



HAL
open science

Securing Investments in Electricity Markets. Three Essays on Market Design

Léopold Monjoie

► **To cite this version:**

Léopold Monjoie. Securing Investments in Electricity Markets. Three Essays on Market Design. Economics and Finance. Université Paris sciences et lettres, 2024. English. NNT : 2024UPSLD005 . tel-04608818

HAL Id: tel-04608818

<https://theses.hal.science/tel-04608818v1>

Submitted on 11 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à Université Paris Dauphine - PSL

Securing Investments in Electricity Markets
Three Essays on Market Design

Soutenue par

Léopold Monjoie

Le 25 mars 2024

École doctorale n°543

SDOSE

Spécialité

Sciences économiques

Composition du jury :

| | |
|---|---------------------------|
| Anna CRETl Professeur, Université Paris Dauphine - PSL | <i>Présidente</i> |
| Carine STAROPOLI Professeur, Université de Rouen Normandie | <i>Rapporteure</i> |
| Matti LISKI Professeur, Aalto University | <i>Rapporteur</i> |
| Mar REGUANT Professeur, Northwestern University | <i>Examinatrice</i> |
| Thomas TANGERÅS Professeur Associé, Stockholm University | <i>Examineur</i> |
| Fabien ROQUES Professeur Associé, Université Paris Dauphine - PSL | <i>Directeur de thèse</i> |

L'Université Paris Dauphine - PSL n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

Securing Investments in Electricity Markets

Three Essays on Market Design

Léopold Monjoie

Remerciements

Je tiens à exprimer ma profonde gratitude à mon directeur de thèse, Fabien Roques, pour son soutien inestimable tout au long de mon parcours académique. Ses conseils avisés et son engagement sans faille ont été les piliers de cette aventure intellectuelle. Sa patience et son expertise ont été une source d'inspiration et m'ont permis de repousser mes limites et d'atteindre de nouveaux sommets dans mes recherches. Je lui suis reconnaissant de la confiance qu'il m'a accordée et de l'opportunité qu'il m'a donnée de contribuer à la recherche.

Je tiens également à remercier chaleureusement Carine Staropoli et Matti Liski d'avoir accepté d'être les rapporteurs de cette thèse. I am fully aware of the magnitude of this task and the time and investment it requires. I am, therefore, all the more grateful for your commitment and dedication to this project. Finally, I'd like to thank Mar Reguant and Thomas Tangerås for agreeing to take part in this jury and shed light on my research work. I am deeply honored.

Cette thèse n'aurait certainement pas vu le jour sans le soutien d'Anna Creti et de Patrice Geoffron, qui m'accompagnent depuis quasiment huit ans, de ma première année de Master jusqu'à mes CST. L'intérêt que vous portez à mon travail et votre engagement à toujours me faire les retours les plus pertinents ont été une source de motivation pour continuer à aller plus loin. Merci également pour votre gentillesse et votre disponibilité.

Merci à tous les membres du Laboratoire d'Economie de Dauphine. Tout d'abord, l'équilibre du CGEMP, sans qui je n'aurais pas fait ce master, et encore moins cette thèse : Jan Horst Keppler, Sophie Meritet, Yannick LePen et Marie Bessec, Stéphanie Monjon. La Chaire CEEM m'a également été d'un grand soutien durant cette thèse, à la fois sur le plan matériel mais surtout au niveau des nombreux séminaires qu'elle a organisés, permettant des débats riches sur les sujets abordés durant cette thèse en particulier. Et surtout, un grand merci à Steve pour sa patience.

Je tiens à remercier les personnes qui m'ont permis de réaliser cette thèse dans un environnement des plus favorables. En effet, cette thèse n'aurait pas été possible sans le soutien de RTE à travers ce contrat CIFRE. Un grand merci donc à Vincent, qui a été une source d'inspiration pendant ma thèse, et qui m'a accompagné jusqu'ici. Merci à toute l'équipe de la DIESE et de PEE, avec qui j'ai passé des moments incroyables : Kevin, Jean Baptiste, Adrienne, Bianka, Jules, Cédric...

Je remercie également les membres du Laboratoire d'Economie de Dauphine. Merci à Bertrand de m'avoir fait confiance pour enseigner l'économie industrielle et d'avoir

écouté mes théories farfelues dans cette thèse. Un grand merci aux *djeuns*, Romain pour les innombrables pauses, Arthur, pour tes commentaires lors de mes présentations, Loïc et Laura pour votre soutien lors du job market, c'est toujours très important pour le moral d'avoir des gens comme vous. Un grand merci également à Daniel pour m'avoir accompagné tout au long de cette thèse. Merci à Vanina pour avoir supporté ma phobie administrative (oui, ça existe).

Une dédicace spéciale à Toan et Sanni pour m'avoir accompagné lors de ma visite à Melbourne et pour m'avoir donné probablement mon seul coup de soleil le jour de Noël. Une pensée particulière pour David Byrne, sans qui cette visite n'aurait pas eu lieu.

Un grand merci à mes chers amis de P206Bis, cela ne fait même pas un an que nous sommes ensemble dans ce bureau, mais cela me semble déjà une éternité. Ils ont eu le courage et la patience de me supporter pendant la fin de ma thèse, et je leur en serai éternellement reconnaissant. Juju pour ta soif de connaissance et ton goût infatigable pour la musique. Côme, toujours si bien habillé, mais aussi la rigueur et la gentillesse incarnée, avec qui on pourrait sauter en parachute les yeux bandés. Inès, la terre est basse, mais on trouvera toujours le moyen de se relever et de s'extasier du monde.

Cédric, j'espère qu'on repartira en vacances ensemble à Chypre. Benoît, nous avons commencé ensemble, nous avons fait le même chemin, et je suis sûr que nous nous reverrons. Adèle, le rayon de soleil tant nécessaire pour le C104gauche. Merci à Fatou d'avoir pimenté ma vie à Dauphine. Eléonore, j'espère qu'on se retrouvera au paradis. Mathilde, même si la thèse ne me casse pas trois patés à un crabe, ton rire me manquera. Coline, je suis si heureuse que les vents du nord t'aient conduite jusqu'à moi. Cécile, depuis le temps que j'en entends parler, j'ai hâte de venir chez toi. Paul-Emile, j'aurais bien aimé une raclette à ce moment précis, mais je crois qu'il est un peu tard.

Merci à tous les doctorants qui ont eu la gentillesse de supporter mon humour : Matteo (désolé de l'oubli), Chloé, Malte, Eugénie, Arthur, Anatole, Anthonin, Paul B., Peio, Zélie. Et à tous les doctorants et anciens doctorants que je pourrais oublier. Une pensée particulière pour tous ceux que j'ai rencontrés lors de conférences et qui, je l'espère, continueront à faire partie de mon parcours. Une pensée particulière pour Jacqueline et Alicia.

Un grand merci à mes meilleurs amis, Camille, je suis si heureux de t'avoir dans ma vie. Antoine, Joris, Luc, Mélo, Livia, merci de m'avoir montré que la vie vaut la peine

d'être vécue et de m'avoir soutenue pendant tout ce temps. Hugo, je suis enfin disponible pour faire ce que tu veux.

Et bien sûr, ma famille. Merci Alexis de faire partie de ma vie et d'être avec moi jusqu'à ce jour. Tu sais comme moi qu'une thèse n'est pas un chemin facile, et tu as tout fait pour que ces dernières semaines soient les plus agréables possibles. Une pensée particulière pour mes cousines, Clotilde, Jeanne et Adèle, qui font partie de ma vie depuis si longtemps. Dédicace à mon cousin Octave, ne change rien. Merci à ma Catherine et à ma Sylvie pour leur précieuse chaleur. Merci à Dominique pour tes discussions et ton humour. Je n'oublie pas François, qui m'a accompagné depuis mes premiers pas et pour qui j'aurai toujours la plus grande affection. Et surtout les plus importants, ma mère et mon père, sans qui je ne serais rien. Vous avez toujours été les premiers à croire en moi, et vous n'avez jamais douté de cette aventure. Cette thèse est aussi la leur. Merci à tous.

Résumé en français

L'électricité est au cœur de notre société moderne. Elle est centrale pour le développement économique, fournissant l'énergie nécessaire pour alimenter les industries, les entreprises et les ménages. Malgré son rôle dans l'économie, et quel que soit le mode de gouvernance envisagé, l'électricité est confrontée à une contrainte physique importante. À tout moment, la quantité consommée doit être égale à la quantité injectée dans le système. Par conséquent, le défi d'avoir toujours une capacité de production suffisante pour répondre à une demande donnée est crucial lorsque le stockage est insuffisant. Dans ce contexte, **comment pouvons-nous garantir que l'équilibre entre la production et la consommation est maintenu aussi efficacement que possible ?** Cette question est d'autant plus importante que l'électricité est appelée à jouer un rôle central dans la transition énergétique, que ce soit du côté de la production, en raison du passage nécessaire d'une production carbonée à une production décarbonée, ou du côté de la consommation, en raison de l'électrification de nos usages tels que les transports, le chauffage, ou de la numérisation de l'économie. La figure 1 illustre, par exemple, l'investissement nécessaire selon deux scénarios à l'horizon 2040 par rapport aux investissements réalisés au cours de la dernière décennie. La figure 2 illustre une situation dans laquelle le Texas a subi des phénomènes météorologiques hivernaux extrêmes qui ont entraîné des coupures de courant importantes. En d'autres termes, la nécessité de mettre en place des institutions efficaces pour garantir des investissements suffisants découle non seulement de la nature même de l'électricité, mais aussi de son rôle central dans la transition énergétique.

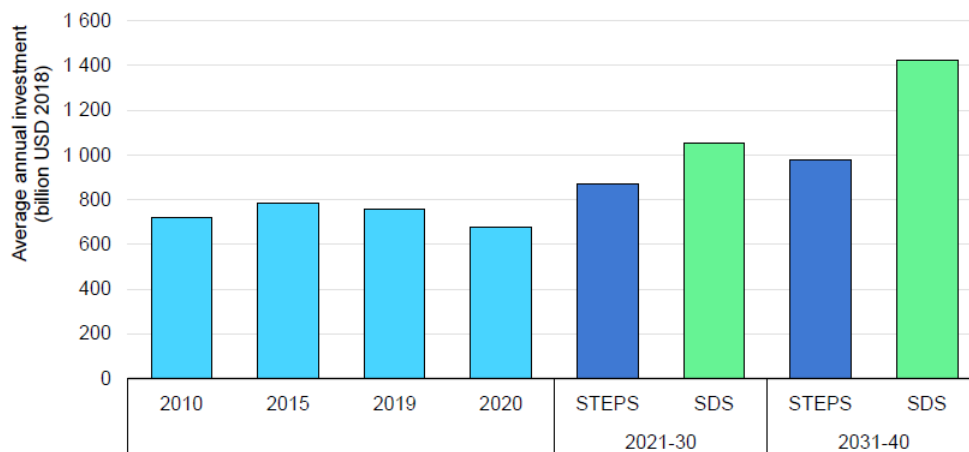


Figure 1: Investissement annuel moyen dans le secteur de l'électricité au niveau mondial, 2010-20, et besoins jusqu'en 2040. (IEA, 2020)

Les marchés de l'électricité présentent de nombreux problèmes analysés par l'économie moderne. D'un point de vue microéconomique et d'économie publique, plusieurs décennies

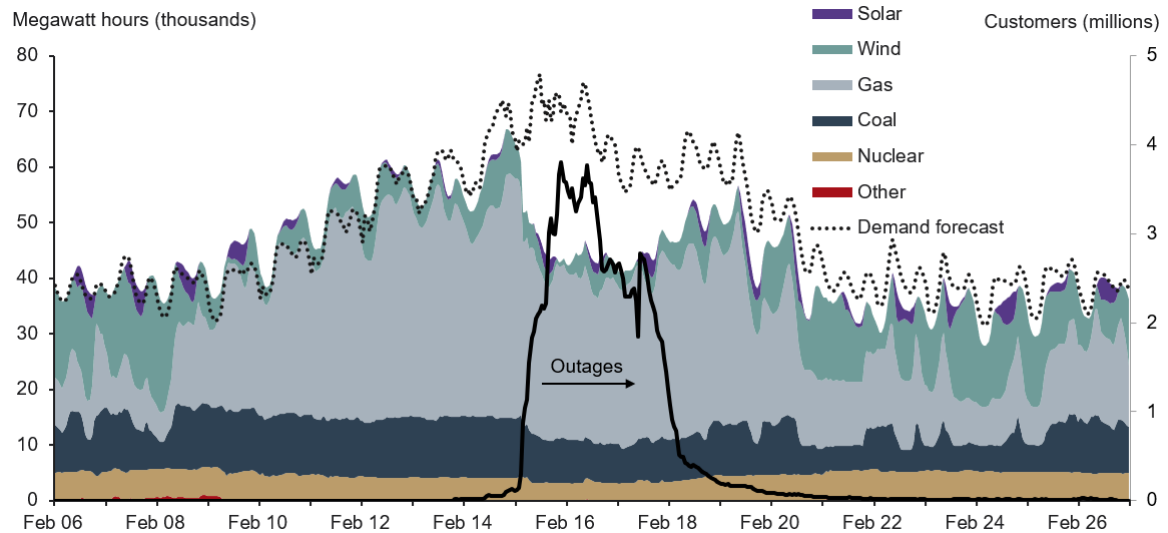


Figure 2: Production d'électricité de l'ERCOT par source, demande et pannes pendant la crise énergétique de 2021 au Texas ([FED, 2021](#))

d'études ont souligné ses nombreuses défaillances de marché : les externalités dues à sa production, son transport et sa consommation ; la nature de bien public de ses infrastructures ; son incomplétude, notamment en ce qui concerne la gestion des risques ; sa spécificité, qui en fait un terrain de jeu idéal pour l'exercice d'un pouvoir de marché. L'existence d'un certain nombre de contraintes physiques pose également un certain nombre de défis aux économistes, qui doivent adapter leur analyse aux lois physiques régissant l'ensemble de la chaîne de valeur de l'électricité. Par ailleurs, la récente crise COVID-19 et la guerre en Ukraine ont également mis en évidence le lien entre l'électricité et les questions macroéconomiques, notamment l'inflation. Dans le chapitre introductif, je décris notamment plus en détail pourquoi nous avons besoin de suffisamment d'investissements. Je discute également des principales idées théoriques qui ont conduit à la libéralisation et à la mise en place des marchés de l'électricité. Je rappelle que si la théorie affirme que le niveau d'investissement peut être atteint efficacement grâce aux marchés, il existe un grand nombre de problèmes qui empêchent ces marchés d'être pleinement efficaces. Ce constat motive d'autant plus l'étude de la forme de gouvernance et d'institution que nous pouvons construire pour fournir les bonnes incitations.

C'est dans ce contexte que cette thèse cherche à répondre à un certain nombre de questions sous l'angle de la théorie économique. En particulier, j'ai choisi d'étudier le lien entre le fonctionnement des marchés de l'électricité et les décisions d'investissement. En d'autres termes, **comment s'assurer que les institutions mises en place pour produire, échanger et consommer l'électricité fournissent des incitations suff-**

isantes pour investir efficacement ?¹ Bien que les questions d'investissement ne soient pas spécifiques à l'électricité, le manque actuel de capacité de stockage à une échelle suffisante et à un coût relativement faible signifie que tout écart dans la disponibilité des investissements par rapport aux niveaux de consommation comporte un risque important. Dans le meilleur des cas, ces risques se traduisent par un rationnement involontaire de la demande, également connu sous le nom de "rolling blackout", et, dans le pire des cas, par un effondrement total du système. Cette question est d'autant plus critique que l'un des principaux objectifs de la politique publique dans le secteur de l'électricité est le développement massif des énergies renouvelables.

Comme je le soulignerai dans cette conclusion générale, j'ai abordé ce problème d'une manière spécifique. L'idée centrale de cette thèse est d'essentialiser et de conceptualiser les particularités du lien entre les décisions d'investissement et de production. Ce lien est bien décrit dans le concept d'externalités d'allocation utilisé dans la littérature économique : les décisions individuelles d'investissement dépendent des équilibres du marché de l'électricité, et les équilibres dépendent des investissements. En fin de compte, l'interdépendance entre les investissements et les équilibres déterminera le bien-être. Une fois ce lien clarifié, j'ai entrepris de répondre à la question suivante concernant la conception du marché : **quelles sont les règles d'organisation des marchés qui maximiseront l'efficacité dans ce cadre particulier ?** En d'autres termes, je tente de répondre à la question pratique de savoir comment fournir les meilleures incitations à l'investissement afin que le niveau atteint soit celui souhaité mais aussi que la manière dont il est atteint ne génère pas trop de coûts indirects et négatifs pour la société.

Le point de départ de cette thèse repose sur le constat suivant. Le secteur de l'électricité est constitué d'une succession d'essais et d'erreurs sur la manière d'organiser efficacement ces marchés. Il offre alors un vaste terrain de jeu pour analyser les politiques de mise en place de diverses règles de marché, tant d'un point de vue positif que normatif. Je prends comme point de départ une approche spécifique : celle des marchés de capacité. Ces marchés se caractérisent par le fait que les producteurs promettent d'être disponibles dans le futur en échange d'une rémunération. Plutôt que de se demander si

¹Les décisions d'investissement couvrent un large éventail de problématiques: (1) d'un point de vue statique, l'efficacité d'une décision d'investissement peut être définie par un niveau et une composition adéquats, et (2) d'un point de vue dynamique, il convient de s'assurer que les décisions sont prises en temps opportun. En particulier, les décisions d'entrée et de sortie, y compris la décision de maintenir ou de mettre sous cocon certaines capacités. Dans cette thèse, je me concentrerai sur l'efficacité du niveau d'investissement, et son maintien, plutôt que sur la composition ou l'efficacité dynamique.

les marchés de capacité sont plus efficaces que d'autres formes d'organisation, ma thèse s'intéresse à la manière dont ils peuvent être mis en place de la manière la plus optimale.

Bien que les contributions empiriques aux marchés de capacité soient restées limitées, les économistes ont utilisé une grande diversité d'outils de modélisation, allant de la représentation de systèmes d'ingénierie complexes à des modèles théoriques classiques, analytiques et stylisés, chacun ayant ses avantages et ses inconvénients. Mes premières années de doctorat ont été consacrées à une analyse approfondie de la littérature quantitative sur les marchés de capacité. Cependant, je me suis rendu compte que pour aborder des questions spécifiques de conception de marché qui étaient très pratiques à première vue, il était nécessaire de revenir aux fondements théoriques des marchés de l'électricité. Par conséquent, mes choix de modélisation se sont concentrés sur une représentation très simplifiée de l'environnement industriel afin d'étudier ce que signifie (i) prendre une décision d'investissement et (ii) participer aux marchés, que ce soit d'un point de vue à long terme avec les marchés de capacité ou sur les marchés de gros à court terme. Par conséquent, chaque chapitre de cette thèse vise à réduire la description de la réalité à ses éléments essentiels afin de mieux comprendre les mécanismes derrière lesquels les choix de conception du marché ont des effets positifs ou négatifs en termes de bien-être.

Résultat Principal

Chaque chapitre de cette thèse met en lumière des questions bien connues en matière de conception de marché et d'économie industrielle. Il y a toujours des compromis fondamentaux qui doivent être compris, et les règles du marché entraînent toujours des conséquences non identifiées, bonnes ou mauvaises. La contribution centrale de cette thèse est de se concentrer sur un sujet très spécifique : dans un environnement aussi particulier que l'électricité, un niveau d'investissement suffisant est crucial et souvent imposé par les décideurs politiques. Cependant, les décisions d'investissement sont prises par des acteurs privés. Dès lors, comment s'assurer que les marchés sont conçus pour garantir la rencontre la plus vertueuse entre les valeurs privées et les besoins sociaux ? En représentant le comportement des acteurs par le biais de la théorie économique et en mettant l'accent sur la modélisation des décisions d'investissement et leur lien avec le reste du marché, la thèse propose de formaliser des effets encore peu abordés ainsi que de donner de nouvelles recommandations sur la façon de concevoir des marchés pour l'investissement.

Au-delà de la contribution que je décris plus en détail pour chaque chapitre à la fin de ce résumé. Ce manuscrit, je l’espère, ouvre la voie à de nouvelles recherches. Pour commencer, il est nécessaire de rappeler l’importance cruciale de l’approche empirique en économie. Les marchés de capacité existent depuis une vingtaine d’années dans le monde, et même si les données ne sont pas aussi robustes que pour les autres marchés de l’électricité, je pense qu’il est tout à fait possible de creuser dans cette direction. Cette extension de ma recherche pourrait impliquer à la fois une approche économétrique classique et l’utilisation de données pour construire des contrefactuels sur la base de modèles plus techniques, donnant une image plus réaliste du fonctionnement d’un système électrique. Cela permettrait également de mesurer la valeur réelle des compromis mis en évidence dans cette thèse. Les investissements jouant un rôle crucial sur les marchés de l’électricité, tant en raison de leur valeur que de leur coût, une telle approche serait utile au débat politique sur l’acceptation des nouvelles technologies et réglementations.

Les observations faites dans le chapitre d’introduction de cette thèse constituent une autre limite importante de ce travail. En effet, j’ai décrit comment l’efficacité d’un marché de capacité est mesurée non seulement en termes de niveau d’investissement, mais aussi en termes d’autres décisions industrielles étroitement liées, telles que les décisions d’entrée et de sortie du marché, ainsi que les décisions de maintenir un certain niveau de capacité. De plus, dans le contexte de la transition énergétique, il est crucial de se poser la question de la composition du mix énergétique. Il serait alors envisageable d’étendre les travaux présentés dans chaque chapitre pour prendre en compte ces autres dynamiques. L’objectif de ces extensions n’est pas simplement d’ajouter un degré de réalisme aux modèles théoriques mais de comprendre comment les comportements et les nouveaux arbitrages soulevés par ces questions s’ajoutent ou s’opposent aux résultats trouvés dans cette thèse. À titre d’exemple, je prendrai les énergies renouvelables qui, contrairement aux moyens de production plus conventionnels, ont un niveau de capacité incertain mais un coût marginal certain.² Dans les sections suivantes, je donne une vue plus détaillée des principaux résultats et des extensions possibles à chaque chapitre.

Enfin, j’aimerais souligner que la contribution de cette thèse va au-delà des marchés de l’électricité. L’approche théorique stylisée utilisée dans cette thèse permet d’aborder des questions plus larges, notamment celle d’assurer des investissements suffisants pour un ensemble de *biens essentiels*. Ces biens sont caractérisés par la nature de bien public

²Voir par exemple [Fabra and Llobet \(2023\)](#) qui montrent que les enchères dans ce contexte peuvent avoir un résultat différent de celui des modèles canoniques.

de l'investissement disponible lorsque l'offre est rare. Dans ces secteurs, la demande et l'offre fluctuent de manière imprévisible, et si une demande dépasse la capacité disponible et ne peut être rationnée efficacement, elle génère des pertes de bien-être significatives. L'électricité est le cas d'école, mais d'autres marchés présentent des caractéristiques similaires : les transports avec congestion ou les biens médicaux pour lesquels la rareté de l'offre peut conduire à la propagation d'une maladie et à la congestion des installations médicales. Plus précisément, je développe un programme de recherche qui examine comment nous pouvons mettre en œuvre et concevoir des institutions, telles que les marchés, qui répondent efficacement à la demande et à l'offre de ces investissements. Cette recherche découle du fait que les décideurs politiques et les consommateurs accordent une grande importance à l'investissement et à la capacité de production, il existe une demande sociale significative pour mettre en lumière ces questions. Le reste de la conclusion générale se compose d'un résumé des trois chapitres.

Chapitre 1 - Comment offrir sur les marchés de capacité et quelle est la relation avec les caractéristiques du produit de capacité ?

Résultat Principal

L'efficacité du marché de capacité dépendant du signal de prix qu'il envoie, les concepteurs des marchés de capacité ne doivent pas négliger le choix des caractéristiques du produit de capacité et, notamment, de la période de disponibilité. En particulier, nous montrons dans ce chapitre que lorsque le risque et la flexibilité sont pris en compte, le résultat peut être différent de celui obtenu avec les hypothèses de modélisation traditionnelles.

Cette recherche part d'une question simple mais fondamentale : "Comment les prix émergent-ils sur les marchés de capacité ? Les prix des capacités constituent un revenu supplémentaire pour les producteurs afin de les encourager à augmenter leur capacité disponible. Ils constituent également une charge supplémentaire pour les consommateurs qui doivent payer pour cette augmentation de capacité. Il est donc essentiel de connaître le bon prix de capacité pour mesurer l'efficacité d'un marché de capacité. La théorie économique souligne depuis longtemps que les prix peuvent différer de leurs fondamentaux. En effet, tout prix est la conséquence d'une intersection entre l'offre et la demande. Sur les marchés de capacité, cette courbe d'offre est l'agrégation des offres de différents acteurs ayant des caractéristiques différentes. Il existe donc toujours une marge de manœuvre importante pour s'écarter du prix efficient.

Cette observation appelle une deuxième question : "Comment les producteurs offrent-ils sur les marchés de capacité ? Je réponds à cette question en affirmant que la participation à un marché de capacité implique un coût d'opportunité marginal spécifique pour les producteurs, ce qui nécessite une méthode d'évaluation particulière. En effet, la vente de capacité implique que le producteur reste disponible sur des périodes futures déterminées dans le contrat vendu. Cette promesse d'offre génère des gains et des pertes spécifiques pour le producteur. Dans ce cas, il doit échanger les revenus escomptés sur la durée de l'approvisionnement, déduction faite du coût fixe associé à la décision de rester disponible. La première contribution de ce chapitre consiste à étudier l'offre comme une valeur d'option associée à la possibilité de fermer temporairement mais irréversiblement

pour éviter les coûts fixes. Cette approche diffère sensiblement de l'approche standard de la valeur actuelle nette, dans laquelle le producteur n'offre que le coût d'opportunité attendu. Elle nous permet notamment de reconnaître la flexibilité de gestion d'un investissement qui peut réagir aux conditions futures du marché. Je propose donc une nouvelle approche pour conceptualiser les offres de capacité en utilisant la théorie des options réelles, où le coût d'opportunité est représenté comme une option sur le revenu qui détermine la rentabilité de l'usine. Tout d'abord, je définis une offre sur un marché de capacité à une période comme une option européenne. Ensuite, je passe à un cadre multipériodique dans lequel les offres de capacité peuvent être évaluées comme une option à panier modifiée.

J'utilise cette nouvelle approche pour répondre à une question concrète sur la conception du marché : "La durée du contrat vendu sur le marché de capacité a-t-elle une importance ?" Une analyse des différentes mises en œuvre des marchés de capacité dans le monde a révélé une grande variété de durées, allant de quelques semaines à plusieurs années. L'hypothèse essentielle de ce chapitre est la suivante : si la durée du contrat modifie le bénéfice escompté associé à la promesse de rester disponible, elle devrait modifier le coût marginal de la participation au marché de capacité. Par conséquent, la courbe d'offre change également, ce qui, en fin de compte, a une incidence sur la formation des prix. Par conséquent, la conception du produit vendu sur un marché de capacité peut définir, en partie, son efficacité.

Les principaux résultats de ce chapitre sont les implications de la modification de la durée des contrats sur le comportement des producteurs sur un marché de capacité, en supposant que le coût marginal soit analysé sous l'angle de la théorie des options réelles. Je constate que cela diffère considérablement du cadre de la valeur actuelle nette. Premièrement, les offres sont toujours plus élevées dans le cadre des options réelles, ce qui signifie que les producteurs accordent une valeur positive à la possibilité de fermer pour éviter certains coûts. Deuxièmement, les facteurs à l'origine des offres ont des effets différents sur leur valeur par rapport au cadre de la valeur actualisée nette. Je fournis une statique comparative sur la valeur des offres et la différence entre les deux cadres. Je constate que la durée du contrat augmente constamment l'offre lorsque l'on utilise la théorie des options réelles, alors qu'elle a un effet ambigu sur l'offre de valeur actualisée nette. La volatilité du marché de gros et le temps d'attente entre la vente du contrat et le début de la disponibilité sont également analysés. Ils ont tous deux des effets ambigus sur les offres de capacité en fonction d'un ensemble de conditions sur les facteurs d'offre. Enfin, je constate l'effet inverse pour la dimension de la conception du produit, avec

une offre plus élevée pour une durée de contrat plus longue que la somme des offres attendues pour des produits plus courts. Je teste les résultats en calibrant le modèle sur le système électrique français. J'utilise également des données réalisées pour une centrale électrique CCGT (gaz) afin de simuler une offre sur le marché de capacité et de comparer les résultats avec les prix réalisés observés sur le marché de capacité français. Sur la base du cadre théorique et de l'illustration numérique, je propose enfin une discussion pour la conception des marchés de capacité : (i) Sur le rôle de la pénalité sur le marché de capacité imposée aux producteurs qui choisissent de ne pas être disponibles pendant la durée de l'appel d'offres bien qu'ils aient vendu un produit de capacité. (ii) La différence entre les capacités existantes et les nouvelles capacités participant aux marchés de capacité.

Le principe de ce chapitre est de jeter les bases d'une meilleure compréhension de la manière dont les producteurs proposent leurs investissements sur les marchés de capacité. En particulier, en décrivant la relation entre leurs offres et les caractéristiques des contrats échangés sur le marché, nous obtenons une meilleure vision de ce qui devrait émerger sur les marchés de capacité et des compromis qui existent en termes de conception du marché. Ce chapitre ouvre également un certain nombre de portes pour des recherches futures. La première consiste à intégrer ces résultats dans un cadre plus systémique. En particulier, il faudrait concilier l'approche d'évaluation du coût marginal avec un processus de formation des prix. En d'autres termes, on pourrait modéliser un ensemble de producteurs hétérogènes et construire des courbes d'offre en utilisant la théorie des options réelles. L'ajout de décisions dynamiques séquentielles permettrait également d'endogénéiser l'ensemble des valeurs et, en particulier, d'analyser plus finement les décisions d'investissement. Une calibration plus fine des paramètres du modèle permettrait également une meilleure comparaison entre ce qui est attendu du point de vue des fondamentaux du marché et le comportement réel des agents. Une deuxième approche consiste à prendre en compte la demande sur les marchés de capacité. En effet, l'ensemble du chapitre se concentre sur la valeur d'un investissement pour un producteur et du point de vue de la théorie des options réelles. Mais une question se pose : quelle est la valeur d'une capacité disponible pour les consommateurs ? On pourrait alors réaliser une étude similaire, en analysant la contrepartie de cette valeur d'option mais pour la demande. La valeur pour la société d'une capacité supplémentaire est donc la conjonction entre la valeur de l'option de demande et la valeur de l'option d'offre.

Chapitre 2 - Les conséquences indirectes de la conception de la demande

Résultat Principal

Ce chapitre montre que la façon dont la conception de la demande sur les marchés de capacité peut avoir un impact sur l'ensemble du système. Nous démontrons notamment qu'elle peut modifier la façon dont les consommateurs achètent l'électricité, ce qui peut avoir des conséquences sur le surplus du consommateur et, indirectement, sur les décisions d'investissement. Selon les caractéristiques du système, ces effets indirects peuvent avoir des conséquences positives ou négatives en termes de bien-être.

La motivation initiale de ce chapitre est de décrire l'arbitrage entre un marché de capacité centralisé ou décentralisé. Dans cette thèse, un marché centralisé est caractérisé par le fait qu'un seul acheteur, généralement une entreprise réglementée, telle qu'un opérateur de système, achète la quantité totale de capacité. D'autre part, un marché décentralisé est basé sur la participation obligatoire de différents acteurs du côté de la demande, principalement des fournisseurs, qui achètent eux-mêmes de la capacité en fonction de leur consommation prévue. Entre les deux, il existe des modèles de marché hybrides dans lesquels les consommateurs et les fournisseurs participent indirectement aux marchés de capacité. Le degré de participation des consommateurs est appelé la conception de la demande des marchés de capacité. Cette recherche découle du fait que la demande de capacité nécessite une intervention réglementaire. Alors que l'offre émerge naturellement sur ces marchés, la nature de bien public de l'investissement pendant les périodes de forte demande implique que les consommateurs ne sont pas disposés à acheter des capacités sur les marchés de capacité. Le régulateur doit donc définir administrativement la fonction de la demande pour que le marché s'équilibre et fournisse les prix des capacités des producteurs.

Ce chapitre vise à jeter les bases permettant de comprendre les avantages et les inconvénients des différentes options qui s'offrent au régulateur. En particulier, il montre les conséquences des différentes options sur un modèle représentant le marché des capacités, les décisions d'investissement, la production et l'équilibre de la consomma-

tion. Enfin, le chapitre décrit le bien-être d'un marché de l'électricité en fonction de différentes hypothèses sur la conception du marché, le comportement des acteurs et l'environnement. Je me concentre sur deux questions interdépendantes qui concernent (i) le régime d'allocation des coûts, c'est-à-dire la manière dont un acheteur unique répartit le prix de la capacité entre les acheteurs de capacité et les consommateurs finaux, et (ii) la mesure dans laquelle la demande réalisée par les consommateurs finaux est prise en compte dans la conception de l'allocation du marché.

La méthodologie utilisée pour répondre à cette question repose sur la théorie canonique de la tarification de pointe pour un bien homogène caractérisé par une demande variable dans le temps, qui décrit la relation entre la production à court terme et les décisions d'investissement à long terme. Le modèle élargit la littérature en fournissant un nouveau cadre analytique qui inclut un équilibre du marché de capacité en plus des décisions d'investissement et de court terme. La proposition de modélisation est centrale car tout effet indirect généré par le marché de capacité peut affecter le revenu attendu par les producteurs et peut être indirectement pris en compte lors de la formation de la fonction d'offre sur le marché de capacité.

Le premier régime de marché étudié est le marché de capacité canonique. Je m'appuie sur la littérature précédente, qui repose sur l'hypothèse qu'un marché de capacité n'a pas d'autre effet que d'augmenter le niveau d'investissement. Ce régime canonique est similaire à un régime d'allocation des coûts basé sur une taxe forfaitaire. Dans ce cas, même si l'on considère la fonction d'offre endogène sur le marché de capacité, l'équilibre de la conception du marché rétablit toujours l'optimum compte tenu des inefficacités du système. J'étudie ensuite le cas où le prix de la capacité a un impact à la marge sur les consommateurs. Dans cette hypothèse, le régime alloue le prix de la capacité comme une taxe unitaire. Je montre que l'existence du marché de capacité affecte indirectement le marché de gros en redistribuant les différents états du monde lorsque la capacité n'est pas contraignante ou au contraire lorsqu'elle contraint, et en diminuant le surplus du consommateur. Par conséquent, je démontre que le bien-être à l'équilibre sous ce régime est toujours inférieur à celui du régime canonique. Je compare ensuite les deux régimes d'allocation des coûts de capacité en incluant un rationnement inefficace. Lorsqu'un plafond de prix est atteint, la disponibilité des investissements devient un bien public car la demande devient inélastique. En raison de l'impossibilité de rationner efficacement les consommateurs, ceux-ci subissent une perte de bien-être significative. Sous cette nouvelle hypothèse, je constate que l'effet indirect créé par l'attribution du prix de la capacité sur une base unitaire est désormais ambigu pour le bien-être social.

J'élargis l'analyse à la mise en œuvre d'un régime dans lequel le régulateur répartit le coût sur la base des parts de marché réelles des fournisseurs. Je montre d'abord comment ce modèle affecte marginalement les détaillants qui jouent "à la Cournot" sur le marché de détail. Ensuite, j'intègre le nouvel équilibre dans le modèle avec les décisions d'investissement et le marché de capacité. Je constate que cette répartition crée un résultat intermédiaire entre la taxe unitaire et la taxe forfaitaire. Enfin, j'analyse le cas d'un marché de capacité entièrement basé sur le niveau de demande réalisé. Pour ce faire, les détaillants sont contraints de couvrir la quantité vendue sur le marché de détail en achetant directement sur le marché de capacité, compte tenu d'un système de pénalités. Je me concentre sur la manière dont les stratégies des détaillants peuvent former une fonction de demande agrégée sur le marché de capacité, et j'analyse la capacité optimale achetée par les détaillants sur le marché de capacité. Je constate qu'une telle approche de la fonction de demande peut fournir le niveau optimal d'investissement dans des conditions spécifiques.

Ce chapitre développe une série d'extensions du modèle canonique représentant les décisions d'investissement et de production sur les marchés de l'électricité. Il commence par le cas le plus simple, avec un plafond de prix et un acheteur unique. Ensuite, j'ai formulé davantage d'hypothèses concernant les inefficacités et le comportement des acteurs. Plutôt que de fournir une comparaison claire entre un marché centralisé et un marché décentralisé, le chapitre cherche à décrire comment nous pouvons interpréter ces modèles en premier lieu. En effet, un marché centralisé est avant tout une question de répartition des coûts, tandis qu'un marché décentralisé dépend de la valeur d'un investissement pour la demande. En gardant ces résultats à l'esprit, je pense qu'il est désormais possible de procéder à une comparaison plus approfondie. Comme le soulignent certains documents qualitatifs, les régimes peuvent différer principalement en raison des informations différentes que chaque agent peut posséder. Par exemple, les détaillants peuvent disposer de meilleures informations concernant leur portefeuille de consommateurs. D'autre part, l'acheteur unique peut disposer de meilleurs outils pour prédire l'évolution de la demande globale future. Par conséquent, l'acheteur unique est susceptible de commettre des erreurs de prévision lors de l'agrégation des informations. Néanmoins, le processus d'agrégation d'informations dispersées par le biais d'un mécanisme décentralisé laisse également une grande marge de manœuvre aux détaillants pour se comporter de manière inefficace. Un autre domaine de recherche futur consisterait à approfondir la représentation de la conception du marché de capacité décentralisé, que les économistes de l'électricité n'ont pas modélisé quantitativement. Les progrès récents de la théorie des

jeux et de l'organisation industrielle pourraient apporter des éclaircissements significatifs sur ce que signifie négocier et échanger des produits de capacité entre les producteurs et les consommateurs.

Chapitre 3 - Information privée, consommateurs et décisions d'investissement

Résultat Principal

Il existe un arbitrage fondamental lorsque l'on conçoit des marchés générant des revenus pour assurer un niveau d'investissement et maximiser le surplus du consommateur : les prix permettent de couvrir les coûts d'investissement tout en guidant les décisions de consommation. Dans ce chapitre, nous étudions cette question en supposant que l'utilité dérivée de la consommation est partiellement inconnue. Il examine comment un concepteur de marché peut choisir les prix et prendre des décisions d'investissement en fonction d'un ensemble de contraintes. En particulier, ce chapitre montre que, sous certaines hypothèses, atteindre un certain niveau d'investissement peut entraîner des problèmes de répartition.

La plupart des discussions sur la conception des marchés de capacité et, plus généralement, sur le niveau d'investissement dans les marchés de l'électricité ont été centrées sur des questions relatives à l'offre. Il s'agit de savoir comment donner les bonnes incitations aux producteurs pour qu'ils atteignent le niveau souhaité. Cependant, il semble tout aussi essentiel de savoir sur quelle base ce niveau doit être déterminé. Cette question de recherche est étroitement liée aux deux chapitres précédents, car elle reflète la question de la valeur de la demande de capacité mentionnée au chapitre 1, et constitue une première étape dans la construction de la fonction de demande du chapitre 2.

La principale contribution de ce chapitre est d'examiner les implications de la prise en compte de la demande lorsqu'il s'agit de garantir un niveau d'investissement efficace. Il étudie un environnement dans lequel il existe un certain nombre de consommateurs ayant des caractéristiques différentes et dont l'utilité de la consommation d'électricité est en partie une information privée. En d'autres termes, le chapitre cherche à savoir comment concevoir les marchés de manière à ce que les consommateurs révèlent leur

utilité pour leur consommation d'électricité et, par conséquent, indirectement pour le niveau d'investissement. Dans cette recherche, les marchés ne sont pas seulement utilisés pour fournir de la consommation et des revenus suffisants pour réaliser des investissements, mais aussi pour filtrer les caractéristiques non observables afin de garantir un niveau d'investissement adéquat. Une deuxième approche que j'aborde dans ce chapitre concerne les questions de redistribution associées à la conception d'un marché optimal dans ce contexte. Plus précisément, je montre que la mise en œuvre du marché le plus efficace pour atteindre le niveau d'investissement n'est pas toujours une amélioration de Pareto pour chaque consommateur.

Pour ce faire, j'étudie les tensions inhérentes à la mise en œuvre d'un mécanisme d'allocation qui (i) dicte la manière dont les agents consomment les biens et (ii) génère des revenus pour financer de nouveaux investissements dans un cadre d'information incomplète. Je développe un cadre théorique dans lequel un concepteur de marché choisit séquentiellement un niveau d'investissement et propose un mécanisme d'allocation aux consommateurs, suivi d'une phase de consommation. Le mécanisme d'allocation définit le transfert monétaire par unité et la quantité pour un ensemble de consommateurs au cours de la phase de consommation, sous réserve d'une contrainte de capacité. Le concepteur du marché utilise l'allocation pour maximiser le surplus du consommateur et financer le coût de l'investissement ; il est donc également soumis à une contrainte budgétaire. Je suppose que lorsque le concepteur du marché prend des décisions d'investissement et d'allocation, il est confronté à des consommateurs hétérogènes qui disposent d'informations privées sur leur niveau de demande et qui appartiennent à une catégorie observée publiquement.

Le premier cas d'information incomplète représente les inefficacités associées à l'information privée lorsque le concepteur du marché est contraint dans son choix d'allocation. Dans ce cas, certains consommateurs surconsomment par rapport à leur type, tandis que d'autres sousconsomment. J'étudie tout d'abord le cas du prix unique lorsque le concepteur du marché ne peut pas faire de distinction entre les catégories de consommateurs. Dans ce cas, le barème de prix augmente avec le niveau d'investissement. Je mets ensuite en œuvre la possibilité de discriminer entre les catégories de consommateurs. Je montre que pour la catégorie de consommateurs dont le niveau de demande moyen est le plus faible, le prix optimal diminue d'abord et augmente ensuite avec le niveau d'investissement. En revanche, le prix pour la catégorie des consommateurs plus importants augmente toujours avec la capacité. Je constate également que le prix de la première catégorie est supérieur à celui de la seconde pour des valeurs de capacité relativement faibles, puis que le classement s'inverse pour des valeurs plus élevées. Ces

non-monotonicités peuvent s'expliquer par l'effet opposé auquel le concepteur du marché est confronté en termes de surplus du consommateur et d'effet de revenu lorsqu'il choisit les prix.

Dans la dernière section, j'étudie une configuration de mechanism design où le concepteur de marché n'est plus contraint dans le calendrier des prix et des quantités qu'il peut offrir aux consommateurs. Il doit maintenant faire face à des contraintes d'incitations et de participation. Je décris d'abord pour quel niveau d'investissement le concepteur de marché est contraint par le revenu utilisé pour couvrir les coûts fixes et la rente d'information qu'il doit fournir aux consommateurs pour qu'ils se comportent de manière sincère. Je constate que le concepteur de marché ne peut fournir une allocation de premier choix sans contrainte que pour des valeurs faibles du niveau d'investissement, en raison de la concavité de l'utilité virtuelle du consommateur par rapport à la capacité. Pour des valeurs plus élevées, les gains d'utilité supplémentaires attendus de l'augmentation de la capacité ne peuvent pas compenser les coûts d'investissement. Ensuite, je montre que le comportement de l'allocation optimale dépend de l'état du monde considéré et du type de consommateurs. Plus précisément, lorsque le niveau d'investissement augmente, les consommateurs sont toujours confrontés à une diminution de la quantité optimale allouée pendant les heures creuses. Pour les périodes de pointe, le changement de quantité dépend du type de consommateur. Enfin, je montre que seuls les gros consommateurs des deux catégories peuvent bénéficier d'une augmentation du niveau d'investissement. Ce résultat s'explique par le lien entre le niveau d'investissement et la rente d'information accordée aux consommateurs pour qu'ils adoptent un comportement sincère.

Cette recherche a plusieurs extensions. Pour commencer, ce travail a principalement dérivé une limite supérieure et inférieure sur les conséquences des différentes conceptions du marché. La première série de résultats dans l'environnement d'information incomplète décrit la limite inférieure *actuelle* dans laquelle le concepteur du marché est fortement contraint et ne peut extraire aucune information. En revanche, le deuxième résultat décrit la limite supérieure *théorique*. En effet, en présence d'informations privées et de consommateurs stratégiques, il s'agit du meilleur modèle de marché que le concepteur peut mettre en œuvre s'il cherche à ce que les consommateurs révèlent leurs types. Toutefois, une grande partie de la littérature sur la conception des marchés et des mécanismes a montré que la solution théorique ne peut pas toujours être mise en œuvre, soit pour des raisons techniques, en raison de la non-linéarité de l'allocation optimale par exemple, soit pour des raisons d'acceptation sociale, étant donné que l'allocation optimale implique une discrimination. Une autre recherche pratique consisterait à tester le comportement

de certaines formes de marchés de capacité en termes de révélation d'informations. En d'autres termes, il serait possible de construire un modèle en ajoutant au chapitre un module décrivant les mécanismes à long terme. Dans ce cas, le concepteur du marché serait confronté à un ensemble spécifique de contraintes, et la résolution d'un tel modèle permettrait une comparaison avec les limites supérieures et inférieures décrites dans ce chapitre. Enfin, le chapitre souligne l'existence de problèmes de redistribution lors de la mise en œuvre du plan d'allocation pour atteindre un certain niveau d'investissement. Il existe donc une piste intéressante qui consisterait à étudier l'arbitrage associé à la restriction sur cette allocation pour éviter les problèmes de redistribution. À première vue, cet arbitrage consisterait à réduire les revenus disponibles pour de nouveaux investissements. Cependant, dans un contexte de difficultés croissantes à financer la transition énergétique, je pense que cette recherche est d'un intérêt significatif.

Bibliography

- Fabra, N. and Llobet, G. (2023). Auctions with privately known capacities: Understanding competition among renewables. *The Economic Journal*, 133(651):1106–1146.
- FED, D. (2021). Cost of texas' 2021 deep freeze justifies weatherization.
- IEA, P. (2020). Power systems in transition. Technical report, IEA.

Contents

| | |
|--|------------|
| Remerciements | iii |
| Résumé en français | vii |
| General Introduction | 1 |
| 1 The role of investments in electricity markets. | 3 |
| 1.1 The adequacy problem | 5 |
| 1.2 Spot Pricing: a theoretical solution | 7 |
| 1.3 The limit to the Spot Pricing theory | 13 |
| 2 Economic instruments for capacity adequacy | 19 |
| 2.1 Theoretical foundation | 19 |
| 2.2 Long-term mechanisms | 21 |
| 3 A case for capacity markets: principles and limits | 22 |
| 1 Bidding Behavior and Product Definition in Capacity Markets | 37 |
| 1 Introduction | 42 |
| 2 Model assumptions | 49 |
| 2.1 Investment and wholesale market | 49 |
| 2.2 Producer behavior | 51 |

| | | |
|----------|---|-----------|
| 2.3 | The capacity market | 52 |
| 3 | The net present value framework | 53 |
| 3.1 | The inframarginal rent and the opportunity cost | 54 |
| 3.2 | A simple example | 56 |
| 4 | The real options framework | 58 |
| 4.1 | A capacity bid as a European Put Option | 58 |
| 4.2 | A capacity bid as a Basket Option | 59 |
| 4.3 | The sum of expected capacity bids | 60 |
| 5 | Comparative statics | 61 |
| 5.1 | Bids value | 62 |
| 5.2 | Flexibility value | 67 |
| 5.3 | Product design | 69 |
| 6 | Case study | 71 |
| 6.1 | Bids in a capacity market | 73 |
| 6.2 | The effect of the waiting time and the volatility | 75 |
| 6.3 | Product design | 77 |
| 7 | Extensions | 79 |
| 7.1 | Bids for new entrants | 79 |
| 7.2 | Penalty and mothballing costs | 80 |
| 8 | Conclusion | 82 |
| 2 | Designing the Demand Side of Capacity Markets | 89 |
| 1 | Introduction | 93 |
| 2 | Benchmark model with capacity market | 98 |

| | | |
|----------|--|------------|
| 2.1 | Environment | 98 |
| 2.2 | Market equilibrium with a capacity market | 100 |
| 2.3 | First-best without a capacity market | 105 |
| 2.4 | Second-best without a capacity market | 106 |
| 3 | Allocation under a price cap | 107 |
| 3.1 | Exogenous allocation | 107 |
| 3.2 | Endogenous allocation | 108 |
| 4 | Allocation with inefficient rationing | 115 |
| 4.1 | Second-best investment and welfare | 115 |
| 4.2 | Exogenous vs. Endogenous allocation | 121 |
| 5 | Extensions - Retailers allocation | 126 |
| 5.1 | Retailers Market Share allocation | 127 |
| 5.2 | Decentralized capacity market | 133 |
| 6 | Conclusion | 139 |
| 3 | Designing Markets for Reliability with Incomplete Information | 147 |
| 1 | Introduction | 151 |
| 2 | Environment | 156 |
| 2.1 | Consumers Preferences | 157 |
| 2.2 | Allocation design | 158 |
| 2.3 | Supply side | 159 |
| 2.4 | Timing | 160 |
| 3 | Complete Information | 161 |
| 3.1 | Optimal allocation proposal | 161 |

| | | |
|-----|---|------------|
| 3.2 | Long-term vs. Short-term consumer surplus | 164 |
| 4 | Incomplete Information - Fixed price | 172 |
| 4.1 | Single price policy | 174 |
| 4.2 | Category-price policy | 177 |
| 4.3 | Individual welfare effect | 183 |
| 5 | Incomplete Information - Mechanism Design | 186 |
| 5.1 | Optimal allocation | 186 |
| 5.2 | Optimal allocation and investment level | 190 |
| 5.3 | Individual welfare effect | 197 |
| 6 | Conclusion | 202 |
| | General Conclusion | 207 |
| | A Literature review - Conceptualizing Behavior in Capacity Markets | 223 |
| 1 | Description of the literature review | 226 |
| 2 | Conceptual definition of actor behavior | 226 |
| 2.1 | True value and offered value | 227 |
| 2.2 | The opportunity cost of capacity markets | 228 |
| 3 | Litterature review of the <i>true value</i> | 230 |
| 3.1 | Risk and uncertainty | 231 |
| 3.2 | Industrial decisions | 233 |
| 3.3 | Product definition | 237 |
| 4 | Litterature review of the <i>offered value</i> | 239 |
| 4.1 | Market power and asymmetric information | 240 |

| | | |
|-------------------------------|---|------------|
| 4.2 | Auction theory | 242 |
| 5 | Litterature reviews of market design issues in capacity markets | 243 |
| 5.1 | Sequentiality of transactions | 243 |
| 5.2 | Decentralization of demand | 245 |
| 5.3 | Price multiplicity | 247 |
| 5.4 | Transaction decentralization | 248 |
| 5.5 | The time horizon covered by the traded product | 250 |
| 6 | Conclusion | 253 |
| B Appendix - Chapter 1 | | 263 |
| 1 | Proof of Proposition 1.2 | 265 |
| 2 | Proof of Proposition 1.3 | 267 |
| 3 | Proof of Proposition 1.5 | 269 |
| 4 | Proof of Lemma 1.1 | 271 |
| 5 | Proof of Proposition 1.7 | 273 |
| 6 | Proof of Lemma 1.4 | 274 |
| C Appendix - Chapter 2 | | 277 |
| 1 | Market equilibrium with Cournot | 279 |
| 2 | Proof of Lemma 2.1 | 285 |
| 3 | Proof of Proposition 2.1 | 285 |
| 4 | Proof of Lemma 2.2 | 286 |
| 5 | Proof of Lemma 2.3 | 288 |
| 6 | Proof of Proposition 2.2 | 291 |

| | | |
|----|------------------------------------|-----|
| 7 | Proof of Lemma 2.4 | 295 |
| 8 | Proof of Lemma 2.5 | 296 |
| 9 | Proof of Proposition 2.3 | 299 |
| 10 | Proof of Lemma 2.6 | 302 |
| 11 | Proof of Proposition 2.4 | 304 |
| 12 | Proof of Corollary 2.1 | 307 |
| 13 | Proof of Lemma 2.7 | 309 |
| 14 | Proof of Proposition 2.5 | 311 |

D Appendix - Chapter 3 315

| | | |
|----|------------------------------------|-----|
| 1 | Proof of Proposition 3.1 | 317 |
| 2 | Proof of Proposition 3.2 | 321 |
| 3 | Proof of Corollary 3.1 | 324 |
| 4 | Proof of Lemma 3.1 | 325 |
| 5 | Proof of Proposition 3.3 | 327 |
| 6 | Proof of Proposition 3.4 | 332 |
| 7 | Proof of Lemma 3.2 | 345 |
| 8 | Proof of Proposition 3.5 | 348 |
| 9 | Proof of Proposition 3.6 | 351 |
| 10 | Proof of Lemma 3.3 | 357 |

General Introduction

Reading guide for the thesis. It is organized as follows. The general introduction presents the economic fundamentals of the subject addressed in the dissertation. A second introductory chapter in the Appendix proposes a frame of thought around a literature review that presents a more detailed methodology and a general philosophy. The main body of the thesis is composed of three chapters. Each chapter is built in such a way as to be self-explanatory. Finally, the general conclusion summarizes the cross-cutting takeaway and the three main chapters. The rest of the Appendix at the end of the manuscript describes the technical details and the proof of each chapter.



This introduction aims to describe the general economic framework in which this thesis is situated and its motivation. First, I will recall why it is crucial to study investment decisions in electricity markets. Due to its intrinsic nature and also because additional capacities are needed for a successful energy transition, economists need to make sure that we implement the right institutions to provide them at the right cost. I will notably summarize what the benchmark model is to think about how private agents can provide investment, but also why this canonical framework might fail. Then, in the face of those limitations, I will describe diverse policy responses to the need to ensure enough available investment. Due to the vast qualitative and applied literature that already exists on the subject, I will focus on a theoretical and conceptual discussion.

1 The role of investments in electricity markets.

Electricity is at the heart of our modern society. It is the backbone of economic development, providing the energy required to fuel industries, businesses, and households.

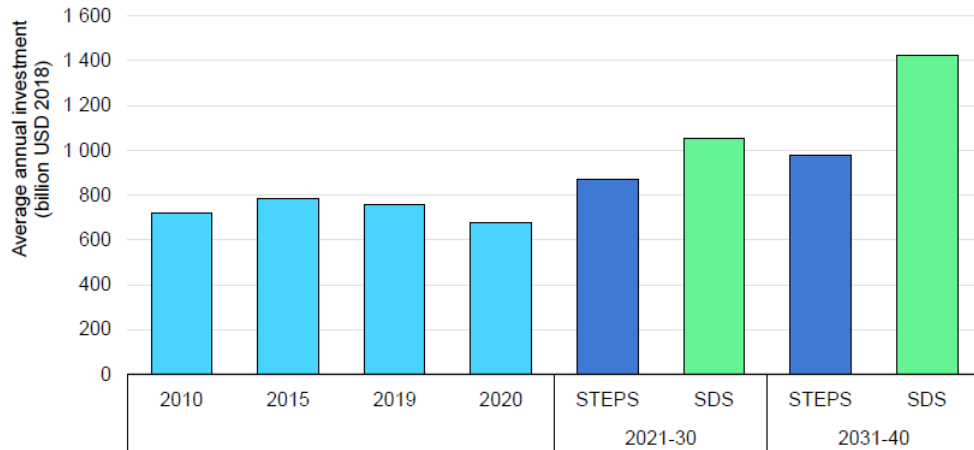


Figure 3: Average annual global power sector investment, 2010-20, and needs to 2040. (IEA, 2020)

Despite its central role in the economy, and whatever the mode of governance envisaged, electricity faces a significant physical constraint. At any given point in time, the quantity consumed must be equal to the quantity injected into the system. Consequently, the challenge of always having sufficient production capacity to meet a given demand is crucial whenever storage is insufficient.³ In this context, **how can we ensure that the balance between production and consumption is maintained as efficiently as possible?** This issue is all the more important as electricity is set to play a central role in the energy transition, whether on the production side, due to the necessary transition from carbon-based production to decarbonized production, or on the consumption side, due to the electrification of our uses such as transport, heating, or the digitization of the economy. Figure 3 illustrates, for instance, the investment needed under two scenarios at the 2040 horizons compared to realized investments in the last decade.

In other words, the need to set up effective institutions to ensure sufficient investment stems not only from the very nature of electricity but also from its central role in the energy transition. I will discuss in more detail in the rest of the section the importance of investment in electricity, how markets can theoretically address this question, and why it is essential to keep thinking about the proper institutional framework to make sure that those questions are dealt with efficiently.

³Imports also play an essential role in power markets. In this thesis, I do not make any distinction in the stylized model between production capacity and imports.

1.1 The adequacy problem

The constraint described above translates into the obligations of the power system to deliver electricity to consumers reliably ("a good quality service" (Petitet, 2016)). This obligation takes the form of a constraint characteristic of the electricity sector: *reliability*. It is technically defined as :

"A general term encompassing all the measures of the ability of the system to deliver electrical energy to all points of utilization within acceptable standards and in the amounts desired" (CIGRE, 1987)

This term encompasses both a short-term horizon concerning the assurance that, at any given moment, the quantities of electricity injected and withdrawn are equal. This obligation, or *security of supply*, is defined as an operational constraint:

"the measure of how an electric power system can withstand sudden disturbances such as electric short circuits or unanticipated loss of system components" (CIGRE, 1987);

And a long-term horizon *adequacy*, which is the system's ability to support a certain level of consumption:

"of a bulk power system to supply the aggregate electric power and energy requirements of the customers within component ratings and voltage limits, taking into account scheduled and unscheduled outages of system components and the operating constraints imposed by operations" (CIGRE, 1987).

Even if the time horizons differ, these constraints both raise the question of the sizing and composition of the generation and consumption mix. Both need to be able to adapt to medium/long-term changes in demand (*Adequacy*), but also to withstand extreme events, mainly to avoid short-term imbalances (*Security of Supply*). Moreover, there is also a wide variety of technologies for generating electricity. Frequent changes in electricity consumption mean that at certain times, to meet a specific demand, certain technologies are considered more economically and technically appropriate. The role of each technology in the system is defined by its overall characteristics, in particular

its cost structure and operational functioning. To be economically efficient and meet all demands, a capacity must, therefore, have a specific technological composition. An investment is capital-intensive and has a relatively long lifespan. This temporal and financial dimension of investment requires a long-term approach. Therefore, having the right institution with respect to investment in this context covers a wide range of issues: (1) from a static perspective, efficiency can be defined by having the right level and composition, and (2) from a dynamic perspective, one should make sure that decisions are made in a timely manner.⁴ **In this thesis, I will focus on the efficiency of the mix level rather than the composition or the dynamic efficiency.**

The lack of sufficient investment to address this adequacy issue results in three main types of economic damage (IEA, 2020):

- Load shedding
- Cascading blackouts/black system events
- Long rationing periods of electricity

The first type of consequence arises when the relative level of investment is not significantly different from the level of consumption. In this case, the economic cost appears as an involuntary rationing of demand. As I will discuss in greater detail throughout Chapter 2, the main idea behind this involuntary rationing of demand is that, due to a number of physical, economic, and political constraints, it is not possible to effectively reduce demand, i.e., to individually target those who should in theory reduce or stop consuming. This impossibility, in fact, generates a welfare cost. A recent example of this is California in 2020 or Texas in 2021 (IEA, 2020). Figure 4 illustrates the Texas extreme winter weather events that lead to severe outages. The second type is directly linked to the physical characteristics of electricity when the gap between production and consumption is significant, and load shedding is not sufficient. In this case, all the agents on a network suffer damage, ranging from the shutdown of transport systems to the shutdown of the cold chain. Finally, when the level of investment is inadequate from a long-term point of view, the power system needs to limit electricity consumption over the long term. The damage incurred is then directly linked to the central function of electricity in economic development, and it causes damage right up to the macroeconomic level.

⁴In particular, entry and exit decisions, including the decision to maintain or mothball certain capacity.

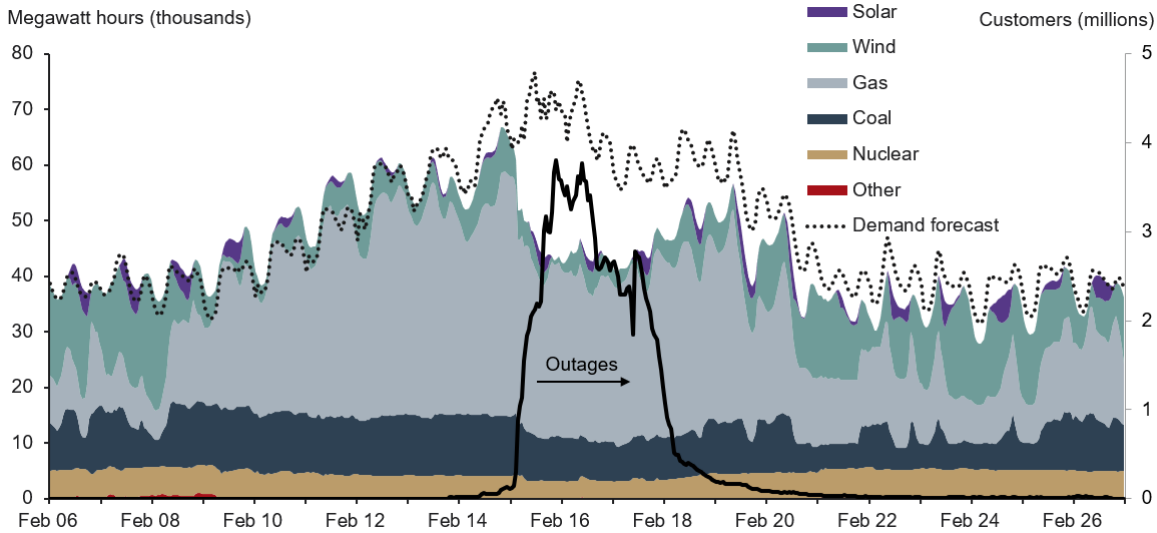


Figure 4: ERCOT electricity generation by source, demand, and outages during Texas Deep Freeze (FED, 2021)

This thesis aims to understand how specific public policies address investment issues in the context of managing constraints of electricity markets. More specifically, I will study economic instruments that aim to provide private agents with sufficient incentives to manage the supply-demand balance adequately over the medium/long term. Although short-term security of supply issues are also crucial, they require a different approach to that used in this thesis. Consequently, these issues will not be addressed in the thesis.

1.2 Spot Pricing: a theoretical solution

Historically, power systems were vertically integrated. The investments and the transport of electricity to the end consumer were managed by a single company, very often controlled by the state. This was referred to as the "control-command" model: economic production decisions to meet demand were made by the regulated company, as were the technical adjustments made to deal with short-term equilibrium (Saguan, 2007). Investment in capacity, and therefore the choice and timing of the level and composition of the electricity mix, is also decided by the public entity and the company. Consequently, in this model, long-term equilibrium issues are handled by a single entity acting as a benevolent planner.

The transition from a vertically integrated model to one based on competition between several agents has challenged the way the power system operates. In this new

model, the production and investment decisions are no longer managed by a single entity but by a set of agents whose objective is to maximize their profit. To simplify, supply-side agents are investors who own generation assets and base their industrial decisions on estimated future profits. Demand-side agents are retailers with a portfolio of customers or relatively large customers who can buy directly in wholesale markets. Each actor has its characteristics: demand-side agents can be differentiated according to various consumption patterns, which can be measured in particular by their demand/price elasticity, i.e., their ability and willingness to modulate their consumption according to electricity prices. Supply-side agents are characterized by a diverse structure, with companies that are more or less vertically and horizontally integrated, of varying sizes, and with a more or less diversified portfolio of customers/technologies.

The initial economic theory associated with this transition is based on the principle that a single electricity market is identical to, or even more efficient than, a "control-command" model. This is the theory of spot pricing, developed in particular by [Caramanis et al. \(1982\)](#). One of the results of this theory is that the price of electricity solves short-term equilibrium problems by itself ([Littlechild, 1988](#)) and leads to an optimal generation mix ([Caramanis et al., 1982](#)).⁵ They show that agents perfectly replicate optimization models in which a benevolent entity minimizes production costs. In fact, agents in the energy market maximize the total surplus of the system, which is composed of equal weighting of consumer surplus and producer surplus. In other words, the market price of energy sends both the right short-term signals (on the operational dispatch of energy, i.e., which technology should produce according to its marginal cost and which actor should consume according to its marginal utility) and the right long-term signals (on the composition and optimal level of the production mix, and the usage patterns for the consumption mix).

It's important to note that even if marginal units set the price, the model implies perfect coverage of fixed costs. Indeed, from an individual point of view, it can be understood as follows: When demand is lower than the total capacity, the marginal production unit sets the price and generates a rent known as "inframarginal rent" for the most efficient producers. During periods of tension, which can be defined as the conjunction between the existence of consumers whose willingness to pay is higher than the marginal cost of the peak plant and the binding of production capacities that are

⁵From an implementation perspective, the corresponding mechanism of this theory is called the energy-only market. In this context, an energy market is a mechanism in which electricity is traded similarly to any other commodity with a price corresponding to a quantity of electricity.

therefore unable to satisfy this demand, the price rises to a level at which consumers prefer to stop consuming. During these periods, the entire production mix receives a rent known as "scarcity rent," which, combined with the inframarginal rent, covers all fixed costs. In a dynamic vision of the model, if the sum of the two rents does not cover fixed costs or, on the contrary, generates an over-profit, the model assumes free exit as well as free entry. The assumptions behind the model are the classic ones for a situation of pure and perfect competition: no transaction costs, perfect information, and atomicity of agents enabling continuous convergence towards an optimal equilibrium (Vassilopoulos, 2007).

Figure 5 is a simplified representation of each period during which all supply-side agents receive either the inframarginal rent or the scarcity rent. The first graph illustrates a situation outside a period of scarcity, where the price is set by the marginal cost of one of the technologies present in the system. At this point, the entire production mix is able to cover all demand. In the second graph, demand cannot be met by the available capacity. Consequently, the price is set at a level at which demand voluntarily decides to stop consuming. This adjustment equates consumption with the total installed capacity of the production mix. According to spot price theory, the quantity of each technology present in the system is established in anticipation of these two types of rent. More precisely, at equilibrium, the sum of these two rents for each technology must cover all fixed costs.

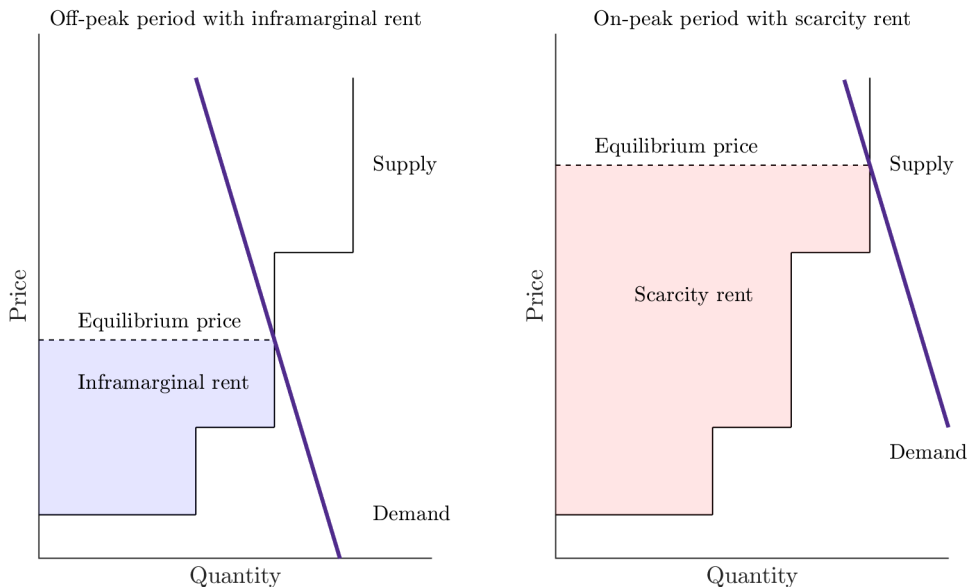


Figure 5: Rents in energy-only markets with off-peak and on-peak periods

In the following Box 1.2, I also provide a simple toy model from [Joskow and Tirole \(2007\)](#), illustrating how this peak load pricing theory works in a stylized framework. Note that it will also be the reference model used in Chapter 2 and Chapter 3.⁶ To summarize, studying the investment adequacy issue from a system point of view can be understood as a tradeoff between :

- The costs associated with the electricity not served. Under the toy model, it is simply the marginal gross consumer surplus net of production costs. With more complex representation, notably with price inelastic consumer, the value of this energy that should have been consumed under optimal technical conditions is measured by *VoLL* (*Value of Loss Load*) and the associated quantity by *ENS*, (*Energy Not Served*). In other words, VoLL can be understood as the price at which demand voluntarily stops consuming. and;
- The CONE ("Cost Of New Entry"), which corresponds to the cost of the most efficient peak power plant that would meet demand. Its formal definition is the income required to ensure that the annual profit from its participation in the energy markets is zero over its lifetime. Under the toy model, it is simply the investment cost net of expected profit.

In the first-best environment, market equilibrium then corresponds to equality between the marginal cost of additional capacity and the expected gain from avoided involuntary rationing (measured as its cost), with the LOLE being the probability that over a given period, demand will exceed installed capacity:

$$CONE = VoLL \times LOLE$$

⁶For a more detailed discussion of the dynamic equilibrium between the short-term production decision and the long-term investment decisions, see [Vassilopoulos \(2007\)](#).

Toy model for peak load pricing theory

Assume that there is a representative consumer of electricity. He is characterized by a price-sensitive demand $D(p, s)$, p being the electricity price, and s being the different states of the world (one can think about weather or macroeconomic shocks) such that $D_s(s) > 0$. It follows a distribution function $f(s)$. Let's also denote $CS(p, s)$ the consumer gross surplus.

To produce electricity, there exists a continuum of investment indexed by the marginal cost of production c . Let $I(c)$ denote the investment cost of a plant producing one unit of electricity at marginal cost c . $G(c) \geq 0$ will be the cumulative distribution function of plants. Therefore, the total investment cost is $\int_0^\infty I(c)dG(c)$, and production cost $\int_0^\infty u(c, s)cdG(c)$, where the utilization rate $u(c, s)$ belongs to $[0, 1]$ such that $\int_0^\infty u(c, s)dG(c) = q(s)$.

In that context, a benevolent social planner would maximize the welfare comprised of the gross consumer surplus net of the production cost:

$$\max_{p(s)} \int_s \left(CS(p, s) - \int_0^\infty u(c, s)cdG(c) \right) dF(s) - \int_0^\infty I(c)dG(c)$$

Such that $q(s) = D(p, s)$. The first-best solutions to this problem yield:

- The *short-term* solution: $u(c, s) = 1$ for $c < p(s)$ and $u(c, s) = 0$ for $c > p(s)$. Namely, only plants whose marginal cost is smaller than the price $p(s)$ are dispatched in the state s .
- The *long-term* solution: $I(c) = \int_s (p(s) - c)dF(s)$. Namely, the investment decision is such that the investment cost equals the expected rent made whenever the price is above the marginal cost.

Therefore, overcapacity can then be interpreted by a level of CONE that is too high or by a cost of avoiding rationing that is too high (if the peak power plant is too expensive compared to the utility of preventing involuntary rationing). The opposite reasoning for a situation of under-capacity is identical (the cost of adding a power plant generates a

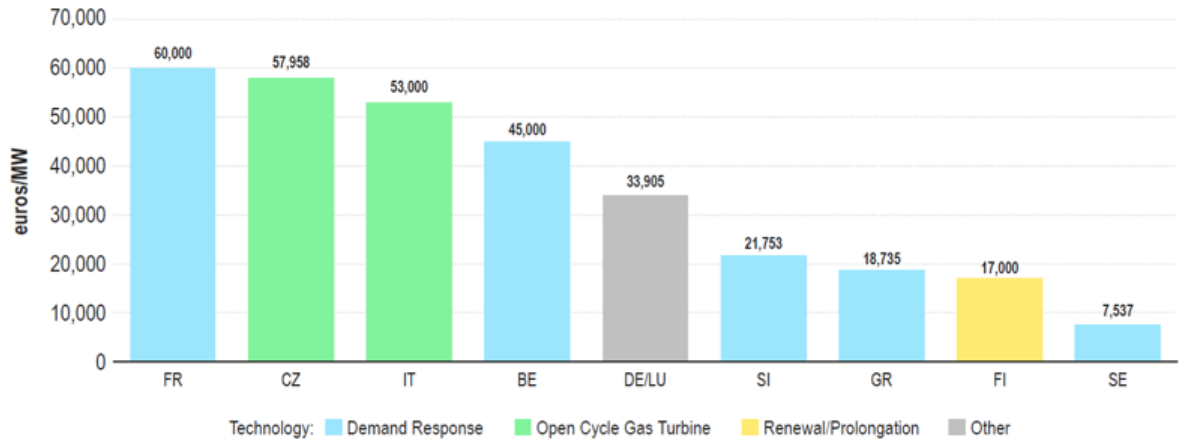


Figure 6: Fixed CONE and technology defining the reliability standard (ACER, 2021)

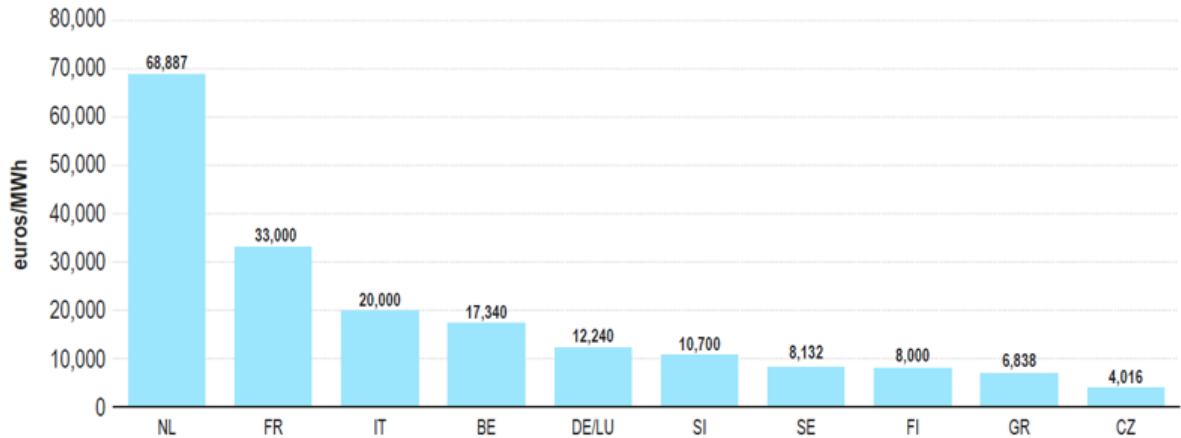


Figure 7: Single VoLL for the calculation of the reliability standard (ACER, 2021)

net gain compared with the initial cost of failure). In all cases, a deviation from the optimal level of the mix leads to significant losses. The usefulness of this system view of adequacy is that it is independent of the system's mode of regulation (whether based on a *control-command* mode or competition). Figures 6, 7 and 8 provide the recent values of the CONE, VoLL, and LOLE used in the European national system from ACER (2021). In this report, the reliability standard corresponds to the LOLE value. For an order of magnitude, the average household price for electricity was 289 euros/MWh (Eurostat, 2024).

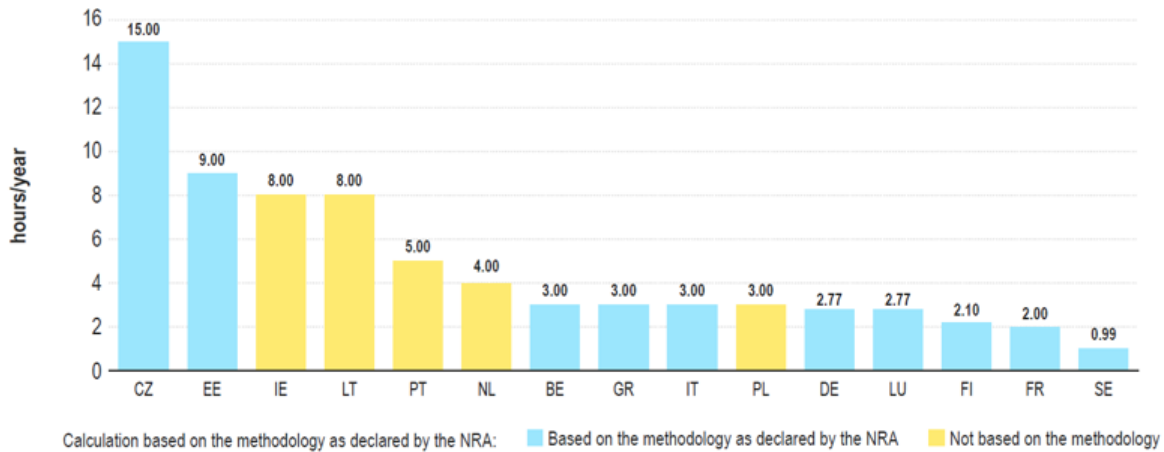


Figure 8: Reliability standard as LOLE (ACER, 2021)

1.3 The limit to the Spot Pricing theory

However, a number of economic, political, and technical constraints prevent energy markets from delivering the optimal generation mix defined in the previous section. Numerous authors have examined the reasons why the current design of electricity markets does not encourage producers to invest sufficiently. The following list presents the main reasons for this sub-optimality:

Explicit price regulation.

It is commonly referred to in the literature as a price cap. To lead to inefficient energy markets, these price caps must be below the price level that would generate sufficient scarcity rent (Joskow, 2008) (Figure 9). This lack of adequate scarcity rent means that the initial production mix changes over the periods following the introduction of a price cap. As a result, producers are unable to cover part of their fixed costs, leading the system to acquire less production capacity. This situation, combined with the impossibility of prices to reach the VoLL, also generates involuntary demand rationing.

In Chapter 2, I successively describe the consequences of having a price cap that first generates a limitation of revenue to cover fixed costs and then a price cap that also generates involuntary rationing of demand. Indeed, if the price cannot exceed the price cap value, it prevents its rationing signal so that elastic demand becomes inelastic. I show in the Box 1.3 the consequences of having a price cap in the toy model.

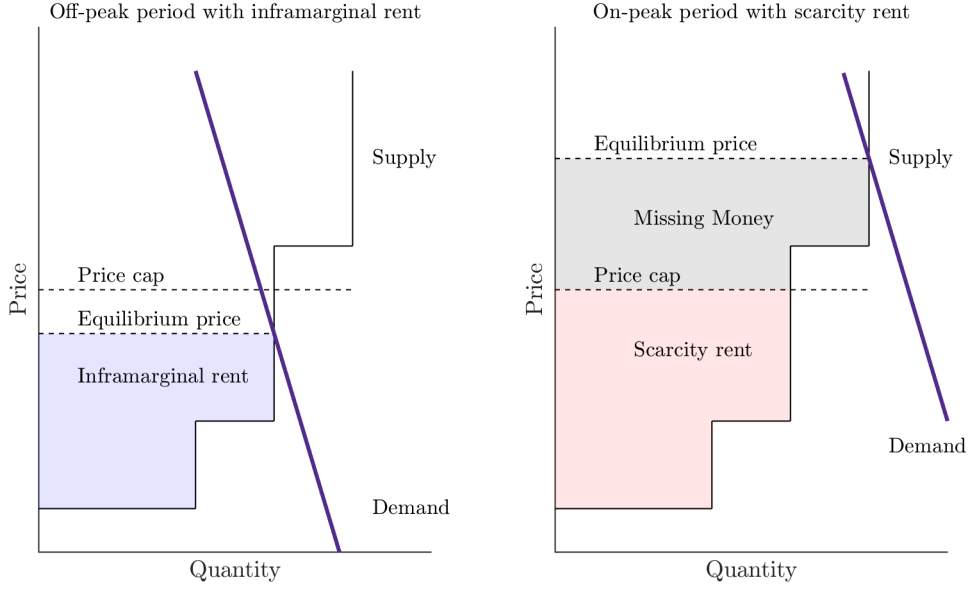


Figure 9: Effects of price caps on rents. The price cap decreases the scarcity rents during on-peak periods. Hence, the expected revenue used to cover fixed costs decreases.

Toy model with a price cap

Assume that there exists a price cap \hat{p} such that the electricity price cannot go beyond $p(s) \leq \hat{p}$ for some s . Then the *price cap long-term* solution is second-best and is defined by:

$$I(c) = \int_0^{\hat{s}} (p(s) - c) dF(s) + \int_{\hat{s}}^{\infty} (\hat{p} - c) dF(s)$$

With \hat{s} , the first state of the world is such that the first-best price equals the price cap. It is straightforward to see that $\int_{\hat{s}}^{\infty} (\hat{p} - c) dF(s) < \int_{\hat{s}}^{\infty} (p(s) - c) dF(s)$. That is, the expected revenue to cover investment costs is lower under a price cap regulation.

The inability of markets to reveal true value for adequacy.

Current technical constraints mean that each consumer can only, with difficulty, reveal the price for which he is prepared to disconnect, i.e., not to consume voluntarily. Indeed,

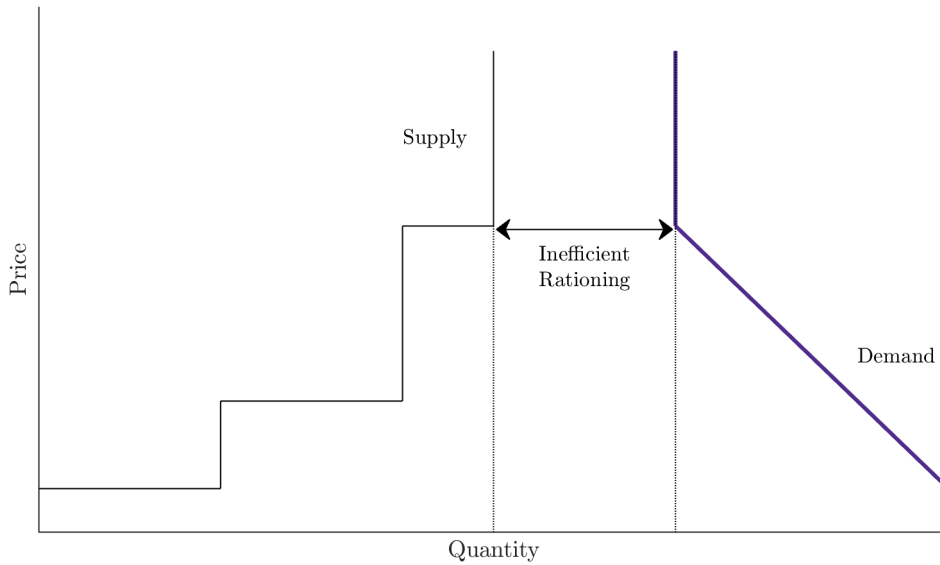


Figure 10: Illustration of the market inability to find the optimum. Capacity constraint implies that supply is inelastic. If demand is also inelastic, then for some periods, prices cannot emerge even in an efficient market.

it is still impossible for the grid operator to disconnect individual consumers for short periods, even if they wish not to consume. This problem lies not only in the market’s inability to avoid involuntary rationing but in its failure to optimize the duration of such rationing (Cramton et al., 2013). Spot-price theory assumes that the level of capacity to be achieved is the result of a tradeoff between the CONE and the VoLL. However, the fundamental and current problem of energy markets is their inability to generate value for VoLL during periods of tension. To function correctly, the system must then turn to an exogenous value, most often determined by a regulated entity (Cramton and Stoft, 2006). The inelasticity of the supply curve (due in particular to the current impossibility of providing storage at an affordable price) and the demand curve (due to the lack of demand response) over the short-term means that the system manager has to use involuntary demand rationing. Figure 10 illustrates this problem.

I partly explore this question within Chapter 3. I describe an incomplete information framework in which a regulated firm making investment decisions does not know perfectly the utility derived from electricity consumption. I describe in the Box 1.3 the consequence of having private information in the toy model. Namely, not knowing the value of electricity implies that when prices cannot ration demand, consumers sustain

higher losses. The central rationale is that consumers who highly value electricity are disconnected instead of consumers who have a lower willingness to pay.

Toy model with random inefficient rationing

Assume that instead of a representative consumer, there is now a unit mass of consumers demanding a single unit of electricity that varies in their willingness to pay θ (i.e., their type). This willingness follows a distribution $G(\theta)$. Let $D(p, s)$ be the demand function aggregating their willingness to pay. In other terms, $D(p, s) = 1 - F(\theta)$, the demand represents the measure of consumer for which their willingness to pay is above the price p . Also, assume that θ is private information, i.e., only the consumer knows its type.

A regulated entity is in charge of lowering consumption (i.e., *disconnect*) whenever the level of available investment k is too low. The entity is also facing a price cap p^r . Following the previous notation, whenever $s > \hat{s}$, quantity is given by $\hat{q}(s) = D(\hat{p}, s)$. Now assume that for every $s > \hat{s}$, $\hat{q}(s) < k$. That is, the quantity asked by all consumers for which their willingness to pay is above the price cap is always higher than the available investment.

Due to private information and in the absence of any complex mechanism to extract information, the regulated entity is left with randomly disconnecting consumers. In that case, the welfare the entity has to maximize is given by:

$$\max_{p(s)} \int_s \left(CS^r(p, s) - \int_0^\infty u(c, s)cdG(c) \right) dF(s) - \int_0^\infty I(c)dG(c)$$

With $\int_s CS^r(p, s)dF(s) = \int_0^{\hat{s}} CS(p, s)dF(s) + \int_0^{\hat{s}} \frac{k}{\hat{q}(s)}CS(p, s)dF(s)$ which is clearly lower than the first-best consumer surplus. This additional cost stems from the impossibility of distinguishing between low-type and high-type consumers.

Uncertainty and risk aversion.

Risk is often cited as a factor that significantly modifies the level of investment (Scouffaire, 2019). Even if uncertainty is naturally present in power markets and, therefore, cannot be excluded from the optimal theoretical model, some authors point out that its level could be significantly reduced, notably through better management of public energy policies (De Sisternes and Parsons, 2016). Moreover, it is also likely that the current environment is leading producers to be overly risk-averse, generating additional costs and thus increasing the sub-optimality of current systems (e.g. Ehrenmann and Smeers (2011); Peluchon (2019, 2021)).

Incomplete markets.

There are two problems with the current electricity system, both of which imply that producers and consumers cannot correctly exchange a number of goods:

Externalities The first type of externality is linked to CO₂ and other forms of pollution. Electricity systems include fossil fuel investment, and the absence of a mechanism that correctly takes into account the costs of CO₂ means that the level and composition of the production mix deviate from its optimal situation (Cepeda and Saguan, 2016). The second type of externality is linked to the "public good" nature of the adequacy of the electricity system. Indeed, the fact that it is technically impossible for the grid operator to make peak-period electricity consumption "excludable" makes it difficult to set a value on adequacy. The externalities associated with public goods and applied to investment issues in the electricity sector have been studied by Oren (2003) and then by Kiessling and Giberson (2004). These authors point out that the public nature of adequacy necessarily implies a risk of *stowaway behavior*, leading to market sub-optimality. On the contrary, Keppler (2017) points out that even with virtuous behavior, externalities, due to the transaction costs of setting up individual contracts between actors not to consume, prevent the market from reaching its optimum. In all cases, the market's sub-optimality is indeed due to the impossibility of actors to carry out their transactions fully.

Risk coverage Actors do not necessarily find counterparties to cover part of their risk, either because they do not exist (demand does not necessarily have the same risk-hedging objective as supply (Finon, 2008, 2011)) or because regulations do not allow risk

hedging for competitive reasons (Genoese et al., 2016). The sub-optimality of a market with no possibility of hedging risk has been illustrated in Newbery and Stiglitz (1984).

Other reasons.

The specific nature of investments (in particular, their size and increment) can lead to opportunistic behaviors that encourage producers to always under-invest and to generate inefficient investment cycles and delays (Keppler, 2017; Stevens et al., 2023). The myopic behavior of certain investors and those generally considered non-rational (gregarious behavior, for example) lead to production mixes that are far from optimal (Finon et al., 2004; White, 2006). This was the case in the early 2000s with excessive investment in gas-fired power plants (Roques et al., 2006). Due to short-term technical constraints, the grid operator may also be forced to intervene in a non-economically optimal way on the markets, creating an implicit price cap on the energy markets (Joskow and Tirole, 2007; Roques, 2008).

How imperfections translate to inefficient investment levels.

To sum up, there are a number of constraints due to market failures, but also to the very nature of power markets, which prevent producers and consumers from ensuring optimal adequacy levels. From a practical point of view, these constraints can be grouped into two categories:

- The impossibility for producers to cover all their fixed costs (a concept known as "missing money" (Joskow, 2008)); in this case, producers cannot properly recover their fixed costs via inframarginal rents and scarcity rents.
- The impossibility for producers to consider their income sufficient, even though they theoretically could be (a relatively less developed concept in power system literature, known as the "missing market") (Newbery, 2016).

A third category, supported in particular by Leautier (2016), is based on competition issues. The risk of under-investment would not be due to the impossibility of producers to cover their costs but would be the result of strategic behavior due to the high level of concentration in electricity markets. The fundamental difference between under-investment created by *missing money* or *missing market* and underinvestment caused

by strategic behavior lies in the characteristics of the policies employed to solve these problems: in the first case, all investors are affected, and it is indeed the market that is structurally inefficient due to technical, political and social constraints; in the second case, a competition policy would make it possible to solve part of the adequacy problems significantly.

To date, the main reason for adequacy risk in power markets remains debated. In the case of the thesis, I will focus solely on public policies aimed directly at *missing money* and *missing market* issues. As we shall see, the majority of mechanisms put in place to resolve the situation of market sub-optimality aim to ensure additional remuneration for producers, which cannot be directly assimilated to a competition policy, particularly in terms of surplus redistribution [Leautier \(2016\)](#).⁷

2 Economic instruments for capacity adequacy

2.1 Theoretical foundation

Two types of public intervention are generally possible when seeking to reduce the inefficiency of a market: the adoption of regulatory measures or the implementation of economic instruments leading economic agents to modify their behavior ([RTE, 2014](#)). In the first case, the problem is managed directly by the authorities, who, in the case of adequacy, directly decide to invest in additional capacity. In the current context of power markets (at least in the European Union), the use of economic instruments remains favored, both for practical reasons - the main agents producing and consuming electricity remain private actors, and for theoretical reasons - the transition from a vertically integrated model to a decentralized model based on competition relies on the same arguments that favor the use of economic instruments.

To understand the various economic instruments available to solve adequacy, one can use the effect of a price cap ([Astier and Lambin, 2019](#)). Based on classical microeconomic theory, the impact of the policies materializes through two axes: a supply-based axis, i.e., investment in capacity; a demand-based axis, and in particular, demand-response.

⁷The main idea is that if under-investment is due to imperfect competition, then producers already receives a surplus which is higher than the optimal level. Therefore, implementing mechanisms that transfer surplus towards producers to increase investment will only reinforce imperfect competition.

