Firm Behaviour in Markets with Capacity Constraints

Thesis by

David Young

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended August 5th, 2008)

For my parents, Tom Young and Wendy Stratton

Acknowledgements

I would never have completed this thesis without the help of the many people who supported, encouraged, and pushed me along the way.

First and foremost, it is difficult to overstate my gratitude to my adviser, John Ledyard. John's constant support, whether through economic advice, career advice, or funding, made the difference in the ups and downs of my research. It is not John's style to dictate the course of his students' research; rather he encouraged my good ideas and ruthlessly demolished the bad ones. I have learned more economic intuition from John than from any half dozen classes. I would also like to thank Preston McAfee, whose keen mind and encyclopaedic knowledge of the economic literature has proven invaluable time and time again, and Matias Iaryczower, who took the trouble to learn and guide my research when nothing seemed to work.

Many other people at Caltech supported me one way or the other. Chapter 4 of this thesis relies heavily on the experimental infrastructure created with the support of the NSF EPNES program award #0323819. I am very grateful to Mani Chandy, Daniel Zimmerman, Walter Yuan, Sera Linardi, and several Caltech SURF students, who created the network software and experimental interface on which I ran these experiments. I owe thanks to my fellow graduate student and office-mate Laurent Mathevet, whose advice and companionship throughout the last five years made a huge difference, and to the members of the Caltech Salsa Club, whose friendship and love of dancing tempted me out of my office. I would also like to thank Laurel Auchampaugh, Patricia Hamad, Tiffany Kim, and Sabrina Boschetti, who helped procure resources, processed paperwork, and patiently navigated me through Caltech's bureaucracy.

I would like to remember those who made my quest for a Ph.D possible in the first place. My undergraduate adviser, Seamus Hogan, and his colleagues at the University of Canterbury were instrumental in my even considering studying for a Ph.D. Mike Wilkinson, my friend at Canterbury, has been inspirational in his constant enthusiasm for all things economic. Fulbright New Zealand provided funding, and trained me extensively to cope with American culture and customs.

Finally, I would like to thank my family. My wife, Laura Panattoni, brightened my life throughout my five years here. My brothers, scattered around the world, have always helped when I needed them. And my parents, although ten thousand miles away, have never ceased their encouragement, and their practice of sending regular food parcels. To them I dedicate this thesis.

Abstract

I study firms' behaviour in markets where firms' long-run capacity decisions, made in the presence of uncertain demand, constrains short-run competition. In Chapter 2, I analyse firms' investment and pricing incentives in a differentiated products framework with uncertain demand. Firms choose production capacities before observing demand and choose prices after demand is realised. Unlike previous models, when firms are identical, symmetric pure-strategy equilibria exist, even in the presence of very low capacity costs. Furthermore, industry capacity in these symmetric equilibria is strictly greater than the equivalent Cournot equilibrium industry capacity for low costs, and equal to the Cournot industry capacity for higher costs. Subsidies on capacity costs have a greater positive impact on equilibrium capacity than an equivalent subsidy on production costs. In Chapter 3, I use this model to analyse how the market changes when firms practice 'withholding'. This is when firms withdraw capacity from the market in the short-run, after demand is realised, in the hope of making greater profits. I show that withholding is an optimal strategy for firms in these markets, and that compared to the no-withholding case, equilibrium output is lower in low demand states and higher in high demand states. Equilibrium capacity strictly increases. I discuss why it is hard to find real world examples of withholding in action, despite the increased profits. Chapter 4 looks at the specific case of the electricity industry. Electricity markets are a good example where capacity constraints and random demand affect competitive outcomes. However, trade in electricity is subject to additional constraints caused by the transmission of electricity through a network. Network constraints are well understood to cause considerable non-convexities in firms' optimisation problems; thus theoretical models have limited use in analysing the behaviour of electricity generating firms. An alternative approach, economic experiments, has become an important tool to study these markets, but questions remain on whether subjects can really imitate large firms in the presence of such complexity. This chapter provides evidence in the affirmative,

specifically showing that experimental subjects can understand loop flows in the presence of Kirchoff's Laws, a key physical constraint, and how this affects firms' pricing decisions. The results suggest that electricity market experiments could be scaled up successfully to more realistic networks.

Contents

List of Figures

Chapter 1

Introduction

There are many natural settings in which firms must invest in costly capacity before seeing the true state of demand, and then face price competition after demand is revealed. Common examples include electricity markets, where generation is fixed in the short-run but prices vary widely as demand changes, hotels, who face seasonal demand swings but have only a fixed number of rooms, and indeed any industry which produces a non-storable good or service, and faces fluctuating demand. In such markets, the nature of competition in the short-run is heavily influenced by firms' long-run capacity choices. Excess industry capacity relative to demand may result in vigorous price competition and consequent depressed prices and profits. Too little industry capacity has the opposite effects. Given the uncertain demand, firms thus face a trade-off when deciding how much capacity to build. Over-investing causes costly capacity to lie idle when demand is low, but under-investing leaves profitable demand unserved when demand is high.

The study of these capacity-price decisions has many practical implications. Consider the example of electricity markets. In the last twenty years, many countries have deregulated their electricity industries, privatising generators and allowing them to sell electricity directly to retailers within a preset market structure. This is, in effect, a giant experiment, with different countries and states introducing different market designs. The choice of market design is of crucial importance to the economic outcomes, as the 2000-1 California electricity crisis showed. A poor design can lead to poor outcomes, such as extremely high or volatile prices, or blackouts caused when demand exceeds industry capacity. Thus there is an important role for economists to play in creating better market designs, the design of which is necessarily predicated on understanding firms' investment incentives, and how these capacity choices influence short-run competition.

Since Edgeworth's modification of Bertrand's price game, economists have studied how capacity constraints can change the nature of short-run competition. These studies have tended to remain theoretical in nature. The seminal 1983 work of Kreps and Scheinkman [12] founded a literature of what I call 'capacity-price' models. These models are based on two-stage games, games where firms choose capacity at stage one, then price at stage two after all firms observe capacity choices. The work of Kreps and Scheinkman was extended to different rationing rules (Davidson and Deneckere [5]), uncertain demand in the short-run (Reynolds and Wilson [14]), and to the differentiated products setting (Friedman [8]). However, most of this work has remained fixedly theoretical, for good reason. Under Kreps and Scheinkman's standard assumptions, pure-strategy equilibria only exist under particular conditions, such as assuming the efficient rationing rule. This limits the use of comparative statics in any possible applications.[1] Furthermore, in this literature, only Reynolds and Wilson tackle the issue of uncertain demand, which is at the heart of the economic dilemma outlined above.

In the specific case of electricity, theory is even more problematic, as electricity markets are not only subject to capacity constraints on the production of electricity, but also to capacity constraints on its transmission. This abundance of non-linear network constraints has seen economists turn away from theory in favour of economic experiments. Such work is relatively new, but promising: economic experiments avoid the intractability that bedevils theoreticians. However, current experiments are usually limited to three or four node networks; hardly a good approximation of, say, the California electricity grid.

In this dissertation, I study some of these issues. I create a new capacity-price model with uncertain demand that does have pure strategy equilibria, and use this model in an extensive application to capacity withholding. In the specific case of electricity markets, I perform a series of economic experiments designed to prove that human subjects can handle the complexity of Kirchoff's laws in electricity networks; results that imply the current small scale experiments in the literature could be scaled up to more realistic networks. I briefly outline each chapter.

In Chapter 2, I create a new capacity-price model with uncertain demand, with the valuable

[1] A notable exception is the work of Maggi [13], who modified Friedman's work to show a nice application to International Trade.

property that pure-strategy equilibria exist, even in the symmetric case. In most previous models, economists have followed Kreps and Scheinkman's lead, assuming that the price subgame will be Bertrand in nature. In practice, this assumes consumers have full information of all prices charged by all firms at all times, and are willing to switch firms instantaneously if one firm raises or lowers its price. This is the approach of Reynolds and Wilson, whose paper in the Journal of Economic Theory is the only example of a capacity-price model with uncertain demand that I know of. However, an alternative assumption is available. Friedman introduced the idea of replacing the Bertrand assumption with the differentiated products pricing assumption in the price subgame. He showed this leads to pure strategy equilibria in the game without uncertain demand, and the usefulness of his approach for applied research was proved by Maggi [13]. In this chapter, I modify Friedman's model to allow for uncertain demand, and prove the existence of pure-strategy equilibria in the resulting game.

Chapter 3 utilises the model created in Chapter 2 to study a form of firm behaviour I call 'capacity withholding'. Capacity withholding is easiest to explain with an example. Consider the market for car ferry sailings across Lake Michigan. The market is contested by two firms, each of whom owns one car ferry. In a single 24 hour period, the ferries can make four and six crossings of Lake Michigan respectively. In this example, the ferries themselves are the firm's long-run capacity, and full capacity would be four/six daily crossings of Lake Michigan. In summer, when demand is high, both firms do run at full capacity, however in the off-season, they dramatically cut the number of sailings per day, while maintaining pretty much the same prices. This is a short-run reduction in capacity, practiced by both firms, which appears to have the net effect of maintaining higher prices and presumably profits. In this chapter, I show it is advantageous for firms to practice withholding, and that when firms withhold, they have an incentive to increase equilibrium capacity. However, it is surprisingly difficult to find real world examples of withholding in action. I also show that the successful practice of withholding requires three factors: commitment to the lower capacity level, observable choices by each firm, and some modicum of divisibility.

The final chapter, Chapter 4, takes a different tack. Electricity is a classic example of a non-storable good with random demand and capacity constraints. However, as I mentioned above, electricity markets face additional constraints not allowed for by the theoretical model in Chapter 2. Indeed, there is little if any theoretical work on capacity-price competition on electricity networks.[2] Thus in recent years, economists have turned to economic experiments to study firm behaviour on electricity networks, and that is the approach I take in this chapter. Since this field is relatively new, there are still many unanswered questions, such as whether human subjects can truly imitate big electricity firms in their bidding behaviour on electricity networks. Do they really understand the constraints and maximise profits? In this chapter, I show that human subjects can understand quite complicated network externalities, specifically Kirchoff's laws, and consistently maximise profit despite the extra constraints imposed by these physical laws. My approach is to conduct repeated experiments, using experienced subjects, on four different networks: two where Kirchoff's laws influence outcomes, and thus profits, and two where Kirchoff's laws have no effect. I then search for variations between the two treatments, finding no significant differences. My results give confidence that experiments with human subjects on more complicated networks are possible.

[2] There is some theoretical work on pricing in electricity networks, one example is the paper by Wu, Variaya, Spiller, and Oren [21]. However, this assumes firms have fixed capacities.

Chapter 2

Endogenous Investment

2.1 Introduction

There are many natural settings in which firms must invest in costly capacity before seeing the true state of demand, and then face price competition after demand is revealed. An example is electricity markets, where capacity decisions are fixed in the short term but prices may vary drastically as demand varies, particularly when demand is close to industry capacity. Firms in this market structure face a trade-off. Holding more capacity allows a firm to serve more consumers in high demand states, but the extra capacity can drive down prices in both high and low demand states. Such trade-offs are often studied in applied theory using variations on the Cournot model, as in Hogendorn [11]. Yet the relationship between the Cournot model, where firms choose capacity and the price clears the market, and these 'capacity-price' models, where firms choose both capacity and price, is poorly understood. With sufficiently high capacity costs, the equilibrium capacity predictions in the existing capacity-price models match those of the Cournot model. However for lower costs, capacity-price models usually have no pure-strategy equilibria, due to the assumption of Bertrand-Edgeworth pricing, and make little or no comparison to the Cournot model. In this chapter, I construct a new model using the differentiated products paradigm in which firms choose capacity before the realisation of demand, and choose price afterwards. I show that even when capacity costs are low, pure-strategy symmetric equilibria exist. In this case, equilibrium capacity is greater than the Cournot model's prediction.

Two-stage games where firms choose capacity, then price, are an appealing way to model market structure. In these games, prices are determined by firms' actions rather than the 'invisible auctioneer' of the Cournot model. Thus Kreps and Scheinkman's seminal 1983 paper [12], which showed that a capacity-price model could yield Cournot outcomes, caused considerable interest. As Tirole [17] points out, this paper was seen as a justification of the use of the more tractable Cournot model. The Kreps-Scheinkman model had two firms, choosing capacity, and then price according to the Bertrand-Edgeworth paradigm. The model assumed the efficient rationing rule. When demand was concave, marginal production costs were zero, and capacity costs were convex, the unique equilibrium outcome matched the Cournot outcome, both for capacity and market price.

However, Kreps and Scheinkman's result is not robust, either to the choice of rationing rule (Davidson and Deneckere [5]) or to the assumption that demand is known (Reynolds and Wilson [14]). Davidson and Deneckere took the Kreps-Scheinkman model and altered the rationing rule. They found that under alternative rationing rules, the Cournot outcome only occurred in equilibrium when capacity costs were sufficiently high. Otherwise, they found only mixed strategy symmetric equilibria. In a specific example, they demonstrated the existence of an asymmetric pure-strategy equilibrium, and showed that firm capacities in this equilibrium were significantly greater than the equivalent symmetric Cournot equilibrium. Davidson and Deneckere concluded that

> "This suggests that the equilibrium [in the capacity-price model] is more competitive than the Cournot model would predict." [p412]

Reynolds and Wilson added demand uncertainty to the Kreps-Scheinkman model.[1] Again, they discovered that for sufficiently high costs, the unique equilibrium outcomes matched those of the equivalent Cournot output game (Cournot with random demand). For lower costs, they proved no pure strategy symmetric Nash equilibria existed. Similar to Davidson and Deneckere, they showed the existence of an asymmetric pure-strategy equilibrium for a specific example. They did not compare this equilibrium to the Cournot model.

All the papers mentioned above share the assumption of Bertrand-Edgeworth pricing in the price subgame. There are a few papers investigating similar structures under the differentiated products paradigm. Friedman [8] created a model, where n firms chose capacity, then price in a two-stage model, with known demand. Friedman found sufficient conditions for pure-strategy equilibria to exist. He did not characterise any of these equilibria. A later

[1] Reynolds and Wilson added additional assumptions above and beyond those made by Kreps and Scheinkman to prove existence.

paper by Maggi [13] used a linear duopoly version of the same model (with an added parameter) in an International Trade application. Amongst his other results, Maggi pointed out that the use of a capacity-price model had richer implications than the equivalent Cournot model. In his model, production costs affected both firms' reduced payoff functions directly, whereas firm i's capacity cost only directly affected firm i's payoff function. Maggi argued this implied "Output policies can do everything that capacity policies can." [Proposition 4, p248].[2] In the equivalent Cournot model, capacity costs are identical to production costs, and thus such distinctions cannot be studied.

I find the lack of papers assuming differentiated products pricing surprising, for two reasons. Firstly, the use of capacity-constrained differentiated products pricing (known as Bertrand-Edgeworth-Chamberlain equilibria) in the subgame significantly ameliorates the lack of pure-strategy equilibria problem present when using the Bertrand-Edgeworth paradigm[3]. If we are looking for tractability in a capacity-price model, surely this is an appropriate choice. Secondly, the differentiated products paradigm is arguably applicable to a wide variety of real world industries. Products need not even be differentiated physically to satisfy the underlying assumptions. It is enough either that people perceive goods from different firms as different (think imported goods versus domestically produced goods), or that some friction exists which prevents people from switching from one firm to the other with complete ease. A paper by Carlson and McAfee [2] gives a good example of the latter. They create a model where consumers face search costs to find the lowest price amongst firms selling identical goods. The resulting demand functions for each firms are in linear differentiated products form. Bertrand pricing on the other hand, is a fairly restrictive assumption. It implies that goods are homogeneous, and that consumers are fully rational, have full information on all prices offered, and respond instantaneously to changes in price. There are few industries in which one could argue all three conditions apply. The two paradigms are mutually exclusive, and while differentiated products may not be the right assumption for every industry, surely we should consider how it changes capacity-price models.

I introduce demand uncertainty into Friedman's model. I assume that at stage one, firms

[2] Maggi's model does not have random demand, so the result cannot be compared to my result that capacity policies have a stronger effect than output policies.

[3] Pure strategy in the B-E-C equilibrium are not guaranteed, though. Vives [20] gives more details.

must choose capacities before seeing the true level of demand. After installing capacity, firms observe the realisation of demand, and choose prices in the subgame. I also extend the model to explicitly account for production costs. Unlike Reynolds and Wilson's model, I show that pure strategy symmetric equilibria may indeed exist when capacity costs are very low. I prove this in two cases. The first is when capacity acts as a strategic complement. Since I require gross substitutes in the goods market, this is somewhat restrictive, though, as I will show, satisfied by a CES-type demand function under a certain range of parameters. This does not mean that there are no pure strategy equilibria when capacity is not a strategic complement, merely that the payoff functions do not satisfy concavity or decreasing differences, properties needed to use conventional existence results. Thus I also show that in the linear case, existence is guaranteed under additional restrictions on the distribution of the random variable, by showing quasiconcavity of the payoff functions. Both results also apply when costs are asymmetric.

When a symmetric equilibrium exists, and an equivalent Cournot equilibrium exists, I make direct comparisons with the Cournot benchmark. In my model, firms facing low demand will not choose market clearing prices, preferring to maximise profits in the price subgame by choosing a higher price. These higher prices feed back to the capacity game, encouraging firms to choose greater capacity levels. Unlike most previous capacity-price models, I can characterise symmetric equilibria in my model. Using this characterisation, I show that equilibrium capacity in my model is greater than or equal to equilibrium capacity in the Cournot model. When capacity costs are low, this inequality is strict, becoming an equality when capacity costs are sufficiently high. I also show that a change in capacity costs has a stronger positive effect on equilibrium capacity (in the symmetric case) than an equivalent change in production costs. This implies that policymakers wishing to raise industry capacity should focus on capacity costs over production costs.

The use of the differentiated products framework, as opposed to the Bertrand-Edgeworth framework, does carry some costs. Much as using Bertrand-Edgeworth requires assuming a rationing rule, I must place assumptions on the demand functions specifying the substitutability of the goods. Furthermore in the differentiated products framework, it is possible for the market demand parameter to affect one firm differently from the other, as opposed to the market demand function used by Reynolds and Wilson. To rule out such effects, I assume that demand uncertainty enters the demand functions multiplicatively, so that any exogenous change in demand affects each firm proportionally. This can be relaxed somewhat, but at the cost of more stringent assumptions on the demand functions.

Now I turn to the details of the model.

2.2 A Model of Capacity Choice

Consider two firms that produce symmetrically differentiated products. Demand for each firm's product is a function of both firms' prices, and of a random variable λ representing the level of market demand. Firms play a two-stage game. At stage one, each firm i chooses a capacity level K_i, at a cost of k_i per unit. At this stage, firms do not know the true value of λ, but they do know that λ will be the realisation of a random variable with a particular distribution $F(\lambda)$ which is twice continuously differentiable, and whose support is contained in $\left[\underline{\lambda}, \overline{\lambda}\right]$. At the beginning of stage two, λ is revealed, and each firm i chooses a price p_i at which it wishes to sell its product. I denote the demand for firm i's product given p_i, p_{-i}, and λ by $x_i(p_i, p_{-i}, \lambda)$. Since λ is intended to represent market demand, I assume λ enters demand multiplicatively. Each firm i pays a cost of c_i for every unit it produces, but cannot produce more than K_i even if its demand is higher. After stage two, each firm i realises demand, and receives profits $(p_i - c_i)\,\mathrm{Min}\left[K_i, x_i\left(p_i, p_{-i}, \lambda\right)\right]$.

I will maintain the following assumptions on the demand and cost functions throughout the paper. These mimic Friedman's assumptions, extended to allow explicitly for a constant marginal cost of production and the demand level variable, λ. Assumption 1 outlines standard regularity conditions on demand functions, and defines the cost functions for the two firms. Assumptions 2(a - e) impose the requirement that the differentiated goods be gross substitutes, and that a firm's demand and marginal profit functions react more strongly to a change in the firm's own price than to a change in the competitor's price. Assumption 2(f) implies that if p_{-i} increases, then firm i's optimal point on it's residual demand curve gives rise to both a higher price and quantity in the price subgame. We will see this assumption provides important structure, essential for Lemma 1, but not for uniqueness in the price game.

Assumption **1.** *The functions $x_i(p_i, p_{-i}, \lambda)$ are defined, continuous, single-valued and non-negative for all $(p_1, p_2) \in \Re^2_+$, $i = 1, 2$. Each firm pays a constant cost per unit produced equal to $c_i > 0$, and a constant cost for each unit of capacity installed equal to $k_i > 0$.*

There exists a finite, positive price, p_i^0 such that $x_i(p_i, p_{-i}, \lambda) = 0$ for all $p_i \geq p_i^0$, for all p_{-i}. Since each firm i's demand will be zero if its price exceeds p_i^0, we can limit the price domain to $\Omega = \{(p_1, p_2) \mid 0 \leq p_1 \leq p_1^0, 0 \leq p_2 \leq p_2^0\}$. Further let Ω_i be the subset of Ω on which $x_i(p_i, p_{-i}) > 0$.

Assumption **2.** *The following derivatives exist and satisfy these conditions:*

(a) $\frac{\partial x_i}{\partial p_i} < 0$ *on* Ω_i *and* 0 *otherwise*

(b) $\frac{\partial x_i}{\partial p_{-i}} > 0$ *on* $\Omega_i \cap \Omega_{-i}$ *and* 0 *otherwise*

(c) $\frac{\partial x_i}{\partial p_i} + \frac{\partial x_i}{\partial p_{-i}} < 0$ *everywhere on* Ω_i

(d) $2\frac{\partial x_i}{\partial p_i} + (p_i - c_i)\frac{\partial^2 x_i}{\partial p_i^2} < 0$ *everywhere on* Ω_i

(e) $2\frac{\partial x_i}{\partial p_i} + (p_i - c_i)\frac{\partial^2 x_i}{\partial p_i^2} + \left| \frac{\partial x_i}{\partial p_{-i}} + (p_i - c_i)\frac{\partial^2 x_i}{\partial p_i \partial p_{-i}} \right| < 0$ *everywhere on* Ω_i

(f) $\frac{-\frac{\partial x_i}{\partial p_{-i}}}{\frac{\partial x_i}{\partial x_i}} \geq \frac{-\left(\frac{\partial x_i}{\partial p_{-i}} + (p_i - c_i)\frac{\partial^2 x_i}{\partial p_i \partial p_{-i}}\right)}{2\frac{\partial x_i}{\partial p_i} + (p_i - c_i)\frac{\partial^2 x_i}{\partial p_i^2}}$ *everywhere on* Ω_i

Assumptions 1, 2(a), and 2(b) imply a natural bound on firms' capacity levels. Firm i's demand is maximised when its own price is 0, the other firm charges p_{-i}^0, and $\lambda = \overline{\lambda}$. Define $K_i^0 = \overline{\lambda} x_i(0, p_{-i}^0)$. K_i^0 is finite because p_{-i}^0 and $\overline{\lambda}$ are finite. Firm i will never choose K_i greater than K_i^0, because such choices lead to strictly higher capacity costs with no possibility of any extra revenue from the excess capacity. Thus I may restrict each firm's capacity choice to the set $[0, K_i^0]$.

