utility derived from the optimal allocation net of the information rent. Under our framework, it can be interpreted as the feasible gain in utility from the allocation after having remunerated the consumers to behave truthfully. Then, we can express the derivative of the $R-IR$ constraint for the first set of optimal quantities.

$$\overbrace{\sum_i \mu_i \int_{s_1(k)}^{\bar{s}} \int_{\theta_i} J_{i,2} \frac{\partial q^*_{i,2}(\theta, s)}{\partial k} dG_i(\theta) dF(s)}^{\text{aggregate expected marginal virtual revenue}} - r$$

Essentially, the constraint starts binding when the aggregated marginal virtual revenue from the mechanism during the on-peak period equals the marginal investment. Note that both the marginal virtual utility and the derivative of the quantity are, in that case, positive. Under our framework, $\frac{\partial q^*_{1,2}}{\partial k}$ is equal to 1, so to ensure that $\sum_i \mu_i \frac{\partial q^*_{1,2}}{\partial k} = 1$. Therefore, an increase of k generates an ambiguous effect on the constraint: (i) it increases the virtual surplus during on-peak periods, and (ii) it increases the investment costs. However, the expected surplus from consumers is concave. Indeed, note that the derivative of the marginal virtual utility with respect to k is equal to $\frac{\partial J_{i,k}}{\partial k} = -\frac{\partial q^*_{i,k}}{\partial k}$, meaning that if an increase of the investment increases the optimal quantity, then it decreases the possible marginal utility net of information rent. This effect also accumulates with the change in occurrence between off-peak and on-peak. As k increases, the capacity binds less in expectation, implying a decrease in the positive first part of the expression above. This second-order effect, combined with the increase in investment costs, implies that the constraint crosses binds only once.

We illustrate our findings in Figure 6. We plot the $R - IR$ constraint under the optimal allocation set $\{q^*_{i,1}, q^*_{i,2}\}$ for different values of k. The black curve shows the constraint. We decompose it with the blue curve only representing the aggregate expected consumer virtual utility and the red curve representing the investment costs. As previously described, the blue curve is concave in k with an increasing value and then a decreasing part. Note that for sufficiently high values of k, the utility function is even independent of k because, in expectation, the capacity is never binding. When adding the increasing investment costs, the difference between the two has to exhibit, at one point, a decreasing behavior. Finally, we have represented the threshold value with the vertical dashed line. For lower values of k, the $R-IR$ constraint is positive, meaning that fixed costs and the information constraint are not binding. Above this value, the value is negative,

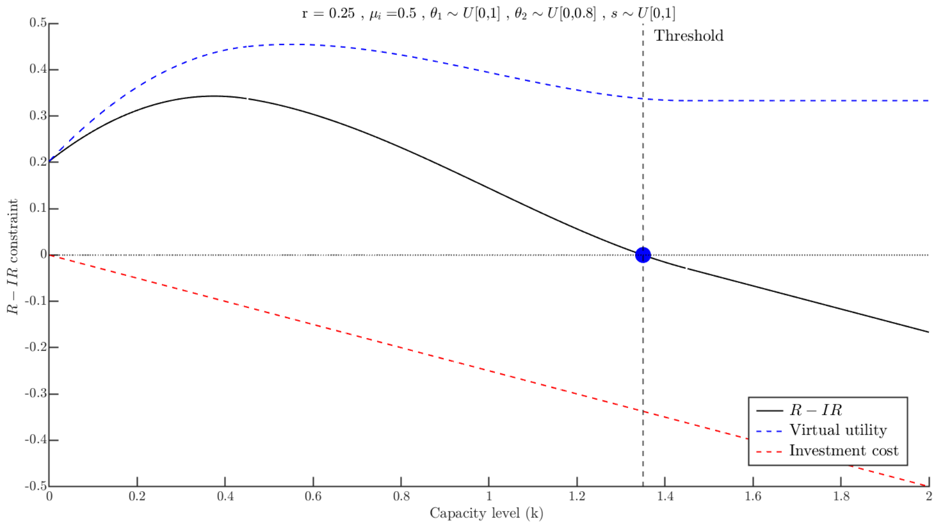


Figure 6: $R - IR$ constraints and its component with respect to the investment level.

so the market designer needs to distort the allocation so that $R - IR = 0$. We now describe how the optimal allocation depends on the investment level when the $R - IR$ binds. We summarize our main findings in the following proposition.

