PWP-040r

The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry

Severin Borenstein, James Bushnell, and Steven Stoft

October 1999

This paper is part of the working papers series of the Program on Workable Energy Regulation (POWER). POWER is a program of the University of California Energy Institute, a multicampus research unit of the University of California, located on the Berkeley campus.

University of California Energy Institute
2539 Channing Way
Berkeley, California 94720-5180
www.ucei.berkeley.edu/ucei

Forthcoming in *RAND Journal of Economics*

The competitive effects of transmission capacity in a deregulated electricity industry

Severin Borenstein*
James Bushnell**
and Steven Stoft***

In an unregulated electricity generation market, the capacity of transmission lines will determine the degree to which generators in different locations compete with one another. We show, however, that there may be no relationship between the effect of a transmission line in spurring competition and the actual electricity that flows on the line in equilibrium. We also demonstrate that limited transmission capacity can give a firm the incentive to restrict its output in order to congest transmission into its area of dominance. As a result, relatively small investments in transmission may yield surprisingly large payoffs in terms of increased competition. We demonstrate these effects in the context of the deregulated California electricity market.

*University of California Energy Institute, U.C. Berkeley Haas School of Business, and National Bureau of Economic Research; borenste@haas.berkeley.edu
**University of California Energy Institute; jimb@ieor.berkeley.edu
***University of California Energy Institute; stoft@planetnow.com

Acknowledgments: For helpful comments and discussions, the authors would like to thank Carl Blumstein, Glenn Ellison, Joe Farrell, Paul Joskow, Ed Kahn, Bart McGuire, Shmuel Oren, Scott Stern, Matt White, Catherine Wolfram, two anonymous referees, Kyle Bagwell (Editor), participants in the 1996 Berkeley-Stanford IOfest, and seminar participants at U.C. Berkeley and MIT. Christopher Knittel and Steve Puller provided excellent research assistance. This research was partially funded by the California Energy Commission under contract #700-93-003, amendment 2.

1. Introduction

Electricity transmission facilities have long been recognized as central elements in the efficient planning and operation of electricity systems. Traditionally, the role of large, inter-utility transmission paths has been to permit transactions between utilities that exploit regional differences in consumption seasonality and generation costs. As the electricity generation industry is deregulated in the U.S. and elsewhere, however, transmission facilities will also provide important competitive links between potentially isolated markets, thus mitigating the potential for exercise of market power.

The way that the generation firms themselves view transmission capacity will likely change as the industry becomes less regulated. A regulated firm would have little or no interest in inducing congestion into its market, where the term "congestion" is used here and throughout the paper in the electrical engineering sense: a line is congested when the flow of power is equal to the capacity of the line, as determined by various engineering standards. A profit maximizing firm, however, may find it quite profitable to induce congestion into its area, thereby becoming a monopolist on any residual demand left unserved by imports from other regions.

In this paper, we examine the incentives of unregulated generation firms and the competition-enhancing role of transmission facilities.[1] The model we use, of two geographically distinct electricity markets, can be thought of as a first-level approximation to the market distinctions between northern and southern California. In the situation where each market is dominated by a single supplier, the benefits of increased transmission capacity manifest as greater competition, in some cases with less actual power flow. The mere threat of competitive entry that is provided by additional transmission capacity acts as a restraining influence on the dominant supplier in each market, causing each to produce nearer to competitive levels even though the threatened imports are not in fact realized.

In order to achieve the full benefits of competition, the transmission capacity between these two markets must be large enough such that each firm prefers to compete over the larger market rather than to act as a residual monopolist in its own, local market. We find that the amount of transmission capacity necessary to achieve this result can be surprisingly small in some cases. Interestingly, in a deterministic world, additional transmission capacity above that critical "threshold" level is of no value. We also find that use of the transmission line may decline when it is expanded to sufficient capacity to support competition.

The analysis of competition between two dominant supply firms separated by a (potentially) constrained transmission link yields several important insights that have immediate policy implications. First, in regions with a few dominant suppliers, it seems likely that the incidence of congestion on transmission links will increase as the regulation of these firms is relaxed. This increased likelihood of congestion implies that the relevant geographic markets considered for market power analysis may be smaller than those indicated by historical congestion patterns or by simulation of perfectly competitive markets. The Federal Energy Regulatory Commission (FERC) has relied upon historical transmission flow and congestion data to determine the geographic scope of markets in order to analyze proposed mergers and has recently proposed the use of simulation

[1] We do not study the incentives of transmission owners, who remain subject to some form of regulation in nearly all markets. We take the capacity of transmission as determined exogenously.

models that assume output levels are set according to a principle of "least-cost" production.[2] Our results indicate that such approaches may greatly understate the potential for market power in some markets.

Second, in a deregulated electricity market, the social benefits provided by transmission facilities may no longer be closely linked to the usage level of those facilities. Historically, before making an investment in transmission facilities, an investor-owned utility would have to demonstrate the "public convenience and necessity" of the new line (see Baldick and Kahn, 1992). The most common measure of necessity was the level of congestion along the path served by that line and the forecast utilization of additional transmission capacity. Our results demonstrate that, in some circumstances, modest additions to a transmission network can yield extremely large social benefits, even though there would be little actual power flowing over the network.

Currently, however, discussions of standards for the planning and construction of new transmission facilities have focused almost exclusively on the reliability enhancing aspects of transmission expansion (see Western Systems Coordinating Council, 1997). Even when "economic" benefits are considered, they are usually limited to assessments of the benefits of reduced curtailments in service, increased imports or exports over the new facilities, or revenues derived from charging for the usage of the new capacity.[3] This emphasis on the reliability and cost benefits of relieving transmission "bottlenecks" could lead to poor policy decisions, since even unused capacity on a transmission path may still be providing important benefits by discouraging the exercise of market power.

In Section 2, we describe the issues and choices faced by firms in separate markets that have the potential to be geographically distinct. In Section 3, we introduce a model of two identical, but geographically distinct, markets that are linked through a single transmission path. We show that the capacity of this transmission path plays a crucial role in determining the market outcomes and derive an expression for the threshold transmission capacity that is sufficient for completely integrating these two markets. We show that for transmission capacities less than this threshold capacity, no pure-strategy Cournot equilibrium outcome can exist, and we discuss briefly the nature of the mixed-strategy equilibria. Section 4 discusses several insights, extensions and generalizations from these results, including the important extension to asymmetric markets. In Section 5, these results are applied to a model of California's electricity market, which is undergoing restructuring. In this section, we demonstrate the potentially very significant impact on this market of the capacity of the major north-south transmission path. Our conclusions are summarized in Section 6.

2. Geographic market distinctions

At first glance, the electricity industry resembles any other in which there are spatially separated markets between which trade is possible. Consider two markets that are identical in every respect. Assume that each has a monopoly supplier with identical costs. If these regions were completely separate, there would be identical monopolistic solutions in the two regions. If, on the other hand, each producer were able costlessly to ship very large quantities into the other's market, the

[2] See Joskow and Frame (1997) for an example of the former and Borenstein, Bushnell, and Knittel (1999) for a discussion of the latter.