2.2.1 Pricing Subgame

I begin by solving for the equilibrium in the pricing subgame, fixing both firms' capacities and λ, the demand shock. Each firm i chooses price p_i to maximise its objective function $(p_i - c_i)\,\mathrm{Min}\left[x_i(p_i, p_{-i}, \lambda), K_i\right]$. Fixing p_{-i}, recall that $x_i(p_i, p_{-i}, \lambda)$ is decreasing in p_i. Thus when p_i is high, $x_i < K_i$ and firm i's objective function becomes $(p_i - c_i)\,x_i$, which is concave by assumption 2(d). For low p_i, $x_i > K_i$ and firm i's objective function reduces to $(p_i - c_i) K_i$, which is an increasing linear function of p_i. In the latter case, firm i has an incentive to increase the price until $x_i = K_i$. In the former case, the objective function $(p_i - c_i) x_i$ will either obtain its maximum within the region $x_i \in [0, K_i]$, or it must be strictly decreasing in this region, in which case, firm i has an incentive to decrease the price until $x_i = K_i$. Thus, for fixed p_{-i}, firm i has two choices. Choose the price p_i^C that clears demand, i.e., p_i^C such that $x_i \left(p_i^C, p_{-i}, \lambda\right) = K_i$, or choose the price p_i^B equal to $\text{argmax}_{p_i} (p_i - c_i) x_i$. Firm i's best response is always to choose the higher of these two prices.

The equilibrium of the price subgame is given by the intersection of the two reaction functions, $p_i = \text{Max}\left\{p_i^C (p_{-i}, \lambda, K_i), p_i^B (p_{-i}, \lambda)\right\}$. Existence and uniqueness are guaranteed by assumptions 1 and 2(a - e), which imply increasing best replies with slope between 0 and 1 on a compact set[4].

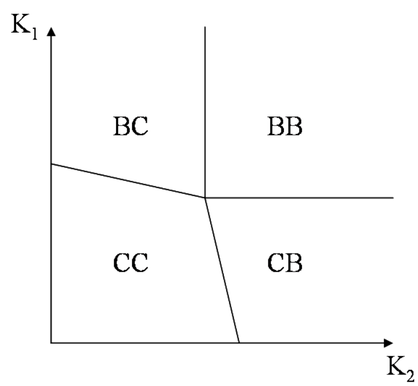


Figure 2.1: Price equilibria for fixed spot market capacity and demand.

Figure 2.1 indicates that for fixed capacities and demand, one of four types of price equilibria may occur. If both firms have relatively low capacity, then both firms will choose clearing prices. I denote this case the CC equilibrium, where the first C indicates firm 1 is choosing a clearing price, and the second C indicates firm 2 is also choosing a clearing price. If firm 1 has relatively low capacity, and firm 2 has relatively high capacity, then I denote the resulting equilibrium the CB equilibrium. This means firm 2 is now choosing the Bertrand price, but firm 1 is still charging the clearing price. The BB and BC equilibria are computed symmetrically. Throughout this paper, p_i^{CC} will denote the price charged by firm i when

[4]A proof of this when production costs are zero may be found in Friedman [8], Theorem 2. The extension, using my assumptions, is immediate.

both firms are charging clearing prices, R_i^{CC} will denote short-run revenue gathered by firm i when both firms charge clearing prices, and π_i^{CC} similarly will denote short-run profit.[5] Equivalent definitions apply for the BC, CB, and BB equilibria.

For fixed K_1 and K_2, the type of price equilibria may change as λ changes. An increase in λ increases demand at all capacity levels. This in turn increases the chance that demand will exceed capacity at Bertrand prices. The effect in figure 2.1 is to push the curves up and to the right. This change is depicted in figure 2.2, where the bold curves represent a higher λ - λ_{HIGH}.

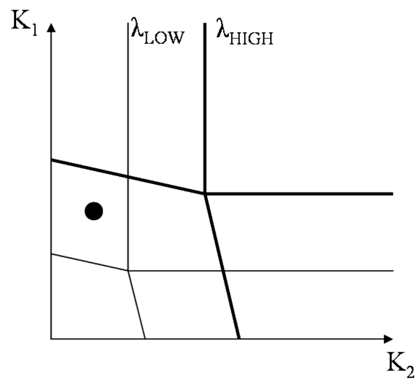


Figure 2.2: Changes in price equilibria for fixed capacities when demand changes.

Assume firms' capacities are represented by the dot in figure 2.2. When $\lambda = \lambda_{HIGH}$, this point sits in the CC region, so both firms choose clearing prices in the price subgame equilibrium. When $\lambda = \lambda_{LOW}$, this capacity point now lies in the BC region, where firm 1 chooses the Bertrand price in the price equilibrium. Between the two, there is a critical value of λ where firm 1 flips from choosing the clearing price to choosing the Bertrand price. I denote this critical value by λ_a. If λ continued to fall then eventually firm 2 would switch to the Bertrand price as well. This second point I denote λ_{aa}.

If both firms have identical production costs and $K_1 > K_2$, then an increase in λ will affect firm 2 before firm 1, causing firm 2 to switch first to the clearing price. This is the case in figure 2.2. This is both because firm 2 has less capacity, and because λ enters demand multiplicatively, implying that firm 2 will face a capacity constraint first. This switching

[5] More formally, short-run revenue here is $R_i^{CC} = p_i^{CC} x_i\left(p_i^{CC}, p_{-i}^{CC}, \lambda\right)$, and short-run profit is $\pi_i^{CC} = \left(p_i^{CC} - c_i\right) x_i\left(p_i^{CC}, p_{-i}^{CC}, \lambda\right)$. Since prices are functions of capacities and the demand shock, so too are revenue and profit.

point is λ_{aa}. As λ continues to rise, eventually firm 1 will switch too, at the point λ_a. If instead $K_2 > K_1$, firm 1 would switch first, choosing the clearing price instead of the Bertrand price in equilibrium, at a point I will call λ_{bb}, followed by firm 2, at a point I will call λ_b. There are several ways to define these values of λ; the following definition is the most convenient.

$$\lambda_a : p_1^{CC} \equiv p_1^{BC} \left(\Leftrightarrow p_2^{CC} \equiv p_2^{BC}\right)$$
$$\lambda_{aa} : p_1^{BC} \equiv p_1^{BB} \left(\Leftrightarrow p_2^{BC} \equiv p_2^{BB}\right)$$

$$\lambda_b : p_2^{CC} \equiv p_2^{CB} \left(\Leftrightarrow p_1^{CC} \equiv p_1^{CB}\right)$$
$$\lambda_{bb} : p_2^{CB} \equiv p_2^{BB} \left(\Leftrightarrow p_1^{CB} \equiv p_1^{BB}\right).$$

I will term these values of λ collectively as the *boundary values.*

When costs are asymmetric, the intuition is similar. Define g by $K_1 = g(K_2)$ when $\lambda_a = \lambda_{aa}$. Further note that when $\lambda_a = \lambda_{aa}$, $p_1^{CC} = p_1^{BB}$, and also $p_2^{CC} = p_2^{BB}$ so $\lambda_b = \lambda_{bb}$ and $\lambda_b = \lambda_a$. Therefore all the properties of the price equilibria I described for the symmetric case carry over to the asymmetric case, with $K_1 > K_2$ replaced by $K_1 > g(K_2)$ and vice versa. For linear demand functions, g is a linear function.

Since the BB price equilibrium occurs when both firms are unconstrained by capacity, equilibrium prices and demands in this region are not functions of capacity. Similarly, prices in the BC region are not dependent on K_1, and prices in the CB region are not dependent on K_2. Therefore these regions will not affect firms' respective marginal payoff as a function of their own capacities. This implies the boundary values play a crucial role in the firms' expected profit functions in the capacity game.

Assumptions 1 and 2 provide a great deal of structure on the price equilibria and the boundary values. The next lemma summarises important features of this structure; features I will need to prove the main results of this paper.

LEMMA **1.** *Under assumptions 1 and 2, the following holds.*

(i) The derivatives of λ_a, λ_b, λ_{aa}, and λ_{bb} with respect to K_1 and K_2 are all positive.

(ii) $\frac{\partial \lambda_a}{\partial K_1} > \frac{\partial \lambda_{aa}}{\partial K_1}$ *and* $\frac{\partial \lambda_{bb}}{\partial K_1} > \frac{\partial \lambda_b}{\partial K_1}$.

Proof. See the Appendix. □

Before turning to firms' capacity choices, I briefly derive an important property of a firm's marginal profit in capacities in the CC equilibrium. Recall that firm 1 will choose the higher of p_1^C and p_1^B in the price equilibrium. At λ_a, these two prices are equal. Substituting in the first-order conditions, $p_1^C \equiv p_1^B$ if and only if $K_1 \equiv -\left(p_1^{CC} - c_1\right)\frac{\partial x_1}{\partial p_1}$. Written in this way, I now show a firm's marginal return to capacity is negative at λ_a.

$$
\begin{aligned}
\left.\frac{\partial \pi_1^{CC}}{\partial K_1}\right|_{(K_1,K_2,\lambda_a)} &= p_1^{CC} + K_1\frac{\partial p_1^{CC}}{\partial K_1} - c_1 \\
&= p_1^{CC} + -p_1^{CC}\frac{\partial x_1}{\partial p_1}\frac{\partial p_1^{CC}}{\partial K_1} + c_1\frac{\partial x_1}{\partial p_1}\frac{\partial p_1^{CC}}{\partial K_1} - c_1 \\
&= \left(p_1^{CC} - c_1\right)\left(1 - \frac{\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2}}{\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}}\right) \\
&= \left(p_1^{CC} - c_1\right)\left(\frac{-\frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}}{\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}}\right)
\end{aligned} \tag{2.1}
$$

The denominator in the brackets is positive by assumption 2(c). The numerator is negative by assumption 2(b). Clearing prices are always greater than Bertrand prices (in equilibrium) which must be greater than production costs. This confirms firm 1's marginal return to capacity at λ_a is negative. Note that when cross-price effects are zero i.e. when all firms are monopolies, this equation reduces to zero. A monopolist will choose the cutoff point between Bertrand and clearing prices to extract all the surplus. As the degree of competition, represented by the cross derivatives, increases, firms choose clearing prices for more values of λ, even though this reduces their profit. A symmetric result holds for firm 2 and λ_b.

2.2.2 Cournot Benchmark

Here I lay down the equivalent Cournot type model to which I will compare my model. I will refer to this henceforth as the *Cournot Benchmark*. Formally, consider a Cournot output choice game in which outputs are chosen before the demand level realisation, and assumptions 1 and 2 above are satisfied. This is essentially my model, except firms always choose the CC equilibrium, even if it is not the best response for either firm to do so. Firm i's first order condition given K_{-i} is

$$\int_{\underline{\lambda}}^{\overline{\lambda}} \left(\frac{\partial R_i^{CC}(K_1, K_2, \lambda)}{\partial K_i} - c_i \right) dF(\lambda) - k_i. \tag{2.2}$$

Assumptions 1 and 2 are not sufficient to guarantee existence in this Cournot game. I will later assume a contraction condition that guarantees existence and uniqueness, primarily to draw comparisons between my model and the Cournot benchmark. This assumption is

Assumption **3.** $\frac{\partial^2 R_i^{CC}(K_1,K_2,\lambda)}{\partial K_i^2} + \left| \frac{\partial^2 R_i^{CC}(K_1,K_2,\lambda)}{\partial K_i \partial K_{-i}} \right| < 0$[6]

Assumption 3 is the differentiated products paradigm equivalent to Reynolds and Wilson's Cournot benchmark case [14] (p126).

2.2.3 Symmetric Capacity Choice in the Whole Game

I now find conditions under which subgame-perfect equilibria of the whole game exist. I will elaborate from the point of view of firm 1, noting that the argument for firm 2 is symmetric. Firm 1's payoff function is defined as the expectation with respect to λ over the reduced profit functions. This payoff function is differentiable with respect to K_1, but the first derivatives are only piecewise differentiable with respect to K_1, because price equilibria are only piecewise differentiable in K_1. This gives rise to a number of special cases. The most important distinction is if K_1 is greater than or less than $g(K_2)$.[7] Fixing K_2, first assume $K_1 > g(K_2)$. Then firm 1's first order condition takes the form

$$\int_{\lambda_a}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}(K_1, K_2, \lambda)}{\partial K_1} - c_1 \right) dF(\lambda) - k_1. \tag{2.3}$$

However, if K_1 is sufficiently high such that $\lambda_a(K_1, K_2) \geq \overline{\lambda}$, then the integral in equation 2.3 reduces to zero, and the first order condition is $-k_1$. Conversely, if K_1 is sufficiently low such that $\lambda_a(K_1, K_2) \leq \underline{\lambda}$, then equation 2.3 reverts to the Cournot benchmark condition, equation 2.2. Define $\overline{K_1^a}(K_2)$ by $\lambda_a(\overline{K_1^a}, K_2) \equiv \overline{\lambda}$, and $\underline{K_1^a}(K_2)$ by $\lambda_a(\underline{K_1^a}, K_2) \equiv \underline{\lambda}$, to represent these critical points.

Suppose instead $K_1 < g(K_2)$. Now firm 1's first order condition typically takes the form

[6] This is taken from Vives [20], chapter 6.

[7] I defined g in the previous section, by $K_1 = g(K_2)$ when $\lambda_a = \lambda_{aa}$. When costs are symmetric, g is the identify function.

$$\int_{\lambda_b}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_1, K_2, \lambda\right)}{\partial K_1} - c_1 \right) dF\left(\lambda\right) + \int_{\lambda_{bb}}^{\lambda_b} \left(\frac{\partial R_1^{CB}\left(K_1, \lambda\right)}{\partial K_1} - c_1 \right) dF\left(\lambda\right) - k_1. \tag{2.4}$$

Here again, if K_1 is sufficiently high such that $\lambda_{bb}\left(K_1, K_2\right) \geq \overline{\lambda}$, then the integrals in equation 2.4 reduce to zero, and the first order condition is $-k_1$. Conversely, if K_1 is sufficiently low such that $\lambda_b\left(K_1, K_2\right) \leq \underline{\lambda}$, then equation 2.4 reverts to the Cournot type first order condition described by equation 2.2[8]. As above, define $\overline{K_1^{bb}}(K_2)$ by $\lambda_{bb}(\overline{K_1^{bb}}, K_2) \equiv \overline{\lambda}$, and $\underline{K_1^b}(K_2)$ by $\lambda_b(\underline{K_1^b}, K_2) \equiv \underline{\lambda}$. In addition, there will be a number of intermediate cases, for example when $\lambda_b > \overline{\lambda}$ but $\lambda_{bb} \in (\underline{\lambda}, \overline{\lambda})$, or when $\lambda_{bb} < \underline{\lambda}$, but $\lambda_b \in (\underline{\lambda}, \overline{\lambda})$. To allow for these cases, I additionally define $\overline{K_1^b}(K_2)$ and $\underline{K_1^{bb}}(K_2)$ in the same way. The primary reason for this notation is to account for all possible cases in the following proofs of existence. The extra cases take the form

$$\int_{\lambda_b}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_1, K_2, \lambda\right)}{\partial K_1} - c_1 \right) dF\left(\lambda\right) + \int_{\underline{\lambda}}^{\lambda_b} \left(\frac{\partial R_1^{CB}\left(K_1, \lambda\right)}{\partial K_1} - c_1 \right) dF\left(\lambda\right) - k_1, \tag{2.5}$$

or

$$\int_{\lambda_{bb}}^{\overline{\lambda}} \left(\frac{\partial R_1^{CB}\left(K_1, \lambda\right)}{\partial K_1} - c_1 \right) dF\left(\lambda\right) - k_1. \tag{2.6}$$

In the symmetric case (when marginal costs are the same for both firms), g is the identity function, and if $K_1 = K_2$, equations 2.3 and 2.4 are equivalent, and simplify to the following condition.

$$Z\left(K\right) = \int_{\hat{\lambda}}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_1, K_2, \lambda\right)}{\partial K_1} - c \right) dF\left(\lambda\right) - k \tag{2.7}$$

Once again, this equation will have two special cases. When K is sufficiently high, the first order condition is $-k$. When K is sufficiently low, the first order condition is equation 2.2, evaluated at $K = K_1 = K_2$.

When K_1 is high enough, all the first order conditions reduce to $-k_1$. This allows me to restrict the region in which firms' wish to choose capacity.

[8]Note by definition that when $K_1 < g(K_2), \lambda_b > \lambda_{bb}$.

Lemma **2.** *Let A be the set defined by (K_1, K_2) such that $0 \leq K_1 \leq \overline{K_1^a}$ and $0 \leq K_2 \leq \overline{K_2^b}$. Any point outside this set is strictly dominated by a point inside the set, for at least one of the firms.*

Proof. Any point outside this set results in at least one firm choosing the Bertrand price for all possible values of λ. This firm thus has excess capacity in all states, so could receive strictly higher payoffs by reducing its own capacity until demand equals capacity in at least one state, i.e. when $\lambda = \overline{\lambda}$, since demand is increasing in λ. □

The first proposition concerns existence of equilibria for capacity choices when capacity is a strategic complement for fixed λ. Assumptions 1 and 2 are insufficient to ensure strategic complementarity,[9] so I need a third assumption.

Assumption **4.** *The derivative $\frac{\partial^2 R_i^{CC}(K_1,K_2,\lambda)}{\partial K_i \partial K_{-i}}$ exists and is positive $\forall K_1$, K_2, λ, and $i = 1, 2$.*

When K_2 increases, there are two forces at play affecting the first order conditions. The first is conventional – for every value of λ, increasing K_2 will increase firm 1's marginal payoff (assuming strategic complements). But an increase in K_2 has a secondary effect – a larger K_2 will decrease the range of λ for which firm 2 chooses clearing prices, and it will decrease the range of λ for which firm 1 chooses clearing prices. This effect will always have a positive influence on firm 1's marginal payoff. In the strategic complements case, these two effects reinforce each other. In this case, I can show the increasing differences property holds and an equilibrium exists quite generally.

Proposition **1.** *Under assumptions 1, 2, and 4, there exists a Nash equilibrium in pure strategies in this game.*

Proof. For full details see the Appendix. I show that under assumption 4, each firm's payoff function satisfies the increasing differences property. I then apply well-known results from Topkis [18] to get the result. □

Strategic complements is a strong assumption given that we have assumed goods are substitutes in the product market. However, there are demand functions that satisfy all three

[9] Linear demand functions satisfy assumptions 1 and 2, but capacity is a strategic substitute for fixed λ.

assumptions. One such demand function is $x_1 = p_1^{-\epsilon_1} p_2^{\epsilon_2} + \eta$, when $\epsilon_1 > 1$, $\epsilon_1^2 - \epsilon_2^2 - \epsilon_1 < 0$ and η is a small positive real number.[10]

If instead capacity acted as a strategic substitute, then for fixed λ, increasing K_2 would decrease firm 1's marginal payoff. However, this has no effect on the secondary effect described above; this remains positive. Thus we have opposing signs. It is straightforward to construct examples where the first outweighs the second and vice-versa, so assuming capacity acts as a strategic substitute for fixed λ does not imply the decreasing differences property. Therefore I cannot find the converse of Proposition 1.

However, equilibria do exist when capacity is not a strategic complement, even if the standard theorems do not apply. In the linear case, with restrictions on the distribution of λ, I prove that payoff functions are quasiconcave, and thus an equilibrium exists. This proof relies on the fact that the second derivatives of the first order conditions are convex in a firm's own capacity, under sufficient conditions on the distribution function.

PROPOSITION **2.** *When demand is linear, i.e.* $x_i = \lambda (a - b_1 p_i + b_2 p_{-i})$, *and under restrictions on the distribution function, there exists a Nash equilibrium in pure strategies.*

Proof. See the Appendix. The exact restrictions on the distribution function are summarised by equations 2.12 - 2.14, in the proof. □

If both firms have the same costs, then the assumptions of Proposition 2 also imply that the resulting symmetric equilibrium is unique.

COROLLARY **1.** *If costs are symmetric in Proposition 2, then there exists a unique symmetric equilibrium.*

Proof. Symmetric equilibria are characterised by equation 2.7. Under the conditions of Proposition 2, equation 2.7 is a convex function on a compact interval,[11] and is negative at the upper bound of this compact interval. These two properties ensure that equation 2.7 can only cross zero once. Therefore there exists a unique symmetric equilibrium (which could be at $K = 0$). Existence is guaranteed by Proposition 2. □

---

[10] Vives [20] discusses this topic briefly (p96 and p151).

[11] Equation 2.7 is convex only if $\frac{\hat{\lambda} F''(\hat{\lambda})}{F'(\hat{\lambda})} \geq -\frac{2b_1^2 + b_1 b_2}{b_2^2}$ $\forall K$. This condition is precisely equation 2.12 (in Proposition 2) evaluated at $K = K_1 = K_2$. Since Proposition 2 requires equation 2.12 to be true for all K_1 and K_2, it is certainly true when $K_1 = K_2$. Thus equation 2.7 is convex.

There is a special case of my model where existence can also be proven under an assumption of strategic substitutes (the converse of Assumption 4). This is when the derivatives of x_i with respect to p_{-i} are equal to zero, i.e., each firm is a monopoly. Here equation 2.1 simplifies to zero, and the derivative of p_1^{CC} with respect to K_1 equals the derivative of p_1^{CB} with respect to K_1. These results imply the secondary effect described above disappears.

2.3 Discussion

2.3.1 Comparison to the Cournot Benchmark

Several authors in the capacity-price model literature have drawn comparisons with the Cournot outcomes against their models. In the Bertrand-Edgeworth paradigm, they generally agree that when capacity costs are high, the equilibrium outcomes will be identical to those of the output game. For lower costs, the only symmetric equilibria in most of these models are in mixed strategies, so direct comparisons can not be made. The Kreps-Scheinkman model is the exception.

In my capacity-price model, the characterisation of symmetric equilibria given by equation 2.7 enables a general, intuitive comparison with the Cournot Benchmark characterisation (equation 2.2). Consider first the symmetric linear case. Let K^* be the unique symmetric equilibrium in my model, and let K^C be the unique symmetric Cournot equilibrium[12]. Then

$$\begin{aligned}
&\int_{\underline{\lambda}}^{\overline{\lambda}} \left(\frac{\lambda a \left(b_1 + b_2\right) - \left(2b_1 + b_2\right) K}{\left(b_1^2 - b_2^2\right) \lambda} - c \right) dF\left(\lambda\right) - k \\
&\leq \int_{\hat{\lambda}(K)}^{\overline{\lambda}} \left(\frac{\lambda a \left(b_1 + b_2\right) - \left(2b_1 + b_2\right) K}{\left(b_1^2 - b_2^2\right) \lambda} - c \right) dF\left(\lambda\right) - k
\end{aligned}$$

$\forall K$, because the integrand is increasing in λ and is negative at $\hat{\lambda}\left(K\right)$. Since the left hand side is decreasing in K, this inequality implies that for any K less than K^C, the equilibrium condition will be positive, thus K^* can only lie above K^C.