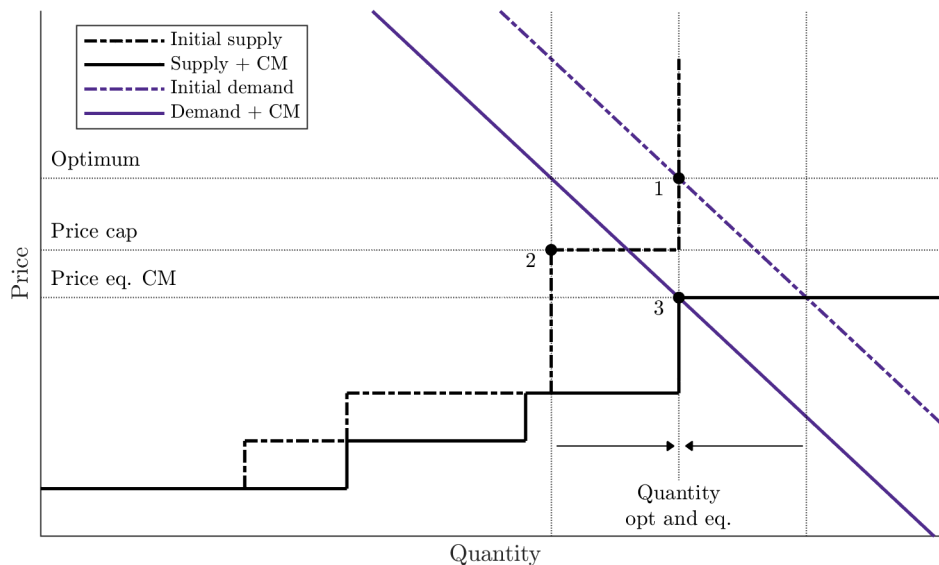


Figure 11: Illustration of economic policy tools to solve adequacy problem. Implementing them can lead to a shift downward of both the supply and the demand depending on who receives incentives. The initial optimum is point (1). The price cap implements an inefficient equilibrium of point (2). The incentives to consumers and producers allow for reaching the initial optimal quantity at point (3).

To illustrate this approach, Figure 11 represents the effect of a price cap on energy markets in the traditional way to what is achieved in theoretical economics. The two dotted curves represent the initial levels of energy supply and demand before the introduction of the price cap on the markets (for simplicity, I have not modeled the impact on the production mix of the introduction of the price cap). The amount of energy supplied to the system corresponds to the first intersection between the supply curve and the price cap level.

The two economic instruments available consist of modifying either the supply curve or the demand curve, or both: to increase the quantity of energy supplied to the economy, the supply curve must be shifted to the right or the demand curve to the left. In economic terms, this means providing the producers with additional remuneration. In classical producer-consumer theory, the supply curve is made up of producers' (marginal) costs and the demand curve of consumers' (marginal) utility. Additional remuneration to producers is identical to a reduction in their costs. On the demand-side, the instrument is more complex: we need to provide an incentive not to consume during peak periods. However, any purchase of energy by the demand-side is associated with a non-zero gain

in terms of utility (gain materialized by consumer surplus). Consequently, economic theory assumes that for a demand-side actor to be indifferent between consuming and not consuming, it is necessary to provide a remuneration equal to the surplus generated by consumption.

The translation of these two curves is shown on the same graph as the solid curves. An ideal calibration of the two economic instruments can then lead to a quantity supplied identical to that without a price cap, which corresponds to the intersection between the dotted supply and demand curves. The cost of implementing these instruments via additional remuneration corresponds to the price difference observed in the energy market.

This highly simplified model applies to both *missing money* and *missing market* issues, even if their associated policies differ. Indeed, the existence of externalities or the presence of too much risk implies that agents (on the supply or demand-side) do not take into account the characteristics of their environment.⁸ Consequently, this boils down to modeling two curves (for both supply and demand): one representing sub-optimality and one taking into account externalities and risk management.

2.2 Long-term mechanisms

An analysis of the various system reforms around the world shows a diverse set of mechanisms capable of ensuring supply-demand adequacy. These can be grouped into two categories: **capacity remuneration mechanism modules**, and **long-term contract modules**. In simple terms, the former was set up to solve *missing money* problems by providing direct remuneration to producers. The second modules provide more certain remuneration than in energy markets, which corresponds above all to a form of risk hedging for producers. Such a solution reduces the problems of incomplete and imperfect information, which is in line with the *missing market* issues. However, due to the specific nature of capacity remuneration mechanisms, that is, over more extended time frames than short-term energy markets, it is also possible that these mechanisms may have an impact on risk issues (De Maere d’Aertrycke et al., 2017). In addition, under specific conditions, some capacity remuneration mechanisms can be analyzed in a similar way to long-term contracts (Leautier, 2016; Schneider et al., 2017).

⁸See for instance Meunier (2013) for the description of the impact of risk aversion on investment decision. The author translates risk aversion into an additional cost for producers. He then implements this approach in a model of Screening Curve.

It is important to note that capacity payment mechanisms and long-term contracts are not the only options that can be considered to solve underinvestment problems. [Cramton and Stoft \(2006\)](#) describes a series of alternative mechanisms that could potentially solve these problems. The authors group them under the term Energy-Only Mechanism, as their implementation does not require the creation of secondary mechanisms in addition to the energy market.

For this thesis, I will focus solely on the design of capacity remuneration mechanisms.⁹ An abundant literature on long-term contract modules already exists both on the analysis of the effects of contracts on risk hedging ([Joskow and Schmalensee, 1983](#); [Finon, 2008](#); [Finon and Roques, 2013](#)) and their impact on producers and their investment decisions ([Roques, 2008](#); [Abada et al., 2019](#)). Nor will I address "Energy-Only" approaches to adequacy. By not focusing on *missing money* or *missing market* issues, these mechanisms run the risk of failing to achieve the optimal level of investment. [Cramton and Stoft \(2006\)](#) point out that the main problem with energy markets lies in the impossibility of private producers and consumers to reveal their preference for adequacy. Consequently, the only answer to this constraint is to set an exogenous, non-market level for adequacy. Even mechanisms based on energy markets and short-term reserves, such as Hogan's approach [Hogan et al. \(2005\)](#), require government intervention. In view of this and of the need to limit this study to a relevant number of mechanisms, these alternative approaches will not be studied in this thesis.

3 A case for capacity markets: principles and limits

From a system point of view, the constraints on current energy markets lead private equilibrium to fall outside the social optimum, as represented in [Figure 12](#). In this case, providing adequate additional remuneration in line with the level of investment would help insure against inefficient rationing-related costs ([De Vries and Ospina, 2012](#)). More precisely, the additional remuneration associated with a capacity market can then be understood as a valuation of capacity given the current adequacy level. [Figure 12](#) highlights the link between the value of current capacity and the opportunity cost of its

⁹The denomination of capacity remuneration in this manuscript covers an extensive range of current mechanisms: capacity markets (centralized or decentralized), strategic reserves, and reliability options. The models developed in the subsequent chapters do not sufficiently go into detail to allow for comparisons between the different implementations, which is outside the scope of the thesis. Moreover, there already exists a vast literature on this subject, for instance ([Holmberg et al., 2021](#)). For ease of presentation, I will use the term capacity market instead of capacity remuneration mechanism.

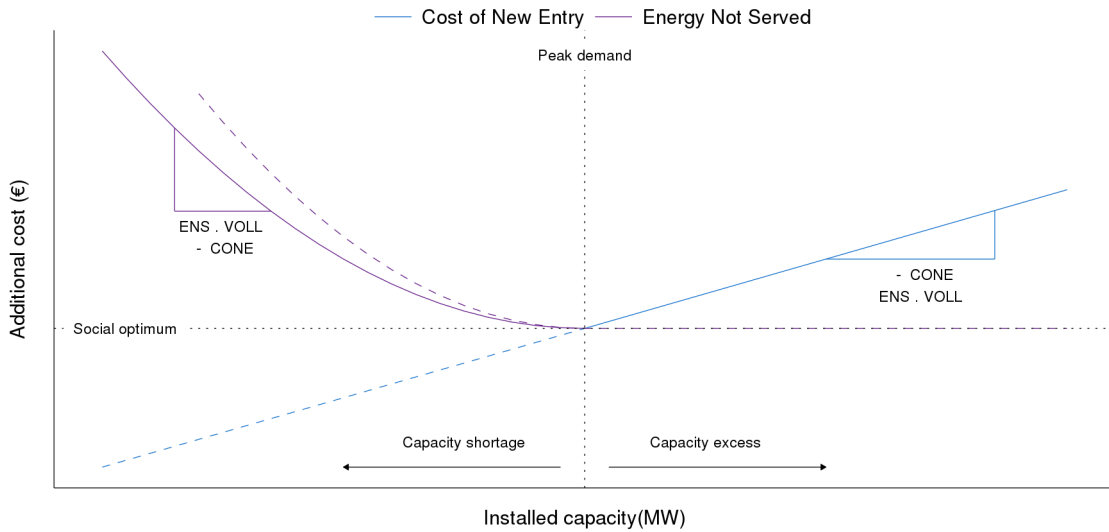


Figure 12: Social cost of electricity shortages and excess capacity

absence. An optimal mechanism should, therefore, be able to provide an incentive to manage adequacy at a cost close to the value of additional capacity.

In their canonical model, [Astier and Lambin \(2019\)](#), establish the link between both this system view (the VoLL / CONE tradeoff) and the actor view (the fixed-cost / rent tradeoff), but also between the different supply and demand-based approaches. From an actor's point of view, and in the presence of a price cap, a capacity market must be able to restore investment incentives by transferring *missing money* from demand to supply. However, [Astier and Lambin \(2019\)](#) highlights the necessary conditions for such a transfer:

”In order to restore allocative efficiency, a producer (resp. a consumer) should only supply (resp. consume) power during a peak state if his marginal cost is below (resp. her willingness-to-pay is above) the social marginal cost of power. Consistently, models of capacity adequacy mechanisms conclude that in order to be supplied in peak states, consumers, willing to serve at all times when the cost of production is below their willingness to pay, should pay the missing money transfer.”

From a system point of view, a capacity market reveals the value of additional capacity for the system, which was not possible with an Energy-Only market. From an actor's point of view, this same mechanism boils down to transferring rents between the

various agents. This additional remuneration can then be understood as a signal sent to agents to modify their incentives and lead to optimal behavior.

At this point, we need to re-establish the link between managing the overall *reliability* issue and capacity markets. As we have seen, this issue can be distinguished between a short-term and a long-term constraint. In particular, although it is explicit that such mechanisms must restore incentives for investment and optimal consumption over the long term, it must also be stressed that there is a significant link between these mechanisms and the short-term horizon. This implicit link is illustrated in Figure 11. Indeed, the introduction of a capacity market is necessarily accompanied by a change in the way the energy market operates. Consequently, such a mechanism must, when accompanied by the transfer of rents between agents, enable supply and demand to meet in the most economically optimal way.

However, the main limitations of these markets stem precisely from why they were set up in the first place. Indeed, these mechanisms rely on the participation of private and public actors in a market process. This is based on the assumption that they reveal more efficiently the true value of additional capacity as well as the willingness to pay for adequacy. However, this principle is inherently fragile. Indeed, the implementation of a capacity market only consists in creating conditions that guarantee sufficient capacity to enable an optimal match between electricity supply and demand, both in the short and long term. Höschle (2018) emphasizes this aspect of these mechanisms:

”A CM [capacity market] is often characterized as a mechanism that ”guarantees” generation adequacy. However, a guarantee cannot be provided by a market design that depends on individual decision-making. A CM is rather a mechanism that reduces the risk of having insufficient capacity by providing an adequate market signal.”

Indeed, capacity markets are based on the formation of a price, which is the confrontation of supply and demand. The supply-side is decentralized and composed of producers owning capacity.¹⁰ On the other hand, the demand-side can be either centralized, decentralized, or hybrid. In the centralized case, a single thesis, usually a regulated firm, forms the demand on the capacity market. In a decentralized case, consumers (or their retailers) purchase directly the capacity. The hybrid case encompasses a form of

¹⁰Administratively, some consumers can sometimes participate as supply-side actors. Under the conceptual and stylized framework, I do not make such a distinction.

mechanism where consumers indirectly participate in the mechanism.¹¹ Therefore, the core ideas of an efficient capacity market are that (i) the participation of producers leads to the correct expression of their private value for capacity, (ii) the demand-side (either the single buyer or consumers) is able to express the aggregate willingness to pay for capacity, (iii) the price signal sent to both supply-side and demand-side leads to the right decisions (both from an investment and consumption perspective). Therefore, capacity markets create an additional risk that participants will not behave as expected. Such a deviation between the public policy objective via the mechanism (in terms of adequacy level) and the level actually achieved is at the core of this thesis's main contribution.

Contribution 1

The economic efficiency of the capacity markets depends entirely on the quality of the signals they send out. Understanding how these participants behave and interact is therefore essential if we are to study the welfare derived from them. In this thesis, I deepen the understanding of capacity markets by representing, using a diverse set of stylized frameworks, the interaction between investment decisions and participation in capacity and energy markets.

Using a literature review on the modeling of the capacity market, I propose in the Appendix Chapter A a thought framework of how we can study the behavior in capacity markets. With respect to this contribution, the three chapters of this thesis can be understood respectively as focusing on (i) solely on the supply-side in Chapter 1, (ii) solely on the demand-side in Chapter 2, and (iii) combining the two issues in Chapter 3. More precisely:

- **Chapter 1** describes how producers bid in a capacity market using Real Option Theory.
- **Chapter 2** studies the interaction between the supply-side and demand-side of electricity markets when a capacity market is implemented.

¹¹For instance, a regulated Firm can aggregate information by implementing a mechanism to make consumer and retailers reveal their future consumption. There is still a unique single buyer in the capacity market, but the behavior of the consumers constrains the regulated firm.

- **Chapter 3** proposes a framework for consumer participation in capacity markets with incomplete information.¹²

I have described the efficiency of a capacity market lies in its ability to provide adequate incentives for producers. In the context of investment decisions, those incentives cover a wide variety of decisions: (i) choosing the level of investment and its technology when making an entry decision, but also (ii) choosing the right entry time as well as the right exit time, and at at least but not last, (iii) weighing the pros and cons of mothballing or refurbishing a power plant. In this thesis, I will limit the analysis to a static analysis of capacity market efficiency. Namely, Chapters 1, 2, and 3 deal only with the level of investment in one-shot games with a single technology.

The simplification of the approach allows us to focus on the second main contribution of the thesis. Indeed, the existence of capacity markets raises the question of their appropriate design. In electricity markets, this exercise is essential since poor design can lead to high costs for society and inefficient behavior on the part of players or even to the creation of a system that is less efficient than if it were vertically integrated (Woo et al., 2003).

Contribution 2

Given the importance of the behavior in a capacity market, the thesis derives new theoretical and policy results with respect to how participants behave in certain market designs and their welfare implications when investment decisions are considered.

In the Appendix Chapter A, I discuss some questions studied by the modeling literature related to the design of the capacity market. With respect to this contribution:

- **Chapter 1** compares the different capacity product designs and how they impact the bidding behavior. I mainly study the length of the product.

¹²In fact, capacity markets are not explicitly represented in Chapter 3. However, the chapter can be understood as deriving the theoretical limits of what a central authority can do when facing consumers with private information and considering investment decisions. In other words, if we assume a perfectly functioning capacity market that allows consumers to reveal their willingness to pay for adequacy truthfully, it cannot perform better than what is found in the chapter.

- **Chapter 2** shows that depending on different assumptions, consumers' and retailers' participation in a capacity market can have a significant effect. I analyze how the cost allocation of the capacity market changes the equilibrium of electricity markets and how retailers can be considered in the design.
- **Chapter 3** provides the aggregate and individual welfare effects of different market designs when a regulated firm needs to generate revenue to make investment decisions. The core idea of the chapter is to study the different designs when a regulated firm faces different constraints and, in particular, revenue, incentive, and participation constraints.

In the rest of the introduction, I provide a short summary of each chapter's primary results. A more complete overview can be found in the General Conclusion 3 of this thesis.

Designing efficient capacity markets: Bidding behavior and product definition.

This research originates from two questions: "How do producers bid in capacity markets?" and "Does the length of the contract sold on the capacity market matter?". I answer those questions by stating that participation in a capacity market implies a specific marginal opportunity cost for the bidders, which requires a particular method of valuation. Namely, selling capacity entails the producer staying available over future periods determined in the sold contract. The first main contribution lies in studying the bid as an option value associated with the possibility of closing temporarily but irreversibly to avoid fixed costs. This significantly differs from the standard approach of the net present value approach, where the producer only offers the expected opportunity cost. Thus, I propose a novel way to conceptualize capacity bids using real options theory, where the opportunity cost is represented as an option on the spread that drives the profitability of the plant. First, I define a bid in a one-period capacity market as a European Put Option. Then, I expand to a multi-period setting in which capacity bids can be evaluated as a modified Basket Option.

The central results of this chapter are the implications of the change in contract duration on the bidding behavior in a capacity market, assuming that the marginal cost is analyzed through the lens of real-option theory. I find that it significantly differs from the net present value framework. First, bids are always higher under the real options

framework, meaning that producers place a positive value on the possibility of closing to avoid some costs. Second, the drivers behind the bids have different effects on their value compared to the net present value framework. I provide comparative statics on the bid value and the difference between the two frameworks. I find that the length of the contract constantly increases the bid when using the real options theory while having an ambiguous effect on the net present value bid. The volatility in the wholesale market and the waiting time between the sale of the contract and the start of the availability are also analyzed. Finally, I find the reverse effect for the product design dimension with a higher bid with a longer contract duration than the sum of expected bids with shorter products. I test the results by calibrating the model to the French electricity system using realized data for a CCGT (gas) power plant to simulate a bid in the capacity market and compare the outcomes with realized prices observed on the French capacity market. I finally provide a policy discussion for the design of capacity markets: (i) On the role of penalty in the capacity market and (ii) On the difference between existing and new capacity participating in capacity markets.

Securing investment for electricity markets. How do we design the demand-side of capacity markets?

This chapter aims to understand the pros and cons of the different options available to the regulator when designing the demand-side of a capacity market. In particular, I show the consequences of different options on a model representing capacity market, investment decisions, generation, and consumption equilibrium. Ultimately, the chapter describes the welfare of a power market as a function of different assumptions about market design, the behavior of players, and the environment. I focus on two interrelated questions that relate to (i) the cost allocation regime, that is, how a single buyer allocates the capacity price between capacity buyers and final consumers, and (ii) the degree to which the final consumers' realized demand is accounted in the market allocation design.

The first market design regime studied is the canonical capacity market. This regime is similar to having a cost allocation regime based on a lump-sum tax. In this case, even when considering the supply function in the capacity market, the equilibrium of the market design always restores the first-best optimum given the system inefficiencies. I then investigate the case in which the capacity price impacts consumers at the margin. Under this assumption, the regime similarly allocates the capacity price as a unitary tax. I demonstrate that the welfare outcome at the equilibrium under this regime is

always lower than under the canonical regime. I then compare the two capacity cost allocation regimes by including inefficient rationing. When a price cap is reached, the investment availability becomes a public good as the demand becomes inelastic. Due to the impossibility of efficiently rationing consumers, they incur a significant welfare loss. Under this new assumption, I find that the indirect effect created by allocating the capacity price on a unitary basis is now ambiguous for social welfare. In other words, if correctly designed, a capacity market can indirectly and imperfectly recreate a price signal that a price cap has suppressed. In the second part of the paper, I study two extensions of the model that represent different attempts to base the capacity market demand-side on the realized consumption. The first regime analyzes how the allocation of the capacity cost to retailers based on their realized market share affects welfare. I show that this market design leads to an intermediary outcome between the lump-sum and the unitary market design and that the level of imperfect competition has ambiguous impacts. The second regime describes a decentralized capacity market in which retailers form the demand function. I notably address the central question of the penalty value and the outcome of the market design.

Designing Markets for Reliability with Incomplete Information

This chapter's main contribution is to discuss the implications of considering the demand-side when it comes to ensuring an efficient level of investment. It studies an environment in which there are a number of consumers with different characteristics whose utility from electricity consumption is partly private information. In other words, it seeks to know how to design markets so that consumers reveal their utility for their electricity consumption and, therefore, indirectly for the investment level. In that research, markets are not only used to provide consumption and sufficient revenue to make investments but also to screen for unobservable characteristics to ensure the proper investment level. A second approach that I tackle in this chapter is the distribution issues associated with the design of an optimal market in this context. Namely, I show that implementing the most efficient market to reach the most efficient investment level is not always Pareto improving for every consumer.

The first case with incomplete information represents the inefficiencies associated with private information when the market designer is constrained in the allocation choice. Namely, some consumers over-consume with respect to their type, while others under-consume. I first study the single-price case when the market designer cannot discriminate between categories of consumers. In that case, the price schedule increases with the level

of investment. I then implement the possibility of discriminating between consumer categories. I show that for the category of consumers with the smaller average private shock, the optimal price first decreases and then increases with the investment level. On the other hand, the price for the category of higher consumers is always increasing with the capacity. I also find that the price for the former category is above the latter for relatively low capacity values, and then the ranking reverses for higher values. These non-monotonicities can be explained by the opposite effect the market designer faces in terms of consumer surplus and revenue effect when choosing prices.

In the last section, I study a mechanism design setup where the market designer is no longer constrained in the prices and quantities schedule he can offer consumers. He now faces incentive compatibility and individual rationality constraints. I first describe for which level of investment the market designer is constrained by the revenue used to cover the fixed costs and the information rent that he needs to provide to consumers so they behave truthfully. I find that the market designer can provide an unconstrained first-best allocation only for low values of the investment level due to the concavity of the virtual consumer utility with respect to the capacity. For higher values, the additional expected utility gains from increased capacity cannot compensate the investment costs. Then, I show that the behavior of the optimal allocation depends on the state of the world considered and the type of consumers. Namely, as the investment level increases, consumers always face a decrease in the optimal quantity allocated during off-peak. For on-peak periods, the quantity change depends on the consumer's type. Finally, I show that only the large consumers from both categories can gain from an increase in the investment level. The rationale behind this result lies in the link between the investment level and the information rent given to consumers in order for them to behave truthfully.

Bibliography

Abada, I., de Maere d'Aertrycke, G., Ehrenmann, A., and Smeers, Y. (2019). What models tell us about long-term contracts in times of the energy transition. *Economics of Energy & Environmental Policy*, 8(1):163–182.

ACER (2021). Security of eu electricity supply in 2021: Report on member states approaches to assess and ensure adequacy. Technical report, ACER.

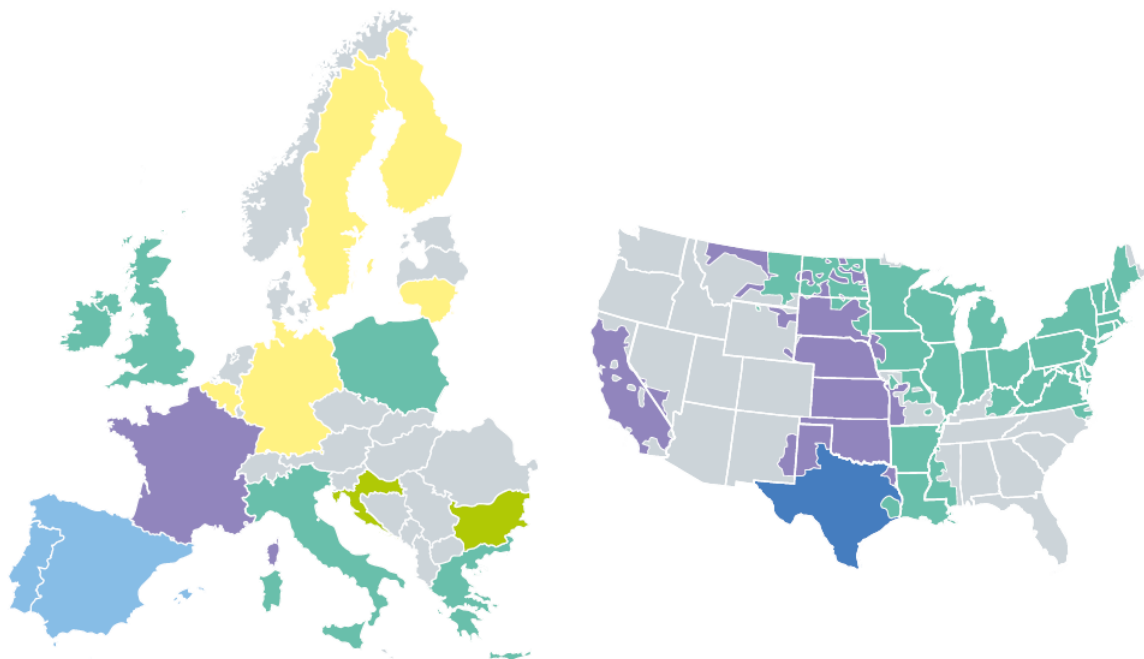
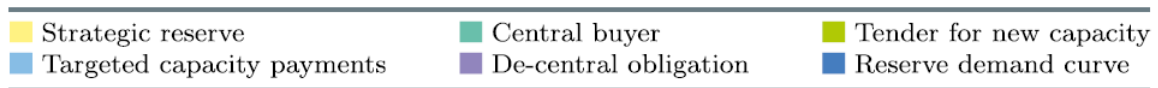


Figure 13: Current market design across the world (Bublitz et al., 2019). Note that the denomination of capacity markets includes the centralized and decentralized design as well as the strategic reserves and reliability options.

- Astier, N. and Lambin, X. (2019). Ensuring capacity adequacy in liberalised electricity markets. *Energy Journal*, 40(3):227–242.
- Bublitz, A., Keles, D., Zimmermann, F., Fraunholz, C., and Fichtner, W. (2019). A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms. *Energy Economics*, 80:1059–1078.
- Caramanis, M. C., Bohn, R. E., and Schweppe, F. C. (1982). Optimal spot pricing: Practice and theory. *IEEE Transactions on Power Apparatus and Systems*, PAS-101(9):3234–3245.
- Cepeda, M. and Saguan, M. (2016). Assessing long-term effects of demand response policies in wholesale electricity markets. *International Journal of Electrical Power & Energy Systems*, 74:142–152.
- CIGRE (1987). Power system reliability analysis. Technical report, CIGRE.
- Cramton, P., Ockenfels, A., and Stoft, S. (2013). Capacity Market Fundamentals. *Economics of Energy & Environmental Policy*.
- Cramton, P. and Stoft, S. (2006). The convergence of market designs for adequate generating capacity. Papers of peter cramton, University of Maryland, Department of Economics - Peter Cramton.
- De Maere d’Aertrycke, G., Ehrenmann, A., and Smeers, Y. (2017). Investment with incomplete markets for risk: The need for long-term contracts. *Energy Policy*, 105:571–583.
- De Sisternes, F. J. and Parsons, J. E. (2016). The impact of uncertainty on the need and design of capacity remuneration mechanisms in low-carbon power systems. *MIT Center for Energy and Environmental Policy Research (February 2016)*.
- De Vries, L. J. and Ospina, J. R. (2012). European security of electricity supply policy in the context of increasing volumes of intermittent generation. In *12th IAEE European Energy Conference*, pages 1–16.
- Ehrenmann, A. and Smeers, Y. (2011). Generation capacity expansion in a risky environment: a stochastic equilibrium analysis. *Operations research*, 59(6):1332–1346.
- Eurostat (2024). Electricity and gas prices stabilise in 2023.
- FED, D. (2021). Cost of texas’ 2021 deep freeze justifies weatherization.

- Finon, D. (2008). Investment risk allocation in decentralised electricity markets. the need of long-term contracts and vertical integration. *OPEC Energy Review*, 32(2):150–183.
- Finon, D. (2011). Investment and Competition in Decentralized Electricity Markets: How to Overcome Market Failure by Market Imperfections? In Glachant, J.-M., Finon, D., and de Hauteclocque, A., editors, *Competition, Contracts and Electricity Markets*, Chapters, chapter 3. Edward Elgar Publishing.
- Finon, D., Johnsen, T. A., and Midttun, A. (2004). Challenges when electricity markets face the investment phase.
- Finon, D. and Roques, F. (2013). European electricity market reforms: The “visible hand” of public coordination. *Economics of Energy & Environmental Policy*, 2(2):107–124.
- Genoese, F., Drabik, E., and Egenhofer, C. (2016). The eu power sector needs long-term price signals. *CEPS Special Report*, 135.
- Hogan, W. W. et al. (2005). On an “energy only” electricity market design for resource adequacy.
- Holmberg, P., Tangerås, T. P., et al. (2021). *Strategic reserves versus market-wide capacity mechanisms*. JSTOR.
- Höschle, H. (2018). Capacity Mechanisms in Future Electricity Markets. *Capaciteitsmechanismen in toekomstige elektriciteitsmarkten*.
- IEA, P. (2020). Power systems in transition. Technical report, IEA.
- Joskow, P. and Tirole, J. (2007). Reliability and competitive electricity markets. *The Rand Journal of Economics*, 38(1):60–84.
- Joskow, P. L. (2008). Capacity payments in imperfect electricity markets: Need and design. *Utilities Policy*, 16(3):159–170.
- Joskow, P. L. and Schmalensee, R. (1983). *Markets for power: an analysis of electric utility deregulation*. Massachusetts Inst. of Technology Press, Cambridge, MA.
- Keppler, J. H. (2017). Rationales for capacity remuneration mechanisms: Security of supply externalities and asymmetric investment incentives. *Energy Policy*, 105:562–570.

- Kiessling, L. and Giberson, M. (2004). Is network reliability a public good? In *Proceedings of the Paper Submitted for Publication to CMU Conference on Electricity Transmission in Deregulated Markets. Carnegie-Mellon Electricity Industry Center, Pittsburgh. 15–16 December.*
- Leautier, T.-O. (2016). The visible hand: ensuring optimal investment in electric power generation. *The Energy Journal*, 37(2).
- Littlechild, S. (1988). Spot pricing of electricity arguments and prospects. *Energy Policy*, 16(4):398–403. Spot pricing.
- Meunier, G. (2013). Risk aversion and technology mix in an electricity market. *Energy Economics*, 40:866–874.
- Newbery, D. (2016). Missing money and missing markets: Reliability, capacity auctions and interconnectors. *Energy Policy*, 94:401–410.
- Newbery, D. M. G. and Stiglitz, J. E. (1984). Pareto inferior trade. *The Review of Economic Studies*, 51(1):1–12.
- Oren, S. S. (2003). Ensuring Generation Adequacy in Competitive Electricity Markets. ., pages 388–413.
- Peluchon, B. (2019). Working Paper: Market design and the cost of capital for generation capacity investment. *Chaire European Electricity Markets.*, pages 1–36.
- Peluchon, B. (2021). Decarbonizing the electricity sector efficiently requires a change of market design. Technical report, CEEM Working Paper 51. Dauphine University, Paris.
- Petitot, M. (2016). *Long-term dynamics of investment decisions in electricity markets with variable renewables development and adequacy objectives.* PhD thesis, PSL - Paris Dauphine University.
- Roques, F. A. (2008). Market design for generation adequacy: Healing causes rather than symptoms. *Utilities Policy*, 16(3):171–183.
- Roques, F. A., Nuttall, W. J., Newbery, D. M., de Neufville, R., and Connors, S. (2006). Nuclear power: A hedge against uncertain gas and carbon prices? *The Energy Journal*, 27(4):1–23.

- RTE (2014). Rapport d'accompagnement de la proposition de règles du mécanisme de capacité. Technical report, RTE.
- Saguan, M. (2007). *L'analyse économique des architectures de marché électrique : application au market design du temps réel*. PhD thesis, Paris 9.
- Schneider, I., Botterud, A., and Roozbehani, M. (2017). Forward contracts and generator market power: How externalities reduce benefits in equilibrium. In *Riding the Energy Cycles, 35th USAEE/IAEE North American Conference, Nov 12-15, 2017*. International Association for Energy Economics.
- Scoufflaire, C. (2019). *Capacity remuneration mechanisms: analytical assessment of contemporary experiences and lessons for the future design of electricity markets*. PhD thesis, PSL - Paris Dauphine University.
- Stevens, N., Smeers, Y., and Papavasiliou, A. (2023). Indivisibilities in investment and the role of a capacity market. *WP*.
- Vassilopoulos, P. (2007). *Les prix des marchés de gros de l'électricité donnent-ils les bons signaux et les bonnes incitations pour l'investissement en capacité de production électrique ?* PhD thesis, Paris 11. Thèse de doctorat dirigée par Chevalier, Jean-Marie Sciences économiques Paris 9 2007.
- White (2006). Financing new nuclear generation. Technical report, Climate Change Capital.
- Woo, C.-K., Lloyd, D., and Tishler, A. (2003). Electricity market reform failures: Uk, norway, alberta and california. *Energy policy*, 31(11):1103–1115.

Chapter 1

Designing efficient capacity markets

Bidding behavior and product definition

Contents of Chapter 1

| | | |
|-----|---|----|
| 1 | Introduction | 42 |
| 2 | Model assumptions | 49 |
| 2.1 | Investment and wholesale market | 49 |
| 2.2 | Producer behavior | 51 |
| 2.3 | The capacity market | 52 |
| 3 | The net present value framework | 53 |
| 3.1 | The inframarginal rent and the opportunity cost | 54 |
| 3.2 | A simple example | 56 |
| 4 | The real options framework | 58 |
| 4.1 | A capacity bid as a European Put Option | 58 |
| 4.2 | A capacity bid as a Basket Option | 59 |
| 4.3 | The sum of expected capacity bids | 60 |
| 5 | Comparative statics | 61 |
| 5.1 | Bids value | 62 |
| 5.2 | Flexibility value | 67 |
| 5.3 | Product design | 69 |
| 6 | Case study | 71 |
| 6.1 | Bids in a capacity market | 73 |
| 6.2 | The effect of the waiting time and the volatility | 75 |
| 6.3 | Product design | 77 |
| 7 | Extensions | 79 |
| 7.1 | Bids for new entrants | 79 |

| | | |
|-----|---|----|
| 7.2 | Penalty and mothballing costs | 80 |
| 8 | Conclusion | 82 |

Abstract

In many countries, capacity markets have been put in place to supplement wholesale market revenues to ensure an adequate generation capacity to maintain the security of supply. This chapter studies the bidding behavior in those markets and how it can be affected by different capacity product designs. A capacity market allows producers to lock in revenues in advance in exchange for their commitment to being available over a future period on wholesale markets. Producers' participation depends on the opportunity cost of making the investment available. When the commitment is made, the profitability of the plant is uncertain. The canonical framework is based on a net present value model, where the capacity bid is equal to the expected loss on the energy market. However, this does not recognize managerial flexibility and assumes that the plant cannot react to future market conditions. Thus, we propose a novel approach to conceptualize capacity bids using real options theory, where the opportunity cost is represented as an option on the spread that drives the profitability of the plant. First, we define a bid in a one-period capacity market as a European Put Option. Then, we expand to a multi-period setting in which capacity bids can be evaluated as a modified Basket Option. This model provides new insights into the interplay between the product duration and the capacity bid. Using the real options approach, the model presents a first attempt to untangle the different drivers of the opportunity cost for providing capacity availability. We analyze the determinants of the option value concomitantly with the length of the procurement and deduce some policy implications for the product's design. Finally, we provide a numerical illustration of this issue using data from the French power system.

1 Introduction

In current power systems, producers do not always receive enough revenue to cover their production costs even though they are deemed necessary to reach the first best investment mix. Electricity prices can be constrained due to political reasons with price caps (Leautier, 2016) or can send distorted price signals due to technical and noneconomic interventions on the market (Joskow, 2008). Other reasons can be found in that electricity prices do not consider the correct value of an additional capacity, for instance, due to the public good nature of capacities during high demand periods (Holmberg and Ritz, 2020) or because some externalities are not correctly internalized (Keppler, 2017). At the same time, the risk of not having enough investment poses a significant threat. Indeed, the absence of adequacy between the capacity installed and the electricity demand, combined with the difficulty of implementing efficient rationing, leads to high system costs. It has been illustrated by the rolling blackouts in the Texas system or during hot summers in California (IEA, 2020).

One solution to restore the right level of investment could be the implementation of capacity remuneration mechanisms. They provide producers with an additional remuneration stream to increase and maintain the optimal level of investment. There are currently various implementations ranging from capacity payments paid directly to the producers to more complex designs with actual markets, where the price emerging from the confrontation between a supply and a demand for capacity makes the additional remuneration. They are usually denominated as capacity markets. Each participating producer makes a price-quantity offer for a capacity on the supply-side of those competition-based mechanisms. If a producer sells a capacity, he receives an additional price, and it legally forces the investment to be available over a specific period in the future.

In this chapter, we mainly investigate two research questions: (i) how to model capacity bids in the context of uncertainty and managerial flexibility to operate or close the plant and (ii) how bids depend on multiple key design features. We tackle those issues by stating that participation in a capacity market implies a specific opportunity cost for the bidder, which is the fundamental driver for its bidding behavior. To do so, we analyze the opportunity cost determinants associated with the decision to be available, allowing a more detailed comparison with the marginal value of an available capacity independent of the product design. Therefore, the subsequent analysis sheds

light on effectively setting up a mechanism based on competition, where the price signal improves economic efficiency. In this chapter, the capacity price encourages producers to invest and stay open when wholesale markets cannot send the proper price signal. Therefore, any deviation of the price from the actual value of an additional capacity for the system can cause an adverse effect. We stress that both market design theory and practitioners must consider the practical limits imposed by the actors' behavior in the face of specific rules.

To our knowledge, we are the first to use a methodology other than the net present value framework to analyze the bidding behavior in a perfectly competitive capacity market. Namely, we state that the opportunity cost of participating in a capacity market is equal to the option value of the availability decision. Such conceptualization sheds new light on how prices emerge in a capacity market. There also has been no formal analysis of the link between the bids in a capacity market and the duration of a capacity product. Therefore, the chapter deepens policy perspectives for the practical implementation of capacity markets.

The opportunity cost of participating in the capacity market is well-known in the literature ([Wilson, 2010](#)), and some papers have highlighted the need to grasp the role of product design better when assessing the efficiency of those mechanisms ([Bushnell et al., 2017](#)). They have underlined the necessity of understanding the opportunity cost drivers when selling an availability to refine the study of capacity prices and help choose the right product design. In this chapter, we underline the multidimensional aspect of this issue with two rationales: (i) the interdependence between the wholesale market and the capacity market and (ii) the managerial flexibility the investment encompasses.

We start the model by recalling the fundamentals behind a single power system investment decision from a private producer perspective. Then, we introduce a simplified capacity market where the representative producer can bid a capacity product in an auction mechanism that forces the investment to be open during a specific period determined before the auction is set.

First, we use a net present approach where the producer only offers the expected opportunity cost associated with the capacity product. In this case, he bids the expected revenues over the procurement duration net of the fixed cost associated with the decision to stay available. We show that a longer product always implies a lower or equal bid than the sum of expected bids for shorter products. In both cases, the bids are always equal to the expected loss. Otherwise, the producer makes a null bid. This first approach

implies for the producer a comparison between only two alternatives: (i) being available during the whole procurement period or (ii) closing during the same periods.

The main contribution lies in studying the bid as an option value associated with the possibility of closing temporarily but irreversibly to avoid fixed costs. First, we use the standard option pricing theory to value a simplified version of a capacity market where the period during which the plant has to be available, called the transaction phase, covers only a single wholesale market clearing. In this case, the capacity product is equivalent to a European Put Option where the exercising date is the transaction phase, the underlying being the wholesale profit, and the strike price is the fixed cost associated with the decision to stay open. Under the real options framework, the bid on the capacity market is strictly equal to the option value. Then, we expand this analysis to a multi-period transaction phase, and we treat the capacity product as a form of Basket Option where the asset price portfolio is the expected revenue generated over the procurement period. It allows us to compare this option value with the sum of the option value for shorter products. Using the real options framework to assess the bidding behavior in a capacity market, we find that it significantly differs from the net present value framework. First, bids are always higher under the real options framework, meaning that producers place a positive value on the possibility of closing to avoid some costs. Second, the drivers behind the bids have different effects on their value compared to the net present value framework. We provide comparative statics on the bid value and the difference between the two frameworks. We find that the length of the transaction phase constantly increases the bid when using the real options theory while having an ambiguous effect on the net present value bid. The volatility in the wholesale market, the policy instrument, and the waiting time between the capacity product sales are also analyzed. They both have ambiguous effects on the capacity bids depending on a set of conditions on the bid drivers. Finally, we find the reverse effect for the product design dimension with a higher bid with a longer transaction phase than the sum of expected bids with shorter products.

We test the results by calibrating the model to the French electricity system.¹ We use realized data for a CCGT (gas) power plant to simulate a bid in the capacity market and compare the outcomes with realized prices observed on the French capacity market. While the results are highly sensitive to the assumptions regarding the drivers of the bids, we find that the real options framework is relatively close to the auction outcomes. The model also stresses that a change of volatility for the investment revenue, due, for

¹We use data before 2022 to avoid taking into account the turmoil in electricity markets.

instance, to the increasing share of renewable, can significantly affect the bidding behavior in a capacity market. Similarly, choosing the duration between the auction date and the transaction phase when designing the capacity market has important implications when looking for the least cost design.

Using both the theoretical framework and the numerical illustration, we provide a policy discussion for the design of capacity markets. Namely, using a real options framework sheds light on the role of penalty in the capacity market imposed onto producers who choose not to be available during the procurement duration despite having sold a capacity product. Similarly, the cost associated with the decision to close can also be included in the analysis. Another crucial point can be made regarding the difference between existing and new capacity. While for the former, the opportunity cost of being available is only made concerning a single-capacity product, for the latter, the decision to enter is more complex. Indeed, it is based on the expected revenue made during the investment lifetime, including future capacity prices. In this case, the effect of different procurement durations can be significant.

Related Literature

A vast amount of literature has studied the effect of capacity markets on investment decisions. Such assessment has been realized in simplified models such as classical Nash equilibrium models (Fabra, 2018), with sometimes a representation of strategic actors (McRae and Wolak, 2019), and stochastic optimization models where the market is mimicked using a minimization cost function (de Maere d’Aertrycke et al., 2017). Other models tried to replicate the complex environment in which those mechanisms have been implemented by representing different fluxes between agents and their decisions’ implications. System Dynamics models study dynamically the effect of capacity markets on investment decisions (Cepeda and Finon, 2011), while Agent base Models use a bottom-up approach to analyze the interactions of specific agents in the power system (Bhagwat et al., 2017). In most papers, regardless of the types of models, they find that capacity markets significantly improve energy markets’ efficiency by increasing investment value and reducing the capacity adequacy issue.

Taking a different angle, we base this work on *single project valuation models*, which are less used in this context. The advantage of this type of model relies on the possibility of finely representing the components of investment value for a producer, their evolution, and technological constraints. More specifically, it allows for a hypothetical investment to represent both its future revenues and the impact of the additional remuneration

on its value dynamically. On the other hand, the model lacks a system view, with no representation of market feedback, technology competition, and market power. Most of the single project valuation model stream applied to power investment focused on real options analysis to study the different values of potential managerial decisions, such as investment in renewable energy under price uncertainty (Fleten et al. (2007), Fuss et al. (2012)), conventional investments under policy and finance uncertainty (Kettunen et al. (2011)) or the effect of different support mechanisms for renewable on investment decision (Fuss et al. (2012)). To our knowledge, only one paper has taken the single project valuation to capacity markets: Hach and Spinler (2016) found that exogenous capacity payments significantly modify the value of new gas power plants, especially when the quantity of renewable is high. Therefore, we expand this approach by endogenizing the payments while using actual data to deduce the investment value and studying different product designs.

The fundamental driver behind bid formation in capacity markets, developed, for instance, by Wilson (2010), is that participation in such a market creates an obligation to be available in a future period on the energy market. Therefore, selling a capacity generates an indirect cost, which could be described as an opportunity cost. The opportunity cost of participating in a mechanism is the cost of being available during a predefined future period, which would not have been incurred if the investment is not producing during the same period. Failures and constraints can lead to insufficiently high prices to cover their costs, even though they are necessary for the system. Consequently, forcing an actor to produce when it is potentially at a loss entails a positive opportunity cost but allows the energy to be efficiently dispatched.

Some papers seeking to reproduce the interdependence of the actors and the different production decisions in power markets are based on this principle (Abani et al., 2016; Bhagwat et al., 2017; Teirilä and Ritz, 2018). Creti and Fabra (2007) shows, for instance, how a monopoly offers on a capacity market when the latter has to give up exporting profit in a foreign market whose price is higher than the price on the national market due to the obligation to be available. The offer on the capacity market is made at a price equivalent to the loss of opportunity to make a profit on the foreign market. Brown (2012) proposes the term of *allocation externality* to characterize the link between capacity bids and energy profits. In his setup, incumbents are dumping capacity prices to avoid new entry into energy markets. Because new entry is made possible with the capacity market, the energy profit could be lower due to higher competition. Therefore, it can be strategic to make losses on capacity markets to prevent more significant losses

in the energy market. The few papers modeling the reliability option markets are also enlightening about this approach.² During periods of scarcity of demand, that is, when the plant is needed, producers undertake to pay back on demand the difference between the energy price received on the energy markets and the strike price of the obligation (Cramton et al., 2013). In those models, the opportunity costs, and thus bids for these options, are equal to the amount transferred on demand (Mastropietro et al., 2016; Meyer and Gore, 2014).

In a similar approach to this chapter, Andreis et al. (2020) uses the real options theory to analyze the bidding behavior in a Reliability Option mechanism. They describe multiple complex frameworks to derive the opportunity cost of participating in those mechanisms. However, the fundamentals for the bids on those mechanisms are different from the setup³ and they do not address the product design dimension. Finally, this work is close to the paper of Matthäus et al. (2021). They provide new insights on bidding behavior for renewable auctions by also using real options. However, they again study a different framework from this chapter.⁴

Current debates on the design of capacity markets have not yet determined the optimal capacity product if it exists. There is a coexistence of these products in most current markets, which underlines the importance of modeling their potential effects on investment values. Table 1.1 recapitulates the different variations and illustrates some example of current capacity markets and their relative product design. Note that some markets also include a distinction between new capacity and existing capacity. The former can either buy the long or the short product, while the latter is usually only allowed to buy the shorter product.

The impact of the length of the contractual period is even less discussed formally in the literature. Bushnell et al. (2017) emphasizes the importance of carrying out such analysis to improve the understanding of a capacity market. To our knowledge, Bialek and Unel (2019) and Bialek and Unel (2020) are the only ones to have addressed this issue qualitatively and quantitatively. In Bialek and Unel (2019), they underline the unsettled tradeoff between the financing costs and complexity costs of shorter products, and the

²These mechanisms, close to capacity markets, are based on the exchange of financial options between the actors holding the investment and demand. Initially held by the players, these options are sold on a market, which constitutes remuneration for their capacity.

³For instance, the strike price is explicit in reliability options mechanisms, and the comparison of different transaction phases does not entail the same implications.

⁴Producers bid for the price they will receive once the investment is made without knowing their production costs. Moreover, the option covers a single period.

| Transaction phase | | | |
|-------------------|-----------|---------|-----------|
| Monthly | Quarterly | Yearly | Multiyear |
| - | - | France | France |
| CAISO | - | CAISO | - |
| - | SPP | - | - |
| - | - | PJM | - |
| - | - | UK | UK |
| - | - | Poland | Poland |
| - | - | Belgium | - |
| - | - | ISO NE | - |
| NYSO | - | NYSO | - |
| - | - | - | Ireland |
| - | - | - | Italy |
| - | - | Greece | - |

Table 1.1: Product designs used in the model and actual implementations

costs of capacity over procurement and costs of excluding flexible generators of longer products. In [Bialek and Unel \(2020\)](#), they investigate the implications of the length of capacity products procured when there is seasonal variation in both the electricity load and the electricity generation. Using a Nash-equilibrium approach with investments and bidding behavior, they illustrate the efficiency tradeoffs associated with introducing multiple shorter capacity products instead of procuring a single annual capacity product and derive the optimal length of a capacity product. This model has the same spirit, but leaving aside the market representation, we focus on the coexistence between engaging in the capacity market and on managerial option (closing) and adapting the model with more detailed technological characteristics.

Similarly, [Abani et al. \(2018\)](#) build a complex model using the System Dynamic approach to understand how closing can modify the implementation effect of a capacity market regarding an initial sub-optimal energy-only market. However, contrary to this model, the author does not expand his analysis on multiple product designs. We also allow a high degree of flexibility in the mothballing decision by representing different closing periods.

We provide in section 2 the main model assumptions of this chapter. Section 3 recalls the key ideas of the net present value framework. Section 4 provides the real options framework. In section 5, we discuss some comparative statics on the capacity market

bids for both environments and the case study using the French data in 6. Finally, we study two extensions in section 7, and conclude in section 8

2 Model assumptions

2.1 Investment and wholesale market

We focus on a hypothetical setup with a single risk-neutral producer. He can invest in a unique power plant of a specific technology used to sell electricity at a future price on the wholesale market at date t and with a price p_t . If the producer enters the market by building his investment, he sustains an initial investment cost of c^I . Every \bar{n}^{om} dates, he can choose to stay open during a following period of length n^{om} , called the closing period. If the producer decides to remain open, he sustains a fixed cost of c^{om} called the periodic fixed cost. Those costs typically include operation and maintenance costs, leases, or wages. He can also produce whenever the wholesale price is above the marginal production cost c^v and sell its electricity on the wholesale market. The variable costs usually include fuel costs and carbon costs. Otherwise, if the investor chooses to close temporarily, he avoids the fixed cost but cannot produce. We normalize the capacity level, so one unit of capacity produces one unit of electricity. It is similar to assuming an absence of economies of scale, where producers with discrete capacity value would make piece-wise bids.

We define the inframarginal rent collected at date t as the net wholesale revenue as $\pi_t = (p_t - c^v)^+$. We assume it is uncertain for the investor at any date prior to t . We model this uncertainty using a stochastic process $(\pi_t)_{t \geq 0}$. This stochastic process follows a Geometric Brownian Motion such as it satisfies the stochastic differential equation (Andreis et al., 2020) :

$$\Delta\pi = \mu\pi_t dt + \sigma\pi_t \Delta Z_t \tag{1.1}$$

With μ and σ , respectively, the drift and the volatility of the Brownian Motion and ΔZ_t are the increments of a standard Brownian motion. This assumption regarding the uncertainty of the profit drivers is commonplace in commodity markets, especially when studying investment decisions in the electricity sector (see, for instance, Kettunen et al.

(2011), and Matthäus et al. (2021)). It allows capturing the randomness of the future variable cost, which follows the price of other commodities such as oil and gas, and the intrinsic uncertainty of electricity prices, which depends, for instance, on weather conditions, demand patterns, and carbon prices.

We consider a risk-neutral investor, so we define a constant risk-free interest rate r , which is also used as a discount rate in this model. Therefore, the rent process is defined by :

$$\Delta\pi^* = r\pi_t^*dt + \sigma\pi_t^*\Delta Z_t^* \quad (1.2)$$

We follow the canonical notation where ΔZ_t^* is the increments of the Brownian motion under the equivalent martingale measure \mathbb{Q} . We also want to study the sum of inframarginal rents' distribution; we make the following assumption.

Assumption 1.1. *The sum of inframarginal rent collected by a producer is log-normally distributed.*

More precisely, if $\int_{i=0}^n \pi_t$ represents the sum of the stochastic process values over a period n with μ and σ respectively, the drift and the volatility, then the sum follows a log-normal distribution. Alternatively, this is similar to say that if $\int_0^n \pi_t = e^X$ then:

$$X \sim \mathcal{N}(m, v)$$

with $m = 2 \ln [M] - 0.5 \ln [V^2]$ and $v^2 = \ln [V^2] - 2 \ln [M]$ such that M is the expected value and V^2 the second-order moment of the sum. We use this assumption as there is no explicit analytic expression of the distribution of the sum of Geometric Brownian Motion. This analytic approximation is commonplace in finance theory and relies on approximating the unknown distribution by another tractable one (see, for instance, Levy (1992) and Ju (2002)). More specifically, we use a moment-matching method where the moments of the sum distribution are matched with the moments of the log-normal distribution. In the Appendix, we provide more details for these assumptions when demonstrating the results of Proposition 1.3. Finally, we make the following assumption regarding the relationship between the different closing periods.

Assumption 1.2. *A closing decision for a specific period does not affect the profit or the producers' cost for other periods.*

For instance, closing the investment does not raise the production cost over the following periods nor decrease the revenue perceived on the wholesale market.

2.2 Producer behavior

Following the canonical theory for investment decisions in the electricity sector, the sum of inframarginal rents collected during the entire lifetime T of the power plant should cover the initial investment cost incurred at $t = 0$. However, when we take into account the periodic fixed cost, the producer enters the market only if the following equality holds :

$$\overbrace{\int_0^T e^{-rt} \mathbb{E}_0^* [\pi_t] dt}^{\text{revenue}} = c^I + \overbrace{c^{om} \sum_{i=0}^{\bar{n}^{om}-1} \int_0^{n^{om}} e^{-r(t+i \times n^{om})} dt}^{\text{fixed cost}} \quad (1.3)$$

With \mathbb{E}_0^* the expectation operator at date $t = 0$ with respect to the equivalent martingale measure \mathbb{Q} , T the power plant's lifetime, \bar{n}^{om} the number of time the producer has to choose to stay open and to pay the periodic fixed cost. The first left term represents the sum of net expected revenue made on the wholesale market, the first right term is the investment cost, and the last right term represents the actualized sum of the periodic fixed cost. For tractability, we made c^{om} periodic occurred at each period t , but recall that it is always sustained whenever the investment stays opened during a period of n^{om} . The equality in Equation 1.3 is similar to stating that the investment NPV is at least null. The investment NPV at time $t = 0$ can be defined as follows:

$$\Pi_0 = \int_0^T e^{-rt} \mathbb{E}_0^* [\pi_t] dt - c^I - c^{om} \sum_{i=0}^{\bar{n}^{om}-1} \int_0^{n^{om}} e^{-r(t+i \times n^{om})} dt$$

Following a similar approach, the power plant chooses to be available whenever it is profitable. The condition for an opening for each period n^{om} is given by the following equality.

$$\overbrace{\int_0^{n^{om}} e^{-rt} \mathbb{E}_0^* [\pi_t] dt}^{\text{revenue}} = \overbrace{\int_0^{n^{om}} e^{-rt} c^{om} dt}^{\text{fixed cost}} \quad (1.4)$$

For tractability, we assume that the decision is made at the beginning of a closing period.

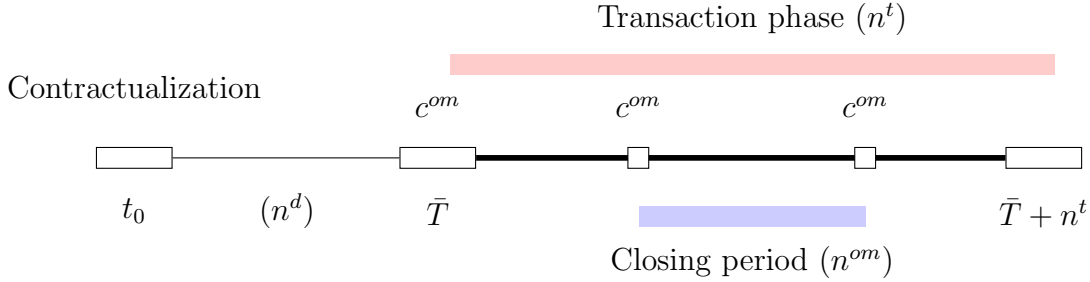
While the canonical theory states that a producer should be producing and, therefore, be available each time the electricity price is above its marginal cost, the introduction of the period fixed costs can induce a risk of inefficiency in the power system. Namely, assuming that the power plant is indeed necessary for the system, as soon as the power plant closes because the inframarginal period collected on a specific period is below the corresponding fixed cost, it generates a net welfare loss for the system.⁵ Given the previous assumptions on the investment, one can compute the optimal number of periods t over the lifetime during which $\pi_t \geq 0$ and compare it to the number of periods over the lifetime during which the investment is open and $\pi_t \geq 0$. A difference between those results would show that the power plant behaves inefficiently. Such inefficiency could be due to a price cap on the energy market, the effect of the non-economic intervention of the system operator, or unpriced externalities.

2.3 The capacity market

If a regulator decides that this power plant is necessary for the system, he implements a capacity market to encourage the producer to invest and be available. To do so, he defines a capacity product with a specific duration called the *transaction period* and notes n^t . He organizes the transaction of this product via a market mechanism such as an auction at a contractualization date noted t_0 . Once the producer has sold the product, he is legally bound to be available during the transaction phase, that is, to be on the market during a transaction phase of length n^t . This period starts at a predefined date noted \bar{T} , with n^d the distance between the auction date and the starting date of the transaction phase. The regulator can use multiple instruments to check for availability, such as unannounced tests or verifying book orders on the energy market. One can note that the transaction period is not necessarily equal to the closing period, which is investment-specific. It can either be lower, equal, or superior. The following figure

⁵This approach to the market efficiency implies that while the wholesale price is below an optimal value that covers the fixed cost, it optimally sends short term signals. To say it differently, if the wholesale price would have been optimal, then occurrences of prices below the marginal cost are the same as the occurrences with the inefficient price. This analysis could be extended to the case where prices are also inefficient concerning the marginal cost. Still, it implies additional assumptions to differentiate between periods when both optimal and inefficient prices are above the marginal cost and when they are not.

illustrates the design of a capacity market where the transaction period implies three closing decisions for the investment.



For the producer, the capacity price received in the capacity market enters its profit as a second stream of remuneration in addition to the revenue made on the wholesale market. Similarly to the closing periods, we define \bar{n}^t the number of times a capacity auction is set during the investment lifetime. Using the NPV of the investment over its lifetime, the final NPV with the capacity market is equal to :

$$\Pi_0^{cm} = \Pi_0 + \overbrace{\sum_{i=0}^{\bar{n}^t} e^{-r(i \times n^t)} p_i^c}^{\text{capacity market revenue}}$$

With p^c , the capacity price is received at every capacity auction. By construction, auctions are set up at an interval of n^t . As we do not model the competitive process in the chapter, we simplify the analysis by assuming that the bids the investor makes on the capacity market are equal to the price he receives. Such assumption holds under the case the investment is always the marginal bidder in a uniform auction, or if he bids truthfully in a pay-as-bid type auction or in a bilateral marketplace.⁶

3 The net present value framework

We start the analysis by describing the bidding behavior of a producer who offers only its net present opportunity cost on the capacity market associated with an existing investment. It allows precise definitions and the bid's rationales in a capacity market.

⁶See for instance [Matthäus et al. \(2021\)](#) for a discussion on the truthful behavior in competition based mechanisms for investments in electricity production.

We discuss the relation between the product design and the expected inframarginal rent net of the periodic fixed cost. This is a basis for the canonical approach to model bids in the capacity market. Those results also serve as a reference value to compare the bidding behavior when the option value is taken into account using a real options framework.

3.1 The inframarginal rent and the opportunity cost

The opportunity cost associated with participation in a capacity market is based on the dichotomy between sunk and non-sunk fixed costs incurred when the producer decides to produce. It is crucial as some fixed costs could be considered sunk before participating in a capacity market. Indeed, recall that fixed costs are decomposed into two parts : (i) Investment costs, which incur the power plant's first activation, and (ii) periodic fixed costs, which incur periodically and irrevocably. When considering entering the market, investment and periodic fixed costs are still pending and avoidable. Consequently, the time horizon used to compute the opportunity cost of entering the market should be based on the entire project lifetime. Indeed, when the producer compares the decision to enter the market at date $t = 0$, he faces a trade-off of closing (i) receiving the asset value and (ii) never entering the market, which translates into a null value. On the other hand, if the producer has already invested, periodic fixed costs are the only fixed costs that are avoidable, and the time horizon is limited to the closing period. Therefore, when the producer forecasts the decision to participate in the capacity market at date t_0 , he faces a tradeoff between having to open the power plant and potentially incurring net losses or leaving the market temporarily at no cost.⁷

The cost associated with such an opportunity over a transaction period is the difference between the sum of the periodic costs linked to the decision to stay open during the obligation to produce and the profits made only during the period covered by the capacity product. Formally, we note B_0 the initial bid made before the investment is made, and b_{t_0} the bids made at every auction date t_0 during the investment lifetime. For instance, assume a product with a transaction period of n^t sold at t_0 with a periodic fixed cost sustained over an identical period ($n^{om} = n^t$); if we assume that the producer offers its net present opportunity cost on the capacity market, then the following equality based on the condition 1.4 must hold :

⁷We assume in the model extensions a cost associated with the possibility of closing temporarily.

$$\int_0^{n^t} e^{-rt} \mathbb{E}_0^* [\pi_t] dt + p^c = \int_0^{n^t} e^{-rt} c^{om} dt$$

Having assumed that producers truthfully bid, then the bid value for an existing plant on a capacity market is :

$$b_{t_0} = e^{-rn^d} \left[\int_0^{n^t} e^{-rt} c^{om} dt - \int_0^{n^t} e^{-rt} \mathbb{E}_0^* [\pi_t] dt \right]^+$$

Under the net present value framework, the opportunity cost of participating in a capacity market is equal to the expected short-term *Missing Money*, that is, the expected loss of staying available due to the existence of fixed periodic costs. Given the bids for an existing power plant, we can now define the bid for a new investment. It is based on the wholesale revenue but also the expected bids on the capacity market. The following equality based on the condition 1.3 must hold and implies that the investment NPV given the capacity prices is null :

$$\int_0^T e^{-rt} \mathbb{E}_0^* [\pi_t] dt + \sum_{i=0}^{\bar{n}^t-1} e^{-r(i \times n^t)} b_{i \times n^t} + p^c = c^I + c^{om} \sum_{i=0}^{\bar{n}^{om}-1} \int_0^{n^{om}} e^{-r(t+i \times n^{om})}$$

Having assumed that producers truthfully bid, then the bid value for an existing plant on a capacity market is :

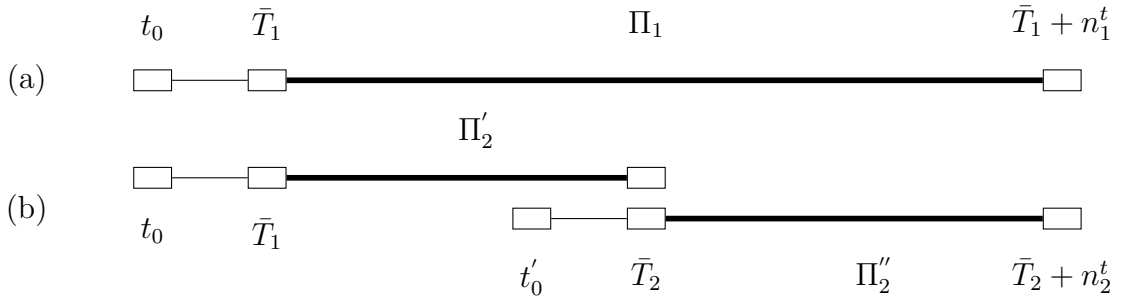
$$B_0 = \left[c^I + c^{om} \sum_{i=0}^{\bar{n}^{om}-1} \int_0^{n^{om}} e^{-r(t+i \times n^{om})} - \int_0^T e^{-rt} \mathbb{E}_0^* [\pi_t] dt - \sum_{i=0}^{\bar{n}^t-1} e^{-r(i \times n^t)} b_{i \times n^t} \right]^+ \quad (1.5)$$

With \bar{n}^t , the number of times a capacity market auction is implemented. Under the net present value framework, the opportunity cost of participating in a capacity market is equal to the expected long-term *Missing Money*, which is equal to the investment fixed cost and the sum of the periodic fixed costs net of the revenue earned on the wholesale and capacity market. This last point is particularly relevant as the chapter aims to understand the link between product design and capacity bids. Therefore, if the costs

and profit are held equal, a different product design should bring different long-term *Missing Money*, hence different first bids. We provide a more detailed discussion for a new entrant in the extension section. In the rest of the analysis, we focus on bids for existing investments.

3.2 A simple example

We start to illustrate the analysis with an example of bidding behavior with an existing plant. We assume a first product design implementation (case (a)) with a single transaction phase of n_1^t . A second implementation (case (b)) is based on two shorter products of the same length n^t . For simplicity, we assume that $n_1^t = 2n^t$. We denote \bar{T}_1 the start of the transaction phase for the single product of case (a) and the first product of case (b), and $\bar{T}_2 = \bar{T}_1 + n_1^t$ the start of the second product of case (b). The periodic fixed cost is incurred at \bar{T}_1 and \bar{T}_2 , meaning that we have $n^t = n^{om}$. We denote the total profit collected on the whole period as Π_1 , while we denote the profit collected on the first sub-period Π'_2 , and on the second sub-period Π''_2 . Finally, we denote t_0 , the date when the auction for the single product of case (a) and the first product of case (b), and t'_0 the auction date for the second product of case (b). For any case, the period between an auction and the starting date is equal and noted n^d . We illustrate the two implementations in the following figure.



Following the framework the expected bid at t_0 in the case (a) noted b^1 is equal to:

$$b^1 = e^{-rn^d} \left[\mathbb{E}_0^* \left[c^{om} \int_0^{2n^t} e^{-rt} dt - \int_{n^d}^{2n^t+n^d} \pi_t dt \right] \right]^+$$

Which gives when rearranged :

$$b^1 = e^{-rn^d} \left[c^{om} \int_0^{n^t} e^{-rt} dt + c^{om} \int_0^{n^t} e^{-r(t+n^t)} dt - 2n^t e^{rnd} \pi_0 \right]^+$$

While the sum of the two bids in the case(b) noted b^2 is :

$$b^2 = e^{-rn^d} \left[\mathbb{E}_0^* \left[c^{om} \int_0^{n^t} e^{-rt} dt - \int_{n^d}^{n^t+n^d} e^{-rt} \pi_t dt \right] \right]^+ \\ + e^{-r(n^t+n^d)} \left[\mathbb{E}_0^* \left[c^{om} \int_0^{n^t} e^{-rt} dt - \int_{i=n^d+n^t}^{2n^t+n^d} e^{-rt} \pi_t dt \right] \right]^+$$

Which gives when rearranged :

$$b^2 = e^{-rn^d} \left([c^{om} - n^t e^{rnd} \pi_0]^+ + [c^{om} e^{-rn^t} - n^t e^{rnd} \pi_0]^+ \right)$$

Proposition 1.1 states that the following inequality always holds for any value of the expected inframarginal rent and periodic fixed cost: $b^2 \geq b^1$.

Proposition 1.1. *Assuming the absence of risk aversion and for an existing investment, a product with a longer transaction phase will always lead to a lower bid than the sum of the bids for products with a shorter transaction phase.*

Proof. The proof is straightforward and is given by the triangle-inequality like of the maximum function: $\max(x, 0) + \max(y, 0) \geq \max(x+y, 0)$. Using the previous example, x takes the value of : $c^{om} - n^t e^{rnd} \pi_0$; and y takes the value of : $c^{om} e^{-rn^t} - n^t e^{rnd} \pi_0$. \square

It should be noted that when the closing period differs from the transaction period, the producers' offer can be significantly affected without impacting the previous proposal. When $n^{om} > n^t$, then the opportunity cost is estimated based on the n^{om} period with the implication that in subsequent auctions, a closing period that overlaps two transaction periods is not taken into account in the offer for the second auction, the opportunity cost being zero because the plant is already open. Similarly, when $n^{om} < n^t$, the opportunity cost is calculated based on a period kn^{om} such that k determines the smallest period greater than the transaction period ($k \equiv \min(kn^{om} - n^t)$ s.t $k \in \mathbb{Z}^+$). When n^t is a multiple of n^{om} , this does not change the offer made by the producer (see, for instance, the

previous example). Otherwise, it is similar to the case $n^{om} > n^t$. Using these observations, we will not study complex cases where the closing periods and transaction phases are not equal in the rest of this chapter. While they can explain some actual bidding behavior in volatility and magnitude, they do not change the fundamental intuitions.

4 The real options framework

4.1 A capacity bid as a European Put Option

The previous framework provides the marginalist intuition behind the bidding behavior on a capacity market. However, it does not consider all the possible rationales, especially when the transaction phase of a capacity product is associated with irreversible managerial decisions. In this section, we conceptualize the capacity product as a real options that allows the option value to not be available over a closing period to avoid potential losses. We model the option value using the canonical option pricing theory. Following the most simple case with a transaction over one period t with $n^t = n^{om} = 1$, then the availability decision associated with the capacity product for an existing investment is a European Put Option with a payout profile of $\max(c^{om} - \pi_t, 0)$. In this case, the periodic fixed cost can be compared to the strike price of a financial option, and the expected inframarginal rent can be compared to the underlying asset. Following the standard approach and the marginalist assumption, Proposition 1.2 states the bid on a capacity market is equal to the option value of being available :

Proposition 1.2. *Given the payout profile associated with the capacity product, a bid noted b^{opt} in an auction set at $t_0 = 0$ for a unique transaction period at starting n^d periods after the auction and a periodic fixed cost of c^{om} is:*

$$b^{opt}(\pi_0, c^{om}) = -\pi_0 \phi(z) + e^{-rn^d} (c^{om} \phi(z + \sigma \sqrt{n^d}))$$

$$z := -\frac{\ln[\pi_0] - \ln[c^{om}] + (r + \frac{\sigma^2}{2})n^d}{\sigma \sqrt{n^d}}$$

With ϕ , the cumulative distribution function of a standard normal distribution.

Proof. See Appendix 1. □

4.2 A capacity bid as a Basket Option

We now expand this approach to the more complex case where a transaction period covers multiple uncertain inframarginal rents. We start with the first case by hanging up the setup with the financial theory for exotic derivatives. We assume that when the transaction period is expanded over multiple periods, the European Put Option becomes a modified Basket Option. In finance, a Basket Option is defined by a payoff profile dependent on the value of a portfolio of assets, each following a stochastic process such as a Geometric Brownian Motion, which can be correlated or independent. Hence, the availability decision associated with a capacity product is similar to exercising a Basket Option, in the sense that the irrevocable decision to stay open at a date \bar{T} implies collecting each inframarginal rent during the transaction period, which individually follows a Geometric Brownian Motion. With a Basket Option, its exercise would have meant the collection of the individual stock prices. The primary constraint associated with pricing such an option is the absence of a closed-form representation of the price since a sum of log-normally distribution random variable is not log-normal. However, we use the approximation stated in Assumption 1.1 to define an analytic approximation of the option price. It allows deriving the following proposition regarding the bid on a capacity market when the transaction period covers multiple inframarginal rent periods.

Proposition 1.3. *Given the payout profile associated with the capacity product, a bid noted b^{opt} in an auction set at $t_0 = 0$ for a transaction period starting of length n^t , starting n^d periods after the auction is set and with an equal closing period of n^t is:*

$$b^{opt}(\pi_0, c^{om}) = -\pi_0 n^t \phi(z) + C^{om} \phi(z + v) \quad (1.6)$$

$$z = -\frac{m - \ln \left[c^{om} \int_0^{n^t} e^{-rt} dt \right] + v^2}{v}$$

$$C^{om} = e^{-rn^d} c^{om} \int_0^{n^t} e^{-rt} dt$$

With ϕ the cumulative distribution function of a standard normal distribution, m and v^2 defined as follow :

$$m = 2 \ln \left[\pi_0 \int_0^{n^t} e^{r(i+n^d)} dt \right] - 0.5 \ln \left[\pi_0^2 \int_0^{n^t} \int_0^{n^t} e^{r(t+s+n^d)+(t+n^d)\sigma^2} dt ds \right]$$

$$v^2 = \ln \left[\pi_0^2 \int_0^{n^t} \int_0^{n^t} e^{r(t+s+n^d)+(t+n^d)\sigma^2} dt ds \right] - 2 \ln \left[\pi_0 \int_0^{n^t} e^{r(i+n^d)} dt \right]$$

Proof. See Appendix 2. □

This definition of a bid on a capacity market can be understood as follows. First, note that the inframarginal rent term $\pi_0 n^t$ is linked to the Geometric Brownian Motion and the risk-free version of the inframarginal rent process.⁸ The fixed cost term is the sum of the actualized periodic fixed costs associated with the decision to stay open. Finally, the value z and $z + v$ are linked with Assumption 1, with the first term in the logarithm of m (second term in v^2) representing the mean of the sum of the inframarginal rent and the second term in the logarithm of m (first term in v^2) is the second moment of the sum of the inframarginal rent.

4.3 The sum of expected capacity bids

Once we derive the bid for a single product, we can now analyze the sum of the bids for multiple capacity products. Assumption 2 states that even though some correlation exists between inframarginal rents over the investment lifetime, the decision to close during one period does not modify the rent value in a subsequent period. It allows defining the sum of the bids for multiple capacity products as the sum of the value of their options estimated at a single date; for simplicity, here, the auction at which the longer product is sold, or the first shorter product is sold.

Proposition 1.4. *Given the payout profile associated with a capacity product of length $\frac{n^t}{k}$ with the same closing period of $\frac{n^t}{k}$, the sum of expected bids noted b^{opt} made during k successive auctions is:*

⁸Under the equivalent martingale measure \mathbb{Q} , the drift of the inframarginal rent is equal to the risk-free rate, meaning that any actualized expected value of the rent is equal to its initial value π_0 .

$$b^{opt}(\pi_0, c^{om}) = -\pi_0 \frac{n^t}{k} \sum_{j=1}^k \phi(z_j) dt + c^{om} e^{-rn^d} \int_0^{\frac{n^t}{k}} e^{-rt} dt \sum_{j=1}^k e^{-r(j-1)\frac{n^t}{k}} \phi(z_j + v_j)$$

With :

$$z_j = -\frac{m_j - \ln \left[c^{om} \int_0^{\frac{n^t}{k}} e^{-rt} \right] + v_j^2}{v_j}$$

With m_j and v_j^2 defined as follow :

$$m_j = 2 \ln \left[\pi_0 \int_{n^t(j-1)}^{n^t j} e^{r(t+n^d)} dt \right] - 0.5 \ln \left[\pi_0^2 \int_{n^t(j-1)}^{n^t j} \int_{n^t(j-1)}^{n^t j} e^{r(s+t+n^d)+(t+n^d)\sigma^2} dt ds \right]$$

$$v_j^2 = \ln \left[\pi_0^2 \int_{n^t(j-1)}^{n^t j} \int_{n^t(j-1)}^{n^t j} e^{r(s+t+n^d)+(t+n^d)\sigma^2} dt ds \right] - 2 \ln \left[\pi_0 \int_{n^t(j-1)}^{n^t j} e^{r(t+n^d)} dt \right]$$

Proof. With no correlation between the decision to close and the profits made during other periods, the different capacity products could be conceptualized as options on different assets (See [Trigeorgis \(1993\)](#)). \square

5 Comparative statics

Using the results of the previous section, we compare the bidding behavior in a capacity market depending on (i) the drivers behind the opportunity cost, (ii) if producers follow the net present value framework or the real options framework, and (iii) the design of the transaction phase for a given period.

We start by studying the evolution of the bids under the real options framework, especially with respect to the length of the product. We also describe the effect of the main variables on the bid. This step is particularly relevant due to the length of the

transaction phase, which significantly complexifies the analysis. Then, we show how the real options bidders value the possibility of closing to avoid the fixed costs. Namely, we describe the flexibility associated with the option to close and how its value is impacted by the length of the transaction phase and other variables. Finally, we discuss the impact of segmenting a capacity product in multiple shorter transaction phases. Similarly to the two previous analyses, we also provide the effect of the drivers on the delta for the bids between different capacity product designs.

5.1 Bids value

This analysis sheds light on how a capacity bid varies with its fundamentals. One of the critical variables in this research is the length of the product n^t . We also look at the two main variables of the opportunity cost as described in the net present value section, namely the initial value of the inframarginal rent π_0 and the fixed cost value c^{om} . Next, we study the regulatory parameters chosen when the capacity market is implemented n^d , representing the waiting time before the transaction phase. However, we left for future work a deeper analysis of this parameter on market efficiency⁹, this analysis still provides some insight into its role in the bidding behavior on the capacity market. Finally, we also study the impact of different volatility levels of the inframarginal rent σ on the capacity bid. Indeed, one of the current policies in the power sector relies on significantly increasing the share of renewables in the production mix. A key consequence would be an immediate increase in the volatility of the wholesale price (Fontini et al., 2021). In turn, it also has an indirect effect on the bidding behavior in the capacity market.¹⁰

For relevancy and simplicity, we assume that the periodic fixed cost occurs simultaneously as the inframarginal rent in this section, namely every t period. Therefore, when the transaction phase increases by one period, the producer gains an uncertain inframarginal rent and sustains an additional unitary periodic fixed cost. We provide in Proposition 1.5 an overview of the effect of increasing the transaction period on the bids for the net present value case and the real options case. Note that for the real options framework, we could not find a closed-form solution in terms of the value of the variables for which a clear-cut answer on the sign of the derivatives exists. However, we

⁹The waiting time has been numerously cited as a critical regulatory parameter when designing capacity markets. However, to our knowledge, very few papers have looked into this issue from a modeling perspective.

¹⁰A second indirect effect is the decrease of the average wholesale price due to a merit order effect, which translates into a lower initial value π_0 .

can provide sufficient general conditions so that such a solution exists. The conditions presented in this section are usually given on the derivatives of the variable z , which represents the threshold on the probability that the sum of the actualized fixed cost (the strike price) is above the sum of uncertain inframarginal rent (the underlying asset), and on specific ratios between the density functions $\phi(z)$ and $\phi(z + v)$, which express the probability of a standard normal distribution at their respective value.

Proposition 1.5. NPV: *The bid is a concave function with respect to the length of the transaction phase n^t . The threshold in terms of fixed cost between an increasing bid and a decreasing bid is given by: $c^{om} = \pi_0 e^{r(nd+nt)}$. **RO:** *An increase in the transaction phase increases the capacity bid if the following conditions hold:**

- *Fixed costs :* $c^{com} \int_0^{n^t} e^{-rt} dt \geq \sqrt{\pi_0} V e^{v^2(2r\frac{\partial v^2}{\partial n^t} - 1)}$
- *Cdf ratio :* $\frac{S_1 C^{om}}{\pi_0} \geq R_0$
- *Df ratio :* $\frac{C^{om}}{n^t \pi_0} \geq R_1$

With $R_0 = \frac{\phi(z)}{\phi(z+v)}$, $\geq R_1 = \frac{\varphi(z)}{\varphi(z+v)}$, C^{om} the actualised sum of the fixed cost and $S_1 = e^{-rn^t} / \int_0^{n^t} e^{-rt} dt$.

Proof. See Appendix 3. □

The intuition behind this result is as follows. For the net present value bid, it is straightforward because the bids are equal to $\pi_0 n^t - C^{om}$. Therefore, the bid starts decreasing as soon as the marginal value of the expected profit at the end of the transaction phase exceeds the additional fixed marginal cost.

For the option value, we provide the derivative of the bid value with respect to n^t in the following expression:

$$\frac{\partial b^{opt}}{\partial n^t} = -\pi_0 \left(\phi(z) + n^t \frac{\partial z}{\partial n^t} \varphi(z) \right) + C^{om} \left(S_1 \phi(z + v) + \frac{\partial z + v}{\partial n^t} \varphi(z + v) \right)$$

The option value is composed of two distinct parts : (i) a negative part which stands for the expected sum of inframarginal rent adjusted by the cumulative distribution function of the standardized normal distribution: $-\pi_0 n^t \phi(z)$; (ii) a positive part which stands

for the strike price represented through the cost of staying available, also adjusted by the cumulative distribution function: $C^{om}\phi(z + v)$; with C^{om} the actualized sum of the fixed cost. For both parts, an increase of n^t has two effects: a direct effect linked to the marginal increase in the profit and the fixed cost and an indirect effect via a change in cumulative distribution function value. The sign of this indirect effect depends on the sign of the derivatives of z and $z + v$ with respect to n^t . We now analyze how the conditions can be widened or tightened to have a clear-cut effect on the sign of n^t on the bid.

The first effect is straightforward: an increase in n^t also increases the profit and fixed costs. For the indirect effect, we start with the fixed cost part. Recall that the cumulative distribution function represents the probability that the periodic fixed cost is above the inframarginal rent. The sign of the corresponding derivative with respect to n^t is positive only if the following condition on the fixed cost in the proposition holds. In this case, an increase of n^t always implies an increase in the probability that the sum of the actualized fixed cost is above the sum of uncertain inframarginal rent. Hence, both effects of part (ii) are always positive, implying that an increase of n^t increases the option value.

However, the effect can be ambiguous for some values of n^t due to part (i). The sign of the corresponding derivative with respect to n^t is negative only if the following condition in the proposition on the fixed cost holds. When the derivative is negative, it decreases the probability that the inframarginal rent is above the fixed cost given by $\phi(z)$. Hence, it lowers the bid value. This value is decreasing with n^t . Therefore, the possibility that the bid is positively impacted by n^t increases with n^t .¹¹

We turn now to the other drivers for the capacity bid, and we provide the result in Lemma 1.1. Again, as there are no closed-formed solutions, we provide conditions for which the drivers have clear-cut signs on the bid value.

Lemma 1.1. *The value of the bid under real options :*

1. *decreases with the inframarginal rent π_0*

¹¹Note that the ratios R_0 and R_1 are similar to other ratios that can be found in the model. For instance, the constrain that the option value is always positive implies that we have the following condition :

$$\frac{C^{om}}{n^t\pi_0} \geq \frac{\phi(z)}{\phi(z + v)}$$

2. increases with the periodic fixed cost c^{om}
3. is ambiguous with the waiting time n^d
4. is ambiguous with the inframarginal rent volatility σ

Results (1) and (2) hold if the condition $\frac{C^{om}}{n^t \pi_0} \geq R_1$ is satisfied.

For Result (3), n^d always decreases the bid when $c^{om} \int_0^{n^t} e^{-rt} dt \geq V e^{-M \frac{r+\sigma}{r-\sigma}}$, $\frac{C^{om}}{n^t \pi_0} \geq R_1$ and $r > \sigma$ are satisfied. If only the first condition holds and that $-r\phi(z+v) \geq \frac{\partial v}{\partial n^d} \varphi(z+v)$, then n^d always decreases the bid.

For Result (3), n^d always increases the bid when $c^{om} \int_0^{n^t} e^{-rt} dt \geq V e^{-M \frac{r+\sigma}{r-\sigma}}$ and $\frac{\partial v}{\partial n^d} \varphi(z+v) \geq -r\phi(z+v)$. Otherwise n^d has an undetermined effect on the bid price and depend on the relation between $\frac{C^{om}}{n^t \pi_0}$ and the ratio $R_3 = \frac{\frac{\partial z}{\partial n^d} \phi(z)}{-r\phi(z+v) + \frac{\partial v}{\partial n^d} \varphi(z+v)}$

For Results (4), σ always increases the bid if $c^{om} \int_0^{n^t} e^{-rt} ft \leq \frac{M^2}{V}$ and $\frac{C^{om}}{n^t \pi_0} \geq R_1$ are satisfied. Otherwise if $c^{om} \int_0^{n^t} \leq V$, then σ always increases the bid if $\frac{C^{om}}{n^t \pi_0} \geq \frac{\varphi(z) \frac{\partial z}{\partial \sigma}}{\varphi(z+v) \frac{\partial z+v}{\partial \sigma}}$. If the conditions are not respected, then σ is always decreasing the bid.

Proof. See Appendix 4. □

The first and second results are standard regarding real options theory (see, for instance, [Matthäus et al. \(2021\)](#)). They have an opposite interpretation: an increase in the initial inframarginal rent signals that future revenues will also increase. Consequently, it decreases the value to close to avoid the fixed cost, as those costs are more likely to be covered by the inframarginal rent. On the other hand, as the fixed cost increases, then the value increases.

The analysis of n^d is less intuitive and relies on stronger conditions on the fixed cost. The effect of n^d is unique on the inframarginal rent part of the bid and depends on the sign of the corresponding derivative. If the derivative of this part is negative, it implies that if n^d increases, then the probability that the rent is above the fixed cost decreases. Otherwise, the reverse effect happens. It is given by the first condition on the fixed costs for n^d to have a negative impact on the bid. To highlight the effect of n^d on the fixed cost part of the option value, we rearrange the conditions of the result (3).

$$\frac{\partial b^{opt}}{\partial n^d} = C^{om}(\varphi(v+v) \frac{\partial v}{\partial n^d} - r\phi(z+v)) + \frac{\partial z}{\partial n^d} \Delta \varphi$$

with $\Delta\varphi = (C^{om}\varphi(z+v) - n^t\pi_0\varphi(z))$. This value is central in the analysis, and it is always positive as long as $\frac{C^{om}}{n^t\pi_0} \geq R_1$ holds and can be found in the four derivatives of the bid value. It gives the second condition for n^d to have a negative effect on the bid. The derivative represents the net marginal change in terms of the bid value when the variables marginally impact z .

We turn now to the first part of the rearranged derivative. When n^d increases, it always decreases the value of the fixed cost because of a discounting effect (second negative term) while also changing the volatility of the revenue (first term). Following the analysis of $\frac{\partial z}{\partial n^d}$, it is sufficient for its sign to be negative and also to have a negative sign $\frac{\partial \sigma}{\partial n^d}$ to have the decreasing effect of n^d on the bid. It has straightforward intuitions: if an increase of n^d decreases the value of the fixed costs, the volatility, and the probability that the inframarginal rent is below the fixed costs, then the option value also decreases, hence the bid. The derivative in the first term is negative if and only if $\frac{\partial v}{\partial n^d}$ is also negative. That is, an increase of n^d decreases the volatility of the total revenue. This is the case only when the risk-free rate (r) is above the volatility of the inframarginal rent (σ). It gives the third condition for n^d to have a negative effect on the bid. Therefore, the first group of conditions for n^d states the conditions under which every part of the derivative is negative.

Finally, when at least one derivative is not negative, then the sign is ambiguous and depends on the magnitude of each part of the derivative. For instance, when $\frac{\partial z}{\partial n^d}$ is still negative, but an increase of n^d decreases the volatility of the sum of inframarginal rent (i.e., $\frac{\partial v}{\partial n^d}$ is positive), it is sufficient that the negative effect of the inframarginal rent is higher than the gains in terms of volatility to ensure that the sign of n^d is negative. Therefore, it gives the second set of conditions. The opposite case when $\frac{\partial z}{\partial n^d}$ is positive gives the third set of conditions.

To conclude, we state that the effect of σ is also counter-intuitive when the conditions do not hold. We rearrange the derivative, which gives:

$$\frac{\partial b^{opt}}{\partial \sigma} = C^{om}\varphi(v+v)\frac{\partial v}{\partial \sigma} + \frac{\partial z}{\partial \sigma}\Delta\varphi$$

First, note that the first derivative is positive as an increase in the volatility of the periodic revenue always increases the volatility of the total revenue. Therefore, the ambiguity of σ on the bid value only depends on the sign of the second derivative $\frac{\partial z}{\partial \sigma}$ and on the magnitude of the positive parts. When analyzing the sign of the derivative ($\frac{\partial z}{\partial \sigma}$),

we find that it requires one condition to be negative. This condition can be found in the lemma. Recall that M^2 and V^2 are the first and second moments of the distribution of the inframarginal rent sum during the transaction phase.¹² Such a condition highlights the critical role of the difference between the fixed costs and the expected value and distribution of the inframarginal rent. When the fixed costs are relatively high compared to the mean value adjusted by the risk of the total revenue (i.e., the conditions are satisfied), then a marginal increase of the volatility always implies a loss for the option value: it increases the occurrence of having the sum of inframarginal being above.

5.2 Flexibility value

We turn to the analysis of the difference between a bid under a net present value framework and a bid under a real options framework. As in the canonical real options theory, the possibility of the managerial option always creates additional value for the producers. Consequently, when comparing the difference between bidding the missing money and bidding the option value associated with the possibility of closing, we have the following proposition.

Proposition 1.6. *Under the same market design, the bid in a capacity market when producers consider the option value is always higher or equal to the bid using only a net present value approach.*

$$b^{opt} \geq b^{npv}$$

Proof. The proof for the proposition is straightforward and comes from the definition of a net present value and a real options bid. Under the first framework, producers bid the maximum of their expected Missing Money. On the other hand, under the second framework, producers bid their option value, namely the expected maximum of their missing money. Therefore we always have $b^{opt} \geq b^{npv}$. \square

Using this result, we can now analyze the value of flexibility, which is the difference between the two bids. We define Γ as the value of the flexibility such as $\Gamma = b^{opt} - b^{npv}$. It is defined in the following equation :

$$\Gamma = -\pi_0 n^t (\phi(z) - 1) + C^{om} (\phi(z + v) - 1)$$

¹²Note the resemblance with the Sharp Ratio used in finance, which defines the performance of an investment compared to a risk-free asset after adjusting for its risk.

The analysis of the evolution of Γ with respect to the main variables is similar to the comparative statics made in the previous section. Indeed, for any variable x note that $\frac{\partial \Gamma}{\partial x} = \frac{\partial b^{opt}}{\partial x} - \frac{\partial b^{npv}}{\partial x}$. Therefore, the net effect of the previous results can either be increased or decreased depending on how the bid under the net present value framework behaves. We summarise the main results in Proposition 1.7.

Proposition 1.7. *The length of the transaction phase has an ambiguous effect on the value of the flexibility :*

- Γ is increasing in n^t when $c^{om} \leq \pi_0 e^{r(nd+nt)}$

When the condition does not hold, given the following ratio

$$R_4 = \frac{\frac{1}{n^t} \phi(z) + \frac{\partial z}{\partial n^t} \varphi(z) - \frac{1}{n^t}}{S_1 \phi(z+v) + \frac{\partial z+v}{\partial n^t} \varphi(z+v) - S_1}$$

- Γ is increasing in n^t if $S_1(\phi(z+v) - 1) + \frac{\partial z+v}{\partial n^t} \varphi(z+v) > 0$ and the following condition holds : $\frac{C^{om}}{n^t \pi_0} \leq R_4$
- Otherwise, Γ is increasing in n^t if $S_1(\phi(z+v) - 1) + \frac{\partial z+v}{\partial n^t} \varphi(z+v) < 0$ and the following condition holds : $\frac{C^{om}}{n^t \pi_0} \geq R_4$.
- When the conditions do not hold, Γ is decreasing in n^t

Proof. The proof is similar to the Proposition 1.5. The results follow directly from the derivative of the bid with respect to the variable. \square

The ambiguity of n^t on the value of flexibility comes from the concavity of the net present value bid as described by the Proposition 1.5 and the monotonicity of the real options bid. When the former decreases with respect to n^t , the flexibility is always increasing as the latter increases. However, when the net present value bid increases, the sign of the flexibility derivative depends on the magnitude between a marginal rise in the net present value bids and the marginal increase of the real options bids. Those opposing effects are materialized in the ratio R_4 with the negative terms $-\frac{1}{n^t}$ and $-S_1$, which represent the marginal increase of respectively the profit and the fixed cost part. Given the results expressed in Proposition 1.5 and Proposition 1.7, we find that the

condition under which the flexibility increases with respect to n^t is weaker than the condition under which the real options bid rises with respect to n^t .