[3] In a recent assessment of transmission capacity additions, (California ISO, 1998) the California Independent System Operator identified six needs that such projects fulfill. The only need related to the competitiveness of the generation market is the reduction of reliance upon "must-run" contracts. There is no consideration of the benefits of enhanced competition amongst suppliers in the broader California market.

markets could no longer be considered geographically distinct. There would instead be a single market served by a duopoly.

When one considers the unusual physical properties of "transporting" electricity, this standard trade problem exhibits some surprising characteristics. First, electricity can be rerouted or re-shipped almost costlessly. A marketer can buy electricity and, if lines are not congested, have it delivered to (or consumed at) a different location at virtually no cost. This means that price differences can be easily arbitraged. Second, electricity is perfectly standardized so the market deals in a completely homogeneous good. As a result, if the producer in market 1 tries to ship 10 MW to market 2, and the other producer decides to ship 9 MW to 1, only 1 MW of power actually flows over the line. In contrast, a Ford that is exported to Japan is actually shipped, not netted out from the number of Toyota's exported to the U.S. Lastly, most studies of geographic trade barriers, which consider shipping, highway, or rail transportation, do not assume constraints on shipping *capacity.* It is simply not a binding constraint for most spatially distinct markets.[4]

Homogeneity of the product, costless shipping, and symmetric markets lead to a somewhat surprising result: the transmission line causes prices to decline and outputs to increase, but no power actually flows on the connecting line. This means that although the line is not used, it is still very useful, because is keeps prices low. Even if the two markets were not perfectly symmetric, there may be very little flow on the line relative to the additional output that would be sold in each market. The *threat* of competition is all that is really needed, and the line (if it is big enough) provides that threat. Thus, if a connecting line is of sufficient capacity to reduce market power as much as possible, it may appear to be over built and under used.[5]

If we assume these same two markets are connected by an extremely low-capacity (*i.e.,* thin) line, then each supplier would almost ignore the other, because each would know that the other could affect it very little. Thus, the result would be something close to the two-monopoly solution in spite of the connecting line. The actual outcome, however, is more complicated than this, because it is not an equilibrium for both firms to leave the line unused. This raises the question of how big a line is needed to induce duopolistic, instead of monopolistic, behavior. Clearly, the answer need not be related to the actual power flow on the line. More generally, we would like to know the effect of any given size line on the degree of competition between the firms at each end of the line. We pursue these questions in the next section.

3. A symmetric two-firm model

We begin with a model of two identical markets, which we call North (N) and South (S). Demand in each market is characterized by identical inverse demand functions, $P^S(Q^S)$ and $P^N(Q^N)$, where Q^N and Q^S are the quantity consumed in markets N and S, respectively, and $P^S(.) \equiv P^N(.) \equiv P(.)$. Each market has a single supplier, firms n and s, with output q^n and

[4] Natural gas markets in some areas may also fit our model. These results would also apply to natural gas markets to the extent that there are firms with market power. Krishna (1989) discusses an example of trade quotas that is related to our analysis, as we point out below. Our work contrasts with Brander and Krugman (1983), in which firms inefficiently cross-haul goods between countries and increase competition in both markets. In that model, each firm makes a separate output decisions for each market, and there are no capacity constraints on shipping.

[5] Throughout the analysis in this paper, we ignore two further complexities of electricity transmission: loop flow and reactive power. These are no doubt important factors, but including them in the analysis would greatly complicate the model without changing our basic points. It is unlikely that the strategic importance of transmission capacity would decrease if these other factors were explicitly considered.

q^s, respectively, and the two firms have identical production costs, $C(q)$. For either firm, i, let $\pi(q_i, q_j) = P(q_i + q_j)q_i - C(q_i)$ be the profit of a firm selling in only one market when it sells q_i in that market and the other firm sells q_j in that market. If the two markets were effectively merged, then the profit of firm i could be expressed as $\Pi(q_i, q_j) = P(\frac{q_i+q_j}{2})q_i - C(q_i)$, which is the profit from selling quantity q_i at a price that obtains when the total quantity produced is split evenly between the two (identical) markets.

To be concrete about the form of competition and assure existence of a unique equilibrium in the absence of transmission congestion, we make the following assumptions:

Assumption 1: The firms employ quantity strategies (Cournot).

Assumption 2: Firm marginal costs are non-decreasing in firm output, $C'' \geq 0$.

Assumption 3: $\pi_{11}(q_i, q_j) < 0, \quad \pi_{12}(q_i, q_j) < 0, \quad \Pi_{11}(q_i, q_j) < 0, \quad \Pi_{12}(q_i, q_j) < 0.$

Assumption 4: $|\Pi_{11}(q_i, q_j)| > |\Pi_{12}(q_i, q_j)|$

where the subscripts 1 and 2 on the profit functions refer to the derivatives with respect to their first and second arguments respectively.

We assume quantity setting due to the severe constraints that the need to commit generation capacity in advance puts on production. In Section 4, we discuss the Cournot assumption at greater length and consider how the analysis changes if firms are assumed to compete in prices instead.[6] Assumptions 3 and 4 guarantee the existence and uniqueness of a Cournot equilibrium and are utilized in later results (see Tirole, 1988, pp. 224-226).

To complete the characterization of the markets, we must also describe how transmission is allocated and priced. We assume that the transmission grid is operated by an entity that attempts to maximize social welfare by providing price signals to induce efficient use of the grid, which is just the line between N and S in this case. The result is known as "nodal pricing." See Schweppe et al. (1988) for a description and analysis of nodal pricing. Under this assumption, if there is no congestion, electricity prices are the same in both markets. When there is congestion in the network, transmission capacity is rationed with price. The effective "cost" of transmission is simply the nodal price difference between the node in which the power is injected into the network and the node where it is consumed. Thus, the grid operator collects the congestion rents associated with a transmission line, but never deters line use with a positive price when the line is uncongested. Even if the producers wished to charge different prices in the two markets, the grid operator, or some third party, would arbitrage the price differences as long as it was feasible to ship power between the two markets.

[6]Wolak and Patrick (1997) provides evidence that electricity generators in the U.K. have used plant availability to compete in quantities in the short run. Klemperer and Meyer (1989) present a model of supply curve competition, which Green and Newbery (1992) apply to the UK market. Wolfram's (1999) analysis finds that U.K. prices are below those that the Cournot model would predict, but argues that the explanation might be the threat of regulation. von der Fehr and Harbord (1993) suggest a multi-unit auction approach to evaluating competition in electricity generation markets. While these approaches are more complex than the Cournot analysis, we see no reason that the basic effects we identify here wouldn't also obtain in other oligopolistic models.