Proposition 6. *The investment level directly affects the optimal allocation.*

- *For every consumer, the optimal quantity during off-peak $q^*_{i,3}$ is always decreasing with k.*
- *For the optimal allocation $q^*_{i,4}$, there exists for each category a unique threshold $\theta^*_i(k)$. For consumers of a category i, if his type is below $\theta^*_i(k)$, his allocation $q^*_{i,4}$ decreases with k. If his type is above $\theta^*_i(k)$, his allocation $q^*_{i,4}$ increases with k. Moreover, $\theta^*_i(k)$ is increasing with k.*

The proposition states that for a higher level of investment, every consumer should receive less electricity during off-peak periods. During on-peak, the change of quantity depends on the types. For lower types of both categories, consumers should also receive less electricity. On the other hand, higher types always receive more electricity as capacity expands. When the capacity is not binding, the effect on the quantity is captured in the equation below:

$$\frac{\partial q^*_{i,3}}{\partial k} = \frac{J_{i,3}}{1+\zeta}\frac{\partial \zeta}{\partial k}$$

The derivative is derived from the first-order condition from Lemma 4. Namely, it can be rewritten such that: $u(q^*_{i,k}, \theta, s)(1+\zeta) - \zeta\Gamma_i(\theta) - \varepsilon = u(q^*_{i,k}, \theta, s) + \zeta J_{i,3} - \varepsilon$. When capacity is not binding, we have $\varepsilon = 0$. Hence, the marginal virtual utility at the optimal allocation during off-peak is always negative. Under our framework, and similarly to the previous sections, we know that the budget constraint behaves convexly with respect to k: a higher capacity level implies a higher need for revenue. Thus it implies that $\frac{\partial \zeta}{\partial k} > 0$. The two observations lead to a negative derivative. The economic intuitions can be understood as follows: as k expands, this does not directly generate any additional quantity for consumers during off-peak, as, by definition, the capacity is not binding. On the other hand, the need for revenue is increasing. Combining the absent surplus effect and the negative budget effect implies that the optimal quantities for all consumers are decreasing. For the on-peak allocation, the initial derivative is expressed as follows:

$$\frac{\partial q^*_{i,4}}{\partial k} = \left[J_{i,4} \frac{\partial \zeta}{\partial k} - \frac{\partial \varepsilon}{\partial k} \right] \frac{1}{1+\zeta}$$

As quantity expands, the willingness to pay for less binding constraint decreases, implying that $\frac{\partial \varepsilon}{\partial k} < 0$. Therefore, the sign of the derivative is ambiguous and notably depends on the sign of $J_{i,4}$. Contrary to the off-peak allocation, the initial first-order condition when $\varepsilon > 0$ does not allow a clear-cut answer for the sign of the virtual marginal utility. Using the constraint from the market design problem, we can express the derivative of the Lagrange multiplier ε associated with the capacity constraint as a function of the derivative of ζ with respect to k. Namely, after simplification, we find that the derivative of the optimal quantity can be expressed as follows:

$$\frac{\partial q^*_{i,4}}{\partial k} = [J_{i,4} - \mathbb{E}J_4] \frac{\partial \zeta}{\partial k} \frac{1}{1+\zeta} + 1$$