This ambiguity renders the analysis of the value of flexibility with respect to other drivers significantly complex. While we had intuitive results and mild conditions for the two fundamental drivers of the bids in Proposition 1.5, namely π_0 and c^{om} . The marginal increase of the net present value bids is always superior to the direct effect observed in the real options bids when the two drivers increase. For instance, recall that an increase of π_0 both directly lowers the value of the option due to a marginal decrease of $-n^t\phi(v)$ while indirectly modifying the option value with a change in the probability of the maximum function: $-\pi_0 n^t \frac{\partial z}{\partial \pi_0} \varphi(v)$. When we introduce the change in the net present value bids, the marginal decreases when π_0 is equal to $-n^t$. It implies that we always have $n^t \geq n^t\phi(v)$, as $\phi(v)$ is a cumulative density function. Therefore, the net effect on the value of the flexibility is always going to be dependent on the magnitude between the net direct effect $n^t(1 - \phi(v))$ and $-\pi_0 n^t \frac{\partial z}{\partial \pi_0} \varphi(v)$, and not anymore on the sign of the derivative.

However, for the volatility of the inframarginal rent, we find a strict identical effect between the real options bid and the flexibility value. Indeed, the net present value bid is independent of the volatility, hence $\frac{\partial \Gamma}{\partial \sigma} = \frac{\partial \pi_0}{\partial \sigma}$. Therefore, the sign and the conditions discussed in Lemma 1.1 can be applied to the analysis of the value of the flexibility with respect to σ .

5.3 Product design

In this section, we analyze the effect of segmenting the capacity product into multiple shorter products. More precisely, we assess the difference between the bid for a single capacity auction considered as a "long" product covering a transaction phase of n^t periods with the sum of expected bids for k successive capacity auctions where producers can sell "short" product of length $\frac{n^t}{k}$. The difference in the cost of a capacity market noted Δb^{opt} , given this configuration is defined in the following equation and uses the Proposition 1.3 and Proposition 1.4 :

$$\Delta b^{opt} = -\pi_{t_0} n^t \left(\phi(z) - \frac{1}{k} \sum_{i=1}^k \phi(z_i) \right) + C^{om} \left(\phi(z + v) - S_2 \sum_{i=1}^k e^{-r(i-1)\frac{n^t}{k}} \phi(z_i + v_i) \right) \quad (1.7)$$

$$\text{With } S_2 = \frac{\int_0^{\frac{n^t}{k}} e^{-rt} f t}{\int_0^{n^t} e^{-rt} f t}.$$

This equation is critical to understanding the impact of different product designs on bidding behavior under the real options framework. Indeed, note that both for the profit and the fixed cost part of the difference, the equation shows that the sign of Δb^{opt} depends on the relation between the cumulative distribution function of the longest product ($\phi(z)$ and $\phi(z+v)$) with an average value of the cumulative distribution function of the shortest product. This average is shown directly with $\frac{1}{k}$ or indirectly with the value of S_2 , which takes into account the discounting effect of the periodic fixed cost. Therefore, it is sufficient for the average impact for the profit part to dominate (resp. to be dominated) $\phi(z)$ and the second average effect to be dominated (res. dominate) $\phi(z+v)$ to have an increase (resp. decrease) of the bid when segmenting the capacity product into a shorter product. Again, we do not have a closed-form solution that guarantees a value for the variables to give a clear-cut answer on the sign of this difference. However, we provide in Lemma 1.2 sufficient conditions that allow such a clear-cut answer to existing.

Lemma 1.2. *The sum of expected bids of shorter products is always lower or equal to the individual bid for the longer product when each threshold z_i is lower or equal to the unique threshold z and when each threshold $z_i + v_i$ is above or equal to the unique threshold $z + v$.*

Proof. The proof is straightforward and stems from the definition of ϕ as the cumulative density function of a standard normal distribution □

Those conditions imply that the probabilities (and the average probability) that the sum of inframarginal rent is below the fixed cost (i.e., $\phi(z+v)$ and $\phi(z_i+v_i)$) are consistently higher under the product design with short term products. Naturally, this condition states that it should also decrease the curtailed expected value given in part by the expressions $\phi(z)$ and $\phi(z_i)$ for the shorter period product design.

From the definition of Δb^{opt} , it is easy to deduce its marginal change with respect to a change of the main drivers. The comparative statics on the difference between the two market designs encompass both the analysis provided in the first section on the value of the bid and the average component present in Equation 1.7. When we derive this value with respect to the set of variables (namely, $\pi_0 c^{om} n^d$ and σ), we find that it depends again on some conditions on the fixed costs and the sign of the derivatives. More precisely, it relies on the difference between the marginal effect of the first section and the average marginal effect of the sum of expected bids. Both elements can be analyzed

separately using the results in Lemma 1.1.¹³ To see this, we provide in the following equation the marginal change of Δb^{opt} with respect to the volatility of the inframarginal rent σ

$$\begin{aligned} \frac{\partial \Delta b^{opt}}{\partial \sigma} = & -\pi_0 n^t \left(\frac{\partial z}{\partial \sigma} \varphi(z) - \frac{1}{k} \sum_1^k \frac{\partial z_i}{\partial \sigma} \varphi(z_i) \right) \\ & + C^{om} \left(\frac{\partial z + v}{\partial \sigma} \varphi(z + v) - S \sum_{i=1}^k \frac{\partial z_i + v_i}{\partial \sigma} \varphi(z_i) \right) \end{aligned}$$

Therefore, an increase in the volatility of the inframarginal rent positively increases the difference between the bid for the longer product and the sum of the expected bid for shorter products if and only if the average effect of the latter is above (resp. below) the marginal change of the former for the profit part (resp. periodic fixed cost) of the option value.

6 Case study

We illustrate the model by simulating a power plant participating in the French capacity market. This mechanism has been recently implemented, with a first auction held in 2016 for a transaction phase starting in 2017. The main characteristics regarding the supply-side rely on a 4-year quasi-continuous forward market. Each capacity product covers a year, with an obligation of being available concentrated between January and March.¹⁴ They can be traded up to four years before the delivery year, either through multiple auctions or bilateral tradings. Figure 1.1 shows the clearing price in the French capacity market for the corresponding yearly period of the transaction phase and each auction before the starting date. Excluding the specific 2017 and 2018 period, the capacity price is on average equal to 23 191 €/MW, with a maximum value of 47 400 €/MW and a minimum value of 13 000 €/MW. We use the average price for the transaction phase in 2022 as a comparative basis when simulating the output using the previous results. The average price is equal to 25 314 €/MW.

¹³Each component of the sum for shorter products is independent, which allows adding the marginal change of each element.

¹⁴In particular, this rule raises the question of the difference between what the contract is called (annual) and the obligations effectively implemented in detail.

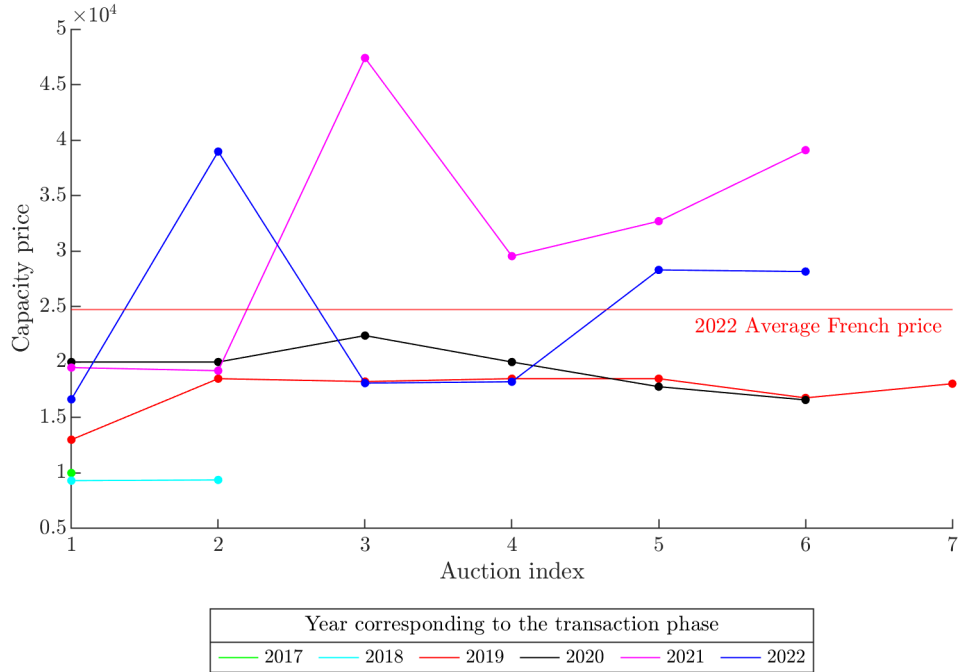


Figure 1.1: Auction results for the French capacity market

We are considering an investment in a CCGT gas power plant, the lifetime of which is 30 years, and we normalize the capacity to 1MW. The production costs have been taken from the consultation report made by RTE, the transport and system operator, which had to build the rules for the capacity market. For an existing power plant, the fixed operating cost is equal to $32.5\text{€}/\text{kMW}\cdot\text{yr}$,¹⁵ which translates into a periodic fixed cost value of $98.63\text{€}/\text{MW}\cdot\text{day}$. The variable production cost includes the fuel cost and the carbon costs and is equal to $25\text{€}/\text{MWh}$. We consider them fixed during the lifetime of the investment.

For the French system, we assume the risk-free yearly rate is 2.32 %. It is the average interest rate of the 30-year government bonds for France between the years 2009 - 2021. It implies a daily value of 0.64%. Then, we estimate the stochastic process. First, we analyze the forward Y1 traded on the French power exchange between the years 2010 - 2015. We find that the average daily electricity price over the period is equal to $47.15\text{€}/\text{MWh}$, with a maximum value of $61.65\text{€}/\text{MWh}$ and a minimum value of $33.50\text{€}/\text{MWh}$. A gas power plant can be considered peak technology or semi-peak technology in the French system, and it does not receive an inframarginal rent every hour during its

¹⁵It is the mean value for a range between 30 and 35 $\text{€}/\text{kMW}\cdot\text{yr}$.

lifetime. Therefore, we first compute an average price duration curve, which gives the proportion of time for which the price exceeds a specific value. Then, we use the data on the marginality duration of a gas power plant given by the yearly report *Functioning of the wholesale electricity* from the CRE, the French electricity regulator. We find that, on average, such investment is either the marginal or an inframarginal bidder for 57% of the time in a year, with a high deviation between years ranging from 5% to 85% over the years 2010 - 2019. Given this significant range, and using the marginal production cost and the price duration curve, we find an interval of daily inframarginal rents from 59 €/day and 949 €/day with a mean value of 641 €/day for an investment selling all the time.¹⁶ We acknowledge this value is highly uncertain and dependent on the actual plant. Therefore, we use those values as a comparative order of magnitude rather than real input.¹⁷ For the volatility of the inframarginal rent, we use the forward Y1 traded on the French power exchange between the years 2010 - 2015. We find a volatility value of 0.00578, close to the range used in [Fontini et al. \(2021\)](#) for the Italian market.¹⁸

6.1 Bids in a capacity market

We first analyze how the bid in a capacity market can vary with respect to the length of the transaction phase. We provide the results both under the net present value framework and the real options framework. Using the French capacity market as the reference design, we use an initial value for the transaction phase (n^t) of one year with a waiting phase (n^d) of four years. The results of the simulation are presented in [Figure 1.2](#).

As shown in [Proposition 1.5](#), a capacity bid always increases with the length of the transaction phase under the real options framework. On the other hand, the net present value bid can decrease for relatively high values of the inframarginal rent. For the initial values of n^d and n^t , beyond an initial value of 87.83 €/MWh, the bid is constantly decreasing with respect to n^t . The figure also shows that for a given initial value of the inframarginal rent π_0 , the bid under the real options framework is always above the bid under the net present value framework, as shown in [Proposition 1.6](#). We also provide

¹⁶Therefore, we also consider the variable π_0 as the average inframarginal rent being available all the time on the wholesale market.

¹⁷A study made by RTE has found that between 2010 - 2018, the annual inframarginal rent for a CCGT ranges daily from 28 €/MW to 188 €/MW. The numerical simulation gives the same order of magnitude with respect to the initial value π_0 .

¹⁸Note that the value of the volatility is consistent with the threshold found in [Levy \(1992\)](#) regarding the approximation condition of [Assumption 1.1](#).

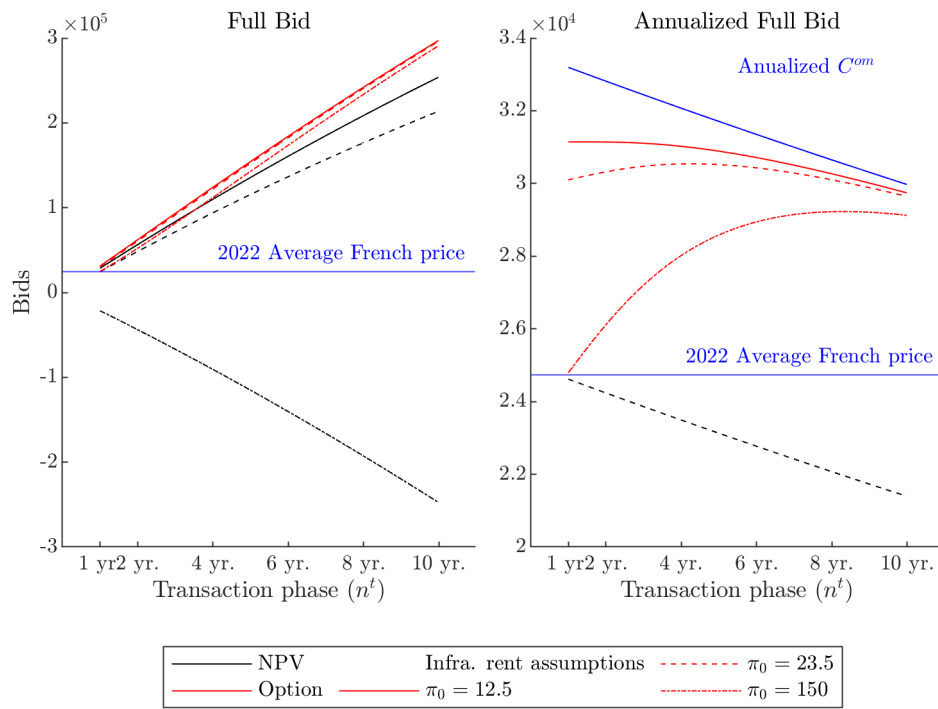


Figure 1.2: Single capacity bid for different transaction phases value. The bids under the real options framework are always increasing in n^t , while for some values of π_0 , they can be decreasing under the net present value framework. Note that the bids are expressed in $\text{€}/\text{MW}$ and π_0 in $\text{€}/\text{day}$.

the annualized value of the capacity bid for more clarity. In the second figure, we show the corresponding hypothesis for the two frameworks that leads to the same bid as the average one observed in the French Capacity market for the transaction phase of the year 2022. Under the net present value framework, the initial value for the inframarginal rent is equal to 12.5 €/MWh, which is significantly below the lower range of the value found using the forward data. It corresponds to being marginal only 1% during a year on the electricity market. Under the real options framework, the initial value is 130 €/MW, almost ten times higher than the previous value. This high value is within the range of the marginality and corresponds to being marginal at least 11% of the time. Note that the capacity price under the real options framework converges towards the periodic fixed cost, as expected when analyzing the derivative of the capacity bid. Therefore, the real options framework always converges toward the canonical behavior of a capacity market. That is, the bids should be equal to the fixed costs. Finally, the data used in the numerical illustration shows that the real options bids tend to increase more rapidly than the net present value bids, which implies that the flexibility rises with respect to the length of the transaction phase.

6.2 The effect of the waiting time and the volatility

We analyze in this section how the bid on a capacity market can be modified by choosing a policy instrument, namely the waiting time between the capacity auction and the start of the transaction phase, and by the volatility of the inframarginal rent, a key variable to understand industrial decisions in the power sector. Figure 1.3 provides the result of the numerical simulation:

Regarding the waiting time, we underline the ambiguous effect of this variable on the capacity bid. Beyond a specific value of n^d represented by the black dots in the first figure, an increase in the value of n^d continuously decreases the capacity bids. Below this value, the waiting time always increases the capacity bid. This threshold depends on the assumption concerning the initial value of the inframarginal rent. A higher initial value implies a higher threshold. To say it differently, when the producer forecasts a more profitable investment, it decreases the capacity bid and reduces the potential negative effect of n^d on the bid. The reason for such results is as follows. First, given the initial data, the bid part relative to the inframarginal rent is always positively impacted by n^t . It means that the sign of the derivative of z with respect to n^d is always negative.

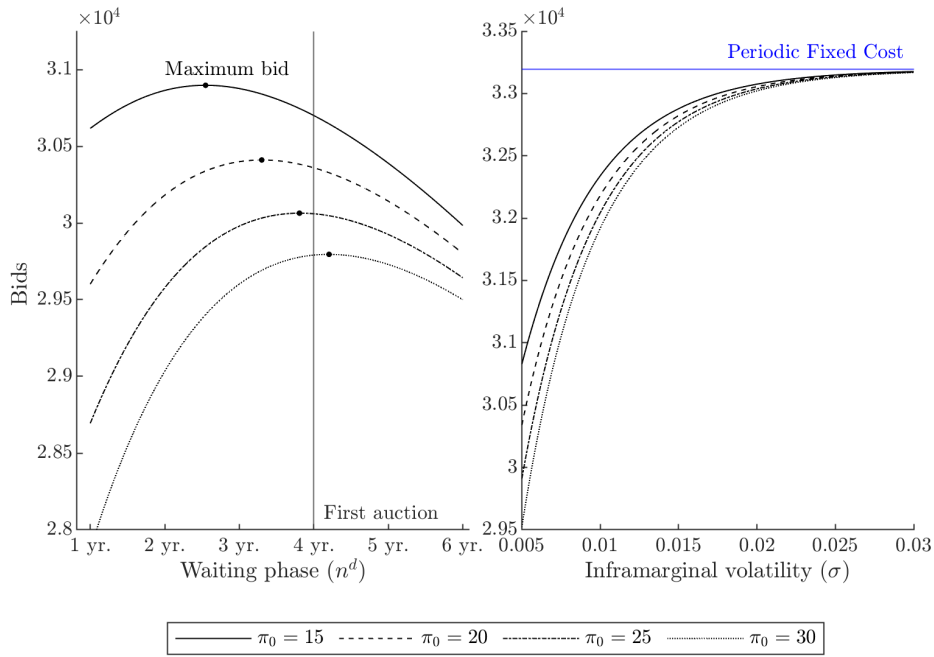


Figure 1.3: Comparative statics on the capacity bid. The waiting phase has an ambiguous effect on the capacity market bids, while an increase in volatility always increases the bids. Note that the bids are expressed in €/MW and π_0 in €/day.

On the other hand, the sign of the second part relative to the fixed costs is mostly negative only when the initial value of the inframarginal rent is low. Recall that this part is composed of a negative value due to the risk-free rate effect and an ambiguous part due to the change in the cumulative distribution function. We find that the first negative part is almost not affected by a change of n^d , while on the contrary, the second part is always positive. Still, it significantly decreases with n^d with a lower value for a higher π_0 . Indeed, the second part is linked to the probability of the fixed cost above the inframarginal rent. Hence, a higher rent value always means a lower probability. All in all, the lower the probability, the higher the negative effect of the risk-free rate and, hence, the potentiality of n^d for having a negative impact on the capacity bid. From a policy perspective, it seems less costly to set up a short waiting phase for a profitable existing power plant and potentially allow a longer waiting time for less profitable investment.

An increase in the volatility of the inframarginal rent always increases the capacity bid in the numerical simulation. As illustrated in the second figure, higher volatility makes the capacity bids converge toward the periodic fixed cost of the investment, even though the initial value regarding the inframarginal rent is different. We also observe

a diminishing marginal effect of the volatility on the capacity bid, meaning that it is sufficient for a slight increase from the current volatility to affect the bid significantly. Those results stem from the fact that the effect of the volatility on the inframarginal part of the bid is relatively small compared to the impact on periodic fixed costs. Indeed, the periodic fixed cost part directly includes the effect of σ on the total volatility of the revenue made during the transaction phase v , which is always positive. However, this effect is significant only for a low value of σ , which explains this diminishing margin effect. It has important implications, as, given this result, we should expect a rapid increase in the bids in capacity markets when the first effect of the introduction of renewable in the system starts to be significant. When renewables have a sufficient share in the production mix, the capacity price is assumed to be relatively stabilized.

6.3 Product design

We conclude the case study by studying the effect of segmenting a given capacity product into successive products with shorter transaction phases. We use an initial long product covering five years as the reference product, and we split this period into shorter periods. We use the same initial value of n^d of four years. We provide in Figure 1.4 the results of the numerical simulation for the net present value and the real options. The point on the two Figures 1.4 represents four different product designs and k the number of successive products, with five yr. a single product covering the five years ($k = 1$), 1 yr. A yearly product ($k = 5$), semester a transaction phase covering six months ($k = 10$), and quarter a transaction phase covering three months ($k = 20$).

This simulation confirms the reverse effect of the choice of the product for the two frameworks: under the real options framework, the bids are, in expectation, lower with shorter products, while under the net present value framework, a more extended transaction phase always implies lower bids. For the latter case, this is explained by the possibility of a more extended transaction phase having positive revenue for a specific period, covering the potential loss incurred during another period in the transaction phase because of the fixed cost.

This smoothing of the opportunity cost of participating in the capacity market is impossible for shorter products as they imply different opportunity costs and decisions. On the other hand, we do not find the same effect with the real options framework. Indeed, under this approach, the segmentation directly impacts the expectation of the option

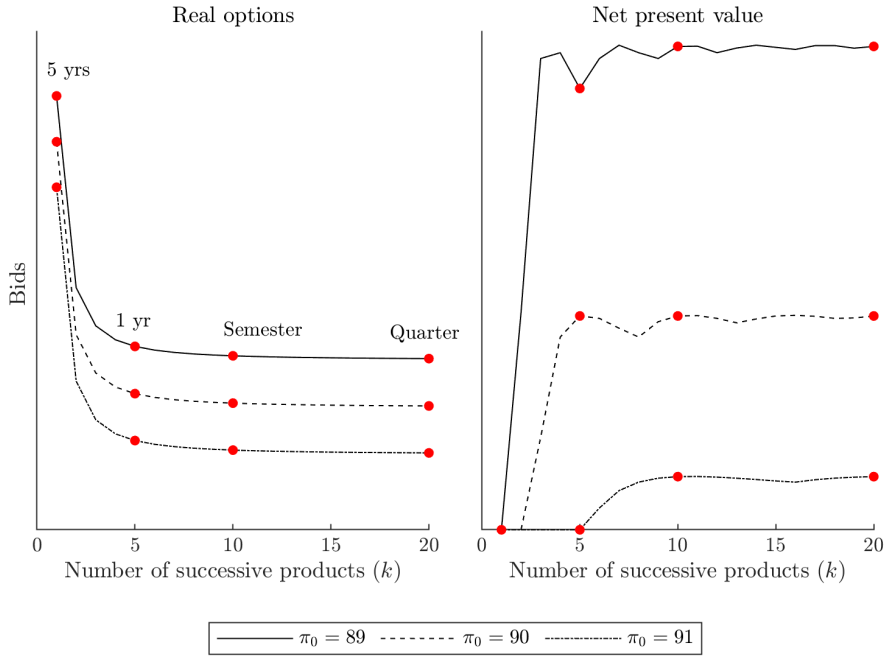


Figure 1.4: Evolution of the expected bids for a single long product and the sum of the expected bids for shorter periods. Note that the bids are expressed in €/MW and π_0 in €/day.

value for shorter periods, which is not the case under the net present value framework. To say it differently, the distribution characteristics of the total inframarginal rents over a long period are different from the sum of the distributions of the total inframarginal rents for successive shorter products.

Given the numerical data, we find that segmenting the bids into shorter periods negatively impacts the inframarginal part and positively the fixed cost part of the bid. Given that the former is always negative and the latter is always positive, this segmentation continuously decreases the cost of a capacity market under the real options framework. Finally, the results also show that the product design choice is different from a marginal perspective by exhibiting a diminishing marginal effect for the two frameworks. We find that segmenting from five years to a single-year period is sufficient to significantly decrease the bid under the real options framework or increase the bid under the net present value framework. Therefore, we show that it is unnecessary to make the capacity market over-complex under the real options framework by having many short products.

7 Extensions

In this section, we provide a discussion on two relevant policy issues associated with the research questions of this chapter. We first look at the implications of the analysis for bids in the capacity market for new power plants. Then, we discuss the effect of the costs associated with the closing decision, which can only be assessed using the real options framework.

7.1 Bids for new entrants

The relation between product design in this chapter and bids for new entrants relies on the definition of the opportunity cost of entering the market as a new entrant. Indeed, as shown in Equation 1.5, the bids should be equal to the NVP of the investment over the whole lifetime, including the two fixed costs (investment and operation) and the two sources of revenues (wholesale and capacity market). Therefore, assuming that the costs and the wholesale revenue are not impacted by the capacity market product design, any changes in the value of the bids when the investment is already in the market will impact the first bid, even though we do not model the competition in the capacity market. It has clear policy implications. Indeed, if policymakers wish for more new entrants, they should aim to increase the probability of those investments being retained when they first bid into the capacity market. Therefore, lower entry bids make this more likely to happen.

Following the previous analysis, we state in Lemma 1.3 this link between product design and bids for the new entry :

Lemma 1.3. *Under the net present value framework, shorter products imply that the initial bid B_0 for new investments is always lower or equal to the bid with a longer product. Under the real options framework, shorter products imply that the initial bid B_0 for new investments is always higher or equal to the bid with a longer product.*

Proof. The proof is straightforward and is given by the results of Proposition 1.1 and when the conditions of Lemma 1.2 hold. \square

To illustrate this discussion, we simulate the bid for a new entrant for a CCGT power plant using the previous data. We provide in Figure 1.5 the relation between the initial

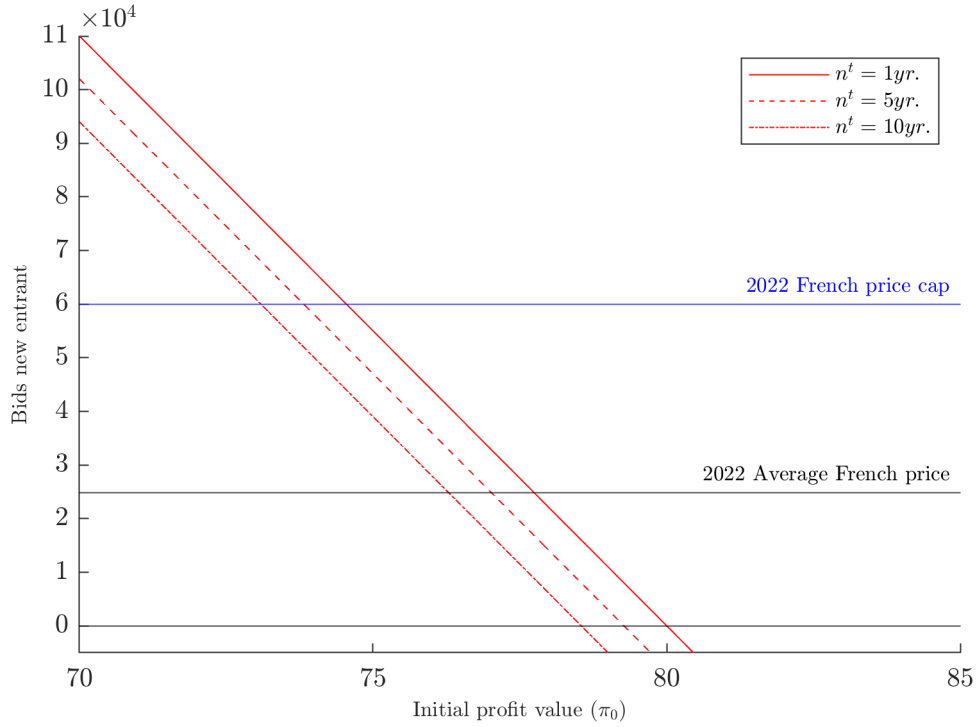


Figure 1.5: Capacity bid for a new entrant with respect to the initial inframarginal rent. Note that the bids are expressed in €/MW and π_0 in €/day.

value of the inframarginal rent and the capacity bid for a new entry and different capacity product designs. As expected, the figure shows that for shorter products, the bids in the capacity market is higher.¹⁹ This figure also illustrates the sensitivity of a new power plant to enter the market when competing with existing investments. Indeed, we find that to provide a price even below the current price cap on the French capacity market. The producer needs to assume an initial value for π_0 of around 74 €/MWh for the three capacity products. On the other hand, as soon as the assumed inframarginal rent is above a value of 79 €/day, a new entrant always makes a null bid.

7.2 Penalty and mothballing costs

Finally, we discuss the effects of two drivers that increase the cost associated with the availability to close to avoid the periodic fixed cost. They both recover two distinct issues regarding the operation of an investment. However, they are identical in their

¹⁹As a higher inframarginal rent implies a lower bid, if for the same bid, the threshold value π_0 is higher, it means that the capacity bids are higher for the same threshold.

conceptualization concerning the capacity bid analysis : (i) a policy instrument being the penalty associated with the failure to respect the obligation of being available by voluntary closing the investment, (ii) the closing costs associated with the temporary shutdown of the power plant. They also have in common that a net present value framework cannot be considered in the capacity bid analysis. We summarize in Lemma 1.4 their effects on the capacity bid under the real options framework.

Lemma 1.4. *Setting a penalty for the failure of not being available when a capacity product has been sold, or the existence of closing costs always leads to a lower bid in a capacity market. An increase in their value decreases the capacity bid.*

Proof. See Appendix 5. □

The intuition behind those results is that the penalty or the closing cost decreases the value associated with closing to avoid the fixed cost. Therefore, they decrease the option value, hence the bid in the capacity market. Regarding the penalty value, it should be stressed that we consider in this section only the case when the power plant deliberately decides not to stay available after observing a too-low inframarginal rent, for instance. On the other hand, we left for future work when the penalty was applied because the power plant failed to stay available due to technical reasons. In this situation, a penalty increases the opportunity cost associated with participating in a capacity market, which increases the bid.²⁰

We provide in Figure 1.6 a numerical illustration of different values of the penalty on the bidding behavior in the capacity market. We assume that the closing costs are equal to 25% of the periodic fixed cost (Abani, 2019).

As expected, an increase in the penalty value decreases the bid in a capacity market. We show the current value in the French capacity market of 40000 €/MW in addition to the closing in the figure. To achieve the same price observed for the 2022 delivery year, we find that the initial value regarding the inframarginal rent needs to be equal to 40.5 €/day, which is almost half of the initial value to reach the same price without the penalty or the closing cost. The numerical illustration shows the significant sensitivity of the choice of the penalty value when implementing the capacity market.

²⁰Such refinement of the model can be analyzed using the net present value framework and has already been studied on a Reliability Option mechanism by Mastropietro et al. (2016).

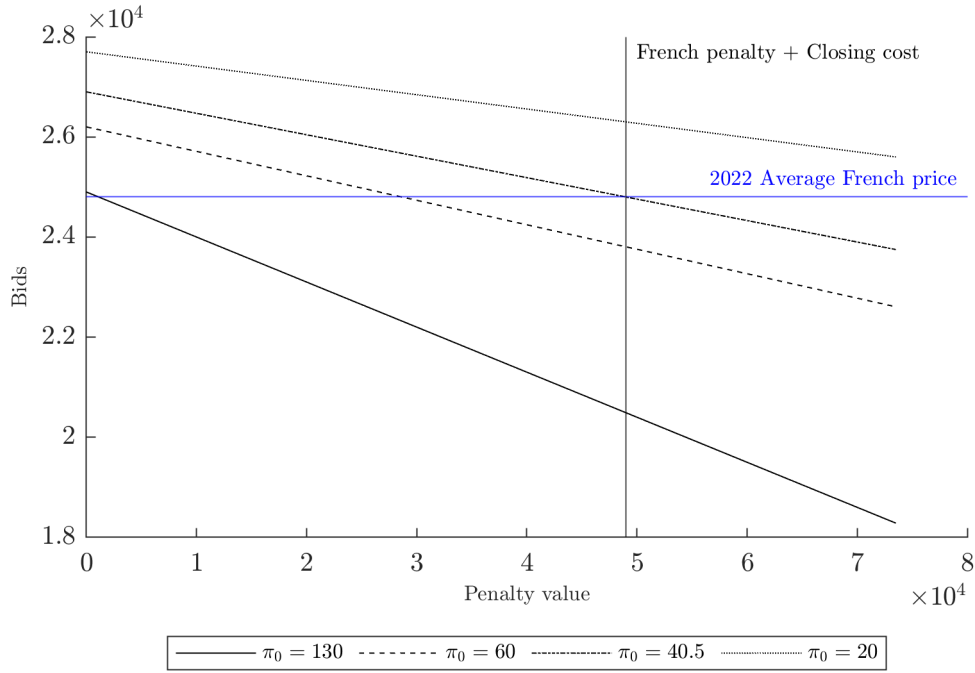


Figure 1.6: Evolution of the capacity price with respect to different penalty values. Note that the bids are expressed in €/MW and π_0 in €/day.

8 Conclusion

In this chapter, we provide a novel approach to analyzing the bidding behavior in capacity markets. We distance ourselves from the model of the net present value, which evaluates the bids on a capacity market as the net expected loss associated with the obligation to have the investment available on the wholesale market. While this framework provides the fundamental rationales to understand the bidding behavior in capacity markets, it does not consider the value associated with the flexibility embedded in the investment. Using a real options framework allows us to consider both the uncertainty regarding the future revenue and costs of the investment and the intrinsic value associated with the alternative of participating in a capacity market, that is, leaving the market to avoid some fixed costs temporarily. We define the bid in a capacity market as the option value associated with this closing option, and we apply a pricing methodology of a Basket Option, an exotic derivative, to evaluate the real options value. Indeed, there is a similarity between this financial derivative, which allows receiving a basket of different assets prices against a strike price, compared to the managerial decision to stay open,

receive revenues from the wholesale market periodically, and sustain an irrevocable fixed cost.

We use this framework to assess two issues related to the implementation and the design of capacity markets. First, we deepen our understanding of capacity markets and how prices are emerging on those competition-based mechanisms. Indeed, the use of a novel framework allows us to assess the determinant of the bids differently, as we have shown, for instance, for some drivers, such as the waiting time between the auction and the beginning of the transaction phase or the volatility of the wholesale revenue. Therefore, any deviation from the actual value of the opportunity cost of participating in a capacity market can be better understood. It is particularly relevant in the current energy policy perspective as capacity markets are usually criticized for their additional burden on consumers.

Then, we analyze the interplay between the product/commitment duration and the opportunity cost for providing capacity availability. We show that the choice of a product design can significantly affect the bids in a capacity market. Indeed, we find that a longer transaction necessarily implies a higher bid than shorter products under a real options framework. On the other hand, it is not always observed under a net present value framework. We also compare opposite market design regimes between having a long product sold in a unique auction or shorter products sold successively in the same period. We find that the outcome depends on various factors, but ultimately, it is more likely that the sum of bids for shorter transaction phases is lower than the individual bid for the long product. The opposite effect is observed for the net present value framework, making the real options all the more relevant.

This chapter provides foundations for future work regarding the analysis of capacity markets. First, it would bring interesting empirical results to integrate this real options approach in a competition model. For instance, [Matthäus et al. \(2021\)](#) has demonstrated that auction theory combined with real options can shed light on market outcomes, especially in the power sector. Second, capacity markets are closely related to the increase in new entries. A significant number of studies have tackled this issue, using real options theory, for instance, to assess the option value to enter the wholesale market ([Fontini et al., 2021](#)). However, to our knowledge, none have combined analysis of the entry decision with the participation in a capacity market with an endogenous bidding behavior under a real options framework. Finally, we have assumed a simplified representative investment with a single source of uncertainty: the inframarginal rent. On the other

hand, it exists in current electricity markets various investments with different operational characteristics. For instance, renewables have high uncertainties regarding their output, while peak technologies face uncertainty in their merit order and production cost. New technologies, such as demand responses and batteries that are pushed to be integrated into capacity markets, also exhibit different uncertainty and operations. A refinement of the model by considering all of these characteristics would make it possible to underline future capacity market implementations better.

Bibliography

- Abani, A., Hary, N., Rious, V., and Saguan, M. (2018). The impact of investors' risk aversion on the performances of capacity remuneration mechanisms. *Energy Policy*.
- Abani, A. O. (2019). *Electricity market design for long-term capacity adequacy in a context of energy transition*. PhD thesis, MINES ParisTech.
- Abani, A. O., Hary, N., Saguan, M., and Rious, V. (2016). Risk aversion and generation adequacy in liberalized electricity markets: Benefits of capacity markets. In *International Conference on the European Energy Market, EEM*.
- Andreis, L., Flora, M., Fontini, F., and Vargiolu, T. (2020). Pricing reliability options under different electricity price regimes. *Energy Economics*, 87:104705.
- Bhagwat, P. C., Iychettira, K. K., Richstein, J. C., Chappin, E. J., and De Vries, L. J. (2017). The effectiveness of capacity markets in the presence of a high portfolio share of renewable energy sources. *Utilities Policy*, 48:76–91.
- Bialek, S. and Unel, B. (2019). Will you be there for me the whole time? On the importance of obligation periods in design of capacity markets. *Electricity Journal*, 32(2):21–26.
- Bialek, S. and Unel, B. A. (2020). Committed But for How Long? On the Optimal Obligation Periods in Capacity Markets - Working Paper. .
- Brown, D. P. (2012). Non-Cooperative Entry Deterrence in a Uniform Price Multi-Unit Capacity Auction. .
- Bushnell, J., Flagg, M., and Mansur, E. (2017). Capacity Markets at a Crossroads. *Energy Institute at Haas*. .

- Cepeda, M. and Finon, D. (2011). Generation capacity adequacy in interdependent electricity markets. *Energy Policy*, 39(6):3128–3143.
- Cramton, P., Ockenfels, A., and Stoft, S. (2013). Capacity Market Fundamentals. *Economics of Energy & Environmental Policy*.
- Creti, A. and Fabra, N. (2007). Supply security and short-run capacity markets for electricity. *Energy Economics*, 29(2):259–276.
- de Maere d’Aertrycke, G., Ehrenmann, A., and Smeers, Y. (2017). Investment with incomplete markets for risk: The need for long-term contracts. *Energy Policy*, 105(June 2016):571–583.
- Fabra, N. (2018). A primer on capacity mechanisms. *Energy Economics*, 75:323–335.
- Fleten, S. E., Maribu, K. M., and Wangensteen, I. (2007). Optimal investment strategies in decentralized renewable power generation under uncertainty. *Energy*, 32(5):803–815.
- Fontini, F., Vargiolu, T., and Zormpas, D. (2021). Investing in electricity production under a reliability options scheme. *Journal of Economic Dynamics and Control*, 126:104004.
- Fuss, S., Szolgayová, J., Khabarov, N., and Obersteiner, M. (2012). Renewables and climate change mitigation: Irreversible energy investment under uncertainty and portfolio effects. *Energy Policy*, 40(1):59–68.
- Hach, D. and Spinler, S. (2016). Capacity payment impact on gas-fired generation investments under rising renewable feed-in - A real options analysis. *Energy Economics*, 53:270–280.
- Holmberg, P. and Ritz, R. A. (2020). Optimal capacity mechanisms for competitive electricity markets. *The Energy Journal*, 41(Special Issue).
- IEA, P. (2020). Power systems in transition. Technical report, IEA.
- Joskow, P. L. (2008). Capacity payments in imperfect electricity markets: Need and design. *Utilities Policy*, 16(3):159–170.
- Ju, N. (2002). Pricing asian and basket options via taylor expansion. *Journal of Computational Finance*, 5(3):79–103.

- Keppler, J. H. (2017). Rationales for capacity remuneration mechanisms: Security of supply externalities and asymmetric investment incentives. *Energy Policy*, 105:562–570.
- Kettunen, J., Bunn, D. W., and Myth, W. (2011). Investment propensities under carbon policy uncertainty. *Energy Journal*, 32(1):77–117.
- Leautier, T. O. (2016). The visible hand: Ensuring optimal investment in electric power generation. *Energy Journal*, 37(2):89–109.
- Levy, E. (1992). Pricing european average rate currency options. *Journal of International Money and Finance*, 11(5):474–491.
- Mastropietro, P., Herrero, I., Rodilla, P., and Batlle, C. (2016). A model-based analysis on the impact of explicit penalty schemes in capacity mechanisms. *Applied Energy*, 168:406–417.
- Matthäus, D., Schwenen, S., and Wozabal, D. (2021). Renewable auctions: Bidding for real options. *European Journal of Operational Research*, 291(3):1091–1105.
- McRae, S. D. and Wolak, F. A. (2019). Market power and incentive-based capacity payment mechanisms. ., pages 1–44.
- Meyer, R. and Gore, O. (2014). Cross-border effects of capacity mechanisms: Do uncoordinated market design policies countervail the goals of European market integration? *Energy Econ.*, pages 9–20.
- Teirilä, J. and Ritz, R. A. (2018). Strategic behavior in a capacity market ? .
- Trigeorgis, L. (1993). The nature of option interactions and the valuation of investments with multiple real options. *Journal of Financial and quantitative Analysis*, 28(1):1–20.
- Wilson, J. F. (2010). Forward Capacity Market CONEfusion. *Electricity Journal*, 23(9):25–40.

Chapter 2

Designing the Demand Side of Capacity Markets

Contents of Chapter 2

| | | |
|-----|---|-----|
| 1 | Introduction | 93 |
| 2 | Benchmark model with capacity market | 98 |
| 2.1 | Environment | 98 |
| 2.2 | Market equilibrium with a capacity market | 100 |
| 2.3 | First-best without a capacity market | 105 |
| 2.4 | Second-best without a capacity market | 106 |
| 3 | Allocation under a price cap | 107 |
| 3.1 | Exogenous allocation | 107 |
| 3.2 | Endogenous allocation | 108 |
| 4 | Allocation with inefficient rationing | 115 |
| 4.1 | Second-best investment and welfare | 115 |
| 4.2 | Exogenous vs. Endogenous allocation | 121 |
| 5 | Extensions - Retailers allocation | 126 |
| 5.1 | Retailers Market Share allocation | 127 |
| 5.2 | Decentralized capacity market | 133 |
| 6 | Conclusion | 139 |

Abstract

This chapter studies the provision of electricity as a homogeneous good with time-varying uncertain stochastic demand and capacity-constrained producers. Due to market failures and public interventions, private agents typically under-procure investments. This prompts some policymakers to implement long-term capacity markets where producers can sell their investment availability to restore efficient investment levels. This chapter studies the effect of different demand-side market designs on equilibrium and welfare. To do so, we provide a novel sequential analytical framework of the capacity market followed by short-term markets (wholesale and retail) where a regulator chooses the level of investment to maximize expected welfare. Depending on a set of assumptions with respect to the market, we show that the regulator is constrained in the welfare he can reach due to the distinct indirect effect a specific capacity market design can generate. In other words, we discuss under which conditions a market design can provide more welfare than others. First, we develop the model regarding the implementation of a single buyer on the capacity market, which needs to choose the cost allocation regime for the demand-side. Then, we extend the model to study how the realized demand is accounted for in the market design.

1 Introduction

For some goods with demand varying over time, such as electricity, wholesale markets' private incentives are not sufficient to ensure that producers make enough investments to meet peak demand in advance of the time when the peak demand materializes.¹ In such industries, due to the critical importance of these goods, policymakers tend to intervene and implement price caps or other types of regulation that distort the price signal and undermine investment incentives. Moreover, the availability of the production capacity for these goods can be characterized as public goods during scarcity periods, for instance, during a cold wave with peak electricity demand or a pandemic with peak demand for medicine or medical equipment. In such circumstances, the absence of adequacy between the capacity and the peak demand, combined with the difficulty of implementing efficient rationing, leads to high costs for society.²

One solution to restore the optimal level of investment lies in implementing a mandatory capacity market in which producers commit to having capacity available to meet the expected peak demand collectively. Current implementations of such mechanisms have been the prerogative of the electricity sector under the name capacity remuneration mechanism.³ However, the COVID pandemic prompted interest in setting up such mechanisms for vaccine supply. For instance, in [Ahuja et al. \(2021\)](#), they study the implementation of a capacity price for the procurement of vaccines and state that *"to accelerate the vaccine delivery timetable, buyers should directly fund manufacturing capacity"*. [Ockenfels \(2021\)](#) proposes a hybrid mechanism that combines a capacity remuneration mechanism with guaranteed prices. On the supply-side of those mechanisms, each participating producer makes a price-quantity offer for a capacity. If a producer sells capacity in this capacity market, he receives a capacity price and commits to being available to produce over future periods.

While the supply emerges naturally in those markets, the capacity demand requires a regulatory intervention. Indeed, the public-good nature of investment during high-

¹Our framework fits into the more general analysis of industries in which a form of competition follows long-run investments, such as electricity markets ([De Frutos and Fabra, 2011](#)), communication network ([Acemoglu et al., 2009](#)), or radio spectrum ([Yan, 2020](#)).

²The COVID-19 crisis offers a recent example of systemic cost induced by the lack of productive capacity. The subject is well known in the electricity sector while remaining a current matter, as illustrated by blackouts experienced in China and Texas ([IEA, 2020](#)). Congestion in transport infrastructure can also be directly linked to the discussion in this chapter([de Palma et al., 2017](#)).

³See for instance [Doorman et al. \(2016\)](#) for a technical description of potential implementations.

demand periods implies that consumers are unwilling to buy capacities in capacity markets.⁴ Hence, the regulator must define the demand function administratively so the market clears and provides producers' capacity prices. This chapter establishes a framework describing the impacts of different demand-side designs for capacity markets and their policy implications. We focus on two interrelated questions that relate to (i) the cost allocation regime, that is, how a single buyer⁵ allocates the capacity price between capacity buyers and final consumers, and (ii) the degree to which the final consumers' realized demand is accounted for in the market allocation design. In this chapter, we describe the channel through which each possible market design impacts the equilibrium. We show that specific market design can affect the demand-side of short-term markets, which also has some feedback effects, positive or negative, on capacity market equilibrium and constrains the regulator in choosing the efficient investment level. In other words, the welfare-maximizing investment level depends on the expected market equilibrium but also on the indirect effect a market design can have on the equilibrium.

The direct effect of an additional stream of remuneration on investment decisions is well understood. The current literature covers a significant range of issues : (i) the outcomes in investment decisions with and without capacity markets, (ii) the effect of market power on the capacity price determination, (iii) the relation with risk and business cycle, (iv) the discrimination between different investment technologies.⁶ In this chapter, we take a step back from the supply-side-focused approach and develop a model that sheds light on the complex interactions between the capacity market design and the demand-side. To our knowledge, there has been no formal analysis of different capacity markets' demand-side designs, the incentive properties of these alternative approaches, and their ability to restore the socially optimal level of investment beyond the direct effect of the increase of the marginal investment value due to the additional capacity price. On the other hand, the importance of the demand function design in the capacity market is well known.⁷ However, those papers still only consider the effect of the capacity market

⁴Transaction cost and asymmetric information prevent adequate transactions up to the optimal level; see, for instance, [Keppler et al. \(2021\)](#) for a discussion in electricity markets. The insurance of having enough capacity has a private value (how much each consumer is willing to pay to avoid inadequacy) and a social value, as an increase in investment reduces the probability of systemic costs ([Fabra, 2018](#)). Furthermore, knowing the willingness to pay for this insurance is sometimes technologically, socially, and economically impossible.

⁵In this chapter, we do not make any assumptions about the identity of the regulator as it is outside its scope. We assume that its objective is to maximize the expected welfare, which comprises all agent individual surplus. In practice, it is usually a regulated firm.

⁶See [Bublitz et al. \(2019\)](#) for a detailed literature review on the theory and implementation issues of capacity markets in electricity markets.

⁷See, for instance, [Hobbs et al. \(2007\)](#) and [Bushnell et al. \(2017\)](#) [Fabra \(2018\)](#) [Brown \(2018\)](#).

directly on the supply-side. In contrast, the chapter underlines the indirect effect of this instrument on retailers and consumers, which in turn impacts producers. [Scoufflaire \(2019\)](#) is the first paper to represent retailers' strategies in the capacity market. The author develops a theoretical model to analyze the preferences regarding information precision for uncertain future demand. The author models heterogeneous price taker producers and homogeneous buyers competing for *à la Cournot* under uncertainty on their level of capacity obligation. Finally, the chapter proposes a conceptualization close to that developed in [Boomhower and Davis \(2020\)](#). The authors analyze on the basis of simulations how the temporal allocation of the capacity price to consumers can increase the economic value of investments in energy efficiency. The central idea is that allocating the cost of capacity markets to peak hours would enable these investments to provide more savings compared to a uniform allocation over every hour. This chapter does not propose such a detailed analysis of the effects of allocating the costs of capacity markets. However, it formalizes these effects on the aggregate welfare of a market with investment decisions.

The central mechanism of this chapter is the relation between the outcome in terms of investment level and expected welfare at the equilibrium and the choice of a particular demand-side market design. To do so, we extend the canonical benchmark model for a homogeneous good characterized by time-varying demand, which describes the relationship between short-term production and long-term investment decisions.⁸ Producers make long-run investments in a single technology in the upstream market to produce a homogeneous good subsequently, given an uncertain future demand. Then, the downstream retailers aggregate and resell the goods at no cost to the final consumers. Our model extends the literature by providing a novel analytical framework that includes a capacity market equilibrium in addition to investment and short-term decisions. In terms of the sequence of decisions, the regulator chooses the demand for capacity in the capacity market to maximize expected social welfare. Given the capacity price and the expected revenue on short-term markets, producers choose the level of investment. Finally, short-term markets are cleared. For an equilibrium to emerge, we derive an endogenous supply function in the capacity market. Namely, following the main theoretical view for capacity markets⁹, we assume that producers offer their marginal opportunity cost of providing additional capacity. This opportunity cost equals the marginal loss of

⁸This model was first developed in a regulated context by [Boiteux \(1949\)](#) for the electricity sector; it was then extrapolated to a market with private producers by [Crew and Kleindorfer \(1976\)](#). This model is widely used to highlight the risk of underinvestment in production capacity.

⁹See for instance [Creti and Fabra \(2007\)](#).

revenue incurred by the investment level beyond the profit-maximizing equilibrium. Our modeling proposition is central as any indirect effects generated by the capacity market can affect the expected revenue made by the producers and can indirectly be captured during the formation of the supply function in the capacity market.

We start the analysis by introducing a price cap regulation, which can be interpreted as representing the effect of different types of market distortions induced by a range of market failures and regulatory interventions.¹⁰ Such a price cap reduces the expected revenues of producers and undermines investment compared to the level needed to reach the welfare-maximizing level of installed capacity. The first market design regime studied is the canonical capacity market. We build on the previous literature and the design found in Léautier (2016) and Holmberg and Ritz (2020), which relies on the assumption that the capacity market does not have any effect beyond increasing the investment level. In the rest of this chapter, we will refer to this market design as the **exogenous regime**. This canonical regime is similar to having a cost allocation regime based on a lump-sum tax. In this case, even when considering the endogenous supply function in the capacity market and the demand function chosen by the regulator¹¹, the equilibrium of the market design always restores the first-best optimum given the inefficiencies. We then investigate the case in which the capacity price impacts the consumers at the margin. In this case, the regime similarly allocates the capacity price as a unitary tax. However, the main difference is that the marginal effect is endogenously determined at the equilibrium because the capacity equilibrium price causes it. In the rest of this chapter, we will refer to this market design as the **endogenous regime**. We show that the existence of the capacity market indirectly affects the wholesale market by redistributing the different states of the world when the capacity does not bind and bind and by lowering the consumer's surplus. Therefore, we demonstrate that the welfare outcome at the equilibrium under this regime is always lower than under the canonical regime.

¹⁰In their seminal paper, Joskow and Tirole (2007) demonstrate that wholesale markets with a price cap cannot lead to the first-best solution. Following this approach, which also serves as a reference model in the chapter, Zöttl (2011) developed a theoretical result on investments under *Cournot* oligopoly with discrete investment and a price cap. Using the same model, Léautier (2016) showed that market power from producers could also be a significant cause of underinvestment. He also introduces a capacity market in the benchmark model where producers can exercise market power. This paper serves as a reference for the implementation of the capacity market.

¹¹In this chapter and unlike Hobbs et al. (2007), we do not analyze the risk of having regulatory errors.

We then compare the two capacity cost allocation regimes by including inefficient rationing. When the price cap is reached, the investment availability becomes a public good as the demand becomes inelastic. Due to the impossibility of efficiently rationing consumers, they incur a significant welfare loss.¹² This additional assumption regarding inefficient rationing has significant implications for comparing the two market designs at their respective equilibria. Indeed, under this new assumption, we find that the indirect effect created by allocating the capacity price on a unitary basis is now ambiguous for social welfare. Under canonical model specification, we find that this market design constantly brings more social welfare at equilibrium than the initial allocation regime. This is due to the interaction between each market design's effects and the equilibrium investment level.

As a third step, we extend the analysis to implementing a regime where the regulator allocates the cost based on actual retailers' market shares, which will be referred to as the **retailer market share regime**.¹³ It allows us to consider the realized demand and analyzes the effect of retailers' market power in the model. We first show how this design affects at the margin the retailers who play '*à la Cournot*' on the retail market, and then we integrate the new equilibrium into the model with investment decisions and the capacity market. We find that this allocation creates an intermediary outcome between the exogenous and endogenous regimes. We also provide comparative statics on the effect of the retail market structure on the equilibrium outcomes of the model.¹⁴ We find that increasing the degree of competition is not always welfare-enhancing.

Finally, we analyze the case of a capacity market entirely based on the realized demand level, which will be referred to as the **decentralized regime**. To do so, retailers are forced to cover the quantity sold on the retail market by buying directly on the capacity market, given a penalty system. We focus on how retailers can form an aggregated demand function in the capacity market, and we analyze the equilibrium capacity emerging in the capacity market. While the regulator cannot directly set the level of investment, we discuss how the penalty impacts the equilibrium investment and the re-

¹²Using the same initial model [Holmberg and Ritz \(2020\)](#) showed that additional capacity payment is necessary when the environment includes the public-good nature of the investments. Indeed, the inadequacy between capacity and consumption generates negative externalities. Hence, to fully internalize the effect of capacity inadequacy, it is necessary to create an adder on the wholesale price. We also use in this chapter the same representation of the public-good nature of the investment. This effect of a price cap is also closely related to the concept of reliability externality described by [Wolak \(2021\)](#).

¹³This is, for instance, the regime implemented in the capacity market of the UK.

¹⁴We do not consider market power on the supply-side in the chapter, as it is well documented in the literature, see, for instance [Zöttl \(2011\)](#) and [Leautier \(2018\)](#) for its effect on investment decision with a price cap, see [Léautier \(2016\)](#) for its impact with a capacity market.

sulting welfare of a decentralized regime. We find that for high penalty values, the regime can provide more welfare than exogenous or endogenous regimes as it gives incentives to retailers to lower their consumption and avoid inefficient rationing.

We provide in section 2 a reminder of the benchmark model that describes investment decisions in generation capacity. We implement the capacity market and build the theoretical supply function in the same section. The different allocation methods are studied in section 3 and in section 4. Section 5 provides the analysis of the retailers' participation in the capacity market. To conclude, we discuss possible extensions of the model in section 6.

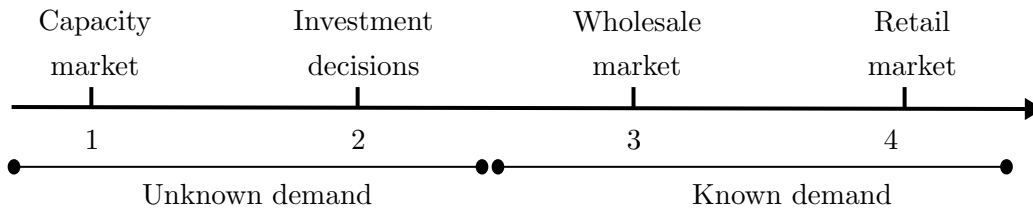
2 Benchmark model with capacity market

2.1 Environment

We consider an initial economic environment with four types of agents: a regulator, producers, retailers, and final consumers. Producers participate in the capacity market and invest in capacities to produce a homogeneous good. They sell the goods to retailers on a wholesale upstream market. Then, retailers resell it to consumers on a downstream retail market. In section 5.2, they participate in the capacity market. Under the sections 3, 4 and 5.1, the regulator chooses the demand function in the capacity market, and in section 5.2, he chooses the penalty level.

Model stages. We consider a four-stage non-cooperative game. First, the capacity market clears. Second, producers choose the level of investment. Third, the wholesale market clears. Fourth, the retail market clears.¹⁵ All decisions during the stages are publicly known and are made simultaneously. We assume the final consumers' demand is uncertain for all agents when making investment decisions. On the other hand, the demand is known when the producers and retailers sell the goods. Those two stages can be interpreted as a repetition of multiple states of the world over a given period (for example, one year), drawn from the distribution (Leautier, 2018). Every agent is to be risk-neutral and maximize expected profit. The game is solved by backward induction.

¹⁵As discussed later, the stage order between capacity market and investment decisions does not matter in this model.



Producers. We assume perfect competition on the supply-side. Producers use a single technology to produce the good. It is characterized by a variable unitary cost of c and a fixed unitary investment cost of r . We normalized the capacity level, so one unit of capacity allows us to produce one unit of the good. The total level of capacity installed after the first stage is k .

Retailers. We allow retailers to be either perfectly competitive or to compete *à la Cournot* to resell the goods to final consumers. However, they do not behave as an oligopsony in the wholesale market or the capacity market. The imperfect competition is modeled using a finite number of retailers n . We model the retail market as perfectly competitive in section 3 and 4 to keep the analysis tractable. In section 5, we introduce the effect of imperfect competition. The use of a finite number is always explicitly indicated. We assume that retailers incur no cost when reselling from the wholesale market to the retail market apart from the wholesale price. Therefore, perfect competition implies that prices are strictly equal in the wholesale and retail markets.

Demand. The following assumptions characterize final consumers in the retail market. They have the same individual uncertain demand with an aggregate demand function $D(p, s)$, s being the state of the world. s can be understood as the demand level affecting only the intercept of the demand function, not its slope. The demand uncertainty is a random variable characterized by a distribution function $f(s)$ and a cumulative distribution function $F(s)$, which is common knowledge. For deriving clear-cut answers, we assume that s follows a uniform distribution with the support $[0, \bar{s}]$.¹⁶ The inverse demand function is linear and equal to $p(q, s) = a_0 + s - bq$, with q the quantity sold on the retail market, such that $D(p(q, s), s) = q$.¹⁷ For convenience, we assume that $p^s(q, s)$ is the price on the wholesale market, and $p(q, s)$ is the price on the retail market. Note that the form of the inverse demand function has the following properties that ensure

¹⁶We obtain similar results with exponential distribution, which represents more realistic consumption variability but prevents us from having a clear-cut answer. See, for instance, [Leautier \(2018\)](#).

¹⁷Similar to the uniform distribution, this assumption is used for both tractability and illustrative purpose.

the initial framework has nice behavior: (i) additive demand shock, which implies that $p_{qs} = p_{sq} = 0$. (ii) states of the world are ordered ($p_s > 0$), (iii) decreasing inverse demand with respect to q ($p_q < 0$), and (iv) decreasing marginal return. To ensure producers invest in capacities, we also need $a_0 > c + r$.¹⁸

Price Cap. there are two main reasons private investors may not provide sufficient capacities: (1) the revenue collected on the market is insufficient to cover their production and investment costs, (2) prices do not consider the positive externalities implied by their availability during high demand periods. For the first rationale, we derived the inefficiency that typically characterized essential goods such as electricity: the suboptimality of the wholesale price modeled via a price cap.¹⁹ Further in this chapter, we present two other rationales: the public-good nature of capacity during peak-demand states of the world and a concentrated retail market represented via retailers' market power. We implement a price cap denoted p^w . To create inefficiencies, the price cap must be binding for some states of the world, so it needs to be below the highest price during the highest demand period; $p^w < a_0 + \bar{s}$. However, to allow for investment, we also need the price cap above the total unitary cost: $p^w > r + c$.

2.2 Market equilibrium with a capacity market

We now describe the equilibrium of the game that consists of a series of equilibria for each stage : (i) the retail market, (ii) the wholesale market, (iii) investment decisions, and (iv) the capacity market. For now, we assume only a direct supply-side effect of the capacity market on the game equilibrium via an increase in the producers' profit. In the rest of the chapter, the analysis of the different market designs follows the same backward induction process.

Fourth stage - Retail market. We assume that symmetric retailers can act strategically 'à la Cournot' in the retail market, and they take the wholesale price as given.

¹⁸For most of the functions $f(x, y)$, $f_x(x, y) = \frac{\partial f}{\partial x}(x, y)$, $f_{xx}(x, y) = \frac{\partial^2 f}{\partial x^2}(x, y)$, $f_{xy}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y)$

¹⁹This modeling approach can represent both an explicit and implicit price cap. In the latter case, political interventions due to the essential nature of the good can artificially alter the price. For instance, when the power system operator needs to carry out technical interventions to avoid system failures. Those policy interventions, such as price caps and non-economic distortions made by a public entity, lead to a *Missing Money* issue that prevents sufficient revenue from being collected to cover costs (Joskow and Tirole, 2007). The effect of price cap regulation was illustrated during the COVID-19 crisis in Italy when the government introduced a 50-cent cap on sales price per mask, which eliminated the incentives to reconvert plants or increase production (Fabra et al., 2022).

The retailer's profit made on the retail market is $\Pi_i^r(s) = q_i(p(q, s) - p^s)$. The first-order condition gives the equality between the marginal revenue and the marginal cost. In case of imperfect competition, the inverse demand function of retailers on the wholesale market is a downward rotation at the intercept of the final consumer demand function $p^s(q, s) = p(q, s) + \frac{q}{n}p_q$. When the retail market is perfectly competitive, we have straightforwardly: $p^s(q, s) = p(q, s)$. For notation clarity, we assume first perfect competition and use $p(q, s)$ as general notation for demand.

Third stage - Wholesale market. Producers know the demand at this stage, so the retailers' inverse demand function is certain. The price is determined by the total investment level of k . We assume perfectly competitive producers, so when k is not binding, the price is equal to the marginal cost c (*off-peak periods*). When k is binding, the price rises above marginal to ensure supply equals demand (*on-peak periods*). We denote s_0 as the first state of the world when capacity is binding, that is, when the price at the capacity level is equal to the marginal cost: $p(k, s_0) = c$. We also define q_0 as the quantity bought by final consumers when the retail price is equal to the marginal cost, such that $p(q_0, s) = c$. During off-peak periods, when $s_0 \geq s$, the price on the wholesale market is the marginal cost c , and the price on the retail market is equal to $p(q_0)$. During peak periods, when $s > s_0$, the demand function determines the price with $p(k, s)$.

$$q^s = \begin{cases} q_0 & \text{if } s \in [0, s_0] \\ k & \text{if } s \in [s_0, \bar{s}] \end{cases} \quad p^s = \begin{cases} c & \text{if } s \in [0, s_0] \\ p(k, s) & \text{if } s \in [s_0, \bar{s}] \end{cases}$$

Second stage - Investment decisions. At this stage, final consumer demand is unknown, and so is the wholesale and retail price. The expected profit of producers is defined as the sum of the expected profit made on the wholesale market and, if implemented, the realized profit on the capacity market minus the investment cost :

$$\Pi^s(k) = \Pi^w(k) + \Pi^c(k) - rk = \int_s q^s(s)(p^s(s) - c)dF(s) + p^c(k)k - rk$$

The market equilibrium in terms of investment decisions with a perfect competitive framework is given by following the first-order condition:

$$\frac{\partial \Pi^w}{\partial k}(k) + p^c(k) - r = 0$$

The crucial term to determine the market equilibrium is the net expected marginal revenue made on the wholesale market. We define $\phi(k)$ for this value for general notation, and it is found by taking the derivative of the wholesale profit: $\phi(k) = \frac{\partial \Pi^w}{\partial k}(k)$. During off-peak periods, producers are perfectly competitive, and prices equal marginal cost; therefore, the marginal revenue is null. When the capacity is binding, it is the difference between the wholesale price and the marginal production cost. The following equation formally defines this rent under the full efficiency assumption (with no price cap):

$$\phi_0(k) = \int_{s_0}^{\bar{s}} \underbrace{(p(k, s) - c)}_{\text{on-peak } k \text{ rent}} dF(s) \quad (2.1)$$

We turn now to the framework with a price cap. We introduce a second threshold s_0^w . It is the first state of the world when the price cap is binding, that is when the price at the capacity level is equal to the price cap: $p(k, s_0^w) = p^w$. We also define q_0^w as the quantity bought by retailers (or consumers under perfect competition) when the price is equal to the price cap, such that $p(q_0^w, s) = p^w$.

$$\phi_0^w(k) = \int_{s_0}^{s_0^w} \underbrace{(p(k, s) - c)}_{\text{on-peak } k \text{ rent}} dF(s) + \int_{s_0^w}^{\bar{s}} \underbrace{(p^w - c)}_{\text{on-peak } p^w \text{ rent}} dF(s) \quad (2.2)$$

The conditions on p^w relatively to the marginal cost c ensure that $s_0^w > s_0$.²⁰

First stage - Capacity Market. We turn to the definition of the equilibrium p^c of the capacity market. For such equilibrium to exist, we impose a market condition following [Léautier \(2016\)](#): There are no short sells, meaning that producers cannot sell more capacity than they own. The existence of the capacity market and the no short-sell assumption leads to the following observations: (i) Decision timing does not matter given the current setting: results still hold if the capacity market is set before or after

²⁰Under the framework of a single producing technology and without market power on the supply-side, the price cap is only binding during on-peak periods. See [Zöttl \(2011\)](#) and [Leautier \(2018\)](#) for a study of price caps with market power and multiple technologies.

the investment decision as long as it is before the final consumers' demand is known. (ii) It is optimal for producers to offer all their capacities if the first condition holds.²¹

The market equilibrium is found via the intersection between the demand and supply functions offered by producers. For now, we remain agnostic on determining the demand function. Except for section 5.2, the demand function is assumed entirely exogenous in the sense that it is determined by the regulator that seeks to maximize welfare and consists of a vertical line.

We build the supply function based on the assumption that producers offer their marginal profit loss associated with the capacity market's participation. The standard approach in the literature represents the cost of investing beyond the optimal capacity level. However, to our knowledge, this is the first time a supply function in a capacity market is directly modeled using the benchmark framework. As we assume perfect competition in the wholesale market compared to Léautier (2016) and Zöttl (2011), capacity choices have no marginal effect on the rent. Indeed, the rent appears only when total capacity is constraining under perfect competition.²² The full profit with a capacity market for a producer is $\Pi^s(k) = \phi(k)k - rk + p^c(k)k$. Under perfect competition, the first-order condition gives $\phi(k) - r + p^c(k) = 0$. Therefore, the capacity market's supply function equals the marginal cost associated with the deviation from the market investment level \bar{k} , which would have been made without the capacity market.

Definition 2.1. *We denote the supply function $X(k)$ and the inverse supply function $X^{-1}(p^c)$ such that $X^{-1}(X(k)) = k$. The supply function on the capacity market is defined as follows:*

$$X(k) = \begin{cases} 0 & \text{if } k \leq \bar{k} \\ r - \phi(k) & k > \bar{k} \end{cases} \quad (2.3)$$

With \bar{k} the market equilibrium given by $r = \phi(\bar{k})$.

²¹The intuitions behind the extension of Léautier (2016)'s proposition in the chapter are as follows: without a direct link between the quantity exchanged in the capacity market and the investment level (i.e., short sell condition); the former does not alter the producer's marginal profit with respect to the latter. Hence, the capacity market does not have any effect on the investment decision. For observation (i), the result is straightforward as we do not include any specification in terms of investment dynamics (e.g., the time to build the investment) and information structure (e.g., the uncertainty of the demand level can reduce when the investment decision is closer to the wholesale market). For observation (ii), the proof relies on the result that the supply and demand function's outcome in the capacity market leads to a unique symmetric equilibrium as the profit function is also concave.

²²Under imperfect competition on the supply-side, the rent also exists due to market power and can appear before the total capacity is binding.

Below \bar{k} , the marginal cost is positive, and the supply is null. Indeed, as the wholesale market's profit function is concave, any marginal revenues on the left side of the equilibrium are above the marginal cost of r . The marginal revenue is below the marginal cost on the right side of the optimal investment level. Therefore, any deviation to the right creates a positive opportunity cost.²³ This approach is particularly relevant as it fully characterizes the effect of different market design regimes in the economic environment. In other words, if a regime changes the expected revenue made in the wholesale and retail market, we can consider its feedback effect on the supply function in the capacity market. For a given demand function, the equilibrium of the capacity market is simply found by equalizing the demand function to the supply function. Figure 2.1 gives an example of how supply functions are built in the capacity market for different assumptions (without or with a price cap and for different values of the price cap). Note the kinks at the bottom for curves; they represent the level of investment that maximizes expected profits. Therefore, on the left, the profit is concave, which implies a null supply function. For sufficiently high values of value k , the capacity level does not bind with positive probability. Hence, no rent is generated on the wholesale market. This explains the convergence towards the marginal investment costs.

Therefore, the equilibrium of the game comprises : (i) a wholesale demand function adjusted for the retailers' market power, if implemented, (ii) a wholesale schedule of prices and quantity for each state of the world, (iii) an investment decision based on the expected wholesale market revenue and from collected capacity market revenue (iv) a capacity market equilibrium price originating from a supply function made by producers and a demand function corresponding to a specific level of investment chosen by the regulator.

In the rest of this section, we describe two benchmark equilibria: the first-best level without any form of inefficiency and the second-best market equilibrium with a binding price cap in the absence of a capacity market.

²³Our approach to the supply function in the capacity market is similar to the theory of supply function equilibria where bidders offer a function such that each point on this function maximizes their profit/utility (Klemperer and Meyer, 1989). In the chapter, the supply function in the capacity market is built such that each producer is indifferent between providing their investment market equilibrium or any investment on the curve in return for the corresponding capacity price.

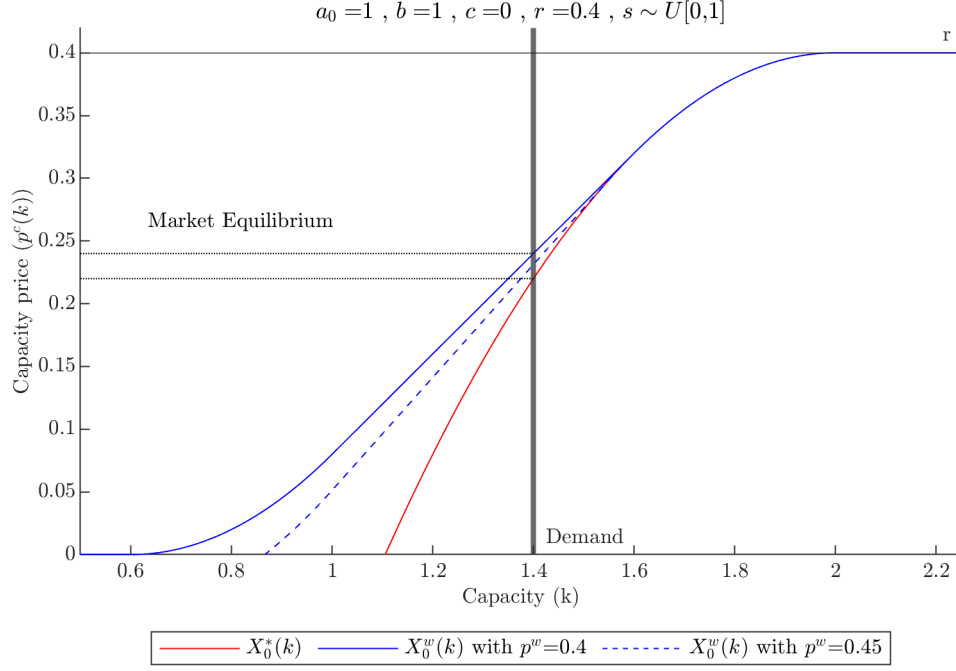


Figure 2.1: Illustration of the capacity market equilibrium given a demand k with the linear continuous model. The market equilibrium is given by the intersection of the supply function and the vertical demand made by the regulator. An increase in the price cap increases the supply function as it lowers expected revenue.

2.3 First-best without a capacity market

We find the optimal first-best investment level as the value of k that maximizes the expected social welfare without any form of inefficiency. For general notation, we define $W(k)$ as the expected social welfare, comprising the consumer, producer, and retailer's surplus. Under the full efficiency assumption, we define k_0^* as $k_0^* = \max_k W_0(k)$, with $W_0(k)$ formally define as follow.

$$W_0(k) = \underbrace{\int_0^{s_0} \int_0^{q_0} (p(q, s) - c) dq dF(s)}_{\text{off-peak welfare}} + \underbrace{\int_{s_0}^{\bar{s}} \int_0^k (p(q, s) - c) dq dF(s)}_{\text{on-peak welfare}} - rk$$

The maximum k_0^* is found by equalizing the marginal surplus gain from an increase of capacity to the marginal cost :

$$\phi_0(k) = \int_{s_0}^{\bar{s}} (p(k, s) - c) dF(s) = r \quad (2.4)$$

Under the initial assumptions, the expected social welfare is concave with respect to the level of investment k , which ensures the existence of a maximum ($\frac{\partial \phi_0}{\partial k} \leq 0$). In the absence of a price cap or any other inefficiency, it is straightforward that the market equilibrium is the first-best solution to maximizing the expected social welfare as the private marginal revenue equals the marginal social revenue.

2.4 Second-best without a capacity market

We now define the market equilibrium with a price cap but without a capacity market. The price cap does not change the social welfare function at the marginal, equal to $W_0(k)$, as it only redistributes surpluses between consumers, producers, and retailers. Without a capacity market, the market equilibrium k_0^w is found by equalizing the expected marginal private profit made on the wholesale market to the marginal investment cost:

$$\phi_0^w(k) = \int_{s_0}^{s_0^w} (p(k, s) - c) dF(s) + \int_{s_0^w}^{\bar{s}} (p^w - c) dF(s) = r \quad (2.5)$$

The Lemma 2.1 shows that a price cap in the wholesale market lowers the market investment level and increases inefficiency. We also provide the optimal payment associated with restoring the optimal investment level. It is equal to the expected difference between what should have been the wholesale price and the price cap when it is binding — this is commonly known as the "Missing Money".

Lemma 2.1. *A binding price cap leads to a lower installed capacity than the optimal investment level given by the social welfare maximization: $k_0^w \leq k_0^*$ as well as lower expected welfare: $W(k_0^*) \leq W(k_0^w)$. The optimal capacity payment $z^w(k)$ is :*

$$z^w(k) = \int_{s_0^w}^{\bar{s}} (p(k, s) - p^w) dF(s) \quad (2.6)$$

Proof. See Appendix 2. □

3 Allocation under a price cap

In this section, we study the effects of the allocation of the capacity price on the equilibrium described previously. From an implementation perspective, we assume that the regulator procures all the capacity on the capacity market. Then, it can choose two general regimes to pass through the procurement cost to the consumers²⁴: (i) lump sum tax, which boils down to assuming an equilibrium kept unchanged by the allocation of the capacity price (**exogenous design**) (ii) a variable tax, which increases the price of the good and generates a specific effect on the equilibrium (**endogenous design**). This section proposes a way of solving the new equilibrium and compares the two outcomes with solely missing money inefficiencies created by the price cap. Section 4 extends this analysis by including inefficient rationing.

3.1 Exogenous allocation

We assume that the regulator forecasts the future expected demand of final consumers, and then it builds the demand function in the capacity market to maximize the expected social welfare. In the analysis, this demand function corresponds to a vertical line equal to the investment level that maximizes the expected social welfare. Indeed, perfect competition on the supply-side implies that producers always offer their marginal cost, and the shape of the demand function does not matter. Finally, to balance its budget, it transfers the full purchasing cost to the retailers using an exogenous ratio or directly to consumers via a lump-sum tax. We formally describe this market design regime as follows:

Assumption 2.1. *The regulator chooses a level of investment to buy on the capacity market that maximizes $W(k)$ at a price $p^c(k)$ given by the supply function described in Equation 2.3. Then, it allocates the full cost $kp^c(k)$ to the retailers or directly to the consumers without any dependence on the expected and realized final demand level.*

This assumption corresponds to the traditional approach used in the literature on the capacity market. We call this market design the exogenous regime because (i) the allocation of capacity costs does not depend, for instance, on retailers' realized strategy but instead on exogenous factors such as their past market share (ii) the design does not

²⁴We do not differentiate retailers or consumers in this section as we assume perfect competition.

depend on realized demand for the final good. In other words, this regime only describes the capacity markets' direct effect via the incentive to invest by the capacity price. There is no effect on the final demand because this remuneration is simply a surplus transfer from consumers to producers. This approach's result is that the capacity price equals the optimal payment, allowing the restoration of an optimal level of capacity when the vertical demand function for capacity is calibrated to k_0^* . Whatever the type of inefficiency is considered. This result is described in the following Proposition. It implies that the cost of a capacity market is strictly equal to the transfer necessary to restore the optimal capacity level.

Proposition 2.1. *Under an exogenous allocation market design, the clearing price in the capacity market at the first-best investment level k_0^* given by the supply function $X_0^w(k)$ is always equal to the optimal payment $z^w(k)$ needed to restore efficiency.*

Proof. See Appendix 3. □

This result highlights the discussion between implementing a price or a quantity instrument to resolve the market inefficiencies or constraints (Weitzman, 1974; Holmberg and Ritz, 2020). We show in this Proposition that the outcome of the capacity market is strictly equivalent to a capacity price set by the regulator defined in equation 2.6. Under this regime, the exogenous approach is optimal because it provides the right investment level, given the inefficiencies.

We illustrate Proposition 2.1 with the model specification. In Figure 2.2, we show two supply functions on a capacity market. When a price cap is introduced, the marginal cost of providing an additional capacity increases, which shifts the supply curve to the left (blue curves), and the market investment level k_0^w decreases compared to the first-best investment level k_0^* . We also include the function associated with the optimal payment as described in equation 2.6 (black curves). As demonstrated in Proposition 2.1, the capacity price at the optimal investment level k_0^* equals the optimal payment.

3.2 Endogenous allocation

We introduce a new allocation regime for the capacity market. In this case, capacity prices marginally impact the final consumer demand via the allocation of the capacity price. The setting is similar to the previous one, with the regulator forecasting the future

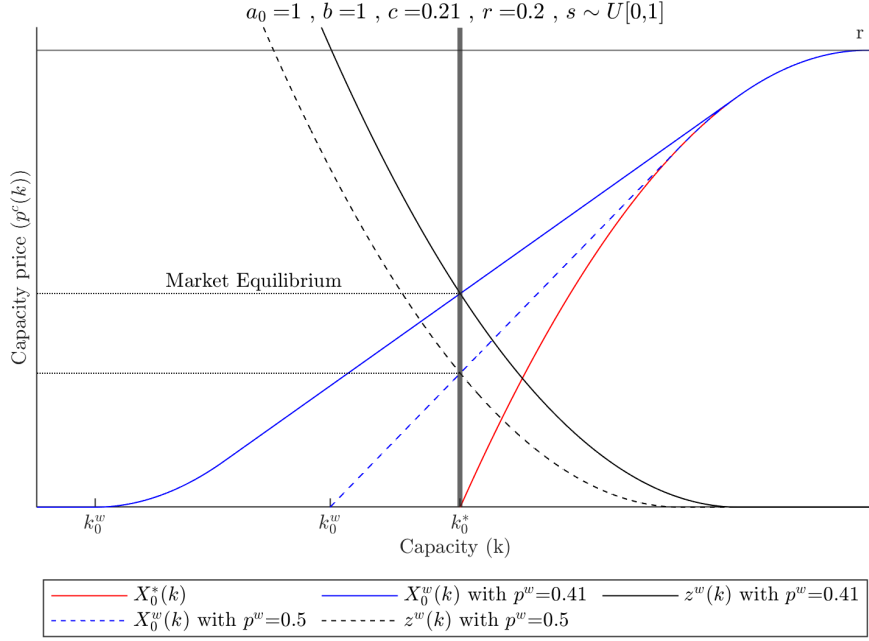


Figure 2.2: Supply function and equilibrium capacity prices on a capacity market under an exogenous regime with linear demand and exponential distribution. The optimal payment corresponds to the market equilibrium.

expected demand and building the demand function in the capacity market. We formally describe this market design regime as follows:

Assumption 2.2. *The regulator chooses a level of investment to buy on the capacity market that maximizes $W(k)$ at a price $p^c(k)$ given by the supply function described in Equation 2.3. Then it allocates the full cost $kp^c(k)$ either to the retailers or directly to the consumers such that the new final demand for the good is equal to $p(q, s) - p^c(k)$.*²⁵

Compared with the previous setting without this indirect effect, the second case in this subsection can be understood as the implementation of a unitary tax.²⁶ However, the main difference with a price instrument such as a tax is that the capacity price and investment decisions emerge from profit maximization from the producers and the regulator's choice of the demand function. Therefore, we formally demonstrate the existence

²⁵This allocation rule is compatible with the budget constraint of the regulator. Another way of expressing this allocation rule would be using the value $p^c(k)\frac{q}{k}$. While this allocation changes the numerical result, it does not impact the fundamental results of this section.

²⁶Note that a possible third regime is an allocation that takes a hybrid form with both a fixed and a variable part. In that case, its effect on the system would be between the two regimes presented in this section.

of the indirect effect by repeating the steps of the previous model and using backward induction.

Fourth stage - Retail market. Let $p^c(k)$ (or p^c for notation clarity) be the capacity price adder for final consumers, identical to a unitary consumption tax as the value is sunk at this period. The final consumer demand function shifts downward with its new value equal to $p(q, s) - p^c(k)$. k is still the quantity bought on the capacity market by the entity at a price $p^c(k)$. We denote $s_1(p^c)$ and $q_1(p^c)$ the new thresholds for respectively the states of the world between on-peak/off-peak periods such that $p(k, s_1) - p^c(k) = c$, and the corresponding quantity such that $p(q_1, s) - p^c(k) = c$. We also define the thresholds for the price cap with $s_1^w(p^c)$ the first state of the world when the price cap is bidding under the endogenous design, that is $p(k, s_1^w) - p^c(k) = p^w$. We also define $q_1^w(p^c)$ as the quantity when the price is equal to the price cap, such that $p(q_1^w, s) - p^c(k) = p^w$. For now, we assume that the capacity price exists. We formally demonstrate it in the proof of Lemma 2.3.

Third stage - Wholesale market While the demand is always lower or equal to the initial demand function, the impact on the expected social welfare is not trivial. The Lemma 2.2 summarizes the main insight and states that the new welfare function is always lower or equal to the exogenous case.

Lemma 2.2. *Allocating the capacity price at the margin only affects the share between on-peak and off-peak periods and the expected surplus size during off-peak periods. Namely, only the occurrence of the two periods s_0 and the intersection between the demand function and the marginal cost q_0 change, the welfare function becomes:*

$$W_1(k, p^c) = \int_0^{s_1(p^c)} \int_0^{q_1(p^c)} (p(q, s) - c) dq dF(s) + \int_{s_1(p^c)}^{\bar{s}} \int_0^k (p(q, s) - c) dq dF(s) - rk$$

Proof. See Appendix 4. □

We can rewrite the equation by showing the initial welfare function without endogeneity: $\Delta W_1(k, p^c) = W_0(k) - W_1(k, p^c)$. With :

$$\Delta W_1(k, p^c) = \int_0^{s_0} \underbrace{\int_{q_1(p^c)}^{q_0} (p(q, s) - c) dq}_{\Delta \text{ in surplus}} dF(s) + \int_{s_0}^{s_1(p^c)} \underbrace{\int_{q_1(p^c)}^k (p(q, s) - c) dq}_{\Delta \text{ in occurrence}} dF(s) > 0$$

The first part of $\Delta W_1(k, p^c)$ represents the loss when it is off-peak periods for both cases (indeed, we have $s_0 \leq s_1(p^c)$ as lower demand always means a higher chance of being off-peak): the consumers fully support the loss as producers receive the marginal cost. The second part represents the loss when the capacity level is such that it is an off-peak period with the endogenous case and an on-peak for the other case. Therefore, the loss is shared between consumers and producers, the former sustaining a higher price and receiving a lower margin. There is no loss when both cases are in peak periods, as the quantity on the market is strictly equal to the capacity installed. Hence, recovering the capacity cost allocation only during peak periods does not generate a deadweight loss.

We continue the endogenous regime analysis by defining the main equilibrium variables. While Proposition 2.2 and Lemma 2.2 underline the effects of this regime on welfare, we show in the following analysis its effect on market equilibrium, namely the outcome in terms of bidding behavior in the capacity market.

Second stage - Investment decisions Producers make their investment decisions based on the expected net revenue, composed of the expected rent and the capacity revenue. The net revenue is similar to the exogenous case, except for the new state of the world thresholds and the wholesale price. It is defined in the following equation.

$$\phi_1^w(k, p^c) + p^c = \int_{s_1(p^c)}^{s_1^w(p^c)} \underbrace{(p(k, s) - p^c - c)}_{\text{net demand}} dF(s) + \int_{s_1^w(p^c)}^{\bar{s}} (p^w - c) dF(s) + p^c \quad (2.7)$$

A sufficient condition for market equilibrium in terms of investment decisions is $\frac{\partial p^c}{\partial k} \geq 0$. In that case, the expected profit is concave, and the second order is satisfied. The derivative of the capacity price depends on underlying assumptions we discuss in the next stage. The initial assumptions with respect to the model assumption ensure that $\frac{\partial p^c}{\partial k} \geq 0$. Note that at the second-best market equilibrium with a price cap k_0^* , the supply function

is identical under the two market design regimes ($\phi_0^w(k_0^w, p^c) = \phi_1^w(k_0^w, p^c)$). Indeed, there is no opportunity cost of being at k_0^w . The capacity price is null, which implies no indirect effect.

First stage - capacity market When a producer participates in the capacity market, it bids its marginal opportunity cost without the capacity revenue (but it takes into account its indirect effect on the demand) equal to $r - \phi_1^w(k, p^c)$. Therefore, following the previous stage, the equilibrium is defined with the equality $X(k) = r - \phi_1^w(k, p^c)$. Lemma 2.3 states how the equilibria are found. It underlines the endogenous nature of this regime, where the choice of capacity changes the bidding behavior in the capacity market compared to the exogenous case.

Lemma 2.3. (i) For any values of $k \in [k_0^w, +\infty)$, there exists a value p^c such that we have $X_1(k, p^c) = p^c$. $X_1(k, p^c)$ is the endogenous supply function in the capacity market given by $X_1(k, p^c(k)) = r - \phi_1^w(k, p^c)$. (ii) Moreover, the supply function is always higher under the endogenous regime than under the exogenous regime: $X_1(k) \geq X_0(k)$

Proof. See Appendix 5. □

Figure 2.3 describes the change in the supply function when considering the indirect effect of the equilibrium capacity price on the consumers. The proof relies on the observation that for a relatively high level of capacity price, the demand is decreased such that at one point, the capacity and the price cap never bind in expectation. In that case, the existence of a solution is always ensured. For other values of the price cap, a solution might exist, but the shape of the supply function prevents the derivation of sufficient conditions.

We have demonstrated the effect of the endogenous market design on the retail and wholesale market and the condition for an investment market equilibrium and a supply function to be well defined. To fully describe the equilibrium, we analyze the level of investment bought on the capacity market by the regulator (and therefore installed by producers). The Proposition 2.2 describes the new optimal investment level that maximizes the expected social welfare given this endogenous regime. It has a strong implication as we state that this regime also modifies the objective for the regulator in

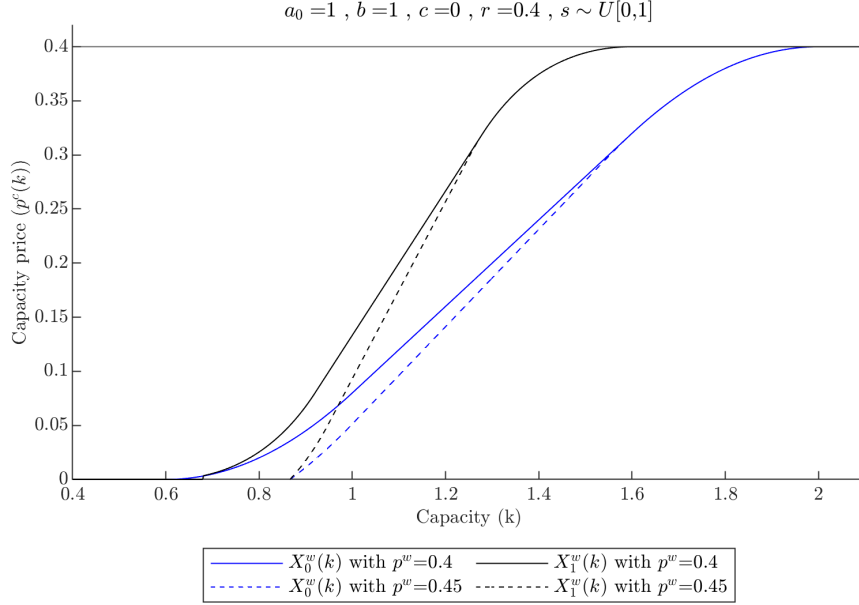


Figure 2.3: Supply functions in the capacity market for the exogenous and endogenous market regime and different values of the price cap. The implementation of the endogenous regime increases the supply function due to the endogenous depressing effect of the capacity price on the demand.

terms of the final investment level.²⁷ Moreover, we find that the endogenous regime is always worse than the exogenous regime regarding social welfare.

Proposition 2.2. (i) *The second-best solution in terms of investment level under the endogenous regime always exists. (ii) It solves $k_1^* = \{k : \phi_1(k) = r\}$, with $\phi_1(k)$ define as follow*

$$\phi_1(k) = \underbrace{\int_0^{s_1} \frac{\partial q_1}{\partial k} p^c(k) dF(s)}_{\text{price effect -}} + \underbrace{\int_{s_1}^{\bar{s}} (p(k, s) - c) dF(s)}_{\text{quantity effect +}}$$

(iii) k_1^* is always lower than the first-best solution under the exogenous level ($k_1^* \leq k_0^*$). The social welfare at the optimal investment level is also always lower than the social welfare at the optimal investment level under the exogenous regime ($W_1(k_1^*) \leq W_0(k_0^*)$).

Proof. See Appendix 6 □

²⁷Therefore, the regulator also needs to take into account the indirect effect while choosing the demand function.