Thus, under nodal pricing, power will flow to the market with the lowest price, subject to flow limits.[7] In the context of this two-market model, with a connecting transmission line of capacity k, nodal pricing determines the functional mapping from the output of the two firms to the quantities consumed in the two markets. Let $N : \Re^2_+ \mapsto \Re^2_+$ denote this function.

$$(Q^N, Q^S) = N(q^n, q^s) = \begin{cases} (q^n + k, q^s - k) & \text{if } q^n < q^s - 2k \\ (\frac{1}{2}(q^n + q^s), \frac{1}{2}(q^n + q^s)) & \text{if } q^s - 2k < q^n < q^s + 2k \\ (q^n - k, q^s + k) & \text{if } q^n > q^s + 2k \end{cases} \quad (1)$$

Under nodal pricing each firm receives its local price for all its output, *whether it is consumed locally or not.*[8] Figure 1 illustrates a demand function faced by a single firm, firm n. For a transmission line of capacity k, if firm n produces a quantity less than $q^s - 2k$, then equation (1) indicates that the quantity k of power is shipped from S to N and firm n faces the region N demand curve shifted left by k. As n increases its output, once it reaches $q^n = q^s - 2k$, the line becomes decongested and the markets are effectively merged, with equal prices and quantities sold in the two markets. The demand curve seen by n for this range of output is the sum of the two market demand curves minus the output of firm s. This will be true until n has increased its output to $q^n = q^s + 2k$ units and k units are being shipped to market S, just barely congesting the line. When $q^n > q^s + 2k$, increases in q^n have no effect on P^S and will cause a decline in P^N, which is the price n will receive for all of its output. For output levels in this range, firm n faces a demand curve equal to its local demand curve shifted to the right by k, the capacity of the line.

The objective of the firm is to maximize its profits by choosing output. For firm n:

$$\Omega_n(q^n, q^s) = \begin{cases} q^n P(q^n + k) - C(q^n) & = \pi(q^n, k) & \text{if } q^n < q^s - 2k \\ q^n P(\frac{1}{2}(q^n + q^s)) - C(q^n) & = \Pi(q^n, q^s) & \text{if } q^s - 2k < q^n < q^s + 2k \\ q^n P(q^n - k) - C(q^n) & = \pi(q^n, -k) & \text{if } q^n > q^s + 2k \end{cases} \quad (2)$$

and similarly for firm s.

With no connecting line, each firm will produce the single-market monopoly quantity, q_m, and the price in each market will be P_m, the monopoly price. With a sufficiently large line, the prices in the two markets will be equal to one another and will equal the unconstrained Cournot duopoly equilibrium price for the combined market with two identical firms, which we will call P_c.[9] Each firm will then produce q_c, its (unique) pure-strategy unconstrained Cournot quantity. Total production in both markets in the duopoly case will therefore be $2q_c$.

[7] For now, and throughout most of this analysis, we assume that there are no line losses — energy can be transported costlessly. We discuss line losses in Section 4.

[8] Variations of nodal pricing have been adopted or proposed in several electricity markets throughout the world. Although we adopt the conventions of nodal pricing, our results do not depend upon this method of transmission pricing. An earlier version of this paper (Borenstein, Bushnell and Stoft, 1998) treated transmission capacity as a costless, congestible public good. Our analysis does assume that prices can differ in the two markets. If price were required to be uniform across markets, an entirely different analysis would apply. We require only that transmission capacity be allocated efficiently, which implies that the *cost* of transmission is equal to the difference between locations in the price of power. Chao and Peck (1996) have proposed an alternative system of transmission pricing that also meets these conditions.

[9] "Uncongested" or "unconstrained" here and throughout the paper means that the capacity of the transmission line does not impose a binding constraint that affects either firms' quantity choice.

The competitive effect of a transmission line

We begin examining the competitive effect of transmission lines by looking at the impact of building a line between the two markets that were previously separate (identical) monopolies.

Lemma 1. Under Assumptions 1-4, when two symmetric monopoly markets are connected by a transmission line, either firm would prefer to increase its output from its monopoly level, q_m, if the other continues to produce q_m.

Proof. To see that it would always be profitable for one firm to expand its output in the presence of even a very thin line (if the other firm did not), note that if neither firm changed its output then the line would be uncongested. If the line is uncongested, then each firm faces (locally) the same problem it would face if there were no transmission congestion, the duopoly market conditions. In that case either firm would want to expand it output to its Cournot duopoly best response. In fact, this observation applies to any pair of firm outputs that would leave the transmission line uncongested. *Q.E.D.*

Lemma 2. Under Assumptions 1-4, the only possible pure -strategy equilibrium in which the line is uncongested is (q_c, q_c).

Proof. When the transmission line is uncongested, each firms' relevant profit function for marginal changes is $\Pi(.,.)$. Under Assumptions 2 and 3, the only equilibrium can be (q_c, q_c). *Q.E.D.*

Of course, with a very thin line, a firm would be able to increase its output only slightly before it would congest the line. More generally, one can derive the best-response (in quantity) functions of each firm in the presence of a line of size k. Figure 2 shows the best response function of firm s when k is small. When n is producing nothing, the best response of s is to produce its optimal quantity given that the line will be congested from S to N.[10] Under a nodal pricing regime, this quantity will be the monopoly output for the firm when it is faced with its native demand in the south shifted to the right by k for any price less than $P(k)$. We now define that quantity.

Definition: Let the quantity $q_m^+(k)$ represent the profit maximizing output for a firm facing an inverse demand curve $P(Q-k)$.[11] This is the "optimal aggressive output" for a firm when the other firm is producing less than $q_m^+(k) - 2k$, causing the transmission line to be congested into the market of that low-output firm:

$$q_m^+(k) = \arg\max_{Q} \quad \pi(Q, -k)$$

When n is producing up to $q_m^+(k) - 2k$, the best response of s is to produce $q_m^+(k)$. As n's output rises above $q_m^+(k) - 2k$, the best response of s, at least locally, will be to produce $2k$ units more than n,[12] firm s's optimal aggressive output response. As n's output continues to rise, however, the profits to s from maintaining an aggressive response decrease.

[10] This assumes that s would want to produce $q^s > 2k$ if there were no capacity constraints on the line and n produced nothing.

[11] By Assumption 3, $\pi_{12} < 0$, which is sufficient to assure that $q_m < q_m^+(k)$ for any $k > 0$.

[12] Note that if s does not raise its output to keep its output $2k$ above n's, the line will become decongested. With a decongested line, s would want to increase its output to get to its Cournot best response. With a sufficiently thin line, s's output increase will congest the line before it gets to its Cournot best response.

As n's output continues to rise, eventually it reaches the point – which we define below as q_{sw} – at which it becomes more profitable for s to "switch" to a much less aggressive output response: s will then allow n to export k units to market S and will maximize its own profit given the residual demand it faces, $P(Q+k)$, which is s's native demand shifted leftward by k. Firm s is then effectively a monopolist on this residual demand curve. We call this the "optimal passive output."[13] These observations are formalized by the following definitions and results.

Definition: Let $q_m^-(k)$ represent the profit maximizing output for a firm facing an inverse demand curve $P(Q+k)$. This is the "optimal passive output":[14]

$$q_m^-(k) = \arg\max_{Q} \quad \pi(Q,k)$$

When it follows the optimal passive output response, the passive firm does not "fully accommodate" imports.

Lemma 3. When the quantity k is shipped into the market of a monopolist, its optimal passive output response results in total market output (including the quantity imported, k) above the monopoly level, *i.e.*, $q_m^-(k) + k > q_m$.