Notice that as k rises, K^* decreases, so $\hat{\lambda}(K^*)$ decreases. Eventually $\hat{\lambda}(K^*)$ will decrease to

[12] Uniqueness in my model follows from the corollary to proposition 2. Uniqueness in the Cournot model is verified by Vives [20]

$\underline{\lambda}$, at which point the equilibrium condition becomes equivalent to the Cournot one. Hence for sufficiently high k, K^* will equal K^C. This accords with previous authors using the Bertrand-Edgeworth paradigm.

I can also place an upper bound on the equilibrium capacity. Let $\acute{\lambda}$ denote the value of λ that sets marginal payoffs equal to zero. Then

$$\int_{\hat{\lambda}(K)}^{\overline{\lambda}} \left(\frac{\lambda a (b_1 + b_2) - (2b_1 + b_2) K}{(b_1^2 - b_2^2)\lambda} - c \right) dF(\lambda) - k$$
$$\leq \int_{\acute{\lambda}(K)}^{\overline{\lambda}} \left(\frac{\lambda a (b_1 + b_2) - (2b_1 + b_2) K}{(b_1^2 - b_2^2)\lambda} - c \right) dF(\lambda) - k,$$

$\forall K$ since by rearranging equation 2.1, $\acute{\lambda} \geq \hat{\lambda}$, the integrands are increasing in λ and are negative at $\hat{\lambda}(K)$. The right hand term here is strictly decreasing in K. Let K^M denote the unique K that sets the right hand equation equal to zero. By similar reasoning for the previous inequality, $K^M \geq K^*$. This right hand term is best thought of as the equilibrium condition if both firms were acting as monopolists (as if cross effects were zero).

Recall the equilibrium condition will generally not be decreasing, even in this linear case. These two bounds however, do restrict the range of potential equilibria, since they bound the equilibrium condition between two decreasing functions and is equal to one or the other of them at the upper and lower limits of K.

This model has another implication that does not accord with the Cournot model. Continuing with the linear case, I show that a decrease in capacity costs will increase K^* by more than an equivalent decrease in production costs. This is easily seen from the derivatives. The derivative of the equilibrium condition with respect to k is simply -1. The derivative of the equilibrium condition with respect to c is

$$-\left(1 - F\left(\hat{\lambda}\right)\right) + \frac{b_2^2 K}{b_1 (b_1^2 - b_2^2)\hat{\lambda}} \frac{\partial \hat{\lambda}}{\partial c} F'\left(\hat{\lambda}\right).$$

The left hand term is negative, but greater than -1. The right hand term is positive. Thus the sign is strictly greater than -1. This suggests that a policy maker who wishes to increase industry capacity should target subsidies on capacity costs rather than production costs. The Cournot model, by contrast, makes no distinction between the two types of costs.

All of the above results will generalise. If we know that marginal revenue in capacity under clearing prices (a) is negative, (b) satisfies assumption 3, and (c) is increasing in λ, and if an equilibrium exists in my model, then the above results hold. Conditions (a) and (b) are standard; the third condition is certainly true in the linear case, and it also holds for the CES-type function I used as an example of Proposition 1. Indeed, since λ affects demand multiplicatively, the derivatives of price with respect to λ are positive, so I would expect condition (c) to be satisfied more often than not.

2.3.2 Concluding Remarks

There are a couple of existing models whose intuition hews closely to mine. I have already mentioned Reynolds and Wilson [14]. Their model is broadly similar to mine in its intuition, but my model does have the advantage of a complete characterisation of symmetric equilibria, allowing a broader discussion of comparative statics and other properties of the equilibria. This should allow for greater use in potential applications, provided the assumptions on demand are met. Reynolds and Wilson do find an example where an asymmetric equilibrium exists (with symmetric costs and demand). It would be interesting to analyse my model to see if this is a general property of capacity-price models, or is just a function of the discontinuities of the Bertrand-Edgeworth paradigm. The models analysed by Gabszewicz and Poddar, and by Grimm and Zoettl [10], are also similar to mine, except that firms choose outputs rather than prices in the final stage subgame. While these models still assume an 'invisible auctioneer', they do provide an alternative (and perhaps more tractable) way to mimic the intuition of capacity-price models with random demand.

The Cournot model is tractable, widely used and understood. It has been used as a proxy for pricing games, even when profit-maximising firms would not choose such prices. Previous authors have found that the Cournot model does not match the predictions of equivalent capacity-price models in specific examples. I extend this comparison by providing a new capacity-price model with random demand, in which symmetric pure-strategy equilibria can be fully characterised. I use this characterisation to show that the Cournot model indeed predicts a less competitive (lower capacity) outcome than my model, and I derive some policy implications in my model that can not be shown in the Cournot model.

2.4 Appendix

Proof of Lemma 1

To simplify the following equations, let $M\pi_i = x_i + (p_i - c_i)\frac{\partial x_i}{\partial p_i}$, be the marginal (unconstrained) profit to firm i given prices p_i and p_{-i}.

Part (i)

For all these derivatives to be positive, we need three things to be true

1. $\frac{\partial p_1^{CC}}{\partial \lambda} \geq \frac{\partial p_1^{CB}}{\partial \lambda}$
2. $\frac{\partial p_1^{CC}}{\partial \lambda} \geq \frac{\partial p_1^{BC}}{\partial \lambda}$
3. $\frac{\partial p_1^{CC}}{\partial K_1} \leq \frac{\partial p_1^{CB}}{\partial K_1}$.

The first inequality requires

$$\frac{\frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial \lambda} - \frac{\partial x_1}{\partial \lambda}\frac{\partial x_2}{\partial p_2}}{\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}} \geq \frac{\frac{\partial x_1}{\partial p_2}\frac{\partial M\pi_2}{\partial \lambda} - \frac{\partial x_1}{\partial \lambda}\frac{\partial M\pi_2}{\partial p_2}}{\frac{\partial x_1}{\partial p_1}\frac{\partial M\pi_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial M\pi_2}{\partial p_1}}.$$

Note that since λ enters demand multiplicatively, and because p_1^{CB} solves $M\pi_2 = 0$, we have $\frac{\partial M\pi_2}{\partial \lambda} = 0$. Using this, cross multiplying and canceling where possible, we get that the above is true if and only if

$$\frac{\partial x_2}{\partial \lambda}\left(\frac{\partial x_1}{\partial p_1}\frac{\partial M\pi_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial M\pi_2}{\partial p_1}\right) \geq \frac{\partial x_1}{\partial \lambda}\left(-\frac{\partial x_2}{\partial p_2}\frac{\partial M\pi_2}{\partial p_1} + \frac{\partial M\pi_2}{\partial p_2}\frac{\partial x_2}{\partial p_1}\right).$$

The left hand side is positive given assumptions 2(c) and 2(e). Assumption 2(f) assures us the right hand side is negative, so this inequality is always true.

The proof of the second inequality is symmetric to that of the first inequality. The inequality requires

$$\frac{\frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial \lambda} - \frac{\partial x_1}{\partial \lambda}\frac{\partial x_2}{\partial p_2}}{\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}} \geq \frac{\frac{\partial x_2}{\partial \lambda}\frac{\partial M\pi_1}{\partial p_2} - \frac{\partial x_2}{\partial p_2}\frac{\partial M\pi_1}{\partial \lambda}}{\frac{\partial M\pi_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial M\pi_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}}.$$

Now we have $\frac{\partial M\pi_1}{\partial \lambda} = 0$ and the above inequality is true if and only if

$$\frac{\partial x_2}{\partial \lambda}\left(\frac{\partial x_1}{\partial p_2}\frac{\partial M\pi_1}{\partial p_1} - \frac{\partial x_1}{\partial p_1}\frac{\partial M\pi_2}{\partial p_2}\right) \leq \frac{\partial x_1}{\partial \lambda}\left(\frac{\partial x_2}{\partial p_2}\frac{\partial M\pi_1}{\partial p_1} - \frac{\partial M\pi_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}\right).$$

Again, the right hand side is positive by assumptions 2(c) and 2(e) and the left hand side is negative by assumption 2(f).

The third inequality is a direct consequence of assumption 2(f). The result is true if and only if

$$\Leftrightarrow \frac{\frac{\partial x_2}{\partial p_1}}{\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}} \leq \frac{\frac{\partial M\pi_2}{\partial p_2}}{\frac{\partial x_1}{\partial p_1}\frac{\partial M\pi_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial M\pi_2}{\partial p_1}}$$

$$\Leftrightarrow \frac{\frac{\partial x_2}{\partial p_2}\left(\frac{\partial x_1}{\partial p_1}\frac{\partial M\pi_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial M\pi_2}{\partial p_1}\right) - \frac{\partial M\pi_2}{\partial p_2}\left(\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}\right)}{\left(\frac{\partial x_1}{\partial p_1}\frac{\partial M\pi_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial M\pi_2}{\partial p_1}\right)\left(\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}\right)} \leq 0,$$

assuming assumptions 2(a - e). The denominator is positive under these assumption so (with some cancellation of common terms) this statement is true if and only if

$$-\frac{\partial x_2}{\partial p_2}\left(\frac{\partial M\pi_2}{\partial p_1}\right) \leq -\frac{\partial x_2}{\partial p_1}\left(\frac{\partial M\pi_2}{\partial p_2}\right),$$

which is true by assumption 2(f). This proves the third statement.

Putting the three inequalities together, we have $\frac{\partial \lambda_a}{\partial K_1} > 0$ and $\frac{\partial \lambda_b}{\partial K_1} > 0$. The derivatives with respect to K_2 follow symmetrically.

Part (ii)

It is immediate that $\frac{\partial \lambda_a}{\partial K_1} > \frac{\partial \lambda_{aa}}{\partial K_1}$ because the former term is positive by lemma 1(i) and the latter is zero (since λ_{aa} is defined by $p_1^{BC} \equiv p_1^{BB}$, neither term which contains K_1.

For the other case, $\frac{\partial \lambda_{bb}}{\partial K_1} > \frac{\partial \lambda_b}{\partial K_1}$ if and only if

$$\frac{-\frac{\partial p_1^{CB}}{\partial K_1}}{\frac{\partial p_1^{CB}}{\partial \lambda}} > \frac{\frac{\partial p_1^{CB}}{\partial K_1} - \frac{\partial p_1^{CC}}{\partial K_1}}{\frac{\partial p_1^{CC}}{\partial \lambda} - \frac{\partial p_1^{CB}}{\partial \lambda}},$$

which can be simplified to

$$\frac{\partial M\pi_2}{\partial p_2}\frac{\partial x_2}{\partial \lambda}\frac{\partial x_1}{\partial p_2} < 0.$$

This last equation is always true (assumption 2), giving the required result.

Proof of Proposition 1

I show that under assumptions 1, 2, and 4, this is a supermodular game.

Step 1. The set of actions is a compact lattice

Each firm can choose $K_i \in \left[0, K_i^0\right] \subset \Re$. This is a compact subset of $\Re$, which is trivially a compact lattice.

Step 2. Payoff functions are upper semi-continuous and supermodular in each firm's capacity

The payoff functions are differentiable and thus continuous. Since each firm chooses an action in $\Re$ when the other firm's capacity is fixed, the payoff functions are supermodular. (See Topkis [18] p46.)

Step 3. The payoff functions exhibit increasing differences in (K_1, K_2)

Consider firm 1. Suppose first that $K_1 \geq g(K_2)$ and $K_1 < \underline{K_1^a}$. Then the cross-partial derivative of firm 1's payoff function is

$$\int_{\underline{\lambda}}^{\overline{\lambda}} \left(\frac{\partial^2 R_1^{CC}(K_1, K_2, \lambda)}{\partial K_1 \partial K_2} \right) dF(\lambda).$$

This integral is positive by assumption 4. If instead $\underline{K_1^a} \leq K_1 < \overline{K_1^a}$, then the cross-partial derivative is of the form

$$\int_{\lambda_a}^{\overline{\lambda}} \left(\frac{\partial^2 R_1^{CC}(K_1, K_2, \lambda)}{\partial K_1 \partial K_2} \right) dF(\lambda) - \left(\frac{\partial R_1^{CC}(K_1, K_2, \lambda_a)}{\partial K_1} - c_1 \right) F'(\lambda_a) \frac{\partial \lambda_a}{\partial K_2}.$$

The integral on the left is again positive by assumption 4. Since $F'(\lambda) > 0$, $\frac{\partial \lambda_a}{\partial K_2} > 0$ (by Lemma 1) and marginal profit is negative at λ_a (from equation 2.1), it follows the right hand term is also positive. Finally, if $K_1 \geq \overline{K_1^a}$, then the cross-partial derivative is just zero. Thus in all possible cases, the cross-partial derivative is non-negative.

When $K_1 < g(K_2)$, we must consider a few more cases. The cases $K_1 \geq \overline{K_1^{bb}}$ and $K_1 < \underline{K_1^b}$ are parallel to above. For the case $\underline{K_1^b} \leq K_1 < \overline{K_1^b}$ [13] the cross-partial derivative is

[13] Some of these cases will be empty sets at certain values of K_2.

$$\int_{\lambda_b}^{\overline{\lambda}} \left(\frac{\partial^2 R_1^{CC}}{\partial K_1 \partial K_2} \right) dF(\lambda) - K_1 \left(\frac{\partial p_1^{CC}}{\partial K_1} - \frac{\partial p_1^{CB}}{\partial K_1} \right) F'(\lambda_b) \frac{\partial \lambda_b}{\partial K_2}.$$

Once again, the integral is positive. For the right hand term, it is easiest to rewrite this expression (referring to Lemma 1) as

$$-K_1 \frac{\partial p_1^{CC}}{\partial K_2} F'(\lambda_b) \frac{\partial \lambda_b}{\partial K_1}.$$

The term $\frac{\partial \lambda_b}{\partial K_1}$ is positive (it is symmetric to $\frac{\partial \lambda_a}{\partial K_2}$), and $\frac{\partial p_1^{CC}}{\partial K_2}$ is negative, allowing me to sign the whole term positive.

There is another possible case where $K_1 > \overline{K_1^b}$ but $K_1 < \underline{K_1^{bb}}$. In this case, the cross-partial derivative is zero. The final possibility is $K_1 > \overline{K_1^b}$ and $\underline{K_1^{bb}} \leq K_1 \leq \overline{K_1^{bb}}$. Here the cross-partial derivative is

$$\int_{\lambda_{bb}}^{\overline{\lambda}} \left(\frac{\partial^2 R_1^{CC}}{\partial K_1 \partial K_2} \right) dF(\lambda),$$

which is positive. Here too, all cases are non-negative.

Thus I have $\frac{\partial \pi_1}{\partial K_1}$ is weakly increasing in K_2, so by Topkis [18] p42, firm 1's payoff function has the increasing differences property. The argument for firm 2 is symmetric.

Steps 1, 2, and 3 prove this is a supermodular game, so existence of a pure strategy Nash equilibrium is guaranteed (see Vives [20] Thm 2.5).

Proof of Proposition 2

I show that for all K_2, firm 1's payoff function is quasiconcave in the set A.[14] The proof of this comprises three parts. First I give conditions under which equations 2.2 - 2.5 are convex in K_1. Then I show that for fixed K_2, firm 1's first order condition is globally convex on the resulting compact interval in set A. Finally, I show that firm 1's first order condition is negative on the upper bound of this interval. The proof then follows by recognising that a convex function on a compact interval of the real line that ends at a negative point can only cross zero once.

Fix $K_2 \in \left[0, K_2^0\right]$. Let $Z_1(K_1, K_2)$ denote firm 1's first order condition.

[14] A is defined in Lemma 2. Recall that any point outside A is dominated by a point inside A for at least one of the firms.

Part I

The derivatives of equations 2.2 - 2.5 respectively with respect to K_1 are (in the linear case)

$$\int_{\underline{\lambda}}^{\overline{\lambda}} \frac{-2b_1}{\left(b_1^2 - b_2^2\right)\lambda} dF(\lambda) \tag{2.8}$$

$$\int_{\lambda_a}^{\overline{\lambda}} \frac{-2b_1}{\left(b_1^2 - b_2^2\right)\lambda} dF(\lambda) + \frac{b_2^2 K_1}{b_1\left(b_1^2 - b_2^2\right)\lambda_a} \frac{\partial \lambda_a}{\partial K_1} F'(\lambda_a), \tag{2.9}$$

$$\begin{aligned} &\int_{\lambda_b}^{\overline{\lambda}} \frac{-2b_1}{\left(b_1^2 - b_2^2\right)\lambda} dF(\lambda) + \int_{\lambda_{bb}}^{\lambda_b} \frac{-4b_1}{\left(2b_1^2 - b_2^2\right)\lambda} dF(\lambda) \\ &+ \frac{b_1 b_2^2 K_1}{\left(b_1^2 - b_2^2\right)\left(2b_1^2 - b_2^2\right)\lambda_b} \frac{\partial \lambda_b}{\partial K_1} F'(\lambda_b) + \frac{b_2^2 K_1}{b_1\left(2b_1^2 - b_2^2\right)\lambda_{bb}} \frac{\partial \lambda_{bb}}{\partial K_1} F'(\lambda_{bb}), \end{aligned} \tag{2.10}$$

and

$$\begin{aligned} &\int_{\lambda_b}^{\overline{\lambda}} \frac{-2b_1}{\left(b_1^2 - b_2^2\right)\lambda} dF(\lambda) + \int_{\underline{\lambda}}^{\lambda_b} \frac{-4b_1}{\left(2b_1^2 - b_2^2\right)\lambda} dF(\lambda) \\ &\qquad + \frac{b_1 b_2^2 K_1}{\left(b_1^2 - b_2^2\right)\left(2b_1^2 - b_2^2\right)\lambda_b} \frac{\partial \lambda_b}{\partial K_1} F'(\lambda_b). \end{aligned} \tag{2.11}$$

The derivative of equation 2.8 with respect to K_1 is zero, and thus equation 2.2 is always convex in the linear case. When I take the derivative with respect to K_1 of equations 2.9 - 2.11, the integrals disappear and I get terms involving $F''(\lambda_a)$, $F''(\lambda_b)$, and $F''(\lambda_{bb})$. With some rearranging, I reduce these derivatives to three conditions, which, if true for all K_1 and K_2, guarantee the derivatives of equations 2.9 - 2.11 with respect to K_1 are positive.

$$\frac{\lambda_a F''(\lambda_a)}{F'(\lambda_a)} \geq -\frac{2b_1^2 K_1 + b_1 b_2 K_2}{b_2^2 K_1} \tag{2.12}$$

$$\frac{\lambda_b F''(\lambda_b)}{F'(\lambda_b)} \geq -\frac{3\left(2b_1^2 - b_2^2\right) K_2 + 2b_1 b_2 K_1}{b_1 b_2 K_1} \tag{2.13}$$

$$\frac{\lambda_{bb} F''(\lambda_{bb})}{F'(\lambda_{bb})} \geq -\frac{4b_1^2}{b_2^2} \tag{2.14}$$

The right hand sides of each of these expressions are always negative. Clearly, any increasing distribution will satisfy these conditions, as will the uniform distribution. Decreasing distributions are fine provided the slope is not too steep (variance is sufficiently high). For example, the exponential distribution will satisfy this for sufficiently low β.

When conditions 2.12 - 2.14 hold, equations 2.2 - 2.5 are convex. This completes Part I.

Part II

I consider three cases.

Case 1: $\hat{\lambda}\left(g(K_2), K_2\right) < \underline{\lambda}$

First suppose $K_1 < g(K_2)$. Then $\lambda_b\left(K_1, K_2\right) < \hat{\lambda} < \underline{\lambda}$. Here Z_1 is just the Cournot condition, equation 2.2. Now suppose $K_1 \geq g(K_2)$. For $g(K_2) \leq K_1 < \underline{K_1^a}$, Z_1 is still equation 2.2. If $\underline{K_1^a} \leq K_1 \leq \overline{K_1^a}$, then Z_1 equals equation 2.3. $\overline{K_1^a}$ is the boundary of the set A. Equations 2.2 and 2.3 are convex by Part I of this proof.

To show the function is globally convex between 0 and $\overline{K_1^a}$, I just need to show that the slope of equation 2.2 is less than the slope of equation 2.3 at the boundary point where equation 2.2 changes to equation 2.3. This point is $\underline{K_1^a}$. But this is trivial, because at $\underline{K_1^a}$, equation 2.9 (the derivative of equation 2.3) equals equation 2.8 (the derivative of equation 2.2) plus an extra positive term. This proves the convexity of Z_1 in this case.

Case 2: $\hat{\lambda}\left(g(K_2), K_2\right) > \overline{\lambda}$

First suppose $K_1 > g(K_2)$. Then $\lambda_a(K_1, K_2) > \hat{\lambda} > \overline{\lambda}$, so (K_1, K_2) lies outside the set A. Thus we need not consider this case. Suppose instead $K_1 \leq g(K_2)$. When $0 \leq K_1 \leq \underline{K_1^b}$, Z_1 is the Cournot condition (equation 2.2). If $\underline{K_1^b} < K_1 \leq \underline{K_1^{bb}}$, Z_1 is equation 2.5, and if $\underline{K_1^{bb}} < K_1 \leq \overline{K_1^b}$, Z_1 is equation 2.4. $\overline{K_1^b}$ is the boundary of the set A. By Part I, equations 2.2, 2.4, and 2.5 are convex.

To show the function is convex between 0 and $\overline{K_1^b}$, I need to show that at the points $\underline{K_1^b}$ and $\underline{K_1^{bb}}$, the slope on the left is less than the slope on the right. At the former point, the derivative on the right is given by equation 2.11, where $\lambda_b = \underline{\lambda}$. This is the same as the derivative of equation 2.2, plus a positive term. At the latter point, the derivative on the right is given by equation 2.10, which at this point has an extra term than equation 2.11. Thus the first order condition is convex on $[0, \overline{K_1^b}]$.

Case 3: $\underline{\lambda} \leq \hat{\lambda}\left(g(K_2), K_2\right) \leq \overline{\lambda}$

First suppose $K_1 \geq g(K_2)$. When $g(K_2) \leq K_1 \leq \overline{K_1^a}$, Z_1 equals equation 2.3. $\overline{K_1^a}$ is the boundary of the set A. Suppose instead that $K_1 < g(K_2)$. When $\underline{K_1^{bb}} \leq K_1 < g(K_2)$, Z_1 equals equation 2.4, and when $\underline{K_1^b} \leq K_1 < \underline{K_1^{bb}}$, Z_1 equals equation 2.5. For lower K_1, Z_1 is equation 2.2. All of these equations are convex by Part I.