The condition in part (i) of the proposition relates to the concavity of the expected social welfare and the transmission channels of the capacity price in the expected welfare function. The first derivative, represented in $\phi_1(k)$, shows that when the capacity level increases: (a) it decreases the unconstrained quantity q_1 , where at this value, the social surplus is equal to the capacity price ($p(q_1, s) - c = p^c(k)$), this term is, therefore, negative, (b) it generates an additional surplus during on-peak periods which is equal to $p(k, s) - c$, which is positive, (c) to invest in an additional surplus, one needs to sustain the investment cost r . All other marginal effects at the state of the world thresholds cancel each other. This marginal value needs to decrease in k to induce a concave expected social welfare. The proof relies on studying the equilibrium quantity q_1 . We show that it is relative to the shape of the demand function concerning the uncertainty and the distribution function $f(s)$. Namely, the marginal loss sustained during off-peak periods, accounting for the indirect effect of the endogenous regime, is decreasing. A convex capacity market supply function is a sufficient condition which is ensured by having an expectation factor²⁸ to be increasing with k . It has different implications depending on the assumptions regarding $f(s)$ and the effect of s on the inverse demand function. For instance, if we assume an exponential distribution instead of the uniform distribution, it implies that $f(s_1^w) < f(s_1)$, it is less likely for the price cap to bind compared to the capacity k . On the other hand, if $p_{tt} < 0$, that is, the inverse demand function increases less for higher states of the world (at the margin), it implies that s_1^w increases more than s_1 , as it is more likely to bind (at the margin). If the second effect dominates the first, then the condition holds. The policy result of (iii) stems from the analysis of the derivative of $\Delta W_1(k)$ with respect to the level of investment k , which is always positive. The condition for the existence of a first-best investment level is sufficient and implies that $\phi_1(k)$ is decreasing with respect to k (i.e., the expected social welfare $W_1(k)$ is concave).

²⁸ $\Delta F_1(k) = \int_{s_1}^{s_1^w} dF(s)$

4 Allocation with inefficient rationing

4.1 Second-best investment and welfare

We now introduce the public-good nature of capacity during peak demand via inefficient rationing. In this section, we use this rationale to revise the comparison between the endogenous and exogenous designs for capacity markets. In section 5, we will also keep the inefficiency to analyze other options in the design of capacity markets.

When binding at the price cap level, the price-elastic demand becomes inelastic.²⁹ Therefore, we face the same rationing problem as in the literature with limited production capacities and inelastic consumers (see, for instance, [Joskow and Tirole \(2007\)](#)).³⁰ The absence of efficient discrimination between consumers with a heterogeneous willingness to pay implies that investment availability is a public good when the price cap is binding. Therefore, it is underprovided by producers when they make their investment decisions. The literature describes the cost of involuntary rationing in various ways. [Joskow and Tirole \(2007\)](#) shows that it depends on whether the rationing is anticipated or not. [Leautier \(2018\)](#) finds that the effect of involuntary rationing can be different if it impacts the expected demand level. From a modeling perspective, [Holmberg and Ritz \(2020\)](#) uses a general function $J(\cdot)$ to represent this negative externality. The function depends on the delta between the quantity bought at a price equal to the price cap and the investment level. For general notation, we note this cost $M(k)$, defines as follow :

$$M(k) = \int_{s^w}^{\bar{s}} J(s, k) dF(s) \quad (2.8)$$

With Δk a function of the difference between the installed capacity k and the quantity bought by retailers at the price cap q^w (q_0^w or $q_1^w(s)$ depending on the chosen regime). To illustrate the inefficient rationing cost, we use the following assumption.

²⁹The introduction of retailers into the model in the rest of the chapter does not change the intuition. At a price p^w , the *Cournot* competition between the retailers incites them to ask for an equilibrium quantity above the investment value. Then we assume that a regulator ration the retailers, such as their final profit and realized sales, are physically constrained by the investment level.

³⁰This inefficiency is associated with the existence of a *Missing Market* issue under which producers consider their revenue insufficient to invest optimally ([Newbery, 2016](#)). This can be caused by hedging markets being incomplete ([De Maere d'Aertrycke et al., 2017](#)), or because of externalities associated with the public-good nature of investment and consumption choices ([Holmberg and Ritz, 2020](#)), innovation spillovers, and climate change.

Assumption 2.3. *Suppose that consumers sustain an additional cost proportional to the share of consumers selected indifferently who are forced to stop consuming based on their expected surplus. The expected cost is equal to :*

$$M(k) = \int_{s^w}^{\bar{s}} \frac{q^w - k}{q^w} \int_0^{q^w} (p(q, s) - p^w) dq dF(s) \quad (2.9)$$

In this case, the function $J(s, k)$ can be decomposed into two components: the rationing ratio $\frac{q^w - k}{q^w}$ and the consumer welfare at the quantity asked at the price cap (and excluding capacity payments) $\int_0^{q^w} (p(q, s) - p^w) dq$. This example resembles the rationing model used in (Léautier, 2014). It can also be interpreted as follows. Assume that there exists a continuum of consumers such that each point on the inverse demand function $p(q, s)$ represents its marginal willingness to pay for the good. In that case, inefficient rationing implies that each consumer sustains the same cost proportionally to its marginal willingness to pay. This illustration leads to $J(k, s^w) = 0$. Without any difference between the capacity and quantity values, inefficient rationing implies no cost. Regarding the sign of the cost and its derivatives, it seems natural to have $\frac{\partial M(\cdot)}{\partial k} \leq 0$, such that the closer the capacity level is to q^w , the lower is the cost. The second derivative of the expression also matters. For instance, Holmberg and Ritz (2020) assumes a convex function : $\frac{\partial M(\cdot)}{\partial k} \geq 0$. However, as discussed below, the assumption is insufficient to satisfy this condition.

We start with extending the benchmark model with inefficient rationing, which also describes the exogenous regime as it does not generate indirect effects. Then, we compare the outcomes under both regimes. First, we express the expected welfare under the exogenous design with inefficient rationing:

$$\begin{aligned} W_0^{bo}(k) = & \int_0^{s_0} \overbrace{\int_0^{q_0} (p(q, s) - c) dq dF(s)}^{\text{off-peak welfare}} + \int_{s_0}^{s_0^w} \overbrace{\int_0^k (p(q, s) - c) dq dF(s)}^{\text{on-peak } k \text{ welfare}} \\ & + \underbrace{\int_{s_0^w}^{\bar{s}} \int_0^k (p^w - c) dq dF(s)}_{\text{producer surplus}} + \underbrace{\int_{s_0^w}^{\bar{s}} \frac{k}{q^w} \int_0^{q_0^w} (p(q, s) - p^w) dq dF(s)}_{\text{consumer surplus} - M_0(k)} - rk \end{aligned}$$

With $M_0(k)$ being the rationing cost under the exogenous design. The terms of the first line represent the expected social welfare for any state of the world where the price cap is not binding. The first term in the second line is the producers' expected revenue, as

we assume no rationing cost on the supply-side. Finally, the last term on the second line is the expected welfare net of the rationing cost and the producer revenue. Compared to the initial inefficiency of a price cap, the social cost of rationing directly affects the social welfare function. On the other hand, producers' expected marginal rent collected on the wholesale market remains unchanged when we include inefficient rationing, which only affects consumers' welfare.³¹

The absence of direct supply-side effects of inefficient rationing allows us to study the environment's last stage directly. That is, how the regulator chooses the level of investment in the capacity market. We denote k_0^{bo} the optimal level of investment that maximizes the expected social welfare, such that $k_0^{bo} = \{k : \phi_0^{bo}(k) = r\}$, with $\phi_0^{bo}(k)$ as usually defined as the marginal expected social welfare with respect to the investment level k :

$$\phi_0^{bo}(k) = \tilde{\phi}_0 + \underbrace{\int_{s_0^w}^{\bar{s}} \frac{1}{q_0^w} \int_0^{q_0^w} (p(q, s) - p^w) dq dF(s)}_{\Delta \text{ in net consumer surplus}}$$

$\tilde{\phi}_0$ is a term common to $\phi_0(k)$ ³². The terms in $\tilde{\phi}_0$ show that an increase of k (1) allows a marginal gain when the capacity is binding for consumers and producers, and (2) allows an additional rent for the producers when the price cap is binding. The second term in $\phi_0^{bo}(k)$ represents consumer gains when the price cap binds the net of the marginal rationing cost.

The following lemma concludes on the difference between the initial inefficiency caused by a price cap and the consequences of inefficient rationing. The optimal payment to restore the first-best solution equals the marginal value of an additional capacity for the market, which decreases the cost of involuntary rationing.

Lemma 2.4. *When the price cap induces involuntary rationing, the inefficiency is greater than with voluntary rationing. The optimal investment level is greater $k_0^{bo} \geq k_0^*$, and the expected social welfare at the optimum is lower $W(k_0^*) \geq W_0^{bo}(k_0^{bo})$. The optimal capacity payment $z^{bo}(k)$ is : $z^{bo}(k) = -\frac{\partial M(k)}{\partial k}$*

³¹Some authors do include those costs in the producer profit, using a fixed reputational cost (Llobet and Padilla, 2018) or a market shutdown during which producers also lose profit (Fabra, 2018)

³²

$$\tilde{\phi}_0 = \int_{s_0}^{s_0^w} (p(k, s) - c) dF(s) + \int_{s_0^w}^{\bar{s}} (p^w - c) dF(s)$$

Proof. See Appendix 7. □

We now turn to analyzing the model under the endogenous regime given the rationing cost. We first discuss the differences with the exogenous case, and then we describe the implications for the regulator. The expected welfare function under the endogenous regime is equal to the following:

$$W_1^{bo}(k) = \tilde{W}_1 + \underbrace{\int_{s_1^w}^{\bar{s}} \frac{k}{q_1^w} \int_0^{q_1^w} (p(q, s) - p^w) dq dF(s)}_{\text{consumer surplus} - M_1(k)} - rk$$

With \tilde{W}_1 being a term common to $W_1(k)$.³³ We denote k_1^{bo} the optimal level of investment such that $k_1^{bo} = \{k : \phi_1^{bo}(k) = r\}$, with ϕ_1^{bo} defined as the marginal expected social welfare with respect to the investment level k :

$$\phi_1^{bo}(k) = \tilde{\phi}_1 + \int_{s_1^w}^{\bar{s}} \left(\frac{1}{q_1^w} \int_0^{q_1^w} (p(q, s) - p^w) dq - \frac{k}{(q_1^w)^2} \frac{\partial q_1^w}{\partial k} \int_0^{q_1^w} p(q, s) - p(q_1^w, s) dq \right) dF(s)$$

$\tilde{\phi}_1$ is common to $\phi_1(k)$.³⁴ The terms in $\tilde{\phi}_1$ show that an increase of k (1) increases the capacity price, which decreases the expected surplus during off-peak periods, (2) allows a marginal gain for consumers and producers when the capacity is binding, and (3) allows an additional rent for producers when the price cap is binding. The second and third terms of $\phi_1^{bo}(k)$ represent the interaction between the consumer surplus and the rationing cost. To illustrate the effect of the inefficient rationing, we develop below the derivative of $M_1(k)$ with respect to k . We illustrate it with the variation of the rationing costs due to the variation of the investment level k with $k' > k$ in Figure 2.4

33

$$\tilde{W}_1 = \int_0^{s_1} \int_0^{q_1} (p(q, s) - c) dq dF(s) + \int_{s_1}^{s_1^w} \int_0^k (p(q, s) - c) dq dF(s) + \int_{s_1^w}^{\bar{s}} \int_0^k (p^w - c) dq dF(s)$$

34

$$\tilde{\phi}_1 = \int_0^{s_1} \frac{\partial q_1}{\partial k} p^c(k) dF(s) + \int_{s_1}^{s_1^w} (p(k, s) - c) dF(s) + \int_{s_1^w}^{\bar{s}} (p^w - c) dF(s)$$

$$\frac{\partial M_1(k)}{\partial k} = \int_{s_1^w}^{\bar{s}} \left(\underbrace{\frac{-q_1^w - k \frac{\partial q_1^w}{\partial k}}{(q_1^w)^2}}_{\Delta \text{ in rationing ratio}} \int_0^{q_1^w} (p(q, s) - p^w) dq - \underbrace{\frac{q_1^w - k}{q_1^w}}_{\text{price effect}} p^c(k) \right) dF(s)$$

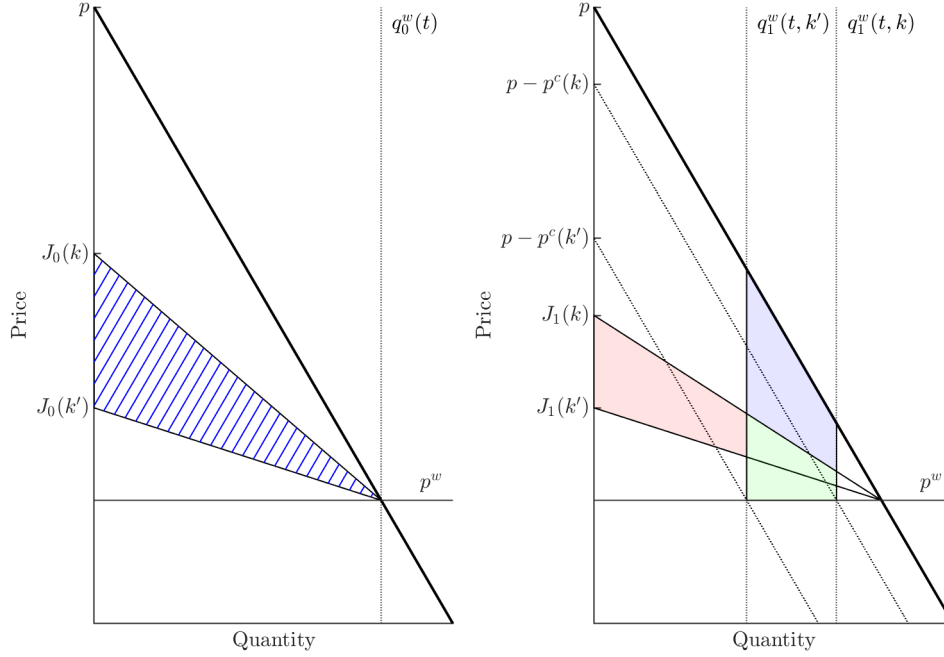


Figure 2.4: Change in surplus without (left panel) and with inefficient rationing (right panel) under the endogenous regime with respect to a change of k with $k' > k$. Under the exogenous regime, an increase of k lowers the rationing ratio. Under the endogenous regime, it also depresses quantity and the size of the consumer surplus.

The first panel shows the delta between the rationing costs for the exogenous case. The initial rationing cost is the integral (up to q_0^w) below $J_0(k)$ while the new cost is the integral below $J_0(k')$. Therefore, the delta is the hatched area. It stems from the fact that increasing the level of investment reduces the rationing ratio. Note the proportionality of the rationing cost via the downward rotation of the inverse demand function $p(q, s)$. The second panel represents the endogenous case. First, note that the quantity demanded by consumers at the price cap (i) is based on the inverse demand adjusted by the price cap $p(q, s) - p^c(k)$ and not on the inverse demand $p(q, s)$, (ii) decreases with the increase of the investment level, due to the increase the capacity price. The red zone represents the first term in $\frac{\partial M_1(k)}{\partial k}$ and corresponds to the change in the proportional ratio $\frac{q_1^w - k}{q_1^w}$.

This is a similar effect represented in the hatched area of the first panel. The green zone represents the change in the rationing cost due to the indirect effect of the capacity price on the demand function. It is the second part of the derivative $\frac{\partial M_1(k)}{\partial k}$. By increasing the price, the capacity market decreases the demand and the rationing cost size. Finally, the sum of the blue and green zones represents the decrease in consumer surplus (without capacity cost) due to the decreases in the quantity demanded at the price cap. The assumption concerning the form of the rationing costs leads to a net total effect that is always positive (red zone - blue zone), as illustrated in the second term of $\phi_1^{bo}(k)$. The following lemma describes a sufficient condition for the existence of an endogenous regime equilibrium. That is, the expected social welfare is concave.³⁵ Define the per unit consumer surplus at s_1^w and for a quantity of k associated with its probability:

$$\text{CS} = \frac{1}{k} \int_0^k (p(q, s_1^w) - p(k, s_1^w)) dq$$

Then,

Lemma 2.5. *If the following condition holds*

$$\frac{\partial}{\partial k} \int_{s_1^w}^{\bar{s}} \text{CS} dF(s) > \int_{s_1^w}^{\bar{s}} \frac{\partial \text{CS}}{\partial k} dF(s)$$

Then a unique maximum of the expected social welfare $W_1^{bo}(k)$ exists.

Proof. See Appendix 8. □

The condition implies that if the variation of the per-unit consumer surplus gains at k when k increases are higher than the expected variation of the per-unit consumer surplus, then a unique maximum exists. In other words, when the level of capacity increases, it decreases the occurrence of states of the world when the price cap binds. It states this marginal decrease of occurrence at the capacity level should be higher than the marginal expected gain of capacity during those states of the world by a factor of two. From an economic perspective, it ensures that when expanding the capacity level, the gains from reducing the inefficient rationing cost do not increase with the capacity level.

³⁵Holmberg and Ritz (2020), while not providing a specific form of $M(k)$, assume a convex rationing cost, which is a sufficient condition for the concavity of expected welfare. In the specification, this is not the case.

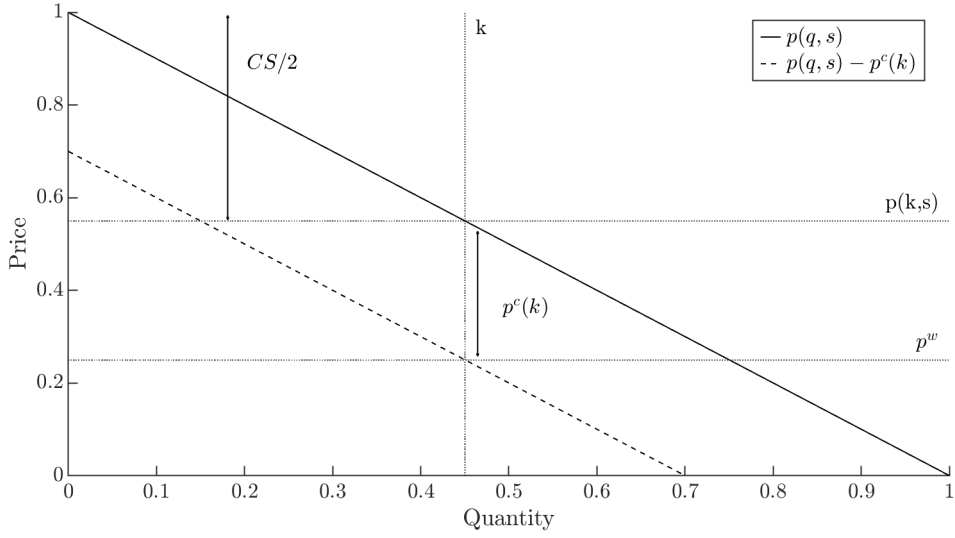


Figure 2.5: Illustration of the expression \mathbb{CS} with a linear demand.

4.2 Exogenous vs. Endogenous allocation

The main difference between the two regimes is that the capacity price indirectly affects q^w , which now depends on the level of investment. This has significant implications for the comparative statics between the two regimes. We start by expressing the difference in terms of welfare $\Delta W_1^{bo}(k) = W_1^{bo}(k) - W_0^{bo}(k)$:

$$\Delta W_1^{bo}(k) = \Delta \text{Welfare} + \underbrace{\int_{s_0^w}^{s_1^w} J_0(s, k) dF(s)}_{\text{Avoided rationing cost}} + \underbrace{\int_{s_1^w}^{\bar{s}} J_0(s, k) - J_1(s, k) dF(s)}_{\Delta \text{ in rationing cost}}$$

$\Delta \text{Welfare}$ represents the endogenous regime's negative effect presented in section 3.³⁶ However, the rationing cost specification is based on the quantity consumed at the price cap. Hence, with rationing, the endogenous regime also indirectly affects periods during which the price cap binds and not only the off-peak periods. The first term stands for

³⁶Formally

$$\begin{aligned} \Delta \text{Welfare} = & - \int_0^{s_0} \int_{q_0}^{q_1} (p(q, s) - c) dq dF(s) - \int_{s_0}^{s_1} \int_{q_0}^k (p(q, s) - c) dq dF(s) \\ & - \int_{s_0^w}^{s_1^w} \int_k^{q_0^w} (p(q, s) - p^w) dq dF(s) - \int_{s_1^w}^{\bar{s}} \int_{q_1^w}^{q_0^w} (p(q, s) - p^w) dq dF(s) \end{aligned}$$

the lower occurrence of periods during which the price cap is binding due to the lower demand. In that case, welfare trades a rationing cost against lower welfare (corresponding to the first term in the second line of $\Delta\text{Welfare}$). The second term represents the change in rationing cost due to a change in the quantity consumed at the price cap. We express this term $\Delta J(s, k) = J_0(s, k) - J_1(s, k)$ below after rearrangement :

$$\Delta J(s, k) = k \left(\frac{1}{q_1^w} - \frac{1}{q_0^w} \right) \int_0^{q_1^w} (p(q, s) - p^w) dq + \frac{k - q_0^w}{q_0^w} \int_{q_1^w}^{q_0^w} (p(q, s) - p^w) dq$$

We have $\Delta J(s, k) > 0$ from the observations that $q_1^w < q_0^w$ due to the negative price effect on the quantity of the endogenous regime. Therefore, for similar states, the rationing cost is always lower under the endogenous regime than in the exogenous case. However, it does not render the comparative statics of the equilibrium straightforward, especially when describing (i) the ranking between the equilibrium investment level and (ii) the ranking between the equilibrium welfare.

We start with the ranking between the different investment levels. Section 3 showed that we always have a lower investment level under endogenous regime $k_1^* \geq k_0^*$ compared to the exogenous regime under a sole price cap inefficiency. With inefficient rationing, Lemma 2.4 showed that the investment level is higher than the first best: $k_0^* \leq k_0^{bo}$. However, due to the opposite effects an endogenous regime exhibits between the negative price effect and the decrease of rationing costs, the ranking between k_1^{bo} and the first-best k_0^* is a priori unclear. Proposition 2.3 provides a ranking between the investment level under the linear assumptions.

Proposition 2.3. *If the following condition holds:*

$$-\frac{\partial}{\partial k} \int_{t_1^w}^{\bar{s}} (\text{CS} + p^c(k)) dF(s) > (\text{CS} + p^c(k)) f(s_1^w) \frac{\partial s_1^w}{\partial k}$$

Then, there is a unique ranking between the investment equilibrium such that: $k_1^ \leq k_0^* \leq k_1^{bo} \leq k_0^{bo}$*

Proof. See Appendix 9. □

Previous analysis shows that introducing rationing costs increases the investment level, and having an endogenous regime (excluding rationing) decreases it. Hence, the ranking between k_0^{bo} and k_1^{bo} is straightforward, as the indirect effect reduces the rationing costs, hence the need for investment. The proof analyzes how the investment level under the endogenous regime k_1^{bo} behaves compared to the first-best level k_0^* . The core of the proof relies on the variation of k_1^{bo} with respect to p^w which is given by the implicit function theorem: $\frac{\partial k_1^{bo}(p^w)}{\partial p^w} = -\frac{\partial^2 W_1^{bo}}{\partial k \partial p^w} / \frac{\partial^2 W_1^{bo}}{\partial k^2}$. On the other hand, the first-best k_0^* does not depend on p^w , which allows focusing only on $\frac{\partial k_1^*(p^w)}{\partial p^w}$ for the ranking. The condition in the proposition has a similar interpretation to the condition for the existence of a maximum. That is, for a cross derivative to be negative, one needs to have the change in expected individual consumer surplus, to which is added the capacity price to be bound from below by a strictly positive value equal to the same value weighted by the derivative of the CDF at the threshold.

We conclude the comparison by studying the ranking between the welfare at the different investment equilibria. Similarly, we have proven that the exogenous regime provides the highest welfare with only a price cap inefficiency and that inefficient rationing decreases welfare: $W_0(k_0^*) \geq W_0^{bo}(k_0^{bo})$ and $W_1(k_1^*) \geq W_1^{bo}(k_1^{bo})$. We are left to study the difference between the exogenous and endogenous regimes at the equilibrium level, that is, $W_0^{bo}(k_0^{bo})$ and $W_1^{bo}(k_1^{bo})$. While not a priori straightforward, we find that there is also a ranking between the expected welfare at the equilibrium, which is described in Claim 2.1.

Claim 2.1. *Under the model specifications, there is a unique ranking between the welfare equilibrium such that: $W_0(k_0^{bo}) \leq W_1(k_1^{bo}) \leq W_1(k_1^*) \leq W_0(k_0^*)$*

Proof. The intuition of the claim relies on the comparison between $W_0^{bo}(k_0^{bo})$ and $W_1^{bo}(k_1^{bo})$ and stems on three observations : (i), from Proposition 2.3, for a given p^w , we have $k_1^{bo} \leq k_0^{bo}$; (ii) the two functions are increasing in p^w , and (iii) ΔW_1^{bo} is decreasing and concave in p^w . \square

Figure 2.6 illustrates the results. The solid black line gives the first-best welfare function, corresponding to the exogenous equilibrium with only a price cap inefficiency. The black diamond represents the first-best investment level k_0^* . Involuntary rationing is added to represent the public-good nature of the investments. The blue curves represent the new expected social welfare with inefficient rationing under an exogenous regime. The red dashed curve encompasses the effects of the endogenous regime with

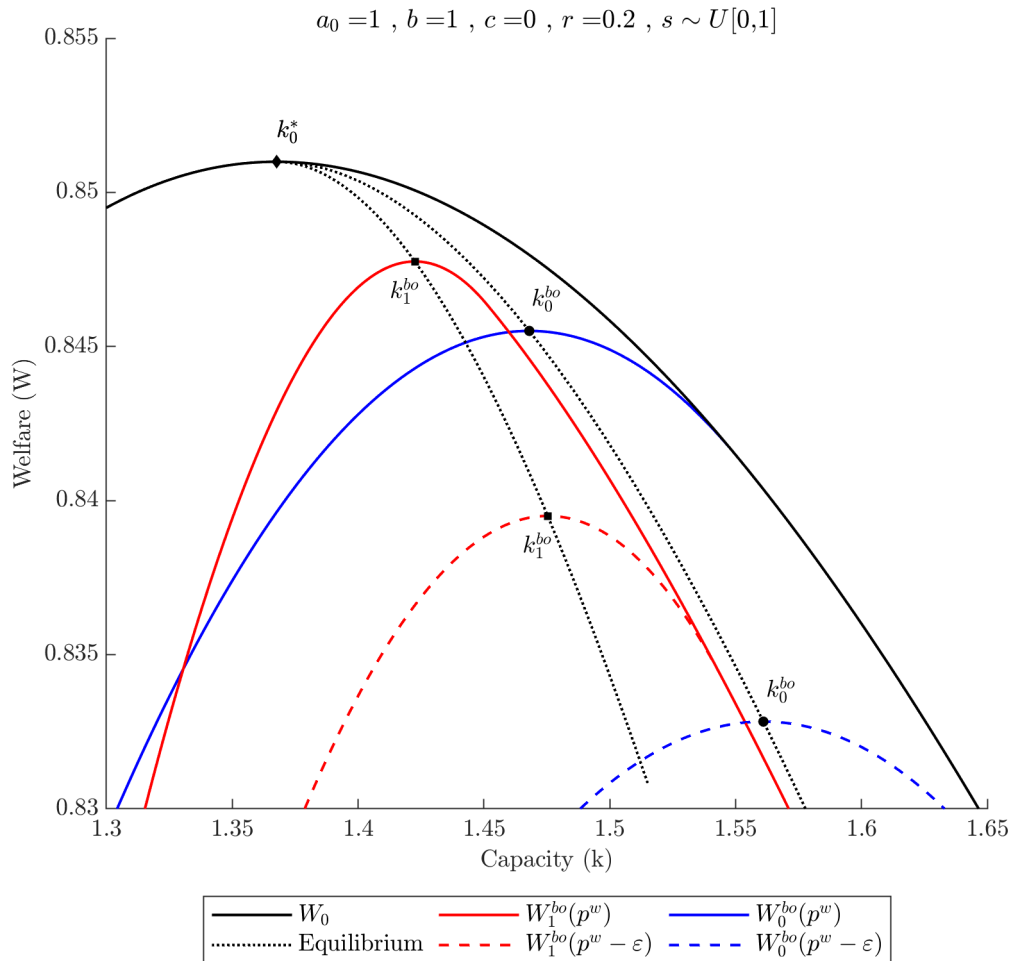


Figure 2.6: Expected social welfare given different capacity market designs. The black curve corresponds to the full efficiency case but also to the exogenous case with only a missing money inefficiency. The blue curve is the exogenous regime with inefficient rationing. The red curve is the endogenous regime with inefficient rationing.

inefficient rationing.³⁷ We then vary the price cap level, with a higher value for solid curves compared to the dashed curves. The equilibrium values, which are also the maximum expected welfare reachable within the two market designs, are represented by black squares and circles for the endogenous and exogenous regimes. We also represented the path for each equilibrium for continuous values of the price cap with the dotted line. As expected, following the results in the different Propositions, for a given price cap, the level of investment is always lower under an endogenous regime. The values of the expected welfare at the black squares are always higher than those at the black circles. Hence, within the framework, the endogenous regime always provides higher welfare than the exogenous regime under inefficient rationing.

The rationale behind the results lies in two distinct causes : (1) the change of welfare due to the indirect effect, for a given investment level, it reduces rationing costs and leads to a negative price effect. (2) the change of investment equilibrium, with previous Proposition 2.3 stating that $k_1^{bo} \leq k_0^{bo}$. The change in investment level between the two market designs can then be decomposed into two components : (i) the relative comparison between the negative price effect and the reduced rationing costs, and (ii) the relative investment costs between the two market design equilibria. For the second component, the effect is always positive for the endogenous regime, as a lower investment level always implies lower investment costs. The dominance ranking between the opposite effects is ambiguous for the first component. The endogenous reduced rationing costs for a given equilibrium investment level do not always overcome the negative price effect it generates. We even find that having a lower investment level can penalize the endogenous regime regarding the net effect. However, as shown in Claim 2.1, saving due to lower investment costs always implies higher welfare under the endogenous regime at the equilibrium.

Figure 2.7 illustrates the interaction between those opposite effects. The dashed curves represent them under a unique investment level (as described by $\Delta W_1^{bo}(k)$), which is arbitrarily taken equal to the equilibrium under the exogenous market design k_0^{bo} . As expected, the difference in investment costs (black curves) is null. For low values of the price cap, the reduced rationing costs dominate the negative price effect. As the price cap increases, there is a switch, and the negative price effect dominates. We then show the actual comparison between the two regimes at their respective equilibrium with the solid curves. In that case, the gain in avoided rationing costs decreases, and the negative price

³⁷Note the convergence for the different curves to the right. Above a specific value of k , the price cap never binds in expectation, and inefficient rationing ceases to exist. It only remains the negative price effect of the endogenous regime for the blue curves.

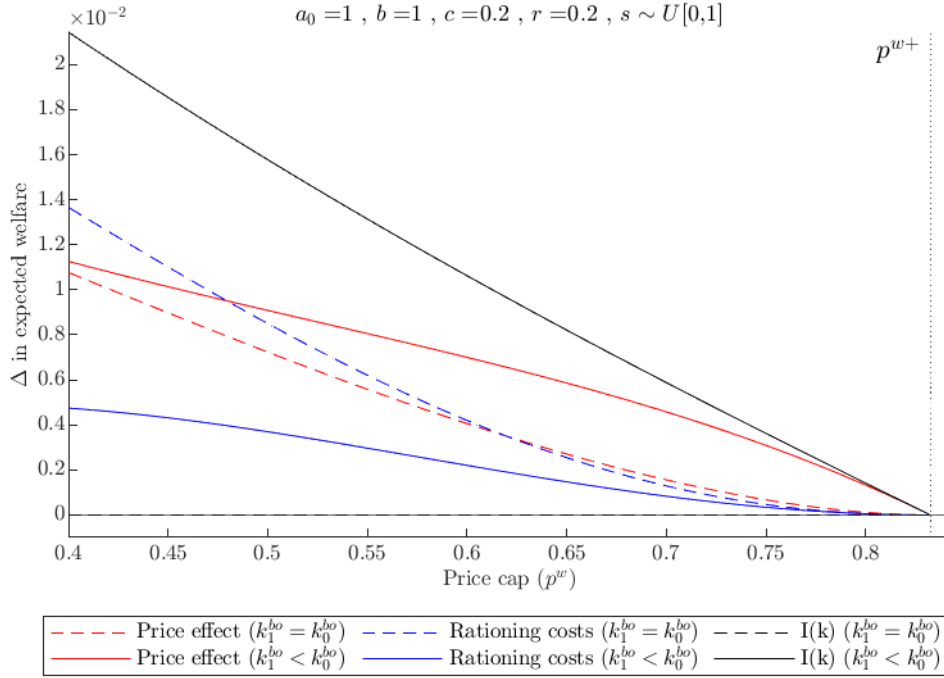


Figure 2.7: Decomposition of the change in welfare at the equilibrium between endogenous and exogenous market design with inefficient rationing

effect increases. However, the lower investment level saves investment costs, represented by the solid black curves. In that case, the sum of the black and blue curves always dominates the red curve, which implies higher welfare under the endogenous regime at the equilibrium.

5 Extensions - Retailers allocation

In this section, we provide an initial analysis of a more complex market design for the capacity market using the previous framework. Under those regimes, the capacity market cost allocation becomes dependent on the consumption realization. The primary motivation for those extensions is to formalize the idea that capacity markets are not only supply-side instruments. In other words, by including the demand-side in the design, capacity markets can provide better price signals to increase efficiency.³⁸ Therefore, we apply the initial model to two settings : (i) when the allocation depends on the realized

³⁸This idea was described in the following report aimed at justifying the decentralization of the capacity market (RTE, 2014).

market share of retailers playing 'à la Cournot' and (ii) when the capacity market is fully decentralized where retailers by themselves the capacity. In the first regime, we study the link between the model outcome and the degree of competition in the retail market. In the second regime, we derive the equilibrium emerging in the capacity market without a regulator. We show that the positive effect of a decentralized capacity market can be monitored by the regulator using the penalty system.

5.1 Retailers Market Share allocation

Under this market design, the capacity allocation depends on the retailers' realized quantity sold to the final consumers. To represent retailers' market share, we assume in this section that there is imperfect competition in the retail market. We underline that having different degrees of competition in the retail market has a direct effect on the capacity cost allocation sustained by final consumers. It also indirectly impacts the efficiency of the market in a somewhat different way than imperfect competition on the supply-side.³⁹ We describe in Appendix 1 the main effect of having a Cournot competition in the retail market on the investment level that maximizes expected welfare. We introduce the capacity market by describing the new market design regime as follows.

Assumption 2.4. *The regulator builds a demand function in the capacity market and buys a level of capacity k for a price $p^c(k)$ given the supply function described in Equation 2.3. Then, it allocates the full cost $kp^c(k)$ directly to the retailers. The share of the capacity cost is based on their realized market share in the retail market. For a retailer i , this share is defined as $\frac{q_i}{q_i+q_{-i}}$ with q_i its quantity sold on the retail market and q_{-i} the quantity sold by its competitors.*

The first implication of ex-post allocation concerns the last stage, which is when the retail market clears. We rewrite the retailers' profit function by including an endogenous ratio in the retailer profit function, as shown in the following equation.

$$\Pi_i^r(q_i, k) = q_i(p(q) - p^s) - \underbrace{p^c(k)k \frac{q_i}{q_i + q_{-i}}}_{\text{capacity market cost}}$$

³⁹See, for instance, (Léautier, 2016) or Zöttl (2011) for an analysis of market power on the supply-side.

Contrary to the previous section, we do not need to assume any tariff hypothesis for the capacity cost allocation as it directly affects retailers' profit at the margin. We focus the analysis on symmetric equilibrium. We drop the notation with s as the state of the world is known at this stage. With $q_{-i} = \sum_{j \neq i}^n q_j$. We find the best-response function of a retailer i with the first-order conditions :

$$BR_i(q_j) = \max_{q_i} \Pi_i^r(q_i, k) \iff p(q) + q_i p_q - p^s - p^c(k)k \frac{q_j}{(q_i + q_j)^2} = 0$$

The main results for the existence of equilibrium are stated in the following lemma:

Lemma 2.6. *At the symmetric equilibrium, retailers' profit function is concave if the capacity cost is not too important, that is, if the following condition holds :*

$$-q^2 p_q \left(\frac{n+1}{n-1} \right) > p^c k$$

The quantities are strategic substitutes whenever the following condition holds :

$$k p^c(k) \left(\frac{n-2}{n} \right) \frac{1}{q^2} \geq -p_q$$

Proof. See Appendix 9 □

Note that when $n = 2$, the equilibrium is always unique and stable. This observation comes from the classical decreasing marginal returns of the Cournot literature [Vives \(1999\)](#). When $n > 2$, the lemma states that the capacity market allocation design induces a stricter condition on the marginal returns, which needs to consider the additional cost in the retailer's profit. Using the first-order conditions and the symmetry between the retailers, the Cournot equilibrium in the retail markets allows us to define the endogenous retailer demand function in the wholesale market:

$$\tilde{p}(q) = p(q) - mp(q) - p^c(k)k \frac{1}{q} \frac{n-1}{n}$$

With $mp(q)$ the markup associated with the market power in the retail market such that $mp(q) = -\frac{q}{n} p_q > 0$ ⁴⁰. The concavity condition stated in Lemma 2.6 is the same as

⁴⁰We simplify $\frac{\partial mp(q)}{\partial q} = mp_q$ and $\frac{\partial mp(q)}{\partial n} = mp_n(q)$.

assuming that the demand function in the wholesale market is decreasing. The equilibrium in this market design is similar to the endogenous regime in the previous section due to the effect of the capacity price on the final demand. Therefore, we can define the new thresholds between on-peak/off-peak/binding price cap periods. We denote them s_n and s_n^w such that the expected value of $\tilde{p}(q, s)$ is equal to, respectively the marginal cost and the price cap, that is $\tilde{p}(q, s_n) = c$ and $\tilde{p}(q, s_n^w) = p^w$. We denote the corresponding quantity q_n and q_n^w such that $\tilde{p}(q_n, s) = c$ and $\tilde{p}(q_n^w, s) = p^w$.

The indirect effect of this market design can be shown in the following expression of the supply function in the capacity market:⁴¹

$$X_n(k, p^c) = r - \left(\int_{s_n}^{s_n^w} (p(k) - c - \overbrace{mp(k)}^{\text{cournot}} - \underbrace{p^c(k) \frac{n-1}{n}}_{\text{capacity market}}) dF(s) + \int_{s_n^w}^{\bar{s}} (p^w - c) dF(s) \right)$$

This expression can be understood similarly as Equation 2.7. The marginal revenue made on the wholesale market depends on the markup and the depressing effect of the capacity price on the demand. The endogeneity created by the market design stems from the fact that the expression $p^c(k)$ appears in both terms via the demand function in the wholesale market and the threshold value s_n and s_n^w . Note the difference with the requirement in the first part of the integrals, where the capacity cost adder is dependent on n . The following Proposition summarizes the main effect of an ex-post allocation based on the realized market share. Allocating the capacity market cost based on retailers' realized market share provides an intermediate indirect effect between an exogenous regime price and an endogenous regime. In the Appendix, we provide a formal definition of the level of investment (respectively $k_{0,n}^*$ and $k_{1,n}^*$) under the two regimes with imperfect competition.

Proposition 2.4. (i) *If the condition in 2.6 holds, then there is a unique investment level under the market share allocation regime that maximizes the expected social welfare.*
(ii) *If it exists it solves $k_n^* = \{k : \phi_n(k) = r\}$, with $\phi_n(k)$ define as follow*

⁴¹For clarity, we do not expose the full effects of this market design on the expected social welfare nor the first-best solution. Indeed, it has consequences similar to the endogenous design; therefore, it implies a depreciation of the demand for the final good. It changes the expected surplus during off-peak states of the world and the occurrence between off-peak / on-peak periods, and the first-best investment level is lower than an allocation without indirect effect.

$$\phi_n(k) = \int_0^{s_n(k)} \frac{\partial q_n}{\partial k} p^c(k) \frac{n-1}{n} dF(s) + \int_{s_n(k)}^{\bar{s}} (p(k, s) - c) dF(s) \quad (2.10)$$

(iii) The first-best investment level is lower than the first-best under exogenous design and higher than the first-best under the endogenous regime ($k_{1,n}^* \leq k_n^* < k_{0,n}^*$). Moreover, the ranking extends the expected social welfare at the first-best investment level: $W_{1,n}(k_{1,n}^*) \leq W_n(k_n^*) < W_{0,n}(k_{0,n}^*)$

Proof. See Appendix 11. □

The capacity cost adder when $n = 2$ is equal to half of the cost adder of equation $X_1(k, p^c(k))$ increases with n . When $n \rightarrow +\infty$, the capacity cost is entirely allocated to the consumer, mimicking the exogenous equilibrium. This Proposition states that increasing competition in the retail market increases the burden of consumers' capacity prices. Hence, the negative effect observed in the regime with endogenous capacity prices is now shared between retailers and consumers. By extension, we have the same results for the endogenous regime when we take into account inefficient rationing with the ex-post market share allocation. Namely, the depressing effect shown in Proposition 2.4 will both lower the expected surplus and the expected rationing cost, hence having an ambiguous impact on the optimal outcome.

In Figure 2.8, we show the expected social welfare under various assumptions. The black curve represents welfare when there is no inefficiency. The blue solid curve stands for the new welfare function when we assume the retailers are playing *a la Cournot* under the exogenous regime, and the dashed curve represents the endogenous case with imperfect competition. The red curve represents the expected social welfare given the effect of the capacity market design based on realized market share. We conclude this section by studying the impact of the degree of competition in the retail market. In this extension and using the analytical framework, we show that a change in market structure can have ambiguous effects. We summarize the main findings in the following corollary.

Corollary 2.1. (i) *With imperfect competition and without inefficient rationing, an increase in n increases the investment level that maximizes expected welfare under the exogenous and endogenous regime.* (ii) *Introducing either inefficient rationing or the retailer market share allocation leads to an ambiguous effect of an increase in n .*

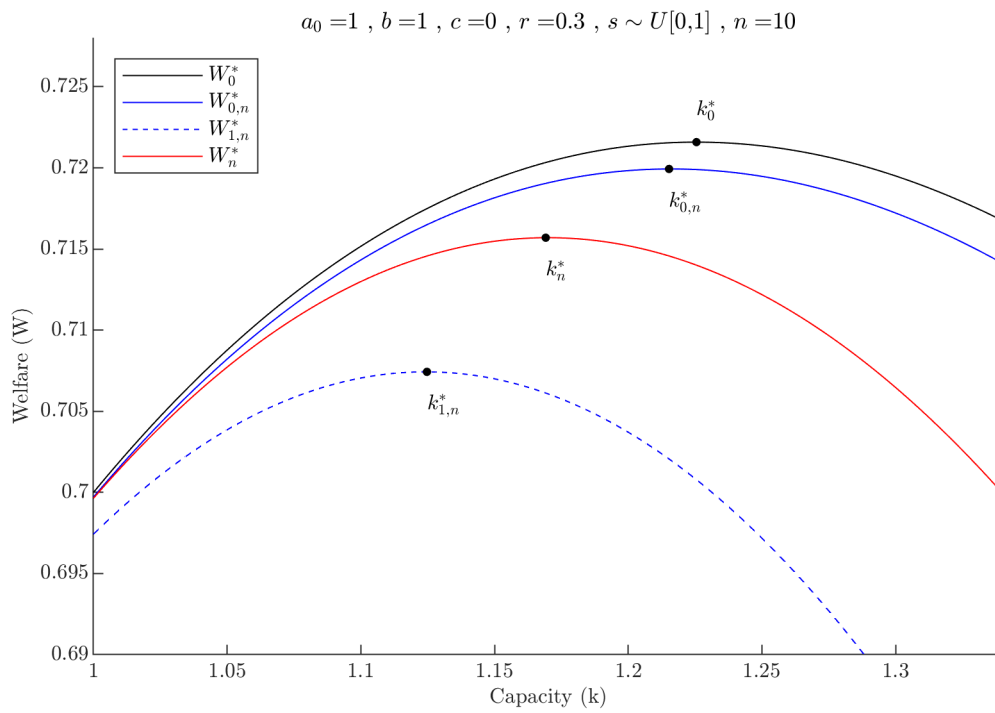


Figure 2.8: Expected social welfare under the market share allocation compared to the first-best welfare and the exogenous and endogenous regime with imperfect competition. The retailer market share allocation provides an intermediate welfare compared to the two previous regimes.

Proof. See Appendix 1 for a discussion on the regimes of the previous sections. See Appendix 12 for the formal proof with respect to retailer market share allocation and inefficient rationing. \square

Without inefficient rationing, increasing the number of retailers has two effects: (i) it decreases the market power in the retail market, which raises demand and, by extension, both the investment level and the expected welfare, and (ii) it indirectly affects the equilibrium in the capacity market. Namely, by providing a higher demand, it also increases the producers' revenue. Hence, it lowers the supply function in the capacity market, implying that for the same level of investment, the capacity price is lower. Under the endogenous regime, it lowers the negative price effect described in section 3, which creates a second positive channel for an increase in demand. Implementing the retailer's market share allocation has an ambiguous effect. Indeed, Equation 2.10 shows that the indirect impact created by the market design is directly increasing with n as $\frac{n-1}{n}$ also increases with n . Therefore, a third effect is observed under this regime as more competition increases the burden supported by consumers, which in turn decreases their consumption. The negative impact can be illustrated in the following expression:

$$\frac{\partial q_n}{\partial n} = -\frac{1}{\tilde{p}_q} \left(\overbrace{mp_n - \mathbb{C}_{p^c} \frac{\partial p^c}{\partial n}}^{+} \quad \overbrace{- \mathbb{C}_n}^{-} \right)$$

With $\mathbb{C} = p^c(k)k^{\frac{1}{q}}\frac{n-1}{n}$, the cost associated with the capacity market in the demand function \tilde{p} . The two first terms in the numerator show the positive effect of an increase in n : (i) it decreases the markup $mp_n < 0$, and the capacity price decreases with n : $\frac{\partial p^c}{\partial n} < 0$, and the capacity market cost also decreases with the price $\mathbb{C}_{p^c} > 0$. On the other hand, the cost pass-through associated with the capacity market increases, which generates the negative effect: $\mathbb{C}_n < 0$. Similarly, having inefficient rationing creates a new tradeoff. As the degree of competition increases, quantity is also increasing. This, in turn, implies that the price cap is binding more often and that the rationing cost also increases. This negative effect is realized as soon as the demand increases with n , whatever the regime considers. We can illustrate the effect of n on the expected welfare with the following derivative of the expected welfare under the endogenous regime with imperfect competition and with inefficient rationing:

$$\frac{\partial W_{1,n}}{\partial n} = \overbrace{\int_0^{s_{1,n}^1} \frac{\partial q_{1,n}}{\partial n} (q_{1,n} - c) dF(s)}^{+ \text{ competition effect}} - \overbrace{\int_{s_{1,n}^w}^{\bar{s}} \frac{\partial q_{1,n}^w}{\partial n} \frac{k}{(q_{1,n}^w)^2} \int_0^{q_{1,n}^w} (p(q, s) - p(q_{1,n}, s)) dq dF(s)}^{- \text{ rationing effect}}$$

The first term represents the effect of increasing n on the expected social welfare. If an increase in the competition in the retail market translates to a higher quantity (i.e., $\frac{\partial q_{1,n}}{\partial n} > 0$), this term is positive. This is the case under both exogenous and endogenous regimes but not always under the retailer market share allocation. The second term represents the negative effect of having more competition. If the quantity at the price cap increases with n , that is $\frac{\partial q_{1,n}^w}{\partial n} > 0$, then this term is negative. The net effect depends on the model parameters, especially the relative share of off-peak and on-peak periods.

5.2 Decentralized capacity market

This last section provides the first analysis of a fully decentralized capacity market. This market design regime takes the furthest step towards accounting for the final electricity demand in the capacity market allocation. Each retailer must purchase a certain amount of capacity in the capacity market. The regulator only monitors ex-post the level of capacities and compares it to each retailer's sales. A penalty mechanism is implemented if there is any difference between the two quantities. We formally describe this market design regime as follows:

Assumption 2.5. *The regulator mandates the retailers to buy the capacity on the capacity market given their realized sales in the retail market for a price $p^c(k)$ given the supply function described in Equation 2.3. For each retailer, if their individual realized sale quantity on the retail market q_i is above their purchase quantity on the capacity market k_i , the regulator imposes a unitary penalty S , such that the penalty mechanism total cost for a retailer is $S(q_i - k_i)$.*

One of the critical features of this regime concerns the case when a retailer is in negative deviation, i.e., has sold more on the retail market than he has bought capacity in the capacity market. In this case, he suffers a penalty, which results in a payment

from the retailer to the regulator by a unitary amount of S , with $S \geq 0$ being an administratively fixed value.⁴²

We use the following process to describe the implications of this market design regime. We develop the results on the idea that retailers act similarly to producers when given the opportunity to participate in the decentralized capacity market. With imperfect competition and a price cap, their expected profit in the retail market depends on the level of capacity. It means that we can also define an expected unitary rent $\phi^d(k)$, similarly to the producers' rent $\phi(k)$. Namely, the existence of a capacity market implies that they can also choose to invest, with the capacity price acting as an investment cost. We introduce the effect of the penalty on retailers' behavior, punishing retailers for having too little capacity. While having little interest in the initial case with a price cap (and imperfect competition) generating solely missing money for producers, the penalty takes on its full meaning when we assume inefficient rationing. Indeed, in this case, the additional cost sustained by consumers is due to the inadequacy between the quantity consumed (and hence sold) to consumers and the level of investment. Therefore, the penalty acts as a means to make partly responsible retailers for this loss.

In terms of market design, the previous observation translates into the fact that the penalty is sustained only when the price cap is binding, that is, only when inefficient rationing occurs. First, let's denote an additional threshold value s_d such that this is the first state of the world when the price cap is binding after accounting for the depressing effect of the penalty $p_s(k, s_d) - S = p^w$.⁴³ We also denote q_d the corresponding quantity, which can also be interpreted as the *Cournot* equilibrium in the retail given retailers' profit function and penalty mechanism. With this new value, we can distinguish three cases depending on the value of the installed capacity.

- (Case 1) When $k > q_{0,n}^w$, the price cap is never binding, and the effect of a capacity market on the expected retailers' profit is strictly identical to an exogenous regime with imperfect competition.
- (Case 2) *implicit rationing*: For a value of k between q_d and $q_{0,n}^w$, we observe a paradoxical outcome. Rationing should have occurred as soon as k is below $q_{0,n}^w$ without a penalty. It implies that retailers sustain the penalty, which is then passed

⁴²Some remuneration mechanisms can exist (in France, for instance) so as to reward retailers who have provided additional capacity. Still, as we focus on symmetric equilibrium, they do not play a role in the outcome.

⁴³The penalty does impact retailers at the margin as the mechanism is based on the difference between the realized quantity, retailers' strategic variable, and the investment level.

to consumers as a marginal cost, which lowers their demand. However, rationing does not occur as $q_d < k$, which contradicts the demand's decrease due to the penalty. Therefore, we assume that retailers follow the level of investment. To do so, they increase the price of their consumers by a unitary amount of $T(k, s) \leq S$ so that at any state of the world between $s_{0,n}^w$ and s_d , the demand is equal to the capacity k ,⁴⁴ that is we have $p^s(k, s) - T(k, s) = p^w$.

- (Case 3) k is below q_d , it is now optimal for the retailers to keep their strategy at q_d before rationing as it is the profit-maximizing quantity given the penalty mechanism.

Therefore, the core idea of the three cases is that given the penalty value, retailers have two options when the price cap is binding : (i) decreasing its consumption due to T , (ii) letting the consumption level at the price cap, which leads to both inefficient rationing for the market and sustaining the penalty for the retailers. In the simplified framework, case (2) corresponds to the preference for retailers to follow (i) as $T < S$ and case (3) corresponds to (ii) as $T > S$.⁴⁵ Given the three different cases, the expected profit function of retailers becomes⁴⁶:

$$\begin{aligned} \Pi_d^s(k) = & \overbrace{\int_0^{s_{0,n}} q_{0,n}(s)mp(q_{0,n})dF(s) + \int_{s_{0,n}}^{\bar{s}} kmp(k)dF(s)}^{\text{expected revenue}} \\ & + \underbrace{\int_{s_{0,n}^w}^{s_d} kT(k, s)dF(s)}_{\text{implicit rationing}} - \underbrace{\int_{s_d}^{\bar{s}} S(q_d - k)dF(s)}_{\text{penalty cost}} - \underbrace{p^c(k)k}_{\text{capacity cost}} \end{aligned}$$

The first term corresponds to the expected revenue for retailers. Under the framework, it is equal to the markup times the quantity, which is either the quantity at the marginal cost or the investment level. The second term corresponds to what is happening in case (2) when retailers deviate from the Cournot equilibrium to lower their consumption.⁴⁷ The third term corresponds to case (3) when retailers do not change

⁴⁴We could also assume the reverse mechanism where retailers pay consumers $T(k, s)$ to reduce the demand in order to avoid the penalty.

⁴⁵We leave for future work a more complex representation of equilibrium in this framework.

⁴⁶We focus on symmetric equilibrium, so we drop the reference to individual profit.

⁴⁷Note that if we assume that retailers pay the consumers to reduce their consumption, only the sign changes.

their consumption and, therefore, sustain the penalty. Finally, the last term is the capacity cost, as retailers have to pay for the level of investment via the capacity market. We now express the marginal value of an investment for retailers without the capacity price:

$$\begin{aligned} \phi^d(k) = & \overbrace{\int_{s_{0,n}}^{\bar{s}} (mp(k) + kmp_q)dF(s)}^{\text{marginal revenue}} + \overbrace{\int_{s_{0,n}^w}^{s_d} \left(T(k, s) + k \frac{\partial T}{\partial k} \right) dF(s)}^{\text{marginal implicit rationing}} + \\ & + \underbrace{kSf(s_d) \frac{\partial s_d}{\partial k}}_{\Delta \text{ in occurrence}} + \underbrace{\int_{s_d}^{\bar{s}} SdF(s)}_{\text{avoided penalty}} \end{aligned}$$

The first line corresponds to the effect of an additional investment on retailers' profit. When the capacity is constraining, an increase in k increases the quantity for retailers and also changes their market power. In that case, we have $mp_q > 0$. By definition of $T(k, s) = p_s(k, s) - p^w$, we have $\frac{\partial T}{\partial k} = p_q - mp_q$. The first term in the second line is due to the fact that as k increases, it increases at the higher bound s_d the occurrence of periods of implicit rationing, at s_d , we have by definition $T(k, s) = S$, which implies the term. Finally, the last term stems from the fact that an increase of k lowers the difference between the quantity at the price cap and the investment; that is, it lowers the burden of the penalty and increases the level of investment. In the following lemma, we define the condition for having an equilibrium k_d in the capacity market, which defines the level of investment. It boils down in the framework to the equality between the supply function $X_{0,n}$ ⁴⁸ with imperfect competition and the marginal value of an investment for producers.

Lemma 2.7. *If $S \geq \frac{2\bar{s}}{n+1}$, then there exists a unique solution k_d to $\phi^d(k) = X_{0,n}(k)$, with $X_{0,n}(k) = r - \phi_{0,n}^w(k)$ the supply function in the capacity market and $\phi_{0,n}^w(k)$ the private marginal revenue of an investment for producers such that:*

$$\phi_{0,n}^w(k) = \int_{s_{0,n}}^{s_{0,n}^w} (p(k, s) - mp(k) - c)dF(s) + \int_{s_{0,n}^w}^{\bar{s}} (p^w - c)dF(s) \quad (2.11)$$

⁴⁸With imperfect competition, when the capacity is binding but not the price cap, then producers receive the retail price minus the retailer's markup. This corresponds to the first term. The second term is identical to the perfect competition case.

Proof. See Appendix 13. □

The proof relies on ensuring that the demand function on the capacity market formed by the marginal value of investment for retailers is decreasing and at least positive for some values of k . Note that the supply function is increasing and bound by 0 and r . Therefore, there is one intersection between the two functions. $\phi_{0,n}^w$ is given similarly and combines the imperfect competition and the price cap effects. In that case, $r - \phi_{0,n}^w$ is simply the supply function as described in Definition 2.1. The lower condition on S is to ensure the concavity of the retailer profit function with respect to k .

We turn now to the study of the implications of a decentralized capacity market from a social welfare aspect. For clarity, we only analyze the case when we take into account inefficient rationing, as the decentralized capacity market with a penalty mechanism is usually explicitly implemented to account for this inefficiency. The efficiency of a decentralized relies on two channels: (i) the market equilibrium it provides and (ii) the indirect effect it has on the expected social welfare. The new expected welfare function under a decentralized market design and with inefficient rationing is given by the following equation:

$$W_d(k, S) = \tilde{W}_{n,0}(k, S) + CS_d(k, S) - rk \quad (2.12)$$

With the consumer surplus under inefficient rationing:

$$CS_d(k, S) = \int_{s_d}^{\bar{s}} \frac{k}{q_d} \int_0^{q_d} (p(q, s) - p^w) dq dF(s)$$

With $\tilde{W}_{n,0}$ being the social welfare without consumer surplus when there is inefficient rationing. It is similar to the expected welfare with imperfect competition and an exogenous regime $W_{n,0}$.⁴⁹ The core idea is found in the new threshold values: s_d and q_d . As previously explained, the incentives created by the penalty mechanism induce retailers to lower their sales in any state of the world between $s_{0,n}^w$ and s_d (case(2)), such that the

49

$$\tilde{W}_{n,0} = \int_0^{s_{n,0}} \int_0^{q_{n,0}} (p(q, s) - c) dq dF(s) + \int_{s_{n,0}}^{s_d} \int_0^k (p(q, s) - c) dq dF(s) + \int_{s_d}^{\bar{s}} \int_0^k (p^w - c) dq dF(s)$$

quantity sold is strictly equal to installed capacity. Therefore, no inefficient rationing is realized during those periods. As the investment level, and therefore the welfare at the equilibrium, is endogenous to the model, we summarize the link between the penalty and the market outcome in the following proposition.

Proposition 2.5. *(i) Under the decentralized market design, the equilibrium level of investment $k_d(S)$ is increasing in the penalty value. The welfare at the equilibrium $W_d(k_s(S), S)$ is concave in the penalty S . (ii) The optimal penalty value that maximizes expected welfare is such that there is no inefficient rationing.*

Proof. See Appendix 14. □

The absence of closed-form solutions prevents having a clear-cut comparison between the different regimes. However, as shown in the Proposition, we do find that the penalty can play a central role in the efficiency of the decentralized capacity market. From Equation 2.12, we can express the derivative of the welfare at the market equilibrium k_d as follows with respect to the penalty value:

$$\frac{\partial W_d(k_d)}{\partial S} = \overbrace{\frac{\partial CS_d(k_d, S)}{\partial S}}^{+} + \left(\overbrace{\frac{\partial W_d(k_d, S)}{\partial k}}^{+ \text{ then } -} + \frac{\partial CS_d(k_d, S)}{\partial k} \right) \frac{\partial k_d}{\partial S}$$

That is, a change in the penalty value that affects the expected welfare is measured by the direct change in the welfare (first term) and by an indirect change in the equilibrium quantity (second term). The proof in the Appendix shows that this direct effect is composed only by a change in that rationing cost and is always positive. In other words, increasing the penalty always increases welfare due to a reduction in the cost of rationing. This stems from the fact that as the penalty increases, it is less profitable for retailers to sustain the penalty, which incentivizes them to increase the implicit rationing period. The expected welfare is concave, following the standard arguments of the chapter. However, we find that the equilibrium is increasing in the penalty:

$$\frac{\partial k_d}{\partial S} = - \frac{\frac{\partial \phi_d}{\partial s}}{\frac{\partial \phi_{0,n}^w}{\partial k} + \frac{\partial \phi_d}{\partial k}}$$

The penalty does not change the expected profit of producers, as they always have a revenue composed of the price cap and the investment level whenever they are binding,

which is also when the penalty never comes into effect. We show in the Appendix that the penalty positively affects the marginal value of the retailers: $\frac{\partial \phi_d}{\partial s} > 0$. Moreover, the capacity market supply function is equal $r - \phi_{0,n}^w$, hence it is positive: $\frac{\partial \phi_{0,n}^w}{\partial k} < 0$. Similarly, ϕ_d is the capacity market function; hence, it is decreasing. It implies that all terms are positive; hence, the market equilibrium is increasing in S . Finally, the intuition for the second part of the result originates from the observation that as the penalty values increase, it decreases the occurrence of periods during which inefficient rationing occurs. As it also increases capacity, this also lowers the occurrence. At one point, the value is such that there are no inefficient rationing periods ($s_d = \bar{s}$). This implies that at this point, the penalty does not play any role in the market: $\frac{\partial \phi_d}{\partial s} = 0$, which coincides with the maximum of k_d . From a welfare perspective, it is also the case where there is no inefficient rationing anymore, which is also the most efficient situation.⁵⁰ We conclude by providing in Figure 2.9 the illustration of the Proposition. It shows that the penalty value clearly impacts the market equilibrium (in blue). The first-best (with imperfect competition but without inefficient rationing) is represented in black. The two red lines represent the exogenous and endogenous regimes with inefficient rationing based on the previous sections. The second plot shows that for the relatively high value of the penalty, the welfare gained under the decentralized market design is significantly higher than the previous market design. In other terms, the indirect effect created by implicit rationing and the burden of the penalty created more values than the endogenous regime. The blue zone represents this additional welfare. However, note that while we show that there is no inefficient rationing at the optimal penalty level because the investment level is still different from the first-best one, there is still a gap between the first-best welfare and the one gained in the decentralized market design.

6 Conclusion

This chapter built a tractable framework to analyze the design of long-term markets for a good prone to underinvestment, such as electricity or medical supplies. We showed how the investment decisions are affected by those markets, their structure (such as the degree of competition), and, most importantly, their design. Our case study is the capacity markets that were implemented to encourage producers to invest by providing additional remuneration. Most of the literature on capacity markets has focused on the

⁵⁰If there is no inefficient rationing then: $\frac{\partial CS_d(k_d, S)}{\partial S} = 0$. Hence, the derivative of the expected welfare is null, which implies that this is also the maximum.

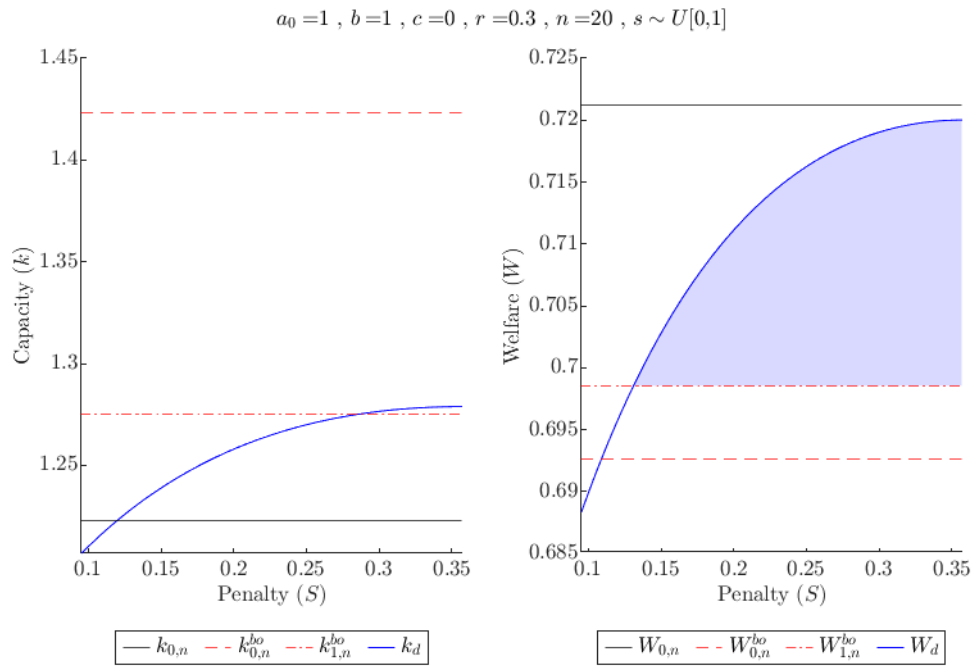


Figure 2.9: Evolution of the market equilibrium and the equilibrium welfare in decentralized capacity market regime with respect to the penalty value S . We compare the outcomes with other regimes (exogenous and endogenous)

supply-side, where producers offer their availability on future transaction periods on the wholesale market. Therefore, the demand-side has been overlooked, even though some efficiency effects are well known. Current implementations show many options regarding the demand-side's design on capacity markets, as consumers do not have proper incentives to buy capacities. Using the framework, we compare multiple market designs and their implications. The first set of regimes is based on differentiating the capacity cost allocation. The second set of regimes is represented by how the design can account for current demand realization. We underline the different parameters that can significantly affect the outcomes of a capacity market on investment decisions. The choice of the design can significantly affect prices and quantities in the three markets. One of the advantages of this framework relies on the possible extensions that we can implement besides providing a simple but complete vision. The rest of the section discusses two issues that could be addressed in future research using this framework.

First, we initially assumed that consumers were fully reactive to retail prices. Such assumptions do not describe the reality yet, as illustrated in the electricity markets, as most small final consumers, such as households, are still under fixed-price contracts. The study of final consumers' heterogeneity and its implications for investment decisions in the power system is an emerging trend. [Léautier \(2014\)](#) and [Léautier \(2016\)](#) provide a relevant model close to the one presented in this chapter. They show the effects of having those two types of consumers with different investment decisions and a capacity market. However, the author does not compare demand design options for capacity markets and does not consider retailers. Therefore, implementing this new extension in the model could shed light on the issue associated with power systems' investment decisions. It could also significantly impact retailers' individual market design options. Indeed, let's consider that some consumers cannot react to price, but retailers are still forced to cover their consumption. The demand function's formation in the capacity market will be significantly impacted.

Finally, we assume that future consumer demand is commonly shared between agents. A single entity, potentially regulated, and retailers could access a different quantity and quality of information. For instance, we can assume that the entity only has a global vision of future demand, and hence, it is prone to make a more significant error forecast than retailers. On the other hand, retailers have private access to more precise information on their client portfolios while sharing common information on the world's future global states. Therefore, introducing these private/common elements in the model could shed new light on the effect of capacity markets and their market design options.

Finally, in some current implementations, the entity based its global forecast on retailers' information. Consequently, the comparison between the various regimes' cases could be analyzed using game theory and signaling.

Bibliography

- Acemoglu, D., Bimpikis, K., and Ozdaglar, A. (2009). Price and capacity competition. *Games and Economic Behavior*, 66(1):1–26.
- Ahuja, A., Athey, S., Baker, A., Budish, E., Castillo, J. C., Glennerster, R., Kominers, S. D., Kremer, M., Lee, J., Prendergast, C., et al. (2021). Preparing for a pandemic: Accelerating vaccine availability. In *AEA Papers and Proceedings*, volume 111, pages 331–35.
- Boiteux, M. (1949). *La tarification des demandes en pointe: Application de la théorie de la vente au coût marginal*. place Henri-Bergson.
- Boomhower, J. and Davis, L. (2020). Do energy efficiency investments deliver at the right time? *American Economic Journal: Applied Economics*, 12(1):115–139.
- Brown, D. P. (2018). Capacity payment mechanisms and investment incentives in re-structured electricity markets. *Energy Economics*, 74:131–142.
- Bublitz, A., Keles, D., Zimmermann, F., Fraunholz, C., and Fichtner, W. (2019). A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms. *Energy Economics*, 80:1059–1078.
- Bushnell, J., Flagg, M., and Mansur, E. (2017). Capacity Markets at a Crossroads. *Energy Institute at Haas*. .
- Creti, A. and Fabra, N. (2007). Supply security and short-run capacity markets for electricity. *Energy Economics*, 29(2):259–276.
- Crew, M. A. and Kleindorfer, P. R. (1976). Peak load pricing with a diverse technology. *The Bell Journal of Economics*, pages 207–231.
- De Frutos, M.-A. and Fabra, N. (2011). Endogenous capacities and price competition: The role of demand uncertainty. *International Journal of Industrial Organization*, 29(4):399–411.

- De Maere d’Aertrycke, G., Ehrenmann, A., and Smeers, Y. (2017). Investment with incomplete markets for risk: The need for long-term contracts. *Energy Policy*, 105:571–583.
- de Palma, A., Lindsey, R., and Monchambert, G. (2017). The economics of crowding in rail transit. *Journal of Urban Economics*, 101:106–122.
- Doorman, G., Barquin, J., Barroso, L., Batlle, C., Cruickshank, A., Dervieux, C., Flanagan, R., Gilmore, J., Greenhalg, J., Höschle, H., et al. (2016). Capacity mechanisms: needs, solutions and state of affairs. *CIGRÉ, Paris*.
- Fabra, N. (2018). A primer on capacity mechanisms. *Energy Economics*, 75:323–335.
- Fabra, N., Motta, M., and Peitz, M. (2022). Learning from electricity markets: How to design a resilience strategy. *Energy Policy*, 168:113116.
- Hobbs, B. F., Hu, M. C., Iñón, J. G., Stoft, S. E., and Bhavaraju, M. P. (2007). A dynamic analysis of a demand curve-based capacity market proposal: The PJM reliability pricing model. *IEEE Transactions on Power Systems*, 22(1):3–14.
- Holmberg, P. and Ritz, R. A. (2020). Optimal capacity mechanisms for competitive electricity markets. *The Energy Journal*, 41(Special Issue).
- IEA (2020). Projected costs of generating electricity. Technical report, IEA.
- Joskow, P. and Tirole, J. (2007). Reliability and competitive electricity markets. *The Rand Journal of Economics*, 38(1):60–84.
- Keppler, J. H., Quemin, S., and Saguan, M. (2021). Why the sustainable provision of low-carbon electricity needs hybrid markets: The conceptual basics. *Hal Science*.
- Klemperer, P. D. and Meyer, M. A. (1989). Supply function equilibria in oligopoly under uncertainty. *Econometrica: Journal of the Econometric Society*, pages 1243–1277.
- Léautier, T.-O. (2014). Is mandating. *The Energy Journal*, 35(4).
- 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.

- Llobet, G. and Padilla, J. (2018). Conventional power plants in liberalized electricity markets with renewable entry. *The Energy Journal*, 39(3).
- Newbery, D. (2016). Missing money and missing markets: Reliability, capacity auctions and interconnectors. *Energy Policy*, 94:401–410.
- Ockenfels, A. (2021). Marktdesign für eine resiliente impfstoff-produktion. *Perspektiven der Wirtschaftspolitik*, 22(3):259–269.
- RTE (2014). Rapport d’accompagnement de la proposition de règles du mécanisme de capacité. .
- Scouflaire, C. (2019). *Capacity remuneration mechanisms: analytical assessment of contemporary experiences and lessons for the future design of electricity markets*. PhD thesis, PSL - Paris Dauphine University.
- Vives, X. (1999). *Oligopoly pricing: old ideas and new tools*. MIT press.
- Weitzman, M. L. (1974). Prices vs. quantities. *The review of economic studies*, 41(4):477–491.
- Wolak, F. A. (2021). Long-term resource adequacy in wholesale electricity markets with significant intermittent renewables. Working Paper 29033, National Bureau of Economic Research.
- Yan, H. (2020). Auctions with quantity externalities and endogenous supply. *International Journal of Industrial Organization*, 71:102638.
- Zöttl, G. (2011). On optimal scarcity prices. *International Journal of Industrial Organization*, 29(5):589–605.

Chapter 3

Designing Markets for Reliability with Incomplete Information

Contents of Chapter 3

| | | |
|-----|---|-----|
| 1 | Introduction | 151 |
| 2 | Environment | 156 |
| 2.1 | Consumers Preferences | 157 |
| 2.2 | Allocation design | 158 |
| 2.3 | Supply side | 159 |
| 2.4 | Timing | 160 |
| 3 | Complete Information | 161 |
| 3.1 | Optimal allocation proposal | 161 |
| 3.2 | Long-term vs. Short-term consumer surplus | 164 |
| 4 | Incomplete Information - Fixed price | 172 |
| 4.1 | Single price policy | 174 |
| 4.2 | Category-price policy | 177 |
| 4.3 | Individual welfare effect | 183 |
| 5 | Incomplete Information - Mechanism Design | 186 |
| 5.1 | Optimal allocation | 186 |
| 5.2 | Optimal allocation and investment level | 190 |
| 5.3 | Individual welfare effect | 197 |
| 6 | Conclusion | 202 |

Abstract

This chapter examines the challenges of allocating a good subject to capacity constraints such as electricity 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, allowing the market designer to distinguish groups of consumers such as households or industries. We show that the optimal allocation implies discriminating against consumers based on their types and categories and that the relative discrimination depends on the level of investment considered. It has significant welfare and distributive implications: an optimal pricing mechanism can minimize the investment cost and lead to a higher aggregate consumer surplus. However, it is not always a Pareto improvement for every consumer, especially for smaller ones. We describe two main environments: the current second-best situation, in which the market designer cannot obtain information about consumers and must choose fixed prices ex-ante, and 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, such as electricity, public transport, or medical goods. 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). It is particularly important in the energy transition context. Indeed, it is crucial to lower the production from fossil fuels but reliable technology and invest massively in carbon-free but intermittent renewables. Moreover, the electrification of end-use consumption also implies that periods of scarcity may occur more often. Therefore, we must carefully design electricity markets by choosing the most efficient pricing mechanisms, allowing for sufficient investment and ensuring demand reacts to scarcity (IEA, 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, increasing contagions and hospital congestion.³ Finally, congestion in transportation systems continues to generate substantial costs (Schrank and Lomax, 2021) and poses challenges for the much-needed modal shift to low-carbon means (ITDP, 2021).⁴

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

²Transportation and distribution infrastructure are also central in current policymakers' debates. 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). 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.

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

⁴Increasing prices for fossil fuel vehicles would reduce pollution and generate revenue to invest in decarbonized means. On the other hand, if we want to encourage consumers to use carbon-free transport solutions, their pricing should also be carefully designed, especially to avoid congestion costs.

In this chapter, we provide a 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 use of consumer heterogeneity allows us to raise the issues of the redistribution generated by an allocation mechanism that is considered more efficient. We find that reaching a certain level of investment that maximizes aggregate consumer surplus and implementing the corresponding allocation mechanism to finance the investment can lead to different welfare levels depending on the consumer types. This issue has been recently studied in several empirical papers in the context of essential goods. For instance, in electricity markets, [Cahana et al. \(2022\)](#) 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 Concerning medical goods, 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 health workers).⁶ Finally, in the case of congestion pricing, [Hall \(2021\)](#) studies the pricing of a lane portion. Using survey and travel time data, the author finds that a fully efficient toll is unnecessary for sufficient welfare and Pareto improvement.

However, most of the recent empirical works study the short-term effect of pricing issues without considering the long-term interactions with the level of investment. On the other hand, numerous theoretical contributions have been made to understand the importance of having sufficient investment.⁷ Nonetheless, they are mainly centered around

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-subsidize adopters of such technologies. One central issue is that the latter are mostly high-income households ([Burger et al., 2020](#)).

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.

⁷Namely, 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. Those mechanisms can range from direct subsidies to the design of more complex competitive markets. An important stream of literature 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](#)).

the supply-side of the problem and consider the demand as given. Therefore, the central contribution of this chapter is to discuss the implications of considering the demand-side when it comes to ensuring an efficient level of investment. We provide a model 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 central mechanism of the chapter is 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 may lead to discrimination of some consumers based on their types. Furthermore, both the existence and the magnitude of the discrimination between consumers depend on the level of investment considered. This means the most efficient mechanism is not always Pareto-improving for every consumer, even considering an increase in available capacity. It has, therefore, significant welfare and distributive implications.

To do so, 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 mechanism, he must consider that demand may exceed the level of available capacity and that specific actions need to be taken to reduce aggregate consumption. This creates an asymmetric effect of the optimal allocation when the capacity is binding or not. Hence, the consideration of the capacity constraint significantly impacts the design of the efficient mechanism and the revenue generated by the mechanism.⁸ We also describe the (potential) incompleteness of the mechanism proposed by the market designer due to implementation constraints.⁹ Finally, note that as the market designer uses the allocation to maximize consumer surplus and finance the investment cost, he is also under a revenue constraint.

⁸In this chapter, 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 blackouts in electricity (Fabra, 2018; Llobet and Padilla, 2018) or congestion costs in transports (Yoshida, 2008; de Palma et al., 2017).

⁹We mean by implementation constraints an environment in which the market designer cannot set the optimal allocation for every demand realization.

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.¹⁰ 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 the 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 environments to highlight the mechanism at play in this framework. 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 revenue 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.

¹⁰The support of the distribution of a consumer type depends on the category to which he belongs.

Related Literature

We build the 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 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. [Holmberg and Ritz \(2020\)](#) study the effect of having inefficient rationing. Consequently, electricity prices do not internalize this additional cost, and the market designer needs to implement an additional stream of revenue for the producers. This 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.¹¹

This chapter is also based on a second stream of papers that is 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 the allocation during the wholesale market at the same time. It is in the same vein as the allocation schedule of this chapter and the probability of being disconnected when demand exceeds 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 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 revenue constraint. However, the focus of

¹¹This 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.

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.¹² In particular, [Akbarpour et al. \(2023a\)](#) and [Akbarpour et al. \(2023b\)](#) provide a framework with consumers' characteristics, such as the private information and the publicly-observed categories, in line with the current chapter. They study the trade-off between allocating certain vaccines on a free but random basis or using prices to discriminate and extract information from consumers. The authors assume 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. This chapter endogenizes the revenue preference by implementing investment decisions with a revenue constraint. We also provide results when the market designer can imperfectly implement prices. Finally, a recent paper by [Crampes et al. \(2023\)](#) studies the implementation of an optimal Pareto income tax schedule *à la Mirrlees* when considering the interaction between consuming energy services (heating, air conditioning, light) and investing in energy efficiency with incomplete information about consumer utility. This chapter has a similar spirit, but the link between consumption and production capacity fundamentally differs.

2 Environment

In this section, we describe the idiosyncratic characteristics of an electricity system for clarity of exposition. Note that while the terminology is specific, the results can be applied, with some modifications, to other essential goods as described in the introduction.

¹²This chapter fits within 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.

(i) The demand-side, which can be interpreted as households, industrial 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 describes how investment and production decisions are made. This current version of the chapter 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 revenue constraint or a set of perfect competitive producers without market failure or public interventions. Finally (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 $[\underline{\theta}_i; \bar{\theta}_i]$. We assume uniform distribution for the type.¹³ 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 follows¹⁴:

Definition 3.1. *If the consumer category i is of a higher type than consumer category j , then $\mu_i \theta_i^{av} > \mu_j \theta_j^{av}$*

¹³While this is a restrictive assumption, we believe that the main results in this chapter hold under different types of distribution functions. For instance, using an exponential distribution will place more weight on periods during which capacity does not bind. It changes the order of magnitude but not the fundamental trade-offs.

¹⁴In this chapter, the terms a higher type consumer and a bigger consumer are interchangeable, as well as a smaller type and smaller.

With θ_i^{av} s the average type for category i : $\theta^{av} = \frac{1}{2}(\theta_i + \bar{\theta}_i)$. On the other hand, we suppose that consumers are also subject to an individual but identical shock represented by s , which will be considered as the state of the world in this chapter. When the shock is realized, 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 $2s$. We assume uniform distribution for the common shock.

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 $F(\cdot)$ [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. 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 $CS(q, \theta, s) = U(q, \theta, s) - tq$. 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 $q_i : [\theta_i, \bar{\theta}_i] \rightarrow \Delta(Q(s))$ 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 $Q(s) = \sum_i Q_i(s) = \sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta)$.

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 chapter, we also provide another pricing mechanism that we name *market allocation*, which is formally defined as follows:

Definition 3.2. *A market allocation for a consumer is defined as follows: (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 use the fact that the market allocation is an *implementable mechanism* mimicking a mechanism design outcome, or on the contrary, as a constraint for the market designer. To present the 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 and makes investment decisions. It encompasses the literature on the management of a public firm or the direct regulation of a private monopolist. The model also describes 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 chapter, we abstract from those details. We assume a market designer making both production and investment decisions.

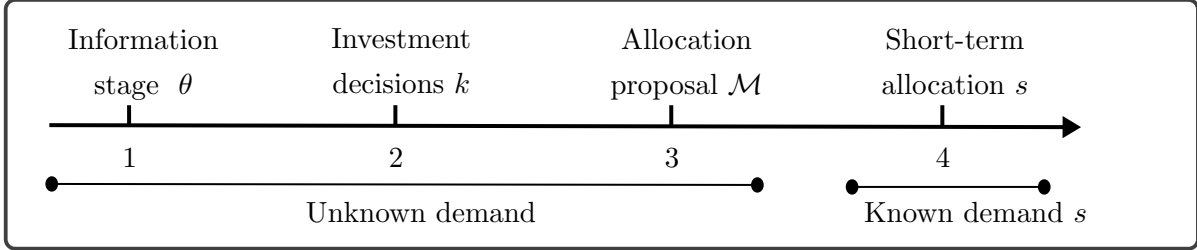
We denote the level of investment k . The investment cost is linear with $I(k) = rk$. The production cost is unitary and normalized to 0. The capacity level k implies a capacity constraint such that for any total quantity allocation Q and any realization of s , we must have at $Q(s) \leq k$.

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 the optimization problem as follows:

$$\begin{aligned} \max_k \max_{\substack{t_i(\theta, s), \\ q_i(\theta, s)}} \quad & CS(k) = \sum_i \mu_i \int_s \int_{\theta_i} (U(q_i(\theta, s), \theta, s) - t_i(\theta, s)q_i(\theta, s)) dG_i(\theta) dF(s) \\ \text{s.t.} \quad & I(k) \leq \sum_i \mu_i \int_s \int_{\theta_i} (t_i(\theta, s)q_i(\theta, s)) dF(s), \quad (\text{R}) \\ & \sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) \leq k, \quad (\text{K}) \end{aligned}$$

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 the 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 3.1: (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's define the inverse demand function:

$$p(q, s) = \sum_i \mu_i \int_s \int_{\theta_i} u(q, \theta, s) dG_i(\theta) dF(s)$$

With s_1 as the first state of the world when the capacity is binding:

$$\sum_i \mu_i \int_{\theta_i} d(0, \theta, s_1) dG_i(\theta) = k$$

Then

Proposition 3.1. (1) *The optimal unity monetary transfer and quantity schedule is defined for each realization of s as follows:*

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

(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:*

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

(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 offer truthfully their demand functions, the market outcome is the first-best allocation.*

Proof. See Appendix 1 □

Solving for the Lagrangean shows that when the capacity is not binding, the optimal allocation is characterized by an expected marginal utility null. On the other hand, when the capacity is binding, the optimal allocation should be equal to the marginal investment cost. 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 this 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 low values of k we have $s_1 = 0$, (ii) the capacity never binds for any s , that is high values of k we have $s_1 = \bar{s}$ (iii) the capacity binds for some s , that is for intermediate values of k we have $s_1 \in [0, \bar{s}]$. For the last case, we can express, for instance, the expected consumer utility under the optimal single-price allocation as follows¹⁵:

¹⁵In the rest of the chapter, 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 the main reference.

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

$p^k = p(k, s)$ 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 pay attention to consumer surplus on a short-term horizon. To represent this behavior, we analyze in this section the consequences of choosing the investment level first and the consumer surplus under a market allocation second.¹⁶ We formally present the assumptions below.

Assumption 3.1.

1. *In a period 0, the market design chooses the level of investment and pays the fixed costs*
2. *In a period 1, the market design chooses and offers the price schedule.*
3. *The market designer implements a market allocation.*¹⁷

¹⁶Recent 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.

¹⁷We omit the proof that the market allocation also implements the corresponding quantity schedule in the short-term allocation, as it is similar to the proof for the first-best.

We define the new objective for the market designer as follows.

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

s.t.

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

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

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

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).¹⁸ Then, it selects the investment level given the revenue made in the second period. 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.¹⁹

Proposition 3.2. *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 $s \in [0, s_1)$ then $t_1^{st}(\theta, s) = t_2^{st}(\theta, s) = 0$
- If $s \in [s_1, s_2)$ 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

¹⁸Without the constraint, the market designer would choose a null price.

¹⁹Formally: $\sum_i \mu_i \int_{\theta_i} d(0, \theta, s_1) dG_i(\theta) = k$, $\mu_1 \int_{\underline{\theta}_1}^{\bar{\theta}_1} d(0, \theta, s_2) dG_1(\theta) + \mu_2 \int_{\underline{\theta}_1}^{\bar{\theta}_2} d(0, \theta, s_2) dG_2(\theta) = k$, $\mu_1 \int_{\underline{\theta}_2}^{\bar{\theta}_1} d(0, \theta, s_3) dG_1(\theta) = k$

$$\mu_2 \int_{\underline{\theta}_2}^{\underline{\theta}_1} d(t_2^{st}(\theta, s), \theta, s) dG_2(\theta) + Q_0^1(s) = k$$

- If $s \in [s_2, s_3)$, 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

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

- For $s \in [s_3, \bar{s}]$ 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.

Proof. See Appendix 2 □

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 $\underline{\theta}_1$.²⁰ 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 drawn 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:

²⁰When $\underline{\theta}_2 = \underline{\theta}_1$ 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 ($s \in [0, s_1]$) 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, ($s \in [s_1, s_3]$), 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: $\sum_i d(t_i^{st}, \theta_i, s) = k$ can be rewritten to express consumer surplus at the capacity constraint only with respect to 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 3.1 for the discrete case.²¹ The red curves represent the higher consumer 1, and the blue curves represent the lower consumer 2. For the (expected) quantity

²¹The shape of the graphs would be the same when plotting the aggregate quantity for each category.

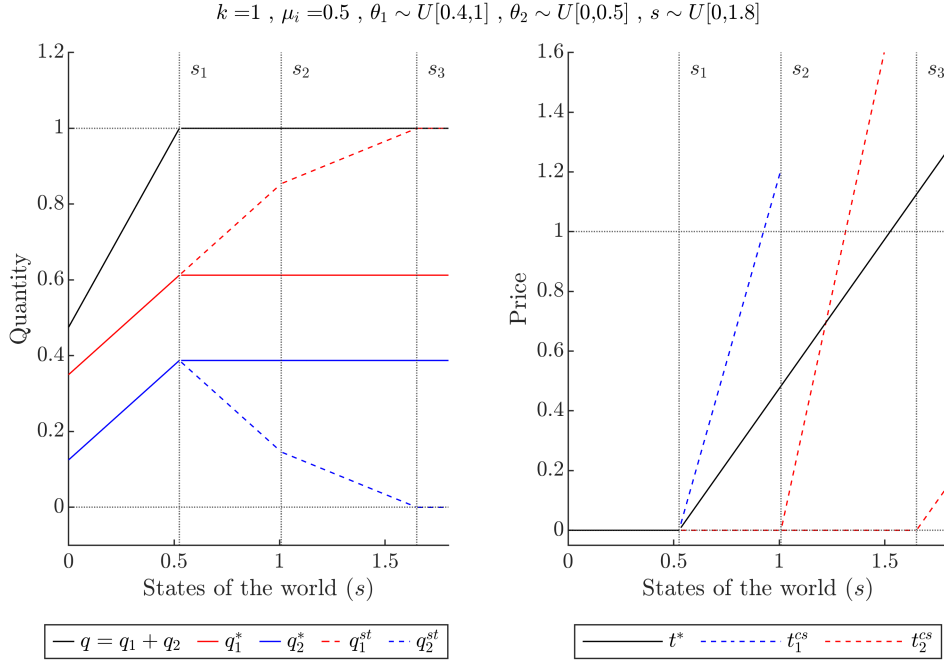


Figure 3.1: Quantity and price schedules under first-best and short-term consumer surplus mechanism. The price schedules on the plot represent only prices associated with a positive quantity. Short-term allocation favors consumers with a higher type and leads to individual price discrimination.

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 consumers. 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. The second plot shows this via individualized (expected) prices. For clarity, we have excluded prices used to exclude consumers. 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 $\bar{\theta}_2$ 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 3.1 a description of the optimal quantity (i.e., 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 and uniform model:

Corollary 3.1. *Under linear marginal utility assumption and uniform distribution.*

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

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

(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:*

$$\begin{aligned} \alpha_1^{st} &= \mu_1 \frac{s + \theta_1^{av}}{k} \quad \text{and} \quad \alpha_2^{st} = 1 - \alpha_1^{st} \quad \forall s \in (s_1, s_2] \\ \alpha_0^{st} &= \frac{1}{k} \left(\frac{\bar{\theta}_1 - \bar{\theta}_2}{\bar{\theta}_1 - \underline{\theta}_1} \right) \left(2s + \frac{\bar{\theta}_1 + \bar{\theta}_2}{4} \right) \quad \text{and} \quad \alpha_{-0}^{st} = 1 - \alpha_0^{st} \quad \forall s \in (s_2, s_3] \end{aligned}$$

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

Proof. See Appendix 3 □

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 3.1 describes the level of investment that maximizes the long-term consumer surplus given the short-term allocation schedule.

Lemma 3.1. *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 five local maxima.*

Proof. See Appendix 4 □

The value of each investment derives from the first-order condition of the consumer surplus under the price and quantity schedule described in Proposition 3.2. It significantly differs from the first-order condition of Proposition 3.1. We express the condition as follows:

$$\begin{aligned}
 r = & \int_{s_1}^{s_2} \underbrace{\mu_2 \int_{\theta_2}^{\bar{\theta}_1} \left(t_2^{st}(\theta, s) \frac{\partial t_2^{st}}{\partial k} \right) dG_2(\theta) dF(s)}_{\text{marg. utility from category 2}} + \int_{s_2}^{s_3} \sum_i \underbrace{\mu_i \int_{\theta_1}^{\bar{\theta}_2} \left(t_i^{st}(\theta, s) \frac{\partial t_i^{st}}{\partial k} \right) dG_i(\theta) dF(s)}_{\text{marg. utility from cat. 1 and 2}} \\
 & + \int_{s_3}^{\bar{s}} \underbrace{\mu_1 \int_{\bar{\theta}_2}^{\bar{\theta}_1} \left(t_1^{st}(\theta, s) \frac{\partial t_1^{st}}{\partial k} \right) dG_1(\theta) dF(s)}_{\text{marg. utility from category 1}}
 \end{aligned}$$

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 multiple possible maxima stems from the boundary conditions due to the capacity constraints. We therefore study the (seven) possible maxima for the different values of the s_j . We find that there are only six possible values.