Proof. Consider the incumbent's position if it were to respond to imports of k by reducing its own output by k from its previous (zero import) optimizing quantity. In that case, the price in its market would be the same as before the imports and the effect on price of selling one more unit would be the same as before, but the incumbent would be selling fewer units. Thus, it's marginal revenue would be greater than before. So long as marginal cost is non-decreasing (Assumption 2), this implies that the incumbent would want to increase its output. *Q.E.D.*

Thus, each firm's best-response function slopes upward over some range and then discontinuously jumps to a smaller quantity if the other firm's output is sufficiently great. At the smaller quantity, the line will be congested carrying power into its region. With these best-response functions, we can show that the functions will not cross if k is sufficiently small. In other words, *for transmission lines that have sufficiently small capacity, there is no pure-strategy equilibrium for the symmetric Cournot duopoly.*

We prove this as a subcase of a later result, but the intuition is straightforward: If firm s is sufficiently aggressive in expanding output in order to be a net exporter of k to the other market,

[13] Krishna (1989) identifies a similar form of "passive" behavior in the context of international trade quotas. In that case, one firm, recognizing the limited quantity that can be imported into its home market, reduces its own output, thereby causing the trade quota to be binding and permitting the firm to maximize profits on the residual demand function that it faces.

[14] One additional technicality must be addressed before taking $q_m^-(k)$ to be the optimal passive output response: the assumption that $q_m^-(k)$ will cause the line to be congested whenever it is the optimal output for a firm. That is, if a firm believes that it is facing inverse demand $P(Q+k)$ and produces its optimal output response, would that necessarily cause the line to be congested (thus, confirming its belief that it faces inverse demand $P(Q+k)$)? The answer is that producing $q_m^-(k)$ would cause the line to be congested in any circumstance in which a firm might choose to produce it. To see this, note that (1) at the quantity choice of the firm that just decongests the line, call this $\tilde{q}$, $\pi(\tilde{q},k) = \Pi(\tilde{q},q^n)$, where q^n is the rival's output, and that (2) $\pi(q,k) < \Pi(q,q^n) \quad \forall q > \tilde{q}$. This means that if $q_m^-(k)$ caused the line to be decongested, it would also generate lower profit (when calculated as if the line were congested) than a best response in the combined market, so the firm would never choose to produce $q_m^-(k)$ in that case. Thus, if the firm ever chose to produce $q_m^-(k)$, on the assumption that it would lead to congestion on the line, that quantity would indeed congest the transmission line.

then firm n's best response will be the optimal passive output, $q_m^-(k)$. But, s would respond to the optimal passive output by producing just $2k$ more than the passive quantity. If k is sufficiently small, n's best response to $q_m^-(k) + 2k$ is to produce $q_m^-(k) + 4k$ and ship k back to the s's market. As this escalates, eventually one firm again finds that it would prefer following the optimal passive output response to further escalation.

As k increases, the problem changes somewhat. While for some k a firm might find it worthwhile to congest the line to the other market by producing $2k$ more than the other firm, there will never be a best response that involves producing more than the unconstrained Cournot duopoly best response, which is, by definition, the optimal response when it is feasible. Figure 3 shows the situation that obtains if k is large enough that the unconstrained Cournot duopoly best response is feasible for some outputs of the other firm. Rather than sloping upward until the point at which the firms switches to the optimal passive output and allows imports, the best response function slopes upward until it hits the unconstrained Cournot duopoly best-response function, and then it coincides with the unconstrained Cournot duopoly best-response function until the discontinuity point.

Definition: Let $UBR[q]$ be a firm's *unconstrained* Cournot best response to an output of q by the other firm.

$$UBR[q] = \arg\max_{Q} \quad \Pi(Q, q)$$

There is a level of output from a rival in the range of $[q_m^+(k) - 2k, \infty]$ at which the firm's best response is to revert to the optimal passive output response, maximizing profits on the residual demand curve while the other firm exports k into its market. We can now formally define this discontinuity point of the reaction function.

Definition: Let $q_{sw}(k) \in [q_m^+(k) - 2k, \infty]$ be the rival's quantity above which it is more profitable for a firm to switch from either the unconstrained best response, if it does not violate the transmission constraint, or the optimal aggressive output, to the optimal passive output with its rival transmitting k into its market. Over the relevant range of $[q_m^+(k) - 2k, \infty], q_{sw}(k)$ is the implicit solution to:

$$\Pi\left(\min(UBR[q_{sw}], q_{sw} + 2k), q_{sw}\right) = \pi(q_m^-(k), k) \tag{3}$$

Lemma 4. q_{sw} exists, is unique, and is increasing in k, the line capacity.

Proof. See the Appendix.

Intuitively, the switch point occurs at a larger output of the opposing firm as k increases, because an increase in k makes the optimal passive output less profitable, so an increase in the opposing firm's output is necessary to make the alternative equally less profitable. If k is large enough, the discontinuity point in each firm's best response function will occur at or beyond the unconstrained Cournot duopoly equilibrium. In this symmetric model, the symmetric Cournot duopoly equilibrium will then result. This is shown in Figure 4. The condition then for the unconstrained symmetric Cournot duopoly equilibrium to obtain is that k is large enough so that each firm will make greater profit in that equilibrium than it would producing the optimal passive output, *i.e.,* allowing the other firm to export the full capacity of the line into its own market and producing its best response to the resulting residual demand. Any additions to the line beyond this point will still change the best-response functions, but will not change the point at which they intersect, which determines the equilibrium. Thus, further increases to the line capacity are of no social value. The k at which the unconstrained Cournot duopoly equilibrium will result is the line capacity that equates the profit each firm earns in the unconstrained Cournot duopoly equilibrium with the profit either would earn producing its optimal passive output response when the other firm produces its unconstrained Cournot duopoly quantity. Clearly, this threshold line capacity depends on the nature of demand.

Definition: Let k^* represent the smallest transmission line capacity that will result in the symmetric unconstrained Cournot duopoly equilibrium.

The capacity k^* is the implicit solution to

$$\pi(q_m^-(k), k) = \pi(q_c, 0). \tag{4}$$

The expression $\pi(q_m^-(k), k)$ is the profit to a single firm from producing its optimal passive output and $\pi(q_c, 0)$ $(\equiv \Pi(q_c, q_c))$ is the profit to a single firm from the unconstrained Cournot duopoly equilibrium. The expression $\pi(q_c, 0)$ is independent of k. The expression $\pi(q_m^-(k), k)$ is the profit earned from optimizing along a demand curve that is a leftward shift of k units from the market demand curve, so the lefthand side is monotonically decreasing in k. At $k = 0$, $\pi(q_m^-(k), k) > \pi(q_c, 0)$ and at $k = \infty$, $\pi(q_m^-(k), k) = 0$. Both expressions are continuous in k, so there is some k that equates the profit of these two strategies.

Theorem 1. Under Assumptions 1-4, the only possible pure strategy equilibrium to this game is the unconstrained Cournot duopoly equilibrium. If $k < k^*$, no pure strategy equilibrium exists.

Proof. If we assume Theorem 1 is not true, then one of three cases must hold. There will either be an asymmetric equilibrium where the line is congested, an asymmetric equilibrium where the line is not congested, or a symmetric equilibrium where both players play some quantity less than the Cournot quantity (and the line is uncongested). By Lemma 2, the only pure strategy equilibrium with an uncongested line is the unconstrained Cournot duopoly equilibrium. We therefore restrict our attention to asymmetric equilibria where the line is congested.

First, we note that there cannot be an asymmetric, congested-line equilibrium unless at least one player is producing its optimal passive output, since the optimal passive output is by definition the best response of the producer in the market into which the congested line is flowing. We now show that there cannot exist an asymmetric pure strategy equilibrium in which one player produces the optimal passive output, q_m^-.