Now consider the boundary points between the regions. I have already considered most of these in case (2) except for the boundary point at $K_1 = g(K_2)$. Here, note that in the linear case, $\frac{\partial \lambda_a}{\partial K_1} > \frac{\partial \lambda_b}{\partial K_1}$. When $K_1 = g(K_2)$, $\lambda_a = \lambda_b = \lambda_{bb} = \hat{\lambda}$. Thus, the difference given by equation 2.9 - equation 2.10 is

$$\frac{b_2^2 K_1 F'(\lambda_a)}{b_1 \left(b_1^2 - b_2^2\right) \lambda_a} \left(\frac{\partial \lambda_a}{\partial K_1} - \frac{\partial \lambda_b}{\partial K_1}\right) - \frac{b_2^2 K_1 F'(\lambda_{bb})}{b_1 \left(2b_1^2 - b_2^2\right) \lambda_{bb}} \left(\frac{\partial \lambda_b}{\partial K_1} - \frac{\partial \lambda_{bb}}{\partial K_1}\right),$$

which is positive, since $\frac{\partial \lambda_a}{\partial K_1} > \frac{\partial \lambda_b}{\partial K_1}$ and $\frac{\partial \lambda_{bb}}{\partial K_1} > \frac{\partial \lambda_b}{\partial K_1}$ (the latter is proved in Lemma 1). Therefore at $K_1 = g(K_2)$, the slope of equation 2.4 is less than the slope of equation 2.3.

Part III

Finally, I show that Z_1 is negative at any boundary point of A. There are two cases

Case 1: $\hat{\lambda}\left(g(K_2), K_2\right) \leq \overline{\lambda}$

In this case, as K_1 rises, the boundary point will be given by $\overline{K_1^a}$. At this point, Z_1 is simply $-k_1$, which is negative.

Case 2: $\hat{\lambda}\left(g(K_2), K_2\right) > \overline{\lambda}$

This case is a little harder. The first order condition here looks like equation 2.6, evaluated at the point $\overline{K_1^a}$. I need to show the integral is negative. Since the integrand is increasing in λ, it is sufficient to show the integrand is negative when $\lambda = \overline{\lambda}$, at $\overline{K_1^a}$. In the linear case, making these substitutions, I can simplify the integrand to

$$\frac{\overline{\lambda} a \left(-2b_1^2 b_2 - 2b_1 b_2^2 - b_2^3\right) - 4b_1^2 b_2 K_2}{\left(2b_1^2 - b_2^2\right)^2 \overline{\lambda}},$$

which is negative.

Thus, in all possible cases, the first order condition on A is a convex function, which is negative on the boundary of A. Thus the first order condition can only cross zero once, thus firm 1's payoff function is quasiconcave. A symmetric argument applies to firm 2. The existence of a pure strategy Nash equilibrium follows by standard theorems (for example, Vives [20] Theorem 2.1).

Chapter 3

Capacity Withholding

3.1 Introduction

In markets with demand uncertainty, capacity constraints, and non-storable goods, firms often end up holding excess capacity when demand is low. This excess capacity can have a dampening effect on prices and profits in the spot market, as firms compete on price to make a return on their costly capacity. A good example often quoted is the hotel industry. In the off-season, hotels end up with a lot of excess capacity, and compete vigorously on price. However, this is not true in all such industries. Consider car ferry sailings in contested markets such as the Lake Michigan crossing or the Cook Strait crossing in New Zealand.[1] In both cases, there is a clear peak in demand during summer, when vacationers and tourists abound, and a correspondingly lull during winter. Intriguingly, the companies on these routes often do not change their prices at all during the off-season. Instead they significantly cut their sailing schedule. For example, Lake Express on the Lake Michigan crossing operates six crossings a day in summer, dropping to four during fall, and cuts service altogether in winter. Its competitor, the S.S. Badger, offers four sailings in the summer, two in the fall, and none in winter. I call this practice *withholding capacity*, after a similar behaviour in the electricity industry. The capacity cuts practiced by both firms in the market apparently enable firms to maintain relatively high prices.

Why might a firm choose to withhold capacity? One answer to this question is the existence of costs with holding unused capacity in the short-run. The ferry case I introduced above is a good example of an industry with low marginal costs of production but high costs of

[1]The Lake Michigan crossing between Michigan and Wisconsin is contested by a duopoly, as is the Cook Strait crossing between Wellington and Picton in New Zealand. I do not consider the monopoly case in this paper, although similar behaviour occurs, for example in many of the Washington State ferry routes.

capacity in the short-run. A ferry sailing uses almost the same amount of fuel and labour regardless of how many passengers are actually on board. Withholding capacity – cancelling ferry sailings – in the short-run can therefore realise considerable savings. Another possible answer is that withholding some of a firm's capacity in the short-run could act as an observable commitment device, which, if matched by the other firms, potentially enables higher prices and profits if they respond in kind. This latter answer has some justification in the economic literature. When goods are substitutes, as in my examples, Singh and Vives [15] showed that firms prefer to compete on quantities than prices. Competition in quantities realizes higher profits than competition in prices, because the price elasticity of demand is greater than the quantity elasticity of demand. If withholding is an observable, committed action, than potentially all firms in the industry can choose their preferred quantity, which would be lower than production at that price otherwise, to realise higher profits. In this paper, I create a model where firms choose capacity in the long-run, price in the short-run, and have the option to withhold capacity. Using this model, I show that both the above explanations are justified, and draw out the consequences of withholding for equilibrium capacity and output. Finally, I explore the fundamental question raised above, why ferry companies practice withholding but hotels do not.

There is little literature studying this type of capacity withholding. The term has been most recently associated with capacity withholding in the electricity industry, where there is considerable empirical evidence of withholding during periods of high demand. Withholding here likely arises from discontinuous pricing caused by uniform price auctions. When demand is low relative to capacity, there are equilibria in which all firms charge marginal cost, but when demand is high relative to demand, it is possible for firms to charge an unlimited high price. Hence there is a strong incentive to withhold when demand is high. Dechenaux and Kovenock [6] provide a more sophisticated analysis in their recent paper. This behaviour contrasts with my examples, where withholding takes place only during troughs in demand. Another use of the term appears in the literature when owners of a natural resource withhold profitable extraction in the short-run to profit more in the long-run. See for example the paper by Gaskins and Haring [9]. Hotels and ferry sailings are examples that fit most naturally into a differentiated products environment, in which withholding has not been studied, to my knowledge.

The key to the study of withholding here is the distinction between firms' long-run and short-run choices. There is a substantial literature with models where firms choose capacity in the long-run, then price or quantity in the short-run. An example of the latter is the paper by Grimm and Zoettl [10], which does not however employ a differentiated products environment, nor does it allow firms to choose prices. There are many examples of the latter. In Chapter 2 of this thesis, I show in markets where firms choose capacity in the long-run and prices in the short-run, firms will under-utilise capacity at the equilibrium in periods of low demand, rather than choose clearing prices as in the Cournot model. This corresponds well with casual observations of the hotel industry, but there of course firms have no option of withholding.

In this paper, I create a differentiated products model, where firms choose capacity in the long-run, and capacity followed by price in the short-run. I show that withholding both reduces costs and increases revenues, making it a dominant strategy for firms provided the assumptions of observability, commitment, and divisibility are met. In section 3.2, I outline this model and prove existence of pure-strategy equilibria in the whole game. In section 3.3, I compare the symmetric equilibrium to the basic capacity-price model without withholding, showing that firms will invest in more capacity in the long-run. Finally in section 3.4, I discuss the intuition behind capacity withholding and why we do not see more of it in practice.

3.2 The Model

I introduce a model with demand uncertainty, where firms may choose both capacity in the long-run, prices in the short-run, and have the option to withhold capacity in the spot market. This is an extension of the capacity-price model in Chapter 2, modified to explicitly allow firms to withhold capacity. I will henceforth refer to this original model as the *basic* or *no-withholding* model, and the new model in this paper as the *withholding* model.

There are two firms, who produce a symmetrically differentiated non-storable product. Demand for each firm's product is a function of both firms' prices, and of a random variable λ representing the level of market demand. In the basic model, firms play a two-stage game. At stage one, each firm i chooses a capacity level K_i, at a cost of k_i per unit. At this stage, firms do not know the true value of λ, but they do know that λ will be the realisation of a random variable with a particular distribution $F(\lambda)$ which is twice continuously differentiable, and whose support is contained in $[\underline{\lambda}, \overline{\lambda}]$. At the beginning of stage two, λ is revealed, and each firm i chooses a price p_i at which it wishes to sell its product. I denote the demand for firm i's product given p_i, p_{-i}, and λ by $x_i(p_i, p_{-i}, \lambda)$. Since λ is intended to represent market demand, I assume λ enters demand multiplicatively. Each firm i pays a cost c_i for every unit it produces, but cannot produce more than K_i even if demand is greater. After stage two, each firm i receives profits $(p_i - c_i)\,\mathrm{Min}\,[K_i, x_i\,(p_i, p_{-i}, \lambda)]$.

I now extend the model to allow for capacity withholding, by adding a third stage to the game. At stage one, each firm i chooses a capacity level K_i as before. The demand shock is then revealed, and at stage two both firms now choose how much capacity they want to make available. I denote their choices q_1 and q_2 respectively. Finally at stage three, both firms choose prices, with the demand constraint now being $x_i \leq q_i$ rather than $x_i \leq K_i$. The firms may incur costs at each stage. I assume henceforth that both firms face identical costs. The production cost c remains the same as in the basic model, but I divide the capacity cost, originally k, into two parts: k^S, a non-negative cost incurred when firms need to use capacity in the short-run – such as start-up costs, maintenance, or staffing costs, and k^L, a positive cost incurred when firms actually build capacity in the long-run. In the equivalent basic model, firms cannot withhold, so they choose $q_i = K_i$, and incur costs k per unit of capacity installed. To keep the equivalence between the models, I assume $k^S + k^L = k$.

I use most of the assumptions on the demand and cost functions from Chapter 2.[2] Assumption 5 outlines standard regularity conditions on demand functions, and defines the cost functions for the two firms. Assumptions 6(a – e) impose the requirement that the differentiated goods be gross substitutes, and that a firm's demand and marginal profit functions react more strongly to a change in the firm's own price than to a change in the competitor's price. Assumption 6(f) implies that if p_{-i} increases, then firm i's optimal point on its residual demand curve gives rise to both a higher price and quantity in the price subgame. Except for the aforementioned differences in costs, these correspond exactly to assumptions 1 and 2 in Chapter 2.

ASSUMPTION **5.** *The functions $x_i\,(p_i, p_{-i}, \lambda)$ are defined, continuous, single-valued and non-*

[2] These assumptions were themselves modified from Friedman (1988) [8] and extended to allow explicitly for a constant marginal cost of production and the presence of the demand level variable, λ.

negative for all $(p_1, p_2) \in \Re^2_+$, $i = 1, 2$. *Each firm pays a constant cost per unit produced equal to* $c > 0$, *a constant cost for each unit of capacity installed equal to* $k^L > 0$, *and a constant cost per unit of capacity actually used in the spot market equal to* $k^S \geq 0$.

There exists a finite, positive price, p_i^0 *such that* $x_i(p_i, p_{-i}, \lambda) = 0$ *for all* $p_i \geq p_i^0$, *for all* p_{-i}. *Since each firm i's demand will be zero if its price exceeds* p_i^0, *we can limit the price domain to* $\Omega = \left\{(p_1, p_2) \mid 0 \leq p_1 \leq p_1^0, 0 \leq p_2 \leq p_2^0\right\}$. *Further let* Ω_i *be the subset of* Ω *on which* $x_i(p_i, p_{-i}) > 0$.

ASSUMPTION **6.** *The following derivatives exist and satisfy these conditions:*

(a) $\frac{\partial x_i}{\partial p_i} < 0$ *on* Ω_i *and* 0 *otherwise*

(b) $\frac{\partial x_i}{\partial p_{-i}} > 0$ *on* $\Omega_i \cap \Omega_{-i}$ *and* 0 *otherwise*

(c) $\frac{\partial x_i}{\partial p_i} + \frac{\partial x_i}{\partial p_{-i}} < 0$ *everywhere on* Ω_i

(d) $2\frac{\partial x_i}{\partial p_i} + (p_i - c)\frac{\partial^2 x_i}{\partial p_i^2} < 0$ *everywhere on* Ω_i

(e) $2\frac{\partial x_i}{\partial p_i} + (p_i - c)\frac{\partial^2 x_i}{\partial p_i^2} + \left|\frac{\partial x_i}{\partial p_{-i}} + (p_i - c)\frac{\partial^2 x_i}{\partial p_i \partial p_{-i}}\right| < 0$ *everywhere on* Ω_i

(f) $\frac{-\frac{\partial x_i}{\partial p_{-i}}}{\frac{\partial x_i}{\partial x_i}} \geq \frac{-\left(\frac{\partial x_i}{\partial p_{-i}} + (p_i - c)\frac{\partial^2 x_i}{\partial p_i \partial p_{-i}}\right)}{2\frac{\partial x_i}{\partial p_i} + (p_i - c)\frac{\partial^2 x_i}{\partial p_i^2}}$ *everywhere on* Ω_i.

In the basic model I used additional assumptions to establish existence of an equilibrium, and I require the same here. Assumption 7 establishes assumptions directly on the price functions that result when both firms charge clearing prices.

ASSUMPTION **7.** *The following derivatives satisfy these conditions:*

(a) $2\frac{\partial p_i^{CC}(q_i, q_{-i}, \lambda)}{\partial q_i} + q_i \frac{\partial^2 p_i^{CC}(q_i, q_{-i}, \lambda)}{\partial q_i^2} < \frac{\partial p_i^{CC}(q_i, q_{-i}, \lambda)}{\partial q_{-i}} + q_i \frac{\partial^2 p_i^{CC}(q_i, q_{-i}, \lambda)}{\partial q_i \partial q_{-i}} < 0$

(b) $\frac{\partial p_i^{CC}(q_i, q_{-i}, \lambda)}{\partial \lambda} + q_i \frac{\partial^2 p_i^{CC}(q_i, q_{-i}, \lambda)}{\partial q_i \partial \lambda} > 0$,

where p_i^{CC} *is the clearing price that sets* $x_i\left(p_i^{CC}, p_{-i}^{CC}\right) \equiv q_i$.

The first of these assumptions is essentially the decreasing marginal revenue assumption and the decreasing differences assumption (that is, capacity is a strategic substitute),[3] and the second simply says marginal revenue increases in the demand shock.

[3]The requirement that capacity is a strategic substitute is needed only for the proof of existence of an equilibrium. Otherwise, cross effects may be positive, so long as in absolute value they are less than own effects.

Assumptions 7(a) and 7(b) are not made on the primitives of the model, the demand functions, but rather on the price functions which result when firms choose prices to clear demand. I make these assumptions for two reasons. First, the decreasing marginal revenue and decreasing differences assumptions are intuitive, well-known assumptions in the literature, and second, these are exactly the assumptions I need to complete the key proofs in the paper. If I instead placed sufficient conditions on the demand functions to obtain these assumptions, I would end up with conditions necessarily more restrictive. This follows the approach of Singh and Vives [15], and Vives [19], who made assumptions both on the demand functions and the resulting profit functions. To give an idea of what assumptions 7 might entail on the demand functions, I now outline some simple sufficient conditions.

First note that assumption 7 is not generally implied by assumptions 5 and 6. An example that satisfies assumptions 5 and 6 but not assumption 7(a) is the CES type demand function $x_1 = \lambda p_1^{-\epsilon_1} p_2^{\epsilon_2} + \eta$, when $\epsilon_1 > 1$, $\epsilon_1^2 - \epsilon_2^2 - \epsilon_1 < 0$ and η is a small positive real number. However, if all the second derivatives of demand with respect to price are zero, then assumptions 5 and 6 do imply all of assumption 7.

When the second derivatives are non-zero, it is difficult to compare marginal revenues in capacities and prices. Thus, in searching for sufficient conditions, I content myself with looking for conditions only on the second derivatives of demand with respect to price. Assuming all these second derivatives are negative is sufficient to ensure all of assumptions 7, except the assumption that $\frac{\partial^2 p_i^{CC}}{\partial q_i^2} < \frac{\partial^2 p_i^{CC}}{\partial q_i \partial q_{-i}}$. This requires an additional assumption. When the cross-partials are zero, this assumption is

$$\begin{aligned} & -\frac{\partial x_2}{\partial p_2}\left(\frac{\partial x_2}{\partial p_2} + \frac{\partial x_1}{\partial p_2}\right)\left(\frac{\partial^2 x_2}{\partial p_1^2}\frac{\partial x_1}{\partial p_2} - \frac{\partial^2 x_1}{\partial p_1^2}\frac{\partial x_2}{\partial p_2}\right) \\ & > \frac{\partial x_2}{\partial p_1}\left(\frac{\partial x_1}{\partial p_1} + \frac{\partial x_2}{\partial p_1}\right)\left(\frac{\partial^2 x_2}{\partial p_2^2}\frac{\partial x_1}{\partial p_2} - \frac{\partial^2 x_1}{\partial p_2^2}\frac{\partial x_2}{\partial p_2}\right). \end{aligned}$$

Adding back the cross-partials puts two extra terms in the right-hand bracket on each side of the equation.

3.2.1 Pricing Subgame

For any fixed non-negative capacities q_1 and q_2, I show in Chapter 2 that under assumptions 5 and 6, there exists a unique pure strategy equilibrium in prices. This equilibrium is characterised by firms choosing either Bertrand prices that maximise $(p_i - c)\, x_i$, or clearing prices that set demand equal to capacity. This gives rise to four regions. If both firms have relatively low capacity, then both firms will choose clearing prices in equilibrium. I call this the CC equilibrium. If firm 1 has relatively high capacity and firm 2 relatively low capacity, then firm 1 will choose a Bertrand price and firm 2 will choose a clearing price. This I call the BC region. The CB and BB regions are similarly defined. I illustrate these regions in figure 3.1.

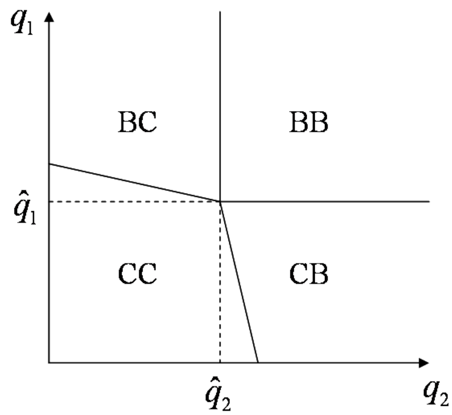


Figure 3.1: Price equilibria for fixed spot market capacity and demand.

For fixed q_2 and λ, as q_1 rises, firm 1 will eventually switch from choosing clearing prices in the subsequent price equilibrium, to choosing Bertrand prices. Define q_1^B by $q_1^B \equiv -\left(p_1^{CC}\left(q_1^B, q_2, \lambda\right) - c\right) \frac{\partial x_1}{\partial p_1}$. This is the critical value of q_1 that makes firm 1 indifferent between the CC equilibrium and the BC equilibrium. Define q_2^B symmetrically. Define $\hat{q}_1$ by $\hat{q}_1 \equiv -\left(p_1^{CB}\left(\hat{q}_1, \lambda\right) - c\right) \frac{\partial x_1}{\partial p_1}$. Depicted in Figure 3.1, this is the value of q_1 that makes firm 1 indifferent between the CB and BB equilibria.[4] When $q_2 \leq \hat{q}_2$, firm 1 switches from clearing to Bertrand prices at the point $q_1 = q_1^B$. When $q_2 > \hat{q}_2$, $q_1 = \hat{q}_1$ is this critical point.

[4] $(\hat{q}_1, \hat{q}_2)$ is also the unique intersection of the equations $q_1 = q_1^B(q_2, \lambda)$ and $q_2 = q_2^B(q_1, \lambda)$.

3.2.2 Withholding Subgame

After the revelation of the true value of λ, each firm seeks to choose capacity q_i to maximise profit subject to the price subgame equilibria. Since there are four possible price equilibria, this maximisation problem is potentially complicated. However, I can establish that in any Nash equilibrium in capacities, both firms will choose clearing prices, even when there is no cost to making spot market capacity available.

LEMMA **3.** *Any Nash equilibrium in spot-market capacities must lie in the CC region.*

Proof. Fix q_2 and λ.

First consider the case $q_2 \leq \hat{q}_2$. In this region, when $q_1 \leq q_1^B$, the price equilibrium is CC, and when $q_1 > q_1^B$, the price equilibrium is BC. In the latter region, profits are decreasing in q_1 if $k^S > 0$ and are constant in q_1 if $k^S = 0$. I now show that at $q_1 = q_1^B$, firm 1's marginal profit in capacity is negative, and thus firm 1's best response is to choose a point inside the CC region.

$$\begin{aligned}
\left.\frac{\partial \pi_1^{CC}}{\partial q_1}\right|_{\left(q_1^B, q_2, \lambda\right)} &= p_1^{CC}\left(q_1^B, q_2, \lambda\right) - c - k^S + q_1^B \frac{\partial p_1^{CC}}{\partial q_1} \\
&= p_1^{CC}\left(q_1^B, q_2, \lambda\right) - c - k^S + q_1^B \left(\frac{\frac{\partial x_2}{\partial p_2}}{\frac{\partial x_1}{\partial p_1}\frac{\partial x_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}}\right) \\
&= p_1^{CC}\left(q_1^B, q_2, \lambda\right) - c - k^S + q_1^B \left(\frac{\partial x_1}{\partial p_1} - \frac{\frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}}{\frac{\partial x_2}{\partial p_2}}\right)^{-1} \\
&< p_1^{CC}\left(q_1^B, q_2, \lambda\right) - c - k^S + q_1^B \left(\frac{\partial x_1}{\partial p_1}\right)^{-1} \\
&\leq p_1^{CC}\left(q_1^B, q_2, \lambda\right) - c + q_1^B \left(\frac{\partial x_1}{\partial p_1}\right)^{-1} \\
&= 0
\end{aligned} \tag{3.1}$$

The inequality follows from assumption 6, and the final equality from the definition of q_1^B.

Now consider the case $q_2 > \hat{q}_2$. In this case, there are three potential price equilibria, but I need only consider the equilibria either side of $q_1 = \hat{q}_1$. When $q_1 > \hat{q}_1$, the price equilibrium is BB, in which case firm 1's profits are weakly decreasing or constant in q_1 as described above. When $q_1 < \hat{q}_1$, the price equilibrium is CB. I now show that at $q_1 = \hat{q}_1$, firm 1's marginal profit in capacity is negative, so firm 1's best response is to choose a point $q_1 < \hat{q}_1$.

$$\begin{aligned}
\left.\frac{\partial \pi_1^{CB}}{\partial q_1}\right|_{(\hat{q}_1,\lambda)} &= p_1^{CB}(\hat{q}_1,\lambda) - c - k^S + \hat{q}_1 \frac{\partial p_1^{CB}}{\partial q_1} \\
&= p_1^{CB}(\hat{q}_1,\lambda) - c - k^S + \hat{q}_1 \left(\frac{\frac{\partial MR_2}{\partial p_2}}{\frac{\partial x_1}{\partial p_1}\frac{\partial MR_2}{\partial p_2} - \frac{\partial x_1}{\partial p_2}\frac{\partial MR_2}{\partial p_1}} \right) \\
&= p_1^{CB}(\hat{q}_1,\lambda) - c - k^S + \hat{q}_1 \left(\frac{\partial x_1}{\partial p_1} - \frac{\frac{\partial x_1}{\partial p_2}\frac{\partial MR_2}{\partial p_1}}{\frac{\partial MR_2}{\partial p_2}} \right)^{-1} \\
&< p_1^{CB}(\hat{q}_1,\lambda) - c - k^S + \hat{q}_1 \left(\frac{\partial x_1}{\partial p_1} \right)^{-1} \\
&\leq p_1^{CB}(\hat{q}_1,\lambda) - c + \hat{q}_1 \left(\frac{\partial x_1}{\partial p_1} \right)^{-1} \\
&= 0 \qquad (3.2)
\end{aligned}$$

The inequality follows from assumption 6, and the final equality from the definition of $\hat{q}_1$.