With $\mathbb{E}J_4 = \sum_i \mu_i \int_{\theta_i} J_{i,4} dG_i(\theta)$ the aggregate marginal virtual utility over every consumer and across all groups. The equation states a sufficient condition for having a positive derivative for a given consumer: If his virtual marginal utility is (sufficiently) higher than the aggregate marginal virtual utility, then its allocation is increasing with k. This condition captures the fundamental trade-off that the market designer faces when there is an information constraint. First, note the value 1 on the right part of the equation. It describes the positive effect of increasing k for consumers when the capacity is binding, which always implies higher utility. Then, let's assume a consumer that, even with on-peak periods, has a negative marginal virtual utility, either because the marginal utility u is also negative or because the information rent $\Gamma_i(\theta)$ is too high. Similarly to the off-peak case and the previous section, allocating a given quantity of electricity to the smaller consumer is always negative (at the margin). Depending on the model parameters, the potential adverse effect of having a negative marginal virtual utility has to be compared to the positive effect of 1 associated with a less binding capacity. Finally, even when a consumer has $J_{i,4} > 0$, the market designer, due to the capacity constraint and the need to cover the fixed costs, has to favor the consumers for which it is less costly to induce an optimal allocation, that is, for consumers of the highest type. This tension is highlighted by the delta $J_{i,4} - \mathbb{E}J_4$. We illustrate our findings in Figure 7.

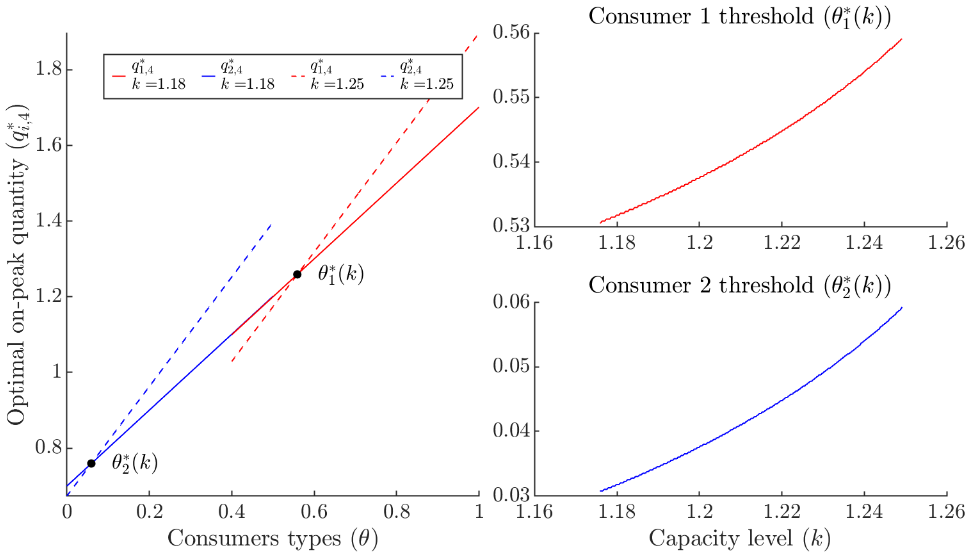


Figure 7: Optimal on-peak allocation with respect to the consumers' type θ, and threshold θ^*_i with respect to k

The left panel shows the optimal on-peak allocation for each consumer depending on their type and for a given realization of s. The solid lines represent a set of quantities given a value of k, and the dashed lines represent the allocation for a higher value of k. As described in the proposition, we observe a rotation of the allocation, with higher types receiving more goods while lower types endure a decrease in their quantity allocated. Interestingly, we do not observe a strict ranking between consumers of different categories. Namely, optimally reducing the quantities given to each consumer concerns the lowest type in each category but not across categories. The rationale behind those results lies in how incentive compatibility and individual rationality contain the market designer. As he can discriminate the consumers based on their category, which is publicly observed, the cost associated with the information rent (partly) depends on the category the consumer belongs to. Therefore, it is less costly to discriminate the consumers of the lowest type negatively.