We know from the definitions above that if an equilibrium exists in which the passive firm is producing q_m^- and the line into its market is congested, then the aggressive firm must be producing at least $q_m^- + 2k$, the minimum quantity that would congest the line. We now show that the aggressive firm would never want to produce more than $q_m^- + 2k$ in response to q_m^-, so it could only be producing exactly $q_m^- + 2k$ in this potential equilibrium.

To see that this is the case, recall that by Lemma 3 $q_m^- > q_m - k$, which implies that $q_m^- + k > q_m$. When the aggressive firm produces $q_m^- + 2k$ and k is shipped to the other market, this leaves the firm producing $q_m^- + k$ for sale in its own market, which we have just shown would be greater than the profit-maximizing quantity for its own market in isolation, q_m. Considering that, under nodal pricing, additional output that is consumed in its own market also drives down the price that the aggressive firm receives for power that is shipped to the other market, it is clear that it would not want to further expand output beyond the quantity $q_m^- + 2k$, which just congests the line.

Thus, the only possible candidate for an asymmetric, pure-strategy equilibrium is one firm producing q_m^- and the other producing $q_m^- + 2k$. For this to be an equilibrium, it must be the case that the optimal passive output, q_m^-, is the best response to the aggressive firm's $q_m^- + 2k$. But production of $q_m^- + 2k$ in response to q_m^- just barely congests the line, *i.e.,* if the passive firm increased its output infinitesimally, it would decongest the line. This means that the profit that the passive firm earns in this outcome lie on its unconstrained profit function, the profit function it would face if each firm produced these quantities, but there were no line constraint. This firm's unconstrained profit function is, by assumption, strictly concave, and its profit maximum is its Cournot best response, which is greater than the optimal passive output. Thus, the optimal passive output cannot be a best response to $q_m^- + 2k$. Therefore, there cannot be an asymmetric pure strategy equilibrium involving one firm producing its optimal passive output. *Q.E.D.*

Transmission lines and unconstrained Cournot equilibria

An interesting and potentially very important area for study is the size of the line necessary to yield the unconstrained Cournot duopoly equilibrium. We have investigated this question in a simple model with constant marginal cost and linear or constant elasticity demand. The results are fairly consistent between these two functional forms. We discuss first the linear demand case.

Consider a case in which the demand in each of two markets is $Q = 10 - P$ and the marginal cost of production is constant at zero. With no line, each firm will produce 5 units of output and the price in each market will be 5. With a very high capacity line between the markets, there will be no transmission constraint and the firms will compete to the Cournot duopoly equilibrium in the combined market, in which each firm produces 6.667 units and a market price of 3.333 results. Each firm then earns profit of 22.22.

Using these results and equation (4), we can calculate that the Cournot duopoly outcome is achieved for any line with $k \geq .57$ (approximately). With $k \approx .57$, either firm would be indifferent between producing 6.667 units and producing 4.715, the optimal passive response, given that the other firm is producing 6.667 units. At any larger k, each firm would strictly prefer the Cournot equilibrium to producing the optimal passive response. The Cournot outcome therefore results when the line capacity is about 17% of the total *increase* in output in the two markets that results from building the line ($k^* \approx .17(2q_c - 2q_m)$). Since selection of quantity and price units is arbitrary, it is easy to demonstrate that these ratios are the same for any linear demand with constant marginal cost. As pointed out earlier, if such a line were built, no power would flow on it, as the equilibrium would be symmetric.

We also have investigated the general constant elasticity demand function $Q = AP^{\epsilon}$ with constant marginal cost, $MC = m$, for a range of values of A, m, and ϵ. For each case, we calculated the same ratio as above, in which the numerator is the size of the line necessary to achieve the unconstrained Cournot duopoly outcome and the denominator is the increase in output associated with a change from separate monopolies to unconstrained Cournot equilibrium when the two markets are fully integrated. The ratio is independent of A and $m(>0)$. It ranges from about 8.5% when $\epsilon = -1.1$ to 14.4% when $\epsilon = -10$, increasing monotonically as ϵ increases in absolute value.

Thin lines and mixed-strategy equilibria

We have also examined what happens to prices and output levels for line capacities that are below the threshold level that induces the Cournot duopoly result, line capacities for which no pure-strategy equilibria obtain.[15]

Even without finding the actual mixed-strategy Cournot-Nash equilibria, it seems likely that the expected price will decline as k increases from zero. With a very thin line, for instance, the expected price must be very close to the monopoly level. If it were not, then either firm could improve its expected profit by simply admitting imports of k and producing the optimal passive output as a pure strategy. With k near k^*, the lower bound on price provided by the optimal passive output response is much weaker and the mixed strategy is more likely to result in a lower expected price.

In numerical examples that we studied – constant marginal cost with either constant elasticity or linear demand – we find that small increases in line capacities can yield expected output increases much larger than the added line capacity. This is consistent with the analytic conclusion that the line capacity necessary to completely merge the two markets is relatively small compared to the added output that such a merging produces. In both models, even small lines produced big benefits in expectation. In fact, the marginal effect of increased line size appears generally to be greatest when the line is very thin, though the slope does not always change monotonically.

4. Extensions and generalizations

Dominant firms with fringe suppliers

Pure monopolies are quite rare, of course, even in the formerly regulated electricity generation business. It is worth noting, however, that the analysis we have presented applies equally well to two symmetric markets in which each includes a dominant firm and a competitive, price-taking fringe. In that situation, the demand functions analyzed above are the residual demand functions in each market, *i.e.,* market demand minus the quantity that the fringe will produce at each price. We have assumed efficient arbitrage between markets throughout the analysis – market prices differ only if the line between the two markets is congested – so the presence of the additional competition in each market has no further effect on the relationship between the markets.

One rival can be more competitive than many

An intriguing result emerges when we compare the outcome from our two-monopoly model with the outcome when one of the markets is already perfectly competitive before the line is built. For example, if the demands in each market are still identical, $P = 10-Q$, and marginal cost is constant at zero, then the competitive market will have a price equal to zero both before and after the line is built. If a line of $k = .57$ is built (the capacity that yields the unconstrained Cournot duopoly equilibrium), then the previous monopolist will produce its best response when an inflow of .57

[15] A detailed description of a simulation exercise in which we solve for mixed-strategy equilibria for various market types is available from the authors at
http://www.ucei.berkeley.edu/ucei/PDF/bbsmixed.pdf.

congests the line, which is to produce $q = 4.715$, the optimal passive output. The total quantity in the previously monopolized market would then be 5.285, much less than the Cournot duopoly quantity that resulted in each market when the same size line was used to connect two previously monopolized markets. That is, the competitive effect of this line on a monopoly market is greater if the market at the other end is also a monopoly than if it is competitive.

The reason for this surprising result is that a more aggressive output choice deters imports when the potential importer is also a monopolist in its home market, while it has no such effect when there is a competitive market at the other end of the line. Thus, with a competitive market at the other end of the line, the Cournot duopoly outcome is not a possibility, and the firm will always choose to produce its optimal passive output, optimizing along the residual demand curve that results when the line is congested with inflowing electricity.