The result for firm 2 is symmetric.

Equation 3.2 implies that any point in the BB region is dominated for both firms, and equation 3.1 implies that any point in the CB region is dominated by a point in the CC region for firm 1, and vice-versa for firm 2. Applying iterated elimination of dominated strategies, any Nash equilibrium in quantities will lie strictly within the CC region. This completes the proof.

□

Lemma 3 implies each firm's maximisation problem at the withholding stage can be simplified. Firm i chooses q_i to maximise $\left(p_i^{CC}(q_i, q_{-i}, \lambda) - c - k^S\right) q_i$ s.t. $q_i \leq K_i$. By assumption 7(a), the objective function is concave, so the best response function is $q_i^r \equiv \text{argmax} \left(p_i^{CC}(q_i, q_{-i}, \lambda) - c - k^S\right) q_i$, except when q_i^r exceeds K_i. In the latter case, choosing $q_i = K_i$ is the best firm can do, so the best response function for the subgame is $q_i^R = \text{Min}\{K_i, q_i^r\}$. Each firm is either constrained by its long-run capacity, or unconstrained. This leads to four possibilities for the equilibrium – either one firm is constrained, both are, or neither are. If firms are constrained, they choose clearing prices, for which

I have already introduced notation – p_i^{CC}. In the case that a firm is unconstrained, the equilibrium prices will differ from the Bertrand prices in the no-withholding model. If a firm ends up in this equilibrium, I will denote this by a 'U'. For example, p_i^{CU} indicates that firm 1 is choosing clearing prices, but firm 2 is choosing unconstrained prices. I will extend this notation to firms' revenues and profits. Define firm i's revenue in the CC region by $R_i^{CC} = p_i^{CC} q_i$, and profit by $\pi_i^{CC} = R_i^{CC} - cq_i - k^S q_i$. Similar notation extends to the other regions.

When firms are unconstrained, assumption 7 implies the best reply functions are strictly downwards sloping and there exists a unique pure strategy equilibrium. Adding capacity constraints does not change this result. If one of the firms is constrained, and cannot obtain the unconstrained equilibrium, then its best reply function becomes constant, and since the other firm's best reply is downward sloping, there still can be but one equilibrium.

For fixed K_1 and K_2, when the demand shock is low, firm i will be unconstrained, and can choose spot market capacity to maximise profit. The resulting spot market capacity is lower than the firm's long-run capacity, so the firm is withholding. When the demand shock is high, firm i's best response is constrained by K_i. Now firm i chooses to produce at full capacity in the subgame equilibrium. Similar to the price subgame, there exists a unique value of λ at which firm i switches from the unconstrained best response to the constrained best response. I will denote this point λ^{C_i}. The firm with the lower long-run capacity will be the first to be constrained as demand rises, so if $K_1 < K_2$, $\lambda^{C_1} < \lambda^{C_2}$.

When $K_i < K_{-i}$, then these points are formally defined by

$$\lambda^{C_i} : K_i \equiv - \left(p_i^{CC}\left(K_i, q_{-i}, \lambda^{C_i}\right) - c - k^S\right) \left(\frac{\partial p_i^{CC}}{\partial q_i}\right)^{-1}, \tag{3.3}$$

$$\lambda^{C_{-i}} : K_{-i} \equiv - \left(p_{-i}^{CC}\left(K_i, K_{-i}, \lambda^{C_{-i}}\right) - c - k^S\right) \left(\frac{\partial p_{-i}^{CC}}{\partial q_{-i}}\right)^{-1}. \tag{3.4}$$

Lemma **4.** *λ^{C_1} and λ^{C_2} are (weakly) increasing in K_1 and K_2.*

Proof. I calculate the derivatives implicitly. Here I show the case $K_1 < K_2$, and the reverse case follows symmetrically.

$$\frac{\partial \lambda^{C_2}}{\partial K_2} = \frac{-\left(2\frac{\partial p_2^{CC}}{\partial q_2} + K_2 \frac{\partial^2 p_2^{CC}}{\partial q_2^2}\right)}{\frac{\partial p_2^{CC}}{\partial \lambda} + K_2 \frac{\partial^2 p_2^{CC}}{\partial q_2 \partial \lambda}}$$

$$\frac{\partial \lambda^{C_2}}{\partial K_1} = \frac{-\left(\frac{\partial p_2^{CC}}{\partial q_1} + K_2 \frac{\partial^2 p_2^{CC}}{\partial q_1 \partial q_2}\right)}{\frac{\partial p_2^{CC}}{\partial \lambda} + K_2 \frac{\partial^2 p_2^{CC}}{\partial q_2 \partial \lambda}}$$

$$\frac{\partial \lambda^{C_1}}{\partial K_1} = \frac{-\left(2\frac{\partial p_1^{CC}}{\partial q_1} + K_1 \frac{\partial^2 p_1^{CC}}{\partial q_1^2}\right)}{\frac{\partial p_1^{CC}}{\partial \lambda} + K_1 \frac{\partial^2 p_1^{CC}}{\partial q_1 \partial \lambda}}$$

$$\frac{\partial \lambda^{C_1}}{\partial K_2} = 0.$$

By assumption 7, the first three derivatives are all positive and the last is zero, giving the result. □

3.2.3 Capacity Choice in the Whole Game

I now find conditions under which subgame-perfect equilibria of the whole game exist. I will elaborate from the point of view of firm 1, noting that the argument for firm 2 is symmetric. Firm 1's payoff function is defined as the expectation with respect to λ over the reduced profit functions. This payoff function is differentiable with respect to K_1, but the first derivatives are only piecewise differentiable with respect to K_1, because the subgame equilibria are only piecewise differentiable in K_1. The most important distinction is if K_1 is greater than or less than K_2. Fixing K_2, first assume $K_1 \geq K_2$. Then firm 1's first order condition takes the form

$$\int_{\lambda^{C_1}}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}(K_1, K_2, \lambda)}{\partial K_1} - c - k^S\right) dF(\lambda) - k^L. \tag{3.5}$$

Now suppose instead $K_1 < K_2$. Then firm 1's first order condition takes the form

$$\begin{aligned}&\int_{\lambda^{C_2}}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}(K_1, K_2, \lambda)}{\partial K_1} - c - k^S\right) dF(\lambda) \\ &+ \int_{\lambda^{C_1}}^{\lambda^{C_2}} \left(\frac{\partial R_1^{CU}(K_1, \lambda)}{\partial K_1} - c - k^S\right) dF(\lambda) - k^L.\end{aligned} \tag{3.6}$$

When $K_1 = K_2 = K$, these two equations collapse to a single equation,

$$E(K) \equiv \int_{\lambda^C(K)}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}(K, K, \lambda)}{\partial K_1} - c - k^S \right) dF(\lambda) - k^L, \tag{3.7}$$

where $\lambda^C(K) = \lambda^{C_2}(K, K) = \lambda^{C_1}(K, K)$. $E(K) = 0$ is the necessary condition for a symmetric equilibrium.

I now briefly demonstrate that $E(K)$ is decreasing in K. $E'(K)$ is

$$\int_{\lambda^C(K)}^{\overline{\lambda}} \frac{\partial}{\partial K} \left(\frac{\partial R_1^{CC}(K, K, \lambda)}{\partial K_1} \right) dF(\lambda) + F'\left(\lambda^C\right) \left(\frac{\partial R_1^{CC}\left(K, K, \lambda^C\right)}{\partial K_1} - c - k^S \right) \frac{\partial \lambda^C}{\partial K}.$$

The right-hand term in this equation is zero, because marginal profit at λ^C is zero by definition of λ^C. The integrand in the left-hand term is negative by assumptions 7(a) and 7(b). Thus, the derivative of $E(K)$ is negative, giving the result.

Proposition **3.** *Under assumptions 5 – 7, there is a unique symmetric Nash equilibrium in pure strategies in this game.*

Proof. A symmetric equilibrium, $K^\star$, must satisfy $E(K^\star) = 0$. Since $E(K)$ is decreasing, uniqueness is straightforward. Under assumptions 5 – 7, firms' payoff functions are quasiconcave at $K^\star$, which proves uniqueness. See the Appendix for a detailed proof. □

3.3 Comparison with no-withholding game

Thus far I have established that under assumptions 5 – 7 there is a unique symmetric equilibrium, in which firms choose to withhold capacity in low demand states. One of my interests in this paper is to compare the outcomes of a market where withholding is allowed against one where it is banned, or impossible for some reason. The same assumptions however, do not guarantee an equilibrium in the no-withholding model, as explained in Chapter 2. Nonetheless, in that chapter I showed that there are cases where assumptions 5 – 7 are satisfied and a unique symmetric equilibrium does exist.[5] In this section, I

[5] See Proposition 2 in Chapter 2.

will proceed under the assumption that a symmetric equilibrium exists in both models. I compare the two equilibria to establish the effect of withholding on equilibrium output, capacity, profits, and welfare.

Denote K_W, P_W, and x_w to be the equilibrium capacity, price, and output in the withholding game. Denote K_N, P_N, and x_N to be the equilibrium capacity, price, and output in the game without withholding. Similarly denote spot market capacity by q_N and q_W, where $q_N = K_N$ and $q_W = x_W$.

3.3.1 Investment in Capacity

When withholding is allowed, equilibrium capacity in the symmetric equilibrium is strictly greater.

PROPOSITION **4.** $K_W > K_N$

Proof. K_W must satisfy the necessary condition

$$\int_{\lambda^C(K_W)}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_W, K_W, \lambda\right)}{\partial K_1} - c - k^S \right) dF\left(\lambda\right) - k^L = 0, \tag{3.8}$$

where $\lambda^C\left(K\right)$ satisfies $K \equiv -\left(p_1^{CC}\left(K, K, \hat{\lambda}\right) - c - k^S\right)\frac{\partial p_1^{CC}}{\partial K_1}$.

When withholding is not an option, the equivalent condition satisfied by K_N is

$$\int_{\hat{\lambda}(K_N)}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_N, K_N, \lambda\right)}{\partial K_1} - c \right) dF\left(\lambda\right) - k = 0, \tag{3.9}$$

where $\hat{\lambda}\left(K\right)$ satisfies $K \equiv -\left(p_1^{CC}\left(K, K, \hat{\lambda}\right) - c\right)\frac{\partial x_1}{\partial p_1}$[6].

I need the following property to complete the proof.

Property A

$$\hat{\lambda}\left(K\right) < \lambda^C\left(K\right), \forall K.$$

I use assumption 7(b), that marginal revenue in capacity is increasing in λ, to prove this. Using a similar argument to Lemma 3, I first show that marginal revenue at $\hat{\lambda}$ is negative.

[6]In Chapter 2, I define $\hat{\lambda}$ from the equilibrium prices. However, this definition is equivalent, and may be obtained directly from the first-order conditions of the price subgame.

$$\begin{aligned}
\left.\frac{\partial \pi_1^{CC}}{\partial K_1}\right|_{(K,K,\hat{\lambda})} &= p_1^{CC}\left(K,K,\hat{\lambda}\right) - c - k^S + K\frac{\partial p_1^{CC}}{\partial K_1} \\
&= p_1^{CC}\left(K,K,\hat{\lambda}\right) - c - k^S + K\left(\frac{\partial x_1}{\partial p_1} - \frac{\frac{\partial x_1}{\partial p_2}\frac{\partial x_2}{\partial p_1}}{\frac{\partial x_2}{\partial p_2}}\right)^{-1} \\
&< p_1^{CC}\left(K,K,\hat{\lambda}\right) - c - k^S + K\left(\frac{\partial x_1}{\partial p_1}\right)^{-1} \\
&\leq p_1^{CC}\left(K,K,\hat{\lambda}\right) - c + K\left(\frac{\partial x_1}{\partial p_1}\right)^{-1} \\
&= 0
\end{aligned} \tag{3.10}$$

The first inequality here follows immediately from assumption 6, and the final equality from the definition of $\hat{\lambda}$. This proves marginal profits are negative at $\hat{\lambda}$, and are equal to zero at λ^C. By assumption 7(b) marginal revenue is increasing in λ, so λ^C must be greater than $\hat{\lambda}$. Now I can show the following.

$$\begin{aligned}
0 &= \int_{\hat{\lambda}(K_N)}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_N,K_N,\lambda\right)}{\partial K_1} - c\right) dF\left(\lambda\right) - k \\
&= \int_{\lambda^C(K_N)}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_N,K_N,\lambda\right)}{\partial K_1} - c - k^S\right) dF\left(\lambda\right) - k^L \\
&\quad + \int_{\hat{\lambda}(K_N)}^{\lambda^C(K_N)} \left(\frac{\partial R_1^{CC}\left(K_N,K_N,\lambda\right)}{\partial K_1} - c - k^S\right) dF\left(\lambda\right) - k^S \int_{\underline{\lambda}}^{\hat{\lambda}(K_N)} dF\left(\lambda\right) \\
&< \int_{\lambda^C(K_N)}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_N,K_N,\lambda\right)}{\partial K_1} - c - k^S\right) dF\left(\lambda\right) - k^L \\
&\quad - k^S \int_{\underline{\lambda}}^{\hat{\lambda}(K_N)} dF\left(\lambda\right) \\
&\leq \int_{\lambda^C(K_N)}^{\overline{\lambda}} \left(\frac{\partial R_1^{CC}\left(K_N,K_N,\lambda\right)}{\partial K_1} - c - k^S\right) dF\left(\lambda\right) - k^L \\
&= E\left(K_N\right)
\end{aligned} \tag{3.11}$$

The first inequality holds because marginal profits are decreasing in λ, and are zero at λ^C. Thus the integral of marginal profit between λ^C and $\hat{\lambda}$ is strictly negative.

I have already shown that $E\left(K\right)$ is decreasing in K. The inequality at 3.11 proves $E\left(K_N\right) >$

0, and by definition $E(K_W) = 0$, therefore $K_W > K_N$. Thus firms choose to hold more capacity in equilibrium when withholding is allowed.

□

The intuition for this is straightforward. When withholding is not allowed, in equilibrium, firms will make negative marginal returns to capacity for some states of demand. When withholding is allowed, firms are able to adjust their output and prices so that they always make a non-negative marginal return to capacity in all states. This difference accounts for the increase in capacity in the withholding model.

3.3.2 Output

When firms have the option of withholding, in the new equilibrium they will increase output in periods of high demand, but decrease it in times of low demand. First consider what happens when the demand shock λ is in the range $[\lambda^C(K_N), \overline{\lambda}]$. When demand is this high, firms will choose to produce at full capacity, with or without the option of withholding. Since $K_W > K_N$, it follows that output will be greater under the withholding option.

At the low end of demand, $\lambda \in [\underline{\lambda}, \hat{\lambda}(K_N))$, firms will choose q to maximise marginal revenue in quantities when withholding is allowed, but choose p to maximise marginal revenue in prices when withholding is prohibited. The equations in Lemma 3 can be quickly modified to show that at the quantity that maximises marginal revenue in capacities, marginal revenue in prices is strictly negative. Thus each firm wants to choose a lower price with a correspondingly higher output. This is true for both firms' best responses, and is thus true at the equilibrium. I conclude that firms produce less when withholding is allowed and demand is low.

Finally, when $\lambda \in [\hat{\lambda}(K_N), \lambda^C(K_W))$, firms in the no-withholding case will be producing at full capacity, but firms in the withholding case will choose to withhold, and produce at less than full capacity. At $\hat{\lambda}(K_N)$, $x_W < x_N = K_N$, and at $\lambda^C(K_W)$, $x_W = K_W > K_N$. So whereas $x_N = K_N$ stays constant in this range, x_W is increasing in λ (by assumption 7(b)), until eventually output under the withholding option surpasses that of the no-withholding option.

It is reasonable to question whether the lower costs in the withholding model drive any of the results for capacity and output. Note however that all my results so far include the special case $k^S = 0$, the case in which there are no lower costs to withholding capacity. In particular, the crucial equations in Proposition 4, equations 3.10 and 3.11, still hold. There are two effects present. Withholding allows firms to increase revenues *and* decrease costs, so even if I take away the cost advantage, the effects on capacity and output are unchanged.

3.3.3 Profits and Welfare

When $k^S > 0$, the differing cost structure of this withholding model from the original capacity-price model prevents me from predicting the effects of withholding on short-run profits, since per-unit short-run costs are different between the models. Similarly, I cannot predict the effect on welfare. However, when $k^S = 0$, I show that withholding increases spot market profits in all states of demand, increases welfare when demand is high and decreases welfare when demand is low.

Consider first spot market profits, that is, profits excluding long-run capacity costs. In the case $\lambda \in [\lambda^C(K_W), \overline{\lambda}]$, firms are constrained whether or not withholding is allowed; and since capacity is higher with withholding, and the spot market profit functions are concave, profits too are higher.

In the low demand set, $\lambda \leq \hat{\lambda}(K_N)$, withholding output is $x_W = q_W$, which is the output that maximises spot market profit $\left(p^{CC}(x_W) - c\right) x_W$. Since spot market profits are concave, this must be greater than $\left(p^{CC}(x_N) - c\right) x_N$. At $x_N < K_N$, the Bertrand price p_N chosen by firms is simply the clearing price that 'clears' x_N, so this last expression is simply short-run profit in the no-withholding case. I conclude in this case that spot market profit is again greater.

In the remaining cases the withholding firm will be unconstrained, so still maximising profit. The non-withholding firm will not be maximising profit, but will be forced to choose a lower capacity at a clearing price, which must be lower.[7]

Since I assume constant marginal costs of production, total welfare in the short-run, measured by the sum of consumer's surplus and firms' profits, and excluding long-run costs, will vary positively in proportion to actual output. When withholding is an option, out-

[7]At some point, the two firms will produce the same output (and make the same profit) before the withholding firm goes on to make more profit by producing more.

put is lower in low demand states, but higher in high demand states, and thus the same is true of welfare. Of greater interest is how expected welfare changes between the no-withholding and withholding models. Unfortunately, expected welfare under withholding does not always increase or decrease compared to the no-withholding model. In the simple case of linear demand with a uniform distribution, I can compute out the equilibria in the two models, and find parameters for which expected welfare increases, and parameters for which expected welfare decreases.[8]

3.4 Discussion

In section 3.2 I established that firms in markets with uncertain demand, non-storable goods, and capacity constraints prefer to withhold capacity when demand is low. Underlying this result is the intuition that firms use withholding to transform the subgame from a competition in prices to competition in quantities. Quantity competition requires the often implicit assumption that prices are set to clear the market, no matter what quantities are chosen. In the model presented in this paper, firms have the option of choosing their own prices. However, in Lemma 3, I proved that when firms had the option to withhold capacity in the spot market, iterated elimination of dominated strategies caused each to choose clearing prices in the subsequent price subgame equilibrium. This established that the short-run subgame under withholding is indeed a competition in quantities. Since I assume firms' goods are substitutes, the elasticity of demand in quantities is lower than the elasticity of demand in prices, therefore firms realise higher revenues and profits. osen. In the model presented in this paper, firms have the option of choosing their own prices. However, in Lemma 3, I proved that when firms had the option to withhold capacity in the spot market, iterated elimination of dominated strategies caused each to choose clearing prices in the subsequent price subgame equilibrium. This established that the short-run subgame under withholding is indeed a competition in quantities. Since I assume firms' goods are substitutes, the elasticity of demand in quantities is lower than the elasticity of

[8]When $x_i = \lambda\,(a - b_1 p_i + b_2 p_{-i})$, choosing $a = 50$, $b_1 = 1$, $b_2 = 0.2$, $c = 0$, $k^L = 10$, and $k^S = 0$ gives $K_N = 34.2241$ and $K_W = 34.2364$, which results in expected welfares equal to 2248.32 and 2248.4 in the basic and withholding models respectively. In this case, the withholding model realises higher expected welfare. Now change b_2 to 0.5, and hold everything else constant. In this case, $K_N = 41.4427$ and $K_W = 42.0552$, and expected welfares are 3526.63 and 3507.33, in the basic and withholding models respectively. Now withholding decreases expected welfare.

demand in prices, therefore firms realise higher revenues and profits.

Given firms' propensity to withhold capacity, it is surprisingly hard to find real world examples of withholding in action. In the introduction, I made a comparison between hotels and ferry sailings – two industries that share the key characteristics of uncertain demand, non-storable goods, and capacity constraints. Ferry companies practice capacity withholding; hoteliers do not, at least not in a way that is easily visible to an outside observer. Why the difference? There are three additional assumptions implicit in my model that explanation this observation. I assume that firms can commit to withholding, that all other firms can observe this withholding, and that capacity is divisible. The three are related. Hotels are unable to commit to withholding capacity, precisely because it is practically impossible to observe whether a hotel puts a wing of rooms out of service or not. Even if it did, there is nothing stopping the firm from using those rooms in the short-run, another unobservable action. The only credible way a hotel can commit to withholding is to close the whole hotel. Such an action creates an all or nothing decision, a lack of divisibility which is outside the scope of my model.

The ability to commit to a given capacity level in the spot market is crucial to the results. Consider the alternative. If at the price stage, a firm could use all its long-run capacity capacity at no extra cost aside from production costs, then the unique equilibrium in the price game would be exactly the same as in the no-withholding game, since the only real constraint is the long-run capacity constraint. In this case, the equilibrium will revert to the no-withholding equilibrium. Thus it is critical firms be able to commit to their capacity withholding to gain the benefits. Ferry companies use at least two commitment tactics. One tactic is to put ferries in the dry-dock for maintenance in the off-season, or even go to the extreme of sending them to other markets. The other tactic is prior scheduling. Ferries publish their sailings month in advance, and it is easy to verify that the firm is indeed sailing, or not sailing when it says it is. It is somewhat ironic that the transparency of the operation is exactly what enables firms to perform this type of 'tacit collusion'.

The bar for firms to successfully withhold are thus quite high. Firms in these capacity-constrained industries with fluctuating demand would indeed prefer to practice withholding; an option which is largely prevented by the lack of commitment mechanisms and divisibility issues. What are the consequences for consumers? When firms are unable to withhold capacity, I showed in section 3.3 that industry capacity was lower. When demand has high, this causes higher prices and lower welfare. On the other hand, as any tourist travelling in the off-season can testify, consumers gain at times of low demand since firms cannot commit to quantity competition. When withholding is possible, the reverse effects apply; consumers benefit at times of high demand but lose when demand is low. Which is the dominant effect? Since the firms' gain at all times when withholding is allowed, my intuition suggests that consumers are likely worse off in expectation. However, the capacity constraints cause a counter effect. In high demand states, consumers too are better off. In section 3.4, I gave an example above where the gain to consumers outweighs the additional deadweight losses in expectation. Thus withholding need not universally make consumers worse off.[9]

I noted in the introduction that this type of capacity withholding is distinct from withholding as practiced in the electricity industry. This is confirmed in the model; electricity firms do not withhold capacity at times of low demand. However, there is one similarity. Both types of withholding are a form of tacit collusion, where firms have a dominant strategy to change the game to achieve equilibria that realise higher profits. These results have another nice interpretation. Capacity-constrained firms in reality almost always compete on price. Yet there are many papers in the literature that assume firms compete on quantities, usually because this is the more tractable assumption. I showed in this paper that firms can compete in quantities in the real world simply by controlling their short-run capacity. This is a possible justification for the assumption of quantity competition, when the commitment assumptions are met.