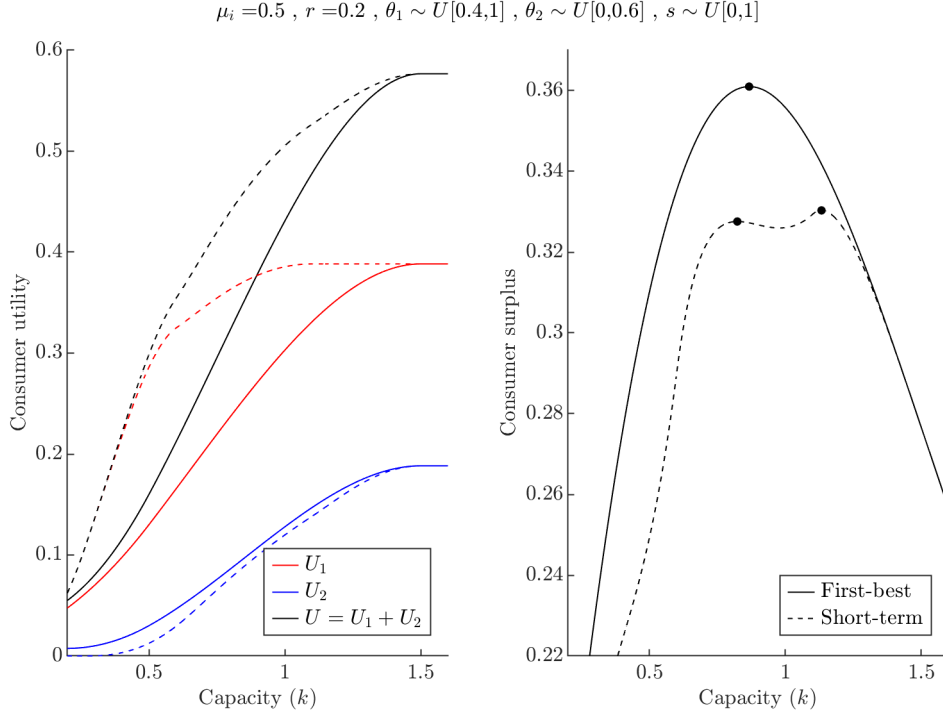


Figure 3.2: Expected consumer utility and surplus under the short-term (ST) and first-best (LT) allocation. The short-term maximizes consumer surplus without taking into account investment costs and favors bigger consumers. On the other, the reverse happens when considering investment costs under the first-best allocation.

Next, the result in Corollary 3.2 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 3.2. *The individualized price system that maximizes the short-term consumer surplus always leads to a lower long-term consumer surplus. Moreover, consumers with lower types are always worse off with the individualized price system.*

Proof. The proof is straightforward, as the quantity and price schedules differ from the first best. In that case, the quantity is lower, and prices are higher for consumers with lower types; their consumer surplus is also lower compared to the first-best.

□

We illustrate the results in Figure 3.2. 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 cover the fixed costs fully, the long-term consumer surplus is consistently lower. Note that the consumer surplus under the single-price system is concave, while under the multi-price system, it exhibits some non-concavity. The dots represent 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:

Assumption 3.2.

1. *The consumer's type is unknown to the market designer.*
2. *The market designer cannot extract any information from consumers.*
3. *The price schedule offered by the market designers is constrained to a unique price for every state of the world.*
4. *Given a set of prices t_i^r , the market designer implements a market allocation until the capacity is not binding, and 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 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 utility²² - (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.²³ In this chapter, 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.²⁴

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.²⁵ The first case allows us to focus the 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. The core idea of the model is that, without any information, a market designer has to choose a price t_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 is adjusted given the price t_i^r . When the capacity is binding, a random allocation is implemented because of incomplete information.²⁶ Following

²²See 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 \(2014\)](#).

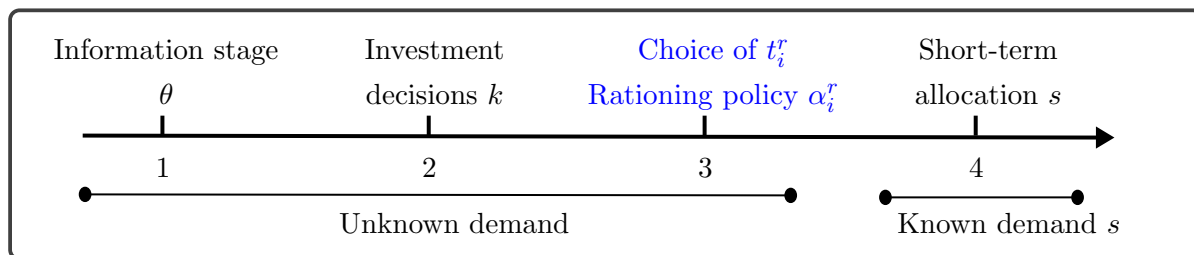
²³In current practice, political and technological constraints imply that the market designer (or any retailer) can only propose a finite number of schedules. See, for instance, [Astier \(2021\)](#) for theoretical and empirical implications for consumer surplus of allocation incompleteness.

²⁴An 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.

²⁵Political 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.

²⁶This 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

the framework description, we modify the actions taken by the market designer. The allocation proposal comprises (i) choosing t_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). Below, we provide 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 t^r is the price chosen by the market designer. The incomplete information set-up in this section has an important implication regarding quantity 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 $d(t^r, \theta, s)$ is the quantity a consumer asks given the price t^r . Let's define s_0^r the first states of the world when the capacity is binding when the price is t^r , that is:

$$\sum_i \mu_i \int_{\theta_i} d(t^r, \theta, s_0^r) dG_i(\theta) = k$$

For any $s \leq s_0^r$, 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 $t^r > 0$, the model does imply an inefficiency similar to the effect of market power. Due to the price being higher to marginal, it prevents some

it randomly to the consumers. We find 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.

Pareto-improving trade from happening. For any $s \geq s_0^r$, 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 a quantity equal to the investment level equally across consumers. 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.

$$\sum_i \mu_i \int_0^{s_0^r} \overbrace{\int_{\theta_i} U(d(t^r, \theta, s), \theta, s) dG_i(\theta) dF(s)}^{\text{off-peak utility}} + \sum_i \mu_i \int_{s_0^r}^{\bar{s}} \overbrace{\int_{\theta_i} U(k, \theta, s) dG_i(\theta) dF(s)}^{\text{on-peak utility}}$$

We now determine the best single-price policy given the framework. Compared to the previous analysis, the optimal price t^r depends not only on the first-best condition but on the revenue constraint. If it exists, the optimal value $t^r(k, t^r)$ satisfies the net revenue condition $R^k(k, t^r) = 0$ with:

$$R^k(k, t^r) := t^r \underbrace{\left(\sum_i \mu_i \int_0^{s_0^r} \int_{\theta_i} d(t^r, \theta, s) dG_i(\theta) dF(s) + \int_{s_0^r}^{\bar{s}} k dF(s) \right)}_{\text{Expected revenue}} - I(k)$$

This observation 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 the 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 3.3 the relation between the investment level and the optimal single-price

Proposition 3.3. *If an optimal single-price $t^r(k)$ exists, it increases in k .*

Proof. See Appendix 5

□

Proposition 3.3 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.²⁷

When choosing the price t^r , the market designer must trade off opposite effects. Indeed, increasing t^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 t^r . Note that the revenue is concave in t^r , meaning that the second-order effects are negative, limiting the market designer's ability to extract revenue from consumers.²⁸ Those effects can be shown by expressing the first derivative of the expected net revenue:

$$\frac{\partial R^r}{\partial t^r} = \underbrace{\int_0^{s_0^r(k)} \left(\overbrace{d_t t^r}^{\text{price effect -}} + \sum_i \mu_i \int_{\theta} \overbrace{d(t^r, \theta, s) dG_i(\theta)}^{\text{quantity effect +}} \right) dF(s)}_{\text{off-peak marg. revenue}} + \underbrace{\int_{s_0^r(k)}^{\bar{s}} \overbrace{k}^{\text{qtt. effect +}} dF(s)}_{\text{on-peak marg. revenue}}$$

With $d_t = \frac{\partial d}{\partial t}$ the derivative of the demand function with respect to prices. Calculation shows that $\frac{\partial s_0^r}{\partial t^r} > 0$, as a higher price implies that consumers decrease their consumption, and the capacity is binding less often. The proof relies on the observation that the second derivative is always negative. Hence, the revenue is concave in t^r . This comes from the (expected) marginal revenue of the off-peak revenue, and the marginal revenue effect cancels each other at $s = s_0^r(k)$, leaving the (negative) price effect. Next, we show how k modifies the expected revenue. An increase in the investment level leads to more investment costs and an increase in the quantity during on-peak periods. In other terms, the gain in on-peak periods cannot compensate for the loss due to the investment costs. Then, we use the fact that $R^r(k)$ is concave in t^r , and we study its behavior at the limit case such that the value k implies that the capacity always binds (i.e., $s_0^r(k) = 0$). In that case, we have $t^r = r$. Moreover, we also have at this limit: $\frac{\partial R}{\partial t^r} > 0$, implying that the revenue is increasing at the limit in t^r . If the function is concave, there could be

²⁷From a policy perspective, the market designer never wants to lower price so as to increase the consumption during off-peak.

²⁸Increasing t^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.

at most two potential values for the optimal value of t^r . However, the 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. As $R^r(k)$ is always decreasing in k , the solution of $R^r(k)$ is also increasing with k .

4.2 Category-price policy

We extend the previous findings by assuming that the market designer imperfectly discriminates between consumers and implements 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. We find that the market 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 given the previous relation. Solving using the first-order condition leads to a capacity share for each consumer belonging to a group i equals to the allocation of Corollary 3.1 under the first best allocation.

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

$$\begin{aligned} \max_k \max_{t_i^r} \quad & CS^r(k, t_1^r, t_2^r) = \sum_i \mu_i \int_s \int_{\theta_i} (U(q_i^r(\theta, s), \theta, s) - t_i^r q_i^r(\theta, s)) dG_i(\theta) \quad dF(s) \\ \text{s.t.} \quad & I(k) \leq \sum_i \mu_i \int_s \int_{\theta_i} t_i^r q_i^r(\theta, s) dF(s), \end{aligned} \quad (\text{R})$$

We drop the capacity constraint as the rationing policy defines it implicitly. To see this, we can redefine the state of the world s_1^r when the capacity starts binding:

$$\sum_i \mu_i \int_{\theta_i} d(t_i^r, \theta, s_1^r) dG_i(\theta) = k$$

Then, the quantity $q_i^r(\theta, s)$ allocated in the market for each consumer of category i is equal to $d(t_i^r, \theta, s)$ when $s < s_1^r$ and $\alpha_i^r k = k + \mu_j(\theta_i^{av} - \theta_j^{av})$ when $s \geq s_1^r$. 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

$$\mathcal{L}^r(k, t_1^r, t_2^r, \gamma^r) = CS^r(k, t_1^r, t_2^r) + \gamma^r R^r(k, t_1^r, t_2^r)$$

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 revenue constraint, and $R^r(k)$ the revenue 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:

$$\frac{\partial t_i^r}{\partial k} = (CS_{ik} + \rho_i CS_{jk} + (R_i - \rho_i R_j)\gamma_k^r + (R_{ik} - \rho_i R_{jk})\gamma^r) \frac{-\mathcal{L}_{jj}}{H^r} \quad (3.1)$$

$$\gamma_k^r = -\frac{1}{bH^r} \left(\sum_i R_i (L_{ij} L_{jk} - L_{jj} L_{ik}) + R_k H^r \right)$$

With $\rho_i = \frac{L_{ij}}{L_{jj}}$. $H^r = \mathcal{L}_{11}\mathcal{L}_{22} - \mathcal{L}_{12}\mathcal{L}_{21}$ 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 t_i^r on the consumer surplus. We summarize the findings in the following proposition.

Proposition 3.4. *Suppose that category 1 consumers are of higher types than category 2 consumers and that the investment cost is not too high then:*

- $t_1^r(k)$ is increasing with k
- $t_2^r(k)$ is first decreasing, then increasing with k .

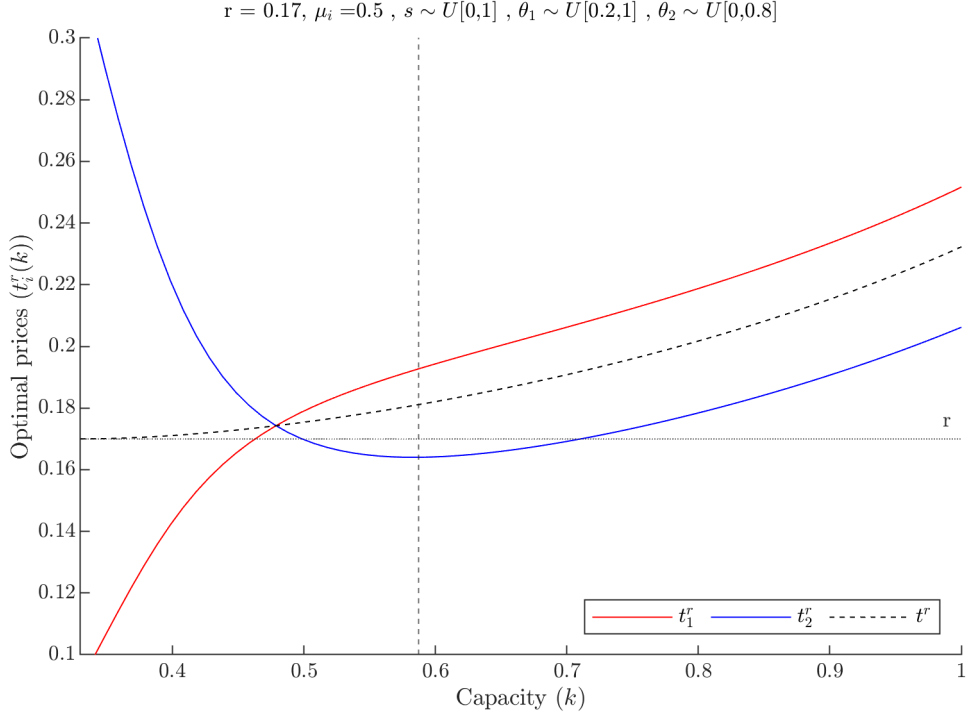


Figure 3.3: Optimal prices under the category-price policy with respect to the investment

Proof. In Appendix 6, we provide the formal proof and the condition under which the results hold. It relies on two Lemmas that ensure that a minimum exists. We also provide a more technical discussion on the rationales for this proposition. \square

Figure 3.3 illustrates the results. The red curve shows $t_1^r(k)$, the blue curve shows $t_2^r(k)$, and the black dashed curve shows the optimal single price $t^r(k)$ 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.

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 revenue constraint. The non-monotonocities 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 effect for higher

values of k . We summarize this tension between consumer and revenue effects in the following equation, which is a more detailed expression of Equation 3.1.

$$\frac{\partial t_i^r}{\partial k} = \frac{\overbrace{-\mathcal{L}_{jj}}^{\leq 0}}{H^r} \left(\underbrace{CS_{ik} + \rho_i CS_{jk} + (R_i - \rho_i R_j) \sum_i (CS_{ik} - \rho_i CS_{jk}) \frac{R_i}{L_{jj} b H^r}}_{\text{effect of } k \text{ on CS with holding } R \text{ fixed}} + \underbrace{\gamma^r \left(R_{ik} - \rho_i R_{jk} + (R_i - \rho_i R_j) \sum_i (R_{ik} - \rho_i R_{jk}) \frac{R_i}{L_{jj} b H^r} \right)}_{\text{effect of } k \text{ on } R} - R_k \frac{H^r}{b H^r} \right)$$

The first step for understanding the results lies in how those opposite effects change with the level of investment. As shown in the previous section, this 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 revenue constraint behaves shows a convex effect with respect to k . It implies that an increase of k leads to the revenue constraint shifting at an increasing rate.²⁹ The switch between the decreasing and increasing parts is associated with the consumer surplus effect dominating the revenue effect first. As k increases, the respective concavity and convexity of the functions lead to the revenue effect dominating the surplus effect.

The increasing prices on the right part of Figure 3.3 can be understood through the results of Proposition 3.3. The ranking between the category prices stems from the preference for discriminating bigger consumers. Increasing t_1^r generates more revenue as they consume, on average, more.

Let us turn towards the left part of Figure 3.3. From a consumer surplus perspective, and as illustrated in Proposition 3.2, the market designer prefers (i) discrimination and (ii) favoring the consumers from the category of the highest type. Preferring discrimi-

²⁹The pure revenue effect of increasing prices is found in the previous section. The intuition is that when k increases, the marginal gain from an increase of capacity during on-peak periods is offset by the increase in investment costs and by the decrease of on-peak periods. In the Appendix, we formally demonstrate that t^r is convex with respect to k .

nation 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. Next, we show why the asymmetry between the two consumers decreases for a higher value of k . Indeed, we find that this is not fully due to the revenue effect. For the sake of clarity, let's assume the revenue does not depend on k . We study the impact of k on the indifference curve of the market planner with respect to the prices t_1^r and t_2^r . We define the marginal rate of substitution between the two prices:

$$MRS_{i \rightarrow j}(k) = \frac{CS_i^r}{CS_j^r}$$

It implies that the MRS changes with respect to k as follows:

$$\frac{\partial MRS_{i \rightarrow j}}{\partial k} = \frac{CS_{ik}^r CS_j^r - CS_{jk}^r CS_i^r}{CS_j^{r2}}$$

Therefore, the decreasing right part of prices can be explained by having $\frac{\partial MRS_{i \rightarrow j}}{\partial k} < 0$. As the level of investment increases, and to keep the same level of consumer surplus, a decrease in t_i^r should lead to a relatively smaller increase of t_j^r . We provide in the Appendix 6 more technical details of this discussion. The economic interpretation of the effect of k can be understood as follows. As k increases, it negatively impacts the (negative) marginal effect of prices on consumer surplus. Because consumers from the bigger category have a higher marginal consumer surplus with respect to price, the negative marginal effect of prices of k is bigger than for the consumers from the smaller category. In other words, as k increases, CS_i decreases faster than CS_j , which implies a negative effect on the MRS. This can be explained by the fact that as k increases, it lowers the occurrences of on-peak periods, which makes consumers more exposed to the negative price effect on the quantity during off-peak periods. This, in turn, incites the market design to have lower prices, which is attained by lowering discrimination.

We illustrate this intuition in Figure 3.4. We represent a fixed revenue constraint as well as two indifference curves with respect to t_1^r and t_2^r for two values of k . As k increases, we show that the indifference curves tend to decrease in the sense that their marginal

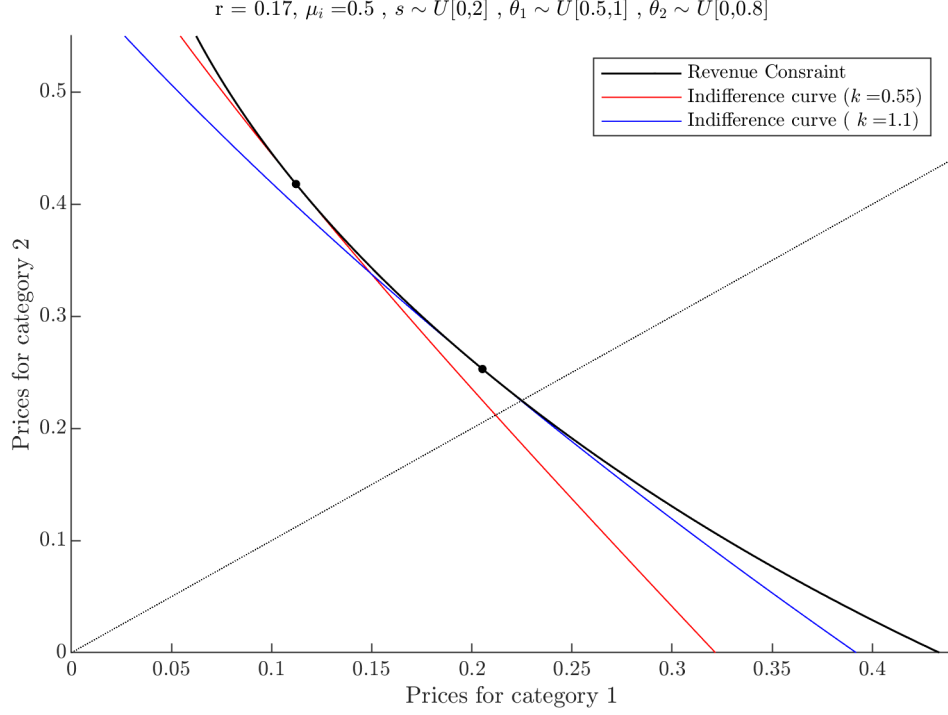


Figure 3.4: Illustration of the decreasing part of t_2^r with respect to k

rate of substitution decreases with k . 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.

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 individual consumer surplus³⁰:

$$\underbrace{\int_{\alpha_i^r k}^{d(t_i^r, \theta, s_1^r)} u(q, \theta, s_1^r) dq}_{\Delta \text{ in quantity btw. off-peak and on-peak}} + \underbrace{\int_{s_1^r}^{\bar{s}} (u(\alpha_i^r k, \theta, s) - t_i^r) dF(s)}_{+ \text{ in on-peak CS}}$$

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

³⁰We use the individual surplus for clarity. The aggregate surplus exhibits similar behavior.

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 does not cancel out. For the revenue, the derivative can be expressed as follows:

$$\underbrace{\sum_i \mu_i \int_{\theta_i}^{t_i^r} \left(d(t_i^r, \theta, s_1^r) - \alpha_i^r k \right) dG_i(\theta)}_{\Delta \text{ in quantity btw. off-peak and on-peak}} + \underbrace{\int_{s_1^r}^{\bar{s}} \sum_i \mu_i t_i^r dF(s)}_{+ \text{ in on-peak rev.}} - r$$

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 show that a level of investment exists that maximizes the expected aggregate consumer surplus, as the consumer surplus and the revenue are concave in k .

4.3 Individual welfare effect

We now compare the outcomes in terms of welfare given the optimal policy for the single-price and category-based prices. We focus the analysis on the individual change in the consumer surplus from a switch from a single-price-based policy to a category-based one.³¹ In this section and for clarity, we focus on the consumer surplus change for a given k and rather on the change based on difference maximizing investment level that might differ from the two policies.

We start by noting that the occurrence of on-peak situations can be higher for both policies depending on the values of prices: under the framework, this boils down to

³¹Note that the individual welfare changes with respect to k for each consumer are not relevant in this section as prices are based on category. Therefore, each consumer in its corresponding category exhibits similar surplus behavior. This motivates the study of this section of the evolution of the change of welfare from the two policies with respect to k .

comparing the average category prices (weighted by the share of consumers) to the single price $s_1^r(t_1^r, t_2^r) - s_0^r(t^r) = \sum_i \mu_i t_i^r - t^r$. For instance, if $t^r - \sum_i \mu_i t_i^r < 0$, then the prices under the category-based policy are relatively higher than under the single-price policy, implying that consumers reduce their consumption during off-peak under the former policy, and capacity binds more often. In any case, the change of individual surplus can be decomposed into three terms composed of two terms. We illustrate the change in the following equation for a consumer of type θ and belonging to category i (with $t^r > \sum_i \mu_i t_i^r$).

$$\begin{aligned} \Delta_i CS(\theta) = & \int_0^{s_1^r(t_1^r, t_2^r)} (\Delta_i U1 + \Delta_i R1) dF(s) + \int_{s_1^r(t_1^r, t_2^r)}^{s_0^r(t^r)} (\Delta_i U2 + \Delta_i R2) dF(s) \\ & + \int_{s_0^r(t^r)}^{\bar{s}} (\Delta_i U3 + \Delta_i R3) dF(s) \end{aligned}$$

ΔU captures the change in utility due to the price effect during off-peak periods on the quantity and the rationing policy under on-peak periods. For instance for a consumer belonging to category i : we have: $\Delta U_1 = U(d(t_i^r, \theta, s), \theta, s) - U(d(t^r, \theta, s), \theta, s)$. ΔR represents the change in the payment from each consumer. That is, for a consumer belonging to category i we have $\Delta R_1 = t^r d(t^r, \theta, s) - t_i^r d(t_i^r, \theta, s)$. Hence, the individual effect from discrimination is captured following the expected change in each term.

The framework prevents us from having any closed-form solution. Therefore, we concentrate the analysis on the main drivers of the welfare effect. Indeed, simulation shows that the second term in the previous equation is relatively small compared to the first and third terms. In other words, while there is a difference in terms of the occurrence of on-peak periods between the two policies, the delta between $s_1^r(t_1^r, t_2^r)$ and $s_0^r(t^r)$ is relatively small compared to other orders of magnitude. Hence, we discard it from the analysis.

We then distinguish two types of change: the ones that depend on the consumer type θ and the ones that are independent. In this environment, $\Delta_i U1$ and $\Delta_i R3$ are independent of θ for every consumer. This is due to the linear marginal utility for $\Delta_i U1$ and to the fact that the quantity during on-peak periods is allocated randomly in both policies for $\Delta_i R3$. In the case of $\Delta_i U1$, the sign of its value is solely dependent on $\Delta_i t^r = t^r - t_i^r$, as we have $\Delta_i U1 = \frac{1}{2}(t^{r2} - t_i^{r2})$. In the case of $\Delta_i R3$, the sign is given by $k\Delta_i t^r - \mu_{-i} t_i^r \Delta_i \theta^{av}$, with $\Delta_i \theta^{av}$ the difference between the average values :

$\Delta_i\theta^{av} = \theta_i^{av} - \theta_{-i}^{av}$.³² On the other hand, Δ_iR1 and Δ_iU3 are dependent on θ , but they can also be decomposed between a dependent and an independent part. In the case of Δ_iR1 , we rewrite the term such that $\Delta_iR1 = \Delta_iR1^\theta + \Delta_iR1^c$ with $\Delta_iR1^\theta = \theta\Delta_it^r$ and $\Delta_iR1^c = (s - \Delta_it^r)\Delta_it^r$, the first term depends on θ while the second does not. In the case of Δ_iU3 , we rewrite the term such that $\Delta_iU3 = \Delta_iU3^\theta + \Delta_iU3^c$ with $\Delta_iU3^\theta = \mu_{-i}\Delta_i\theta^{av}\theta$ and $\Delta_iU3^c = \mu_{-i}\Delta_i\theta^{av}(s - k - 0.5\mu_{-i}\Delta_i\theta^{av})$. We summarize the analysis of the individual change in the following claim. Without loss of generality, we assume that category 1 is of a higher type than category 2.

Claim 3.1. *When the category-based policy is implemented, individual consumer surplus for consumers of category 1 (resp. category 2) increases (resp. decreases) with low values of k . It decreases (resp. increases) with high values of k . Moreover, individual consumer surplus for smaller consumers from category 1 (resp. category 2) sustains smaller (resp. greater) gains and greater (resp. smaller) losses from the change in policy.*

The results in the Claim can be understood as a mirror effect from one category to another both in terms of magnitude and of who gets more impacted by the change of policy. The results are illustrated in Figure 3.5. The within-category comparison is clearly stated in $\Delta_iR1^\theta = \theta\Delta_it^r$ and $\Delta_iU3^\theta = \mu_{-i}\Delta_i\theta^{av}\theta$. Whatever the sign of those terms, the smaller the θ , the smaller the change. Then, we find that the main gains for the larger category come from Δ_iU3 , while it is the main source of losses for the smaller category. Switching from a single-price-based policy to a category-based policy increases the utility of the consumers from category 1 as they are allocated a greater share of quantity when the capacity is binding. However, as k increases, the capacity binds less in terms of occurrences. It implies that the main utility gain for the higher category (and the main loss for the smaller category) decreases as k increases. For Δ_iU1 and Δ_iR1 , the sign (mostly) depends on Δ_it^r . When k increases, we previously showed that there is a switch in terms of ranking between t_1^r and t_2^r , namely that for high values of k we have: $t_1^r > t^r > t_2^r$. In that case we have $\Delta_1U1 < 0$ and $\Delta_2U1 > 0$. For $\Delta_iR3 = k\Delta_it^r - \mu_{-i}t_i^r\Delta_i\theta^{av}$, the sign is ambiguous. However, for higher values of k , the change for the higher category is negative as we have $\Delta_1t^r < 0$ and $\Delta_1\theta^{av} > 0$. In other terms, the main losses for the higher type come from the difference in prices and are also mostly localized during off-periods. As k increases, the price differential increases, and it also increases the occurrences of off-peak periods. All in all, a higher level of investment means more losses for the higher category and more gains for the lower category. Finally, as the common net effect (Δ_iU1, Δ_iR3) is negative for the higher category and positive

³²Note that if we assume that a category i has bigger consumers then $\theta_i^{av} > \theta_{-i}^{av}$.

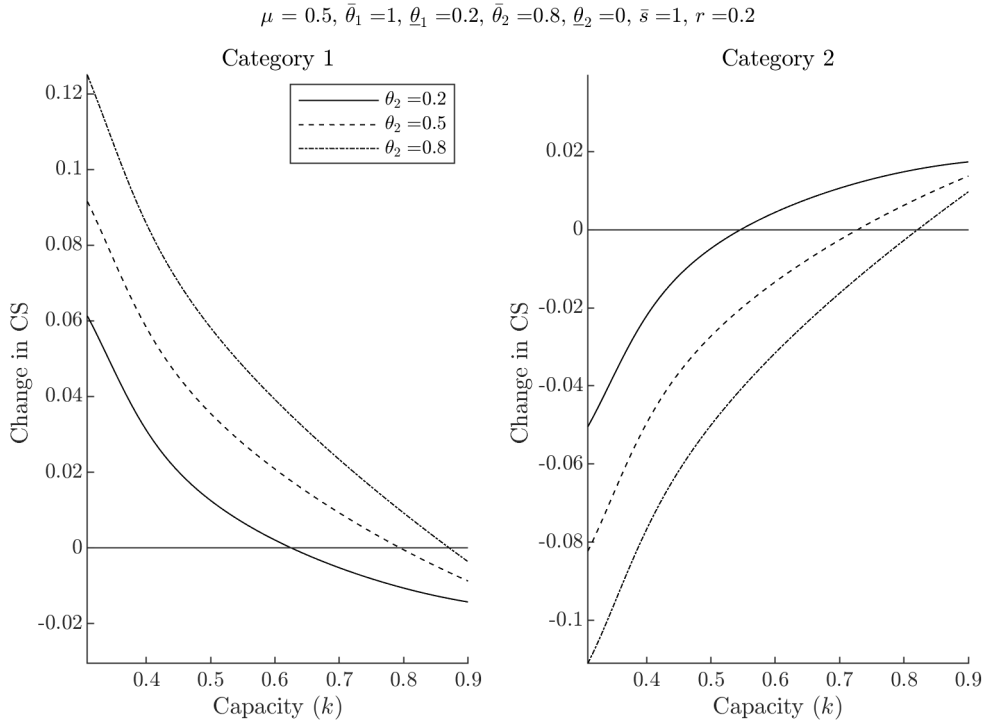


Figure 3.5: Change in consumer surplus with respect to investment level and consumer type. Consumers from the higher category exhibit a high gain from the switch to a category-based policy with a low level of investment. On the other hand, consumers from the lower favor a higher level of investment. Moreover, smaller consumers from the higher category sustain lower gains, which is the opposite for smaller consumers from the lower category.

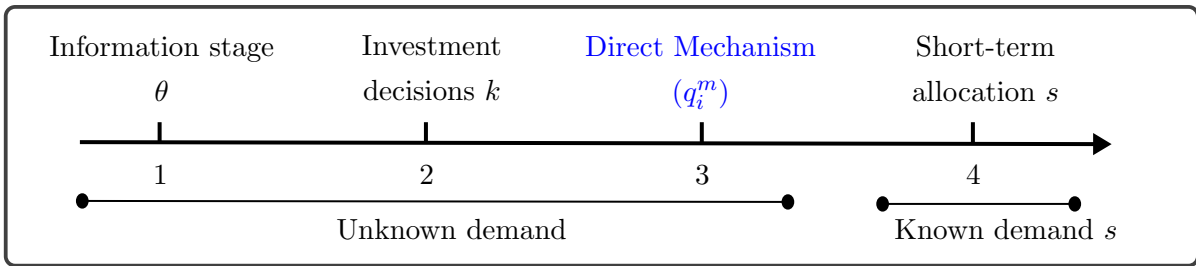
for the lower category, the smaller consumers are more negatively affected for the higher category and positively affected for the smaller category.

5 Incomplete Information - Mechanism Design

5.1 Optimal allocation

We extend the 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^m , leaving the market designer only with the quantity choice.



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

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

While the participation of every consumer implies that:

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

In that case, the mechanism design problem faced by the market designer is given by:

$$\begin{aligned}
\max_k \max_{\substack{t_i(\theta,s), \\ q_i(\theta,s)}} \quad & CS(k) = \sum_i \mu_i \int_s \int_{\theta_i} (U(q_i(\theta,s), \theta, s) - t_i(\theta,s)q_i(\theta,s)) dG_i(\theta) dF(s) \\
\text{s.t.} \quad & I(k) \leq \sum_i \mu_i \int_s \int_{\theta_i} (t_i(\theta,s)q_i(\theta,s)) dG_i(\theta) dF(s), \tag{R} \\
& \sum_i \mu_i \int_s \int_{\theta_i} q_i(\theta,s) dG_i(\theta) dF(s) \leq k, \tag{K} \\
& 0 \leq \int_s (U(q_i(\theta,s), \theta, s) - t_i(\theta,s)q_i(\theta,s)) dF(s), \tag{IR} \\
& \theta = \arg \max_{\hat{\theta}} \int_s (U(q_i(\hat{\theta},s), \theta, s) - t_i(\hat{\theta},s)q_i(\hat{\theta},s)) dF(s), \tag{IC}
\end{aligned}$$

We start describing the optimal allocation schedule given the new constraints and for a given level of investment k . Using the method developed in [Guesnerie and Laffont \(1984\)](#) and [Spulber \(1992a\)](#), we characterize the monetary transfer $t_i(\theta, s)$ in terms of quantity $q_i(\theta, s)$. As the 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 . 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 :

$$t_i(\theta, s)q_i(\theta, s) = U(q_i(\theta, s), \theta, s) - \int_{\hat{\theta}}^{\theta} q_i(\hat{\theta}, s) d\hat{\theta} + Cst$$

Where Cst is an arbitrary constant. The payoff equivalence uses the canonical approach of the Envelope Theorem (see [Milgrom and Segal \(2002\)](#)). Next, we use the approach from [Spulber \(1992a\)](#) to characterize a feasibility constraint that associates the revenue 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 $R - IR$:

$$\sum_i \mu_i \int_s \int_\theta (U(q_i(\theta, s), \theta, s) - \Gamma_i(\theta) \int_s q_i(\theta, s) dF(s)) dG_i(\theta) dF(s) - I(k) \geq 0, \quad (\text{R-IR})$$

With $\Gamma_i(\theta) = \frac{1-G_i(\theta)}{g_i(\theta)}$ the inverse hazard rate. Under the 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 $q_i^m(\theta, s)$.

Lemma 3.2. *Given IC and IR constraints, the optimal allocation $q_{i,l}^m(\theta, s)$ for a consumer of type θ from category i and for a given realization of s satisfies the following condition.*

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

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.

Proof. See Appendix 7 □

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.

$$\frac{\partial q_{i,4}^m}{\partial \theta} = \left(\frac{1}{2} - \frac{\partial u}{\partial \theta} \right) \left[1 - \frac{\zeta}{1 + \zeta} \Gamma_i'(\theta) \right] / \frac{\partial u}{\partial q}$$

Following the model specification, $\frac{\partial u}{\partial \theta} = 1$ and $\frac{\partial u}{\partial q} < 0$, the assumption concerning $\Gamma_i'(\theta) < 0$ implies that $\frac{\partial q_{i,4}^m}{\partial \theta} > 0$. Lemma 3.2 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 $\{q_{i,1}^m(\theta, s), q_{i,2}^m(\theta, s)\}$ 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, $\{q_{i,3}^m(\theta, s), q_{i,4}^m(\theta, s)\}$ is the set

of quantities such that the constraint is binding, implying that the optimal allocation needs to be distorted to cover both fixed costs and participation. Finally, note that the optimal allocation under the mechanism design approach when the $R - IR$ constraint is not binding is identical to the first-best allocation. Namely, when $\zeta = 0$, the condition in Lemma 3.2 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 (Spulber, 1992a). Therefore, we focus the rest of the analysis on the allocations $\{q_{i,3}^m(\theta, s), q_{i,4}^m(\theta, s)\}$.

5.2 Optimal allocation and investment level

We next analyze the threshold between the two sets of quantities. That is, we describe under which value of k the market designer faces a binding $R - IR$. We summarize the findings in the following proposition.

Proposition 3.5. *A unique value of k exists 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.*

Proof. See Appendix 8 □

Proposition 3.5 shows 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: $J_i(q, \theta, s) = u(q_{i,k}^m(\theta, s), \theta, s) - \Gamma_i(\theta)$, which is the marginal utility derived from the optimal allocation net of the information rent. Under the 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.

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

In other terms, the constraint starts binding when the aggregated marginal virtual revenue from the mechanism during the on-peak period equals the marginal investment.

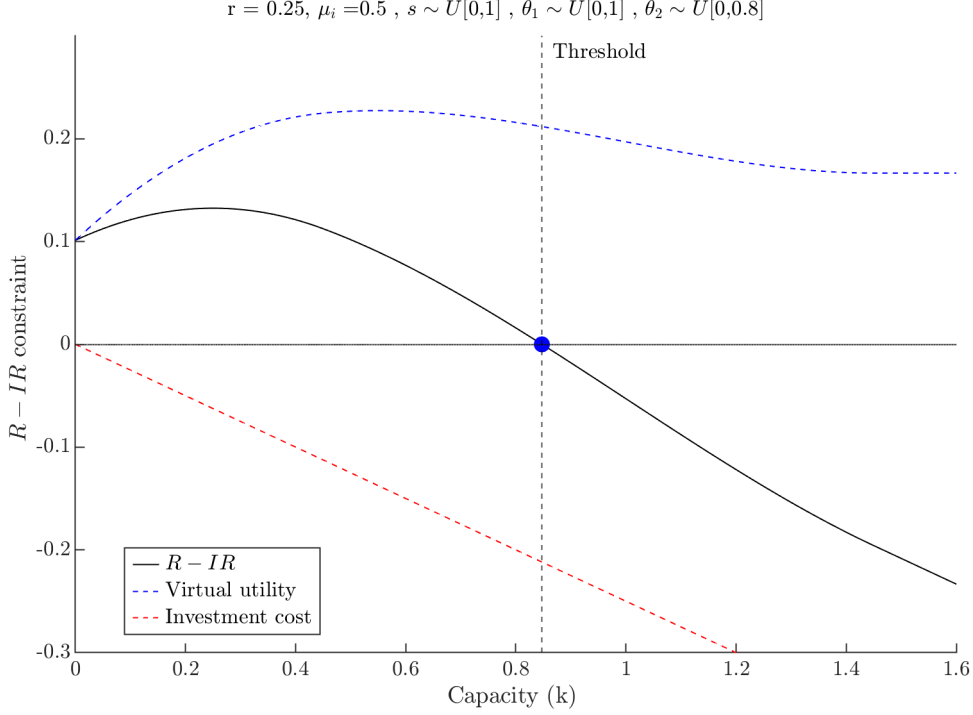


Figure 3.6: $R - IR$ constraints and its component with respect to the investment level.

Note that both the marginal virtual utility and the derivative of the quantity are, in that case, positive. Under the framework, $\frac{\partial q_{1,2}^m}{\partial k}$ is equal to 1, so to ensure that $\sum_i \mu_i \frac{\partial q_{1,2}^m}{\partial k} = 1$. 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 $\frac{\partial J_i}{\partial k} = -\frac{\partial q_{i,k}^m}{\partial k}$, 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 the findings in Figure 3.6. We plot the $R - IR$ constraint under the optimal allocation set $\{q_{i,1}^m, q_{i,2}^m\}$ 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, 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 the main findings in the following proposition. We define the term $\mathbb{A}(k)$ in the Appendix 9 such that we have $\frac{\partial \zeta}{\partial k} = \mathbb{A}(k)(1 + \zeta(k))$. It is independent of the states of the world and the consumer types. It represents the direct marginal effects of an additional level of investment (higher investment costs and more on-peak quantities) weighted by the marginal aggregate indirect effects (due to the change in off-peak and on-peak quantities at the optimum). Then:

Proposition 3.6. *The investment level directly affects the optimal allocation:*

- (Optimal off-peak) $q_{i,3}^m(\theta, s)$ is always decreasing with k for every values of k and for every type.
- (Optimal on-peak) for a consumer belonging to category i and of type θ , if:
 - * $J_i(q_{i,4}^m, \theta, s) > \mathbb{E}J_4 - \frac{1}{\mathbb{A}}$ then $q_{i,4}^m(\theta, s)$ is always increasing with k .
 - * $J_i(q_{i,4}^m, \theta, s) < \mathbb{E}J_4 - \frac{1}{\mathbb{A}}$ then $q_{i,4}^m(\theta, s)$ is always decreasing with k .

Proof. See Appendix 9. □

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:

$$\frac{\partial q_{i,3}^m}{\partial k} = J_i(q_{i,3}^m, \theta, s)\mathbb{A}(k)$$

Under the framework, we find that $\mathbb{A}(k) > 0$. The derivative originates from the first-order condition from Lemma 3.2: the marginal virtual utility at the optimal allocation during off-peak is always negative. Under the framework, and similarly to the previous sections, we know that the revenue 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 revenue effect implies that the optimal quantities for all consumers are decreasing. For the on-peak allocation, the initial derivative is expressed as follows:

$$\frac{\partial q_{i,4}^m}{\partial k} = \left[J_i(q_{i,4}^m, \theta, s) \frac{\partial \zeta}{\partial k} - \frac{\partial \varepsilon}{\partial k} \right] \frac{1}{1 + \zeta}$$

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 $J_i(q_{i,4}^m, \theta, s)$. 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:

$$\frac{\partial q_{i,4}^m}{\partial k} = [J_i(q_{i,4}^m, \theta, s) - \mathbb{E}J_4] \mathbb{A}(k) + 1 \quad (3.2)$$

With

$$\mathbb{E}J_4 = \sum_i \mu_i \int_{\theta_i} J_i(q_{i,4}^m, \theta, s) dG_i(\theta)$$

Being 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 not too low 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. On the other hand, the left part can be negative if the marginal virtual utility is not sufficiently close to the aggregate expression. We find that this is due to the existence of the cost associated with incomplete information.

Illustrative Example When expressing the marginal utility with respect to the optimal allocation, we find the following difference assuming two consumers of the same category, with one having type θ^1 and the other having θ^2 .

$$u(q_{i,4}^m(\theta^1, s), \theta^1, s) - u(q_{i,4}^m(\theta^2, s), \theta^2, s) = (\theta^2 - \theta^1) \frac{\zeta(k)}{1 + \zeta(k)}$$

That is, smaller consumers imply smaller marginal utility at the optimal allocation. We turn now to the difference in terms of information rent:

$$\Gamma_i(\theta^1) - \Gamma_i(\theta^2) = \theta^2 - \theta^1$$

The information rent is higher with smaller consumers. The two combines leads to :

$$J_i(q_{i,4}^m, \theta, s)(q_{i,4}^m(\theta^1, s), \theta^1, s) - J_i(q_{i,4}^m, \theta, s)(q_{i,4}^m(\theta^2, s), \theta^2, s) = (\theta^1 - \theta^2) \frac{1}{1 + \zeta(k)}$$

Consumers with higher types always lead to a higher marginal virtual surplus in the on-peak allocation. In other terms, the negative effect of an increase in capacity is directly due to the existence of the information rent. 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 $J_i(q_{i,4}^m, \theta, s) > 0$, 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 $J_i(q_{i,4}^m, \theta, s) - \mathbb{E}J_4$.

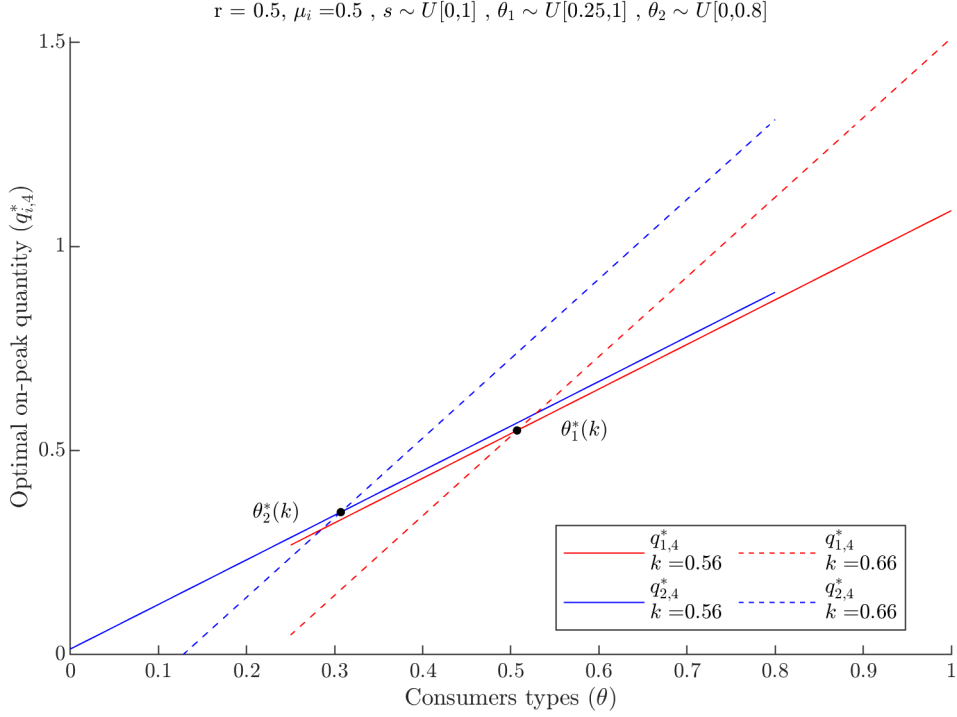


Figure 3.7: Optimal on-peak allocation with respect to the consumers' type θ , and threshold θ_i^m with respect to k

We illustrate the findings in Figure 3.7. It 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 constrain 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.³³ Therefore, it is less costly to discriminate the consumers of the lowest type negatively. We conclude the analysis of the relation between the quantity and the level of investments by describing

³³However, as we will discuss later, the category does not play a crucial role in the results of this section compared to section 4. The same behavior of the optimal allocation with respect to the investment level holds for both categories but with a different magnitude order.

the behavior of the type threshold $\theta_i^m(k)$ such that $\frac{\partial q_{i,4}^m}{\partial k} = 0$ which defines the type for which the derivative of the quantity is null. To do so, we denote Θ_i^m , the fraction threshold that gives the shares of consumers being negatively and positively impacted by the level of investment: $\Theta_i^m(k) = \mu_i \frac{\theta_i^m(k) - \theta_i}{\theta_i - \theta_i}$

Lemma 3.3. *If $\theta_1^m(k)$ and $\theta_2^m(k)$ exist, then (i) they are unique (ii) $\Theta_1^m(k) > \Theta_2^m(k)$ if $\mu_1 > \mu_2$ and $\frac{\mu_1}{\theta_1 - \theta_1} < \frac{\mu_2}{\theta_2 - \theta_2}$ and (iii) The thresholds follows the same behavior with respect to k for every category*

Proof. See Appendix 10. □

The first result ensures that there is a clear ranking between the different consumer types. This stems from the fact that the cross derivative of the quantity is positive with respect to the type. The second result in the Lemma implies that the bigger category usually exhibits a higher share of consumers having an increase in quantity when the level of investment increases compared to the smaller category. Under the symmetry of the number of consumers ($\mu_1 = \mu_2$), it is a sufficient condition to have a higher average type. The third result shows that the shape of the threshold is identical between categories. Namely, the mechanism design does not exhibit different behavior depending on the category compared to the results in section 4. Finally, numerical simulations show that the threshold exhibits a convex shape. It stems from the following observations. First, note that the marginal effect $\frac{\partial q_{i,4}^m}{\partial k}$ can be either increasing or decreasing with respect to k . That is, the individual quantity can either be convex or concave with respect to k . Indeed, the difference between the virtual marginal utility and the aggregate virtual utility can be expressed as follows:

$$J_i(q_{i,4}^m, \theta, s) - \mathbb{E}J_4 = (\theta - \bar{\theta}_i + \sum_i \mu_i(\bar{\theta}_i - \theta_i)) \frac{1}{1 + \zeta(k)}$$

As the revenue/participation constraint is more binding as k increases, the multiplier is also increasing ζ . Therefore, as k increases, the difference (in absolute terms) decreases. The economic intuition is that the change in virtual utility is the opposite of the quantity change. The marginal change with respect to k of the aggregate virtual utility is equal to -1 , as the aggregate quantity is equal to k . The marginal change of the individual virtual utility is $-\frac{\partial q_{i,4}^m}{\partial k}$. Hence, the positive effect of increasing k for consumers when the capacity is binding (right term in Equation 3.2) is offset by the change in the aggregate virtual

utility. This leaves the sign of the derivative of the difference defined as the opposite sign of the same difference. On the other hand, the marginal effect of k materialized by \mathbb{A} is increasing in k .³⁴ Therefore, the second-order effect of k on the quantity is given by the change of the following ratio $\frac{\mathbb{A}(k)}{1+\zeta(k)}$, namely, which part ratio is increasing at a higher rate. For low values of k , in $J_i(q_{i,4}^m, \theta, s) - \mathbb{E}J_4$ is lower than the increase in \mathbb{A} , to compensate the threshold is lower. On the other hand, as k increases, the constraint is binding more than its marginal effect. This explains the convex shape of the threshold.

5.3 Individual welfare effect

We turn now towards the welfare effects of the mechanism design approach in this framework. We start defining the individual consumer surplus given the optimal pricing and quantity functions of the mechanism design problem. Using the standard definition of the consumer surplus and the results from Equation 5.1, we have:

$$CS_i^m(k, \theta) = \int_0^{s_1^m} \int_{\theta_i}^{\theta} q_{i,3}^m(\hat{\theta}, s) d\hat{\theta} dF(s) + \int_{s_1^m}^{\bar{s}} \int_{\theta_i}^{\theta} q_{i,4}^m(\hat{\theta}, s) d\hat{\theta} dF(s) \quad (3.3)$$

Therefore, the relation between the level of investment and the gain/loss in welfare is given by the following derivative:

$$\frac{\partial CS_i^m}{\partial k} = \int_0^{s_1^m} \int_{\theta_i}^{\theta} \frac{\partial q_{i,3}^m}{\partial k} d\hat{\theta} dF(s) + \int_{s_1^m}^{\bar{s}} \int_{\theta_i}^{\theta} \frac{\partial q_{i,4}^m}{\partial k} d\hat{\theta} dF(s)$$

We describe in the next Lemma the conditions under which only a certain group of consumers profits from an increase in the level of investment. To do so, we define the surplus threshold $\tilde{\theta}_i(k)$ such that $\frac{\partial CS_i^m}{\partial k} = 0$.

Lemma 3.4. (i) *There exists for each category a unique threshold $\tilde{\theta}_i(k)$. For consumers of a category i , if his type is below $\tilde{\theta}_i(k)$, an increase in the level of investment decreases its consumer surplus. If his type is above $\tilde{\theta}_i(k)$, its consumer surplus increases with k . If $\tilde{\theta}_i(k)$ exist, then (ii) $\tilde{\theta}_i(k) > \theta_i^m(k)$.*

³⁴Which implies that $\zeta(k)$ is convex in k . This echoes the convex effects of the revenue constraint in the previous section.

Proof. The proof relies on the same arguments as for Lemma 3.3. Note that the cross-derivative of the information rent in Equation 3.3 with respect to k and θ is composed of the expected cross-derivative of the optimal quantity. The cross derivative of the optimal allocation $q_{i,3}^m$ shows that it is equal to the derivative of $q_{i,4}^m$, which we proved is positive in Lemma 3.3. Hence, the cross derivative is positive, so there is a unique threshold. The second result is straightforward as, in the case of the information rent, the expected derivative is based on the integral of the optimal allocations over lower types, which by Lemma 3.3 are smaller. \square

In other words, an increase in the quantity during on-peak is not a sufficient condition for an increase in the consumer surplus, and only higher types may benefit from an increase in the level of investment. The result in Lemma 3.4 directly stems from the form of the consumer surplus given by the mechanism design approach. Namely, the consumer surplus is equal to the information rent given to the consumer by the market designer to behave truthfully. Therefore, the central interpretation of the result is that the expected information rent only increases with k for higher types. The main difference between the quantity and the welfare effect with respect to k boils down to the consideration for a given type θ to all the lower types.

We illustrate the findings in Figure 3.8. We represent the evolution of the (expected) derivative of the off-peak quantities (dashed line) and the on-peak quantities (solid line) with respect to the consumer type. The components of $\frac{\partial CS_i^m}{\partial k}$ are represented by the integrals of the different functions. That is, the red zone corresponds to the off-peak first term, and the blues zone corresponds to the on-peak term. As expected, $\frac{\partial q_{i,3}^m}{\partial k} < 0$ for any values of θ and k . Note also that both $\frac{\partial q_{i,3}^m(\hat{\theta}, s)}{\partial k}$ and $\frac{\partial q_{i,4}^m}{\partial k}$ are increasing in θ . In the first panel, we have represented the case of a consumer of a lower type, such that the information rent is negatively affected by the level of investment. This is due to the fact that the derivatives of the quantities for both off-peak and on-peak periods are always negative. For the second panel, $\theta = \hat{\theta}$ such that its consumer surplus derivative is null at this level of investment. Some of the lower types generate sufficient positive derivatives such that they perfectly compensate for the negative effect from the smallest consumers and the off-peak periods. The last plot illustrates the case of a big consumer having a positive derivative.

The second result implies that consumers who observe an increase in the consumer surplus whenever the level of investment increases also have an increase in their level of investment. This is directly related to the fact that quantities and their derivative

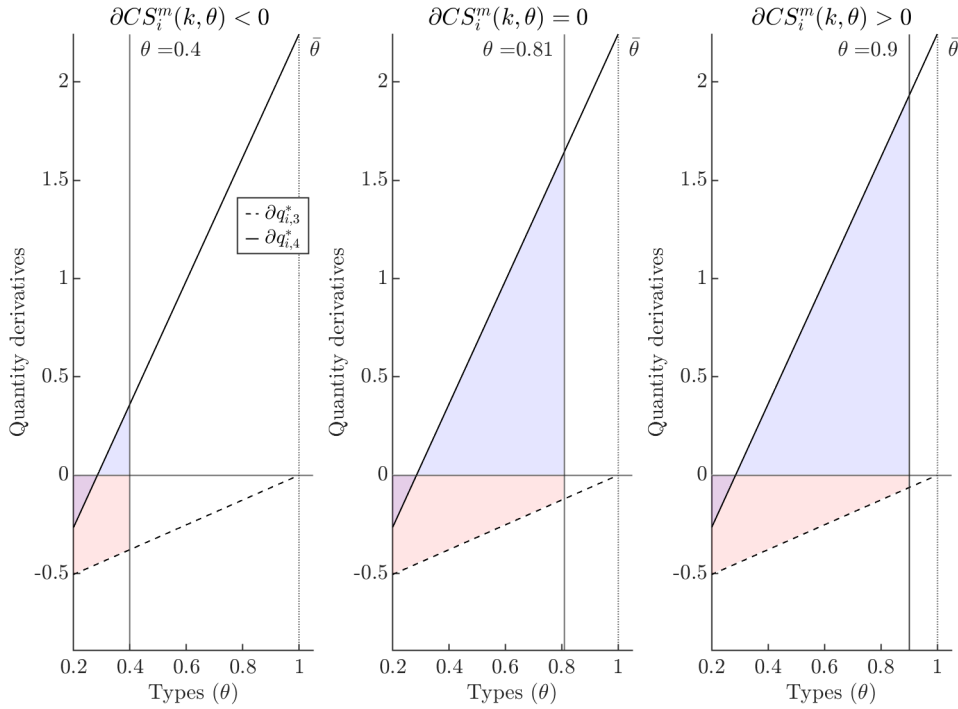


Figure 3.8: Components of the marginal individual consumer surplus for a given level of k with respect to consumers types θ

with respect to k are increasing in the type θ .³⁵ The third result shows that the surplus threshold θ_i^m exhibits similar behavior that the quantity $\tilde{\theta}_i$ with respect to k . Especially one can infer from the condition in Lemma 3.4 that if $\frac{\partial \tilde{\theta}_i}{\partial k} > 0$, then we necessarily have $\frac{\partial \theta_i^m}{\partial k} > 0$, the main difference is that the surplus threshold becomes increasing in k before the quantity threshold. In other terms, beyond a certain level of investment, an increase in k implies a decrease in the share of consumers having a positive marginal surplus. Those consumers excluded are always the smaller consumers before the increase. The economic interpretation of the decrease in the share of consumers positively impacted by k has two origins. (i) An occurrence effect, that is, an increase in the level of investment reduces the occurrence of on-peak periods in favor of off-peak periods. This implies that even if the on-peak quantity increases in k , as on-peak periods occur less often, its positive effect on the consumer surplus decreases. This is materialized by the term left term of the condition in Lemma 3.4. (ii) A second-order effect similar closely related to the one leading to an increase in $\theta_i^m(k)$: for sufficiently high values of k , the quantity is concave in k , implying that the derivative is decreasing in k , therefore reducing the

³⁵A negative derivative for a given type always implies that all derivatives of lower type are also negative

positive effect. This is materialized by the right term in Lemma 3.4. We illustrate those findings in Figure 3.9. The first plot shows how the different components of the information rent change with respect to k . As in Figure 3.8, the dashed lines correspond to the off-peak periods. The red lines correspond to the components with higher levels of investment. As k increases, the expected (negative) marginal derivative increases due to the higher occurrence of off-peak periods. The downward shift of the dashed lines translates this. The effect on the on-peak parts is a priori ambiguous, as illustrated by the rotation of the solid lines: increasing k implies less negative marginal quantities from smaller consumers, hence the positive blue part, but also less positive marginal quantities from bigger consumers, hence the negative red part. In the end, we find that the net effect leads to the results in Lemma 3.4. The second plot shows how the share of consumers having an increase or a decrease in quantity and surplus evolve with k . The solid line corresponds to the quantity threshold $\theta_i^m(k)$ and the dashed line to the surplus threshold $\hat{\theta}_i(k)$. As expected, consumers who receive a higher surplus when k increases are located in the higher part of the distribution. The blue zone corresponds to those consumers. The green zone describes the consumers having an increase in their quantity but not in their surplus, and the red zone corresponds to the consumer for which both quantity and surplus decrease with k .

We end the section with the aggregate effect of the investment level in this incomplete information framework. The second-best investment level is described in the following corollary.

Corollary 3.3. *The market designer chooses the second-best investment level k^m that maximizes the expected consumer surplus. If it exists, it solves the following first-order condition:*

$$\begin{aligned} \sum_i \mu_i \left[\overbrace{\int_0^{s_1^m} \int_{\theta_i} \int_{\theta_i}^{\theta} \frac{\partial q_{i,3}^m}{\partial k} d\hat{\theta} dG_i(\theta) dF(s)}^{\text{off-peak marginal loss}} + \overbrace{\int_{s_1^m}^{\bar{s}} \int_{\theta_i}^{\hat{\theta}_i} \int_{\theta_i}^{\theta} \frac{\partial q_{i,4}^m}{\partial k} d\hat{\theta} dG_i(\theta) dF(s)}^{\text{on-peak marginal loss}} \right] \\ = \underbrace{\sum_i \mu_i \int_{s_1^m}^{\bar{s}} \int_{\theta_i}^{\hat{\theta}_i} \int_{\theta_i}^{\theta} \frac{\partial q_{i,4}^m}{\partial k} d\hat{\theta} dG_i(\theta) dF(s)}_{\text{on-peak marginal gain}} \end{aligned}$$

Proof. The proof is straightforward and is given by deriving the expected information rent in Equation 3.3. □

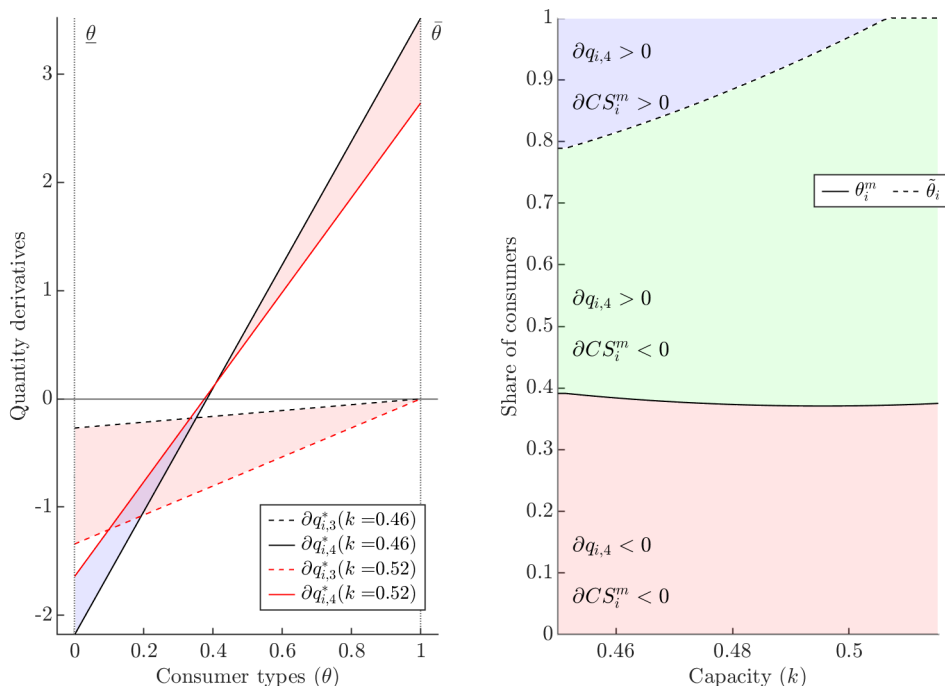


Figure 3.9: Components of the marginal individual consumer surplus for a given level of k with respect to consumers types θ

The mechanism design approach adds a new dimension to the choice of the optimal investment level compared to the complete information framework and the second section. In addition to covering the investment costs, the market designer faces a new tension associated with the participation and incentive compatibility constraint. Due to private information, the market designer has to reward consumers to behave truthfully. In the end, individual consumer surplus is equal to the information rent. Therefore, the objective function of the market design when choosing the level of investment is to maximize the aggregate information rent. The first step in the analysis is to characterize how individual quantities relate to the investment level. We find that all consumers receive less during off-peak periods, and the smaller consumers observe a decrease in their allocated quantity during on-peak periods. Compared to the previous section, the same effect is found across the two categories. Following the standard approach adapted to the setting, the expected information rent of a given type is composed of the integral of the quantity allowed to every smaller consumer below the type. Consequently, to be positively affected by the increase in the level of investment, a consumer needs to have a sufficiently high type to offset the negative effects of off-peak periods and from smaller consumers of on-peak periods. In the end, the level of investment is found to be such

that the marginal decrease in expected information rent from smaller consumers equals the marginal increase in expected information rent from bigger consumers.

6 Conclusion

This chapter 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 chapter 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 chapter provides policy and technical results by adding additional constraints to the initial framework. 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 chapter 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. In an initial complete information framework, we find the market design already prefers discriminating against smaller consumers when prices are only used to ration quantity. We then introduce incomplete information in a contemporaneous setting. That is, we model a market designer who cannot set prices for every state of the world. This framework shows that the market designer faces a tradeoff between generating a higher surplus by discriminating against smaller consumers and generating higher revenue to finance investment by discriminating against bigger consumers. In the end, smaller consumers from the smaller category favor a high level of investment, and bigger consumers from the bigger category favor a low investment level. In the last section, we adopt a standard mechanism design approach to consider incentive and participation constraints. The central highlight is the specific relation between the level of investment and the individual and aggregate information rent. We find that only bigger consumers can face an increase both in quantity and surplus when the level of investment is high as they are the only consumers to face an increasing information rent with respect to the capacity level.

Finally, we plan to extend the 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. The 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 the chapter shows, 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 the framework could highlight a new trade-off.

Bibliography

- Akbarpour, M., Budish, E. B., Dworzak, P., and Kominers, S. D. (2023a). An economic framework for vaccine prioritization. *Available at SSRN 3846931*.
- Akbarpour, M., Dworzak, 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. *CEPR Discussion Paper No. DP17200*.
- 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.
- Crampes, C., Ladoux, N., and Lozachmeur, J.-M. (2023). Pricing energy consumption and residential energy-efficiency investment: An optimal tax approach. *Journal of Public Economic Theory*.
- 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. Technical report, Mimeo.
- 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.

General Conclusion

Motivations

Electricity is a paragon within the realm of economic science. It is one of, if not the central commodity, that has enabled the development of modern societies over the last century. Many public institutions have stressed that access to electricity is a central concern for the growth of developing countries. This is particularly illustrated by numerous theses currently being investigated by my colleagues ([Saliou Barry, 2022](#); [Ly, 2023](#); [Le Picard, 2023](#)).

Electricity markets exhibit many of the problems analyzed by modern economics. From a microeconomic and public economics point of view, several decades of studies have underlined its numerous market failures: externalities due to its production, transport, and consumption; the public good nature of its related infrastructure; its incompleteness, especially with respect to risk management; its specificity, that makes it an ideal playground for exerting market power. The existence of a certain number of physical constraints also poses a number of challenges for economists, who must adapt their analysis to the physical laws governing the entire electricity value chain. Moreover, the recent COVID-19 crisis and the war in Ukraine have also highlighted the link between electricity and macroeconomic issues, especially inflation.

It is in this context that my thesis seeks to answer a number of policy questions under the lens of economics theory. In particular, I have chosen to address the link between the functioning of electricity markets and investment decisions. In other words, how do we ensure that the institutions set up to produce, exchange, and consume electricity provide sufficient incentive to invest efficiently? While investment issues are not specific to electricity, the current lack of storage capacity on an adequate scale and at a relatively low cost means that any shortfall in the availability of investments in relation to consumption levels entails significant risk. In the best-case scenario, these risks translate into involuntary rationing of demand, also known as a rolling blackout, and, in the worst-case scenario, into total system collapse. This issue is all the more critical as one of the main objectives of public policy in the electricity sector is the massive development of renewable energy.

As will be emphasized in this general conclusion, I have addressed this problem in a specific way. The central idea of this thesis is to essentialize and conceptualize the particularities of the link between investment and production decisions. This link is well described in the concept of allocation externalities used in economic literature: individual investment decisions depend on electricity market equilibria, and equilibria depend on

investments. Ultimately, the interdependence between investment and equilibria will determine welfare. Once this link had been clarified, I set out to answer the following market design question: what are the organizational rules that will maximize efficiency given this particular framework? In other words, I attempt to answer the practical question of how to provide the best incentives to invest so that the level reached is the desired one but also so that the way it is reached does not generate too many indirect costs.

The starting of this thesis is based on the following observation. The electricity sector is made up of a succession of trials and errors on how to organize these markets efficiently. It provides a vast playground for analyzing institution-building policies both in a positive and normative way. I take as its starting point a specific approach: that of capacity markets. They are characterized by producers offering a promise to be available in the future in return for compensation. Rather than asking whether capacity markets are more efficient than other forms of organization, my thesis looks at how they can be set up most optimally.

While the empirical contributions to capacity markets have remained limited, electricity economists used a wide diversity of modeling tools ranging from complex engineering system representation to classic analytical and stylized theoretical models, each with its pros and cons. My first years as a PhD student focused on an in-depth analysis of this quantitative literature on capacity markets. However, I realized that in order to address specific market design issues that were very practical at first sight, it was necessary to return to the theoretical fundamentals of electricity markets. As a result, my modeling choices focused on a very simplified representation of the industrial environment in order to study what it means to (i) make an investment decision and (ii) participate in markets, whether from a long-term point of view with capacity markets or in short-term wholesale markets. Therefore, each chapter of this thesis aims to reduce the description of reality to its essentials to better understand the mechanisms behind which market design choices have positive or negative effects in terms of welfare.

General Takeaway

Each chapter of this thesis highlights well-known issues in market design and industrial economics. There are always some fundamental trade-offs that need to be understood, and market rules always have unidentified consequences, both good and bad. The central contribution of this thesis is to focus on a very specific topic: in an environment as particular as electricity, a sufficient level of investment is crucial and often imposed by policymakers. However, investment decisions are made by private actors. So, how can we ensure that markets are designed to ensure the most virtuous encounter between private values and social needs in this context? By representing the behavior of actors through economics theory with a focus on modeling investment decisions and their link with the rest of the market, the thesis proposes to formalize a variety of positive effects as well as give new policy warnings with respect to how to design a market for investment.

Beyond the contribution that I will describe in more detail for each chapter at the end of this summary. This manuscript, I hope, paves the way for further research. To begin with, it is necessary to recall the crucial importance of the empirical approach in economics. Capacity markets have been in place for some twenty years around the world, and even if the data are not as robust as for other electricity markets, I think it's quite possible to dig in this direction. This extension of my research could involve both a classic econometric approach and the use of data to build counterfactuals on the basis of more engineering models, giving a more realistic picture of how a power system works. It would also allow measuring the true value of the trade-off highlighted in this thesis. As investments play a crucial role in the electricity markets, both due to their value and their cost, such an approach would benefit the policy debate around the acceptance of new technologies and regulations.

Another significant limitation of this work is the observations made in the introductory chapter. Indeed, I have described how the efficiency of a capacity market is measured not only in terms of the level of investment but also in terms of other closely related industrial decisions, such as market entry and exit decisions, as well as decisions to maintain a certain level of capacity. Moreover, in the context of the energy transition, it is crucial to ask the question of the composition of the energy mix. It would then be conceivable to extend the work presented in each chapter to take these other dynamics

into account. The aim of these extensions is not simply to add a degree of realism to the theoretical models but to understand how the behaviors and new trade-offs raised by these issues add to or conflict with the results found in this thesis. As an example, I'll take renewable energies, which, unlike more conventional means of production, have an uncertain level of capacity but a certain marginal cost.³⁶ I provide in the next sections a more detailed view of the main results and the possible extensions to each chapter.

Finally, I would like to emphasize that this thesis has a contribution beyond electricity markets. The stylized theoretical approach used in this thesis makes it possible to address broader issues, notably that of ensuring sufficient investment for a set of *essential goods*. Those goods are 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. Electricity is the textbook case, but other markets exhibit similar characteristics: transport with congestion or medical goods for which scarce supply can lead to contagion and congestion of medical facilities. Specifically, I am developing a research program that examines how we can implement and design institutions, such as markets, that efficiently address the demand and supply of these investments. This research stems from the fact that this is still an understudied topic from the academic side. Moreover, as policymakers and consumers place a high value on investment and production capacity, a significant social demand exists to highlight those questions. The rest of the general conclusion is composed of a summary of the three chapters.

³⁶See for instance [Fabra and Llobet \(2023\)](#) which show that auctions in this context might have a different outcome than the canonical models.

What is the cost of participating in a capacity market and the relation with contract characteristics?

Chapter Takeaway

As capacity market efficiency depends on the price signal it sends, policymakers should not overlook the choice of the capacity product characteristics and, notably, the availability period. In particular, we show in this chapter that when risk and flexibility are considered, the outcome can be different than under traditional modeling assumptions.

This research originates from a simple yet fundamental question: "How do prices emerge in capacity markets?". Capacity prices act as supplementary revenue for producers to encourage them to increase their available capacity. It is also an additional burden for consumers who have to pay for this increase in capacity. Therefore, knowing the right capacity price is crucial to measuring the efficiency of any capacity market. Economic theory has long stressed that prices may differ from their fundamentals. Indeed, every price is the consequence of an intersection between supply and demand. In capacity markets, this supply curve is the aggregation of the bids of different actors having different characteristics. Hence, there is always significant room for deviation from the efficient price.

This observation calls for a second question: "How do producers bid in capacity markets?". I answer it by stating that participation in a capacity market implies a specific marginal opportunity cost for the bidders, which requires a particular method of valuation. Namely, selling capacity entails the producer staying available over future periods determined in the sold contract. This offered promise generates specific gains and losses for the producer. In this case, they have to trade expected revenues over the procurement duration net of the fixed cost associated with the decision to stay available. The first main contribution lies in studying the bid as an option value associated with the possibility of closing temporarily but irreversibly to avoid fixed costs. This significantly differs from the standard approach of the net present value approach, where the producer only offers the expected opportunity cost. It notably allows us to recognize the managerial flexibility of an investment that can react to future market conditions.

Thus, I propose a novel approach to conceptualize capacity bids using real options theory, where the opportunity cost is represented as an option on the spread that drives the profitability of the plant. First, I define a bid in a one-period capacity market as a European Put Option. Then, I expand to a multi-period setting in which capacity bids can be evaluated as a modified Basket Option.

I use this novel approach to answer a concrete market design question: "Does the length of the contract sold on the capacity market matter?". An analysis of different implementations of capacity markets around the world revealed a wide variety of durations, ranging from a few weeks to several years. The critical assumption of this chapter is as follows: if the duration of the contract changes the expected profit associated with the promise to stay available, then it should change the marginal cost of participating in the capacity market. Hence, the supply curve changes, too, which ultimately impacts price formation. Therefore, the design of the product sold in a capacity market can define, in part, its efficiency.

The central results of this chapter are the implications of the change in contract duration on the bidding behavior in a capacity market, assuming that the marginal cost is analyzed through the lens of real-option theory. I find that it significantly differs from the net present value framework. First, bids are always higher under the real options framework, meaning that producers place a positive value on the possibility of closing to avoid some costs. Second, the drivers behind the bids have different effects on their value compared to the net present value framework. I provide comparative statics on the bid value and the difference between the two frameworks. I find that the length of the contract constantly increases the bid when using the real options theory while having an ambiguous effect on the net present value bid. The volatility in the wholesale market and the waiting time between the sale of the contract and the start of the availability are also analyzed. They both have ambiguous effects on the capacity bids depending on a set of conditions on the bid drivers. Finally, I find the reverse effect for the product design dimension with a higher bid with a longer contract duration than the sum of expected bids with shorter products. I test the results by calibrating the model to the French electricity system. I also use realized data for a CCGT (gas) power plant to simulate a bid in the capacity market and compare the outcomes with realized prices observed on the French capacity market. Based on both the theoretical framework and the numerical illustration, I finally provide a policy discussion for the design of capacity markets: (i) On the role of penalty in the capacity market imposed on producers who choose not to be

available during the procurement duration despite having sold a capacity product. (ii) On the difference between existing and new capacity participating in capacity markets.

The principle of this chapter is to lay the foundations for a better understanding of how producers offer their investment in capacity markets. In particular, by describing the relationship between their offers and the contract characteristics traded on the market, we gain a better view of what should emerge on capacity markets and what trade-offs exist in terms of market design. This chapter also opens a number of doors for future research. The first venue is to incorporate these results into a more systemic framework. In particular, one should reconcile the valuation approach of the marginal cost with a price formation process. In other words, one could model a set of heterogeneous producers and build supply curves using real options theory. Adding sequential dynamic decisions would also enable us to endogenize the set of values and, in particular, allow a finer analysis of investment decisions. A finer calibration of the model's parameters would also give a better comparison between what is expected from a market fundamentals point of view and the actual behavior of agents. A second approach consists of taking into account the demand in capacity markets. Indeed, the entire chapter focuses on the value of an investment for a producer and from the point of view of real options theory. But a question must be asked: what is the value of an available capacity for consumers? One could then carry out a similar study, analyzing the counterpart of this option value but for demand. The value to society of additional capacity is thus the conjunction between the demand and supply option value.

The indirect consequences of designing the demand-side

Chapter Takeaway

This chapter shows that the way the demand-side of capacity markets is designed can have an impact on the overall system. We notably prove that it can change how consumers buy electricity, which can have consequences on the consumer surplus and indirectly on investment decisions. Depending on the characteristics of the system, these indirect effects may have positive or negative consequences in terms of welfare.

The initial motivation for this chapter is to describe the trade-off between having a centralized or decentralized capacity market. In this thesis, a centralized market is characterized by having a single buyer, usually a regulated firm, such as a System Operator, procuring the total amount of capacity. On the other hand, a decentralized market design is based on the mandatory participation of different demand-side actors, mainly retailers, which buy themselves capacity in the face of their expected consumption. In between, hybrid market designs exist in which consumers and retailers indirectly participate in capacity markets. The degree to which consumers participate is, after that, called the demand-side design of capacity markets. This research stems from the fact that the capacity demand requires a regulatory intervention. While the supply emerges naturally in those markets, the public-good nature of investment during high-demand periods implies that consumers are unwilling to buy capacities in capacity markets. Hence, the regulator must define the demand function administratively so the market clears and provides producers' capacity prices.

This chapter aims to lay the foundation for understanding the pros and cons of the different options available to the regulator. In particular, I show the consequences of different options on a model representing capacity market, investment decisions, generation, and consumption equilibrium. Ultimately, the chapter describes the welfare of a power market as a function of different assumptions about market design, the behavior of players, and the environment. I focus on two interrelated questions that relate to (i) the cost allocation regime, that is, how a single buyer allocates the capacity price between capacity buyers and final consumers, and (ii) the degree to which the final consumers' realized demand is accounted in the market allocation design.

The methodology used to answer this question relies on the canonical peak load pricing theory for a homogeneous good characterized by time-varying demand, which describes the relationship between short-term production and long-term investment decisions. The model extends the literature by providing a novel analytical framework that includes a capacity market equilibrium in addition to investment and short-term decisions. The modeling proposition is central as any indirect effects generated by the capacity market can affect the expected revenue made by the producers and can indirectly be captured during the formation of the supply function in the capacity market.

The first market design regime studied is the canonical capacity market. I build on the previous literature, which relies on the assumption that a capacity market does not have any effect beyond increasing the investment level. This canonical regime is similar to having a cost allocation regime based on a lump-sum tax. In this case, even when considering the endogenous supply function in the capacity market, the equilibrium of the market design always restores the first-best optimum given the system inefficiencies. I then investigate the case in which the capacity price impacts consumers at the margin. Under this assumption, the regime similarly allocates the capacity price as a unitary tax. I show that the existence of the capacity market indirectly affects the wholesale market by redistributing the different states of the world when the capacity does not bind and bind and by lowering the consumer's surplus. Therefore, I demonstrate that the welfare outcome at the equilibrium under this regime is always lower than under the canonical regime. I then compare the two capacity cost allocation regimes by including inefficient rationing. When a price cap is reached, the investment availability becomes a public good as the demand becomes inelastic. Due to the impossibility of efficiently rationing consumers, they incur a significant welfare loss. Under this new assumption, I find that the indirect effect created by allocating the capacity price on a unitary basis is now ambiguous for social welfare.

I extend the analysis to implementing a regime where the regulator allocates the cost based on actual retailers' market shares. I first show how this design marginally affects the retailers who play 'à la Cournot' in the retail market. Then, I integrate the new equilibrium into the model with investment decisions and the capacity market. I find that this allocation creates an intermediary outcome between the unitary tax and the lump-sum tax. Finally, I analyze the case of a capacity market entirely based on the realized demand level. To do so, retailers are forced to cover the quantity sold on the retail market by buying directly on the capacity market, given a penalty system. I focus on how retailers' strategies can form an aggregated demand function in the capacity

market, and I analyze the optimal capacity bought by retailers in the capacity market. I find that such an approach for the demand function can provide the optimal level of investment under specific conditions.

This chapter builds a series of extensions to the canonical model representing investment and production decisions in electricity markets. It starts with the most simple case with a price cap and a single buyer. Then, I had more assumptions concerning inefficiencies and the actors' behavior. Rather than providing a clear comparison between a centralized and a decentralized market design, the chapter seeks to describe how we can interpret those designs in the first place. Namely, a centralized market is, before all, a question of cost allocation, and a decentralized market depends on the value an investment has for the demand-side. With those results in mind, I believe that there is now more precise room to conduct a deeper comparison. As underlined by some qualitative papers, the regimes might differ mainly because of the different information each agent might possess. For instance, retailers can have better information with respect to their consumer portfolio. On the other hand, the single buyer might possess better tools to predict the evolution of future aggregate demand. Therefore, the single buyer is prone to make forecast errors when aggregating information. Still, the process of aggregating dispersed information via a decentralized mechanism also leaves significant room for retailers to behave inefficiently. Another future research area would be to deepen the representation of the decentralized capacity market design, which electricity economists have not quantitatively modeled. Namely, recent advances in games theory and industrial organization could bring significant highlights to what it means to bargain and exchange capacity products between producers and consumers.

Private information, consumers and investment decisions

Chapter Takeaway

There is a trade-off when designing markets to provide revenue for ensuring investments and maximizing consumer surplus: prices allow for covering investment costs while driving consumption. In this chapter, we study this issue under the assumption the utility derived from consumption is partly unknown. It discusses how a policymaker can choose prices and make investment decisions under a set of different constraints. In particular, the chapter shows that under certain assumptions, reaching a certain investment level can lead to distributive issues.

Most of the discussion on the design of capacity markets and, more generally, on the level of investment in electricity markets have been centered around supply-side questions. Namely, how to give the correct incentives to producers to reach a desired level. However, it seems just as essential to know on what basis should be based on this level. This research question is closely related to the two previous chapters, as it mirrors the question of the demand value of capacity mentioned in Chapter 1, and it is a first step to building the demand function of Chapter 2.

This chapter's main contribution is to discuss the implications of considering the demand-side when it comes to ensuring an efficient level of investment. It studies an environment in which there are a number of consumers with different characteristics whose utility from electricity consumption is partly private information. In other words, it seeks to know how to design markets so that consumers reveal their utility for their electricity consumption and, therefore, indirectly for the investment level. In that research, markets are not only used to provide consumption and sufficient revenue to make investments but also to screen for unobservable characteristics to ensure the proper investment level. A second approach that I tackle in this chapter is the distribution issues associated with the design of an optimal market in this context. Namely, I show that implementing the most efficient market to reach the most efficient investment level is not always Pareto improving for every consumer.

To do so, I study the inherent tensions 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. I develop a theoretical framework where a market designer sequentially chooses a level of investment and proposes an allocation mechanism to consumers followed by a consumption stage. 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. The market designer uses the allocation to maximize consumer surplus and finance the investment cost; hence, he is also under a budget constraint. I assume that when the market designer makes investment and allocation decisions, he faces heterogeneous consumers who have private information about their demand level and belong to a publicly observed category.

The first case with incomplete information represents the inefficiencies associated with private information when the market designer is constrained in the allocation choice. Namely, some consumers over-consume with respect to their type, while others under-consume. I first study the single-price case when the market designer cannot discriminate between categories of consumers. In that case, the price schedule increases with the level of investment. I then implement the possibility of discriminating between consumer categories. I show that for the category of consumers with the smaller average private shock, the optimal price first decreases and then increases with the investment level. On the other hand, the price for the category of higher consumers is always increasing with the capacity. I also find that the price for the former category is above the latter for relatively low capacity values, and then the ranking reverses for higher values. These non-monotonicities can be explained by the opposite effect the market designer faces in terms of consumer surplus and revenue effect when choosing prices.

In the last section, I study a mechanism design setup where the market designer is no longer constrained in the prices and quantities schedule he can offer consumers. He now faces incentive compatibility and individual rationality constraints. I first describe for which level of investment the market designer is constrained by the revenue used to cover the fixed costs and the information rent that he needs to provide to consumers so they behave truthfully. I find that the market designer can provide an unconstrained first-best allocation only for low values of the investment level due to the concavity of the virtual consumer utility with respect to the capacity. For higher values, the additional expected utility gains from increased capacity cannot compensate the investment costs. Then, I show that the behavior of the optimal allocation depends on the state of the world considered and the type of consumers. Namely, as the investment level increases, consumers always face a decrease in the optimal quantity allocated during off-peak. For on-peak periods, the quantity change depends on the consumer's type. Finally, I show

that only the large consumers from both categories can gain from an increase in the investment level. The rationale behind this result lies in the link between the investment level and the information rent given to consumers in order for them to behave truthfully.

There are various extensions of this research. To begin, this work has mostly derived an upper and lower bound on the consequences of different market designs given this framework. The first main set of results in the incomplete information environment describes the *current* lower bound in which the market designer is highly constrained and cannot extract any information. On the other hand, the second result describes the *theoretical* upper bound. Namely, with private information and strategic consumers, this is the best market design that the market design can implement if it seeks consumers to reveal their types. However, a significant share of the market and mechanism design literature has shown that the theoretical second best cannot always be implemented, either due to technical reasons, due to the non-linearity of the optimal allocation schedule for instance, or because of social acceptance, as the optimal allocation implies discrimination. Another practical research would be to test how certain forms of capacity markets behave in terms of information disclosure. In other words, one could build a model by adding a module describing long-term mechanisms to the chapter. In that case, the market designer would face a specific set of constraints, and the resolution of such a model would allow a comparison with the upper and lower bound described in this chapter. Finally, the chapter underlines the existence of distribution issues when implementing the second-best allocation schedule to reach a certain level of investment. Hence, there is an exciting venue that would consist of studying the trade-off associated with restricting the market design. At first glance, this trade-off would consist of reducing the available revenue for new investment. However, in the context of the increasing difficulty of financing the energy transition, I believe that such research is of significant interest.

Bibliography

- Fabra, N. and Llobet, G. (2023). Auctions with privately known capacities: Understanding competition among renewables. *The Economic Journal*, 133(651):1106–1146.
- Le Picard, H. (2023). *How Africa is Inventing Decentralized Electricity Networks : An Original Analysis Based on Machine and Deep Learning*. PhD thesis, Paris Dauphine University - PSL. Thèse de doctorat dirigée par Méritet, Sophie Sciences Economiques Université Paris sciences et lettres 2023.

Ly, A. (2023). *Electrification, Environment, and Economic Development in Developing Countries*. PhD thesis, Paris Dauphine University - PSL. Thèse de doctorat dirigée par Creti-Bettoni, Anna et Chakir, Raja Sciences Economiques Dauphine Université Paris sciences et lettres 2023.

Saliou Barry, M. (2022). *Essais sur l'Accès à l'Electricité en Afrique Sub-Saharienne*. PhD thesis, Paris Dauphine University - PSL.

Appendix A

Literature review

Conceptualizing Behavior in Capacity Markets

Contents

| | | |
|-----|---|-----|
| 1 | Description of the literature review | 226 |
| 2 | Conceptual definition of actor behavior | 226 |
| 2.1 | True value and offered value | 227 |
| 2.2 | The opportunity cost of capacity markets | 228 |
| 3 | Litterature review of the <i>true value</i> | 230 |
| 3.1 | Risk and uncertainty | 231 |
| 3.2 | Industrial decisions | 233 |
| 3.3 | Product definition | 237 |
| 4 | Litterature review of the <i>offered value</i> | 239 |
| 4.1 | Market power and asymmetric information | 240 |
| 4.2 | Auction theory | 242 |
| 5 | Litterature reviews of market design issues in capacity markets | 243 |
| 5.1 | Sequentiality of transactions | 243 |
| 5.2 | Decentralization of demand | 245 |
| 5.3 | Price multiplicity | 247 |
| 5.4 | Transaction decentralization | 248 |
| 5.5 | The time horizon covered by the traded product | 250 |
| 6 | Conclusion | 253 |

We will define behavior as "the set of possible actions for an actor when participating in a capacity market, including the decision to participate or not to participate." This definition deliberately encompasses a number of subjects. In the context of the mechanisms studied, we can mention the most obvious ones:

- Whether or not to participate in the mechanism if it is not mandatory;
- The type of technologies offered in the mechanism;
- The pair of price and quantity of capacity offered/purchased;

Since it is the formation of a price that is being studied, and therefore the meeting of supply and demand, actor behavior can be conceptually understood as the way in which actors form supply curves and demand curves, whether the latter are explicit (in the case of auction-type mechanisms, for example) or implicit (in the case of bilateral transactions, for instance). **Such a definition of actor behavior highlight the central question that must be asked before any analysis of the choices made in setting up the mechanisms studied in this thesis:**

"What are the determinants that define actor behavior in capacity markets and which by extension condition its effectiveness?"

To answer this question, we establish a thought framework on the issue of actor behavior in this chapter. This framework refines the economic definition of behavior by studying the factors that can modify and influence the actions available to players in capacity markets. A first approach will enable us to analyze how to represent actors' behavior. Based on a literature review of models applied to capacity markets and terminology specific to these mechanisms, we will define the term actor behavior and highlight the different methodologies employed by the literature.

The following section briefly describes the literature used in this chapter. Section 2, 3, and 4 lay the foundation of the conceptual framework for how we can describe behavior in capacity markets. Section 5 discusses the market design issues of the capacity market.

1 Description of the literature review

This literature review follows on from the reviews carried out by [Bublitz et al. \(2019\)](#) on the current state of our knowledge of capacity markets and by [Scoufflaire \(2019\)](#), who analyzes the results of models comparing energy-only markets and capacity markets.

Electricity economists use a wide variety of models to represent the capacity market and its effect on power markets. They range from stylized equilibrium mode to engineering calibrated dynamic model. A complete description can be found in [Petitet \(2016\)](#); [Höschle \(2018\)](#); [Bublitz et al. \(2019\)](#). To date, the literature review includes 52 models: 10 system dynamics models, 23 equilibrium models, 12 optimization models, and seven agent-based models. Each model has its advantages and disadvantages. More precisely, as they have different objectives in the analysis of an economic system, they represent the behavior of actors in capacity markets in conceptually varied ways. When it comes to the precise representation of these behaviors, this difference highlights both the limitations of each model and the potential improvements to be made. The literature review includes 57 economic models whose distribution between capacity markets is as follows: 36 capacity market models, nine capacity payment models, four strategic reserve models, and eight reliability options models.

2 Conceptual definition of actor behavior

All the models studied focus on representing the behavior of supply-side players. Therefore, we will study, within this framework, sellers as the owners of capacity and the buyers as the consumers. This is due to the fact that all models use centralized demand for capacity with a single buyer. Even though this type of market design is actually implemented in some systems (e.g., PJM, U.K.), other markets have chosen to decentralize demand (CAISO, France). In this case, demand-side actors (such as retailers or large consumers) place bids on the mechanism, which, when aggregated, gives the demand curve. Modeling these behaviors can, therefore, be crucial in assessing price formation on capacity markets.

2.1 True value and offered value

To understand the concept of behavior, one needs to go back to the very notion of a market. That of a place in which an actor on the production side makes an offer, most often a price/quantity pair, and an actor on the demand-side demands the same good, also in the form of a price/quantity pair. When individual offers and demands meet, a price (or several, in the case of bilateral markets) is established, enabling transactions to take place. **With this view, an actor's behavior can be understood as the formation of a market supply.** Economic literature, and in particular auction theory, distinguishes two components of an actor's offer on the market:

- La *true value*: the utility that the actor can derive from participating in a market;
- La *offered value*: the strategy actually implemented by players during trading periods.

These concepts are usually clearly explained in auction models¹. They can also be found in more traditional microeconomic models but in a more implicit way. For example, a profit function can be seen as a *true value* since it determines the utility that the actor derives from the market. The *offered value* in these models is more complex. In fact, it depends on what is known as the conjectural variation, that is, the assumptions that each company makes about the reaction of its competitors (Giocoli, 2005).