Of course, the result does not hold for all size lines. If the line is sufficiently large ($k = 10$ in the linear demand example), then both markets would be driven to the competitive price. Still, if the line is smaller than is necessary to induce an output increase in the monopolized market equal to the output increase when it expands to the Cournot duopoly equilibrium, then a customer in a monopolized market will benefit more from having a monopoly market at the other end of the line than from having a competitive market at the other end. In the present example, if market N were monopolized, then any line smaller than $k = 3.333$ would yield greater benefits to the consumers in market N if market S were also monopolized than if market S were competitive. This doesn't even take into account the fact that it would also benefit the customers in a monopolized S market, while it would have no effect on a competitive market S.

Bertrand competition

We have thus far considered only the case in which the firms compete in quantities.[16] This seems reasonable given the structure of most actual deregulated electricity markets. In such markets, firms generally bid supply curves into a forward or futures market. In real time, however, firms can (and, in California, regularly do) depart from their scheduled output to produce more or less than they have been scheduled for. The system operator adjusts price to equate supply and demand, and pays (or charges) this "real-time" price for deviations from scheduled quantities. While this behavior doesn't track Cournot competition precisely, quantity competition does seem like the best simple representation of the process. Furthermore, Wolak and Patrick (1997) provide evidence that exercise of market power in the UK has taken place through capacity withholding: at high-demand times the two large firms in the UK appear to have made some of their capacity unavailable in order to raise the market-clearing price.

It is possible, however, that even after committing some portfolio of generating units to produce on a given day, a dominant firm will still have significant output flexibility in the short-run. It is therefore worth considering the competitive effect of transmission lines when there is an unconstrained dominant firm at each end of the transmission line and these firms compete in prices. With both firms recognizing a *quantity* constraint in transmission, however, their behavior in choosing prices will necessarily differ somewhat from the standard Bertrand model.

Again, we will assume that the markets and the firms are identical. To ensure existence of a pure-strategy Bertrand equilibrium, we will now also assume constant marginal costs of production

[16] Oren (1997) and Cardell, Hitt, and Hogan (1997) also assume Cournot behavior to examine other questions concerning market power in electricity networks.

for each firm. If there is no transmission line between the markets, then obviously the monopolists in each market will exercise their full market power. If the line is sufficiently large, then the standard Bertrand equilibrium will obtain: the markets will be fully integrated and price will be equal to marginal cost.

Theorem 2. If the two identical firms in identical markets N and S compete in prices and if each firm has constant marginal costs, then
(1) the minimum capacity line that supports the unconstrained Bertrand duopoly equilibrium is equal to the quantity demanded in either market when price is equal to MC: $k_B^* = Q(MC)$, where $Q(\cdot)$ is the demand function in a single market.
(2) the only possible pure strategy equilibrium to this game is the unconstrained Bertrand duopoly equilibrium.

Proof. (1) Since price is equal to marginal cost in the unconstrained Bertrand duopoly equilibrium, firms earn zero operating profit. Thus, if a firm were to make a "passive" price choice — *e.g.,* firm s charging a higher price than firm n and firm n exporting the full capacity of the line into market S — and this left *any* residual demand for s to serve at a price above MC, then s could make higher profit optimizing along a residual demand curve after n congests the line than playing the unconstrained Bertrand strategy. Only if n can force the price in S down to MC (or s can force the price in N down to MC) will the Bertrand outcome obtain. This requires a line equal in capacity to the quantity demanded in either market when price is equal to MC, $k_B^* = Q(MC)$.

(2) If the line capacity is less than this k_B^*, it is clear immediately that there will not be a pure strategy symmetric equilibrium with both firms charging the same price above marginal cost; in that situation either firm could increase profit by cutting price infinitesimally and thereby selling all quantity in its own market and k, the capacity of the line, in the other market.

Nor could there be a pure strategy equilibrium in which the two firms charge different prices. If such an equilibrium did exist, it would have to involve the firm with the lower price, call this firm n for clarity, congesting the line with exports to the other market and the firm with the higher price, firm s, optimizing along the residual demand that is left in its own market. But if firm s makes a passive price choice, allowing the line into its market to be congested, and raises its price to profit maximize along the residual demand it faces, then n will raise its own price to infinitesimally below P^s and still congest the line into S. If their prices are nearly equal, however, s is better off lowering its price infinitesimally to below P^n, allowing s to capture all sales in its market and k in the other market. This will restart the price cutting until one firm again prefers to make a passive price choice and allow the line into its market to be congested. *Q.E.D.*

Thus, as with Cournot, there will be no pure strategy equilibrium with $k < k_B^*$. It is straightforward to show that $k_B^* > k*$, *i.e.* for symmetric markets the threshold line capacity will be larger for Bertrand competition than for Cournot competition. The resulting increase in production relative to the monopoly levels will also be greater with Bertrand competition, however.[17] There still will be at least one mixed-strategy equilibrium with a line of $k < k_B^*$ and, for the reasons described in Section 3, it seems likely that the expected price will decline as k increases from zero.

Transmission capacity under regulation and deregulation

[17]In the linear demand case discussed in Section 3, the increase in output was eight times the line capacity needed to induce that increase under Cournot competition. Under Bertrand competition this ratio is 1 to 1.

The transmission capacity that now exists in most of the U.S. was built to provide the cost-arbitrage and reliability functions described earlier. It was not built to augment competition among power generators. Our analysis leads immediately to the question of how the socially optimal transmission capacity under regulation compares to the social optimum when power generating markets are deregulated. Clearly, it is difficult to make such a theoretical comparison without characterizing the regulation in detail. We will assume here that regulation is optimal, first-best regulation of price, location of generation, and transmission capacity.

The symmetric markets that we have studied thus far demonstrate that situations can exist in which the optimal transmission capacity under regulation would be less than for competition. With symmetric markets, transmission lines fulfill no cost-arbitrage function. The value of transmission for reliability is likely to be just as great under competition as regulation. Furthermore, if each firm produces from a portfolio of generators, as is generally the case, the value of transmission lines in providing reliability diminishes rapidly and is unlikely to justify transmission lines that are of significant capacity compared to market outputs.

Ignoring reliability considerations, there is no value to a transmission line in the symmetric market case under regulation; the optimal line capacity would be zero. In a deregulated market, we have demonstrated that the marginal benefit to transmission capacity could be quite large. Of course, the socially optimal transmission capacity would still depend on the cost of constructing transmission, but in this case it clearly will be (weakly) greater than under regulation.

Unfortunately, even ignoring reliability issues, one cannot make a general statement that the marginal value of transmission is greater in a deregulated environment. For an example of a situation in which the marginal value of transmission would be greater under regulation, consider two markets with identical demand in which the single producer in one market has slightly lower constant marginal costs than the single producer in the other market. From the previous sections, it is clear that there is some line size that will support a Cournot duopoly equilibrium and that any transmission capacity beyond that k^* has no social value. This will be less than the k that would be necessary for the low-cost producer to serve all demand in both markets, which would be the case under optimal regulation that sets price equal to the low cost firm's marginal cost. For transmission capacity above k^*, but below the quantity consumed in the market of the high-cost producer, the marginal social value would be greater in a regulated market. Thus, an optimally regulated market *could* yield greater transmission capacity than would be optimal in a deregulated market.[18]

The effect of line losses

Thus far, we have assumed that the transmission line is lossless. In reality, of course, some electricity will be dissipated as heat in the transmission process. The proportion of power lost in this way increases with the load on the line and can be as great as 5-10% of the flow on the line. In this section, we show that the presence of small losses will lower output quantities in the unconstrained Cournot equilibrium and will lower the line capacity necessary to attain the equilibrium. If losses are sufficiently large, however, they can eliminate the existence of a Cournot equilibrium.