3.5 Appendix

Proof of Proposition 3

Let $K^\star$ satisfying $E(K^\star) = 0$ be the candidate equilibrium, and let $K_2 = K^\star$. Note that such a point exists because $E(K)$ is decreasing in K, $E(0) > 0$ and we can choose K such that at $\overline{\lambda}$, firms choose $q < K$, so at this point $E(K) < 0$. First suppose $K_1 > K_2$. Then the second derivative of firm 1's profit function is

[9] Compare to this to the unconstrained case studied by Singh and Vives, where, in the linear case, quantity competition strictly decreased total welfare.

$$\int_{\lambda^{C_1}}^{\overline{\lambda}} \frac{\partial^2 \pi_1^{CC}}{(\partial K_1)^2} dF(\lambda) + \frac{\partial \lambda^{C_1}}{\partial K_1} \frac{\partial \pi_1^{CC}\left(K_1, K_2, \lambda^{C_1}\right)}{\partial K_1} f\left(\lambda^{C_1}\right). \tag{3.12}$$

Since $\lambda^{C_1}(K_1)$ by definition is the λ that makes K_1 the optimal choice for firm 1, the middle expression in the second term is zero. The first term is negative by assumption 7(a). Thus when $K_1 > K_2$, firm 1's payoff function is concave.

Now suppose $K_1 \leq K_2$. In this region, the profit function is not concave, but following the proof of Grimm & Zoettl [10], I will show that it is quasiconcave. The following properties will be helpful in completing the proof.

Property B

$$\frac{\partial \pi_1^{CU}(K_1, \lambda)}{\partial K_1} \geq 0 \,\text{if}\, \lambda \geq \lambda^{C_1}(K_1)$$

The proof of this is intuitive. Since firm 1 is constrained ($\lambda \geq \lambda^{C_1}$), firm 1 would like to produce more than K_1 for all demand realisations in this region. This is not possible however, due to the capacity constraint.

Property C

$$\frac{\partial \pi_1^{CC}(K_1', K_2, \lambda)}{\partial K_1} \geq \frac{\partial \pi_1^{CC}(K_1, K_2, \lambda)}{\partial K_1} \geq 0 \quad \forall K_1' < K_1, \quad \forall \lambda \geq \lambda^{C_2}(K_1, K_2)$$

The first inequality follows from concavity of the profit functions in the spot market, implied by assumption 7(a). The second inequality is due to the fact that when both firms are constrained and λ is sufficiently high, firm 1 would like to produce more. This is not possible however, due to the capacity constraint.

Now I can use properties B and C, along with Lemma 4, to complete the proof of existence.

$$\begin{aligned}
\frac{\partial \pi_1 (K_1, K_2, \lambda)}{\partial K_1} &= \int_{\lambda^{C_2}(K_1,K_2)}^{\overline{\lambda}} \frac{\partial \pi_1^{CC} (K_1, K_2, \lambda)}{\partial K_1} dF(\lambda) \\
&\quad + \int_{\lambda^{C_1}(K_1)}^{\lambda^{C_2}(K_1,K_2)} \frac{\partial \pi_1^{CU} (K_1, \lambda)}{\partial K_1} dF(\lambda) - k^L \\
&\geq \int_{\lambda^{C_2}(K_1,K_2)}^{\overline{\lambda}} \frac{\partial \pi_1^{CC} (K_1, K_2, \lambda)}{\partial K_1} dF(\lambda) - k^L \\
&= \int_{\lambda^{C_2}(K_1,K_2)}^{\lambda^{C_2}(K_2,K_2)} \frac{\partial \pi_1^{CC} (K_1, K_2, \lambda)}{\partial K_1} dF(\lambda) \\
&\quad + \int_{\lambda^{C_2}(K_2,K_2)}^{\overline{\lambda}} \left[\frac{\partial \pi_1^{CC} (K_1, K_2, \lambda)}{\partial K_1} - \frac{\partial \pi_1^{CC} (K_2, K_2, \lambda)}{\partial K_1} \right] dF(\lambda) \\
&\quad + \int_{\lambda^{C_2}(K_2,K_2)}^{\overline{\lambda}} \frac{\partial \pi_1^{CC} (K_2, K_2, \lambda)}{\partial K_1} dF(\lambda) - k^L
\end{aligned}$$

The inequality holds by property B. In the final equation, the first term is positive by property C and Lemma 4, the second term is positive by property C, and the third term is zero since it is the equilibrium condition.

Thus I conclude that for all $K_1 \leq K_2$, $\frac{\partial \pi_1(K_1,K_2)}{\partial K_1} \geq \frac{\partial \pi_1(K_2,K_2)}{\partial K_1} = 0$. Together with the concavity of π_1 when $K_1 > K_2$, I conclude that π_1 is quasiconcave, and is maximised at $K_2 = K^{\star}$.

Chapter 4

Economic Experiments and Kirchoff's Laws

4.1 Introduction

In the past few years, economists studying electricity markets have been turning to economic experiments to answer questions found to be theoretically intractable. The size of electricity networks combined with the large number of nonlinear physical constraints potentially give experiments the edge in such research. However, experiments conducted to date have mostly been small scale with either no network or at most three to four transmission lines. In part this is due to the lack of complex software required to run more complicated experiments, and in part this is due to a question of understanding. Can human subjects really imitate firms realistically, firms that often devote whole departments to pricing? In this paper, I examine facets of this question experimentally, using infrastructure built for the EPNES project at Caltech.

Previous experiments on electricity networks include Chao and Plott [4], and Backerman, Rassenti, and Smith [1]. Chao and Plott were interested in testing a new decentralised electricity market design, based on a design by Chao and Peck [3]. They were primarily interested in the efficiency of the mechanism. On a three node 'triangle' network (2 firms, 1 demand node) they found the new design to be highly efficient in the presence of a single line constraint.[1] Backerman et al. were more interested in pricing power transmission in the presence of a line constraint. Their network was a modified radial network with three firms either side of a single demand node termed 'London'. They noted that one effect of

[1] One transmission line could potentially constrain the optimal solution when demand was sufficiently high.

the transmission constraint was to give a pricing advantage to the firms not affected by the constraint.

While these experiments succeeded in proving various economic ideas on small networks, in an ideal world economists would like to pre-test designs for electricity markets on experimental replicas of the actual market networks. It would be fantastic if the new designs for the California market could be tested experimentally by human subjects or computerised automata on a simulated California grid, at relatively little cost, before being subjected to the gaming of firms in real-life. In practice, experiments on electricity markets have not moved much past the three node stage. As noted above, there is a big question mark as to whether human subjects playing on large experimental networks can really tell us anything about the real world performance of such networks.

At first glance, it does seem fanciful to expect subjects to make sound profit-maximising decisions when faced with, say, the entire California network. The California grid has hundreds of nodes and thousands of physical constraints. But such complexity is an illusion. Many of the constraints on the system will in fact never bind, or bind only in extreme cases, such as when demand is very high. Of the remainder, only a subset will be binding at any given point in time, depending on the relative balance of demand and supply in the network, which in turn depends upon market bids and pricing. This is important, because in market designs such as the power exchange market, constraints that do not bind are economically irrelevant. They do not affect the prices firms receive and thus do not enter the profit-maximising decision process.

When we examine the network in this light, it becomes apparent that the binding constraints subdivide the network into regions. For example, California is often said to divide into two main regions, North and South. In the summer, demand is particularly high in the South, but generation is high in the North. This implies that constraints on the transmission lines running north to south are important. In summer, such lines are constrained, thus dividing California into two regions. The constraint on electricity flowing from the North gives Southern generators relative market power, and the opportunity to set higher prices. Such a region, where demand is high and local supply is low, is called a load pocket. However for much of the year, the constraint is not binding and thus economically irrelevant.

These regions are of interest because they dramatically simplify the network. A firm does not need to consider the impact of every constraint in the system, only those that could possibly be binding, because these are the constraints that have the potential to confer market power upon the firm. A subject performing an experiment similarly has a simple task; to identify which constraints shape the region their firm is located in and how to manipulate, via bidding and supply, those constraints to maximise market power, i.e., When are you in a load pocket and how can you maximise profit accordingly? Moreover, in real life, firms bid repeatedly on the same network, so there can be no objection to using very experienced subjects. I argue that combining this relative simplicity with sufficient experience should be sufficient for subjects to make good economic decisions on the network.

While there are many types of physical constraints in electricity networks, the important ones for electricity markets are transmission limits. Constraints may arise from thermal and/or stability limits on transmission lines or as a side-effect of a shortage of reactive power. Typically a constraint can be expressed as a maximum megawatt flow through a given line. The exception is thermal limits caused by reactive power in which case the limit is fluid because it depends on the amount of reactive flow. I ignore this last type of constraint because reactive power tends to be generated locally (it cannot be transmitted over long distances), and is typically purchased and used separately from real flows. Thus, in this paper, I consider only fixed megawatt limit constraints on transmission lines.

Whether a constraint will bind or not depends on the relative power flows through the transmission lines. These flows in turn depend critically on Kirchoff's laws, physical rules that govern electricity flow in a network. It follows that any constraint in a network potentially could affect any firm's profit. In a three node network, it is possible to mentally compute approximate power flows for any given inputs/outputs of electricity. In larger networks, this rapidly becomes impossible. It is quite possible that a constraint may be economically important, but that a subject might not recognise this importance immediately. For the purposes of this paper, I distinguish two types of constraint. I will denote a constraint to be *simple* when it affects a transmission line that links two distinct sub-networks. In this case, there is only one way for electricity to flow from one sub-network to the other, and that is through the potentially constrained line. Here it is relatively easy to gauge how the constraint will affect by pricing by calculating relative demand and supply on each side. The other type of constraint I denote *Kirchoff*, short for 'Kirchoff-induced'. Here there is a constraint on a transmission line, but at least one alternative path for the electricity to take between the two nodes connected by the constrained line. Kirchoff's laws require a certain amount of electricity to flow along each line, which can change depending upon how much electricity flows in and out of each node. This type of constraint is a lot trickier to identify just by looking at a network.

Therefore, the focus of this paper is deliberately narrow. I aim to (a) show that human subjects are capable of recognising and profiting from binding constraints and (b) identify any differences in subjects' behaviour between simple constraints and Kirchoff constraints. Although I present the subjects with networks containing only one relevant constraint, it is my contention that if subjects can identify one load pocket, that they can identify multiple load pockets, given sufficient experience on a given network.

In this paper, I present an economic experiment conducted with four firms on eight node networks explicitly designed to test subject's understanding of physical constraints in electricity networks. The networks I create are non-trivial, in the sense that each has two loops and at least one transmission constraint potentially binding some of the time. Two networks have simple constraints, and the other two have Kirchoff constraints. In each network, the constraints are arranged in such a way as to give one subject market power. All subjects have prior experience on similar networks before participating in the actual experiment. The market I use in these experiments is a basic version of a power exchange market, which is widely used around the world in various forms.[2] I discuss the setup of the experiment in sections 4.2 and 4.3, and I outline the results in section 4.4.

4.2 Model Setup

I constructed four distinct networks, each a minor variation on the next. In all networks, I placed four firms at different nodes, along with four demand nodes, all linked by a transmission grid. All firms had the same generating capacity, 100 units, and marginal cost equal to 10 francs per unit. Demand at each demand node was uniform random, between known limits. For simplicity, I assumed all transmission lines to be the same length, but allowed them to differ in their stated maximum capacity.

[2]Stoft [16] pp226-9 discusses various market designs, including the power exchange market.

To test for the effect of a Kirchoff constraint versus a regular constraint, I varied this basic network in the following way. Two of the four networks had a simple constraint which put one firm into a load pocket, giving it considerable market power. I called this firm the *monopolist*. The other two networks had a Kirchoff constraint. This constraint had the same real effect as the simple one – one firm ended up in a load pocket, gaining monopoly power. In both cases I set the constraints and demand values such that the monopolist had the same expected demand with the same support, and the other firms similarly faced the same distribution of demand. This ensured that any differences between the two networks were purely cosmetic, and that any differences in subject behaviour could be attributed to their understanding of Kirchoff's laws.

In either case, I further distinguished between a high demand network and a low demand network.[3] The theory on uniform auctions has different predictions in the static game for low versus high demand, see for example Fabra, von der Fehr and Harbord [7], and I wanted to control for any differences in the data this might make.

On each network, firms bid in a power exchange market, which was repeated for approximately 40 periods.[4] Firms were required to bid one price for all their capacity, their bids being restricted to the set [0, 30], and further forced to bid in every period. Bids were sealed until the end of the period, at which time the true values of demand for that period were revealed. The price at each node was determined by the central authority, in our case a computer algorithm, based on maximising total welfare subject to each firm receiving at least what they bid, this being the basic premise of the power exchange market. Firms were then paid the marginal price at their node.

In any network, if the monopolist chose to bid the maximum possible bid, 30 francs, then the other three firms were effectively bidding against each other in a uniform auction. In the low demand setting, I set total demand between 150 and 200 units, so that the highest bidder out of these firms would supply nothing and make no profit. In the high demand setting, I set demand between 200 and 260 units, so that all firms were guaranteed to supply something, still assuming the monopolist's bid was the highest. However, if the monopolist's bid was not the highest of all four firms, then the transmission constraint no longer held,

[3] I include diagrams of the four networks in Appendix A.

[4] I varied the number of periods in each treatment so that subjects would not anticipate the last period. In the experimental literature, subjects' behaviour can change dramatically in the last iterations of an experiment.

and the market became exactly equivalent to a uniform auction for all four firms. Here, in the low demand case, total demand worked out between 210 and 280 units, so the highest bidder was still guaranteed to supply nothing, while in the high demand case, total demand was between 260 and 340 giving the highest bidder a 50% chance of supplying nothing.

In all cases, the nature of the constraints guaranteed the monopolist would supply a minimum of 60 units, no matter what the bids were. Thus the monopolist could make a minimum profit of 1200 francs per period, just by bidding 30. However, if another firm was bidding 27 or higher, than it became in the monopolist's best interest to undercut that firm, to sell all 100 units at a price 27 or greater, thus realising a higher profit.

4.3 Experimental Procedure

I ran a series of identical experiments using the four networks described above. In each experiment, I divided the subjects into groups of four. Each subject in a group represented one firm, labelled from 0 to 3. Each group performed the same experiment independently. Each group participated in a basic power exchange market on each of the networks, in the order A, C, B, D (see Appendix A).[5] The firms in each network were arranged so that each subject in a group would be the monopolist in just one of the four networks.

Subjects for the experiments were drawn from a pool of experienced subjects who had previously participated in one, but no more than one, practice experiment. I required this training for two reasons. As noted above, I wanted the subjects to have some minimum level of experience, both to imitate real life more closely and to avoid randomness in the data as the subjects learned how to bid in the market. Moreover, due to the complexity of the network, power exchange market and other elements, the instructions took over an hour to complete, including three quizzes to check the subjects' understanding. Thus by pushing the instructions to the practice experiment, I was able to complete the real experiments in a reasonable two hours.[6]

[5] In the very first experiment, the order was intended to be A, B, C, D. However, this put the two low-demand networks first. Relatively inexperienced subjects tended to make little or no money on these networks, and halfway through the first network in the very first experiment, one subject asked to leave because the other subjects were "stupid". Therefore I altered the network order to give the subjects some hope of making money earlier, to try to prevent further active or passive lassitude. I maintained this new order in all subsequent experiments, and no other students asked to leave.

[6] The instructions, like the EPNES software, were designed to be applicable over a wide range of electricity market experiments with minimum modification. Thus, they were longer than strictly necessary for this

Each practice experiment involved the experimenter reading the instructions, and asking the subjects to complete quizzes during the reading. Some of the quizzes required the use of computer software, the same software used in the real experiment, to complete successfully. I checked each subject's answers and explained any mistakes they made. Following the instructions, I asked the subjects to participate in multiple periods of bidding, on two different networks. Subjects were paid a $15 show-up fee, plus up to another $25 possible from profits bidding in the two networks.

During the real experiments, I kept a record on each subject, recording which subjects they had previously been grouped with and which firm they had represented. I allowed and encouraged subjects to attend multiple experiments. However, where possible, I did not let subjects repeat the experiment with any subjects from their previous group(s), nor did I let them represent the same firm twice. My purpose was to weight the experiment yet again against collusion, including the possibility that subjects were talking to each other outside the laboratory. In four days of experiments, there were only three cases where a pair of subjects turned up in the same group twice, and only two cases where a subject played the same firm twice.

In the experiment, I ran markets in each of the four networks consecutively. Before each market opened, I gave all subjects a diagram showing the current network, then allowed them five minutes to use the practice tool before opening the real market. Each market was divided into periods, each period being one bidding round. On average, I allowed about 40 periods per network, which took about 25 minutes to complete. Subjects were given 40 seconds per period to place a bid, which had to be between 0 and 30 francs per unit.

At the end of each market, each subject's profit in francs was displayed on the computer screen. I asked subjects to record this profit on a payoff sheet. At the end of the experiment, I had subjects total up profits from the four markets and divide by the exchange rate to find their net earnings in dollar, with partial amounts rounded up to the nearest dollar. This amount plus the show-up fee of $10 was the subject's total earnings from the experiment. The exchange rates were set at 5000 francs to the dollar for subjects playing firms 0 and 2, and 4600 francs to the dollar for subjects playing firms 1 and 3. The difference was to due to the fact that the former had a chance to be the monopolist in the low demand rounds,

particular experiment. See Appendix B for the full instructions.

when the other firms could expect to be earning very little. Thus ex ante I expected these firms to make more francs than firms 1 and 3, who played the monopolist in one of the high demand rounds.

4.4 Results

To answer the original question, as to whether subjects could understand and incorporate network constraints into their economic decision-making, I analysed the behaviour of the monopolists. Although the monopolists' behaviour is superficially straightforward, my analysis was complicated by two factors of the experiment: random demand and the sealed simultaneous bid format. Thus to discern whether the monopolist understood the constraints I chose a comparative approach, comparing the monopolists' behaviour with that of the other firms.

If the other firms' bids and demand were known, the monopolist would have a simple strategy. Bid high if all the other firms bid sufficiently low, bid low if at least one of the other firms (or at least two firms in the low demand case) bid high enough. The exact cut-off value would depend on the exact value of demand. In reality, the monopolist must incorporate such uncertainty into her bidding strategy, so her responses will not be as clear-cut as the full information case. Furthermore, the monopolist's optimal response will differ when the group is tacitly colluding to keep prices high, as some groups did. Such issues conspired to prevent me from directly analysing whether the monopolist was pursuing a profit-maximising strategy.

By design however, each of the four networks I presented to subjects had one monopolist and three *other* firms. The latter were identical economically speaking. Since the monopolist was in a load pocket, she had significant market power. As a result, I expected to observe the monopolist obtaining both higher prices and higher profits than the other firms, if she understood the nature of the constraint. If she didn't, there was no reason for her to behave any differently from any of the other firms.

Prices and profits are both variables I can measure directly from the data. Profit I found to be a poor comparative measure, because even if the monopolist behaved exactly like one of the other firms, she would still make a higher profit, as whenever she made the highest bid, she would receive an allocation greater than would have been received by another firm. Thus I use price as my measure of comparison. As noted above, the monopolist can sustain on average a considerably higher price than the other firms, especially in the low demand case.