6 Conclusion

This paper built a tractable framework to analyze the role of market designers in finding the most efficient way of consuming an essential good when faced with investment decisions. Most of the literature has focused either on providing additional remuneration streams for producers to increase the level of investment or on designing the second-best pricing schedule for consumers, given informational and technical constraints. This paper provides a unifying framework linking investment decisions and consumer participation. We show an inherent tension when implementing an allocation mechanism to maximize consumer surplus and generate revenue to cover fixed costs. The paper provides policy and technical results by adding to the initial framework additional constraints. We assume that consumers possess private information with respect to their utility level and that the market designer may be constrained in the allocation mechanism he can propose to consumers. The central result of the paper is that, depending on a set of assumptions, some specific and non-intuitive relations exist between the level of investment and the optimal allocation proposed to consumers, which has significant welfare and distributive implications. Namely, when the market designer faces heterogeneous consumers, an increase in the level of investment may affect different consumers depending on their type. For instance, under a mechanism design approach, low types can experience a decrease in their quantity allocated despite an increase in the capacity level.

Finally, we plan to extend our result with two main extensions: (i) study market design constraints with the mechanism design framework. While market designers may wish to implement some information revelation mechanism, as theoretically studied in the third result, practical contractual arrangements between the market designer and consumers may constrain him in the implementable mechanism. It would lead to specific effects, as highlighted in the second set of results. (ii) Implement specific distribution preferences associated with consumer types and categories. Our current framework does not consider welfare weights, which may distort the optimal allocation. Including such parameters would highlight the tension between generating sufficient revenue and maximizing consumer surplus. From a more extreme view, as our paper shows that the allocation can exhibit some non-monotonicities of the optimal quantities and prices, a market designer may want to avoid any decreasing quantities when the level of investment rises. Including such constraints in our framework could highlight a new trade-off.

References

Akbarpour, M., Budish, E. B., Dworczak, P., and Kominers, S. D. (2023a). An economic framework for vaccine prioritization. *Available at SSRN 3846931*.

Akbarpour, M., Dworczak, P., and Kominers, S. D. (2023b). Redistributive allocation mechanisms. *Available at SSRN 3609182*.

Astier, N. (2021). Second-best pricing for incomplete market segments: Application to electricity pricing. *Journal of Public Economic Theory*, 23(6):1287–1311.

Athey, S., Castillo, J. C., Chaudhuri, E., Kremer, M., Simoes Gomes, A., and Snyder, C. M. (2022). Expanding capacity for vaccines against Covid-19 and future pandemics: a review of economic issues. *Oxford Review of Economic Policy*, 38(4):742–770.

Ausubel, L. M. and Cramton, P. (2010). Using forward markets to improve electricity market design. *Utilities Policy*, 18(4):195–200.

Boiteux, M. (1949). *La tarification des demandes en pointe: Application de la théorie de la vente au coût marginal*. place Henri-Bergson.

Burger, S., Knittel, C., and Pérez-Arriaga, I. (2020). Quantifying the distributional impacts of rooftop solar pv adoption under net energy metering.

Cahana, M., Fabra, N., Reguant, M., and Wang, J. (2022). The distributional impacts of real-time pricing.

Chao, H.-p. (2011). Efficient pricing and investment in electricity markets with intermittent resources. *Energy Policy*, 39(7):3945–3953.

Chao, H.-p. (2012). Competitive electricity markets with consumer subscription service in a smart grid. *Journal of Regulatory Economics*, 41(1):155–180.

Chao, H.-p., Oren, S., and Wilson, R. (2022). Priority pricing for clean power under uncertainty. *Current Sustainable/Renewable Energy Reports*, 9(3):52–64.

Chao, H.-p. and Wilson, R. (1987). Priority service: Pricing, investment, and market organization. *The American Economic Review*, pages 899–916.

de Palma, A., Lindsey, R., and Monchambert, G. (2017). The economics of crowding in rail transit. *Journal of Urban Economics*, 101:106–122.

Eurelectric (2021). Are efficient network tariffs the missing piece in the energy transition? Technical report, Eurelectric.