[18]Of course, this comparison is based on the assumption of Cournot behavior in a deregulated market. The outcomes, and the value of transmission, would be the same under optimal regulation and perfect competition.

A standard approximation for losses on a line is $L = rf^2$, where f is the flow on the line and r is a constant based on the physical properties of the line (see Schweppe, 1988, appendix D). This implies that marginal loss is $l = 2rf$, which equals zero when there is no flow on the line. When losses are added to our model, the prices in our two nodes need not be the same even when the flow on the line is less than capacity.

To assess the impact of losses, it is necessary to state explicitly who bears the cost of the losses associated with increased flow on a line. The dominant paradigm for pricing losses in a nodal electricity market is to set nodal prices to reflect the *marginal* impact on losses in the transmission grid. In our two-node model, this implies that the difference in prices between the two nodes must reflect the marginal cost of "shipping" the power from one node to another. In other words, when there is a non-zero flow between the two nodes, the price in the importing node must exceed the price in the exporting node by an amount equal to the marginal cost of losses on the line. Additionally, the power that arrives at the importing node is less than the power that leaves the exporting node. If f MW is exported from node N to node S, then $f - rf^2$ arrives in node S. When there are losses, but no congestion, there is an arbitrage condition that determines the relationship between prices in the two nodes, N and S. Adopting the convention that $f > 0$ for flow from S to N and $f < 0$ for flow from N to S,

$$P(q^s - f - rf^2)(1 + 2rf) = P(q^n + f) \quad \text{if} \quad q^s < q^n \tag{5a}$$

$$P(q^s - f) = P(q^n + f - rf^2)(1 - 2rf) \quad \text{if} \quad q^s > q^n \tag{5b}$$

The right-hand side of (5a) is the profit of an arbitrager from selling a unit of electricity in market N and the left-hand side is the profit from shipping the power to market S and then selling the quantity that remains after the marginal losses imposed $(2rf)$ are deducted from that shipment. For example, when $q^s < q^n$, $P(q^s - f - rf^2)$ is the price in S and $(1 + 2rf)$ (which is less than 1, since $f < 0$) is the claim to power in S that the shipper has after covering the losses that the shipment imposes.

Line losses will change the Cournot equilibrium quantity relative to the case when there were no losses. This is true even though, in the symmetric Cournot equilibrium, there is no flow and therefore there are no losses.

Theorem 3. Assume that total line losses are $L = rf^2$ for a flow of f between nodes, that the price difference between nodes reflects the marginal losses, $2rf$, as set forth in (5), and that the line capacity is very large. For a sufficiently small r, a symmetric Cournot equilibrium will exist and the equilibrium quantity of each firm, q_c^l, will be less than the Cournot equilibrium quantity when there are no losses, q_c.

Proof. See the Appendix.

Since the marginal pricing of losses has the effect of lowering Cournot equilibrium output levels and raising Cournot equilibrium profit, the threshold transmission capacity, k^*, that supports this Cournot equilibrium also changes.

Corollary 1. Under Assumptions 1-4, the threshold line capacity required for a symmetric Cournot equilibrium is lower when there are marginally priced losses than when there are no losses, $k_l^* < k^*$.

Proof: Since the profits from producing the optimal passive output are the same with or without losses, and the profits of the symmetric Cournot equilibrium are greater with losses, a line of capacity k^* would produce strictly greater profits from the Cournot equilibrium than from producing the optimal passive output. Reducing the line size from k^* has no effect on profits from the Cournot equilibrium, but raises the profits from producing the optimal passive output. Thus, a line of capacity less than k^* is necessary to equate the profits of the Cournot equilibrium with the profits from producing the optimal passive output. *Q.E.D.*

Although it is "easier" (in terms of line capacity necessary) to achieve the Cournot outcome when losses are introduced, consumer and total surplus are lower at the with-losses Cournot equilibrium than at the no-losses Cournot equilibrium. Thus, line losses have two offsetting effects: lower capacity investment necessary to reach a Cournot equilibrium and lower total surplus gain from doing so. Which effect dominates will depend on the marginal cost function for line capacity and the shapes of the market demand functions.[19]

If r is large enough, it is clear that these results will not hold. For a very large r, the effective capacity of the line approaches zero, though it never is zero since there are no marginal losses when there is no flow on the line. We saw in Section 3 that no pure-strategy equilibrium exists for very low capacity lines.

Asymmetric firms and markets

So far, we have analyzed symmetric firms and markets, which have the special property that identical pure strategies will produce no flow on the connecting line. While the analysis of symmetric markets is useful as an illustration, interesting real-world applications, including the one in the following section, will invariably be asymmetric to some extent. Fortunately, the intuition and analytical approach developed thus far carries over to the analysis of asymmetric markets.

With a low-capacity transmission line and symmetric markets, we were able to rule out any pure strategy equilibrium; we showed that the unconstrained Cournot equilibrium was not supportable and that no other pure strategy equilibrium was feasible. When markets are asymmetric, however, a pure strategy equilibrium is likely to result with a very thin line.

Theorem 4. Under Assumptions 1-4, a pure strategy equilibrium exists as the transmission line capacity approaches zero if and only if the monopoly equilibria in the two markets (*i.e.,* with no transmission line) produce different prices. In such an equilibrium, the transmission line is congested with flow from the lower price (under monopoly) market to the higher price (under monopoly) market.

Proof: Consider two markets S and N in which the monopoly prices are $P_m^S < P_m^N$. If a transmission line of very low capacity, k, is built between the two markets, then absent a change in output by the firm N, k units of power will be shipped from S to N and firm s will reoptimize by increasing its output. For a sufficiently small k, this would still result in $P^S < P^N$. The question then is whether firm n's best response would be to produce its optimal passive output or to increase its output enough to either decongest the line or congest it with flow in the opposite direction, the aggressive response.

[19] We have not compared the values that a certain line would have with and without losses when it is not sufficiently large to achieve Cournot equilibria and results in mixed-strategy outcomes.

For firm n, the passive choice has a negative effect on its profit (compared to the monopoly equilibrium) that goes to zero as $k \to 0$. In contrast, for firm n to decongest the line or congest it in the opposite direction, even for an arbitrarily small k, requires that n increase output at least enough to lower the price in N to P^S. This has a negative effect on n's profit (compared to the monopoly equilibrium) that does not go to zero as $k \to 0$. Thus, for a sufficiently small k, firm n would prefer to respond by producing the optimal passive output, and a pure-strategy equilibrium exists in which the line is congested with power flow from S to N.