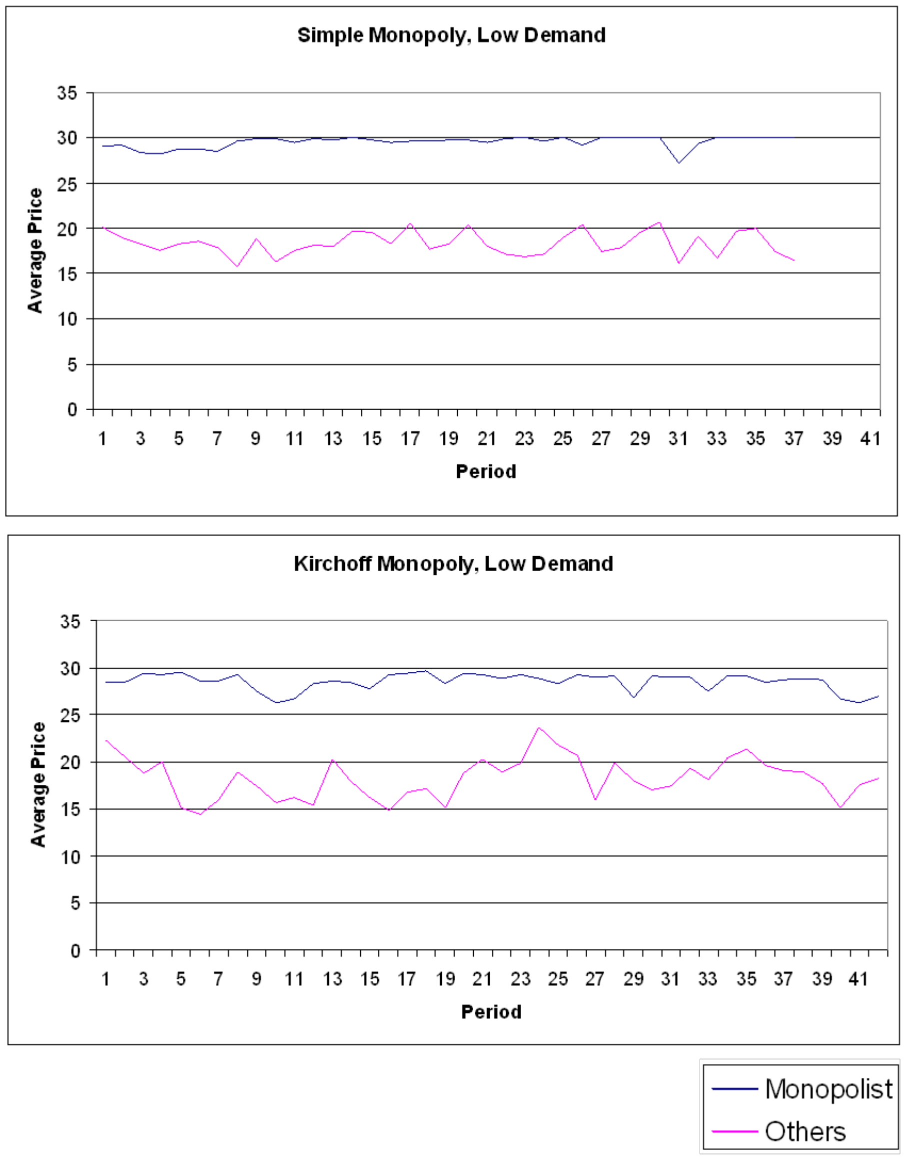


Figure 4.1: Average price received by firms in low demand networks

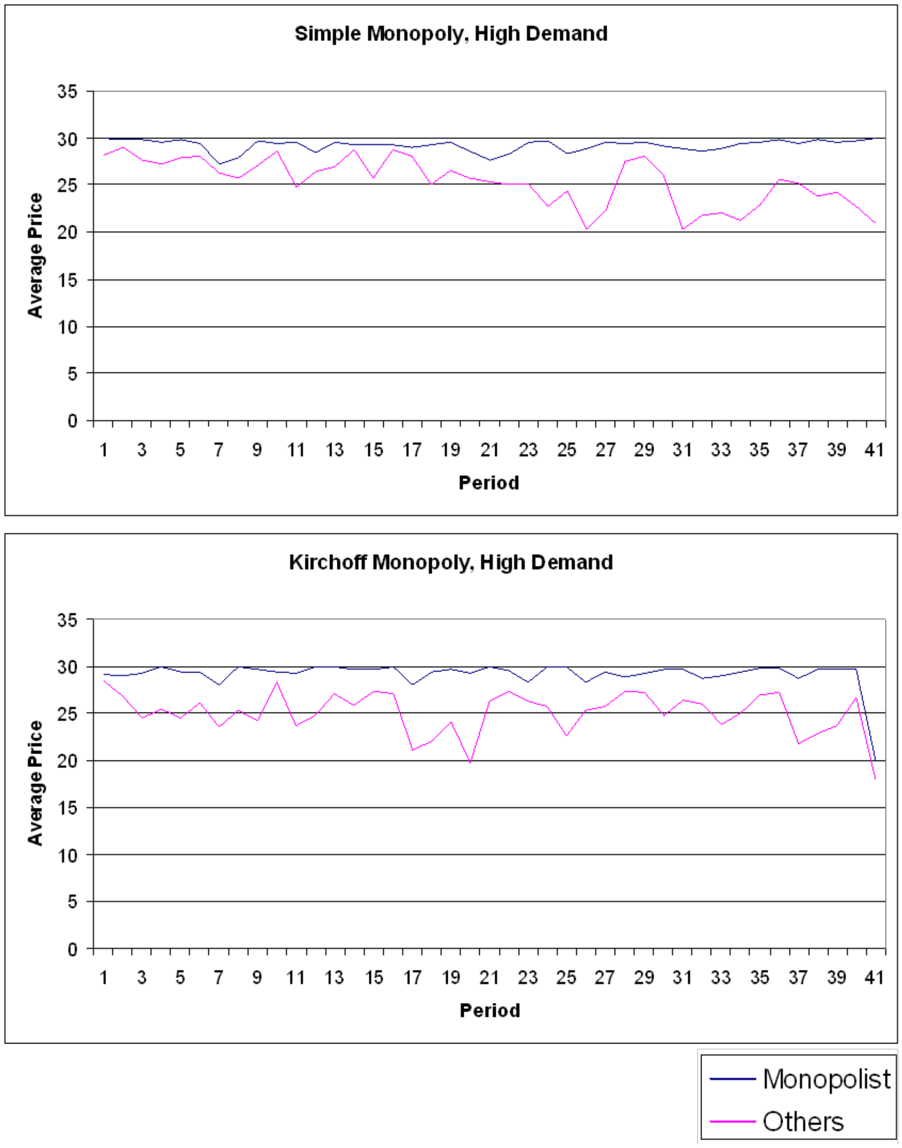


Figure 4.2: Average price received by firms in high demand networks

Figures 4.1 and 4.2 show an enormous difference between the average prices received by the monopolist, and those received by the other firms. In the more competitive low demand case, the difference was extreme. The monopolist on average rarely received a price of less than 27 francs, while the other firms struggled to receive more than 20 francs per unit. Even in the high demand case the monopolist sustained consistently higher prices on average. Thus there was a sustained difference between the behaviour of the monopolist, and that of the other firms. Since the only difference between the two was the existence of the constraint that created the monopolist's load pocket, I conclude that the monopolists, and thus subjects in general, can quickly comprehend the existence of a load pocket and understand how it affects the prices they will receive.[7]

The more interesting question I designed this experiment to answer was if I could identify any difference in subject behaviour between the simple constraint and the more complicated Kirchoff constraint. I hypothesised, based on pilot experiments with inexperienced subjects, that subjects would take longer to take advantage of the Kirchoff constraint, if they identified it at all.

Once again, I compared the effect of the two constraints by analysing the monopolists' behaviour, now comparing the Kirchoff constraint monopolists to the simple constraint monopolists. I found that prices in this case were not a good measure. All the monopolists in both simple and Kirchoff cases maintained average prices in the 27 to 30 franc range. Such figures are potentially consistent with profit-maximising behaviour. Prices consistently below 27 francs would have signalled a lack of understanding, but I did not observe this. Thus to answer this question, I looked instead at monopolists' profit. First I controlled for random demand. Even if all firms make the same bids every period, the monopolist's profit would still vary due to random demand. Thus I recalculated every monopolist's profit by dividing actual profit by the actual realisation of demand. I call the resulting measure 'Profit per available unit'. In figure 4.3 on the next page I present graphs of the monopolist's average profit per available unit for the high and low demand scenarios across the twelve groups.

Figure 4.3 makes clear that there is no significant difference in average profits between the simple and Kirchoff constraints. Although there is a lot of variation, there are no trends that might suggest monopolists earned lower profits under one constraint that the other. The variation that exists can easily be explained by the fact that monopolists' profits depend significantly upon the bids of the other players. I conclude that monopolists, and thus experimental subjects in general, had no more trouble recognising a Kirchoff constraint than they did a simple constraint.

[7] Anecdotally, the subjects I interviewed after the last experiments were well aware of the existence of the monopolist. They called it "the sweet spot" where you could get plenty of money even in a "bad" (low demand) situation.

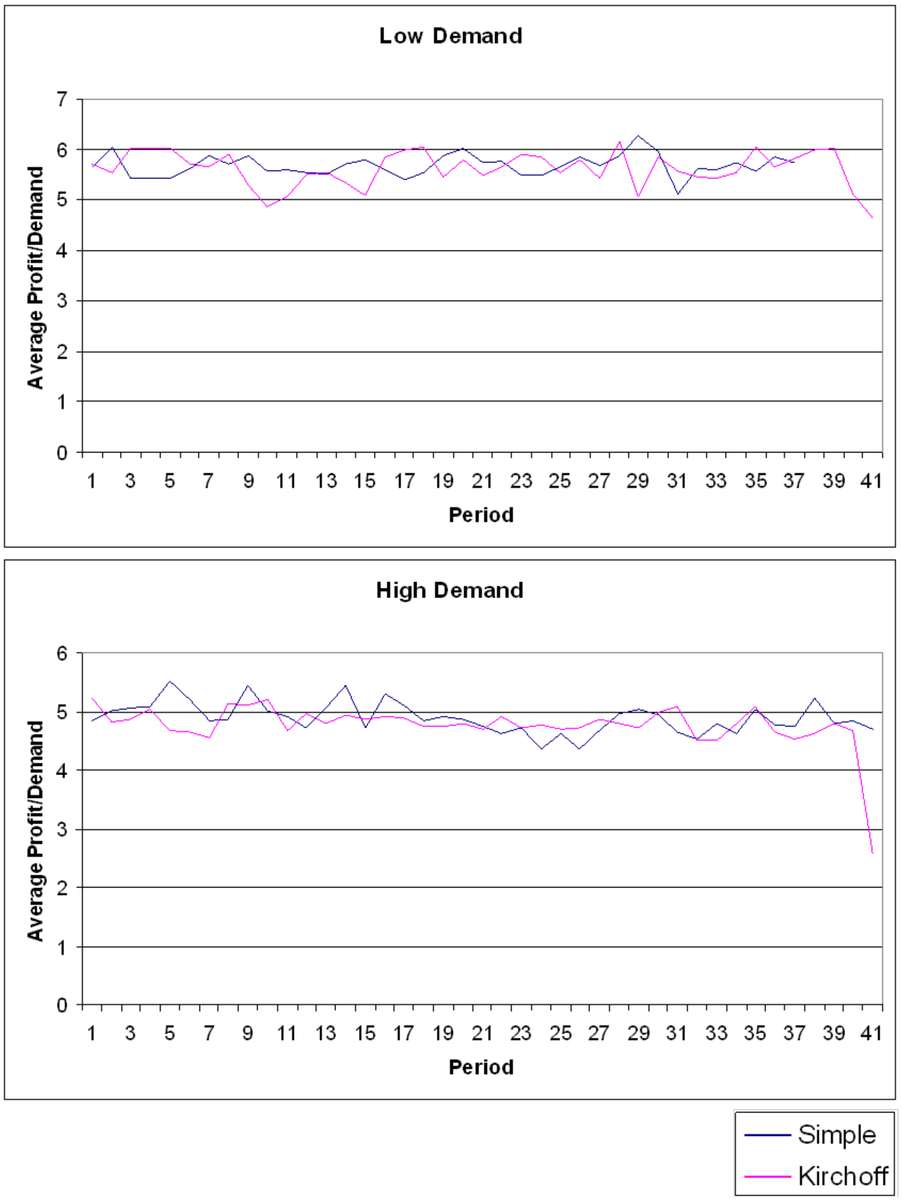


Figure 4.3: Monopolists' profit per available unit

4.5 Conclusion

If economic experiments on electricity markets are to provide results applicable to the real world design of such markets, they will need to attain a level of complexity greater than has previously been attempted. This paper shows that experimental subjects are perfectly capable of rational profit-maximising behaviour in more complicated networks where physical constraints including Kirchoff's laws play a key role in determining economic outcomes. The data is very clear. Presented with two networks that were theoretically identical, subjects behaved almost identically on both networks. They quickly identified the constraint and used it to their advantage. These results should spur researchers to experiment on larger networks, testing different market mechanisms with hope that the results will be applicable to real world markets.

This research leaves a couple of open questions. Firstly, although I went to some trouble to write detailed instructions explaining such constraints, it is likely that some time spent with the practice tool on the given network would be more than sufficient. The presence of such a constraint can be inferred by the results of trying various high and low bids, without any knowledge of how the network works at all. I designed my instructions for even more complicated experiments, and did not test if understanding would change with fewer or no explicit instructions.

The second question arises from the data produced by this experiment. The more experienced groups, groups containing subjects who had done the experiment several times, showed significant signs of collusion, such as all firms bidding 30 francs each period for instance, behaviour which contradicted standard theory for one-shot games. This was despite my weighting the experiments against collusion (see section 4.2). It seems possible that subjects were playing the Nash equilibrium of a repeated game, a result which would have important implications for the power exchange market. I could not test this theory adequately without obtaining more data with more experienced subjects.

4.6 Appendix A: Network Diagrams

These diagrams are the same ones seen by the experimental subjects. Each network consists of eight nodes, labelled from 0 to 7. Some nodes are depicted with an arrow pointing in.

These nodes represent firms, with the capacity of the firm in MW listed next to the arrow. The other nodes (with an arrow pointing out) represent demand nodes, with the numbers in square brackets being the minimum and maximum possible demand at that node. Nodes are connected by transmission lines, with each line's maximum capacity in MW labelled next to the line.[8]

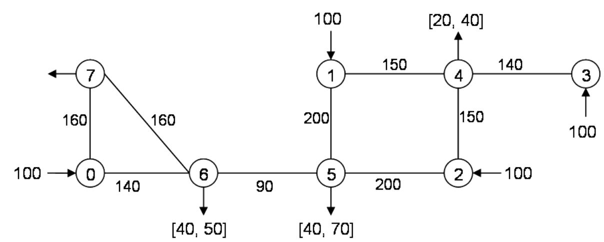


Figure 4.4: Network A (Simple constraint, low demand)

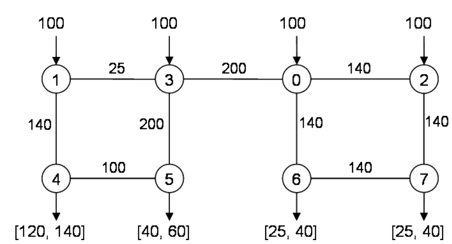


Figure 4.5: Network B (Kirchoff constraint, low demand)

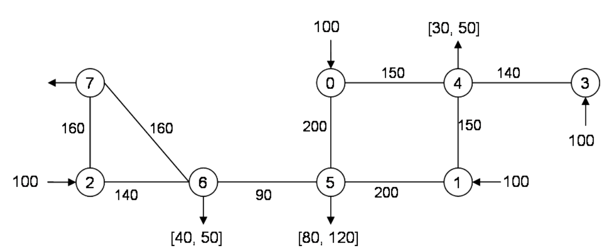


Figure 4.6: Network C (Simple constraint, high demand)

[8]A more detailed description of the networks was given in the instructions, see Appendix B.

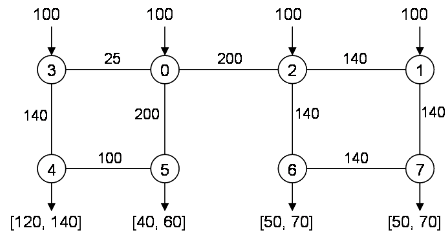


Figure 4.7: Network D (Kirchoff constraint, high demand)

4.7 Appendix B: Instructions

Introduction

This is an experiment in the economics of market decision-making.[9] In this experiment we are going to simulate an electricity market. You will be buying and selling electricity at different points on an electricity network, in order to make money that will be paid to you in cash at the end of the experiment. This experiment should take about two and a half hours.

Today we will be using a simplified model of an electricity network. There will be unusual rules for buying and selling on the network, and for determining the prices paid. These instructions are divided into three sections. The first explains how the network works, the second explains how to buy and sell, and the third explains how prices are determined. At the end of each section, I will test you on your knowledge. The tests are not difficult, and after you complete all three you should have a good idea of how to buy and sell electricity on the network, while making money doing so.

This experiment will be divided into rounds. Each round may be conducted on a different electricity network. Each round is divided into periods. At the beginning of each round, you will be asked to login to the new round. Your login details are written on a sheet in the folder in front of you. When each round finishes, your profit (loss) from the round will be displayed on the computer screen in front of you.

The currency in these markets is francs. At the end of the experiment, your total earnings of francs will be converted to dollars at a rate that will be given to you when the actual

[9]Note that the instructions were accompanied by a slideshow giving screen shots with step by step guides as to how to make bids, and how to interpret the output in the experimental interface.

experiment starts. Please do not show this information or information on your profits to other experiment participants.

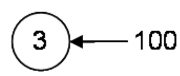
Network Instructions

Here I will explain how a network is modelled. Please pay attention, as your ability to rapidly evaluate any network diagram is important for making money in this experiment.

An electricity network consists of three components.

Supply Nodes

A supply node is a node – a point on the network - where a firm who can produce electricity is located. In the diagrams I will give you, a supply node looks like:

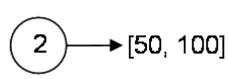


The number inside the circle is the 'name' of the node. This is node 3. Each node has a unique number. When I refer to this firm, I would say 'Firm 3', meaning the firm located at node 3. The arrow pointing in to the circle means that this node is a supply node (it is making electricity which is flowing into the network). The number next to the arrow tells you the capacity of the firm. The capacity of a firm is the maximum units of electricity this firm can produce per period. This firm can produce 100 units per period.

In this experiment, each of you will control one firm on the network.

Demand Nodes

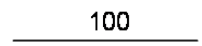
A demand node is a node where a consumer - someone who uses electricity - is located. You can think of a demand node as a city, filled with people who are using electricity. In the diagrams, demand nodes look like this:



This is node '2'. The arrow pointing out of the node indicates that this is a demand node. This means the node is taking electricity out of the network. Demand at demand nodes is random. The numbers inside the brackets indicate the possible range of demand in one period for this node. In each period, we are not sure what the true value of demand is, but we do know it must be between these minimum and maximum values. So in this example demand node 2 will always demand at least 50 units of electricity per period, but no more than 100 units of electricity per period.

Transmission Lines

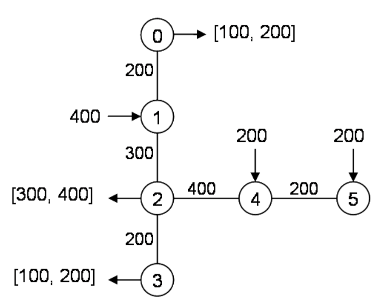
Finally, demand nodes and supply nodes are joined together by transmission lines. A line looks like this:



Electricity flows along these lines, between supply nodes and demand nodes. The number above the line gives the maximum (per period) capacity of the line. This line can only transport up to 100 units of electricity in one period. Please note that as this a simplified model of electricity flow, there are no transmission line losses in this model. No electricity is lost when flowing from one node to the next. Also you should note that Kirchoff's laws do hold in this model. We do not expect you to know Kirchoff's laws, or be able to calculate them. Instead we will give you a computer tool that will work out flows for you. If you want more information, I have included a note about Kirchoff's laws at the end of these instructions to give you some intuition.

Electricity Network

Supply nodes, demand nodes and transmission lines are joined together to form an electricity network. For example:



There is one more important piece of information you need to know about the network, and that is how much it costs each firm to produce electricity. Each firm has a fixed per unit ***cost*** that it must pay every time it makes one unit of electricity. If the cost is 12, and the firm makes 20 units, it must pay $12 \times 20 = 240$ francs to produce the electricity. If it makes 55 units, it must pay $12 \times 55 = 660$ francs to produce the electricity. The ***cost*** of each firm is shown to you in a table next to the diagram. For example, the cost table in the above example might look like:

Costs	1	12
	4	10.5
	5	14

Notice that the firms in the example above are located at nodes 1, 4 and 5. Firm 1 must pay 12 francs for *each* unit of electricity it makes. Firm 4 must pay 10.5 francs per unit, and firm 5 must pay 14 francs per unit.

Every period, every firm will get a chance to sell their electricity via an auction process that I will describe later. After the auction, each firm is told how much electricity to make, and what ***price*** per unit of electricity they will get paid. Firms make profits when the ***price*** they get paid for their electricity is greater than the ***cost*** they pay to make the electricity. Profit is given by $q \times (p - c)$, where p is the ***price*** you are paid, c is the ***cost*** you pay to make the electricity, and q is the amount you are asked to make.

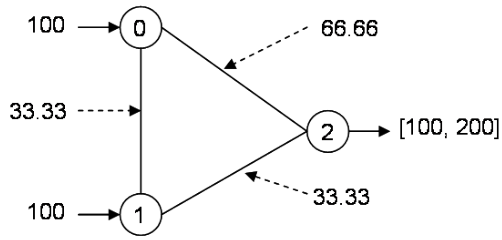
During the experiment, you will see a number of different networks. You will control one firm in each network, but the identity (node number) of your firm may change. For each Round, the firm you control is written on your login sheet. For example:

ELECTRICITY BIDDING INSTRUCTIONS PRACTICE ROUND		
Round	PracticeX	Profit
Player	2 ← **THIS IS YOUR FIRM**	
Password	xxxx	

If you have any questions, please raise your hand now.

Some Notes on Kirchoff's Laws

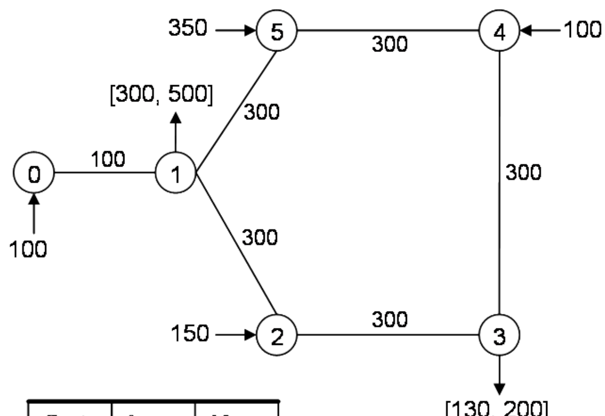
Suppose there is a firm at node 0 and it wants to send 100 units of electricity to a consumer at node 2. Ignore firm 1 – pretend it produces nothing. Then electricity must flow from 0→2, but also down the lines 0→1 and 1→2. Because the flow from 0→1 and 1→2 is twice as long as 0→2 directly, Kirchoff's laws say the flow of electricity is only half as much. In this case, the real flows are:



In general if there are two lines that electricity can take to get from a supply node to a demand node, than the amount of electricity that flows on each path is inversely proportional to the length of the path.

Bidding Instructions

Here I will explain the rules of the market - the process by which you, the firms, can sell your electricity. I will give you examples and a chance to practice as we go through these instructions. All examples will be based upon the following network.



Costs	0	12
	2	11
	4	12
	5	10.5

Each firm on the network is controlled by one person in this room. Remember that demand at each demand node is random between a lower and upper limit in each period. Firms do NOT sell their electricity directly to demand nodes. Instead they sell their electricity to a third entity called the Market Maker. Every period, the Market Maker buys electricity from the firms and sells it to the demand nodes. Because demand is random, and not observable until the *end* of the period, the Market Maker asks firms to make ***bids***. A ***bid*** is the minimum amount the firm wants to be paid before it is willing to produce electricity. After demand is revealed, the Market Maker will then buy enough electricity from the firms to satisfy all demand, according to a specific formula.

1) A period starts.

2) Every firm enters a ***bid***. You cannot see the other firms' bids until the end of the period. Your ***bid*** has to be between 0 and 30 francs, and can have up to 3 decimal places. You will have 40 seconds in which to choose your ***bid***. You may change your ***bid*** while the bidding period is open. Once you are happy with your ***bid***, press the confirm button. The bidding period will end when every firm presses the confirm button, or when 40 seconds is up, whichever comes first. (If you don't make a ***bid*** after 40 seconds the computer will submit the bid from your last period. If it is the first period, the computer will choose a random number between 0 and 30 for you.)

3) The true value of demand at each demand node is realised. This means the computer generates a random number at each demand node.

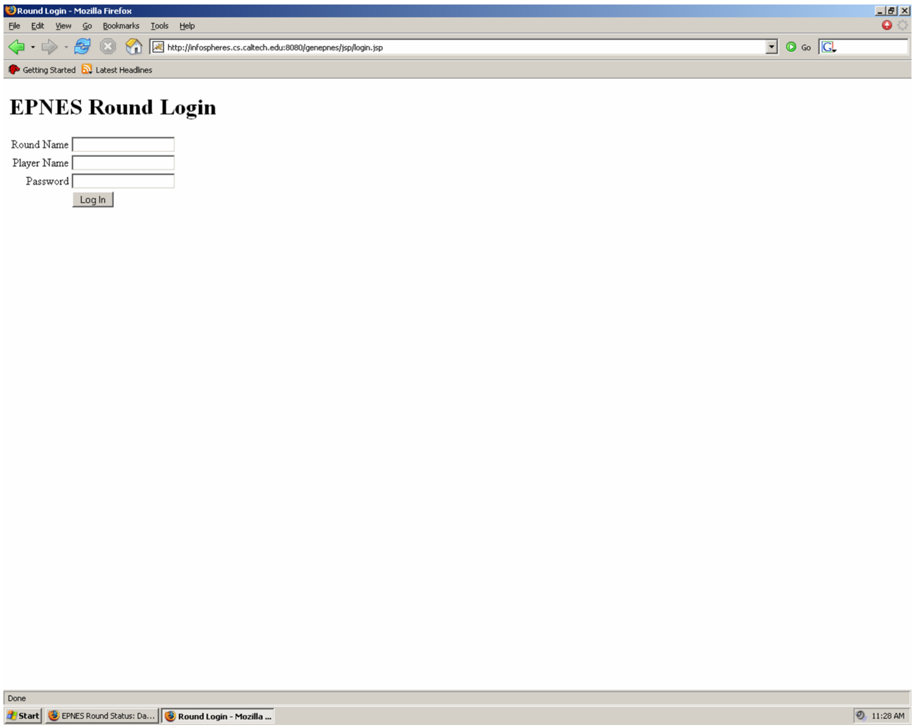
4) The Market Maker takes the firms' bids and the true value of demand. It must buy enough electricity from the firms equal to the total amount of electricity demanded by all the demand nodes. It decides how much electricity each firm will produce (the '*allocation*') and what ***price*** each firm will be paid, according to a formula that we will discuss later.