Fabra, N. (2018). A primer on capacity mechanisms. *Energy Economics*, 75:323–335.

Gilbert, R. J. and Klemperer, P. (2000). An equilibrium theory of rationing. *The RAND Journal of Economics*, pages 1–21.

Guesnerie, R. and Laffont, J.-J. (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(3):329–369.

Hall, J. D. (2021). Can tolling help everyone? estimating the aggregate and distributional consequences of congestion pricing. *Journal of the European Economic Association*, 19(1):441–474.

Holmberg, P. and Ritz, R. A. (2020). Optimal capacity mechanisms for competitive electricity markets. *The Energy Journal*, 41(Special Issue).

IEA (2020). Power systems in transition. Technical report, IEA, Paris.

IEA (2021). Security of clean energy transitions. Technical report, IEA, Paris.

ITDP (2021). The compact city scenario – electrified. Technical report, ITDP, and the University of California, Davis.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review*, 104(2):537–563.

Kominers, S. D. and Tabarrok, A. (2022). Vaccines and the Covid-19 pandemic: lessons from failure and success. *Oxford Review of Economic Policy*, 38(4):719–741.

Léautier, T.-O. (2016). The visible hand: ensuring optimal investment in electric power generation. *The Energy Journal*, 37(2).

Leautier, T.-O. (2018). On the long-term impact price caps: Investment, uncertainty, imperfect competition, and rationing. *International Journal of Industrial Organization*, 61:53–95.

Levinson, A. and Silva, E. (2022). The electric gini: Income redistribution through energy prices. *American Economic Journal: Economic Policy*, 14(2):341–365.

Llobet, G. and Padilla, J. (2018). Conventional power plants in liberalized electricity markets with renewable entry. *The Energy Journal*, 39(3).

Loertscher, S. and Marx, L. M. (2021). Incomplete-information models for industrial organization. Technical report, Working Paper.

Loertscher, S. and Muir, E. V. (2020). Monopoly pricing, optimal rationing, and resale. *Working Pa.*

Loertscher, S. and Muir, E. V. (2021). Wage dispersion, minimum wages and involuntary unemployment: a mechanism design perspective.

Martimort, D. and Stole, L. A. (2020). Nonlinear pricing with average-price bias. *American Economic Review: Insights*, 2(3):375–396.

Milgrom, P. and Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601.

Persad, G., Peek, M. E., and Emanuel, E. J. (2020). Fairly Prioritizing Groups for Access to COVID-19 Vaccines. *JAMA*, 324(16):1601–1602.

Rahmandad, H. (2022). Behavioral responses to risk promote vaccinating high-contact individuals first. *System Dynamics Review*, 38(3):246–263.

Schrank, D. L. and Lomax, T. J. (2021). *Urban mobility report.* Texas Transportation Institute, The Texas A&M University System.

Spulber, D. F. (1992a). Capacity-contingent nonlinear pricing by regulated firms. *Journal of Regulatory Economics*, 4(4):299–319.

Spulber, D. F. (1992b). Optimal nonlinear pricing and contingent contracts. *International Economic Review*, pages 747–772.

Spulber, D. F. (1993). Monopoly pricing of capacity usage under asymmetric information. *The Journal of Industrial Economics*, pages 241–257.

Sudarmawan, A., Zha, H., Tatlow, H., Green, K., Di Folco, M., Hale, T., and Phillips, T. (2022). What have we learned from tracking every government policy on covid-19 for the past two years? Technical report, Blavatnik School of Government Research Note.

Vickrey, W. S. (1963). Pricing in urban and suburban transport. *The American Economic Review*, 53(2):452–465.

Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, pages 251–260.

Yoshida, Y. (2008). Commuter arrivals and optimal service in mass transit: Does queuing behavior at transit stops matter? *Regional Science and Urban Economics*, 38(3):228–251.

Zöttl, G. (2011). On optimal scarcity prices. *International Journal of Industrial Organization*, 29(5):589–605.