As we'll see in the following sections, focusing on the distinction between these two concepts highlights the specific nature of capacity markets. Indeed, the question of the *true value* on these markets is crucial: as Wilson (2010) points out, the first capacity market and their design were based on the idea that players' bids were made on the basis of an average of the total costs incurred by players over the entire lifetime of the investment. However, the first auctions clearly showed that producers constructed their bids in a significantly different way:

The expectation that new capacity would and should be offered into FCCMs at prices based on Net CONE was and remains reflected in the FCCM designs in numerous ways." However, "The FCCMs have not operated according to the expectations described above [...]: Auction clearing prices have not been

¹See, for example, Menezes and Monteiro (2004) for a comprehensive introduction to these models.

stable around Net CONE (or any other price level). The FCCM auction supply curves have not exhibited the anticipated 'flat' segment or any cluster of offers around Net CONE or any other price level, even where the evidence suggests.

Therefore, this approach allows us to reformulate our previous problem. The search for the main factors influencing player behavior in capacity markets would then come down to asking: **"What are the factors that define the *true value*? And how can we model the *offered value*?"**. The literature review in the following sections presents an overview of the different ways this literature uses to model actors' behavior in capacity markets. Based on the above approach, we will outline how the factors have been represented and their associated modeling assumptions.

In the following section, we'll explain how the mechanisms studied depart from traditional good. In section 3, we propose a reflection on the estimation of the *true value* in the markets studied. Finally, in section 4, we present the different approaches adopted by the authors to model *offered value*.

2.2 The opportunity cost of capacity markets

The idea developed by [Wilson \(2010\)](#), and used in the majority of subsequent works is that participation in a capacity market is conditional on the obligation, in the case of the sale of capacity, to be available in a future period on the energy markets; this period is the transaction phase. The sale of capacity on the capacity markets, therefore, entails an indirect cost that could be described as an opportunity cost.

In theory, the opportunity cost of participating in a mechanism is the cost of being available during that future period. In an energy market with no market failures, all players have an opportunity cost of zero since their production is sold without making any losses. However, for players owning electricity capacity, failures and constraints specific to energy markets can result in losses. Consequently, forcing a player to produce during a period when he is potentially making a loss entails a positive cost.

The opportunity cost of the availability obligation is, therefore, the potential loss incurred during the period of availability obligation. Many papers that seek to model the interdependence of players and different economic decisions in power systems are based on this principle of the opportunity cost of availability ([Abani et al., 2018](#); [Bhagwat et al.,](#)

2017a,b; Teirilä and Ritz, 2018). In a simplified model, Creti and Fabra (2007) showed how a monopoly bids on a capacity market when it has to forego exporting electricity to a foreign market whose price is higher than the price on the domestic market due to the obligation to be available. The offer made on the capacity market is made at a price equivalent to the loss of opportunity to make a profit on the external market.

The few papers modeling reliability option markets are also enlightening with regard to this approach. These mechanisms, similar to classic capacity markets, are based on the exchange of financial options between the players holding the investment and the demand. These options, initially held by the players, are sold on a market, which then constitutes remuneration for their capacity. In exchange, during periods of demand scarcity, i.e., when the price exceeds the strike price, the players undertake to pay back to demand the difference between the selling price received on the energy markets and the strike price of the obligation defined in the product (Cramton et al., 2013). In the models, the initial *true value* of these options is consequently equal to the amount paid back to demand, which is indeed the opportunity cost associated with such products (Mastropietro et al., 2016; Meyer and Gore, 2014).

Figure A.1 illustrates a hypothetical situation in which a player participates in a capacity market every year for four years and is, therefore, forced to produce during periods of scarcity. Graph (a) illustrates his annual estimate of net profits on the energy market (the inframarginal rent). It is assumed that this energy market is constrained and inefficient, so the actor also estimates that he will make losses by being forced to produce during transaction phases (i.e., his missing money). As investment costs are stranded, only operation and maintenance costs are taken into account. The opportunity cost associated with participating in the capacity market each year is, therefore, the difference, when positive, between fixed costs on the one hand and inframarginal rents realized on the energy market on the other (i.e., the difference between the price of electricity and the variable cost). Under these assumptions, the bids made on the capacity market are equal to this difference. The sequence of these bids is illustrated in Figure A.1. Note that the bids can vary and that they depend solely on the profit and cost estimates made for the transaction phase of the year concerned by the product sold.

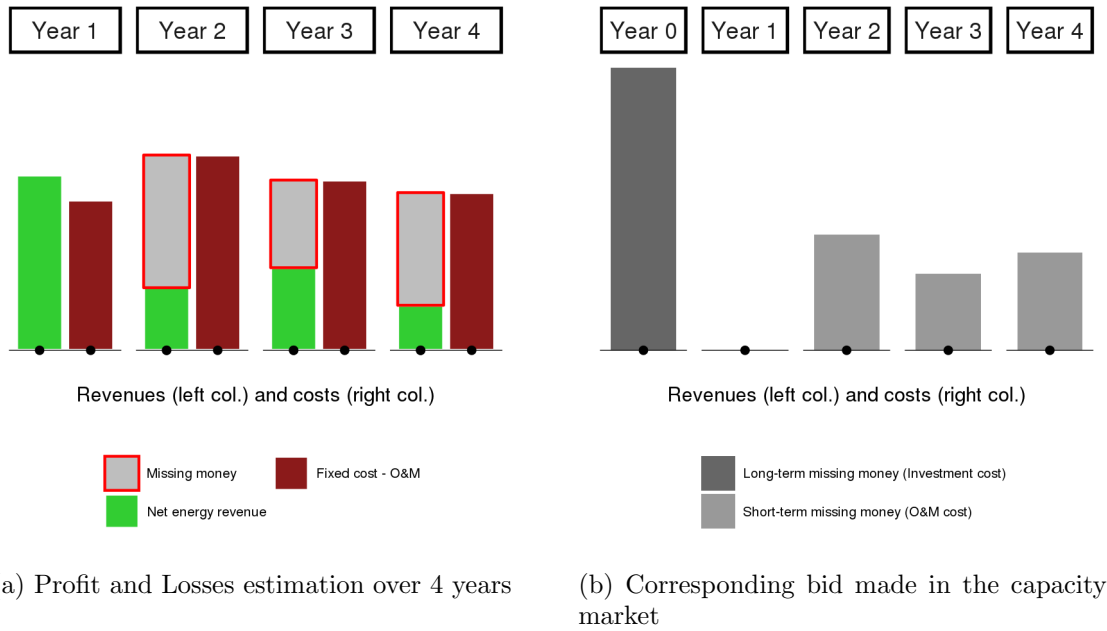


Figure A.1: Illustration of the opportunity cost for an incumbent

3 Litterature review of the *true value*

The fundamental hypothesis described so far is that capacity markets consist of selling an obligation for players to be available. In other words, the product sold on the capacity market is conceived not necessarily as physical capacity but as an assurance that the seller of the product will be present during the periods defined in the transaction phase. To illustrate this concept, it is possible to find rules in the implementation of specific capacity markets according to which the product sold can be either physical or financial. In the latter case, sellers are not necessarily investors owning generation or demand response but maybe traders acting as intermediaries.

In all cases, participation in a capacity market entails an opportunity cost, which can be understood as the difference in profit between producing and not producing (or consuming and not consuming). In this case, the *true value* associated with the capacity market is precisely this cost of participating.

Once this *true value* has been defined, the key question is how players can estimate this value. Indeed, the latter can have a wide variety of characteristics. Moreover, as we will see, market rules necessarily lead to significant differences in the assessment of this value. In connection with the literature review, we will address three topics that

are both fundamental to the understanding of *true value* in the capacity market but also insufficiently addressed in the literature:

- Taking risk into account ;
- The industrial characteristics of the players;
- The characteristics of the product sold on the mechanisms.

As we will see in the following sections, the treatment of opportunity cost has been better studied in quantitative models than in analytical models. Indeed, the opportunity cost approach assumes a particular form of dynamism that static models of the equilibrium or optimization type have difficulty taking into account. The usefulness of these models explains this difference in treatment. Indeed, quantitative models seek to introduce greater realism and take into account the dynamic nature of the models, while analytical models tend to emphasize strategic actor behavior.

3.1 Risk and uncertainty

In the economic literature, there are two types of risk particularly relevant to capacity markets ([Scouflaire, 2019](#)):

- Exogenous risk; and
- Imperfect information specific to public decisions.

In the latter case, although this topic is relevant to an overall analysis of mechanisms, it does not seem necessary to address it in the context of this thesis. Indeed, a priori, there is no reason why private players should include regulatory failures in their estimate of their opportunity cost. To date, a few papers have studied these failures, either in the form of quantitative models ([Winzer, 2013](#)) or analytical models ([Feng and Xu, 2009](#)). The primary source of regulatory error lies in the setting of target levels and the choice of the measure to be adopted. The error would consist of incorrectly forecasting consumption levels, resulting in either an over-capacity or an under-capacity situation.

Exogenous risk is generally well-represented in the literature on capacity markets. The main level of uncertainty in the literature is electricity demand. Some authors

also represent the risk associated with fuels (Bhagwat et al., 2017a,b), the technical constraints of means of generation (Cepeda and Finon, 2011, 2013), or the decisions of authorities (Joung et al., 2009; Mastropietro et al., 2016).

In the majority of cases, the authors analyze the impact of risk in order to compare different types of mechanisms, especially the effect of the presence or absence of a capacity market. Indeed, in perfect information and the lack of risk, theory shows that the impact of a capacity market is identical to that of an energy-only market (De Sisternes and Parsons, 2016). These mechanisms only transfer part of the rents between players, a transfer that takes place directly in an energy-only market. However, when risk is taken into account, that is, when we represent how rents are distributed according to possible scenarios, then capacity markets have non-negligible effects on risk for the various players. In a similar literature review, Scouflaire (2019) finds that, in the totality of papers analyzed, capacity markets all imply a decrease in price volatility, a significant risk factor for producers and consumers.

From a practical point of view, it's essential to understand that risk generates implicit costs for players, whether in production or consumption. However, in modeling terms, this cost can only be taken into account if we assume risk-averse actors. The definition of such aversion is simply a preference for a certain amount whose value is lower than the expectation of the same but risky amount. In a model applied to electricity investment, Meunier (2013) shows that risk can be represented as a cost in the same way as variable and fixed costs. In this case, the level of adequacy achieved corresponds to that of an optimal production mix with technologies whose cost of production is adjusted for risk.

This notion of risk aversion is found mainly in quantitative models due to their need for realism in the interpretation of their results. When applied to capacity markets, the tools used to represent risk are remarkably diverse. The literature uses utility functions such as the Mean-Variance financial theory function (Meunier, 2013; Winzer, 2013) and the more classical exponential function (Hobbs et al., 2007a; Fan et al., 2012; Petitet, 2016). Applications can also be found using tools specific to asset financing: the CAPM (Peluchon, 2019) or the CVAR (Abani et al., 2018). Building on the seminal work of Ehrenmann and Smeers (2011a,b) on modeling investment decisions in an uncertain environment in the form of stochastic optimization models, a very interesting literature attempts to model the effects of different mechanisms (including long-term contracts) on the level of investment and adequacy (de Maere d'Aertrycke et al., 2017; Abada et al., 2019; Peluchon, 2019).

Finally, many system dynamics-type models are based on the notion of bounded rationality (Olsina et al., 2006; Hobbs et al., 2007b; Assili et al., 2008; De Vries and Heijnen, 2008; Hary et al., 2016) to represent the imperfection of actors' information in their estimation of electricity prices and capacity. This model, which is based on relatively arbitrary characteristics, does, however, represent a well-known effect in power sector investments, namely the cycles of undercapacity and overcapacity.

In Chapter 1, we provide a new contribution to the literature by considering the risky environment in which producers evolve when analyzing the bidding behavior in the capacity market. Using real options theory, we study how risk and risk aversion imply different bids compared to canonical risk-neutral approaches to measure the opportunity marginal cost of participating in the capacity market. Therefore, we depart from the previous literature by focusing not on how risk interacts with the capacity market and investment decisions but by describing the link between risk and behavior in a capacity market.

3.2 Industrial decisions

Producers and consumers in current power systems are faced with a large number of industrial decisions. These decisions, which could be described as options (and potentially real options in the case of some that are characterized by a form of irreversibility), describe a set of actions available to players that can significantly modify the value of the opportunity cost in capacity markets.

The most obvious is whether or not to enter the energy market. Hence, a distinction needs to be made between players already present in the markets (incumbents) and those wishing to install a new investment or consumption (new entrants). At this stage, the models have not, to our knowledge, studied portfolio effects that would imply a difference in treatment between a new capacity for a player with existing capacity or a new entrant to the sector. Nevertheless, studies are assuming a difference in efficiency (in terms of costs in particular) between incumbents and new entrants (Brown, 2018a,b). Cross-subsidies could also exist, but the subject has not been addressed either.

Figure A.2 thus illustrates a hypothetical situation in which a new entrant considers the year-0 offer it must make if it participates in the capacity market. Assume its lifetime is five years, and it incurs losses over this period. These losses are calculated from the difference between his estimates of the marginal rent earned on the energy

market, capacity prices in the case where the player is retained, and fixed costs made up of investment costs and operation and maintenance costs. The illustrative values of these components are shown in Figure (a). The fundamental difference in estimating the opportunity cost between this new entrant and a player who has already built capacity lies in the fact that investment costs and future capacity prices are taken into account. This hypothesis assumes that the player is a price-taker in the capacity market. The new entrant's offer is then equal to the sum of fixed costs in the first year and the discounted sum of potential losses in the future (here at an arbitrary rate of 5%). Its value is shown in Figure (b).

Although the model shown in Figure A.2 is illustrative, it does help to show the difference in opportunity cost between a new entrant and an incumbent. Indeed, the *true value* in year 0 is evaluated by the need to cover potential losses in subsequent years. However, the initial model in year 0 does not give us any direct insight into the value for the player once year 0 is over. Revenues linked to capacity markets are only estimated values and by no means represent a certain future. One of the assumptions made by specific authors is that players will continue to reason both in terms of margin and opportunity cost. The crucial difference lies in the time horizon considered when deciding whether or not to participate in the mechanism.

- When an actor decides to enter the market, he takes, as we have seen, estimates of his profit over a relatively long time horizon. His objective, when he participates in the mechanism, is to cover all these fixed costs, taking into account the estimated prices of energy and capacity.
- When the player has already entered the market, his time horizon is greatly reduced. Since it has succeeded, in theory, in covering at least its investment costs, the player must now ensure that it covers only its short-term fixed costs, i.e., its operating and maintenance costs.

This distinction between the two objectives when participating in a capacity market is due to the fact that the player does not base his decision on an estimate of the capacity price when he is already present in the market. His primary objective is to fill the short-term missing money that arises when he is obliged to be available during the product transaction period when he enters the contracting phase. This change of horizon is illustrated in Figure A.1, where the exact estimates of energy revenues and operation and maintenance costs are represented. As investment costs have been covered by the

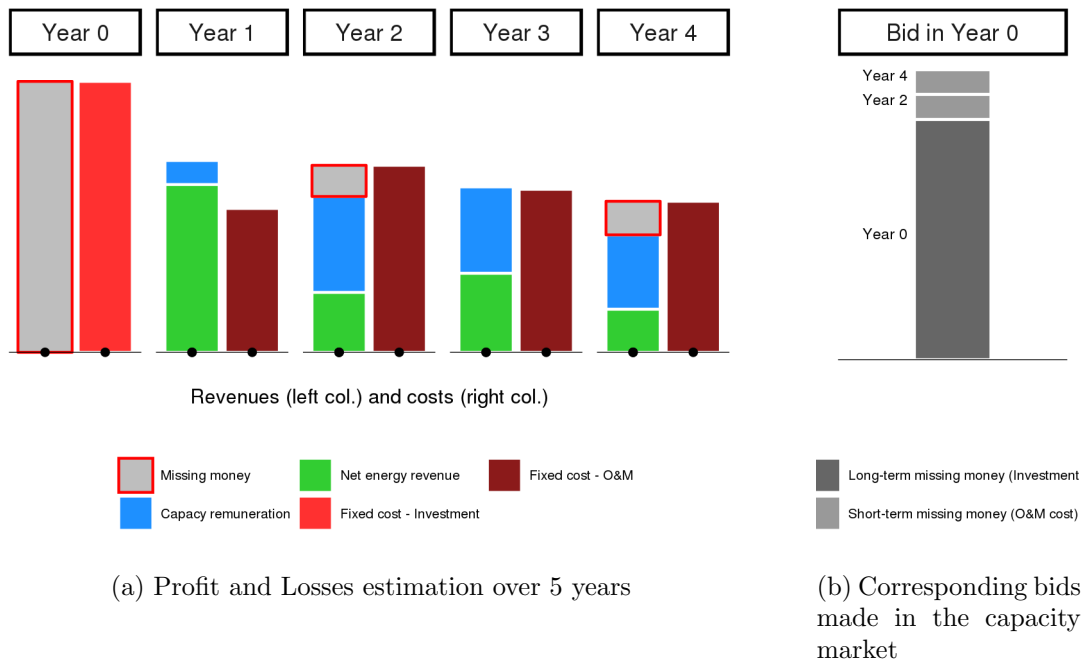


Figure A.2: Illustration of the opportunity cost for a new entrant

transaction in year 0, they no longer appear in the actor's estimates. Moreover, since the player is only looking at the transaction period associated with the product he sells, estimating the future capacity prices is not necessary. Consequently, at each capacity market clearing, the opportunity cost of being available during the transaction phase is equal to the missing money corresponding to the difference between energy revenues and maintenance costs.

Such assumptions about capacity markets are crucial to understanding the economic fundamentals of these mechanisms. In particular, they establish that the price of capacity must adapt at all times to reflect the conditions of adequacy of the power system. According to the model outlined above, when capacity is to be added to the system, the market price must necessarily adjust so that it equals the long-term missing money as well as the sum of the short-term missing money estimates for the most efficient player available as a new entrant. [Wilson \(2010\)](#) summarizes these conditions well:

"The FCCMs [Capacity Markets] were designed based on the concept that clearing prices should reflect the cost of new entry when new entry is needed. [...], with the cost of new entry understood to mean the price level at which sufficient existing and new capacity has been offered to meet reliability re-

quirements. [Natural” clearing of FCCM supply and demand leads to prices that reflect the true state of supply and demand at any time.”

In other words, one cannot expect prices on capacity markets to be stable in the same way as long-term contracts, for example. When looking at incumbents for investment and consumption, the *true value* of each actor will also depend on a set of decisions concerning the transaction phase of the product sold: (i) The decision to continue producing, i.e., to take part in the mechanism and make oneself available; (ii) mothballing during the transaction period, i.e., the power plant does not produce but does not incur a certain number of operating costs; (iii) exiting the market, which corresponds to a total halt in production over all the years under consideration.

The possibility of mothballing a player means forgoing remuneration on capacity markets but also means avoiding operating and maintenance costs. When there is a risk that the player will not be sufficiently remunerated on the capacity market, his offer may be modified by the option of not making himself available on the energy or capacity markets. Quantitative models are the most numerous to have focused on the representation of industrial decisions in capacity markets. [Abani \(2019\)](#) offer a comprehensive analysis of the importance of actor decisions through opportunity cost modeling. This study highlights the great diversity that can emerge when we look at the offers that can appear on these mechanisms. In an agent-based model, [Mastropietro et al. \(2016\)](#) differentiates new entrants from incumbents by including investment costs in offers on a mechanism. Most system dynamics models incorporate this in their representation of actor behavior ([Petitet, 2016](#); [Hary et al., 2016](#); [Abani et al., 2016](#)).

It is worth noting the work that has focused on the trade-off between decisions to invest in plants already installed (refurbishment) and the installation of new capacity. In particular, the effect of the possibility of extending lifetime has been addressed in an optimization model by [Lynchab and Devinea \(2017\)](#) and in a system dynamics model by [Abani et al. \(2016\)](#). In the first paper, the authors model the effect of refurbishment in terms of the availability of players in energy markets, leading to a positive correlation between producer profits and the refurbishment decision. Although there is no implicit representation of the opportunity cost, such an industrial decision provides a better fit for the system. In [Abani et al. \(2016\)](#) paper, players base their trade-off between new capacity or keeping their old capacity, whose O&M cost increases according to their risk aversion. Indeed, an O&M cost is considered relatively fixed and predictable, while the cost of a new plant is risky even if, in the expectation, it remains less costly. In this

case, the effect of refurbishment is analyzed in the context of risk and risk-averse actors. Although the opportunity cost is made explicit in this model, the net impact of the industrial decision is not analyzed.

Finally, when these decisions are irreversible (notably the decision to enter or close), it is possible to apply an analysis specific to real options theory, which highlights the value of this decision-induced flexibility. To date, the effect of industrial decisions analyzed through the prism of real options has only been addressed by [Hach and Spinler \(2016\)](#), in which they model the implementation of a capacity payment.

The analysis of the risk in [Chapter 1](#) is made concomitantly with the managerial flexibility an investment can have. Namely, we use real options theory to analyze the option value of participating in a capacity market when a producer has the option to mothball to avoid O&M costs. Therefore, this chapter also contributes to a better understanding of how industrial decisions can affect behavior in a capacity market. [Chapter 2](#) also relies on the idea that producers make a specific commitment when participating in a capacity market. In an analytical framework, we build a supply function in a capacity market, which is based on the marginal investment cost net of the expected marginal profit made on the energy market. One of the main technical contributions is the endogeneization of the supply function. Namely, we show the different indirect effects of capacity market design regimes that can be captured through this supply function.

3.3 Product definition

The notion of product design relates to the set of characteristics that define the transaction phase of capacity markets. Several key elements define these rules: (i) The type of obligation associated with the commitment made when the product is sold, (ii) the penalties associated with non-compliance with the commitment, and (iii) the duration of the commitment.

These issues are directly linked to the practical implementation of capacity markets around the world. Their definition, therefore, varies from system to system. To date, few studies have focused on analyzing the effects of market rules on actor behavior. Yet the general economic literature, and especially that on auction theory, stresses the importance of market rules on player behavior.

In the context of capacity markets, the importance of penalties for the availability obligation is highlighted in the analytical work of [Joung et al. \(2009\)](#) and in the quantitative work of [Mastropietro et al. \(2016\)](#). To be more precise, the penalty associated with engaging in capacity payment mechanisms can be understood as a form of opportunity cost. Indeed, going back to the hypotheses put forward in the previous sections, when a player commits himself when he participates in a capacity market, he must necessarily produce (or consume) during the transaction phase. However, we did not make any assumptions about whether the actor decided not to produce or was technically unable to do so.

When we introduce this notion of penalty, we then explicitly take into account the cost of unavailability for the actor. Indeed, in a model with uncertainty, if the probability of not being available over the obligation period is positive, then two scenarios can be considered:

- On the one hand, the actor produces and suffers the missing money as modeled in [Figure A.2](#) and [Figure A.1](#);
- On the other hand, the actor does not produce (voluntarily or involuntarily); in this case, he suffers the penalty associated with non-compliance with the conditions of the transaction phase. Without the penalty, the induced cost would be identical to that of mothballing (assuming the actor does not incur O&M costs during unavailability), i.e., relatively low.

As a result, and in expected value terms, the opportunity cost associated with implementing a penalty would increase. [Bublitz et al. \(2019\)](#) highlights the conditions for implementing such rules in capacity markets:

”If a financial penalty is chosen, it needs to be high enough to encourage investors to comply, which, however, increases the risk of investors, and this is reflected in their bids. For the exact amount of the penalty, it is possible to rely on the VoLL, the capacity price, or the Net CONE.”

The rules associated with the obligation to produce determine the magnitude of the rent transfer made during the transaction phase. The characteristics of this obligation can also be varied: the form of the obligation (e.g., on which market players are controlled

when supplying energy) and the level of uncertainty associated with the obligation. To date, no studies have proposed quantitative or analytical analyses on this subject. However, some authors have studied the effect of the obligation to produce on players' strategy, i.e., on the *offered value*. This question is addressed in the following section.

At first glance, the various associated rules seem to have a quantitative effect on capacity markets. In other words, their implementation does not alter the economic fundamentals and functioning of these mechanisms. However, some studies tend to show that they can have a non-negligible effect on the behavior of players through their strategy. Several models attempt to represent this availability obligation in capacity markets and energy markets (Schwenen, 2014). The work of Bialek and Unel (2019) provides an enlightening view of the advantage of flexibilizing products on capacity markets in order to match the obligations provided by players according to the different seasons. Indeed, the cost of availability as well as demand are significantly different depending on exogenous factors such as temperature or luminosity. Allowing the exchange of flexible products could lower costs. However, the authors do not propose any modeling of actor behavior.

Using the approach of combining risky environment and industrial decisions in Chapter 1, we also contribute to the literature by studying the relation between capacity products and behavior in a capacity market. More precisely, we analyze how the length of the transaction phase increases or decreases the bid and if it is better to have a succession of short transition phases instead of having a single product. We are also able to capture the effect of a penalty on the bidding behavior. In Chapter 2, we provide, in an extension, a representation of a decentralized capacity market. In this framework, retailers are forced to cover their expected consumer portfolio consumption. If they do not buy enough capacity, they face a penalty. Therefore, we are able to show that this penalty is a fundamental driver of the aggregate demand function in the capacity market.

4 Litterature review of the *offered value*

The economic analysis of player behavior should not be limited solely to the representation of opportunity cost. Indeed, economic theory underlines that there is no guarantee that a player will actually offer what it seeks to cover by participating in a market. This dichotomy between *true value* and *offered value* is most often described as an actor's strategy.

Economic theories use different terminology to characterize the notion of strategy. In classical microeconomics and industrial organization, this notion refers to opportunistic behavior. Strategy is then measured as the difference between marginal cost and price, which, according to the terms used above, can be considered as *true value* and *offered value*. In contract theory, asymmetric information leads actors to report false values. For example, when an actor has a certain level of availability to satisfy the supply-demand equation, there is no guarantee that this same level will be reported to the authorities when they seek to assess the actual level of availability in the system. The advantage of auction theory, which introduced this distinction, lies in its relative neutrality with regard to the notion of opportunistic behavior.

In this section, we will attempt to analyze the approaches adopted in the economic literature to represent the strategy of players in capacity markets. In this context, strategy can be understood as the difference (positive or negative) between an actor's opportunity cost and the offer he actually makes on the mechanism.

4.1 Market power and asymmetric information

Classical microeconomic models distinguish between situations of pure and perfect competition and those of oligopoly. In reality, the main difference between these two approaches is one of conjecture. Models of pure and perfect competition imply that players offer their marginal costs on markets, which in conjectural terms rests on the principle that their reactions and those of competitors do not affect price formation.

Several models of capacity markets are based on pure and perfect competition. In this case, the offer made by a player is strictly identical to the opportunity cost, whether he is a new entrant or incumbent. Such an approach is mostly used in quantitative models ([Bhagwat et al., 2017a,b](#); [Cepeda and Finon, 2011, 2013](#); [Franco et al., 2015](#); [Hach and Spinler, 2016](#); [Mastropietro et al., 2016](#)).

For oligopoly models, there are variants to the conjectures. The most widely developed in the literature is Cournot's conjecture: players assume that their competitors will not react to their strategy. This model has been commonly used in economic theory applied to energy markets due to the similarities between this model and the characteristics of electricity markets ([Twomey et al., 2005](#)). As part of the literature review, we can also see its use in representing the conjectures of different players in capacity markets ([Joung et al., 2009](#); [Elberg and Kranz, 2013](#); [Brown, 2018a,b](#); [Teirilä and Ritz,](#)

2019). The effects of such models on the representation of actor behavior are well-known and concur with the results of the generalist literature. In a situation of imperfect competition, actors tend to offer their availability at a price higher than their opportunity cost while reducing the quantity provided on the mechanism. The direct consequence of these strategies is to undermine the economic efficiency of capacity markets, both by increasing the cost and by the risk of not satisfying the adequacy criteria.

In order to represent the asymmetry existing between players in electricity markets while remaining within a framework of imperfect competition, some authors have also used other conjectures. The main one is based on modeling a small number of firms facing a multitude of firms behaving as in pure and perfect competition (Joung et al., 2009; Elberg and Kranz, 2013; Brown, 2018a,b; Teirilä and Ritz, 2019). In this case, dominant firms face residual demand and behave identically to a Cournot-style model, while competitive firms offer at their opportunity cost. To our knowledge, there are no Stackelberg-type models directly applied to the formulation of offers. Nevertheless, some authors use this model to introduce a form of sequentiality into either mothballing decisions (Brown, 2018a,b). or investment decisions (Le Cadre and Soubra, 2013),

Finally, in the context of information asymmetry between actors other than conjectures, the models studied have mainly attempted to represent issues between authority/actors rather than between actors themselves (Creti and Fabra, 2007; Joung et al., 2009; Le Cadre and Soubra, 2013). Le Cadre and Soubra (2013) model is particularly interesting, as it focuses on market rules that respect incentive and participation conditions. More specifically, the authors construct a capacity market in such a way that payment induces players to declare their actual availability on the markets in order to avoid strategic behavior.

Both Chapters 1 and 2 rely on perfect competition on the supply-side, as they are well-developed in the literature on electricity markets, both with or without a capacity market. However, in Chapter 2, we contribute to the literature by providing a stylized model where retailers that participate (indirectly or directly) in a capacity market play 'à la Cournot.' We notably describe the effect of allocating the capacity cost on retailers based on their relative share and how this allocation changes their behavior. We also show that imperfect competition plays a role in the decentralized design. In Chapter 3, one of our contributions is allowing consumers to behave strategically when a market designer implements prices to provide revenue to make investment decisions. In that

sense, we show that producers are not the only agents to deviate from optimal behavior when thinking about how we can design markets to ensure an efficient level of investment.

4.2 Auction theory

Other promising work on capacity markets is based on a complex representation of interactions between players. This work is based on the seminal article by [Fabra et al. \(2006\)](#), who developed an auction model in which a pivotal player and the reactions of inframarginal players are represented according to the strategies of this pivotal player and the level of demand.

First applied to energy markets, then to investment decisions in [Fabra et al. \(2011\)](#), this model was extended in the work of [Schwenen \(2014\)](#) as part of a theoretical model studying the effects of a capacity market on the energy market and in [Schwenen \(2015\)](#), in which the author empirically compares the model's results with bids made on the PJM capacity market. The effect of capacity markets on the system is reflected in the obligation to produce on the energy market. By committing these capacities, players participating in a capacity market create a lower limit to the quantity offered in the energy market, which, in the context of the relative concentration of supply in energy markets, can limit capacity withholding strategies.

Other previously cited authors also rely on this model, notably to represent the link between investment subsidies and behavior on capacity markets ([Brown, 2018b](#)), as well as between demand and behavior on capacity markets ([Brown, 2018a](#)). Finally, a particularly interesting model also developed by [Brown \(2012\)](#) represents price dumping strategies by incumbents on capacity markets in order to create a barrier to entry for potential new entrants. In other words, the model assumes that players make offers on the capacity market in such a way that they do not cover all their missing money. Such a strategy can only be profitable if the new entrant does not build new capacity, as entry has potentially led to lower prices in the energy markets and, therefore, a loss of profit for the incumbents.

To our knowledge, there have been no other attempts to apply auction theory to capacity markets. This absence in the academic literature is quite remarkable, both because of the well-known contribution of this theory to the understanding of how markets work, notably in the representation of the effects of the characteristics of market

designs on player behavior, but also from a practical point of view, as the vast majority of capacity markets employ at least one auction in the allocation of capacity.

5 Litterature reviews of market design issues in capacity markets

From a general point of view, it is now recognized that current capacity market designs have corrected specific inefficient rules, particularly on issues related to the definitions of the traded product (Cramton, 2017). For example, the introduction of penalties in U.S. mechanisms has confirmed the need to align the objectives of a capacity market and availability during transaction phases (Bushnell et al., 2017).

An essential point for a better understanding of the issues associated with the market design of competition-based mechanisms is the strong interdependence between specific issues. To date, the literature has already established these connections between different market designs. It seems to us, however, that certain links remain to be made.

5.1 Sequentiality of transactions

Auction sequentiality has not yet been modeled in capacity markets. However, Salant and Stoddard (2008) highlights the importance of auctions in price formation in capacity markets based on theoretical work that has not been applied. They describe several results from the literature on price formation in sequential auctions. In particular, Weber's Martingale Theorem states that under the assumption that the "costs of the bidders participating on both are the same, then the expected prices in both auctions should be the same," then "the expected price does not depend much on the division of the quantity over time, provided there are adequate lead times for all the auctions or requests for offers." However, the martingale does not hold if there is a link between the strategy on a first auction and the quantity offered on the second or if the order of economic efficiency between players is altered, due in particular to uncertainty over costs. If the mechanism is also decentralized at the demand level, but competition is relatively weak at the producer level, then multiplying auctions would also reduce market power.

Auction dynamism in capacity markets is also the focus of discussion by [Harbord and Pagnozzi \(2014\)](#) on the design in Colombia and the U.K. Indeed. The Colombian mechanism is based on descending clock auctions in which bids are disclosed at each increment. As a result, each player observes the bids of its competitors before the market is fully cleared. According to economic theory and in the view of the Colombian authorities, this type of sequentiality reduces players' risk, particularly with regard to the winner's curse. It increases the economic efficiency of the auction. These results depend on a number of assumptions about the market, notably that the product exchanged is of the common value type, i.e., that an actor's observation of one of his competitors' bids modifies the value he has of that good, which is his true value. However, the authors point out that the private value environment does not necessarily characterize capacity markets. Consequently, there is no benefit of learning competitors' bids, leaving room for opportunistic behavior such as tacit collusion. In this case, the results of auctions are similar to those developed in models applied to energy markets. However, the recent introduction of the Irish mechanism based on descending clock auctions underlines the fact that the debate is not yet over. A series of reports from [AESO \(2018\)](#) concludes that the type of auction to choose depends strongly on the characteristics and structure of the market under consideration: In the case where players are relatively concentrated and the uncertainties about fundamentals are low, then a uniform, non-dynamic auction is preferred. In the opposite case, a dynamic descending clock auction may be more appropriate.

Finally, the modeling of a descending clock auction was carried out by [Teirilä and Ritz \(2019\)](#). Their approach is based on an analysis of market power in a reliability options mechanism. It finely represents market concentration, entry-exit issues, and certain industrial decisions. However, the role of information in the auction is not analyzed, and a comparison with other transaction modes is not made.

By definition, the temporal sequentiality of transactions for the same product can only arise when the capacity market clearing is relatively distant from the transaction phase. In other words, such an analysis can only be made in the case of capacity markets and not in the case of spot markets. A second connection can also be made between the problems associated with multiple prices. Indeed, sequencing transactions means producing several prices for the same good. To our knowledge, no study has attempted to draw such a parallel in capacity markets.

5.2 Decentralization of demand

Decentralization is a relatively unexplored topic in the literature, with a large proportion of capacity markets having chosen a centralized mechanism. As the supply of capacity is mainly made up of private players, this question primarily concerns the choice of having a single buyer of capacity or, on the contrary, letting electricity retailers cover their consumption with products traded on the mechanisms. The former is referred to as a centralized mechanism, the latter as a decentralized mechanism.

The decentralization of a capacity market can lead to significant differences. Indeed, the target level in a decentralized mechanism is not necessarily decided at the system level: it may rest with each actor on one of the two sides (in practice, usually the demand-side). The aggregation of individually defined target levels then constitutes the system target level of the public policy. As in a classical market, remuneration on decentralized mechanisms takes the form of transfers between demand and supply actors. A centralized mechanism, on the other hand, involves a single transfer between the authorities and only some of the players. In other words, the inframarginal rents and scarcity rents missing from certain players are centralized at the level of the single buyer before being redistributed to private players, which is not the case with decentralized mechanisms. It is at this level, when the aggregation of individual demands and offers takes place in conjunction with the transfer of rents, that there is potentially a deviation between the two mechanisms.²

According to some papers, the advantages of decentralization are based on greater responsibility on the demand-side to cover demand RTE (2014). Indeed, by definition, demand-side actors have the best access to information on the need to cover their portfolio. Any aggregation of information by a player then entails a risk of information asymmetry and opportunistic behavior (PowerAuctions, 2016).³ De Maere d’Aertrycke

²The notion of centralization can also concern the formation of products on capacity markets. In particular, the different elements of these products (such as the activation period and notably the level of *strike prices*) could emerge from negotiation between the different parties participating in the mechanism (Pöyry, 2015). To date, the majority of systems are based on product standardization decided at the public authority level.

³Note the comment by Harvey et al. (2013): "This has been particularly apparent in the PJM capacity market, which contracted forward for capacity based on load growth forecasts that proved materially inaccurate following the financial crisis, with the cost of keeping the excess capacity in service, or of buying back the capacity obligation at a lower price in an incremental auction, borne by PJM power consumers."

et al. (2017) point out that centralizing bids on capacity markets prevents players with different risk aversions from adequately managing their risk: ”

”In contrast with preceding contracts, FCM does not offer risk-trading capabilities: the risk-taking consumers do not reveal their risk aversion, but the government contracts capacity on their behalf.”

The issue of risk has yet to be studied from this angle. The authors’ article sheds some very interesting light on the subject. Moreover, the problem raised by the authors refers to the more general limits for centralized models to reveal and aggregate the information available within a system. Pöyry (2015) outlines the main advantages of decentralization, notably in terms of economic efficiency, redistribution of rents between players, and better integration of specific technologies.

Finally, the lack of concrete analysis on the formation of a decentralized demand curve in the case of capacity markets prevents a rigorous analysis of their economic efficiency. Indeed, the shape of the demand curve is crucial to understanding price formation in a capacity market (Hobbs et al., 2007b). In particular, several authors emphasize the link between this curve and the exercise of market power (Bublitz et al., 2019; Teirilä and Ritz, 2019). In a very interesting approach, Brown (2018a) seeks the parameters of an optimal administered demand curve that would minimize costs for society in the context of imperfect demand. According to the analysis of specific authors, it would seem that the effectiveness of decentralization of demand also depends on other criteria, in particular, the degree of decentralization of transactions, which will be discussed in greater detail in the following sections. Another practical criterion, absent from the academic debate, is the temporal definition of the individual responsibility associated with the purchase of capacity. Indeed, two design options may exist within a decentralized mode:

- The obligation to cover individual demand is defined before the transaction period, then decentralized to the demand-side⁴ ;
- The obligation is associated with the individual demand realized by each demand actor after the transaction period;

⁴For example, an aggregate adequacy level is defined initially, then shared between each player on the basis of past demand achievements

According to initial feedback provided by [Cigré \(2016\)](#), it would appear that the incentives for actors linked to the second option are much more significant than those generated by the first option. Indeed, in the first case, individual demand is de facto exogenous. Demand-side actors, therefore, have no incentive to modify their behavior in such a way as to increase or decrease their level of contracting. This is not the case for the second option, where players have an incentive to modify their consumption level if, for example, the cost of demand response is lower than the price on capacity markets.

5.3 Price multiplicity

The question of price discrimination in capacity markets relates to two issues raised in the literature:

- The form of the auction (uniform / discriminatory auction) ;
- Discrimination between existing and new capacity, issues specific to these mechanisms.

In the first case, there has been no work comparing the two auctions for the capacity markets. However, [Harbord and Pagnozzi \(2014\)](#) summarizes the results of the literature and the lack of consensus in both theoretical and empirical work applied to the power sector. The relatively less detrimental effect of the uniform auction for small players leads to their choice for this type of design. Indeed, the authorities managing the mechanisms most often advocate uniform auctions ([DECC, 2012](#); [AESO, 2018](#)).

The multiplicity of prices in this form can be analyzed from multiple angles. Aside from discriminating auctions, as mentioned above, sequentiality also introduces a form of discrimination. Some authors, notably [Bower and Bunn \(2000\)](#), assume that a market where transactions are decentralized is identical to a discriminating auction. This hypothesis thus establishes a link between price multiplicity and transaction decentralization.

In the second case, papers on capacity markets are more detailed. The question of discriminating between the two types of capacity is still being debated, as evidenced by the existence of both types of design. In a detailed qualitative report on the California mechanism, [Pfeifenberger et al. \(2012\)](#) argues that discriminating between new

and existing capacity is likely to cause significant problems in allocating capacity in the system:

”Price suppression benefits are temporary, achieved at the expense of existing suppliers - Missed opportunities to procure lower-cost existing assets, D.R., and imports - leads to greater retirement, requiring even more new generation to be built - increased regulatory risk - new generators will demand higher prices to compensate for suppressed prices after contract expiration.”

[Cramton et al. \(2013\)](#) also highlights the limitations of capacity discrimination. The main argument is that the service provided by a capacity, whether new or existing, is strictly identical. This idea is conceptualized by asserting that in an energy-only market, there is no discrimination between capacities, so the mechanisms do not have to either. In a theoretical framework linked to the analysis of players’ strategic behavior, [Salant and Stoddard \(2008\)](#) notes that :

By splitting the purchase of a single undifferentiated product, administrative capacity credits, into two distinct procurements, a buyer in the first procurement can underpay or overpay relative to the second. One problem with differential pricing for existing and new capacity is the fact that this can distort plant retirement and other maintenance decisions. The potential mismatch in prices between the two procurements not only has cost ramifications for purchasers but also indicates an inefficient allocation of capital and, consequently, inefficiency in the market.

Similar conclusions were discussed in an [AESO \(2018\)](#), the aim of which is to shed light on the establishment of a capacity market in the region covered by the institution. To our knowledge, there is no similar analysis of the benefits of price discrimination between existing and new capacity. However, the type of product offered to them may vary, which also leads to discrimination. This topic will be addressed in a subsequent section, which deals explicitly with product differences in capacity markets.

5.4 Transaction decentralization

Recent U.S. federal authority rulings on the market design of capacity markets underline the importance of the issues associated with centralized, bilateral markets. Whether

transitions from a decentralized to a centralized market were rejected or approved by the authorities, they nevertheless gave rise to a number of studies.

To date, the existence of a two-sided market requires the presence of some form of demand decentralization. In their study for PJM, [Pfeifenberger et al. \(2009\)](#) clearly demonstrate the wide variety of designs available, even for decentralized demand. The authors offer an overview of possible designs for capacity markets:

- Bilateral markets (decentralized demand)
- Bilateral markets are cohabiting with centralized markets (with centralized demand).
- Centralized markets (centralized demand)

In listing the advantages of each design, the authors emphasize the importance of taking context into account for the effectiveness of each mechanism. Without going into detail, they refer to issues linked to the structure of the market studied: intensity of competition, size of players, asset financing methods, quantity and quality of information, transaction costs, and so on. In particular, their work points out that a bilateral market would generate too much risk for demand-side players, especially small ones. Transaction costs linked to negotiations and the exercise of market power would result in a sub-optimal mechanism.

In a more practical approach, applied to the Californian decentralized market, [Pfeifenberger et al. \(2009\)](#) proposes two options for the capacity market, without however showing a clear preference. They have the advantage of eliminating discrimination between new and existing capacity:

- Keep demand decentralization when contracting over long-term periods, and allow, via centralized auctions with centralized demand, the exchange of capacity for adjustment before the delivery period.
- Fully centralize the mechanism in the form of uniform auctions ahead of the delivery period. The mechanism would be centralized in a similar way to PJM.

Given the relatively low concentration of demand players, decentralization of the MISO market does not seem to raise any problems for [Newell et al. \(2010\)](#). The co-existence of a decentralized market with the possibility of participating in a centralized

demand auction would lead to the efficient operation of the mechanism. In another study on NYISO, however, [Newell et al. \(2009\)](#) argues in favor of a fully centralized market upstream of the delivery period, due in particular to the existence of a number of small players.

It should be noted, however, that the authors do not propose a rigorous, methodological analysis of each mechanism that could be implemented in each system. It would, therefore, be interesting to establish a robust link between the degree of decentralization of transactions and the efficiency of a capacity market or a *reliability options*. In this respect, it is interesting to note the analysis proposed by [Morrison \(2016\)](#) on the transitions of specific American markets. The author's arguments, which are highly critical of the centralization of capacity markets, are based on the causes of these transitions. Indeed, the arguments put forward by a number of parties are that decentralized markets would encourage abusive behavior by demand-side actors. Indeed, the latter would abuse their monopsony position and practice predatory pricing strategies. Such behavior is particularly encouraged by the possibility of carrying out bilateral transactions. Centralizing transactions would, therefore, limit such risks. However, the author stresses the specific nature of capacity markets, which would make such practices impossible. The quantity of capacity demanded, linked to the exogenous setting of the target level by the authority, cannot, in fact, be modified. Such strategies are only tenable by modifying the quantities offered on the mechanism.

5.5 The time horizon covered by the traded product

The problem of determining the product traded on capacity markets was first addressed in the previous section part of this study. We demonstrated that it is crucial to define a product that enables optimal management of the balance between supply and demand over the short term. This necessity is reflected in the introduction of penalties for non-compliance with the obligation to produce during the transaction phase.

In the context of the market design study, the product traded on capacity markets raises two significant issues that have been, and continue to be, discussed:

- The time horizon between the capacity market clearing and transaction phases. In the case of a spot market, the two phases are very close (on the order of a week or

a month). In the case of a forward market, the two phases may be several years apart.

- The length of the trading phase. A short-term contract on capacity markets refers to a transaction phase lasting one season or one year. A long-term contract, on the other hand, refers to an obligation to produce over several years.

The first issue has been extensively addressed in the practitioner literature and has been the subject of some modeling in academic articles. The central argument put forward for the introduction of forward markets is that they make it easier for producers to install additional capacity. Indeed, by enabling capacity to be sold several years in advance, markets provide new entrants with a guaranteed income, even if they have not yet built their capacity. Once the capacity revenue has been secured, the main risk for a producer lies in the time it takes to build the capacity. In the event of non-delivery, the only cost incurred by the producer is the penalty associated with the delivery phase. This design would, in fact, lower some of the barriers to entry, which, as in the classic literature, would enable greater competition in capacity markets and, hence, lower prices for consumers.

However, one of the limitations of this design also lies in the time difference between the capacity market clearing phase and the transaction phase. As electricity systems are highly uncertain, the need for adequacy, and therefore its value, is based on variations in factors over the long term, but also over the short term. As a result, an exogenous adequacy requirement valued several years in advance generates a significant risk of error, leading to under- or over-investment in the system (Harvey et al., 2013; Bhagwat et al., 2017a). To compensate for this problem, current mechanisms use sequential auctions to allow actors to adjust their positions up to a very short period before the delivery phase. Therefore, this solution reconciles two transaction modes mentioned in this section: issues linked to product formats (spot vs. forward) and matters linked to sequentiality.

In establishing this link between the two issues, some authors point to the risk of seeing the cost to consumers increase as a result of sequentiality, itself made necessary by the move to a forward market. Indeed, Harvey (2015) finds that in the PJM market with centralized demand, prices on intermediate auctions are often lower than the price of the main auction furthest from the transaction phase. This result could be linked to opportunistic arbitrage on the part of players. Sequentiality would also increase the cost of the mechanism for consumers. Indeed, the system operator buys at a high

price initially and is forced to sell at a low price during the intermediate auctions if the initial forecasts turn out to be wrong. The price difference is then passed on directly to consumers. The authors' work underlines the importance of understanding the potential strategic effects of auction sequentiality, which is itself due to the forward nature of the mechanism.

As [Bushnell et al. \(2017\)](#) points out, the question of the temporal framework is still open:

”In both cases, more careful research of the effectiveness of ISO targets and the activity in bilateral contracting would help inform the question of the optimal forward time-frame and commitment time of R.A. markets.”

To date, the only academic modeling of the various options has been carried out by [Bhagwat et al. \(2017a\)](#). In an agent-based model, the authors compare a spot capacity market with a forward capacity market (multi-year contracts) and existing capacity (annual contracts). The results of the model highlight the advantages and disadvantages of each design: the forward market allows for better investments, notably by reducing investment cycles, and has lower capacity price volatility. On the other hand, the spot capacity market offers greater flexibility in the face of exogenous shocks to the power system (such as the 2008 financial crisis). Given the mixed results, the authors do not conclude on the best market design. Moreover, the large number of assumptions used in the article makes it impossible to decide on the specific effects of each design.

The discussion of the importance of the design of capacity products also refers to the traditional issues associated with forward in commodity markets, as presented in the previous section. According to some authors, an energy market associated with a capacity market can be understood as an energy-only market in which players would have massively hedged themselves through energy-based forward contracts ([Schneider et al., 2017](#); [Leautier, 2016](#); [Harvey, 2015](#)). Indeed, contractualization via contracts would provide a strong incentive for players with production capacity to make themselves available during periods of scarcity, effectively replicating the actions of the transaction phases of capacity markets. Based on this observation, several studies have examined the effects of capacity markets on the power system. [Schneider et al. \(2017\)](#) endogenizes the demand for *forward* in the classic model of [Allaz and Vila \(1993\)](#) and shows that demand-side players do not have sufficient incentives to contract themselves via the capacity market. More classically, [Leautier \(2016\)](#) highlights the effects of contracting

via a capacity market on the exercise of market power in energy markets. This approach is often taken up in the literature associated with *reliability options*, with several authors pointing out that such contracts would be very close to a form of mandatory risk hedging, leading to the same effects as *reliability options* (Chao and Wilson, 2004; Cramton and Stoft, 2008; Ausubel and Cramton, 2010). Finally, as we described in the second section, the literature has systemically dealt with risk issues, modeling the effects of the capacity market on the system. In this respect, we note the work of de Maere d’Aertrycke et al. (2017), whose comparison of different mechanisms and designs sheds particular light on this issue.

6 Conclusion

The academic literature assumes that when a player undertakes to be available in a capacity market, it seeks to cover the cost incurred by this obligation to stay available. This opportunity cost then forms the *true value* of capacity markets. Such an approach is certainly not new to the analysis of markets as economic instruments. When we look at actor behavior through the prism of economic theory, every economic actor seeks to cover an opportunity cost. However, this approach, when applied to capacity markets, seems to be all the more relevant as it was not necessarily employed when the mechanisms were implemented and still does not appear in a number of works on that mechanism. Of course, a systematic representation of opportunity cost is also unnecessary when the modeling of actor behavior is not relevant to the analysis. Nevertheless, as these behaviors are fundamental to the functioning and the objective of capacity markets, it is pertinent to verify that their absence does not lead to contradictions in the results of the studies.

In the introduction chapter, we reviewed the economic rationales for capacity markets. Through the exchange of an obligation to be available during a specific period, these mechanisms make it possible to value investment. Consequently, their role is to prevent the system from deviating from an optimum defined in terms of adequacy. The main mechanisms designed around the world are based on the principle of competition: the value of additional capacity or demand response for the system emerges through the confrontation of a supply and a demand for capacity. The price is an additional remuneration for a number of players, inciting them to modify their investment and consumption behavior. In this chapter, we concluded that to understand how capacity

markets function and to measure their economic efficiency, it is necessary to look at how prices are formed on these mechanisms. The primary rationale behind price formation in a market is linked to the way players behave in that market. Before any study can be made of the economic relevance of a capacity market, we need to understand how players trade on this mechanism. To address this introductory problem, we first defined actor behavior conceptually and then presented the results of a literature review on how economic models have represented these actor behaviors.

This chapter aims to present the central problem of this thesis: improving understanding of the effects of specific market designs on the economic efficiency of capacity markets. More precisely, we seek to deepen the theoretical and practical link between actor behavior and the design of the mechanisms in which these same actors evolve. Initially, this observation motivated the need to establish a general analytical framework on how actors behave in capacity markets.

Secondly, it was necessary to specify and present the various issues associated with the designs of capacity markets. We, therefore, decided to limit the market design questions to those dealing with the transaction modes in the mechanisms. This framework enabled us to establish five issues: the sequentiality of transactions, the decentralization of players, the multiplicity of prices, the decentralization of transactions, and the temporal definition of the product. Although closely interrelated, these different issues have been the subject of specific studies by the academic and professional sectors. They have been applied to the energy market from the outset, both qualitatively and normatively, as well as quantitatively, using a wide variety of theoretical and empirical models. The results of this work are clear: the power system has its environment, which requires a specific representation of the sector, and the way in which transaction modes are set up significantly affects the efficiency of markets and mechanisms through changes in player behavior and price formation.

Based on this observation, we have reviewed the various analyses of these issues when applied to the mechanisms studied in this thesis. Despite the absence of models establishing clear and, above all, specific results for capacity markets, we have concluded that the issues relating to the modes of transaction on these mechanisms are significant and can clearly impact their economic efficiency. In fact, this literature review confirms the approach adopted in this thesis, which robustly establishes the link between transaction modes, actor behavior, and the efficiency of capacity markets.

Bibliography

- Abada, I., de Maere d'Aertrycke, G., Ehrenmann, A., and Smeers, Y. (2019). What models tell us about long-term contracts in times of the energy transition. *Economics of Energy & Environmental Policy*, 8(1):163–182.
- Abani, A., Hary, N., Rious, V., and Saguan, M. (2018). The impact of investors' risk aversion on the performances of capacity remuneration mechanisms. *Energy Policy*.
- Abani, A. O. (2019). *Electricity market design for long-term capacity adequacy in a context of energy transition*. PhD thesis, MINES ParisTech.
- Abani, A. O., Hary, N., Saguan, M., and Rious, V. (2016). Risk aversion and generation adequacy in liberalized electricity markets: Benefits of capacity markets. In *2016 13th International Conference on the European Energy Market (EEM)*, pages 1–5. IEEE.
- AESO (2018). Capacity market design alternative sheet. Technical report, AESO.
- Allaz, B. and Vila, J.-L. (1993). Cournot competition, forward markets and efficiency. *Journal of Economic theory*, 59(1):1–16.
- Assili, M., DB, M. H. J., and Ghazi, R. (2008). An improved mechanism for capacity payment based on system dynamics modeling for investment planning in competitive electricity environment. *Energy Policy*, 36(10):3703–3713.
- Ausubel, L. M. and Cramton, P. (2010). Using forward markets to improve electricity market design. *Utilities Policy*, 18(4):195–200.
- Bhagwat, P. C., Marcheselli, A., Richstein, J. C., Chappin, E. J., and De Vries, L. J. (2017a). An analysis of a forward capacity market with long-term contracts. *Energy policy*, 111:255–267.
- Bhagwat, P. C., Richstein, J. C., Chappin, E. J., Iychettira, K. K., and De Vries, L. J. (2017b). Cross-border effects of capacity mechanisms in interconnected power systems. *Utilities Policy*, 46:33–47.
- Bialek, S. and Unel, B. (2019). Will you be there for me the whole time? On the importance of obligation periods in design of capacity markets. *Electricity Journal*, 32(2):21–26.

- Bower, J. and Bunn, D. W. (2000). Model-based comparisons of pool and bilateral markets for electricity. *The energy journal*, 21(3):1–29.
- Brown, D. P. (2012). Non-Cooperative Entry Deterrence in a Uniform Price Multi-Unit Capacity Auction. .
- Brown, D. P. (2018a). Capacity payment mechanisms and investment incentives in restructured electricity markets. *Energy Economics*, 74:131–142.
- Brown, D. P. (2018b). The effect of subsidized entry on capacity auctions and the long-run resource adequacy of electricity markets. *Energy Economics*, 70:205–232.
- Bublitz, A., Keles, D., Zimmermann, F., Fraunholz, C., and Fichtner, W. (2019). A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms. *Energy Economics*, 80:1059–1078.
- Bushnell, J., Flagg, M., and Mansur, E. (2017). Capacity Markets at a Crossroads. *Energy Institute at Haas*. .
- Cepeda, M. and Finon, D. (2011). Generation capacity adequacy in interdependent electricity markets. *Energy Policy*, 39(6):3128–3143.
- Cepeda, M. and Finon, D. (2013). How to correct for long-term externalities of large-scale wind power development by a capacity mechanism? *Energy Policy*, 61:671–685.
- Chao, H.-p. and Wilson, R. (2004). Resource adequacy and market power mitigation via option contracts. *Electric Power Research Institute*.
- Cigré (2016). Capacity Mechanisms : Needs , Solutions and state affairs. Technical report, International Council on Large Electric Systems (Cigré), Working Group C5.17.
- Cramton, P. (2017). Electricity market design. *Oxford Review of Economic Policy*, 33(4):589–612.
- Cramton, P., Ockenfels, A., and Stoft, S. (2013). Capacity Market Fundamentals. *Economics of Energy & Environmental Policy*.
- Cramton, P. and Stoft, S. (2008). Forward reliability markets: Less risk, less market power, more efficiency. *Utilities Policy*, pages 194–201.
- Creti, A. and Fabra, N. (2007). Supply security and short-run capacity markets for electricity. *Energy Economics*, 29(2):259–276.

- de Maere d’Aertrycke, G., Ehrenmann, A., and Smeers, Y. (2017). Investment with incomplete markets for risk: The need for long-term contracts. *Energy Policy*, 105(June 2016):571–583.
- De Maere d’Aertrycke, G., Ehrenmann, A., and Smeers, Y. (2017). Investment with incomplete markets for risk: The need for long-term contracts. *Energy Policy*, 105:571–583.
- De Sisternes, F. J. and Parsons, J. E. (2016). The impact of uncertainty on the need and design of capacity remuneration mechanisms in low-carbon power systems. *MIT Center for Energy and Environmental Policy Research (February 2016)*.
- De Vries, L. and Heijnen, P. (2008). The impact of electricity market design upon investment under uncertainty: The effectiveness of capacity mechanisms. *Utilities Policy*, 16(3):215–227.
- DECC (2012). Electricity market reform – capacity market. Technical report, DECC.
- Ehrenmann, A. and Smeers, Y. (2011a). Generation capacity expansion in a risky environment: a stochastic equilibrium analysis. *Operations research*, 59(6):1332–1346.
- Ehrenmann, A. and Smeers, Y. (2011b). *Stochastic equilibrium models for generation capacity expansion*. Springer.
- Elberg, C. and Kranz, S. (2013). Capacity Mechanisms and Effects on Market Structure - Preliminary Version -. , pages 1–21.
- Fabra, N., von der Fehr, N.-H., and Harbord, D. (2006). Designing electricity auctions. *The RAND Journal of Economics*, 37(1):23–46.
- Fabra, N., von der Fehr, N.-H. M., and de Frutos, M.-Á. (2011). Market design and investment incentives. *The Economic Journal*, 121(557):1340–1360.
- Fan, L., Norman, C. S., and Patt, A. G. (2012). Electricity capacity investment under risk aversion: A case study of coal, gas, and concentrated solar power. *Energy Economics*.
- Feng, D. and Xu, Z. (2009). Risk analysis of volume cheat strategy in a competitive capacity market. In *2009 IEEE Power & Energy Society General Meeting*, pages 1–7. IEEE.
- Franco, C. J., Castaneda, M., and Dyner, I. (2015). Simulating the new british electricity-market reform. *European Journal of Operational Research*, 245(1):273–285.

- Giocoli, N. (2005). The escape from conjectural variations: the consistency condition in duopoly theory from bowley to fellner. *Cambridge Journal of Economics*, 29(4):601–618.
- Hach, D. and Spinler, S. (2016). Capacity payment impact on gas-fired generation investments under rising renewable feed-in - A real options analysis. *Energy Economics*, 53:270–280.
- Harbord, D. and Pagnozzi, M. (2014). Britain’s electricity capacity auctions: lessons from colombia and new england. *The Electricity Journal*, 27(5):54–62.
- Harvey, S. M. (2015). Fundamentals of Capacity Market Design and Performance. In *EUCI Capacity Markets Conference*.
- Harvey, S. M., Hogan, W. W., and Pope, S. L. (2013). Evaluation of the New York Capacity Market. Technical report, Cambridge.
- Hary, N., Rious, V., and Saguan, M. (2016). The electricity generation adequacy problem: Assessing dynamic effects of capacity remuneration mechanisms. *Energy Policy*, 91:113–127.
- Hobbs, B. F., Hu, M. C., Iñón, J. G., Stoft, S. E., and Bhavaraju, M. P. (2007a). A dynamic analysis of a demand curve-based capacity market proposal: The PJM reliability pricing model. *IEEE Transactions on Power Systems*, 22(1):3–14.
- Hobbs, B. F., Hu, M.-C., Iñón, J. G., Stoft, S. E., and Bhavaraju, M. P. (2007b). A dynamic analysis of a demand curve-based capacity market proposal: The pj m reliability pricing model. *IEEE Transactions on Power Systems*, 22(1):3–14.
- Höschle, H. (2018). Capacity Mechanisms in Future Electricity Markets. *Capaciteitsmechanismen in toekomstige elektriciteitsmarkten*.
- Joung, M., Baldick, R., and Kim, J. (2009). Strategic behavior in electricity capacity markets. *Proceedings of the 42nd Annual Hawaii International Conference on System Sciences, HICSS*, pages 1–6.
- Le Cadre, H. and Soubra, M. (2013). *Designing Rules for the Capacity Market*. PhD thesis, Chaire Modélisation prospective au service du développement durable.
- Leautier, T.-O. (2016). The visible hand: ensuring optimal investment in electric power generation. *The Energy Journal*, 37(2).

- Lynchab, M. Á. and Devinea, M. T. (2017). Investment vs. refurbishment: examining capacity payment mechanisms using stochastic mixed complementarity problems. *The Energy Journal*, 38(2):27–52.
- Mastropietro, P., Herrero, I., Rodilla, P., and Batlle, C. (2016). A model-based analysis on the impact of explicit penalty schemes in capacity mechanisms. *Applied Energy*, 168:406–417.
- Menezes, F. M. and Monteiro, P. K. (2004). *An introduction to auction theory*. OUP Oxford.
- Meunier, G. (2013). Risk aversion and technology mix in an electricity market. *Energy Economics*, 40:866–874.
- Meyer, R. and Gore, O. (2014). Cross-border effects of capacity mechanisms: Do uncoordinated market design policies countervail the goals of European market integration? *Energy Econ.*, pages 9–20.
- Morrison, J. (2016). Capacity markets: a path back to resource adequacy. *Energy LJ*, 37:1.
- Newell, S., Bhattacharyya, A., and Madjarov, K. (2009). Cost-benefit analysis of replacing the nyiso’s existing icap market with a forward capacity market. Technical report, The Brattle Group.
- Newell, S., Spees, K., and Hajos, A. (2010). Midwest iso’s resource adequacy construct an evaluation of market design elements. Technical report, The Brattle Group.
- Olsina, F., Garcés, F., and Haubrich, H.-J. (2006). Modeling long-term dynamics of electricity markets. *Energy Policy*, 34(12):1411–1433.
- Peluchon, B. (2019). Working Paper: Market design and the cost of capital for generation capacity investment. *Chaire European Electricity Markets.*, pages 1–36.
- Petitot, M. (2016). *Long-term dynamics of investment decisions in electricity markets with variable renewables development and adequacy objectives*. PhD thesis, PSL - Paris Dauphine University.
- Pfeifenberger, J. P., Spees, K., and Newell, S. (2012). Resource adequacy in california options for improving efficiency and effectiveness. Technical report, The Brattle Group.

- Pfeifenberger, J. P., Spees, K., and Schumacher, A. (2009). A comparison of pjm’s rpm with alternative energy and capacity market designs. Technical report, The Brattle Group.
- PowerAuctions (2016). Evaluation of auctions and capacity market design for electric power. Technical report, PowerAuctions.
- Pöyry (2015). Decentralised reliability options - securing energy markets. Technical report, Pöyry.
- RTE (2014). Rapport d’accompagnement de la proposition de règles du mécanisme de capacité. Technical report, RTE.
- Salant, D. and Stoddard, R. (2008). Auction design for capacity markets. In *Fifth Conference on “The Economics of Energy Markets”*.
- Schneider, I., Botterud, A., and Roozbehani, M. (2017). Forward contracts and generator market power: How externalities reduce benefits in equilibrium. In *Riding the Energy Cycles, 35th USAEE/IAEE North American Conference, Nov 12-15, 2017*. International Association for Energy Economics.
- Schwenen, S. (2014). Market design and supply security in imperfect power markets. *Energy economics*, 43:256–263.
- Schwenen, S. (2015). Strategic bidding in multi-unit auctions with capacity constrained bidders: the new york capacity market. *The RAND Journal of Economics*, 46(4):730–750.
- Scoufflaire, C. (2019). *Capacity remuneration mechanisms: analytical assessment of contemporary experiences and lessons for the future design of electricity markets*. PhD thesis, PSL - Paris Dauphine University.
- Teirilä, J. and Ritz, R. A. (2018). Strategic behavior in a capacity market ? .
- Teirilä, J. and Ritz, R. A. (2019). Strategic behaviour in a capacity market? the new irish electricity market design. *The Energy Journal*, 40(The New Era of Energy Transition).
- Twomey, P., Green, R., Neuhoff, K., and Newbery, D. (2005). A review of the monitoring of market power. *Massachusetts Institute of Technology Center for Energy and Environmental Policy Research*. <http://ideas.repec.org/p/cam/camdae/0504.html>. Consultada el el, 29.

Wilson, J. F. (2010). Forward Capacity Market CONEfusion. *Electricity Journal*, 23(9):25–40.

Winzer, C. (2013). Robustness of various capacity mechanisms to regulatory errors. *International Conference on the European Energy Market, EEM*, pages 1–7.

Appendix B

Appendix - Chapter 1

Contents

| | | |
|---|------------------------------------|-----|
| 1 | Proof of Proposition 1.2 | 265 |
| 2 | Proof of Proposition 1.3 | 267 |
| 3 | Proof of Proposition 1.5 | 269 |
| 4 | Proof of Lemma 1.1 | 271 |
| 5 | Proof of Proposition 1.7 | 273 |
| 6 | Proof of Lemma 1.4 | 274 |

1 Proof of Proposition 1.2

We follow the demonstration from [Matthäus et al. \(2021\)](#) and adapt it to the capacity market framework. The demonstration relies on the existence of a self-financing strategy Y_t between a bond process $D\beta_t = r\beta_t dt$ and the inframarginal rent process π_t and on the assumption of an arbitrage-free market. Equating the coefficients of the self-financing strategy results in the following equation

$$-rb^{opt}(\pi_t, t) + b_t^{opt}(\pi_t, t) + r\pi_t b_x^{opt}(\pi_t, t) + \frac{1}{2}\sigma^2 \pi_t^2 b_{xx}^{opt}(\pi_t, t) = 0$$

It implies that we need to solve the following PDE.

$$-rb^{opt}(w, t) + b_t^{opt}(x, t) + rx b_x^{opt}(x, t) + \frac{1}{2}\sigma^2 x^2 b_{xx}^{opt}(x, t) = 0 \quad (\text{B.1})$$

on the region $(x, t) \in (0, \text{inf}) \times [0, T]$ with boundary condition $W(x, t) = \max(c^{om} - \pi_t, 0)$.

To solve this PDE, we introduce an equivalent risk-neutral measure \mathbb{Q} with $d\mathbb{Q} = Z_{\bar{T}} d\mathbb{P}$. Here \mathbb{P} denotes the natural measure and $dZ_t = (\mu - r)\sigma^{-1} Z_t d\beta$. Girsanov theorem yields $d\beta_t = -(\mu - r)\sigma^{-1} dt + d\beta_t^{\mathbb{Q}}$ and therefore $d\pi_t = r\pi_t dt + \sigma\pi_t d\beta_t^{\mathbb{Q}}$. We can solve the PDE by applying the Feynman-Kac formula. For tractability, we assume that $t = 0$ and therefore $\bar{T} = n^d$. A solution is given by

$$\begin{aligned} b_{opt}(\pi_t, 0) &= \mathbb{E}^* \left(- \int_0^{n^d} e^{rs} ds \max(c^{om} - \pi_t, 0) \right) \\ &= e^{rn^d} \left(c^{om} \int_{-\infty}^{c^{om}} dF(\pi_t) - \int_{-\infty}^{c^{om}} \pi_t dF(\pi_t) \right) \end{aligned}$$

The rest of the demonstration relies on computing the integrals. For that purpose, we note that $Y \mathcal{N}(\omega, \zeta)$ and $X = e^Y$, then the distribution of X is

$$F_x(x) = \phi\left(\frac{\ln(x) - \omega}{\zeta}\right)$$

Where $\phi(\cdot)$ is the cumulative density function of the standard normal distribution. Furthermore, the curtailed expected value of X is given by

$$E[X|X < x] = e^{\omega + \frac{\zeta^2}{2}} \phi\left(\frac{\ln(x) - \omega - \zeta^2}{\sigma}\right)$$

and

$$\ln(\pi_{n^d}) \sim \mathcal{N}\left(\ln(\pi_0) + \left(r - \frac{\sigma^2}{2}\right)n^d, \sigma^2 n^d\right)$$

under the risk-neutral measure \mathbb{Q}

Using the identity $\phi(-x) = 1 - \phi(x)$, we have for the first integral :

$$\int_{-\infty}^{c^{om}} dF(\pi_{n^d}) = F_{\pi_{n^d}}(c^{om}) = \phi\left(-\frac{\ln(\pi_0) - \ln(c^{om}) + \left(r - \frac{\sigma^2}{2}\right)n^d}{\sigma\sqrt{n^d}}\right) = \phi(z + \sigma\sqrt{n^d})$$

And we define $z := -\frac{\ln(\pi_0) - \ln(c^{om}) + \left(r + \frac{\sigma^2}{2}\right)n^d}{\sigma\sqrt{n^d}}$. Which gives :

$$\int_{-\infty}^{c^{om}} dF(\pi_{n^d}) = \phi(z + \sigma\sqrt{n^d})$$

For the second integral :

$$\int_{-\infty}^{c^{om}} dF(\pi_{n^d}) = \mathbb{E}(\pi_{n^d} | \pi_{n^d} < c^{om}) = \pi_0 e^{rn^d} \phi\left(-\frac{\ln(\pi_0) - \ln(c^{om}) + \left(r + \frac{\sigma^2}{2}\right)n^d}{\sigma\sqrt{n^d}}\right)$$

Which gives:

$$\int_{-\infty}^{c^{om}} dF(\pi_{n^d}) = \pi_0 e^{rn^d} \phi(z)$$

With the expression of the integrals, we can express the option value associated with the possibility of closing to avoid the fixed costs:

$$b^{opt}(\pi_0) = -\pi_0 \phi(z) + e^{-rn^d} (c^{om} \phi(z + \sigma\sqrt{n^d}))$$

2 Proof of Proposition 1.3

We use the initial paper by Levy (1992), which approximates the distribution of the basket option by a log-normal distribution, and we apply it to the unique framework of the capacity market.

Let $\pi(t)$ be the inframarginal rent received at the time. We suppose that the sum of the inframarginal rent is determined on the interval $[\bar{T}, \bar{T} + n^t]$, which represents the transaction phase. We define the continuous sum as follows, and we assume that $\bar{T} = 0$:

$$\Pi_0 = \int_0^{n^t} \pi(t) dt$$

We look at characterizing the value of the *modified* basket put option. Using our capacity market framework, with notably the strike price equal to the sum of actualized periodic fixed cost C^{om} , it can be defined as :

$$P[\pi(t), \Pi_0] = e^{-rn^d} \mathbb{E}_0^*(\max(C^{om} - \Pi_0))$$

With E_0^* the expectation operator defined in the model section, which implies that under the risk-adjusted density function, the inframarginal rent process can be described by $d\pi(t) = r\pi(t)dt + \sigma\pi_t dZ^*(t)$. For any value $t > 0$ we know that the value $\ln(\pi(t))$ is normally distributed, with mean $\ln(\pi_0) + (r - \frac{\sigma^2}{2})t$ and standard deviation $\sigma\sqrt{t}$.

The demonstration continues by assuming that the sum of log-normally distributed values Π_0 is indeed following a log-normal distribution, namely that $\ln(\Pi_0)$ is normally distributed with an unknown mean m and variance v^2 . Therefore, we use the moment-generating function to determine those parameters. We define this function as $\Phi_x(k)$ with

$$\Phi_x(k) = \mathbb{E}_0^*(\Pi_0^k) = e^{km + \frac{v^2}{2}k^2}$$

This expression allows us to consider a system of two equations with two unknowns, with the equations being the first two moments and the unknowns being m and v^2 . Solving the system allows the following expressions.

$$m = 2\ln(\mathbb{E}_0^*[\Pi_0]) - \frac{1}{2}\ln(\mathbb{E}_0^*[\Pi_0^2])$$

$$v^2 = \ln(\mathbb{E}_0^*[\Pi_0^2]) - 2\ln(\mathbb{E}_0^*[\Pi_0])$$

With $\mathbb{E}_0^*[\Pi_0]$ and $\mathbb{E}_0^*[\Pi_0^2]$ being the first and second moment of Π_0 . Following our assumption regarding the process of the inframarginal rent, we can find a closed-form expression for the two moments. For the first moment, namely the mean of the sum, we can initially define it as follows:

$$\mathbb{E}_0^*[\Pi_0] = \int_0^{n^t} \pi(t) dt$$

Which gives :

$$\mathbb{E}_0^*[\Pi_0] = \pi_0 \int_0^{n^t} e^{rt} dt$$

For the second moment, we use the initial expression for two variables following a geometric Brownian motion process, say $\pi(t_1)$ and $\pi(t_2)$. In this case, we have $\mathbb{E}_0^*[\pi(t_1)\pi(t_2)] = \pi_0^2 e^{r(t_1+t_2)+\sigma^2 t_1}$. Then, we can expand the expression to a continuous framework and to the sum of the inframarginal rent, which gives :

$$\mathbb{E}_0^*[\Pi_0^2] = \pi_0^2 \int_0^{n^t} \int_0^{n^t} e^{r(t+s+n^d)+(s+n^d)\sigma^2} dt ds$$

When assuming that Π_0 does follow a log-normal distribution, and with a closed-form expression for m and v^2 , we can evaluate the put option $P[\pi_t, \Pi_0]$ using the standard finance theory as shown in the proof of Proposition 1.2:

$$b^{opt}(\pi_0, \Pi_0) = P[\pi_0, \Pi_0] = -\pi_0 n^t \phi(z) + C^{om} \phi(z + v)$$

Where :

$$z = -\frac{m - \ln(C^{om} \int_0^{n^t} e^{-rt}) + v^2}{v}$$

Note that we do not include any discounting factor for the inframarginal as it is already done using $n^t \pi_0$. Compared to the initial basket option, which compares asset price at the same period in time, in our framework, the option is exercised only with respect to the sum of the expected discounted inframarginal rent received during the transaction phase.

3 Proof of Proposition 1.5

For the net present value, the derivative of the bid with respect to n^t is

$$\frac{\partial b^{npv}}{\partial n^t} = -n^t + c^{om} e^{-r(n^t+n^d)}$$

Therefore, the threshold for the sign of n^t on the net present value bid is given by the first order condition such that :

$$-n^t + c^{om} e^{-r(n^t+n^d)} = 0$$

Which implies :

$$c^{om} = \pi_0 e^{r(nd+nt)}$$

For the real option bid, the derivative of the bid with respect to n^t is :

$$\frac{\partial b^{opt}}{\partial n^t} = -\pi_0(\phi(z) + n^t \frac{\partial z}{\partial n^t} \varphi(z)) + c^{om} e^{-rn^d} (e^{-rn^t} \phi(z+v) + \int_0^{n^t} e^{-rt} dt \varphi(z+v))$$

When rearranged :

$$\frac{\partial b^{opt}}{\partial n^t} = -\pi_0(\phi(z) + n^t \frac{\partial z}{\partial n^t} \varphi(z)) + C^{om} (S_1 \phi(z+v) + \varphi(z+v) \frac{\partial z+v}{\partial n^t})$$

with $C^{om} = e^{-rn^d} c^{om} \int_0^{n^t} e^{-rt} dt$ and $S_1 = e^{-rn^t} / \int_0^{n^t} e^{-rt} dt$.

The Cdf ratio and the Df ratio conditions are given by rearranging the equation and by, respectively, the first and second term in brackets :

$$\frac{\partial b^{opt}}{\partial n^t} = [C^{om} S_1 \phi(z+v) - \pi_0(\phi(z))] + \left[\frac{\partial z}{\partial n^t} (C^{om} \varphi(z+v) - n^t \pi_0 \varphi(z)) \right] + C^{om} \varphi(z+v) \frac{\partial v}{\partial n^t}$$

Excluding the sign of $\frac{\partial z}{\partial n^t}$ the first two terms are positive if and only if : Cdf ratio : $\frac{S_1 C^{om}}{\pi_0} \geq R_0 = \frac{\phi(z)}{\phi(z+v)}$ Df ratio : $\frac{C^{om}}{n^t \pi_0} \geq R_1 = \frac{\varphi(z)}{\varphi(z+v)}$

The derivative $\frac{\partial v}{\partial n^t}$ is always positive and is equal to

$$\frac{r e^{nt(r+\sigma)} + \sigma e^{nt(r+\sigma)} - r e^{nt r} - \sigma e^{nt(2r+\sigma)'}}{2 \sqrt{v^2} (e^{nt(r+\sigma)} + e^{nt r} - e^{nt(2r+\sigma)} - 1)}$$

Therefore, we need a third condition given by the sign of $\frac{\partial z}{\partial n^t}$. It can be expressed as follows:

$$\frac{\partial z}{\partial n^t} = -\frac{\frac{\partial v^2}{\partial n^t} + 2r}{2 \sqrt{v^2}} - \frac{\frac{\partial v^2}{\partial n^t} \left(\ln \left(-\frac{\text{com}(e^{-nt r} - 1)}{r} \right) - \frac{\ln(V^2 p_0)}{2} \right)}{2 v^3}$$

The sign of the derivative is given when equating the equation to 0, which gives the following condition on the fixed cost for the derivative to be positive:

$$c^{com} \int_0^{n^t} e^{-rt} dt \geq \sqrt{\pi_0} V e^{v^2(2r \frac{\partial v^2}{\partial n^t} - 1)}$$

The limits of the derivative at its extreme are found by analyzing the behavior of $\phi(z)$, $\phi(z+v)$, $\varphi(z)$ and $\varphi(z+v)$. Note first that $z \rightarrow 0$ when $n^t \rightarrow +\infty$, while $v \rightarrow +\infty$ when $n^t \rightarrow +\infty$.

Then, the density function converges towards 0 when $z \rightarrow 0$, $z+v \rightarrow +\infty$. For the cumulative density function : $\phi(z+v) \rightarrow 1$ when $z+v \rightarrow +\infty$, while $\phi(z) \rightarrow 0$. This implies the first result.

Concerning the case of $n^t \rightarrow 0$, depending on the initial value of π_0 with respect to c^{om} , the value of $\phi(z)$ can either converge to 0 or to 1 when $n^t \rightarrow 0$. Indeed, recall that

the sign of the derivative of z can either be positive or negative. However, in both cases, the value is either 0 or a positive value.

4 Proof of Lemma 1.1

The results of the proposition follow directly from the derivative of the bid with respect to the variables.

For π_0 :

$$\frac{\partial b^{opt}}{\partial \pi_0} = -n^t(\phi(z) + \pi_0 \frac{\partial z}{\partial \pi_0} \varphi(z)) + C^{om} \varphi(z+v) \frac{\partial z+v}{\partial \pi_0}$$

Which gives when rearranged :

$$\frac{\partial b^{opt}}{\partial \pi_0} = -n^t \phi(z) + \frac{\partial z}{\partial \pi_0} (C^{om} \varphi(z+v) - n^t \pi_0 \varphi(z)) + C^{om} \varphi(z+v) \frac{\partial v}{\partial \pi_0}$$

We found that : (i) $\frac{\partial v}{\partial \pi_0}$ is null and that (ii) $\frac{\partial z}{\partial \pi_0}$ is always negative as it is equal to :

$$\frac{\partial z}{\partial \pi_0} = -\frac{1}{p_0 \sqrt{v^2}}$$

Under the condition that $\frac{C^{om}}{n^t \pi_0} \geq R_1 = \frac{\varphi(z)}{\varphi(z+v)}$ (which is the Df ratio) then the derivative is always negative.

For c^{om} :

$$\frac{\partial b^{opt}}{\partial c^{om}} = -n^t \pi_0 \frac{\partial z}{\partial c^{om}} \varphi(z) + e^{-rn^d} \int_0^{n^t} e^{-rt} dt \phi(z+v) + \frac{\partial z+v}{\partial c^{om}} C^{om} \varphi(z+v)$$

Which gives when rearranged :

$$\frac{\partial b^{opt}}{\partial c^{om}} = e^{-rn^d} \int_0^{n^t} e^{-rt} dt \phi(z+v) + \frac{\partial z}{\partial c^{om}} (C^{om} \varphi(z+v) - n^t \pi_0 \varphi(z)) + \frac{\partial v}{\partial c^{om}} C^{om} \varphi(z+v)$$

We found that : (i) $\frac{\partial v}{\partial c^{om}}$ is null and that (ii) $\frac{\partial z}{\partial c^{om}}$ is always positive as it is equal to :

$$\frac{\partial z}{\partial c^{om}} = \frac{1}{com \sqrt{v^2}}$$

Under the condition that $\frac{C^{om}}{n^t \pi_0} \geq R_1 = \frac{\varphi(z)}{\varphi(z+v)}$ (which is the Df ratio) then the derivative is always positive.

For n^d :

$$\frac{\partial b^{opt}}{\partial n^d} = -\pi_0 n^t \frac{\partial z}{\partial n^d} \varphi(z) + C^{om} (-r \phi(z+v) + \frac{\partial z+v}{\partial n^d} \varphi(z+v))$$

Which gives when rearranged :

$$\frac{\partial b^{opt}}{\partial n^d} = C^{om} (\varphi(v+v) \frac{\partial v}{\partial n^d} - r \phi(z+v)) + \frac{\partial z}{\partial n^d} (C^{om} \varphi(z+v) - n^t \pi_0 \varphi(z))$$

The conditions on the derivative of the bid with respect to n^d come straightforwardly. Note that the condition on the risk-free rate and the fixed costs are given by the respective derivative of v and z with respect to n^d

$$\frac{\partial v}{\partial n^d} = -\frac{r - \sigma'}{2 \sqrt{v^2}}$$

and

$$\frac{\partial z}{\partial n^d} = \frac{\left(\ln \left(-\frac{com(e^{-nt} r - 1)}{r} \right) - \frac{\ln(V)}{2} \right) \left(2r - \frac{e^{-nd} \sigma (r + \sigma) (Vr + V\sigma)}{M p_0 (e^{nt} r + nt \sigma - 1)} \right)}{2 v^3} - \frac{e^{-nd} \sigma (r + \sigma) (Vr + V\sigma)}{2 M p_0 (e^{nt} r + nt \sigma - 1) \sqrt{v^2}}$$

5 Proof of Proposition 1.7

Recall that :

$$\frac{\partial b^{npv}}{\partial n^t} = -n^t + c^{om} e^{-r(n^t+n^d)}$$

And that :

$$\frac{\partial b^{opt}}{\partial n^t} = -\pi_0(\phi(z) + n^t \frac{\partial z}{\partial n^t} \varphi(z)) + C^{om}(S_1 \phi(z+v) + \varphi(z+v) \frac{\partial z+v}{\partial n^t})$$

Therefore, the derivative of the flexibility, which is equal to the derivative of the difference between the two previous equations:

$$\frac{\partial \Gamma}{\partial n^t} = -\pi_0(\phi(z) + n^t \frac{\partial z}{\partial n^t} \varphi(z)) + C^{om}(S_1 \phi(z+v) + \varphi(z+v) \frac{\partial z+v}{\partial n^t}) + n^t - c^{om} e^{-r(n^t+n^d)}$$

When rearranged :

$$\frac{\partial \Gamma}{\partial n^t} = -\pi_0((\phi(z) - 1) + n^t \frac{\partial z}{\partial n^t} \varphi(z)) + C^{om}(S_1(\phi(z+v) - 1) + \varphi(z+v) \frac{\partial z+v}{\partial n^t})$$

The conditions and the ratio in the proposition stem directly from equating :

$$\frac{\partial \Gamma}{\partial n^t} = 0$$

And differentiate from the cases where the denominator of the ratio is positive or negative. Which is given by $S_1(\phi(z+v) - 1) + \frac{\partial z+v}{\partial n^t} \varphi(z+v) > 0$ or by $S_1(\phi(z+v) - 1) + \frac{\partial z+v}{\partial n^t} \varphi(z+v) < 0$.

6 Proof of Lemma 1.4

We modify the results of the proof of Proposition 1.2 and apply it directly to the proof of Proposition 1.3. First, note that the new payoff of the basket option is now equal to $\max(C^{om} - \Pi_0, -P)$, with P the costs associated with the closing decision, which are the penalty and the closing costs. In this case, we need to solve the PDE of equation B.1 on the same region but with boundary condition $W(x, t) = \max(C^{om} - \Pi_t, P)$. We can solve the PDE by applying the Feynman-Kac formula. A solution is given by

$$\begin{aligned} b_{opt}(\Pi_t, 0) &= \mathbb{E}^* \left(- \int_0^{n^d} e^{rs} ds \max(C^{om} - \Pi_0) \right) \\ &= e^{rn^d} \left(C^{om} \int_{-\infty}^{C^{om}+P} dF(\Pi_t) - \int_{-\infty}^{C^{om}+P} \Pi_t dF(\Pi_t) - P \int_{C^{om}+P}^{+\infty} dF(\Pi_t) \right) \end{aligned}$$

First, note that for the two integrals, their expression is close to the one in the proof of Proposition 1.2 and 1.3. Namely, we add the closing cost to the periodic fixed cost in the value z :

$$e^{rn^d} \left(C^{om} \int_{-\infty}^{C^{om}+P} dF(\Pi_t) - \int_{-\infty}^{C^{om}+P} \Pi_t dF(\Pi_t) \right) = -\pi_0 n^t \phi(z) + e^{-rn^d} C^{om} \phi(z + v)$$

$$\text{with } z := -\frac{m - \ln(C^{om} \int_0^{n^t} e^{-rt} dt + P) + v^2}{v}.$$

For the third integral, recall that :

$$\int_{C^{om}+P}^{+\infty} dF(\Pi_t) = 1 - \int_{-\infty}^{C^{om}+P} dF(\Pi_t) = 1 - \phi(z + v)$$

Therefore :

$$b_{opt}(\Pi_t, 0) = -\pi_0 n^t \phi(z) + e^{-rn^d} (C^{om}) \phi(z + v) - P(1 - \phi(z + v))$$

When rearranged :

$$b_{opt}(\Pi_t, 0) = -\pi_0 n^t \phi(z) + e^{-rn^d} ((C^{om} + P)\phi(z + v) - P)$$

Bibliography

Levy, E. (1992). Pricing european average rate currency options. *Journal of International Money and Finance*, 11(5):474–491.

Matthäus, D., Schwenen, S., and Wozabal, D. (2021). Renewable auctions: Bidding for real options. *European Journal of Operational Research*, 291(3):1091–1105.

Appendix C

Appendix - Chapter 2

Contents

| | | |
|----|---|-----|
| 1 | Market equilibrium with Cournot | 279 |
| 2 | Proof of Lemma 2.1 | 285 |
| 3 | Proof of Proposition 2.1 | 285 |
| 4 | Proof of Lemma 2.2 | 286 |
| 5 | Proof of Lemma 2.3 | 288 |
| 6 | Proof of Proposition 2.2 | 291 |
| 7 | Proof of Lemma 2.4 | 295 |
| 8 | Proof of Lemma 2.5 | 296 |
| 9 | Proof of Proposition 2.3 | 299 |
| 10 | Proof of Lemma 2.6 | 302 |
| 11 | Proof of Proposition 2.4 | 304 |
| 12 | Proof of Corollary 2.1 | 307 |
| 13 | Proof of Lemma 2.7 | 309 |
| 14 | Proof of Proposition 2.5 | 311 |

Notation guide: For readability, we sometimes drop the references to variables in the expression of functions and their derivatives.

1 Market equilibrium with Cournot

We extend the findings of the endogenous and exogenous regimes to take into account imperfect competition in the retail market. Therefore, we do not directly implement the retailer market share allocation. We denote $mp(q)$ the markup associated with the market power in the retail market such that $mp(q) = -\frac{q}{n}p_q > 0$ ¹. Similarly to the other cases, we denote $k_{0,n}^*$ the investment level that maximize the expected social welfare $W_{0,n}(k)$ under imperfect competition such that $k_{0,n}^* = \{k : \phi_{0,n}(k) = r\}$, with :

$$\phi_{0,n}(k) = \int_{s_{0,n}}^{\bar{s}} (p(k, s) - c)dF(s)$$

With $s_{0,n}$ the threshold value such that this is the first state of the world for which the wholesale price adjusted by retailers' market power equals the marginal cost: $p(k, s_{0,n}) - mp(k) = c$. The initial general assumptions ensure that the expected social welfare is concave under the *Cournot* competition. The market equilibrium, such as the expected marginal revenue, equals the marginal cost of providing an additional investment: $k_0^n = \{k : \phi_0^n(k) = r\}$ with:²

$$\phi_0^n(k) = \int_{s_{0,n}}^{\bar{s}} (p(k, s) - mp(k) - c)dF(s)$$

The following lemma sums up the results: we find that market power in the retail market lowers the investment level beyond the market power's direct effect. The market investment level is different from the optimal investment level even when maximizing the welfare function given the market power in the retail market.

Lemma C.1. *For every $n \in [2, \infty[$, imperfect competition in the retail market leads (i) to a lower first best capacity level compared with the optimal investment level $k_0^* \geq k_{0,n}^*$ (ii) to a lower market equilibrium in terms of investment level $k_{0,n}^* \geq k_0^n$. The optimal*

¹We simplify $\frac{\partial mp(q)}{\partial q} = mp_q$ and $\frac{\partial mp(q)}{\partial n} = mp_n(q)$.

²Recall that under perfect competition, the relation between the wholesale price and retail price is given by $p^s(q, s) = p(q, s) + \frac{q}{n}p_q$

capacity payment z_n is equal to the expected markup of retailers in the retail market during peak periods:

$$z_n(k) = \int_{s_{0,n}}^{\bar{s}} mp(k) dF(s) \quad (\text{C.1})$$

Proof. The welfare function is given by

$$W(k) = \int_0^{s_0} \int_0^{q_0} (p(q, s) - c) dq dF(s) + \int_{s_0}^{+\infty} \int_0^k (p(q, s) - c) dq dF(s) - rk \quad (\text{C.2})$$

The first order condition gives the optimal investment level:

$$\int_{s_0}^{+\infty} \int_0^k (p(q, s) - c) dq dF(s) = r \quad (\text{C.3})$$

We define the inframarginal rent under Cournot competition as follows :

$$\phi(k) = \int_{s_0}^{+\infty} (p^s(k, s) - c) dF(s) \quad (\text{C.4})$$

The market investment level is equal to the optimal investment level only if the solution following equality has the same solution as with the first-order condition of the welfare function:

$$\phi(k) + z_n(k) = r \quad (\text{C.5})$$

Which is the case if and only if :

$$z_n(k) = \phi(k) - \int_{s_0}^{+\infty} \int_0^k (p(q, s) - c) dq dF(s) \quad (\text{C.6})$$

Therefore, the optimal payment is equal to the expected markup of retailers.

$$z_n(k) = \int_{s_0}^{+\infty} mp(k)dF(s) \quad (\text{C.7})$$

□

With imperfect competition, the first best investment level is lower due to a shift in the occurrence of on-peak periods. It can be showed by comparing the expression $\phi_0(k)$ and $\phi_{0,n}(k)$ due to the presence of the threshold value $s_{0,n}$. Similarly, the inefficient market equilibrium is also lower due both to the threshold and to a lower expected price $p(q, s) - mp(q)$ as shown in the expression of $\phi_0^n(k)$.