5) *The* ***price*** *each firm gets will always be greater than the* ***bid*** *that firm made!*

6) The period ends. Each firm makes a profit (or loss) equal to $q \times (p - c)$, where q is the number of units you are asked to make, p is the ***price*** you get, and c is the ***cost*** you pay per unit for making electricity. You make money when the ***price*** you get is higher than the ***cost*** you pay. Remember that the ***price*** you get is always higher than the ***bid*** you make, so if your ***bid*** is higher than your ***cost***, you cannot lose money.

7) Now each firm can see all the other firms' bids and prices, the true value of demand at each demand node, and the actual amount of electricity produced by each firm.

We will now make some bids to get comfortable with the software. You should be looking at the login screen, which looks like this:



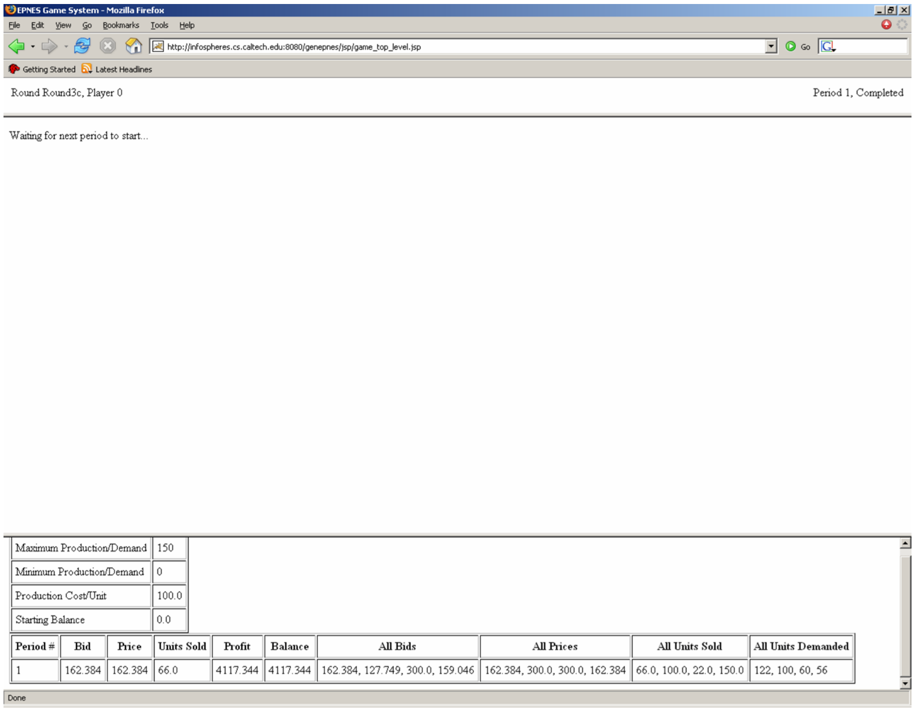
In your folder, pull out the sheet with the login details. Please log into the round labelled "Electricity Bidding Instructions Practice" at the top of your sheet, using the Round name, password and player number written in the box. Remember your player number is the number of the firm you will control for this round.

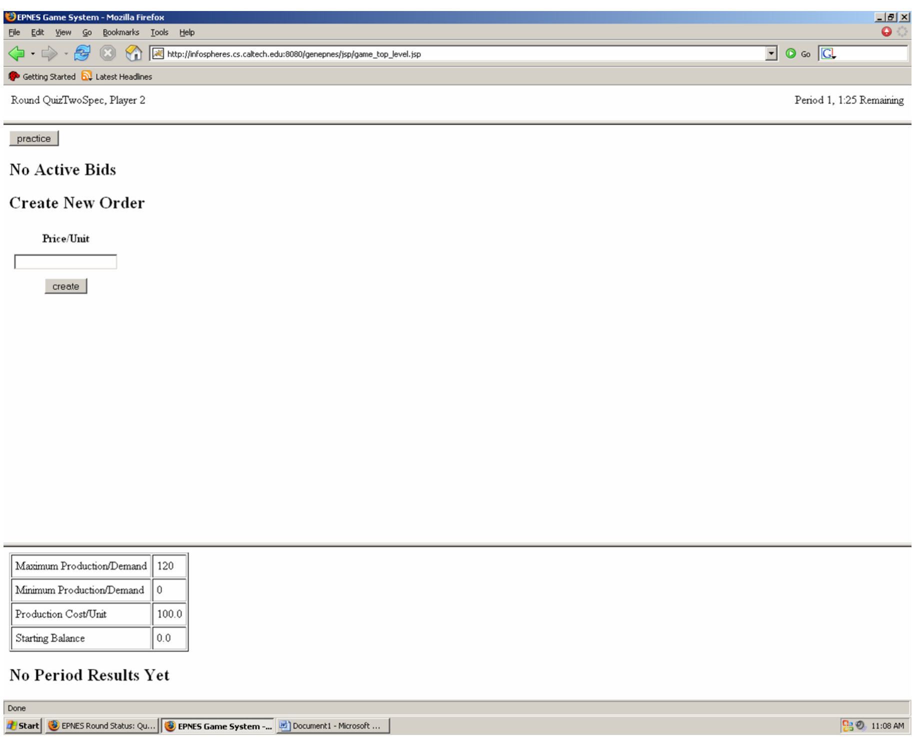
If you have logged in correctly, you should see "Waiting for next period to start" written on the screen, similar to the screenshot at the top of the next page. If you do not see this, please raise your hand.

Once everyone is logged in, we will start the period. Please do not touch the keyboard or mouse yet. You are now looking at the 'Bidding screen'. It should look like the second screenshot on the next page.

First look at the top left corner of the screen. This tells you the Round you are in and the Player number you are. The Round tells you which network you are playing on and the Player number tells you which firm (node) you are at.

Now look at the top right corner of the screen. The timer there tells you how much time is


EPNES Game System - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://infospheres.cs.caltech.edu:8080/genepnes/jsp/game_top_level.jsp
Getting Started Latest Headlines
Round Round3c, Player 0
Period 1, Completed
Waiting for next period to start...
Maximum Production/Demand 150
Minimum Production/Demand 0
Production Cost/Unit 100.0
Starting Balance 0.0
Period # | Bid | Price | Units Sold | Profit | Balance | All Bids | All Prices | All Units Sold | All Units Demanded
1 | 162.384 | 162.384 | 66.0 | 4117.344 | 4117.344 | 162.384, 127.749, 300.0, 159.046 | 162.384, 300.0, 300.0, 162.384 | 66.0, 100.0, 22.0, 150.0 | 122, 100, 60, 56
Done


EPNES Game System - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://infospheres.cs.caltech.edu:8080/genepnes/jsp/game_top_level.jsp
Getting Started Latest Headlines
Round QuizTwoSpec, Player 2
Period 1, 1:25 Remaining
practice
No Active Bids
Create New Order
Price/Unit
create
Maximum Production/Demand 120
Minimum Production/Demand 0
Production Cost/Unit 100.0
Starting Balance 0.0
No Period Results Yet
Done
Start
EPNES Round Status: Qu...
EPNES Game System -...
Document1 - Microsoft ...
11:08 AM

left in the period. In this practice round you are allowed 15 minutes per period in which to decide on your bid. In the real experiment you will only have 40 seconds.

Now notice the 'Practice' button on the left upper side of the screen. This takes you to a practice area where you can enter pretend bids and see what kind of price and allocation you would get. We will use this tool more later on.

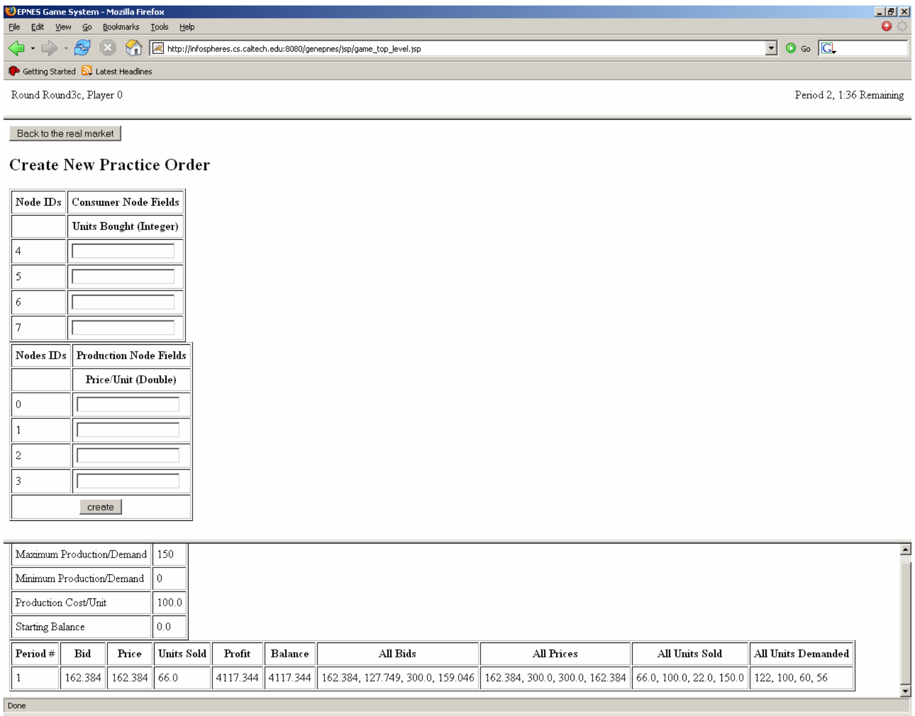
In the middle of your screen is a box where you can make your ***bid***. Please make a ***bid*** now. Press enter or 'create' when you are done. Then press the confirm button. When you have done this you will see a notice "Waiting for all players to finish submitting bids". When everyone has finished submitting a bid, the screen will change and say "Waiting for next period to start".

While you are waiting, take a look at the box at the bottom of the screen. There is a lot of valuable information in this box that tells you how much money you are making, and how the other firms are bidding. Each period this information will be updated, so that you will have a complete history of the market during the round. From left to right, you can see the following. The first box is the period number. The next boxes are the ***bid*** you made, the ***price*** you received, the number of units you were asked to make, your profit in that period and your balance (total profit) in the round up to that period. The following box shows all the bids made by all the firms. This box is always sorted by increasing firm number, from lowest to highest. The firm with the lowest number made the bid given in the first entry of this list and so forth. The boxes containing 'All Prices' and 'All units sold' information are similarly sorted. It is important that you can easily see which firms bid what and how much they made. Finally the last box shows you how many units were demanded in reality at each of the demand nodes for that particular period. Once again, these are sorted by increasing order by node number.

Notice that the other firms' bids will affect the ***price*** you receive! (We'll discuss how later.) Now a new period will start. Try making another ***bid***. Press confirm. When the period ends, notice how the record of each period accumulates in the bottom box. The most recent period's results is at the top of the list.

The Practice Tool

The Practice Tool is an important tool for helping you make good bids. Here you get a chance to enter fake bids for *all* firms and see what kind of profit you might make. You can access it by pressing the 'Practice' button on the main bidding screen. Press the Practice button now. This will take you to a screen that looks like this:



Remember what the network looks like in this example (at the top of these instructions). First enter some values of demand (called 'Consumer Node Fields' on the screen). Your choices of demand have to be within the minimum and maximum bounds on the network diagram. Now enter a fake bid for each firm. Press enter and the computer will tell you the results. Take a look at the results. It is important you can comprehend them quickly. The prices, allocations and profits for each firm are given in a table listed by firm number. The line flows are given in a table below that. Remember that a positive line flow in one direction can also be shown as a negative line flow in the other direction. If you want to make another practice bid, scroll down and enter new information.

Now we are going to do a second test. This will test your ability to use the practice tool and understand the output it gives you. If you have any questions, please raise your hand now.

Market Instructions

In this section I will explain how the Market Maker determines prices and allocations given the bids firms make. There is a two step process to determine prices. The *allocation* (how much electricity each firm makes) is determined first, then the prices are determined in a second step. Remember that the Market Maker is just a computer.

Step 1

The Market Maker takes all of the firms' bids and the actual values of demand and runs an optimisation algorithm. The goal is to minimise the cost to the Market Maker of purchasing enough electricity to satisfy demand assuming that the Market Maker will pay each firm what it bid per unit of electricity. This algorithm finds an optimal solution and returns the amount of electricity each firm should make. This is the optimal 'allocation'. The Market Maker also works out the total cost it has to pay for this allocation. We will call this cost **C**.

Step 2

To determine the actual price each firm gets paid, the computer uses the following process. It takes a firm and pretends that that firm has an extra unit of electricity available at 0 cost. So, if a firm bid

Bid	Units Offered
24	100

This would be the firm's capacity

the computer will pretend that the firm bid

Bid	Units Offered
0	1
24	101

Now the computer repeats Step 1 with this firm making this new bid, and all other firms making their original bids. It finds the new total cost D it has to pay now. Note that the new cost **D** must be less than or equal to the original cost **C** because now this firm is offering a free unit! The difference **D** – **C** is the amount this firm gets paid.

This process is repeated for each firm to find how much each firm gets paid. If a firm was allocated nothing in Step 1, then the price for that firm is set equal to 0.

Note that if a firm is given a free unit, the computer will always take this unit which means that another firm will lose one unit (because demand must equal supply). Therefore the firm will usually (but now always) get a price equal to the *bid of the firm who lost that unit.* This will usually be the firm with the highest bid, but may not be if line constraints prevent this.

I will now go through several examples to help you understand the pricing rule.

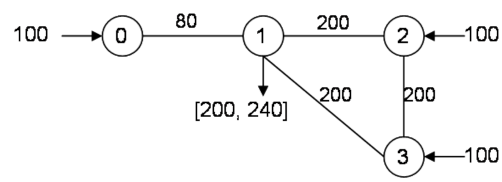
One important consequence of this algorithm is that if none of the line constraints in the network are binding, then all firms must receive the same ***price*** equal to the highest ***bid*** of those firms who are producing electricity. This is because if you give any firm a free unit, then the computer will automatically take one unit away from the firm with the highest bid. The difference in the total cost is then equal to that firm's bid.

Let's look at an example. Suppose we have four firms, each with capacity of 100 units of electricity, in an unconstrained network (none of the line constraints will ever be binding). Suppose the bids are as follows.

Firm	Bid
1	22
2	11.5
3	22.4
4	29.9

If total demand was 220, the ***price*** would be 22.4 francs for all firms, except firm 4. To compute this, first work out the allocation. The minimum cost is to have firm 2 produce 100, then firm 1 produce 100 and finally firm 3 produce 20. This is the optimal allocation. Now work out prices. If you give any firm one free unit, what will the Market Maker do? The Market Maker can minimize total cost by taking away one unit from firm 3, because firm 3's bid was the highest amongst the those firms with positive allocation. Notice that firm 3 must get a ***price*** of at least 22.4 (this is what it ***bid***). Therefore firms 1 and 2 must also get a ***price*** of at least 22.4. As firm 4 is not producing anything, it gets a ***price*** of 0.

Now let's look at a trickier example to see when prices might be different amongst firms. Consider the following network.



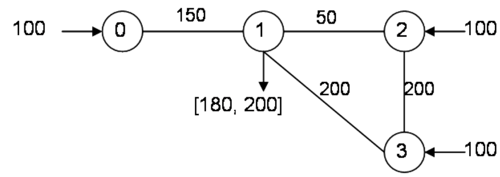
Suppose the true value of demand is 240, and suppose the bids for the three firms are as follows.

Firm	Bid
0	11.2
2	17.8
3	20.1

As normal, the allocation is decided first. Firm 0 has the lowest ***bid***, so it gets to supply 100. Unfortunately, a capacity constraint on the line restricts it to 80 units. Firm 2 has the second lowest ***bid***, it can supply 100 units. Finally firm 3 supplies the remaining 60 units. Firm 3 has the highest ***bid*** out of the firms that are given a positive allocation, so firm 3 must receive a ***price*** of 20.1, equal to its bid.

Notice what happens if you give firm 0 an extra free unit. Which firm loses a unit? It must be firm 1, because there is a line constraint of 80 between node 0 and node 1 (the demand node). If firm 0 gets an extra free unit, it is now producing 81 units, which violates the line constraint. Therefore, the only way to make the solution feasible is to take one unit away from firm 0 – this unit would cost 11.2, so **D** – **C** equals 11.2. There are no line constraints between firms 2 and 3, and node 1 however. So if you give either firm 2 or firm 3 a free unit, the computer will take 1 unit off firm 3 (the firm with the highest bid) so the price for firms 2 and 3 is 20.1.

Now here is a really interesting example. Consider the modified network below. Suppose again that demand is 200 and suppose that the bids are as given.



Firm	Bid
0	14
2	12
3	16

First we will work out the allocation. Kirchoff's laws make the allocation complicated at this point so during the experiment you should rely on the practice tool to compute allocations. But to give you some intuition, I will indicate how it works.

Firm 2 gets first allocation because it has the lowest bid. The electricity firm 2 produces must flow to node 1, because that is the only demand node. By Kirchoff's laws, two-thirds of the electricity it makes will flow on the line 2→1, and the other one-third will flow from 2→3 and 3→1. So firm 2 will initially get allocated 75 units. Next firm 0 gets an allocation, and can produce 100 units. This leaves firm 3. Firm 3 can produce nothing – but now notice that if firm 2 produces 50, then firm 3 can also produce 50. As the Market Maker needs 200 units total, it will do this, so the final allocation is firm 0 gets 100, and firms 2 and 3 get 50. Now look at the prices.

If firm 0 can produce 1 more unit (for free), which firm loses an extra unit? Firm 3 is the obvious pick, but notice that if firm 3 loses an extra unit, then firm 2 can make an extra one-third of a unit but then firm 3 should lose this one-third of a unit (because it is more expensive) and so forth until we find the new optimal solution is to give firm 0 101 units, firm 2 gets 51 units and firm 3 gets 48 units. So take off 16×2 and add 12 to find the new total cost which is 20 francs less! So the price for firm 0 is 20. This price is *not* equal to any of the firms' bids! For firms 2 and 3, because the line 2→1 is constrained, the free unit replaces one of their units, so their price is what they bid.

The moral of the story is that differences in price are caused when transmission lines are at full capacity at the optimal allocation. Kirchoff's laws can make it tricky to identify exactly what the price should be so you need to rely on the practice tool to help you guess prices and profits in any given situation.

In event of a tie in bids, the computer will try and give each firm an allocation that is proportional to their capacity. If constraints prevent this, the computer will try to find an allocation as close to this as possible.

We will now do a final quiz to gain some intuition on how prices are calculated. If you have any questions, please raise your hand now.

Bibliography

[1] S. Backerman, S. Rassenti, and H. Smith. Efficiency and Income Shares in High-Demand Energy Networks: Who Receives the Congestion Rents When a Line is Constrained? *Pacific Economic Review*, 5(3):331–347, 2000.

[2] J. Carlson and R. McAfee. Discrete Equilibrium Price Dispersion. *Journal of Political Economy*, 91(3):480–493, 1983.

[3] H. Chao and S. Peck. A Market Mechanism for Electric Power Transmission. *Journal of Regulatory Economics*, 10:25–59, 1996.

[4] H. Chao and C. Plott. A Competitive Mechanism for the Distribution of Electric Power through a Network and the Determination of Network Capacity: Experimental Tests and Performance. Working paper presented at the California Public Utility Commission ORA Seminar, San Francisco, Feb. 7, 2000.

[5] C. Davidson and R. Deneckere. Long-Run Competition in Capacity, Short-Run Competition in Price, and the Cournot Model. *The RAND Journal of Economics*, 17(3):404–415, 1986.

[6] E. Dechenaux and D. Kovenock. Tacit Collusion and Capacity Withholding in Repeated Uniform Price Auctions. *The RAND Journal of Economics*, 38(4):1044–1069, 2007.

[7] N. Fabra, N. von der Fehr, and D. Harbord. Designing Electricity Auctions. Centre for the Study of Energy Markets Working Paper #122, 2004.

[8] J. Friedman. On the Strategic Importance of Prices Versus Quantities. *The RAND Journal of Economics*, 19(4):607–622, 1988.

[9] D. Gaskins and J. Haring. A Note on Unilateral Withholding. *Land Economics*, 55(1):135–140, 1979.

[10] V. Grimm and G. Zoettl. Capacity Choice under Uncertainty: The Impact of Market Structure. University of Cologne Working Paper Series, No. 23, 2006.

[11] C. Hogendorn. Tacit Collusion in Capacity Investment: The Role of Capacity Exchanges. *The B. E. Journal of Theoretical Economics*, 7(1):1–14, 2007.

[12] D. Kreps and J. Scheinkman. Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *Bell Journal of Economics*, 14(2):326–337, 1983.

[13] G. Maggi. Strategic Trade Policies with Endogenous Mode of Competition. *American Economic Review*, 86(1):237–258, 1996.

[14] S. Reynolds and B. Wilson. Bertrand-Edgeworth Competition, Demand Uncertainty, and Asymmetric Outcomes. *Journal of Economic Theory*, 92(1):122–141, 2000.

[15] N. Singh and X. Vives. Price and Quantity Competition in a Differentiated Duopoly. *The RAND Journal of Economics*, 15(4):546–554, 1984.

[16] S. Stoft. *Power System Economics: Designing Markets for Electricity.* Wiley-IEEE Press, 2002.

[17] J. Tirole. *The Theory of Industrial Organization.* MIT Press, 1988.

[18] D. Topkis. *Supermodularity and Complementarity.* Princeton University Press, 1998.

[19] X. Vives. On the Efficiency of Bertrand and Cournot Equilibria with Product Differentiation. *Journal of Economic Theory*, 36(1):166–175, 1985.

[20] X. Vives. *Oligopoly Pricing: Old Ideas and New Tools.* MIT Press, 1999.

[21] F. Wu, P. Variaya, P. Spiller, and S. Oren. Folk Theorems on Transmission Access: Proofs and Counterexamples. *Journal of Regulatory Economics*, 10(1):5–23, 1996.