We illustrate this result in Figure C.2. We show the expected social welfare for different values of capacity and the cumulative effect of imperfect competition in the system. The black curve represents the case with perfectly competitive retailers. The results in the of Lemma C.1 demonstrate that with imperfect competition in the retail market : (i) the expected social welfare is lower, which is represented by the blue curves and implies a lower first-best investment level, (ii) A decrease of imperfect competition (via an increase of the number of retailers $n' > n$ with $n = 2$ and $n' = 3$) increase both the investment level that maximizes the welfare and the welfare at this level. This result is in agreement with the literature on market power. To see this, the derivative of the expected social welfare in this case can be written as follows:

$$\frac{\partial W_{0,n}(k)}{\partial n} = \int_0^{s_n} \frac{\partial q_{0,n}}{\partial n} mp(q_{0,n})dF(s)$$

This captures the fact that as n varies, it changes the aggregate quantity offered by retailers only during the off-peak periods. Indeed, the on-peak periods are such that the total quantity is equal to k ; hence, a change of n does not affect welfare. Using the definition of $q_{0,n}$, we have:

$$\frac{\partial q_{0,n}}{\partial n} = \frac{mp_n(q_{0,n})}{p_q - mp_q}$$

Note that $mp_q > 0$ (for instance with linear demand function : $mp_q = \frac{b}{n}$), p_q by definition and $mp_n(q_{0,n}) < 0$ as an increase of n decreases retailers market power. Hence $\frac{\partial q_{0,n}}{\partial n} > 0$, which implies that welfare is increasing with n . For the investment level that maximizes consumer surplus, we have:

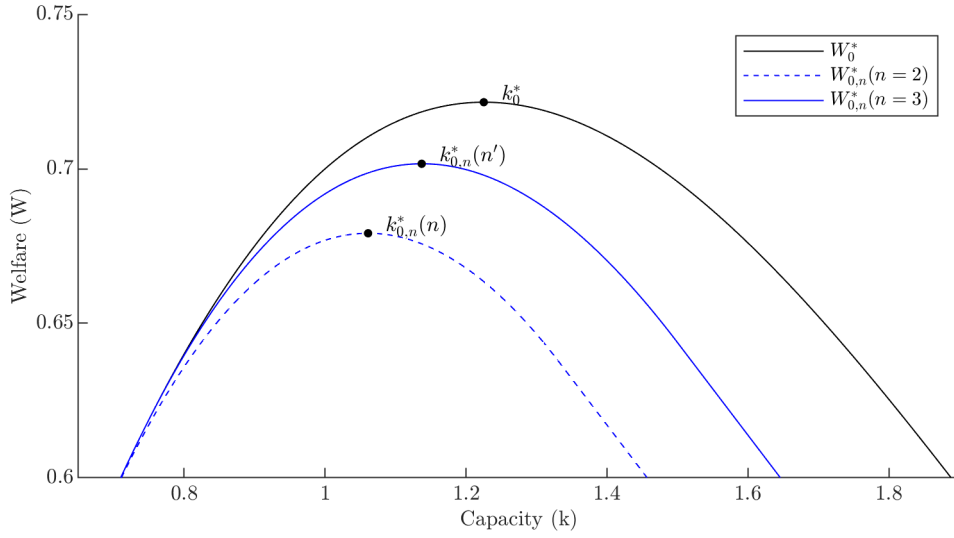


Figure C.1: Expected social welfare, first-best and market investment level with and without imperfect competition in the retail market

$$\frac{\partial^2 W_{0,n}(k)}{\partial k \partial n} = -\frac{\partial s_{0,n}}{\partial n} mp(k) f(s_{0,n})$$

From the definition of $s_{0,n}$, we have $\frac{\partial s_{0,n}}{\partial n} = -\frac{\partial q_{0,n}}{\partial n} / \frac{\partial q_{0,n}}{\partial s}$. We have shown that the first derivative is positive. The second is also positive, as an increase of s increases the demand function. Hence, increasing n decreases the occurrence of off-peak periods. This is intuitive as decreasing imperfect competition raises demand, so capacity binds more often. The envelop theorem implies that the sign of the derivative of $k_{0,n}^*$ with respect to n is positive.³

We now turn to the endogenous case. We denote $k_{1,n}^*$ as the first-best under an endogenous design with imperfect competition but without the allocation based on market share. That is, the capacity price is allocated on a variable basis directly to the final consumers. In that case the marginal expected social welfare such as $k_{1,n}^* = \{k : \phi_{1,n}(k) = r\}$ with:

$$\phi_{1,n}(k) = \int_0^{s_{1,n}} \frac{\partial q_{1,n}}{\partial k} (p^c(k) + mp(q_{1,n})) dF(s) + \int_{s_{1,n}}^{\bar{s}} (p(k, s) - c) dF(s)$$

³The concavity of the expected welfare with respect to k is straightforward. Imperfect competition does not alter its behavior.

Similarly to the previous case, we illustrate this result in Figure C.2. We show the expected social welfare for different values of capacity and the cumulative effect of imperfect competition in the system. The black curve represents the case with perfectly competitive retailers under an exogenous regime. The red curve represents the endogenous case, still without imperfect competition. We then compare the exogenous (dashed) and endogenous (solid) regimes with the blue curves. As expected, the imperfect competition in this case amplifies the negative price effect. The derivative of the expected social welfare with respect to n in this case can be written as follows:

$$\frac{\partial W_{1,n}(k)}{\partial n} = \int_0^{s_{1,n}} \frac{\partial q_{1,n}}{\partial n} (p^c(k) + mp(k)) dF(s)$$

The expression captures the two effects of increasing n : (i) it changes the market-up of retailers as under the exogenous welfare, but (ii) it also changes capacity market equilibrium as the expected profit of producers changes. In this framework, this is captured via the capacity price.

$$\frac{\partial q_{1,n}}{\partial n} = \frac{mp_n(q_{n,1}) + \frac{\partial p^c(k)}{\partial n}}{p_q - mp_q}$$

The denominator is clearly negative. The derivative of the capacity price is given by:

$$\frac{\partial p^c(k)}{\partial n} = \frac{mp_n(k)\Delta_{1,n}F}{1 - \Delta_{1,n}F}$$

Similarly to the expression in the proof of appendix 4, we have $\Delta_{1,n}F = \int_{s_{0,n}}^{s_{1,n}^w} dF(s) > 0$. Hence, $\frac{\partial p^c(k)}{\partial n} > 0$. An increase in n increases the demand in the wholesale market, which, in turn, increases the profit of producers, which lowers their cost of providing more capacity. This lowers the negative price effect on demand. Consequently, under the endogenous regime, an increase of n has a compound positive effect on the demand: $\frac{\partial q_{1,n}}{\partial n} > 0$. It proves that the expected welfare is positively affected by n in this regime. Finally, the cross derivative gives:

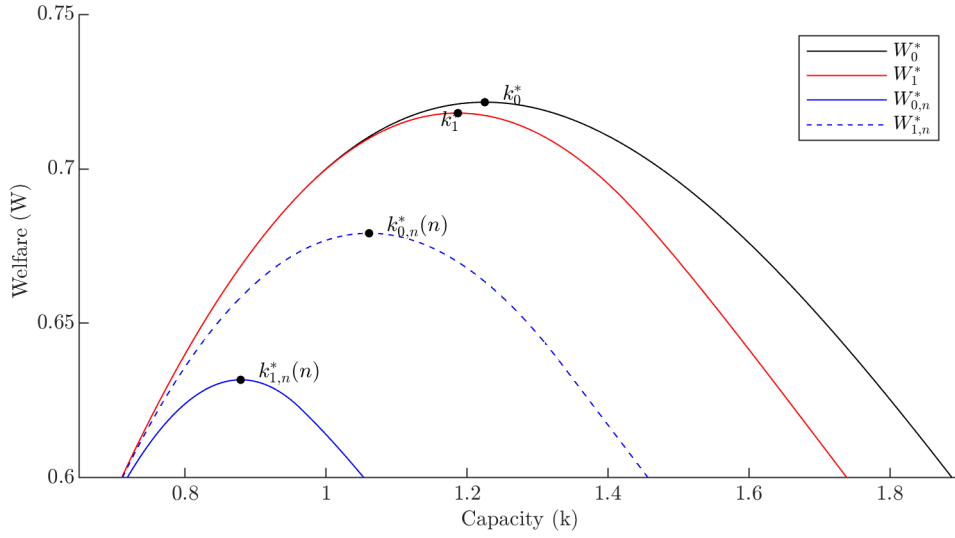


Figure C.2: Expected social welfare, first-best, and market investment level with and without imperfect competition in the retail market under an exogenous and endogenous regime

$$\begin{aligned} \frac{\partial^2 W_{1,n}(k)}{\partial k \partial n} &= \frac{\partial s_{1,n}}{\partial n} \left(\frac{\partial q_{1,n}(s_{1,n})}{\partial k} - 1 \right) (p^c(k) + mp(k)) f(s_{1,n}) \\ &\quad + \int_0^{s_{1,n}} \frac{\partial q_{1,n}}{\partial k} \left(\frac{\partial p^c}{\partial n} + mp_n(q_{1,n}) + \frac{\partial q_{1,n}}{\partial n} mp_q \right) \end{aligned}$$

We have $\frac{\partial s_{1,n}}{\partial n} = mp_n(k) + \frac{\partial p^c}{\partial n} < 0$ and $\frac{\partial q_{1,n}}{\partial k} = \frac{\partial p^c(k)}{\partial k} \frac{1}{p_q - mp_q} < 0$ as the supply function in the capacity market is increasing with k : $\frac{\partial p^c(k)}{\partial k} = -(p_q - mp_q) \frac{\Delta_{1,n} F}{1 - \Delta_{1,n} F} > 0$. Therefore, the first line is positive. The sign of the second line is a priori ambiguous. However, simplification shows that :

$$\begin{aligned} \frac{\partial^2 W_{1,n}(k)}{\partial k \partial n} &= \frac{\partial s_{1,n}}{\partial n} \left(\frac{\partial q_{1,n}(s_{1,n})}{\partial k} - 1 \right) (p^c(k) + mp(k)) f(s_{1,n}) \\ &\quad + \int_0^{s_{1,n}} \frac{\partial q_{1,n}}{\partial k} \left(\frac{\partial p^c}{\partial n} + mp_n(q_{1,n}) \right) \frac{p_q}{p_q - mp_q} \end{aligned}$$

All the terms are negative; hence, the second line is positive. Therefore, using the envelop theorem, we find that the investment level that maximizes expected welfare is also increasing with n

2 Proof of Lemma 2.1

Proof. The market investment level is equal to the optimal investment level only if the solution of the following equality has the same solution as with the first-order condition of the welfare function: $\phi^w(k) + z^w(k) = r$. Which is the case if and only if :

$$z^w(k) = \phi^w(k) - \int_{s_0}^{\bar{s}} \int_0^k (p(q, s) - c) dq dF(s)$$

Therefore, the optimal payment is equal to the expected lost revenue between the optimal price and the price cap:

$$z^w(k) = \int_{s_0^w}^{\bar{s}} (p(k, s) - p^w) dF(s)$$

□

3 Proof of Proposition 2.1

Proof. The supply function in the capacity market, with a price cap in the wholesale market, is positive and is equal to the:

$$X(k) = r - \phi(k) = r - \int_{s_0}^{s_0^w} (p(k, s) - c) dF(s) - \int_{s_0^w}^{\bar{s}} (p^w - c) dF(s)$$

We then find the intersection between the supply function and the optimal payment functions given by the following equations:

$$z^w(k) = \int_{s_0^w}^{\bar{s}} (p(k, s) - p^w) dF(s)$$

For instance, when $X(k) = z^w(k)$, then it gives:

$$r - \int_{s_0}^{s_0^w} (p(k, s) - c) dF(s) - \int_{s_0^w}^{\bar{s}} (p^w - c) f dt = \int_{s_0^w}^{\bar{s}} (p(k, s) - p^w) dF(s)$$

When rearranged, we have:

$$r = \int_{s_1^w}^{\bar{s}} (p(k, s) - c) dF(s)$$

Which is the condition for the first-best to be reached. Therefore, having a demand function at the optimal investment level, given the supply function in the capacity market, is strictly the same as providing producers with a payment whose value is given by the payment function at the optimal investment level. \square

4 Proof of Lemma 2.2

Proof. For clarity, let's denote $p^c(k)$, the capacity price imposed on the final consumers. We assume, for now, that it exists. We formally demonstrate it in the proof of Lemma 2.3. For ease of notation, we sometimes use only p^c .

First, we state in this proof that allocating the capacity price on the consumer as a tax only affects the share between on-peak and off-peak periods and the surplus's size during off-peak periods. We start by defining the expected consumer surplus without a capacity market:

$$\begin{aligned} W^c(k) = & \int_0^{s_0} \int_0^{q_0(s)} (p(q, s) - c) dq dF(s) + \int_{s_0}^{s_0^w} \int_0^k (p(q, s) - p(k, s)) dq dF(s) \\ & + \int_{s_0^w}^{\bar{s}} \int_0^k (p(q, s) - p^w) dq dF(s) \end{aligned}$$

Then, we define both the expected welfare for consumers and producers with a capacity market for three different periods: (i) when the capacity is not binding, (ii) when the capacity is binding but not the price cap, and (iii) when the price cap is binding. We show that for each period, the surplus encompasses, at the same time, the direct welfare loss and gain from the transfer due to the capacity market and the indirect effect due to the capacity price allocation. The expected surplus for the consumer during off-peak periods is:

$$\begin{aligned}
W_{of}^c(k, p^c) &= \int_0^{s_1(p^c)} \left[-p^c(k)(k - q_1(p^c)) + \int_0^{q_1(p^c)} (p(q, s) - c - p^c(k))dq \right] dF(s) \\
&= \int_0^{s_1(p^c)} \left[-p^c(k)k + \int_0^{q_1(p^c)} (p(q, s) - c)dq \right] dF(s)
\end{aligned}$$

Note that the term $-p^c(k)(k - q_1(p^c))$ in the integral ensures that the capacity buyers do not make any losses. For the on-peak periods when the price cap is not binding, the consumer welfare is:

$$\begin{aligned}
W_{on}^c(k, p^c) &= \int_{s_1(p^c)}^{s_1^w(p^c)} \left[-p^c(k)k + \int_0^k (p(q, s) - p(k, s) - p^c(k))dq \right] dF(s) \\
&= \int_{s_1^w(p^c)}^{\bar{s}} \int_0^k (p(q, s) - p(k, s))dq dF(s)
\end{aligned}$$

For the periods when the price cap is binding, the consumer welfare is:

$$W_{cap}^c(k, p^c) = \int_{s_1^w}^{\bar{s}} \left[-p^c(k)k + \int_0^k (p(q, s) - p^w)dq \right] dF(s)$$

On the other hand, the producer welfare during offpeak periods is:

$$W_{of}^p(k, p^c) = \int_0^{s_1(p^c)} p^c(k)k dF(s)$$

For the on-peak periods when the price cap is not binding, the producer welfare is:

$$\begin{aligned}
W_{on}^p(k, p^c) &= \int_{s_1(p^c)}^{s_1^w} [p^c(k)k + k(p(k, s) - p^c(k) - c)] dF(s) \\
&= \int_{s_1(p^c)}^{s_1^w} k(p(k, s) - c) dF(s)
\end{aligned}$$

For the periods when the price cap is binding, the producer welfare is:

$$W_{cap}^c(k, p^c) = \int_{s_1^w}^{\bar{s}} [p^c(k)k + k(p^w - c)] dF(s)$$

When we add the different expected welfare for consumers and producers, we have the $p^c(k)k$ parts canceled and the price cap p^w . It gives the following expected social welfare:

$$W_1(k, p^c) = \int_0^{s_1(p^c)} \int_0^{q_1(p^c)} (p(q, s) - c) dq dF(s) + \int_{s_1(p^c)}^{\bar{s}} \int_0^k (p(q, s) - c) dq dF(s) - rk$$

□

5 Proof of Lemma 2.3

Proof. The endogenous capacity market price is given by solving $X_1(k, p^c) = p^c$. In the endogenous case, its expression is

$$X_1(k) = \begin{cases} 0 & \text{if } k \leq k_0^w \\ r - \phi_1^w(k, p^c) & k > k_0^w \end{cases} \quad (\text{C.8})$$

With k_0^w , the market equilibrium investment level is such that we have $r = \phi_1^w(k, p^c)$. Hence, we focus on proving that a solution exists for any value of $k \in [k_0^w + \infty)$. The expression of $\phi_1^w(k, p^c)$ depends on the value of p^c , similarly to the case of k . That is, there exist some values of p^c such that capacity or price cap might be always or never binding in expectation. The expressions of $\phi_1^w(k, p^c)$ are therefore:

$$\phi_1^w(k, p^c) = \begin{cases} \int_0^{\bar{s}} [p^w - c] dF(s) & \forall p^c \in [-\infty, p_{--}^c] \\ \int_0^{s_1^w(p^c)} [p(k, s) - p^c - c] dF(s) - \int_{s_1^w(p^c)}^{+\infty} [p^w - c] dF(s) & \forall p^c \in [p_{--}^c, p_{-}^c] \\ \int_{s_1(p^c)}^{s_1^w(p^c)} [p(k, s) - p^c - c] dF(s) - \int_{s_1^w(p^c)}^{+\infty} [p^w - c] dF(s) & \forall p^c \in [p_{-}^c, p_{+}^c] \\ \int_{s_1(p^c)}^{\infty} [p(k, s) - p^c - c] dF(s) & \forall p^c \in [p_{+}^c, p_{++}^c] \\ 0 & \forall p^c \in [p_{++}^c, +\infty] \end{cases}$$

With the threshold values defined as follows: $p_{--}^c = \{p^c : s_1^w(p^c) = 0\}$, price cap always binds in expectation, $p_{-}^c = \{p^c : s_1(p^c) = 0\}$, capacity always binds, and sometimes the price cap bind. $p_{+}^c = \{p^c : s_1^w(p^c) = +\infty\}$, the price cap never binds, and the capacity sometimes binds. $p_{++}^c = \{p^c : s_1(p^c) = +\infty\}$, the price cap and the capacity never bind.⁴ The middle case where both capacity and the price cap sometimes bind in expectation is the third line. Note that because of the definition of s_1 and s_1^w , the derivatives with respect to p^c are always increasing. For instance, s_1 solves:

Differentiating with p^c gives $p_s(k, s) \frac{\partial s_1}{\partial p^c} - 1 = 0$, hence: $\frac{\partial s_1}{\partial p^c} = \frac{1}{p_s(k, s)} > 0$. For instance, under the uniform distribution of s : $s_1(p^c) = c - a_0 + p^c + k$ and $s_1^w(p^c) = p^w - a_0 + p^c + k$. If a_0 is sufficiently high, then there exists some low value of p^c such that both terms are negative.

We define the function $g(p^c, k) = X_1(k) - p^c = r - \phi_1^w(k) - p^c$. We take the derivative with respect to p^c of the third expression :

$$\begin{aligned} \frac{\partial g(p^c, k)}{\partial p^c} &= (p(k, s_1) - p^c - c) f(s_1) \frac{\partial s_1(p^c)}{\partial p^c} + \int_{s_1(p^c)}^{s_1^w(p^c)} dF(s) \\ &\quad - f(s_1^w) (p(k, s_1^w(p^c)) - p^c - c) \frac{\partial s_1^w(p^c)}{\partial p^c} + (p^w - c) f(s_1^w) \frac{\partial s_1^w(p^c)}{\partial p^c} - 1 \end{aligned}$$

At $s_1(p^c)$ we have $p(k, s_1(p^c)) - p^c(k) - c = 0$. The third and fourth terms cancel out as $p(k, s_1^w(p^c)) - p^c(k) = p^w$. Therefore, the first derivative extended to all cases are:

$$\frac{\partial g(k, p^c)}{\partial p^c} = \begin{cases} -1 & \forall p^c \in [-\infty, p_{--}^c] \\ \int_0^{s_1^w(p^c)} dF(s) - 1 & \forall p^c \in [p_{--}^c, p_{-}^c] \\ \int_{s_1(p^c)}^{s_1^w(p^c)} dF(s) - 1 & \forall p^c \in [p_{-}^c, p_{+}^c] \\ \int_{s_1(p^c)}^{\bar{s}} dF(s) - 1 & \forall p^c \in [p_{+}^c, p_{++}^c] \\ -1 & \forall p^c \in [p_{++}^c, -\infty] \end{cases}$$

Therefore $g(p^c, k)$ is a necessary decreasing function for any $p^c(k)$ and any expression of ϕ_1^w . Note also that $g(p^c, k)$ is increasing with the investment value k :

⁴In a similar way, one can also define capacity thresholds for which capacity always bind or never binds in expectation for instance.

$$\frac{\partial g(p^c, k)}{\partial k} = - \int_{s_1(p^c)}^{s_1^w(p^c)} p_q dF(s)$$

Again, note that the derivative is also positive for other expressions of ϕ_1^w . By definition we know that at the minimum value of k , k_0^w , the fixed point is $p^c(k) = 0$ because $r = \phi_1^w(k, p^c)$. As $g(p^c, k)$ is increasing with k , we have for any $k > k_0^w$: $g(0, k) > 0$. The second derivative with respect to $p^c(k)$ is :

$$\frac{\partial^2 \phi_1^w(k, p^c)}{\partial p^{c2}} = \begin{cases} 0 & \forall p^c \in [-\infty, p_{--}^c] \\ f(s_1^w) \frac{\partial s_1^w(p^c)}{\partial p^c} & \forall p^c \in [p_{--}^c, p_-^c] \\ f(s_1^w) \frac{\partial s_1^w(p^c)}{\partial p^c} - f(s_1) \frac{\partial s_1(p^c)}{\partial p^c} & \forall p^c \in [p_-^c, p_+^c] \\ -f(s_1) \frac{\partial s_1(p^c)}{\partial p^c} & \forall p^c \in [p_+^c, p_{++}^c] \\ 0 & \forall p^c \in [p_{++}^c, -\infty] \end{cases}$$

From the definition of $s_1(p^c)$ and $s_1^w(p^c)$ we have $\frac{\partial s_1^w(p^c)}{\partial p^c} = \frac{1}{p_s(s_1^w(p^c))} > 0$ and $\frac{\partial s_1(p^c)}{\partial p^c} = \frac{1}{p_s(s_1)} > 0$.

Therefore, the function g depending on the value of k , exhibits a succession of linear, convex, ambiguous, concave, and finally, linear parts. The last expression ensures at least a fixed point for any $k \in [k_0^w + \infty)$, without giving a precise location on which part the fixed point is.

The second part of the proposition comes for the difference between $X_0(k)$ and $X_1(k)$. Note that we necessarily have $s_1^w(p^c) > s_0^w$ and $s_1(p^c) > s_0$ for any positive capacity price. We express the delta as follows:

$$\begin{aligned} X_1(k) - X_0(k) = & - \int_{s_1(p^c)}^{s_1^w(p^c)} [p(k, s) - p^c(k) - c] dF(s) - \int_{s_1^w(p^c)}^{+\infty} [p^w - c] dF(s) \\ & + \int_{s_0}^{s_0^w} [p(k, s) - c] dF(s) + \int_{s_0^w}^{\bar{s}} [p^w - c] dF(s) \end{aligned}$$

Rearranged, and assuming (w.l.o.g.) $s_1(p^c) < s_0^w$, we have :

$$\begin{aligned}
X_1(k) - X_0(k) &= \int_{s_1(p^c)}^{s_1^w(p^c)} p^c(k) dF(s) + \int_{s_0}^{s_1(p^c)} [p(k, s) - c] dF(s) \\
&\quad + \int_{s_0^w}^{s_1^w(p^c)} [p^w - p(k, s) + p^c(k)] dF(s)
\end{aligned}$$

By definition of $s_1^w(p^c)$ such that $p(k, s_1^w(p^c)) - p^c(k) - p^w = 0$, the last term is always positive. Hence for any $k \in [k_0^w + \infty)$ we have $X_1(k) > X_0(k)$. □

6 Proof of Proposition 2.2

Proof. We define $\tilde{k}_1^-, \tilde{k}_1^{w-}, \tilde{k}_1^{w+}, \tilde{k}_1^+$ the threshold investment level such that : (1) capacity always binds: $s_1(\tilde{k}_1^-) = 0$, (2) price cap always binds: $s_1^w(\tilde{k}_1^{w-}) = 0$, (3) price cap never binds: $s_1^w(\tilde{k}_1^{w+}) = \bar{s}$, (4) capacity never binds: $s_1(\tilde{k}_1^+) = \bar{s}$ (recall that $s_1^w < s_1$)

We chose to study the welfare in the thresholds $[\tilde{k}_1^{w-}, \tilde{k}_1^{w+}]$ which corresponds to the case where capacity and the price cap binds (but not always), as the results hold for the other cases. The expected social welfare is equal to the following:

$$W_1(k, p^c) = \int_0^{s_1(p^c)} \int_0^{q_1(p^c)} (p(q, s) - c) dq dF(s) + \int_{s_1(p^c)}^{\bar{s}} \int_0^k (p(q, s) - c) dq dF(s) - rk$$

The first derivative with respect to the level of investment is :

$$\begin{aligned}
\frac{\partial W_1(k, p^c)}{\partial k} &= f(s_1) \frac{\partial s_1}{\partial k} \int_0^{q_1(s_1, p^c)} (p(q, s_1) - c) dq + \int_0^{s_1(p^c)} \frac{\partial q_1}{\partial k} [p(q_1, s) - c] dF(s) \\
&\quad - f(s_1) \frac{\partial s_1}{\partial k} \int_0^k (p(q, s_1) - c) dq + \int_{s_1(p^c)}^{\bar{s}} [p(k, s) - c] dF(s) - r
\end{aligned}$$

At $s_1(p^c)$, we have $q_1(s_1, p^c) = k$. Hence, the first and third terms cancel out. Moreover, by definition $p(q_1, s) - c = p^c(k)$. The first derivative with respect to the level of investment becomes :

$$\frac{\partial W_1(k, p^c)}{\partial k} = \int_0^{s_1(p^c)} p^c(k) \frac{\partial q_1}{\partial k} dF(s) + \int_{s_1(p^c)}^{\bar{s}} [p(k, s) - c] dq dF(s) - r$$

Using the definition of $s_1(p^c)$ such that $p(k, s_1(p^c)) - c = p^c(k)$, the second derivative with respect to the level of investment is :

$$\begin{aligned} \frac{\partial^2 W_1(k, p^c)}{\partial k^2} &= p^c(k) f(s_1) \frac{\partial s_1}{\partial k} \left[\frac{\partial q_1(s_1)}{\partial k} - 1 \right] \\ &+ \int_0^{s_1(p^c)} \left[p^c(k) \frac{\partial^2 q_1}{\partial k^2} + \frac{\partial q_1}{\partial k} \frac{\partial p^c}{\partial k} \right] dF(s) + \int_{s_1(p^c)}^{\bar{s}} p_q dF(s) \end{aligned}$$

We simplify the previous expression using the initial assumption that states of the world only change the demand intercept.

$$\begin{aligned} \frac{\partial^2 W_1(k, p^c)}{\partial k^2} &= p^c(k) f(s_1) \frac{\partial s_1}{\partial k} \left[\frac{\partial}{\partial q_1}(s_1)k - 1 \right] + \int_0^{s_1(p^c)} p^c(k) \frac{\partial^2 q_1}{\partial k^2} dF(s) \\ &+ \int_0^{s_1(p^c)} \frac{\partial q_1}{\partial k} \frac{\partial p^c}{\partial k} dF(s) + \int_{s_1(p^c)}^{\bar{s}} p_q dF(s) \end{aligned}$$

We now analyze the derivative of the capacity price with respect to k . The expression of $p^c(k)$ is given by

$$p^c(k) = r - \int_{s_1(p^c)}^{s_1^w(p^c)} [p(k, s) - p^c(k) - c] dF(s) - \int_{s_1^w(p^c)}^{\bar{s}} [p^w - c] dF(s)$$

Therefore, we have

$$\begin{aligned} \frac{\partial p^c}{\partial k} = & - \int_{s_1(p^c)}^{s_1^w(p^c)} \left[-\frac{\partial p^c}{\partial k} \right] dF(s) + (p^w - c)f(s_1^w) \frac{\partial s_1^w}{\partial k} \\ & (p(k, s_1) - p^c(k) - c)f(s_1) \frac{\partial s_1}{\partial k} - (p(k, s_1^w) - p^c(k) - c)f(s_1^w) \frac{\partial s_1^w}{\partial k} \end{aligned}$$

By definition $p(k, s_1(p^c)) - c = p^c(k)$, and $p(k, s_1^w(p^c)) - p^w = p^c(k)$. Therefore, the third term is null, and the second and fourth terms cancel out. The capacity price is independent of the realized state of the world. Hence we have

$$\frac{\partial p^c}{\partial k} = -p_q \int_{s_1(p^c)}^{s_1^w(p^c)} dF(s) + \frac{\partial p^c}{\partial k} \int_{s_1(p^c)}^{s_1^w(p^c)} dF(s)$$

Therefore,

$$\frac{\partial p^c}{\partial k} = -p_q \frac{\int_{s_1(p^c)}^{s_1^w(p^c)} dF(s)}{1 - \int_{s_1(p^c)}^{s_1^w(p^c)} dF(s)} = -p_q \frac{\Delta_1 F}{1 - \Delta_1 F}$$

With $\Delta_1 F = \int_{s_1(p^c)}^{s_1^w(p^c)} dF(s) = f(s_1^w) - f(s_1)$. Whenever $s_1(p^c)$ and $s_1^w(p^c)$ exists and are positive then $\Delta_1 F > 0$ and $\Delta_1 F < 1$. It implies that $\frac{\partial p^c}{\partial k} > 0$. As expected, it confirms that the supply function in the capacity market is increasing with the level of investment k . An increase of k directly increases the occurrence of off-peak periods and indirectly lowers the demand via a higher capacity price :

$$\frac{\partial q_1}{\partial k} = \frac{\partial p^c}{\partial k} \frac{1}{p_q} = -\frac{\Delta_1 F}{1 - \Delta_1 F} < 0$$

A higher investment level means less demand. The second-order derivative is expressed as follows :

$$\frac{\partial^2 q_1}{\partial k^2} = \frac{1}{p_q} \frac{\partial^2 p^c}{\partial k^2}$$

Therefore, $\frac{\partial^2 q_1}{\partial k^2} < 0$ if and only if the term inside the brackets is positive. Using $\frac{\partial p^c}{\partial k} = -p_q \frac{\Delta_1 F}{1 - \Delta_1 F}$, the second derivative

$$\frac{\partial^2 p^c(k)}{\partial k^2} = -p_q \frac{\partial \Delta_1 F}{\partial k} \frac{1}{(1 - \Delta_1 F)^2}$$

Now we express $\frac{\partial \Delta_1 F}{\partial k}$

$$\frac{\partial \Delta_1 F}{\partial k} = \frac{\partial s_1^w}{\partial k} f(s_1^w) - \frac{\partial s_1}{\partial k} f(s_1)$$

Using the previous expression and noting that $\frac{\partial^2 q_1}{\partial k^2}$ does not depend on s , we again simplify the second derivative of the expected social welfare:

$$\begin{aligned} \frac{\partial^2 W_1(k, p^c)}{\partial k^2} &= p^c(k) f(s_1) \frac{\partial s_1}{\partial k} \left[\frac{\partial q_1(s_1)}{\partial k} - 1 \right] - p^c(k) \frac{\partial \Delta_1 F}{\partial k} \frac{f(s_1)}{(1 - \Delta_1 F)^2} \\ &\quad + \int_0^{s_1(p^c)} \frac{\partial q_1}{\partial k} \frac{\partial p^c}{\partial k} dF(s) + \int_{s_1(p^c)}^{\bar{s}} p_q dF(s) \\ &= -p^c(k) f(s_1) \frac{\partial s_1}{\partial k} \frac{1 - f(s_1^w)}{(1 - \Delta_1 F)^2} - p^c(k) f(s_1^w) \frac{\partial s_1^w}{\partial k} \frac{f(s_1)}{(1 - \Delta_1 F)^2} \\ &\quad + \int_0^{s_1(p^c)} \frac{\partial q_1}{\partial k} \frac{\partial p^c}{\partial k} dF(s) + \int_{s_1(p^c)}^{\bar{s}} p_q dF(s) \end{aligned}$$

The first term is negative as $1 - f(s_1^w) \geq 0$, and the second term is also negative from the definition of the derivative of the threshold as :

$$\frac{\partial s_1}{\partial k} = \frac{-p_q + \frac{\partial p^c}{\partial k}}{p_s(s_1)} = -\frac{p_q}{p_s(s_1)} \frac{1}{1 - \Delta_1 F} > 0$$

Finally, the third term is also negative as $\frac{\partial q_1}{\partial k} \leq 0$ and $\frac{\partial p^c}{\partial k} \geq 0$. Finally, the last term is also negative.

The expression in (ii) of $\phi_1(k)$ is given by $\frac{\partial W_1(k, p^c)}{\partial k}$. The proof of (iii) relies on the analysis of the derivative of $\frac{\partial \Delta W_1(k, p^c)}{\partial k}$. The derivative is equal to

$$\begin{aligned} \frac{\partial \Delta W_1(k, p^c)}{\partial k} &= \frac{\partial s_0}{\partial k} f(s_0) \int_{q_1(s_0, p^c)}^{q_0(s_0)} (p(q, s_0) - c) dq + \int_0^{s_0} -\frac{\partial q_1}{\partial k} dF(s) \\ &\quad - \frac{\partial s_0}{\partial k} f(s_0) \int_{q_1(s_0, p^c)}^k (p(q, s_0) - c) dq + \frac{\partial s_1}{\partial k} f(s_1) \int_{q_1(s_1)}^k (p(q, s_1) - c) dq \\ &\quad + \int_{s_0}^{s_1(p^c)} \left[-\frac{\partial q_1}{\partial k} + (p(k, s) - c) \right] dF(s) \end{aligned}$$

Note that the three terms with the derivatives of s_0 and s_1 are null as : $p(q, s_0) - c = 0$ and $q_1(s_1) = k$. Therefore, the equation boils down to:

$$\frac{\partial \Delta W_1(k, p^c)}{\partial k} = \int_0^{s_0} \left[-\frac{\partial q_1}{\partial k} p(q_1, s) - c \right] dF(s) + \int_{s_0}^{s_1(p^c)} \left[-\frac{\partial q_1}{\partial k} + (p(k, s) - c) \right] dF(s)$$

From the previous analysis, we know that the sign of $\frac{\partial q_1}{\partial k}$ is negative. In turn, the derivative $\frac{\partial \Delta W_1(k, p^c)}{\partial k}$ is fully positive. Hence, the new first-best solution in terms of investment level given under the endogenous regime is always lower or equal to the first-best solution under the exogenous level. The same result applied to the expected social welfare at the first-best investment level.

□

7 Proof of Lemma 2.4

Proof. In this case, the inframarginal rent is the same as the case with a price cap only. On the other hand, the new welfare function is:

$$W^{bo}(k) = W(k) - M(k)$$

The first-order condition is similar to those in the previous proof:

$$\int_{s_0}^{+\infty} \int_0^k (p(q, s) - c) dq f(s) dt - M_k(k) = r$$

The market investment level is equal to the optimal investment level only if the solution of the following equality has the same solution as with the first-order condition of the welfare function. Because the inefficient rationing is due to the price cap, the optimal payment includes de facto $z^w(k)$ to restore the optimal investment level. Here, we derive only the additional part relative to the rationing.

$$\phi(k) + z^{bo}(k) = r$$

Which is the case if and only if :

$$z^{bo}(k) = \phi(k) - \left(\int_{s_0}^{+\infty} \int_0^k (p(q, s) - c) dq f(s) dt - M_k(k) \right)$$

Therefore the optimal payment is equal to the marginal surplus loss due to inefficient rationing.

$$z^{bo}(k) = M_k(k)$$

□

8 Proof of Lemma 2.5

Proof. We start with the definition of the expected welfare under the endogenous regime with inefficient rationing. Using the definition of $W_0(k)$ and the rationing cost $m(k)$, and after rewriting, we have :

$$W_1^{bo}(k) = W_1(k) - \mathbb{A}$$

With

$$\mathbb{A} = \int_{s_1^w}^{\bar{s}} \left(\int_0^k (p(q, s) - p^w) dq - \frac{k}{q_1^w} \int_0^{q_1^w} (p(q, s) - p^w) dq \right) dF(s)$$

Hence, the second order derivative with respect to k is given both by $\frac{\partial^2 W_0(k)}{\partial k^2}$ and by deriving the second term. The first derivative is equal to:

$$\begin{aligned} \frac{\partial \mathbb{A}}{\partial k} = & -f(s_1^w) \frac{\partial s_1^w}{\partial k} \left[\int_0^k (p(q, s_1^w) - p^w) dq - \frac{k}{q_1^w(s_1^w)} \int_0^{q_1^w(s_1^w)} (p(q, s_1^w) - p^w) dq \right] \\ & + \int_{s_1^w}^{\bar{s}} \left((p(k, s) - p^w) - \frac{1}{q_1^w} \int_0^{q_1^w} (p(q, s) - p^w) dq + \frac{\partial q_1^w}{\partial k} \frac{k}{(q_1^w)^2} \int_0^{q_1^w} (p(q, s) - p^w) dq \right. \\ & \left. - \frac{\partial q_1^w}{\partial k} \frac{k}{q_1^w} (p(q_1^w, s) - p^w) dq \right) dF(s) \end{aligned}$$

This equation can be significantly reduced. First, note that at s_1^w by definition, we have $q_1^w = k$. Hence, the first line is null. Then, we use the observation that $\int_0^{q_1^w} (p(q, s) - p^w) dq = \int_0^{q_1^w} (p(q, s) - p(q_1^w, s)) dq + \int_0^{q_1^w} (p(q_1^w, s) - p^w) dq$. This allows the third term of the second line and the term in the third line to cancel out partly. Hence, the derivative is equal to:

$$\begin{aligned} \frac{\partial \mathbb{A}}{\partial k} = & \int_{s_1^w}^{\bar{s}} \left((p(k, s) - p^w) - \frac{1}{q_1^w} \int_0^{q_1^w} (p(q, s) - p^w) dq \right. \\ & \left. + \frac{\partial q_1^w}{\partial k} \frac{k}{(q_1^w)^2} \int_0^{q_1^w} (p(q, s) - p(q_1^w, s)) dq \right) dF(s) \end{aligned}$$

Next, we take the second derivative with respect to k . Following the same previous arguments, we have:

$$\begin{aligned} \frac{\partial^2 \mathbb{A}}{\partial k^2} = & -f(s_1^w) \frac{\partial s_1^w}{\partial k} \left(1 - \frac{\partial q_1^w}{\partial k} \right) \frac{1}{k} \int_0^k (p(q, s_1^w) - p(k, s_1^w)) dq \\ & + \int_{s_1^w}^{\bar{s}} \left(\frac{1}{(q_1^w)^2} \int_0^{q_1^w} (p(q, s) - p(q_1^w, s)) dq \left(1 - \frac{\partial q_1^w}{\partial k} \right) \right) dF(s) \end{aligned}$$

Hence:

$$\frac{\partial^2 \mathbb{A}}{\partial k^2} = \left(1 - \frac{\partial q_1^w}{\partial k}\right) \left[-f(s_1^w) \frac{\partial s_1^w}{\partial k} \frac{1}{k} \int_0^k (p(q, s_1^w) - p(k, s_1^w)) dq \right. \\ \left. + \int_{s_1^w}^{\bar{s}} \frac{1}{(q_1^w)^2} \int_0^{q_1^w} (p(q, s) - p(q_1^w, s)) dq dF(s) \right]$$

Define the per unit consumer surplus at s_1^w associated with its probability:

$$\mathbb{CS}(k) = \frac{1}{k} \int_0^k (p(q, s_1^w) - p(k, s_1^w)) dq$$

And the corresponding derivative with respect to k :

$$\frac{\partial \mathbb{CS}}{\partial k} = -\frac{1}{k^2} \int_0^k (p(q, s_1^w) - p(k, s_1^w)) dq - p_q$$

Hence, the derivative of the expected per-unit consumer surplus when the price cap is binding is such that

$$-\mathbb{CS}(k) f(s_1^w) \frac{\partial s_1^w}{\partial k} + \int_{s_1^w}^{\bar{s}} \left(-\frac{1}{k^2} \int_0^k (p(q, s_1^w) - p(k, s_1^w)) dq - p_q \right) dF(s)$$

Then, the derivative can be rewritten as:

$$\frac{\partial^2 \mathbb{A}}{\partial k^2} = \overbrace{\left(\frac{\partial}{\partial q_1^w} - 1 \right)}^{-} \left[\frac{\partial}{\partial k} \int_{s_1^w}^{\bar{s}} \mathbb{CS}(k) dF(s) + \int_{s_1^w}^{\bar{s}} \frac{\partial \mathbb{CS}}{\partial k} dF(s) \right]$$

The second term has to be negative, hence the good for the concavity of the expected welfare.

□

9 Proof of Proposition 2.3

Proof. Before starting the proof, we define the limit value of the price cap p^{w+} . It is given by the solution of $k_0^w(p^w) = k_0^*$. Recall that $k_0^w(p^w)$ is the equilibrium investment level such that profit is maximized under a price cap and with any capacity market. Hence, p^{w+} is the price cap level such that it does not generate any inefficiency.⁵ The usefulness of this threshold lies in the fact at this value, we have $p^c(k_0^w) = 0$. Indeed, if $k_0^w = k_0^*$, the supply function on the capacity market starts at the first-best solution. Hence, the regulator always chooses k_0^* as the demand level on the capacity market. This result also extends to the endogenous regime. Hence, at p^{w+} we have $k_0^* = k_1^*$. If the price cap never binds, then we do not have any inefficient rationing. hence, we also have $k_0^* = k_1^* = k_0^{bo} = k_1^{bo}$. In this chapter, we do not have closed-form solutions. Therefore, clear-cut results are only given by studying behaviors at the vicinity of p^{w+} . From the previous analysis, we have proved that $k_0^* \geq k_1^*$ and $k_0^{bo} \geq k_1^{bo}$ due to the negative price effect of the capacity market. We know that due to the inefficient rationing cost, we have $k_0^{bo} \geq k_0^*$ and $k_1^{bo} \geq k_1^*$. Finally, by deduction, we also have $k_0^* \geq k_1^*$. We still need to study the ranking between k_0^* and k_1^{bo} .

To do so, note that by definition, k_0^* is independent of p^w as it is the first-best investment level. Therefore, studying the derivative of $k_1^{bo} - k_0^*$ with respect to p^w boils down to studying the derivative of k_1^{bo} with respect to p^w . By definition k_1^{bo} solves the equation $\frac{\partial W_1^{bo}}{\partial k} = 0$. Therefore:

$$\frac{\partial k_1^{bo}}{\partial p^w} = -\frac{\partial^2 W_1^{bo}}{\partial k \partial p^w} / \frac{\partial^2 W_1^{bo}}{\partial k^2}$$

Proposition 2.2 shows that the expected welfare is concave, hence: $\frac{\partial^2 W_1^{bo}}{\partial k^2} \leq 0$. We now express the cross derivative, first without inefficient rationing. using the expression of $W_1(k)$, this gives:

$$\frac{\partial^2 W_1}{\partial k \partial p^w} = p^c(k) \left(\frac{\partial q_1}{\partial k} - 1 \right) f(s_1) \frac{\partial s_1}{\partial p^w} + \int_0^{t_1} \left(\frac{\partial^2 q_1}{\partial k \partial p^w} p^c(k) + \frac{\partial q_1}{\partial k} \frac{\partial p^c}{\partial p^w} \right) dF(s)$$

⁵Equivalently, it can be solved by having $t_0^w(k_0^w(p^w), p^w) = \bar{s}$, that is when the price cap never binds in expectation.

From the definition of s_1 such that $p(k, s) - p^c(k) = c$, the derivative with respect to p^w is given by: $\frac{\partial p^c}{\partial p^w} \frac{1}{p_t}$. From the definition of $p^c(k)$ the solves the equation $p^c(k) = X_1(k, p^c)$, we have :

$$\frac{\partial p^c}{\partial p^w} = -\frac{\int_{t_1^w}^{\bar{s}} 1dF(s)}{1 - \Delta_1 F}$$

Recall that $\Delta_1 F = F(s_1^w) - F(s_1) > 0$ Therefore, the derivative is negative. Hence $\frac{\partial s_1}{\partial p^w} \leq 0$. Using the expression of the derivatives, we can simplify the cross derivative as:

$$\frac{\partial^2 W_1}{\partial k \partial p^w} = \left(\frac{\partial q_1}{\partial k} - 1 \right) \left(\frac{\partial s_1}{\partial p^w} + \frac{\partial s_1^w}{\partial p^w} \right) p^c(k) f(s_1) + \int_0^{t_1} \left(\frac{\partial q_1}{\partial k} \frac{\partial p^c}{\partial p^w} \right) dF(s)$$

From the previous analysis, we know that $W_1(k)$ and $W_0(k)$ are both concave with respect to k , and so is their difference. Moreover, calculations show that the difference decreases with respect to p^w :

$$\int_0^{s_1} -\frac{\partial q_1}{\partial p^w} p^c(k) dF(s) < 0$$

Indeed, we have $\frac{\partial q_1}{\partial p^w} = \frac{\partial p^c}{\partial p^w} \frac{1}{p_q} > 0$. Moreover, we know that at the limit of p^{w+} , we have $k_0^* = k_1^*$. Note also that the derivative is positive at p^{w+} as $p^c(k) = 0$ at the maximizing investment level. Therefore, $\frac{\partial k_1^*}{\partial k} > 0$. Intuitively, as p^w increases, the negative price effect of the capacity decreases itself due to the decreases of the supply function on the capacity market. Therefore, as the price cap increases, the gap between the two welfare functions (exogenous and endogenous) reduces.

On the other hand, the following expression provides the cross derivative for the exogenous case with inefficient rationing. As the initial welfare is independent of the price cap, this measures only the effect of p^w on the rationing costs. After simplifications, we have:

$$\frac{\partial^2 W_0^{bo}}{\partial k \partial p^w} = -\text{CS} + \frac{1}{2}(1 - F(s_0^w))$$

Which is negative if $\mathbb{CS} \geq \frac{1}{2}(1 - F(s_0^w))$. Under our closed-form framework, this condition is met as soon as we have :

$$\frac{\partial}{\partial k} \int_{s_0^w}^{\bar{s}} k dF(s) \leq 0$$

That is, the expected on-peak quantity is decreasing in k . In other terms, the decrease in the occurrence of on-peak periods should be higher than the increase in the expected quantity. In that case, the cross derivative is negative, which implies that the quantity that maximized expected welfare is decreasing in p^w . This is an intuitive relationship, as a higher p^w decreases the occurrence for which the price cap is binding; it implies a lower occurrence of inefficient rationing. Hence, as p^w , the gap between the two welfare functions (without and with inefficient rationing) decreases. Note also that the derivative is negative at p^{w+} as $F(s_0^w) = 1$ at the maximizing investment level.

We turn now to the expression of the cross derivative for the endogenous case with inefficient rationing.

$$\frac{\partial^2 W_1^{bo}}{\partial k \partial p^w} = \frac{\partial^2 W_1}{\partial k \partial p^w} + \left(\frac{\partial q_1^w}{\partial k} - 1 \right) \left(\frac{\partial s_1}{\partial p^w} + \frac{\partial s_1^w}{\partial p^w} \right) \mathbb{CS} f(s_1^w) - \int_{t_1^w}^{\bar{s}} \left(\frac{\partial \mathbb{CS}}{\partial k} \frac{\partial q_1^w}{\partial p^w} \right) dF(s)$$

Note the second term is similar to $\frac{\partial^2 W_1}{\partial k \partial p^w}$. After simplification, under the closed-form framework we have $\frac{\partial q_1^w}{\partial k} = \frac{\partial q_1}{\partial k}$ and $f(s_1^w) = f(s_1)$, which implies:

$$\begin{aligned} \frac{\partial^2 W_1^{bo}}{\partial k \partial p^w} = & \overbrace{\left(\frac{\partial q_1^w}{\partial k} - 1 \right) \left(\frac{\partial s_1}{\partial p^w} + \frac{\partial s_1^w}{\partial p^w} \right)}^{-} \overbrace{(\mathbb{CS} + p^c(k)) f(s_1^w)}^{+} \\ & + \underbrace{\int_0^{t_1} \left(\frac{\partial q_1}{\partial k} \frac{\partial p^c}{\partial p^w} \right) dF(s) - \int_{t_1^w}^{\bar{s}} \left(\frac{\partial \mathbb{CS}}{\partial k} \frac{\partial q_1^w}{\partial p^w} \right) dF(s)}_{+} \end{aligned}$$

The individual consumer surplus in k is defined as increasing with k . Due to the negative price effect on the demand: $\frac{\partial q_1}{\partial k} < 0$. An increase in the price cap leads to $\frac{\partial p^c}{\partial p^w} < 0$ as the marginal cost for the producer decreases. Finally, we have $\frac{\partial q_1^w}{\partial p^w} = \left(1 + \frac{\partial p^c}{\partial p^w} \right) \frac{1}{p_q}$, which is a priori ambiguous: an increase of p^w decrease the negative price effect, which

put upward pressure on quantity, but as q_1^w defines the quantity at the price cap if p^w increases then it should decrease. Observe the previous expression of $\frac{\partial p^c}{\partial p^w}$ in Equation 9 and note that it is necessary below one due to the use of CDF. This implies that the second line is positive. The first term in the first line is negative as $\frac{\partial q_1^w}{\partial k} < 0$. We rewrite the second term such that:

$$\frac{\partial s_1}{\partial p^w} + \frac{\partial s_1^w}{\partial p^w} = 1 + 2 \frac{\partial p^c}{\partial p^w} = - \frac{1 - F(s_1^w) - F(s_1)}{1 - \Delta_1 F}$$

Hence a necessary condition for having $\frac{\partial^2 W_1^{bo}}{\partial k \partial p^w} < 0$, and therefore a decreasing k_1^{bo} , we must have an effect of the price cap lower than $\frac{1}{2}$. It also translates in the condition: $1 - F(s_1^w) - F(s_1) < 0$. This is also the condition such that the price cap has a convex effect on the capacity price. We rewrite the cross derivative such that:

$$\frac{\partial^2 W_1^{bo}}{\partial k \partial p^w} = \left[(\text{CS} + p^c(k)) f(s_1^w) \frac{\partial q_1^w}{\partial k} \left(1 - \frac{1 - F(s_1^w)}{F(s_1)} \right) - \int_{t_1^w}^{\bar{s}} \frac{\partial}{\partial k} (\text{CS} + p^c(k)) dF(s) \right] \frac{\partial q_1^w}{\partial p^w}$$

$\frac{\partial q_1^w}{\partial p^w} < 0$ so the term in brackets need to be positive if $\frac{\partial^2 W_1^{bo}}{\partial k \partial p^w} < 0$. If the cross derivative is negative, then recall that $1 - F(s_1^w) - F(s_1) < 0$. Therefore: $\frac{1 - F(s_1^w)}{F(s_1)} < 1$. Hence, if the following condition holds:

$$- \frac{\partial}{\partial k} \int_{t_1^w}^{\bar{s}} (\text{CS} + p^c(k)) dF(s) > (\text{CS} + p^c(k)) f(s_1^w) \frac{\partial s_1^w}{\partial k}$$

The cross derivative is negative. Due to the concavity of the expected welfare, this also implies that k_1^{bo} is decreasing with p^w . As at the upper limit of p^{w+} $k_1^{bo} = k_0^*$, then for any p^w below the limit we have $k_1^{bo} > k_0^*$.

□

10 Proof of Lemma 2.6

Proof. The profit function of retailers is defined as follows:

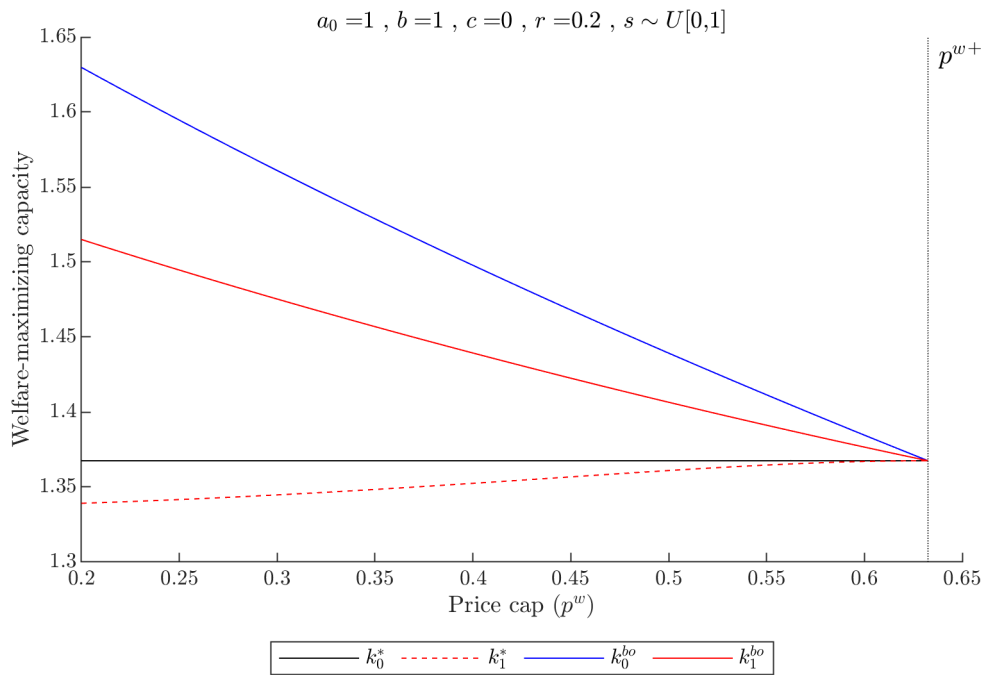


Figure C.3: welfare maximizing investment level with respect to k , different market design regime and assumptions. Note that the first-best is independent of the price cap; under the endogenous regime, the level of investment is increasing. When inefficient rationing is introduced, the investment level under both regimes decreases in the price cap.

$$\pi_i^r(q_i, k) = q_i(p(q) - p^s) - p^c(k)k \frac{q_i}{q_i + q_{-i}}$$

Given the retail market structure, the first-order condition of the profit function under a competition *à la Cournot* implies that

$$p(q) + q_i p_q - p^s - p^c(k)k \frac{q_{-i}}{(q_i + q_{-i})^2} = 0$$

And the second order condition:

$$2p_q - p^c(k)k \frac{q_{-i}}{(q_i + q_{-i})^3} = 0$$

The cross derivative of the profit function with respect to the competitor q_{-i} is:

$$p_q - p^c(k)k \frac{q_{-i} - q_i}{(q_i + q_{-i})^3} = 0$$

With n symmetric retailers and q the total quantity, the last equation becomes:

$$p_q + p^c(k)k \frac{(n-2)}{n} \frac{1}{q^2} = 0$$

The condition for existence requires that the cross derivative be positive, which establishes the condition of the lemma. The stability and the uniqueness of the equilibrium are given by the second-order condition, which is always negative.

□

11 Proof of Proposition 2.4

Proof. The proof follows similar arguments to the proof of Proposition 2.2 in Appendix 6. The expected welfare function is given by:

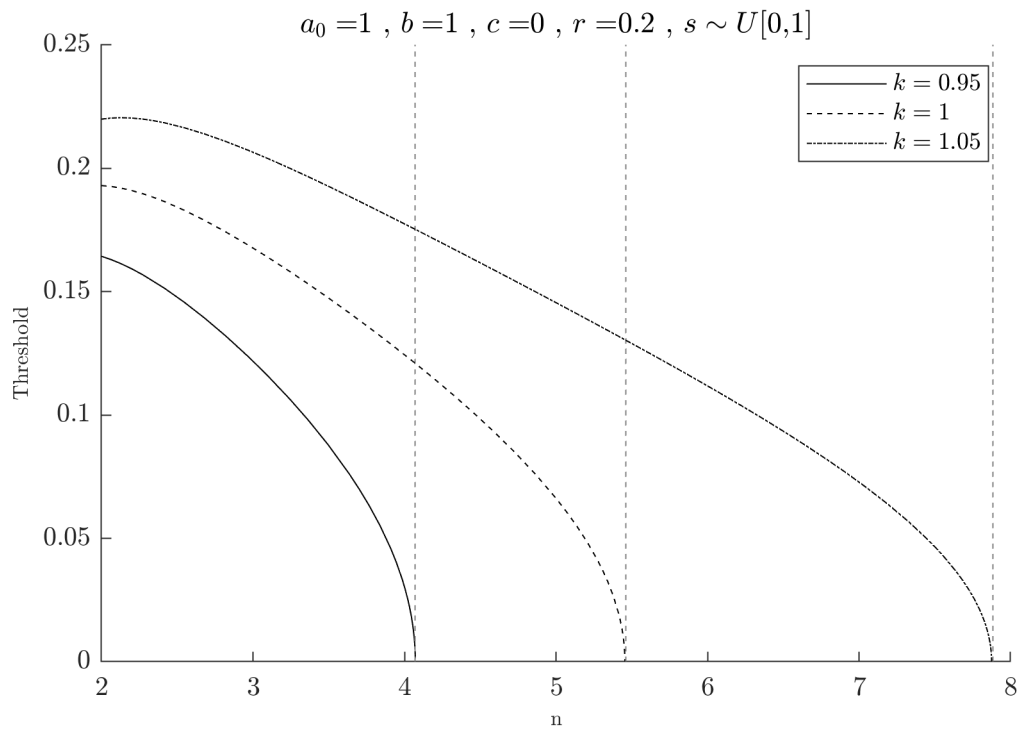


Figure C.4: Illustration of the thresholds in terms of quantity: for a given n , it gives the minimal quantity such that below the demand is increasing in q . The vertical lines are the lower limit such that any n above their value implies a demand function decreasing everywhere for a positive quantity.

$$W_n(k) = \int_0^{s_n} \int_0^{q_n} (p(q, s) - c) dq dF(s) + \int_{s_n}^{\bar{s}} \int_0^k (p(q, s) - c) dq dF(s) - I(k)$$

And the first derivative with respect to k is equal to:

$$\frac{\partial W_n(k)}{\partial k} = \int_0^{s_n} \frac{\partial q_n}{\partial k} (p(q_n, s) - c) dF(s) + \int_{s_n}^{\bar{s}} (p(k, s) - c) dF(s) - r$$

And the second derivative with respect to k :

$$\begin{aligned} \frac{\partial^2 W_n(k)}{\partial k^2} &= \int_0^{s_n} \left(\frac{\partial^2 q_n}{\partial k^2} (p(q_n, s) - c) + p_q \left(\frac{\partial q_n}{\partial k} \right)^2 \right) dF(s) + \int_{s_n}^{\bar{s}} p_q dF(s) \\ &\quad + (p(k, s_n) - c) f(s_n) \left(\frac{\partial q_n}{\partial k} - 1 \right) \frac{\partial s_n}{\partial k} \end{aligned}$$

We have:

$$\frac{\partial s_n}{\partial k} = -\frac{p_q + mp_q}{1 - \mathbb{C}_{p^c}(k) \Delta_n F} > 0$$

In that case $\Delta_n F = F(s_n^w) - F(s_n) \in [0, 1]$. We note $\mathbb{C}(q) = p^c k^{\frac{n-1}{nq}}$ the cost associated to the capacity market on the retailers and $\mathbb{C}_{p^c}(k) = \frac{\partial \mathbb{C}(k)}{\partial p^c}$ the aggregate negative effect of the capacity market on the demand function. Hence $\mathbb{C}_{p^c}(k) = \frac{n-1}{n} \in [0, 1]$. Therefore, the denominator is positive. The numerator is negative. Hence, the term is positive: accounting for all the indirect effects of increasing k still reduces the occurrences of off-peak periods. We have :

$$\frac{\partial q_n}{\partial k} = 1 - \frac{p_q + mp_q}{(p_q + mp_q + \mathbb{C}_k(k)) (1 - \mathbb{C}_{p^c}(k) \Delta_n F)} < 0$$

With $\mathbb{C}_k(k) = \frac{\partial \mathbb{C}(k)}{\partial k}$. Note that $p_q + mp_q + \mathbb{C}_k(k) > 0$ as this is the derivative of the demand function with respect to q . $\mathbb{C}_k(k) = p^c \frac{n-1}{n} > 0$. The denominator and the numerator are negative, and we also have $|(p_q + mp_q + \mathbb{C}_k(k))(1 - \Delta_n F)| < |p_q + mp_q|$. As by definition $(1 - \Delta_n F) < 1$. Hence, the second term is greater than 1, which implies

that the derivative is negative. Therefore, the second line of $\frac{\partial^2 W_n(k)}{\partial k^2}$ is also negative. From the definition of q_n The second derivative with respect to k of the quantity is given by the following condition:

$$\tilde{p}_{qq} \left(\frac{\partial q_n}{\partial k} \right)^2 + 2\tilde{p}_{qk} \frac{\partial q_n}{\partial k} + \tilde{p}_{kk} + \tilde{p}_q \frac{\partial^2 q_n}{\partial k^2} = 0$$

After simplification, we have:

$$\frac{\partial^2 q_n}{\partial k^2} = \frac{\mathbb{C}_{qq} \left(\frac{\partial q_n}{\partial k} \right)^2 + 2 \left(\mathbb{C}_{kq} + \mathbb{C}_{p^c q} \frac{\partial p^c}{\partial k} \right) \frac{\partial q_n}{\partial k} + 2\mathbb{C}_{p^c k} \frac{\partial p^c}{\partial k}}{\tilde{p}_q}$$

With $\mathbb{C}_{qq} = \frac{2kp^c(n-1)}{nq^3} > 0$, $\mathbb{C}_{kq} = -\frac{p^c(n-1)}{nq^2} < 0$, $\mathbb{C}_{p^c q} = -\frac{k(n-1)}{nq^2} < 0$ and $\mathbb{C}_{p^c k} = \frac{n-1}{nq} > 0$. We have shown above that $\frac{\partial p^c}{\partial k} > 0$, $\tilde{p}_q < 0$ and $\frac{\partial q_n}{\partial k} < 0$. Hence, the numerator is positive, which implies that $\frac{\partial^2 q_n}{\partial k^2} < 0$. This completes the proof that $\frac{\partial^2 W_n(k)}{\partial k^2}$ is negative. The expected social welfare is concave, and a unique maximum exists.

Finally, the proof for the ranking is straightforward and relies on the depressing effect generated at the margin on retailers' profit function. In this case, this form of allocation is similar to an increase in the marginal cost of production passed onto consumers.

Similar to the endogenous market design, this additional marginal cost is sustained whatever the realization of the demand level for final consumers is. Therefore, it lowers the quantity bought in the off-peak periods and lowers the prices in the peak periods. However, the effect is not as significant as in the endogenous case, as, under our framework, retailers do take into their welfare part of the capacity cost allocation, and they do not fully transfer this new marginal cost onto the consumers.

□

12 Proof of Corollary 2.1

Proof. The proof for the exogenous and endogenous regime with Cournot competition is discussed in the Technical Appendix 1. The derivative of the expected social welfare with respect to n is equal to:

$$\frac{\partial W_n(k)}{\partial n} = \int_0^{s_n} \frac{\partial q_n}{\partial n} (\tilde{p}(q_n, s) - c) dF(s)$$

Hence, it is sufficient to show that if the quantity at the marginal cost q_n is increasing in n , the expected welfare is also increasing in n . To do so, we derive in the following equation the first derivative with respect to n of the capacity price :

$$\frac{\partial p^c(k)}{\partial n} = -(mp_n(k) - \mathbb{C}_n(k)) \frac{\Delta_n F}{1 - \mathbb{C}_{p^c}(k) \Delta_n F}$$

This clearly shows the ambiguity of n on the equilibrium as under our framework $mp_n = \frac{bk}{n^2} > 0$ and $\mathbb{C}_n = \frac{p^c}{n^2} > 0$, the rest of the terms being positive. However, we do find that $bk > p^c(k)$.⁶ Therefore, the derivative is always negative for any value of k of interest. In other terms, increasing n always decreases the capacity price for a given value of k . The effect of n on the quantity is given is the following expression:

$$\frac{\partial q_n(s)}{\partial n} = - \frac{mp_n(q_n) - \mathbb{C}_n(q_n) - \mathbb{C}_z(q_n) \frac{\partial p^c(k)}{\partial n}}{\tilde{p}_q}$$

We have shown that the derivative of the price is negative, and we have $\mathbb{C}_z(q) = \frac{k}{q} \frac{n-1}{n} > 0$. Therefore, the depressing effect of n on the capacity price is positively impacting the quantity (recall that demand should be decreasing in q , hence $\tilde{p}_q < 0$). However, compared to the analysis of the capacity price, we do not have a clear-cut answer for the term $mp_n(q_n) - \mathbb{C}_n(q_n)$. This creates the ambiguity of n on the quantity at the marginal cost, which in turn can lead to a negative effect on welfare.

⁶Closed-form model gives the level of investment that implies that for any value we have $p^c < bk$ is equal to:

$$\frac{\mathbb{D}}{2p^w - 2c - \bar{s}}$$

The level of investment such that the capacity price is null is given.

$$\frac{n}{n+1} \frac{\mathbb{D}}{p^w - c}$$

With $\mathbb{D} = -(2a_0c - 2a_0p^w + 2c\bar{s} - 2p^w\bar{s} + 2r\bar{s} - c^2 + p^{w2}) \frac{1}{2b}$. Clearly, for any $n \geq 2$, we have the second value greater than the first one. As only the investment level that matters in our analysis is above the second value (the supply function is increasing in k), it implies that we always have $p^c < bk$ for the level of investment of interest.

Finally, the extension to the inefficient rationing case is straightforward, as we have, for instance.

$$\frac{\partial W_{1,n}}{\partial n} = \int_0^{s_{1,n}^{1,n}} \frac{\partial q_{1,n}}{\partial n} (q_{1,n} - c) dF(s) - \int_{s_{1,n}^w}^{\bar{s}} \frac{\partial q_{1,n}^w}{\partial n} \frac{k}{(q_{1,n}^w)^2} \int_0^{q_{1,n}^w} (p(q, s) - p(q_{1,n}, s)) dq dF(s)$$

□

13 Proof of Lemma 2.7

Proof. We start by noting that the supply function is increasing k . Under imperfect competition and exogenous regime, the marginal value of an investment is given by :

$$\phi_{0,n}^w(k) = \int_{s_{0,n}}^{s_{0,n}^w(k)} (p(k, s) - mp(k) - c) dF(s) + \int_{s_{0,n}^w(k)}^{\bar{s}} (p^w - c) dF(s)$$

When the capacity is binding but not the price cap, then the producer receives the retail price minus the retailer's markup. the derivative with respect to k gives:

$$\frac{\partial \phi_{0,n}^w(k)}{\partial k} = \int_{s_{0,n}}^{s_{0,n}^w(k)} (p_q - mp_q) dF(s)$$

As all the values at the integral thresholds cancel out. As p_q and mp_q are negative, the additional investment level decreases the expected marginal revenue. Hence, the supply function $X_{0,n} = r - \phi_{0,n}^w$ in the capacity market is positive. The upper bound r of the supply function comes from the fact that for a sufficiently high value of k : $s_{0,n} = \bar{s}$. Hence at this limit $X_{0,n} = r$. The lower limit of 0 from the fact that if there exists a private (inefficient) equilibrium without a capacity market, then $X_{0,n} = 0$ at this level by definition of the supply functions. For any value below the equilibrium, the supply function is null.

We now turn towards the demand function, which is equal to the expression:

$$\begin{aligned}\phi^d(k) &= \int_{s_{0,n}}^{\bar{s}} (mp(k) + kmp_q)dF(s) + \int_{s_{0,n}^w(k)}^{s_d(k)} \left(T(k, s) + k\frac{\partial T}{\partial k} \right) dF(s) + \\ &+ kSf(s_d)\frac{\partial s_d}{\partial k} + \int_{s_d(k)}^{\bar{s}} SdF(s)\end{aligned}$$

Deriving by k gives

$$\begin{aligned}\frac{\partial \phi^d(k)}{\partial k} &= - (kmp_q + mp(k))f(s_{0,n})\frac{\partial s_{0,n}}{\partial k} + Sf(s_d)\frac{\partial s_d}{\partial k} \\ &+ \int_{s_{0,n}}^{\bar{s}} 2mp_qdF(s) + \int_{s_{0,n}^w}^{s_d} 2(p_q - mp_q)dF(s)\end{aligned}$$

Both lines have an ambiguous sign. Indeed, $\frac{\partial s_{0,n}}{\partial k}$ and $\frac{\partial s_d}{\partial k}$ are positive following the same arguments as in the previous proofs, and $mp_q > 0$ and $p_q < 0$. Within this, to derive sufficient conditions, we compute the derivative of the demand function at a relatively low investment value such that the capacity always binds in expectation: $s_{0,n} = 0$. In that case, the derivative is equal to

$$\mathbb{V} = Sf(s_d)\frac{\partial s_d}{\partial k} + \int_0^{\bar{s}} 2mp_qdF(s) + \int_{s_{0,n}^w}^{s_d} 2(p_q - mp_q)dF(s)$$

Which is equal to $\frac{b}{n\bar{s}}(\bar{s} - S(n+1))$. Therefore, the derivative is negative if and only if $S > \frac{\bar{s}}{n+1}$. Then note that

We rewrite the first derivative when the capacity does not always bind:

$$\frac{\partial \phi^d(k)}{\partial k} = - (kmp_q + mp(k))f(s_{0,n})\frac{\partial s_{0,n}}{\partial k} - \int_{s_{0,n}}^{\bar{s}} 2mp_qdF(s) + \mathbb{V}$$

Hence, if $\mathbb{V} < 0$, the derivative is also always negative. Then note that at $k = 0$, we have:

$$\phi^d(k) = \int_{s_{0,n}^w(k)}^{s_d(k)} T(k, s) dF(s) + \int_{s_d(k)}^{\bar{s}} S dF(s)$$

Which is positive. Hence, if the demand function is decreasing, then a unique value of k exists, which always intersects with the supply function. Note that this value can be such that the capacity price is null.

□

14 Proof of Proposition 2.5

Proof. We start by studying the variation of the capacity market equilibrium equilibrium with respect to the penalty. By definition, it is the solution of the following equality:

$$\phi^d(k) = r - \phi_{0,n}^w$$

From the previous proof, we derive each expression with respect to S :

$$\frac{\partial \phi^d(k)}{\partial S} = \int_{s_d}^{\bar{s}} 1 dF(s)$$

Which is clearly positive. Finally, note from Equation 2.11 in the previous lemma that $\phi_{0,n}^w$ is independent of S . In the same Lemma, we have shown that the supply function is increasing in k , and we provide sufficient conditions to ensure a decreasing demand function. Using the condition that gives the equilibrium k_d , we can deduce the derivative of the equilibrium with respect to s as follows:

$$\frac{\partial k_d}{\partial S} = - \frac{\frac{\partial \phi_d}{\partial s}}{\frac{\partial \phi_{0,n}^w}{\partial k} + \frac{\partial \phi_d}{\partial k}}$$

From the previous observations, this is clearly always positive. Note that by definition: $p_s(k, s_d) - S = p^w$, therefore: $\frac{\partial s_d}{\partial S} = \frac{1}{p_s} > 0$ and $\frac{\partial s_d}{\partial k} = -\frac{p_q}{p_s} > 0$. Hence, an increase of S always increases s_d at the market equilibrium, which implies that $\frac{\partial \phi^d(k)}{\partial S}$ converges

to 0. It also proves that when S is such that $s_d = \bar{s}$, then the derivative of the market equilibrium is null. The second derivative of the market equilibrium with respect to the penalty value is equal to:

$$\frac{\partial^2 k_d}{\partial S^2} = - \frac{\frac{\partial \phi_d}{\partial S \partial k} \frac{\partial k_d}{\partial S} + \frac{\partial^2 \phi_d}{\partial S^2} + \left(\frac{\partial \phi_d}{\partial S \partial k} + \frac{\partial^2 \phi_d}{\partial k^2} \frac{\partial k_d}{\partial S} \right) \frac{\partial k_d}{\partial S}}{\frac{\partial \phi_{0,n}^w}{\partial k} + \frac{\partial \phi_d}{\partial k}}$$

We have shown that $\frac{\partial \phi^d(k)}{\partial S}$ is decreasing both in k and in S . Hence, the numerator is negative, and so is the denominator. Hence, the second derivative is concave. Therefore, when the first derivative with respect to S is null, this is also a maximum. We turn now to the expected welfare with inefficient rationing. We have:

$$\begin{aligned} W_d(k, S) &= \int_0^{s_{n,0}(k)} \int_0^{q_{n,0}} (p(q, s) - c) dq dF(s) \\ &+ \int_{s_{n,0}(k)}^{s_d(k)} \int_0^k (p(q, s) - c) dq dF(s) + \int_{s_d(k)}^{\bar{s}} \int_0^k (p^w - c) dq dF(s) \\ &+ \int_{s_d(k)}^{\bar{s}} \frac{k}{q_{n,0}^w} \int_0^{q_{n,0}^w} (p(q, s) - p^w) dq dF(s) - rk \end{aligned}$$

The derivative with respect to S gives:

$$\frac{\partial W_d(k, S)}{\partial S} = - \int_{s_d}^{\bar{s}} \left(\frac{\partial q_d}{\partial S} \frac{k}{q_d^2} \int_0^{q_d} (p(q, s) - p(q_d, s)) dq \right) dF(s)$$

Recall that the definition of q_d is given by having $p_s(q_d, s) - S = p^w$. Hence, $\frac{\partial q_d}{\partial S} = \frac{1}{p_q - mp_q} < 0$: an increase in the penalty value decreases the quantity asked by consumers. This implies that $\frac{\partial W_d(k, S)}{\partial S} = \frac{\partial CS_d(k, S)}{\partial S} < 0$. Note also that when S reaches the level such that $s_d = \bar{s}$, then the Depression is null. This implies that at this level and the equilibrium, we have :

$$\frac{\partial CS_d(k_d, S)}{\partial S} + \frac{\partial W_d(k_d, S)}{\partial k} \frac{\partial k_d}{\partial S} = 0$$

This also shows that this is the maximum of the expected social welfare.



Appendix D

Appendix - Chapter 3

Contents

| | | |
|----|------------------------------------|-----|
| 1 | Proof of Proposition 3.1 | 317 |
| 2 | Proof of Proposition 3.2 | 321 |
| 3 | Proof of Corollary 3.1 | 324 |
| 4 | Proof of Lemma 3.1 | 325 |
| 5 | Proof of Proposition 3.3 | 327 |
| 6 | Proof of Proposition 3.4 | 332 |
| 7 | Proof of Lemma 3.2 | 345 |
| 8 | Proof of Proposition 3.5 | 348 |
| 9 | Proof of Proposition 3.6 | 351 |
| 10 | Proof of Lemma 3.3 | 357 |

Notation guide: For readability, we sometimes drop the references to variables in the expression of functions and their derivatives.

1 Proof of Proposition 3.1

Proof. The maximization program of the market designer is equal to the following:

$$\begin{aligned} \max_{\substack{t_i(\theta,s) \rightarrow \mathbb{R}^+, \\ q_i(\theta,s) \rightarrow \mathbb{R}^+, \\ k \geq 0}} \quad CS(k) &= \sum_i \mu_i \int_s \int_{\theta_i} (U(q_i(\theta, s), \theta, s) - t_i(\theta, s)q_i(\theta, s)) dG_i(\theta) dF(s) \\ \text{s.t.} \quad I(k) &\leq \sum_i \mu_i \int_s \int_{\theta_i} t_i(\theta, s)q_i(\theta, s) dG_i(\theta) dF(s), \tag{R} \\ \sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) &\leq k, \tag{K} \end{aligned}$$

We simplify the problem by noting that the revenue constraint is always binding, as increasing quantity benefits consumer surplus. Then, we can rewrite the problem as follows:

$$\begin{aligned} \max_{\substack{t_i(\theta,s) \rightarrow \mathbb{R}^+, \\ q_i(\theta,s) \rightarrow \mathbb{R}^+, \\ k \geq 0}} \quad U^*(k) &= \sum_i \mu_i \int_s \int_{\theta_i} U(q_i(\theta, s), \theta, s) dG_i(\theta) dF(s) - I(k) \\ \text{s.t.} \quad \sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) &\leq k, \tag{K} \end{aligned}$$

The corresponding Lagrangian is equal to

$$\mathcal{L} = \sum_i \mu_i \int_s \int_{\theta_i} U(q_i(\theta, s), \theta, s) dG_i(\theta) dF(s) + \varepsilon \left(k - \sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) \right)$$

The associated first-order conditions are equal to

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial k} &= r - \varepsilon = 0 \\ \frac{\partial \mathcal{L}}{\partial q_i} &= \int_s u(q_i(\theta, s), \theta, s) dF(s) - \varepsilon = 0 \\ \frac{\partial \mathcal{L}}{\partial \varepsilon} &= k - \sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) = 0\end{aligned}$$

Which follows that $r = \varepsilon$. Hence, when capacity is not binding $\varepsilon = 0$ it implies that the expected $\int_s u(q_i(\theta, s), \theta, s) dF(s) = 0$. The marginal utility is always positive or null. Therefore, $u(q_i(\theta, s), \theta, s) = 0$ for every allocation when the capacity is not binding.

When capacity is binding $\varepsilon > 0$, it implies that $\int_s u(q_i(\theta, s), \theta, s) = r$. We show next that the optimal quantity is increasing with s . The derivative of the marginal utility at the optimal quantity for off-peak periods is equal to:

$$\frac{\partial u(q_i(\theta, s), \theta, s)}{\partial s} = \frac{\partial u(q_i(\theta, s), \theta, s)}{\partial q_i} \frac{\partial q_i(\theta, s)}{\partial s} + \frac{\partial u(q_i(\theta, s), \theta, s)}{\partial s}$$

From the initial assumptions, we have $\frac{\partial u(q_i(\theta, s), \theta, s)}{\partial q_i} < 0$ and $\frac{\partial u(q_i(\theta, s), \theta, s)}{\partial s} > 0$. Hence, for the first-order condition to hold, we must have $\frac{\partial q_i(\theta, s)}{\partial s} > 0$. We then define s_1 as the world's first state for which the capacity may be binding: when $\varepsilon = 0$ we have: $\sum_i q_i(\theta, s_1) = k$.

The optimal quantities are found as follows. Due to the linear form, when k is not binding: $u(q_i(\theta, s), \theta, s) = 0$, implying under the linear assumption that: $q_i(\theta, s) = \theta + s$. The aggregate quantity for a category i is then equal to $\mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) = \mu_i(\theta_i^{av} + s)$. We can deduce the value of the threshold s_1 such that when $s = s_1$, capacity starts to bind. Total quantity off-peak is equal to $\sum_i \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) = \sum_i \mu_i(\theta_i^{av} + 2s)$. Then at s_1 we have $\sum_i \theta_i^{av} + 2s_1 = k$, implying that $s_1 = \frac{k - \sum_i \mu_i \theta_i^{av}}{2}$. When k is binding, we use both the capacity constraint and the fact that the expected utility is equal for every consumer as $\int_s u(q_i(\theta, s), \theta, s) = r$. The capacity constraint gives: $\int_{\theta_j} q_j(\theta, s) dG_j(\theta) = \frac{1}{\mu_j} \left(k - \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) \right)$. Hence, the expected utility for a consumer from category j can be expressed as :

$$\begin{aligned} \mu_j \int_s \int_{\theta_j} u(q_j(\theta, s), \theta, s) dG_j(\theta) dF(s) &= \mu_j \int_s \int_{\theta_j} (\theta + s - q_j(\theta, s)) dG_j(\theta) dF(s) \\ &= \mu_j \left[\int_s \int_{\theta_j} (\theta + s) dG_j(\theta) dF(s) - \int_s \frac{1}{\mu_j} \left(k - \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) \right) dF(s) \right] \end{aligned}$$

Moreover, we know that $\int_s u(q_i(\theta, s), \theta, s) dF(s) = r$ for every consumer and category, hence we have $\int_s \int_{\theta_i} u(q_i(\theta, s), \theta, s) dG_i(\theta) dF(s) = \int_s \int_{\theta_j} u(q_j(\theta, s), \theta, s) dG_j(\theta) dF(s)$. This, with the previous result, implies that

$$\mu_j \int_s \int_{\theta_i} u(q_i(\theta, s), \theta, s) dG_i(\theta) dF(s) = \mu_j \int_s \int_{\theta_j} u(q_j(\theta, s), \theta, s) dG_j(\theta) dF(s)$$

After calculation:

$$\begin{aligned} \mu_j \int_s \int_{\theta_i} u(q_i(\theta, s), \theta, s) dG_i(\theta) dF(s) &= \mu_j \left[\int_s (\theta_j^{av} + s) dF(s) - \right. \\ &\quad \left. \frac{1}{\mu_j} \int_s \left(k - \mu_i \int_{\theta_i} q_i(\theta, s) dG_i(\theta) \right) dF(s) \right] \end{aligned}$$

Ultimately:

$$\int_s \int_{\theta_i} q_i(\theta, s) dG_i(\theta) dF(s) = k + \mu_j (\theta_i^{av} - \theta_j^{av})$$

It implies that the aggregate quantity is equal to $\mu_i \int_s \int_{\theta_i} q_i(\theta, s) dG_i(\theta) dF(s) = k\mu_i + \mu_i \mu_j (\theta_i^{av} - \theta_j^{av})$. We denote $q_{i,1}^*(\theta, s)$ and $q_{i,2}^*(\theta, k, s)$ the optimal allocation when the capacity is respectively not binding and binding for a consumer from category i . Also note that we do not have a discontinuity between quantities at s_1 : $q_{i,1}^*(\theta, s_1) = q_{i,2}^*(\theta, k, s_1)$. We then rewrite the objective function:

$$U^*(k) = \sum_i \mu_i \left[\int_0^{s_1} \int_{\theta_i} U(q_{i,1}^*(\theta, s), \theta, s) dG_i(\theta) dF(s) + \int_{s_1}^{\bar{s}} \int_{\theta_i} U(q_{i,2}^*(\theta, k, s), \theta, s) dG_i(\theta) dF(s) \right] - I(k)$$

The first-order condition gives the optimal investment level:

$$\begin{aligned} \frac{\partial U^*(k)}{\partial k} &= \sum_i \mu_i \int_{s_1}^{\bar{s}} \int_{\theta_i} u(q_{i,2}^*(\theta, k, s), \theta, s) \frac{\partial q_{i,2}^*(\theta, k, s)}{\partial k} dG_i(\theta) dF(s) - r \\ &= \sum_i \mu_i \int_{s_1}^{\bar{s}} \int_{\theta_i} u(q_{i,2}^*(\theta, k, s), \theta, s) dG_i(\theta) dF(s) - r \end{aligned}$$

This implies that the optimal investment level is found such that the expected aggregate marginal utility should equal the marginal investment costs. We now describe the optimal payment $t_i^*(\theta, s)$, which should be such that

$$\sum_i \mu_i \left[\int_0^{s_1} \int_{\theta_i} t_i^*(\theta, s) q_{i,1}^*(\theta, s) dG_i(\theta) dF(s) + \int_{s_1}^{\bar{s}} \int_{\theta_i} t_i^*(\theta, s) q_{i,2}^*(\theta, k, s) dG_i(\theta) dF(s) \right] = I(k)$$

It turns that if $t_i^*(\theta, s) = \{t_{i,1}^*(\theta, s), t_{i,2}^*(\theta, s)\}$ with $t_{i,1}^*(\theta, s) = 0$ when $s \leq s_1$ and $t_{i,2}^*(\theta, s) = \sum_i \mu_i \int_{\theta_i} u(q_{i,2}^*(\theta, k, s), \theta, s) dG_i(\theta) dF(s)$ when $s > s_1$, then by construction at the optimal investment level, the expected revenue is null.

We end the proof by showing that if we assume a spot market where consumers adjust their quantity given a price, the optimal allocation mechanism is equivalent to the spot market allocation. To see it, note first that the demand function on the spot market is equal to the inverse of the marginal utility function: $d(t, \theta, s) = u^{-1}(q, \theta, s)$. Therefore, when the price is equal to the marginal cost when capacity is not binding, we find that $q_i(\theta, s) = d(0, \theta, s) = \theta + s$, the same as the optimal allocation. When the capacity is binding, the quantity supply curve is a vertical line equal to k . The aggregate inverse demand on the market $p(q, s)$ is given by the solution of $\sum_i \mu_i \int_{\theta_i} d(t, \theta, s) dG_i(\theta) = q$. At k the inverse demand is equal to $\sum_i \mu_i \int_{\theta_i} u(k, \theta, s) dG_i(\theta) dF(s)$ which is the optimal unit transfer found in the optimal allocation at the optimal investment level.

□

2 Proof of Proposition 3.2

Proof. The objective function of the market designer is equal to:

$$\begin{aligned}
 \max_{t_i(\theta, s) \rightarrow \mathbb{R}^+} \quad CS(k) &= \sum_i \mu_i \int_s \int_{\theta_i} (U(d_i(t_i, \theta, s), \theta, s) - t_i(\theta, s)d_i(t_i, \theta, s)) dG_i(\theta) dF(s) \\
 \sum_i \mu_i \int_{\theta_i} d_i(t_i, \theta, s) dG_i(\theta) &\leq k, \tag{K} \\
 0 &\leq d_i(t_i, \theta, s)
 \end{aligned}$$

We start with the discrete case by assuming only two consumers defined by θ_1 and θ_2 . Then, we expand the results to the continuous case. First, note that the consumer surplus is maximized when the price equals the marginal cost. That is when it is null. Hence, we define the first case when the capacity is not binding with an optimal price $t_i^{cs}(\theta, s) = 0$. Optimal quantity is then given by $q_i^{cs}(\theta, s) = \theta_i + s$. This corresponds also to the first-best case. Therefore, the first state of the world when the capacity binds is identical to s_1 . Then, when the capacity binds, we have $k = d(t_1^{cs}, \theta_1, s) + d(t_2^{cs}, \theta_2, s) = \theta_1 + s - t_1^{cs} + \theta_2 + s - t_2^{cs}$, implying that $t_2^{cs}(t_1^{cs}) = \sum_i \theta_i + 2s - t_1^{cs} - k$. We define the realized consumer surplus given a value of s

$$\tilde{CS}(t_1^{cs}) = U(d(t_1, \theta_1, s), \theta_1, s) - t_1^{cs} d(t_1, \theta_1, s) + U(d(t_2^{cs}(t_1^{cs}), \theta_2, s), \theta_2, s) - t_2^{cs} d(t_2^{cs}(t_1^{cs}), \theta_2, s)$$

We then show that this consumer surplus exhibits a U shape with respect to k . The first-order condition is equal to:

$$\begin{aligned}
\frac{\partial \tilde{C}S(t_1^{cs})}{\partial t_1^{cs}} &= d_t u(d(t_1^{cs}, \theta_1, s), \theta_1, s) + \frac{\partial t_2^{cs}}{\partial t_1^{cs}} d_t u(d(t_2^{cs}(t_1^{cs}), \theta_2, s), \theta_2, s) \\
&\quad - d(t_1^{cs}, \theta_1, s) - \frac{\partial t_2^{cs}}{\partial t_1^{cs}} d(t_2^{cs}(t_1^{cs}), \theta_2, s) - t_1^{cs} d_t - t_2^{cs}(t_1^{cs}) \frac{\partial t_2^{cs}}{\partial t_1^{cs}} d_t \\
&= d(t_2^{cs}(t_1^{cs}), \theta_2, s) - d(t_1^{cs}, \theta_1, s)
\end{aligned}$$

With the linear assumption: $u(d(t)) = t$, $d_t = \frac{\partial d(t, \theta, s)}{\partial t} = -1$ and $\frac{\partial t_2^{cs}}{\partial t_1^{cs}} = -1$ and the second order condition :

$$\frac{\partial^2 \tilde{C}S(t_1^{cs})}{\partial t_1^{cs2}} = d_t \frac{\partial t_2^{cs}}{\partial t_1^{cs}} u_q(d(t_2^{cs}(t_1^{cs}), \theta_2, s), \theta_2, s) - d_t u_q(d(t_1^{cs}, \theta_1, s), \theta_1, s) = 2 > 0$$

With $u_q = \frac{\partial u(q, \theta, s)}{\partial q} = -1$. Therefore, the consumer surplus for a given value of s at the capacity constraint is convex. Then, when $t_1^{cs} = 0$, $t_2^{cs}(0) = \sum_i \theta_i + 2s - k$, implying that $\frac{\partial \tilde{C}S(t_1^{cs})}{\partial t_1^{cs}} \Big|_{t_1^{cs}=0} = d(t_2^{cs}(0), \theta_2, s) - d(0, \theta_1, s) = d(t_2^{cs}(0), \theta_2, s) - d(0, \theta_1, s) = \theta_2 + s - t_2^{cs}(0) - \theta_1 - s = k - 2s - 2\theta_1$. At $s_1 = k - (\theta_1 + \theta_2)/2$, which implies that $\frac{\partial \tilde{C}S(t_1^{cs})}{\partial t_1^{cs}} \Big|_{t_1^{cs}=0, s=s_1} = -k + \theta_2 - \theta_1$. Hence, if $\theta_1 > \theta_2$, the derivative is negative at $t_1^{cs} = 0$. Note finally that as s increases, the derivative decreases, so the value is always negative. We then find the upper value of \tilde{t}_1^{cs} such that $t_2^{cs}(\tilde{t}_1^{cs}) = 0$. That is when $t_1^{cs} = \sum_i \theta_i + 2s - k$. In that case, $\frac{\partial \tilde{C}S(t_1^{cs})}{\partial t_1^{cs}} \Big|_{t_1^{cs}=\tilde{t}_1^{cs}} = \theta_2 + s - \theta_1 - s + \tilde{t}_1^{cs} = 2\theta_2 - \theta_1 + 2s - k$. At s_1 : $\frac{\partial \tilde{C}S(t_1^{cs})}{\partial t_1^{cs}} \Big|_{t_1^{cs}=\tilde{t}_1^{cs}, s=s_1} = k + 3/2\theta_2 - 3/2\theta_1$, which is positive only if the value of the lowest type is sufficiently high compared to the highest type. This confirms the U shape of the consumer surplus. We illustrate this observation in Figure D.1. Finally, we compare the consumer surplus for both prices: $\tilde{C}S(0) - \tilde{C}S(\tilde{t}_1^{cs}) = (\theta_1 - \theta_2)(2s - k + \sum_i \theta_i)$. At $s = s_1$ we have : $k(\theta_1 - \theta_2)$ which is always positive when $\theta_1 > \theta_2$, and the derivative is positive with respect to s . Hence, the value is always positive. This proves that it is always more efficient to give a price equal to the marginal cost to the highest type and set $t_2^{cs}(0)$ for the lowest type. To conclude, we study the case when, for a given set of parameters, we have $t_2^{cs}(0) = \theta_2 + s$. In that case, for relatively higher values of s , prices cannot increase as the demand is null for the lowest type. To ensure the capacity constraint, it is now optimal to raise the price of the highest consumer type. In that case, the price for consumer 2 is $\theta_2 + s$. The price for consumer 1 is found by having $k = d(t_1^{cs}, \theta, s)$, that is $t_1^{cs} = \theta_1 + s - k$.

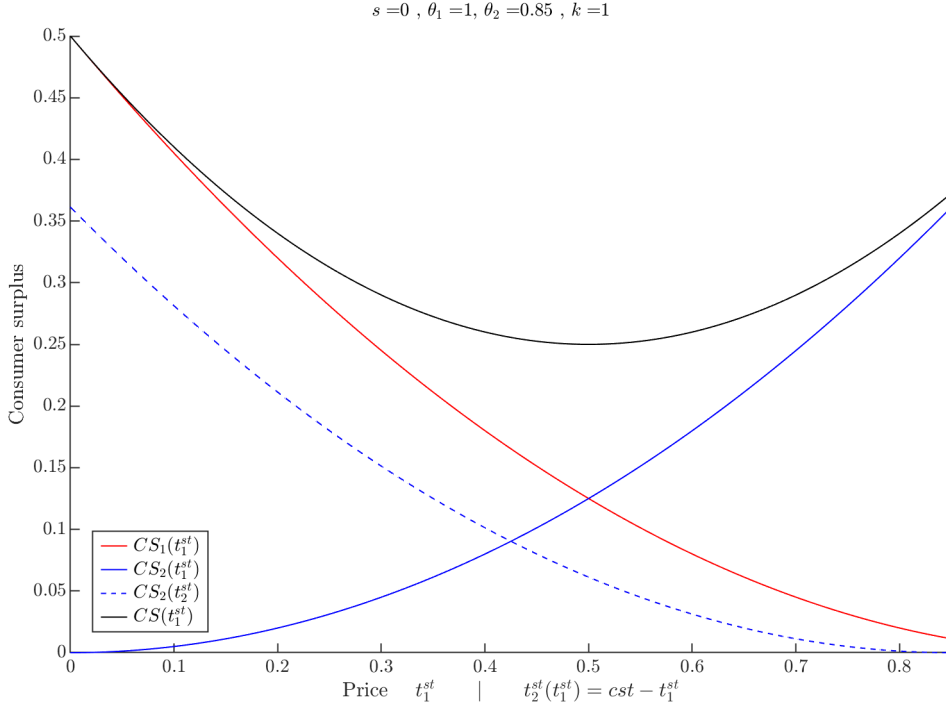


Figure D.1: Illustration of the U shape of the consumer surplus

The extension to the continuous case needs some restrictions. Indeed, we assume a continuous distribution from type θ and the state of the world s . Therefore, contrary to the discrete case, we cannot pin down a unique consumer that needs to be rationed first, followed by other consumers as s increases. To overcome this restriction, we proceed as follows. First, we define the case when the capacity is not binding, which boils down to the first-best allocation. It is defined by the inequality : $\sum_i \int_{\theta_i} d(0, \theta, s) dG_i(\theta) \leq k$.

When the capacity starts to bind for $s > s_1$, we use the definition of category 1 being a higher type of category 2 and the fact that for an equal type, a consumer from both categories brings the same utility. Assume that we have $\underline{\theta}_2 < \underline{\theta}_1 < \bar{\theta}_2 < \bar{\theta}_1$. We start by studying the case for a relatively small increase of s such that $s = s_1 + \epsilon$. This small increase boils down to having the possibility to provide sufficient goods to all consumers from category 1 and from consumers of category 2 having the same type that some of the consumers from category 2:

$$Q_0^1(s_1 + \epsilon) < k < \sum_i \mu_i \int_{\theta_i} d(0, \theta, s_1 + \epsilon) dG_i(\theta)$$

With $Q_0^1(s) = \mu_1 \int_{\theta_1} d(0, \theta, s_1 + \epsilon) dG_1(\theta) + \mu_2 \int_{\bar{\theta}_2}^{\bar{\theta}_1} d(0, \theta, s_1 + \epsilon) dG_2(\theta)$. In that case, and similar to the discrete case, we impose a price above marginal cost to the consumers having lower types from category 2 such that $\mu_2 \int_{\bar{\theta}_2}^{\bar{\theta}_1} d(t_2^{st}(\theta, s), \theta, s) dG_2(\theta) + Q_0^1(s) = k$. As s increases, we have $Q_0^s = k$. It is not possible to ration smaller consumers as their demand is null. Each consumer for every $\theta \in [\theta_2, \bar{\theta}_1]$ receives the following price $t_2cs(\theta, s) = \theta + s$. Therefore, the market designer imposes a price equal to the marginal cost to the consumers of the highest type from category 1: $\theta \in [\bar{\theta}_2, \bar{\theta}_1]$ and a uniform price for the smaller consumers of both categories having $\theta \in [\theta_1, \bar{\theta}_2]$, as they both bring the same utility even though they are from different category. The total quantity asked by the highest consumer having a price equal to marginal costs is equal to $Q_0^2s = \int_{\bar{\theta}_2}^{\bar{\theta}_1} d(0, \theta, s) dG_1(\theta)$. Hence, the situation corresponds to the case where $s = s_1 + \epsilon$ such that:

$$Q_0^2(s_1 + \epsilon) < k < Q_0^1(s_1 + \epsilon)$$

In that case, we impose a price above marginal cost to the consumers having lower types from category 1 and 2 such that $\sum \mu_i \int_{\theta_1}^{\bar{\theta}_2} d(t_i^{st}(\theta, s), \theta, s) dG_i(\theta) + Q_0^2(s) = k$. We conclude with the case when s is sufficiently high such that $Q_0^2(s_1 + \epsilon) = k$. Now, consumers from both categories with a type below $\bar{\theta}_2$ receive the price $\theta + s$, and their quantity equals 0. The price is above the marginal cost for the highest type from consumers of category 1 such that $\int_{\bar{\theta}_2}^{\bar{\theta}_1} d(t_1^{cs}, \theta, s) dG_1(\theta) = k$. The different values of $s_i(k)$ are given by the previous equalities delimiting each case.

□

3 Proof of Corollary 3.1

Proof. From the proof of Proposition 3.1, we know that the aggregate quantity is equal to

$$\mu_i \int_s \int_{\theta_i} q_i(\theta, s) dG_i(\theta) dF(s) = k\mu_i + \mu_i\mu_j(\theta_i^{av} - \theta_j^{av})$$

Which is confirmed by the fact that $k\mu_i + \mu_i\mu_j(\theta_i^{av} - \theta_j^{av}) + k\mu_j + \mu_i\mu_j(\theta_j^{av} - \theta_i^{av}) = k$. This gives the ratio:

$$\alpha_i^*(k) = \mu_i + \frac{\mu_i \mu_j (\theta_i^{av} - \theta_j^{av})}{k}$$

From the proof of Proposition 3.2, we know that the individual price depends on which consumers receive a null price. Namely, if s is sufficiently low, no consumers are rationed; for the middle range of s , either smaller consumers from the smaller category are rationing, followed by the smaller of both categories. If s is sufficiently high, all smaller consumers are driven out of the market, and only the bigger consumers from the bigger consumers are rationed. In the second case, the individual quantity for bigger consumers is equal to $s + \theta_1$. With the continuous case, the aggregate consumer quantity is equal to $\mu_1(s + \theta^{av})$, which directly gives the ratio $\alpha_1^{st}(s, k)$ and the ratio for the rest of the available quantity $\alpha_2^{st}(s, k)$. The following expression gives the second ratio:

$$\mu_i \int_{\bar{\theta}_2}^{\bar{\theta}_1} d(0, \theta, s) dG_1(\theta) = \frac{\bar{\theta}_1 - \bar{\theta}_2}{\bar{\theta}_1 - \bar{\theta}_1} \left(2s + \frac{\bar{\theta}_1 + \bar{\theta}_2}{4} \right)$$

□

4 Proof of Lemma 3.1

Proof. Proposition 3.2 shows that in the short-term framework, there exist three thresholds concerning the states of the world. It implies that depending on the value of k , seven different expressions of the consumer surplus could exist. Note first that $s_1 < s_2(k) < s_3(k)$. Then the seven expressions depends on the values of k such that $s_1 = 0$, $s_2(k) = 0$, $s_3(k) = 0$, and $s_1 = \bar{s}$, $s_2(k) = \bar{s}$, $s_3(k) = \bar{s}$. The case where $s_1 = \bar{s}$, for instance, corresponds to the value of the parameter such that the capacity never begins in expectation and prices are always null. On the other hand, the case where $s_3(k) = 0$ implies that every smaller consumer from both categories is always excluded from the market. The first extreme case leads to the consumer surplus being independent of k . Therefore, only six possible maxima can exist. Note that by construction, there is no discontinuity in terms of consumer welfare between the different cases. The average case, that is, when the three thresholds are both positive and below \bar{s} , is described below:

$$\begin{aligned}
& \mu_2 \int_{s_1}^{s_2(k)} \int_{\theta_2}^{\theta_1} \left(t_2^{st}(\theta, s) \frac{\partial t_2^{st}(\theta, s)}{\partial k} \right) dG_2(\theta) dF(s) \\
& + \sum_i \mu_i \int_{s_2(k)}^{s_3(k)} \int_{\theta_1}^{\bar{\theta}_2} \left(t_i^{st}(\theta, s) \frac{\partial t_i^{st}(\theta, s)}{\partial k} \right) dG_i(\theta) dF(s) \\
& + \mu_1 \int_{s_3(k)}^{\bar{s}} \int_{\bar{\theta}_2}^{\bar{\theta}_1} \left(t_1^{st}(\theta, s) \frac{\partial t_1^{st}(\theta, s)}{\partial k} \right) dG_1(\theta) dF(s) - r
\end{aligned}$$

All four other cases are derived from the above expression by sequentially canceling the different terms. The convexity of consumer welfare in this framework is due to the opposite effect of an increase in the level of investment. In the case when the capacity always binds, for instance, we have:

$$\mu_1 \int_0^{\bar{s}} \int_{\bar{\theta}_2}^{\bar{\theta}_1} \left(t_1^{st}(\theta, s) \frac{\partial t_1^{st}(\theta, s)}{\partial k} \right) dG_1(\theta) dF(s) - r$$

Under uniform distribution and the linear marginal utility, the revenue is concave in k . Indeed, we have $t_1^{st} = s + \frac{\bar{\theta}_1 + \bar{\theta}_2}{2} - 2k \frac{\bar{\theta}_1 - \theta_1}{\theta_1 - \theta_2}$.¹ In that case the derivative of the price with respect to k is negative and equal to $\frac{\partial t_1^{st}(\theta, s)}{\partial k} = -2 \frac{\bar{\theta}_1 - \theta_1}{\theta_1 - \theta_2}$. Differentiating $t_1^{st}(\theta, s) \frac{\partial t_1^{st}(\theta, s)}{\partial k}$ with respect to k leads to the expression $-\frac{2(\bar{\theta}_1 - \theta_1)(2\bar{\theta}_1 - 2\theta_1)}{2(\theta_1 - \theta_2)^2}$, which is also negative. Hence, the above expression is concave. On the other hand, in the case when the capacity binds less, we have

$$\mu_2 \int_{s_1}^{\bar{s}} \int_{\theta_2}^{\theta_1} \left(t_2^{st}(\theta, s) \frac{\partial t_2^{st}(\theta, s)}{\partial k} \right) dG_2(\theta) dF(s) - r$$

The second-order derivative is in:

¹The value of the price is found by equalizing the quantity consumed at t_1^{st} by consumers having a type between $\bar{\theta}_2$ and θ_1 to the level of investment as described in Proposition 3.2.

$$\begin{aligned}
& -\mu_2 f(s_1) \frac{\partial s_1}{\partial k} \int_{\theta_2}^{\theta_1} \left(t_2^{st}(\theta, s_1) \frac{\partial t_2^{st}(\theta, s_1)}{\partial k} \right) dG_2(\theta) \\
& + \mu_2 \int_{s_1}^{\bar{s}} \int_{\theta_2}^{\theta_1} \left(\left(\frac{\partial t_2^{st}(\theta, s)}{\partial k} \right)^2 + t_2^{st}(\theta, s) \frac{\partial^2 t_2^{st}(\theta, s)}{\partial k^2} \right) dG_2(\theta) dF(s)
\end{aligned}$$

Calculations show that the second term is negative.² The first term is positive due to the fact that $\frac{\partial s_1}{\partial k} > 0$. Indeed, s_1 satisfies: $\sum_i \mu_i \int_{\theta_i} d(0, \theta, s) dG_i(\theta) - k = 0$. Therefore, we have;

$$\sum_i \mu_i \int_{\theta_i} \left(\frac{\partial s_1}{\partial k} \frac{\partial d(0, \theta, s)}{\partial s} \right) dG_i(\theta) = 1 \quad \Rightarrow \quad \frac{\partial s_1}{\partial k} = 1$$

As $\frac{\partial d(0, \theta, s)}{\partial s} = 1$. This directly implies that more investment means less binding capacity. We also have $\frac{\partial t_2^{st}(\theta, s)}{\partial k} < 0$. This creates the non-concavity effect in the consumer surplus. Essentially, the convexity arises because as k increases, the concavity of the revenue is reduced due to the decrease in the occurrence of a binding capacity, which is precisely when the concavity of the revenue is observed.

The absence of a closed-form solution to this framework prevents having clear-cut answers on the existence and the coexistence of the different maximma. Figure D.2 illustrates the different maximma that can exist for a given set of parameters and different values of the investment costs. The black line represents the first-best investment level, and the red lines are some maximma under the short-term framework.

□

5 Proof of Proposition 3.3

Proof. The expected revenue from consumers is built from the expected off-peak revenue, during which the quantity bought by consumers depends on the unique price implemented by the market designer, and from the expected on-peak revenue, during which

²We have $t_2^{st}(\theta, s) = (s + \theta) \frac{(2\bar{\theta}_1 - \theta_1 - \theta_2)(4s - 4s_1)}{(\theta_1 - \theta_2)(4k + \theta_1 + \theta_2 - \theta_1 - \theta_2)}$

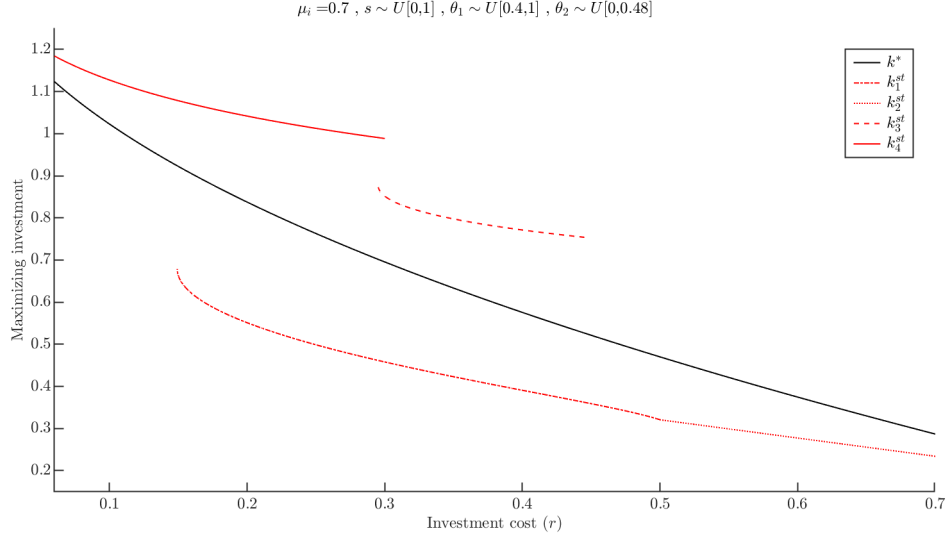


Figure D.2: Consumer surplus maximizing investment level for different values of investment cost

the aggregate quantity is by definition equal to the level of investment. The expected off-peak revenue is equal to :

$$\sum_i \mu_i \int_0^{s_0^r(k, t^r)} \int_{\theta_i} (t^r d(t^r, \theta, s)) dG_i(\theta) dF(s)$$

The expected on-peak revenue is equal to :

$$\sum_i \mu_i \int_{s_0^r(k, t^r)}^{\bar{s}} \int_{\theta_i} (t^r k) dG_i(\theta) dF(s)$$

With $s_0^r(k, t^r)$ defined as the first state of the world such that the capacity is binding:

$$\sum_i \mu_i \int_{\theta_i} d(t^r, \theta, s_0^r(k, t^r)) dG_i(\theta) = k$$

Because there is only one decision variable to choose from, and the market designer maximizes the consumer surplus under revenue constraint, the revenue constraint fully pinned down the optimal transfer such that it solves:

$$R_0^r(k, t^r) = \sum_i \mu_i \left[\int_0^{s_0^r(k, t^r)} \int_{\theta_i} (t^r d(t^r, \theta, s)) dG_i(\theta) dF(s) + \int_{s_0^r(k, t^r)}^{\bar{s}} \int_{\theta_i} (t^r k) dG_i(\theta) dF(s) \right] - I(k) \Rightarrow R_0^r(k, t^r) = 0 \quad (D.1)$$

There exists at most three solutions to the problem depending on the value of k : (i) when the capacity never binds in expectation ($s_0^r(k, t^r) = \bar{s}$), (ii) when the capacity always binds that is ($s_0^r(k, t^r) = 0$), and (iii) the middle case as illustrated in Equation D.1. In case (ii), the solution is straightforward such that:

$$\sum_i \mu_i \int_0^{\bar{s}} \int_{\theta_i} (t^r k) dG_i(\theta) dF(s) = I(k) \Rightarrow t^r = \frac{I(k)}{k} \quad (D.2)$$

The optimal transfer is the average investment cost. Under our framework, it is equal to $t^r = r$ for a value of $k^-(r)$ such that $s_0^r(k^-(r)) = 0$, which is $k^-(r) = \sum_i \mu_i \theta_i^{av} - r$. The other extreme case is given under our framework by the expression:

$$\sum_i \mu_i \int_0^{\bar{s}} \int_{\theta_i} (t^r d(t^r, \theta, s)) dG_i(\theta) dF(s) = I(k) \Rightarrow t^r \left(\frac{\bar{s}}{2} - t^r - \sum_i \mu_i \theta_i^{av} \right) = I(k)$$

The first derivative with respect to t^r is given by:

$$\frac{\partial R_0^r(k, t^r)}{\partial t^r} = \frac{\bar{s}}{2} - 2t^r - \sum_i \mu_i \theta_i^{av} \quad (D.3)$$

And the maximum is given by $t^r = \frac{1}{2} (\frac{\bar{s}}{2} - \sum_i \mu_i \theta_i^{av})$. The second derivative is equal to -2 , which is negative, confirming the expected revenue is concave in t^r . Note also that the expected revenue is independent of k . This implies that the revenue net of investment cost is always decreasing in k . Discarding the case of corner solutions and lower limit of k such that $s_0^r(k, t^r) = \bar{s}$, a solution to the problem exists if and only if the net revenue at the maximum is positive, that is only when $k < \frac{(\frac{\bar{s}}{2} + \sum_i \mu_i \theta_i^{av})^2}{4r}$. The interpretation of this limit is that beyond it, the investment cost is such that the price

to cover it leads to a negative consumer surplus. That is, the reaction to the price from the consumer prevents the financing of such an investment level.

For the middle case, the closed-form solution does not exist. Still, the principles remain the same: a value t^r exists for only a relatively low value of k such that it leads to positive or null demand during off-peak periods. We now study how the optimal value $t^r(k)$ behaves with respect to k . To do so, we use the implicit function theorem, which gives the following:

$$\frac{\partial t^r(k)}{\partial k} = -\frac{\partial R_0^r(k)}{\partial k} / \frac{\partial R_0^r(k)}{\partial t^r}$$

We start with the derivative of the revenue with respect to k :

$$\frac{\partial R_0^r(k)}{\partial k} = \int_{s_0^r(k, t^r)}^{\bar{s}} t^r dF(s) - r$$

Which apparently has an ambiguous sign. We prove below that the derivative has to be negative at the optimal value of t^r . From Equation D.1, we rewrite:

$$R_0^r(k) = \sum_i \mu_i \int_0^{s_0^r(k, t^r)} \int_{\theta_i} (t^r d(t^r, \theta, s)) dG_i(\theta) dF(s) + \left(\sum_i \mu_i \int_{s_0^r(k, t^r)}^{\bar{s}} t^r dF(s) - r \right) k$$

Hence, the second term in parenthesis is precisely the derivative of the revenue with respect to k :

$$R_0^r(k) = \sum_i \mu_i \int_0^{s_0^r(k, t^r)} \int_{\theta_i} (t^r d(t^r, \theta, s)) dG_i(\theta) dF(s) + \frac{\partial R_0^r(k)}{\partial k} k$$

As $t^r d(t^r, \theta, s) > 0$, we must have $\frac{\partial R_0^r(k)}{\partial k} < 0$ to have $R_0^r(k) = 0$. We turn now to the derivatives with respect to t^r ; the first derivative with respect to t^r is equal to :

$$\frac{\partial R_0^r(k)}{\partial t^r} = \sum_i \mu_i \left[\int_0^{s_0^r(k, t^r)} \int_{\theta_i} \left(d(t^r, \theta, s) + t^r \frac{\partial d(t^r, \theta)}{\partial t^r} \right) dG_i(\theta) dF(s) + \int_{s_0^r(k, t^r)}^{\bar{s}} k dF(s) \right]$$

Which has an ambiguous sign similar to Equation D.3 of the extreme case when the capacity never binds. We show next that the derivative at $k^-(r)$, defined in the previous extreme case when the capacity always binds and $t^r = r$, is in we have $s_0^r(k^-(r), t^r) = 0$, hence:

$$\frac{\partial R_0^r(k^-(r))}{\partial t^r} = \sum_i \mu_i \int_0^{\bar{s}} k dF(s) > 0$$

Note that for a given level of investment, it is always better to have the same revenue with a lower price to maximize consumer surplus. Therefore, if R_0^r is concave in t^r and as $R_0^r(k)$ is decreasing in k , the optimal price is decreasing in k . The second derivative with respect to t^r is equal to:

$$\begin{aligned} \frac{\partial^2 R_0^r(k)}{\partial t^{r2}} &= f(s_0^r(k, t^r)) \frac{\partial s_0^r(k, t^r)}{\partial t^r} \left(\sum_i \mu_i \int_{\theta_i} \left(d(t^r, \theta, s_0^r(k, t^r)) + t^r \frac{\partial d(t^r, \theta)}{\partial t^r} \right) dG_i(\theta) - k \right) \\ &\quad + \sum_i \mu_i \left[\int_0^{s_0^r(k, t^r)} \int_{\theta_i} \left(2 \frac{\partial d(t^r, \theta)}{\partial t^r} + t^r \frac{\partial^2 d(t^r, \theta)}{\partial t^{r2}} \right) dG_i(\theta) dF(s) \right] \end{aligned}$$

Under linear marginal utility, we have $\frac{\partial^2 d(t^r, \theta)}{\partial t^{r2}} = 0$, which ensure that the revenue during off-peak is concave in t^r . Next, note that, by definition, we have the equality $\sum_i \mu_i \int_{\theta_i} d(t^r, \theta, s_0^r(k, t^r)) dG_i(\theta) = k$. This allows us to express the following derivative:

$$\sum_i \mu_i \int_{\theta_i} \left(\frac{\partial s_0^r(k, t^r)}{\partial t^r} \frac{\partial d(t^r, \theta)}{\partial s} + \frac{\partial d(t^r, \theta)}{\partial t^r} \right) dG_i(\theta) = 0$$

Hence:

$$\frac{\partial s_0^r(k, t^r)}{\partial k} = -\frac{\partial d(t^r, \theta)}{\partial t^r} / \frac{\partial d(t^r, \theta)}{\partial s} = 1$$

This allows us to rewrite the second derivative as:

$$\frac{\partial^2 R_0^r(k)}{\partial t^{r2}} = t^r f(s_0^r(k, t^r)) \frac{\partial s_0^r(k, t^r)}{\partial t^r} \frac{\partial d(t^r, \theta)}{\partial t^r} + 2 \int_0^{s_0^r(k, t^r)} \frac{\partial d(t^r, \theta)}{\partial t^r} dF(s)$$

Which is negative, so the expected revenue is concave in t^r . Hence $t^r(k)$ is increasing in k .

□

6 Proof of Proposition 3.4

Proof. We denote the consumer surplus under the policy as follows:

$$CS^r(k, t_i^r, t_j^r) = \sum_i \mu_i \left[\int_0^{s_1^r(k)} \int_{\theta_i} (U(d(t_i^r, \theta, s), \theta, s) - t_i^r d(t_i^r, \theta, s)) dG_i(\theta) dF(s) \right. \\ \left. + \sum_i \mu_i \int_{s_1^r(k)}^{\bar{s}} \int_{\theta_i} (U(\alpha_i^r(k)k, \theta, s) - t_i^r \alpha_i^r(k)k) dG_i(\theta) dF(s) \right] \quad (D.4)$$

With $\alpha_i^r(k) = 1 + \frac{\mu_j(\theta_i^{av} - \theta_j^{av})}{k}$ corresponding to the share of capacity allocated to category i . $s_1^r(k)$ is given by solving: $\sum \mu_i \int_{\theta_i} d(t_i^r, \theta, s_1^r(k)) dG_i(\theta) = k$. Which gives :

$$s_1^r(k) = k + \sum_i \mu_i (t_i^r - \theta_i^{av}) \quad (D.5)$$

The expected revenue net of investment costs is defined as:

$$R^r(k, t_i^r, t_j^r) = \sum_i \mu_i \left[\int_0^{s_1^r(k)} \int_{\theta_i} (t_i^r d(t_i^r, \theta, s)) dG_i(\theta) dF(s) \right. \quad (\text{D.6})$$

$$\left. + \sum_i \mu_i \int_{s_1^r(k)}^{\bar{s}} \int_{\theta_i} (t_i^r \alpha_i^r(k) k) dG_i(\theta) dF(s) \right] - I(k) \quad (\text{D.7})$$

Hence, the maximization problem is given in the following expression:

$$\begin{aligned} \max_{\substack{k, \\ t_i^r \rightarrow \mathbb{R}^+}} \quad & CS^r(k, t_i^r, t_j^r) \\ \text{s.t.} \quad & 0 \leq R^r(k, t_i^r, t_j^r), \end{aligned} \quad (\text{R})$$

The Lagrangian associated with the market designer program such that

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

With γ^r , the lagrangian multiplier is associated with the revenue constraint. The first-order condition of the program at the optimal value of $t_i^r(k)$ and $t_j^r(k)$ are equal to:

$$\begin{cases} CS_i^r(k, t_i^r, t_j^r) + \gamma(k) R_i^r(k, t_i^r, t_j^r) = 0 \\ CS_j^r(k, t_i^r, t_j^r) + \gamma(k) R_j^r(k, t_i^r, t_j^r) = 0 \end{cases}$$

With $CS_i^r(k, t_i^r, t_j^r) = \frac{\partial CS^r(k, t_i^r, t_j^r)}{\partial t_i^r}$ and $R_i^r(k, t_i^r, t_j^r) = \frac{\partial R^r(k, t_i^r, t_j^r)}{\partial t_i^r}$. Differentiating with respect to k the first equation and dropping the variables for clarity implies that:

$$CS_{ii}^r \frac{\partial t_i^r(k)}{\partial k} + \gamma(k) R_{ii}^r \frac{\partial t_i^r(k)}{\partial k} + CS_{ij}^r \frac{\partial t_j^r(k)}{\partial k} + \gamma(k) R_{ij}^r \frac{\partial t_j^r(k)}{\partial k} + CS_{ik}^r + \gamma(k) R_{ik}^r + \frac{\partial \gamma(k)}{\partial k} R_i^r = 0$$

With $CS_{ii}^r = \frac{\partial^2 CS^r}{\partial t_i^r{}^2}$, $R_{ii}^r = \frac{\partial^2 R^r}{\partial t_i^r{}^2}$, $CS_{ij}^r = \frac{\partial CS^r}{\partial t_i^r \partial t_j^r}$, $R_{ij}^r = \frac{\partial R^r}{\partial t_i^r \partial t_j^r}$, $CS_{ik}^r = \frac{\partial CS^r}{\partial t_i^r \partial k}$, $R_{ik}^r = \frac{\partial R^r}{\partial t_i^r \partial k}$. The equation simplifies to:

$$\mathcal{L}_{ii}^r(k, t_i^r, t_j^r) \frac{\partial t_i^r(k)}{\partial k} + \mathcal{L}_{ij}^r(k, t_i^r, t_j^r) \frac{\partial t_j^r(k)}{\partial k} + \mathcal{L}_{ik}^r(k, t_i^r, t_j^r) + \frac{\partial \gamma(k)}{\partial k} R_i^r(k, t_i^r, t_j^r) = 0$$

With $\mathcal{L}_{ii}^r, \mathcal{L}_{ij}^r$ and \mathcal{L}_{ik}^r the derivatives of the Lagrangian. Hence, the derivatives of the first-order condition are given by:

$$\begin{cases} \mathcal{L}_{ii}^r(k, t_i^r, t_j^r) \frac{\partial t_i^r(k)}{\partial k} + \mathcal{L}_{ij}^r(k, t_i^r, t_j^r) \frac{\partial t_j^r(k)}{\partial k} + \mathcal{L}_{ik}^r(k, t_i^r, t_j^r) + \frac{\partial \gamma(k)}{\partial k} R_i^r(k, t_i^r, t_j^r) = 0 \\ \mathcal{L}_{jj}^r(k, t_i^r, t_j^r) \frac{\partial t_j^r(k)}{\partial k} + \mathcal{L}_{ji}^r(k, t_i^r, t_j^r) \frac{\partial t_i^r(k)}{\partial k} + \mathcal{L}_{jk}^r(k, t_i^r, t_j^r) + \frac{\partial \gamma(k)}{\partial k} R_j^r(k, t_i^r, t_j^r) = 0 \end{cases}$$

Which implies :

$$\begin{cases} \frac{\partial t_i^r(k)}{\partial k} = - \left(\mathcal{L}_{ij}^r(k, t_i^r, t_j^r) \frac{\partial t_j^r(k)}{\partial k} + \mathcal{L}_{ik}^r(k, t_i^r, t_j^r) + \frac{\partial \gamma(k)}{\partial k} R_i^r(k, t_i^r, t_j^r) \right) \frac{1}{\mathcal{L}_{ii}^r(k, t_i^r, t_j^r)} \\ \frac{\partial t_j^r(k)}{\partial k} = - \left(\mathcal{L}_{ji}^r(k, t_i^r, t_j^r) \frac{\partial t_i^r(k)}{\partial k} + \mathcal{L}_{jk}^r(k, t_i^r, t_j^r) + \frac{\partial \gamma(k)}{\partial k} R_j^r(k, t_i^r, t_j^r) \right) \frac{1}{\mathcal{L}_{jj}^r(k, t_i^r, t_j^r)} \end{cases}$$

Let's note the determinant of the Hessian matrix of the Lagrangian $H^r = \mathcal{L}_{ii}^r \mathcal{L}_{jj}^r - \mathcal{L}_{ij}^r \mathcal{L}_{ji}^r$. Then the derivative of the optimal value of $t_i^r(k)$ with respect to k is equal to:

$$\frac{\partial t_i^r(k)}{\partial k} = \left(\mathcal{L}_{ik}^r(k, t_i^r, t_j^r) + R_i \frac{\partial \gamma(k)}{\partial k} - \rho_i \left(\mathcal{L}_{jk}^r(k, t_i^r, t_j^r) + R_j \frac{\partial \gamma(k)}{\partial k} \right) \right) \frac{-\mathcal{L}_{jj}^r}{H^r} \quad (\text{D.8})$$

With $\rho_i = \frac{L_{ij}}{L_{jj}}$. The revenue constraint gives the expression of the derivative of the Lagrangian multiplier. As it is binding, we have the revenue at the optimal values:

$$R^r(k, t_i^r(k), t_j^r(k)) = 0$$

Hence:

$$R_i^r(k, t_i^r(k), t_j^r(k)) \frac{\partial t_i^r}{\partial k} + R_j^r(k, t_i^r(k), t_j^r(k)) \frac{\partial t_j^r}{\partial k} + R_k^r(k, t_i^r(k), t_j^r(k)) = 0$$

From the previous findings on the derivative of the optimal values $t_i^r(k)$ and $t_j^r(k)$, we have:

$$\begin{aligned} & R_i^r(k, t_i^r(k), t_j^r(k)) \left(\mathcal{L}_{ik}^r(k, t_i^r, t_j^r) - \rho_i \mathcal{L}_{ij}^r(k, t_i^r, t_j^r) + (R_i - \rho_i R_j) \frac{\partial \gamma(k)}{\partial k} \right) \frac{-\mathcal{L}_{jj}}{H^r} \\ & + R_j^r(k, t_i^r(k), t_j^r(k)) \left(\mathcal{L}_{jk}^r(k, t_i^r, t_j^r) - \rho_j \mathcal{L}_{ji}^r(k, t_i^r, t_j^r) + (R_j - \rho_j R_i) \frac{\partial \gamma(k)}{\partial k} \right) \frac{-\mathcal{L}_{ii}}{H^r} \\ & \qquad \qquad \qquad + R_k^r(k, t_i^r(k), t_j^r(k)) = 0 \end{aligned}$$

It implies that:

$$\frac{\partial \gamma(k)}{\partial k} = -\frac{1}{bH^r} \left(\sum_i R_i (\mathcal{L}_{ij}^r \mathcal{L}_{jk}^r - \mathcal{L}_{jj}^r \mathcal{L}_{ik}^r) + R_k H^r \right)$$

In the next step, we formally prove that at least one minimum exists for the price of the smaller category. The critical insight of the proof relies on the behavior of prices in the special cases of the maximization program. That is when the capacity level is such that it always binds or it never binds in expectation. The following results show that the prices are always increasing with a relatively high level of k when the capacity never binds, and the price of the smaller consumer at the level of investment such that it always binds is always decreasing. For the price of the bigger consumer, the result is ambiguous as it may happen that for some values of the parameters, the price might decrease. We also provide a more detailed discussion of the rationale behind such behavior in the proof.

Lemma D.1. *Assume that for a given k , the optimal values $t_i^r(k)$ and $t_j^r(k)$ are such that $s_1^r(k) > \bar{s}$. Then $t_i^r(k)$ and $t_j^r(k)$ are increasing and convex in k . Moreover, if i is the bigger category then we have $t_j^r(k) > t_i^r(k)$.*

Proof. Assuming that the capacity never binds in expected significantly implies the analysis. All the cross derivatives are null, and the second derivatives for both the consumer surplus and the revenue with respect to k are also null. The derivative of the optimal price $t_i^r(k)$ is equal to:

$$\frac{\partial t_i^r(k)}{\partial k} = -\frac{\partial \gamma(k)}{\partial k} \frac{R_i}{\mathcal{L}_{ii}}$$

With $R_i > 0$. Indeed, the revenue is concave in t_i^r : $R_{ii} = -2\mu_i \int_s 1dF(s) < 0$. Hence, there are at most two possible values that can satisfy the revenue constraint. As the market designer always prefers smaller prices for consumers, then the chosen value is the minimum of the two, which, under a concave function, is located in the increasing part of the function. Next, we have $\mathcal{L}_{ii} = \mu_i(1 - 2\gamma(k))$. We prove that $\gamma(k) > 1$, hence $\mathcal{L}_{ii} < 0$. From the maximization problem, we have :

$$\gamma(k) = -\frac{CS_i^r}{R_i}$$

With

$$CS_i^r = -\mu_i \int_s \int_{\theta_i} d(t_i^r, \theta, s) dG_i(\theta) dF(s)$$

And

$$R_i^r = \mu_i \int_s \left(\int_{\theta_i} d(t_i^r, \theta, s) dG_i(\theta) + t_i^r \frac{\partial d(t_i^r, \theta)}{\partial t_i^r} \right) dF(s) \quad (\text{D.9})$$

Which boils down:

$$R_i^r = -CS_i^r + \mu_i \int_s t_i^r \frac{\partial d(t_i^r, \theta)}{\partial t_i^r} dF(s)$$

Hence :

$$\gamma(k) = \frac{CS_i^r}{CS_i^r - \mu_i \int_s t_i^r \frac{\partial d(t_i^r, \theta)}{\partial t_i^r} dF(s)} > 1$$

The derivative of $\gamma(k)$ is such that:

$$\frac{\partial \gamma(k)}{\partial k} = -\frac{R_k^r \mathcal{L}_{ii} \mathcal{L}_{jj}}{bH^r}$$

With bH^r , the determinant of the bordered Hessian matrix is positive at the maximum. $R_k = -r < 0$ as the capacity never binds in expectation. Hence, the revenue is independent of k . Finally $\mathcal{L}_{ii} \mathcal{L}_{jj} = \mu_i \mu_j (2\gamma(k) - 1)^2 > 0$. Therefore, we have: $\frac{\partial \gamma(k)}{\partial k} > 0$, and the price is increasing with k . Calculations give the second derivative with respect to k :

$$\frac{\partial^2 t_i^r(k)}{\partial k^2} = (3R_{jj}(\mathcal{L}_{ii} R_j^r)^2 + (\mathcal{L}_{jj} R_i^r)^2 R_{jj} - 2\mathcal{L}_{jj} R_{ii} \mathcal{L}_{ii} (R_i^r)^2) \frac{R_i^r \mathcal{L}_{jj} R_k^{r2}}{bH^{r3}}$$

The term in parenthesis is equal to

$$2(2\gamma(k) - 1)^2 \sum_i \mu_i^3 \mu_j^2 (2\bar{s} - 2t_i^r + \theta_i^{av})^2$$

Which is positive. We also have: $\mathcal{L}_{jj} = \mu_j (2\gamma(k) - 1) < 0$, and $R_i^r < 0$. Hence, the second derivative is positive, and $t_i^r(k)$ is convex in k . The ranking between prices is given by the closed-form solution of the difference between the two prices that can be expressed as follows:

$$t_i^r - t_j^r = \frac{\Delta \theta^{ab}}{2\mathbb{G}} (\mathbb{G} - (\mathbb{G}^2 - 16kr\mathbb{G})^{0.5}) > 0$$

With $\mathbb{G} = 4 \sum_i \mu_i (\theta_i^{av})^2 + \bar{s}(\bar{s} + 4 \sum_i \mu_i \theta_i^{ab})$ which is positive implying that the difference is also positive.

□

Note that as prices decrease as the level of investment decreases, then it implies that a value of k exists such that the investment starts binding in expectation. It ensures the existence of the middle case with $0 < s_1^r(k) < \bar{s}$. Next, we study the opposite extreme case where k is sufficiently low such that $s_1^r(k) = 0$. We have $s_1^r(k) = k + \sum \mu_i(t_i^r - \theta_i^{av})$. Recall that no price can be negative, and neither can the demand. Hence, we have : $t_i^r \in [0, \underline{\theta}_i]$. Therefore, $s_1^r(k) \leq k$. It implies that a $k^-(r)$ such that $s_1^r(k^-(r)) = 0$ exists. We provide in the following lemma the results with respect to the sign of the prices at $k^-(r)$

Lemma D.2. *When $k = k^-(r)$, the optimal value $t_j^r(k)$ for the smaller category is always decreasing. The optimal price derivative for the bigger category $t_i^r(k)$ is either decreasing or increasing depending on the parameters. The price of the smaller category is always higher than the price of the bigger category.*

Proof. We start by describing the solution of the optimization model at $k^-(r)$. Recall that the first-order condition is $CS_i^r + \lambda^r R_i^r = 0$. When $k = k^-(r)$, we express the consumer surplus derivative from Equation D.4 with $s_1 = 0$ as follows:

$$CS_i^r = \mu_i \mathbb{B} - \mu_i \alpha_i^r(k) k$$

With

$$\mathbb{B} = \sum_i \mu_i \int_{\theta_i} (CS(d(t_i^r, \theta, s_1^r(k)), \theta, s_1^r(k)) - CS(\alpha_i^r(k)k, \theta, s_1^r(k))) dG_i(\theta) f(s)$$

It corresponds to the consumer surplus adjustment between off-peak and on-peak periods as the price level changes. Due to the incomplete information framework, the inefficient rationing during on-peak periods implies a discontinuity of the surplus between the two periods. It is always positive, as $U(d(t_i^r, \theta, s_1^r(k)), \theta, s_1^r(k)) - U(\alpha_i^r(k)k, \theta, s_1^r(k)) = (\mu_i(t_i^r - t_j^r) - \theta + \theta_1^{av})^2 \frac{1}{2} > 0$. That is, increasing the level of investment always increases the on-peak quantity (second term) but also substitutes a lower on-peak surplus with a higher off-peak surplus. Similarly, from Equation D.6. The marginal revenue is given by :

$$R_i^r = \mu_i \mathbb{C} + \mu_i \alpha_i^r(k) k$$

With

$$\mathbb{C} = \sum_i \mu_i \int_{\theta_i} (t_i^r d(t_i^r, \theta, s_1^r(k)) - t_i^r \alpha_i^r(k) k) dG_i(\theta) f(s)$$

This expression is close to \mathbb{B} in the sense that it represents the cost adjustment for consumers due to the discontinuity of the individual consumer surplus. Note that we have $\mathbb{B} = \mathbb{D} - \mathbb{C}$, with $\mathbb{D} = \sum_i \mu_i \int_{\theta_i} (U(d(t_i^r, \theta, s_1^r(k)), \theta, s_1^r(k)) - U(\alpha_i^r(k) k, \theta, s_1^r(k))) dG_i(\theta) f(s)$, which is the utility adjustment. Therefore, we rewrite the Lagrange multiplier such that $\lambda^r = -\frac{CS_i^r}{R_i^r}$. It implies that: $\frac{CS_i^r}{R_i^r} = \frac{CS_j^r}{R_j^r}$. After rewriting the different terms, the optimal prices need to satisfy the following conditions:

$$(\mu_i \alpha_i^r(k) k - \mu_j \alpha_j^r(k) k) \mathbb{D} = 0$$

As only \mathbb{D} depends on the prices, it implies that at the maximum, we have $\mathbb{D} = 0$. Which allows us to rewrite $CS_i^r = -\mu_i \mathbb{C} - \mu_i \alpha_i^r(k) k$. From the expression of λ^r , we therefore have $\lambda^r = 1$. The first-order condition at $k^-(r)$ implies that $CS_i^r = -R_i^r$. Solving for two first-order conditions, plugging them in the function s_1 and solving for k gives:

$$k^-(r) = \frac{1}{2} \left(\theta^{av} - r + ((\theta^{av} - r)^2 - 4(\mu_i \mu_j \sigma)^{0.5} \Delta \theta^{av})^{0.5} \right)$$

With $\theta^{av} = \sum_i \mu_i \theta_i^{av}$ the weighted sum of average type. $\Delta \theta^{av} = \theta_i^{av} - \theta_j^{av} > 0$ as we assume i being the biggest category. And $\sigma = \sum_i \mu_i \sigma_i^2 > 0$, with $\sigma_i = \frac{(\bar{\theta}_i - \theta_i)^2}{12}$. Note that the threshold $k^-(r)$ is decreasing with r , as we also have $\theta^{av} - r > 0$. We now turn to express the values t_i^r and t_j^r under a closed-form solution such that it respects the case that $s_1(k) = 0$. To find such values, we used the two observations that at $k^-(r)$ we both have $R^r = 0$ and $s_1 = 0$. Solving the system and deriving the solution with respect to k gives:

$$\frac{\partial t_i^r}{\partial k} = \frac{1}{\mu_i \Delta \theta^{av}} (2k + r - (\theta_j^{av} + 2\mu_i \Delta \theta^{av}))$$

and

$$\frac{\partial t_j^r}{\partial k} = -\frac{1}{\mu_j \Delta \theta^{av}} (2k + r - (\theta_i^{av} - 2\mu_j \Delta \theta^{av}))$$

This leads to the following observation: if $2k + r - (\theta_j^{av} + 2\mu_i \Delta \theta^{av}) > 0$, it implies that whenever the price of the bigger category is increasing, then the price of the smaller category is always decreasing. Indeed, note that $(\theta_j^{av} + 2\mu_i \Delta \theta^{av}) - (\theta_i^{av} - 2\mu_j \Delta \theta^{av}) = \Delta \theta^{av} > 0$.

It is straightforward to see that the price of the bigger category is convex in k , and the price of the smaller category is concave in k . The rest of the proof is as follows: we show that the investment level that minimizes the price for the bigger category $k_i^-(r)$ is higher than the level $k_j^-(r)$ that maximizes the price for the bigger category and that $k^-(r)$ can be smaller than $k_i^-(r)$ but never smaller than $k_j^-(r)$. Hence, the smaller category always exhibits a negative derivative (as we always have $k^-(r) > k_j^-(r)$), while the price for the bigger category can be either increasing (when $k^-(r) > k_i^-(r)$) or decreasing (when $k^-(r) < k_i^-(r)$). Using the price derivative we have $k_i^-(r) = \frac{1}{2}(\theta_j^{av} + 2\mu_i \Delta \theta^{av} - r)$ and $k_j^-(r) = \frac{1}{2}(\theta_i^{av} - 2\mu_j \Delta \theta^{av} - r)$. Clearly, $k_i^-(r) > k_j^-(r)$ as $k_i^-(r) - k_j^-(r) = \frac{1}{2} \Delta \theta^{av} > 0$. Moreover, both investment levels are decreasing in r .

Simulations shows that the ranking between $k_i^-(r)$ and $k^-(r)$ is ambiguous. However, note that we have shown that $k^-(r)$ is decreasing in r and that the term inside the square root is also decreasing in r . As it must be positive, the limit toward the highest value r^- that can imply a solution is such that $\lim_{r \rightarrow r^-} k^-(r) = \frac{1}{2}(\theta^{av} - r^-)$. We now prove that $\theta^{av} - r^- > 0$. To see this, recall that t_j^r is concave in k and t_i^r is convex in k . Therefore, there are at most two intersections. In that case, those intersections are such that one of them implies that both prices equal r (see Figure D.3 for an illustration).³ As prices also satisfy the condition $s_1 = 0$, from Equation D.5, we can deduce that it implies at the corresponding k of the intersection $\theta^{av} - k - r = 0$. Therefore, for any values of k below the term, it is positive and, by extension, $\theta^{av} - r > 0$. From the expression of $k_j^-(r)$, it is clear that this limit is above it. Therefore, $k^-(r)$ is always above $k_j^-(r)$ and

³Calculations show that if both intersections exist, then one implies that both prices meet at $k = 0$. The other intersection is at an investment level, which is always above $k_i^-(r)$.

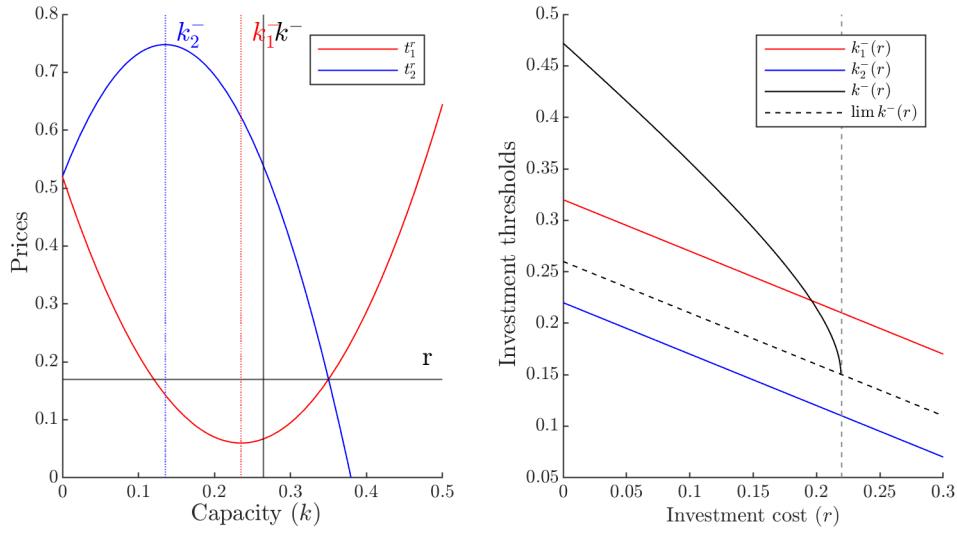


Figure D.3: Optimal prices and investment threshold such that $s_1 = 0$ and $R^r(k, t_i^r, t_j^r) = 0$

there is not intersection. Hence, the level of investment such that $s_1 = 0$ implies that the price for the smaller is always decreasing in k . We illustrate the part of this proof in the second panel of Figure D.3. We conclude the proof by expressing the difference between the two prices:

$$\frac{k}{\mu_i \mu_j \Delta \theta^{av}} (\theta^{av} - k - r)$$

From the discussion on the price intersection, we know that $\theta^{av} - k - r > 0$. Hence, the price difference is positive.

□

Now that we have formally shown that a minimum always exists for the smaller consumers, we provide in the rest of the proof a technical discussion on the rationale behind the behavior of the prices. To do so, we decompose the effect of k on the optimal values between two opposite effects:

- A consumer surplus effect.
- A revenue effect.

The non-monotonicities emerge from the tension between the two effects. Namely, the consumer effect dominates the revenue effect for relatively low values of k . As k increases, the revenue effect takes over the consumer effect. The central idea of the consumer surplus effect comes from the observation that an additional investment is always beneficial for consumers. To see this, we express the derivative of the consumer surplus with respect to k as follows:

$$CS_k^r(k, t_i^r, t_j^r) = \mathbb{B} + \sum_i \mu_i \int_{s_1^r(k)}^{\bar{s}} \int_{\theta_i} (u(\alpha_i^r(k)k, \theta, s) - t_i^r) dG_i(\theta) dF(s) \quad (\text{D.10})$$

The second derivative with respect to k is given by:

$$CS_{kk}^r(k, t_i^r, t_j^r) = - \int_{s_1^r(k)}^{\bar{s}} 1 dG_i(\theta) dF(s)$$

Which is negative. Hence, a higher level of investment supposes a lower marginal gain for the consumer surplus. This explains why the consumer surplus effect is relatively less significant for a higher level of investment. On the other hand, the revenue effect is convex in k . Namely, the revenue constraint is increasingly tighter as k increases, which necessitates higher prices. Two results can illustrate this effect: (i) Proposition 3.3, which shows that for a single price, an increase in k always implies an increase in $t^r(k)$, (ii) Lemma D.1 that shows for the extreme case such that when the capacity never binds in expectation, prices are also increasing and convex in k .

We turn toward the analysis of the behavior and the ranking between prices. Lemma D.1 proves that the prices are increasing in k for high values of k , which is reinforced by the results from the single-price policy. The ranking between the two prices, that is, having a higher price for the bigger category, can be illustrated by the following lemma:

Lemma D.3. *The pair of prices that maximizes the expected revenue when the capacity is always binding is always asymmetric with $t_i(k) > t_j(k)$ whenever the category i is the bigger category compared to category j .*

Proof. Under the assumption that the capacity is always binding, the expected revenue is equal to:

$$R^r(k, t_i^r, t_j^r) = \sum_i \mu_i \int_s \int_{\theta_i} (t_i^r d(t_i^r, \theta, s)) dG_i(\theta) dF(s) - I(k)$$

Which gives the first-order condition in Equation 6. Under the linear and uniform assumption, the derivatives imply that the price \tilde{t}_i^r that maximizes revenue is given by:

$$\tilde{t}_i^r = \frac{\bar{s} + \theta_i^{av}}{2}$$

This clearly implies that if the category i is bigger than j , then its average type is higher, and so is the price \tilde{t}_i^r . \square

We conclude this proof by showing that the consumer surplus effect can explain the decrease in the price for the smaller category. Namely, we find that for a revenue constraint independent of the investment level, the change in the consumer surplus with respect to k is sufficient for generating the decrease in price. This observation is supported by the fact that for a small value of k , the change in the revenue constraint is relatively smaller due to its convexity with respect to k . The decrease in the price of the smaller category is explained by the decrease in the marginal rate of substitution as k increases. The rate of substitution between the price of the bigger category relative to the smaller category for having the same consumer surplus is given by:

$$MRS_{i \rightarrow j}(k) = \frac{CS_i^r}{CS_j^r}$$

It implies that the MRS changes with respect to k as follows:

$$\frac{\partial MRS_{i \rightarrow j}(k)}{\partial k} = \frac{CS_{ik}^r CS_j^r - CS_{jk}^r CS_i^r}{CS_j^{r2}}$$

In the absence of a clear closed-form solution. We provide in the following lemma for the sign of $\frac{\partial MRS_{i \rightarrow j}(k)}{\partial k}$ for the symmetric case, namely when $t_i^r(k) = t_j^r(k)$

Lemma D.4. When $t_i^r(k) = t_j^r(k)$, the marginal rate of substitution $MRS_{i \rightarrow j}$ is decreasing in k .

Proof. The expression of the derivatives in $\frac{\partial MRS_{i \rightarrow j}(k)}{\partial k}$ are described below. For the first derivative with respect to t_i^r :

$$CS_i^r = \mu_i \mathbb{B} - \mathbb{E}q_i$$

With $\mathbb{E}q_i$ being the expected quantity for category i :

$$\mathbb{E}q_i = \mu_i \left[\int_0^{s_1^r(k)} \int_{\theta_i} d(t_i^r, \theta, s) dG_i(\theta) dF(s) + \int_{s_1^r(k)}^{\bar{s}} \alpha_i^r(k) k dG_i(\theta) dF(s) \right]$$

Then the cross derivative is equal to:

$$\begin{aligned} CS_{ik}^r &= -\frac{\partial \mathbb{E}q_i}{\partial k} \\ &= \mu_i \left[\left(\int_{\theta_i} d(t_i^r, \theta, s_1^r(k)) dG_i(\theta) - \alpha_i^r(k) k \right) - \int_{s_1^r(k)}^{\bar{s}} \frac{\partial \alpha_i^r(k) k}{\partial k} dF(s) \right] \\ &= \frac{\partial \mathbb{B}}{\partial t_i^r} - \mathbb{E} \partial_k q_i \end{aligned}$$

We have:

$$\frac{\partial \mathbb{B}}{\partial t_i^r} = \mu_i \mu_j (t_i^r - t_j^r) \frac{1}{\bar{s}}$$

Hence at $t_i^r = t_j^r$ the derivative is null. We then rewrite the MRS :

$$\frac{\partial MRS_{i \rightarrow j}(k)}{\partial k} = \frac{\mathbb{E} \partial_k q_j [(\mu_i - \mu_j) \mathbb{B} + \mathbb{E}q_j - \mathbb{E}q_i]}{CS_j^{r2}}$$

Note that $\mu_j \mathbb{E} \partial_k q_i = \mu_i \mathbb{E} \partial_k q_j = \int_{s_1}^{\bar{s}} \mu_i \mu_j dF(s)$. this allows us to simplify the expression to:

$$\frac{\partial MRS_{i \rightarrow j}(k)}{\partial k} = \frac{\mathbb{E} \partial_k q_i (\mathbb{E} q_j - \mathbb{E} q_i)}{CS_j^{r^2}}$$

If category i is the bigger category, then $\mathbb{E} q_i > \mathbb{E} q_j$ at $t_i^r = t_j^r$. This implies that $\frac{\partial MRS_{i \rightarrow j}(k)}{\partial k} < 0$.

□

□

7 Proof of Lemma 3.2

Proof. The proof summarizes the approach of [Spulber \(1992\)](#). It starts by showing that the participation and the revenue constraints can be merged to form the unique constraint noted $R - IR$. Given a strategy $\hat{\theta}$, an individual surplus of a consumer of type θ is given by:

$$CS(\theta, \hat{\theta}) = \int_s (U(q(\hat{\theta}, s), \theta, s) - t(\hat{\theta}, s)q(\hat{\theta}, s)) dF(s) \quad (\text{D.11})$$

If incentive compatibility holds, then $\hat{\theta} = \theta$. It implies that the first-order condition is such that $\frac{\partial CS}{\partial \hat{\theta}} = 0$ at $\hat{\theta} = \theta$. Using this condition and totally differentiating the first-order condition implies at $\hat{\theta} = \theta$: $\frac{\partial^2 CS}{\partial \hat{\theta}^2} + \frac{\partial^2 CS}{\partial \hat{\theta} \partial \theta} = 0$. If $\hat{\theta}$ is a maximum then $\frac{\partial^2 CS}{\partial \hat{\theta}^2} \leq 0$. Hence, we must also have : $\frac{\partial^2 CS}{\partial \hat{\theta} \partial \theta} \geq 0$. From Equation [D.11](#), we have

$$\frac{\partial^2 CS}{\partial \hat{\theta} \partial \theta} = \int_s \left(\frac{\partial u(q(\hat{\theta}, s), \theta, s)}{\partial \theta} \frac{\partial q(\hat{\theta}, s)}{\partial \hat{\theta}} \right) dF(s) = \int_s \frac{\partial q(\hat{\theta}, s)}{\partial \hat{\theta}} dF(s)$$

Hence, if the derivative of the individual quantity increases with the type, the solution of the first-order condition is also a maximum. The envelope theorem implies that the derivative of the consumer surplus at the maximum with respect to the type is given

solely by the derivative of the expected utility. Therefore, writing $CS(\theta)$ as the surplus when the consumer behaves honestly gives :

$$CS(\theta) = \int_s CS(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta} \frac{\partial U}{\partial \theta} d\tilde{\theta} = CS(\underline{\theta}_i) + \int_s \int_{\underline{\theta}_i}^{\theta} q(\tilde{\theta}, s) d\tilde{\theta} dF(s)$$

Isolating $CS(\underline{\theta}_i)$ gives : $CS(\underline{\theta}_i) = CS(\theta) - \int_s \int_{\underline{\theta}_i}^{\theta} q(\tilde{\theta}, s) d\tilde{\theta} dF(s)$. Note that it is by definition independent of any consumer type, hence: $CS(\underline{\theta}_i) = \int_{\theta_i} CS(\underline{\theta}_i) dG_i(\theta) = \int_{\theta_i} \left(CS(\theta) - \int_s \int_{\underline{\theta}_i}^{\theta} q(\tilde{\theta}, s) d\tilde{\theta} dF(s) \right) dG_i(\theta)$. The smallest type from both groups needs to satisfy the participation constraint, hence $\sum_i \mu_i \int_{\theta_i} CS(\underline{\theta}_i) dG_i(\theta) \geq 0$. Then, it remains to note that the net revenue also needs to be at least positive. Note R^m the expected revenue is given by :

$$R^m = \sum_i \mu_i \int_s \int_{\theta_i} (t_i(\theta, s) q_i(\theta, s)) dG_i(\theta) dF(s) - I(k) \geq 0 \quad (\text{D.12})$$

Combining both conditions leads to the following:

$$\sum_i \mu_i \int_{\theta_i} \left(CS(\theta) - \int_s \int_{\underline{\theta}_i}^{\theta} q(\tilde{\theta}, s) d\tilde{\theta} dF(s) \right) dG_i(\theta) + R^m \geq 0$$

As a consumer surplus is given by $CS(\theta) = \int_s (U(q(\theta, s), \theta, s) - t_i(\theta, s) q_i(\theta, s)) dF(s)$, the consumer cost from the consumer surplus and the payment from the revenue cancel each other. This leaves:

$$\sum_i \mu_i \int_{\theta_i} \int_s \left(U(q(\theta, s), \theta, s) - \int_{\underline{\theta}_i}^{\theta} q(\tilde{\theta}, s) d\tilde{\theta} dF(s) \right) dF(s) dG_i(\theta) - I(k) \geq 0$$

Finally, integration by parts leads to the final expression of the constraint $R - IR$. Then, note the revenue equivalence stems from the observation that the following expression can give the payment schedule:

$$t(\hat{\theta}, s)q(\hat{\theta}, s) = U(q(\hat{\theta}, s), \hat{\theta}, s) - \int_{\theta_i}^{\hat{\theta}} q(\tilde{\theta}, s)d\tilde{\theta} - R - IR$$

Plugged in the consumer surplus and simplified:

$$CS(\theta, \hat{\theta}) = \int_s (U(q(\hat{\theta}, s), \theta, s) - U(q(\hat{\theta}, s), \hat{\theta}, s) + \int_{\theta_i}^{\hat{\theta}} q(\tilde{\theta}, s)d\tilde{\theta} + R - IR)dF(s) \quad (\text{D.13})$$

The derivative with respect to $\hat{\theta}$ leads to:

$$\frac{\partial CS}{\partial \hat{\theta}} = \int_s \left(\frac{\partial U(q(\hat{\theta}, s), \theta, s)}{\partial q} \frac{\partial q}{\partial \hat{\theta}} - \frac{\partial U(q(\hat{\theta}, s), \hat{\theta}, s)}{\partial q} \frac{\partial q}{\partial \hat{\theta}} - \frac{\partial U}{\partial \theta} + q(\hat{\theta}, s) \right) dF(s) \quad (\text{D.14})$$

Recall that $\frac{\partial U}{\partial \theta} = q(\hat{\theta}, s)$. Hence, $\frac{\partial CS}{\partial \hat{\theta}} = 0$ if and only if $\theta = \hat{\theta}$, which ensures the incentive compatibility constraint to be satisfied. The cross derivative with respect to θ is also clearly positive if $\frac{\partial q}{\partial \theta} > 0$. Finally, note that by construction, the expected revenue is also non-negative. Moreover, from Equation D.13, we have $CS(\theta, \theta) = \int_s \int_{\theta_i}^{\hat{\theta}} q(\tilde{\theta}, s)d\tilde{\theta} + R - IR)dF(s)$. Hence, when $\theta = \bar{\theta}_i$, $CS(\bar{\theta}_i, \bar{\theta}_i) = R - IR = 0$, which implies that both revenue and participation constraints are satisfied.

The Lagrangian of the problem can be expressed as follows:

$$\begin{aligned} \mathcal{L}^m &= \sum_i \mu_i \int_s \int_{\theta_i} (U(q_i(\theta, s), \theta, s) - q_i(\theta, s)t_i(\theta, s))dG_i(\theta)dF(s) \\ &+ \varepsilon \left(k - \sum_i \mu_i \int_{\theta_i} q_i(\theta, s)dG_i(\theta) \right) + \zeta R - IR \end{aligned}$$

With ε being the Lagrange multiplier associated with the capacity constraint and ζ the multiplier associated with the new participation/revenue constraint. Now, note that as the market designer seeks to maximize consumer surplus, having a positive net revenue is not optimal. In other terms, at the maximum we have $\sum_i \mu_i \int_s \int_{\theta_i} q_i(\theta, s)t_i(\theta, s)dG_i(\theta) =$

$I(k)$. From the expression of $R - IR$ in Equation D.16, the derivative with respect to q_i gives:

$$\frac{\partial \mathcal{L}^m}{\partial q_i} = \int_s u(q_i(\theta, s), \theta, s) dF(s) - \varepsilon + \zeta \int_s (u(q_i(\theta, s), \theta, s) - \Gamma_i(\theta)) dF(s)$$

Which gives the first-order condition:

$$u(q_i(\theta, s), \theta, s) - \varepsilon + \zeta(u(q_i(\theta, s), \theta, s) - \Gamma_i(\theta)) = 0 \quad (\text{D.15})$$

Note that due to the consumer unit mass assumption, having different categories of consumers does not change the solution to the problem. Namely, the maximization of the expected aggregate consumer surplus can be understood as the individual sum of the maximization of the consumer surplus at the category level.

Finally, we show that the incentive compatibility condition is satisfied with the first-order condition. We derive the condition D.15 with respect to the type θ : $\frac{\partial u}{\partial q} \frac{\partial q_i}{\partial \theta} + \frac{\partial u}{\partial \theta} + \varepsilon + \zeta(\frac{\partial u}{\partial q} \frac{\partial q_i}{\partial \theta} + \frac{\partial u}{\partial \theta} - \frac{\partial \Gamma_i}{\partial \theta}) = 0$. With the linear framework and uniform distribution we have: $\frac{\partial u}{\partial q} = -1$, $\frac{\partial u}{\partial \theta} = 1$ and $\frac{\partial \Gamma_i}{\partial \theta} = -1$. It implies that $-\frac{\partial q_i}{\partial \theta} + 1 + \varepsilon + \zeta(-\frac{\partial q_i}{\partial \theta} + 2) = 0$. That is:

$$\frac{\partial q_i}{\partial \theta} = \frac{1 + 2\zeta}{1 + \zeta} > 0$$

□

8 Proof of Proposition 3.5

Proof. The merged participation and revenue constraint, when it is not binding, is given in the following equation:

$$R^{IR}(k) = \sum_i \mu_i \int_{\theta_i} \left[\int_0^{s_1} (U(q_{i,1}^m(\theta, s), \theta, s) - \Gamma_i(\theta) q_{i,1}^m(\theta, s)) dF(s) + \int_{s_1}^{\bar{s}} (U(q_{i,2}^m(\theta, s), \theta, s) - \Gamma_i(\theta) q_{i,2}^m(\theta, s)) dF(s) \right] dG_i(\theta) - I(k) \quad (\text{D.16})$$

With $q_{i,1}^m$ and $q_{i,2}^m$ being respectively the allocation when the capacity is not binding and the capacity is binding. As it is the same allocation as the first-best, the first state of the world such that the capacity binds is the same as under the first-best: s_1 . Similar to other proves, we start by stating that depending on the value of k , the expression of R^{IR} can take three forms: (1) when the capacity never binds $s_1 \geq \bar{s}$, (2) when it always binds $s_1 \leq 0$, (3) when it both binds and not binds as given by equation D.16. We have:

$$s_1 = k - \theta^{av}$$

In this case, situation (1) is realized whenever the investment level k is above the threshold $k^+ = \theta^{av} + \bar{s}$. Situation (2) is realized whenever the investment level k is below the threshold $k^- = \theta^{av}$. Finally, note that the equation D.16 is continuous at both threshold.

We start with the lowercase: $k \geq k^-$ such that $s_1 = 0$. Equation D.16 becomes:

$$R^{IR}(k) = \sum_i \mu_i \int_{\theta_i} \int_s (U(q_{i,2}^m(\theta, s), \theta, s) - \Gamma_i(\theta)q_{i,2}^m(\theta, s))dF(s)dG_i(\theta) - I(k)$$

With the linear and uniform assumption, we have $q_{i,2}^m(\theta, s) = k + s + \theta - \sum_i \mu_i (s + \theta_i^{av})$ and $\Gamma_i(\theta) = \bar{\theta}_i - \theta$. In that case:

$$R^{IR}(0) = \frac{1}{8} \left(\left(\sum_i \mu_i (\bar{\theta}_i - \theta_i) \right)^2 + 4\mu_i \mu_j (\bar{\theta}_i - \bar{\theta}_j)^2 \right)$$

Which is always positive. Next, we compute the first derivative with respect to k , which gives:

$$\begin{aligned} \frac{\partial R^{IR}}{\partial k} &= \sum_i \mu_i \int_{\theta_i} \int_s (u(q_{i,2}^m(\theta, s), \theta, s) \frac{\partial q_{i,2}^m}{\partial k} - \Gamma_i(\theta) \frac{\partial q_{i,2}^m}{\partial k})dF(s)dG_i(\theta) - r \\ &\quad \sum_i \mu_i \int_{\theta_i} \int_s (u(q_{i,2}^m(\theta, s), \theta, s) - \Gamma_i(\theta)(\theta, s))dF(s)dG_i(\theta) - r \end{aligned}$$

As $\frac{\partial q_{i,2}^m}{\partial k} = 1$. And the second order derivative with respect to k

$$\frac{\partial^2 R^{IR}}{\partial k^2} = \sum_i \mu_i \int_{\theta_i} \int_s \frac{\partial u}{\partial q} \frac{\partial q_{i,2}^m}{\partial k} dF(s) dG_i(\theta) = -1$$

As $\frac{\partial u}{\partial q} = -1$. This shows that the expression R^{IR} is concave in k . This, and the fact that the R^{IR} is positive in $k = 0$, implies that there can only be one solution to $R^{IR}(k) = 0$ from case (1).

We now study the middle case (3). From the first-order condition, we have $q_{i,1}^m(\theta, s) = \theta + s$ which does not depend on k . As the function in the interval is continuous in s , the marginal effect at s_1 cancels out. Hence, the first derivative with respect to k gives:

$$\frac{\partial R^{IR}}{\partial k} = \sum_i \mu_i \int_{\theta_i} \int_{s_1}^{\bar{s}} (u(q_{i,2}^m(\theta, s), \theta, s) - \Gamma_i(\theta)) dF(s) dG_i(\theta) - r$$

Deriving again by k :

$$\frac{\partial^2 R^{IR}}{\partial k^2} = \sum_i \mu_i \int_{\theta_i} \left[-\frac{\partial s_1}{\partial k} (u(q_{i,2}^m(\theta, s_1), \theta, s_1) - \Gamma_i(\theta)) f(s) - \int_{s_1}^{\bar{s}} 1 dF(s) \right] dG_i(\theta)$$

Using the value of $u(q_{i,2}^m(\theta, s_1), \theta, s_1)$ and $\Gamma_i(\theta)$, we have $\int_{\theta_i} \sum_i \mu_i (u(q_{i,2}^m(\theta, s_1), \theta, s_1) - \Gamma_i(\theta)) = -\sum_i \mu_i \frac{\bar{\theta}_i - \theta_i}{2} < 0$. As $\frac{\partial s_1}{\partial k} = 1$, the first term is positive. Hence, the sign of the second derivative is ambiguous. This is clear as at k^+ , we have $s_1(k^+) = \bar{s}$. Hence, the second term in the second derivative is null, and the expression $R^{IR}(k^+)$ is convex in k . However, the first derivative with respect to k is also negative at k^+ , as it is equal to $-r$. Finally, the third derivative is positive and equal to a constant, ensuring that there are no multiple combinations of concave/convexity from R^{IR} . In other words, in the case (3), R^{IR} is first concave and then convex. Continuity with case (1) and the negative derivative at the threshold with case (2) ensures that there is at most one solution to $R^{IR}(k) = 0$.

The existence of a solution comes from the observation that $R^{IR}(0) > 0$ and that the derivative of $R^{IR}(k)$ with respect to k whenever $k \geq k^r$ is negative and equal to $-r$.

□

9 Proof of Proposition 3.6

Proof. We start with the first derivative with respect to k . The first-order condition of the Lagrangian is equal to:

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

We rewrite the equation such that:

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

With $J_i(q_{i,l}^m, \theta, s) = u(q_{i,l}^m, \theta, s) - \Gamma_i(\theta)$. It corresponds to the marginal virtual utility of a consumer. We start with the optimal off-peak allocation: $q_{i,3}$. In that case, given a realization of s , the capacity does not bind, and $\varepsilon = 0$. Hence

$$u(q_{i,3}^m, \theta, s) + \zeta J_i(q_{i,3}^m, \theta, s) = 0$$

This implies that at the optimum, we have: (i) $u(q_{i,3}^m, \theta, s) \geq 0$ and (ii) $J_i(q_{i,3}^m, \theta, s) \leq 0$. The derivative with respect to k of the first-order condition gives:

$$\frac{\partial u}{\partial q} \frac{\partial q_{i,3}^m}{\partial k} + \frac{\partial \zeta}{\partial k} J_i(q_{i,3}^m, \theta, s) + \zeta \frac{\partial J}{\partial q} \frac{\partial q_{i,3}^m}{\partial k} = 0$$

With $\frac{\partial u}{\partial q} = -1$, which implies that $\frac{\partial J}{\partial q} = -1$. Hence, this gives:

$$\frac{\partial q_{i,3}^m}{\partial k} = \frac{\partial \zeta}{\partial k} \frac{J_i(q_{i,3}^m, \theta, s)}{1 + \zeta}$$

If $\frac{\partial \zeta}{\partial k} > 0$, and because we have $J_i(q_{i,3}^m, \theta, s) < 0$, then $\frac{\partial q_{i,3}^m}{\partial k} < 0$. For the on-peak quantity, the derivative of the first-order condition gives:

$$\frac{\partial u}{\partial q} \frac{\partial q_{i,4}^m}{\partial k} + \frac{\partial \zeta}{\partial k} J_i(q_{i,4}^m, \theta, s) + \zeta \frac{\partial J}{\partial q} \frac{\partial q_{i,4}^m}{\partial k} + \frac{\partial \varepsilon}{\partial k} = 0$$

Hence:

$$\frac{\partial q_{i,4}^*}{\partial k} = \left[\frac{\partial \zeta}{\partial k} J_i(q_{i,4}^m, \theta, s) - \frac{\partial \varepsilon}{\partial k} \right] \frac{1}{1 + \zeta}$$

We turn now to the derivative of the multipliers. We start with the capacity constraint:

$$\sum_i \mu_i \int_{\theta_i} q_{i,4}^m(\zeta, \varepsilon) dG_i(\theta) = k$$

The derivative with respect to k gives:

$$\sum_i \mu_i \int_{\theta_i} \left(\frac{\partial q_{i,4}^m}{\partial \zeta} \frac{\partial \zeta}{\partial k} + \frac{\partial q_{i,4}^*}{\partial \varepsilon} \frac{\partial \varepsilon}{\partial k} \right) dG_i(\theta) = 1 \quad (\text{D.17})$$

Using the first-order conditions gives:

$$\begin{aligned} \frac{\partial q_{i,4}^*}{\partial \zeta} &= \frac{J_i(q_{i,4}^m, \theta, s)}{1 + \zeta} \\ \frac{\partial q_{i,4}^*}{\partial \varepsilon} &= -\frac{1}{1 + \zeta} \end{aligned}$$

Plugged in Equation D.17 gives:

$$\sum_i \mu_i \int_{\theta_i} \frac{1}{1 + \zeta} \left(J_i(q_{i,4}^m, \theta, s) \frac{\partial \zeta}{\partial k} - \frac{\partial \varepsilon}{\partial k} \right) dG_i(\theta) = 1$$

Therefore :

$$\frac{\partial \varepsilon}{\partial k} = \mathbb{E}_\theta J_i(q_{i,4}^m, \theta, s) \frac{\partial \zeta}{\partial k} - (1 + \zeta)$$

With $\mathbb{E}_\theta J_i(q_{i,4}^m, \theta, s) = \sum_i \mu_i \int_{\theta_i} J_i(q_{i,4}^m, \theta, s)$ the aggregate virtual marginal utility across all consumers. We turn toward the R^{IR} constraint:

$$\begin{aligned} & \sum_i \mu_i \int_{\theta_i} \left[\int_0^{s_1^m(k)} (U(q_{i,3}^m(\zeta), \theta, s) - \Gamma_i(\theta)q_{i,3}^m(\zeta)) dF(s) \right. \\ & \left. + \int_{s_1^m(k)}^{\bar{s}} (U(q_{i,4}^m(\zeta, \varepsilon), \theta, s) - \Gamma_i(\theta)q_{i,4}^m(\zeta, \varepsilon)) dF(s) \right] dG_i(\theta) - I(k) = 0 \end{aligned}$$

The derivative is equal to

$$\begin{aligned} & \sum_i \mu_i \int_{\theta_i} \left[\int_0^{s_1^m(k)} \left(J(q_{i,3}^m(\zeta), \theta, s) \frac{\partial q_{i,3}^m}{\partial \zeta} \frac{\partial \zeta}{\partial k} \right) dF(s) \right. \\ & \left. + \int_{s_1^m(k)}^{\bar{s}} \left(J(q_{i,4}^m(\zeta), \theta, s) \left(\frac{\partial q_{i,4}^m}{\partial \zeta} \frac{\partial \zeta}{\partial k} + \frac{\partial q_{i,4}^m}{\partial \varepsilon} \frac{\partial \varepsilon}{\partial k} \right) \right) dF(s) \right] dG_i(\theta) - r = 0 \end{aligned} \quad (D.18)$$

From previous calculations, this gives after rewriting :

$$\begin{aligned} & \sum_i \mu_i \int_{\theta_i} \left[\int_0^{s_1^m(k)} \left(\frac{(J_i(q_{i,3}^m, \theta, s))^2}{1 + \zeta} \frac{\partial \zeta}{\partial k} \right) dF(s) + \int_{s_1^m(k)}^{\bar{s}} \left(\frac{(J(q_{i,4}^m(\zeta), \theta, s))^2}{1 + \zeta} \frac{\partial \zeta}{\partial k} \right) dF(s) \right. \\ & \left. - \int_{s_1^m(k)}^{\bar{s}} \left(\frac{J(q_{i,4}^m(\zeta), \theta, s) \mathbb{E}_\theta J_i(q_{i,4}^m, \theta, s)}{1 + \zeta} \frac{\partial \zeta}{\partial k} \right) dF(s) + \int_{s_1^m(k)}^{\bar{s}} J(q_{i,4}^m(\zeta), \theta, s) dF(s) \right] dG_i(\theta) - r = 0 \end{aligned}$$

Define

$$\mathbb{K} = r - \sum_i \mu_i \int_{\theta_i} \int_{s_1^m(k)}^{\bar{s}} J(q_{i,4}^m(\zeta), \theta, s) dF(s) dG_i(\theta)$$

and

$$\mathbb{M} = \sum_i \mu_i \int_{\theta_i} \left[\int_s (J_i(q_{i,l}^m, \theta, s))^2 dF(s) dG_i(\theta) - \int_{s_1^m(k)}^{\bar{s}} (J(q_{i,4}^m, \theta, s) \mathbb{E}_\theta J_i(q_{i,4}^m, \theta, s)) dF(s) \right]$$

Then, define:

$$\mathbb{A} = \frac{\mathbb{K}}{\mathbb{M}}$$

Therefore, from [D.18](#), we isolate the derivative of the multiplier which gives:

$$\frac{\partial \zeta}{\partial k} = (1 + \zeta) \mathbb{A}$$

We show now that $\mathbb{A} > 0$.

First note that : $\sum_i \mu_i \int_{\theta_i} J(q_{i,4}^m, \theta, s) dG_i(\theta) = s - k + \sum_i \mu_i \underline{\theta}_i$, this can be seen from the fact that by definition $\sum_i \mu_i q_{i,4}^m = k$ and $\Gamma_i(\theta) = \hat{\theta}_i - \theta$ under uniform distribution. Then, we show that $\mathbb{K} > 0$ for every k in the case that the capacity always binds (i.e., $s_1^m(k) = 0$). Using the expression of the aggregate virtual marginal utility, we have $\mathbb{K} = k + r - \frac{\bar{s}}{2} - \sum_i \mu_i \underline{\theta}_i$. Next, we use [Proposition 3.5](#) to show that it is always positive. Indeed, the maximum of the non-binding constraint R^{IR} described in [equation D.16](#) and noted k^{IR} is equal to $k^{IR} r - \frac{\bar{s}}{2} - \sum_i \mu_i \underline{\theta}_i$. Hence we can rewrite the $\mathbb{K} = k - k^{IR}$. As $R^{IR}(k)$ is concave in k and $R^{IR}(0) > 0$, therefore the solution of $R^{IR}(k) = 0$ is always greater than the maximum k^{IR} . Because we study the case when $R - IR$ is binding, the level of investment of interest is such that k is above the solution of $R^{IR}(k) = 0$. Hence, the k of interest is also greater than k^{IR} . Therefore, $\mathbb{K} \geq 0$. First note that : $\sum_i \mu_i \int_{\theta_i} J(q_{i,4}^m, \theta, s) dG_i(\theta) = s - k + \sum_i \mu_i \underline{\theta}_i$, this can be seen from the fact that by definition $\sum_i \mu_i q_{i,4}^m = k$ and $\Gamma_i(\theta) = \hat{\theta}_i - \theta$ under uniform distribution.

Then, we show that $\mathbb{K} > 0$ for every k in the case that the capacity always binds (i.e., $s_1^m(k) = 0$). Using the expression of the aggregate virtual marginal utility, we have $\mathbb{K} = k + r - \frac{\bar{s}}{2} - \sum_i \mu_i \theta_i$. Next, we use Proposition 3.5 to show that it is always positive. Indeed, the maximum of the non-binding constraint R^{IR} described in equation D.16 and noted k^{IR} is equal to $k^{IR}r - \frac{\bar{s}}{2} - \sum_i \mu_i \theta_i$. Hence we can rewrite the $\mathbb{K} = k - k^{IR}$. As $R^{IR}(k)$ is concave in k and $R^{IR}(0) > 0$, therefore the solution of $R^{IR}(k) = 0$ is always greater than the maximum k^{IR} . Because we study the case when $R - IR$ is binding, the level of investment of interest is such that k is above the solution of $R^{IR}(k) = 0$. Hence, the k of interest is also greater than k^{IR} . Therefore, $\mathbb{K} \geq 0$. First note that : $\sum_i \mu_i \int_{\theta_i} J(q_{i,4}^m, \theta, s) dG_i(\theta) = s - k + \sum_i \mu_i \theta_i$, this can be seen from the fact that by definition $\sum_i \mu_i q_{i,4}^m = k$ and $\Gamma_i(\theta) = \hat{\theta}_i - \theta$ under uniform distribution. Then, we show that $\mathbb{K} > 0$ for every k in the case that the capacity always binds (i.e., $s_1^m(k) = 0$). Using the expression of the aggregate virtual marginal utility, we have $\mathbb{K} = k + r - \frac{\bar{s}}{2} - \sum_i \mu_i \theta_i$. Next, we use Proposition 3.5 to show that it is always positive. Indeed, the maximum of the non-binding constraint R^{IR} described in equation D.16 and noted k^{IR} is equal to $k^{IR}r - \frac{\bar{s}}{2} - \sum_i \mu_i \theta_i$. Hence we can rewrite the $\mathbb{K} = k - k^{IR}$. As $R^{IR}(k)$ is concave in k and $R^{IR}(0) > 0$, therefore the solution of $R^{IR}(k) = 0$ is always greater than the maximum k^{IR} . Because we study the case when $R - IR$ is binding, the level of investment of interest is such that k is above the solution of $R^{IR}(k) = 0$. Hence, the k of interest is also greater than k^{IR} . Therefore, $\mathbb{K} \geq 0$.

It remains to show that the level of investment that solves $R^{IR}(k) = 0$ for the middle case (3) is always higher than the one of the case when the capacity always binds. If this is the case, it would imply that \mathbb{K} under the middle case is greater than the \mathbb{K} of the lower case. As we proved that the latter is always positive, then the former will also be positive. We prove it by studying the difference between the two expressions of R^{IR} for a given k , that is, when the capacity always binds as in case (1) and sometimes binds as in case (3). Note that the difference is null at the level of investment k^- that satisfies $s_1 = 0$ due to the continuity between the two expressions. Therefore, we are only interested in the sign of the difference at any level of investment greater than k^- . The difference is given by

$$\sum_i \mu_i \int_{\theta_i} \int_0^{s_1} (U(q_{i,1}^m, \theta, s) - U(q_{i,2}^m, \theta, s) - \Gamma_i(\theta)(q_{i,1}^m - q_{i,2}^m)) dF(s) dG_i(\theta)$$

As only the on-peak allocation $q_{i,2}^m$ depends on k , the first derivative with respect to k gives:

$$\sum_i \mu_i \int_{\theta_i} \int_0^{s_1} (-u(q_{i,2}^m, \theta, s) + \Gamma_i(\theta)) dF(s) dG_i(\theta)$$

Note that at k^- , the derivative is null. Finally, the second derivative with respect to k gives:

$$\sum_i \mu_i \int_{\theta_i} \left[-\frac{\partial s_1}{\partial k} (u(q_{i,2}^m(\theta, s_1), \theta, s_1) - \Gamma_i(\theta)) f(s) + \int_0^{s_1} 1 dF(s) \right] dG_i(\theta)$$

We previously showed that the first term is positive. The second term is clearly also positive. Hence, the difference is convex in k . This and the previous observation implies that the solution of $R^{IR}(k) = 0$ is higher for the middle case, hence \mathbb{K} is also positive in that case. The last case when the capacity never binds is straightforward as $\mathbb{K} = r$. Finally, \mathbb{M} is also positive. We rewrite the expression:

$$\begin{aligned} \mathbb{M} &= \sum_i \mu_i \int_{\theta_i} \int_s (J_i(q_{i,2}^m, \theta, s))^2 dF(s) dG_i(\theta) \\ &+ \sum_i \mu_i \int_{\theta_i} \int_{s_1^m(k)}^{\bar{s}} ((J_i(q_{i,4}^m, \theta, s))^2 - J(q_{i,4}^m, \theta, s) \mathbb{E}_\theta J_i(q_{i,4}^m, \theta, s)) dF(s) dG_i(\theta) \end{aligned}$$

The first term is positive. Dropping the reference to the state of the world s , the second term can be rewritten as :

$$\sum_i \mu_i \int_{\theta_i} (J_i(q_{i,4}^m, \theta, s))^2 dG_i(\theta) - \left[\sum_i \mu_i \int_{\theta_i} J(q_{i,4}^m, \theta, s) \theta_i dG_i(\theta) \right]^2$$

The closed-form solution gives : $\mu_i(4\mu_j + \mu_i)\sigma_i^2 + \mu_j(4\mu_i + \mu_j)\sigma_j^2 - 6\mu_i\mu_j\sigma_i\sigma_j$. Recall that $\sigma_i = \frac{\bar{\theta}_i - \theta_i}{12}$. With $\mu_i \in [0, 1]$ and $\mu_j = 1 - \mu_i$, the term is positive. Hence \mathbb{M} is always positive.

□

10 Proof of Lemma 3.3

Proof. We recall the derivative of the optimal allocation with respect to k :

$$\frac{\partial q_{i,4}^m}{\partial k} = [J_i(q_{i,4}^m, \theta_i^m, s) - \mathbb{E}J_4] \mathbb{A}(k) + 1 \quad (\text{D.19})$$

Hence, if we derive with respect to the type, this boils down to the following:

$$\frac{\partial^2 q_{i,4}^m}{\partial k \partial \theta} = \frac{\partial J_i}{\partial q} \frac{\partial q_{i,4}^m}{\partial \theta} + \frac{\partial J_i}{\partial \theta} = -1 \frac{2\zeta + 1}{1 + \zeta} + 2 = \frac{1}{1 + \zeta}$$

Therefore, the cross derivative increases in the type. If, for the lowest type, the derivative is negative, the threshold exists, and it is unique.

The condition for having a higher threshold is given by the expression of θ_i^m :

$$\theta_i^m = \frac{1}{\mathbb{A}}(1 + \zeta) - \bar{\theta}_1 + \sum_i (\bar{\theta}_i - \underline{\theta}_i)$$

Then from the definition of $\Theta_i^m(k)$:

$$\Theta_i^m(k) = \mu_i \frac{\theta_i^m(k) - \underline{\theta}_i}{\bar{\theta}_i - \underline{\theta}_i}$$

The comparison with category j gives sufficient conditions. Then, we can observe that the derivative of the threshold is independent of the type:

$$\frac{\partial \theta_i^m}{\partial k} = \frac{\frac{\partial \zeta}{\partial k} \mathbb{A} - (1 + \zeta) \frac{\partial \mathbb{A}}{\partial k}}{\mathbb{A}^2} = \frac{(1 + \zeta)(\mathbb{A}^2 - \frac{\partial \mathbb{A}}{\partial k})}{\mathbb{A}^2}$$

Therefore, the threshold for every category behaves the same, and only the level changes. Note that both terms in the numerator are positive. Hence, the sign of the

threshold is ambiguous. Numerical simulation shows that the threshold can be either increasing or decreasing with respect to k .

□

Bibliography

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

RÉSUMÉ

L'électricité est au cœur du fonctionnement de notre économie moderne, par conséquent un défi majeur pour les économistes est de s'assurer d'avoir suffisamment d'investissement. Actuellement dans la plupart des pays développés, ce sont majoritairement des acteurs privés qui prennent les décisions à la fois d'investir et de consommer. C'est dans ce contexte que cette thèse s'intéresse à comprendre comment déterminer des architectures de marché permettant de donner des incitations vertueuses aux acteurs. L'objectif est alors qu'ils prennent des décisions d'investissement et de consommation efficaces. La méthodologie employée dans cette thèse repose sur une représentation théorique des comportements d'acteur dans l'ensemble des marchés électriques et s'attèle à étudier comment ces comportements interagissent avec les règles édictées sur ces mêmes marchés. Le premier chapitre examine le comportement des producteurs sur les marchés de capacité. Un marché de capacité permet aux producteurs de générer des revenus à l'avance en échange de leur engagement à être disponible, ce qui doit les inciter à suffisamment investir. Le premier chapitre propose une nouvelle approche pour conceptualiser les offres sur ces marchés en utilisant la théorie des options réelles. Ce modèle décrit notamment l'interaction entre les caractéristiques du contrat vendu sur le marché de capacité, notamment sa durée, et les offres faites par les producteurs. Ainsi, le chapitre apporte un nouvel éclairage sur la formation des prix sur ces marchés. Le second chapitre souligne l'importance de bien choisir comment la demande sur les marchés de capacité est mise en place. En effet, il démontre qu'en fonction des caractéristiques des marchés et des acteurs, certaines façons de prendre en compte la demande dans ces marchés peuvent avoir des effets inattendus sur le surplus. Ces effets peuvent à la fois être positifs ou négatifs. Enfin, le dernier chapitre pose la question de savoir comment s'assurer d'avoir suffisamment d'investissement lorsque l'on ne connaît pas la demande des consommateurs. Il décrit ainsi l'arbitrage entre financer des investissements et maximiser la consommation d'électricité. Le chapitre souligne notamment que la mise en place de marchés permettant d'atteindre un niveau d'investissement peut poser des questions de redistribution, avec certains consommateurs se retrouvant lésés même si le bien-être global est maximisé.

MOTS CLÉS

Design de Marchés, Economie Industrielle, Décisions d'Investissement, Marchés d'électricité, Régulation.

ABSTRACT

Because electricity is at the heart of our modern economy, a major challenge for economists is to ensure sufficient investment. Currently, in most developed countries, it is predominantly private actors who make both investment and consumption decisions. It is in this context that this thesis is concerned with understanding how to design markets that provide virtuous incentives to a diverse set of actors. The aim is then to induce efficient investment and consumption decisions. The methodology employed in this thesis is based on a theoretical representation of actor behavior in electricity markets and studies how this behavior interacts with the rules enacted in these same markets. The first chapter examines the behavior of producers in capacity markets. A capacity market allows generators to generate income in advance in exchange for their commitment to availability, which should provide incentives to invest sufficiently. The first chapter proposes a new approach to conceptualizing offers on these markets, using real options theory. In particular, this model describes the interaction between the characteristics of the contract sold on the capacity market, notably its duration, and the bids made by producers. In this way, the chapter sheds new light on price formation in these markets. The second chapter highlights the importance of choosing the right way to choose the demand on capacity markets. Indeed, it shows that, depending on the characteristics of the markets and the players involved, certain ways of taking demand into account in these markets can have unexpected effects on surplus. These effects can be both positive and negative. Finally, the last chapter raises the question of how to ensure sufficient investment when consumer demand is unknown. It describes the trade-off between financing investment and maximizing welfare. In particular, the chapter points out that setting up markets to achieve a certain level of investment can raise questions of redistribution, with some consumers being harmed even if overall welfare is maximized.

KEYWORDS

Market Design, Industrial organization, Investment Decisions, Electricity markets, Regulation.