The component of the proof that shows $P_m^S = P_m^N$ implies that no pure-strategy equilibrium exists even for small k is the same as in the proof of Theorem 1. *Q.E.D.*

Virtually any real pair of markets will embody asymmetries that result in different monopoly prices and thus will have such a pure strategy "passive/aggressive equilibrium" for sufficiently small lines.[20] This is illustrated in Figure 5 for two markets in which N is larger than S, the two firms, n and s, have equal and constant marginal costs, and $P_m^N > P_m^S$.[21] As the capacity of the line increases, exports from the low-price market, S, increase. This shifts rightward the demand that s faces in market S and drives up the price in S.[22]

As exports into N increase with the increase in k, firm n will cut back production, but by less than the increase in imports to N, so the price in market N will fall (see Lemma 3 in Section 3). The increase in k makes it less attractive for n to allow the line to be congested into its market and just optimize along the residual demand. For line capacity greater than some level, which we call $\hat{k}$, firm n is better off acting more aggressively, which eliminates the passive/aggressive equilibrium. Also as k increases, eventually a point must be reached at which a pure-strategy unconstrained Cournot duopoly equilibrium can be supported (under Assumptions 1-4). This occurs when each firm weakly prefers the unconstrained Cournot equilibrium outcome to producing its optimal passive output response when the other firm produces its Cournot equilibrium quantity. Once again, we use k^* to represent the smallest k that can support the unconstrained Cournot equilibrium. As the line capacity approaches k^*, there are two possible outcomes depending on whether the largest line capacity that supports a passive/aggressive equilibrium is greater or less than the smallest line capacity that supports the unconstrained Cournot equilibrium.

Theorem 5. Consider asymmetric markets, N and S with monopoly producers n and s, respectively. Assume that $P_m^S < P_m^N$. As the capacity of a transmission line between the markets, k, increases from zero, one of two possible outcomes will obtain.

CASE 1:	$0 < k < \hat{k}$	passive/aggressive equilibrium exists
	$\hat{k} < k < k^*$	no pure-strategy equilibrium exists
	$k^* < k$	unconstrained Cournot equilibrium exists

[20]There are some asymmetric examples that exhibit equal monopoly prices and therefore produce only non-degenerate mixed-strategy equilibria as $k \to 0$. For instance if one market is just a scaled-up version of the other, the monopoly prices will be equal. Additionally, there are examples where the two markets have different shaped cost and demand functions but still have the same monopoly price. These too produce mixed strategies in the thin-line limit.

[21]We continue here to assume no line losses. Line losses will change the unconstrained Cournot equilibrium by lowering the output of the firm that would be a net exporter in a lossless framework and raising the output of the firm in the importing market.

[22]The markets illustrated in Figure 5 have different demands, while the firms illustrated have identical costs, so the unconstrained Cournot equilibrium is still symmetric.

or

CASE 2:	$0 < k < k^*$	passive/aggressive equilibrium exists
	$k^* < k < \hat{k}$	both passive/aggressive and unconstrained Cournot equilibria exist
	$\hat{k} < k$	unconstrained Cournot equilibrium exists

Proof. See the Appendix.

Case 1 is illustrated in Figure 6. After a certain line capacity is reached, there is a region where only mixed strategy equilibria exist up to the threshold line capacity that induces the producer in N, the higher-price market, to switch from producing its optimal passive output to playing its unconstrained Cournot responses. In Case 2, which we call the "overlap case" and which seems to arise in more asymmetric markets than does Case 1,[23] there is at least one pure-strategy equilibrium for all line capacities. Unlike the symmetric case, n is content to produce its optimal passive output when s is producing q_m^{s+}, s's profit-maximizing output when it is congesting he line to N. In fact, even when the line capacity is large enough that firm n would play an unconstrained Cournot response to s's unconstrained Cournot equilibrium quantity, n would still prefer to allow the line to be congested into its market if its opponent maintains its output at q_m^{s+}. Case 2 with a line size $k^* < k < \hat{k}$ is illustrated in Figure 7.[24] Of course, we would like to know when each of these cases will arise. A relatively simple rule can be a guide:

Theorem 6. Case 2 of Theorem 5 (the "overlap" case) will arise if and only if at $k = k^*$, the unconstrained Cournot equilibrium output of firm s is less than s's best response when firm n produces its optimal passive output, *i.e.,* $q_m^{s+} > q_c^s$.

Proof: Consider the threshold line size k^* which would support the unconstrained Cournot duopoly equilibrium. By definition, k^* is the line size at which (at least) one firm, call it firm n, is indifferent between the Cournot equilibrium and playing its optimal passive output. Assume that $k = k^*$, firm n chooses to produce its optimal passive output and firm s responds by increasing its output (from q_c^s, the south firm's unconstrained Cournot equilibrium quantity). An increase in firm s's output has no effect on firm n's profit from choosing it optimal passive output, since the line was congested anyway. But it lowers the profitability of firm n responding with an unconstrained Cournot best response, because the profit of one firm in a Cournot duopoly game is decreasing in the output of the other firm. Since firm n was previously just indifferent between producing its optimal passive output and its unconstrained Cournot best response, it would now strictly prefer to produce its optimal passive output. Thus, if firm s's response to firm n producing its optimal passive output is to increase its own output, this reinforces firm n's preference for the optimal passive output. Since both firms are then playing their best responses to the other, this passive/aggressive outcome would be a Nash equilibrium.

Next consider the case in which firm n chooses to produce its optimal passive output when $k = k^*$ and firm s responds by *decreasing* its output. A large enough decrease in the firm s's output would decongest the line. This would yield a decongested line with each firm producing less than its Cournot best response, which cannot be an equilibrium, by the arguments in the text

[23]By more asymmetric, we mean that, starting from separate markets, a proportionally larger amount of power would have to be shipped from one market to the other in order to equalize the prices across markets. This is a function of both the cost functions of the firms and the demand functions in the markets.

[24]In this case where two pure-strategy equilibria exist, there will also be a mixed-strategy equilibrium (Fudenberg and Tirole, 1991, pp. 479-480). We have not examined this equilibrium.

and the proof of Theorem 1. A small decrease in firm s's output would not decongest the line. This would have no effect on the n's profit from producing its optimal passive output, since the line remains congested. But it would increase firm n's profitability of responding to firm s with an unconstrained Cournot best response. Since firm n was previously just indifferent between the producing its optimal passive output and its unconstrained Cournot best response, it would now strictly prefer the Cournot best response. Thus, the passive/aggressive outcome would not be an equilibrium. *Q.E.D.*

Thus, although the assumption of asymmetry does lead to a qualitative change in our analysis, it does not fundamentally change the picture. Increasing line capacity gradually increases competition, and can lead through a region of mixed strategies to a standard Cournot duopoly equilibrium. With a sufficiently large line, the Cournot duopoly equilibrium exists, and with that or, possibly, a larger line, the Cournot duopoly equilibrium is the only pure-strategy equilibrium that exists.

5. The California electricity market

We now demonstrate empirically the practical relevance of these results by applying the analysis of the previous sections to a model of California's electricity market. In this section, we demonstrate how strategic use of transmission limits can increase congestion and how a modest increase in that capacity can greatly impact firm behavior and the resulting power flows. We examine the California market under the generation asset ownership structure at the time of deregulation, early in 1998.[25]

Congestion on the north-south transmission lines often divides the state into at least two distinct geographic markets. These two regions are connected through what is known as the Path 15 transmission constraint. As of 1998, Pacific Gas & Electric (PG&E) was the dominant supplier north of Path 15 with about 60% of generation capacity in the region and Southern California Edison (SCE) was the largest electricity producer south of Path 15 with about 45% of that region's capacity.[26] Peak demand in the south is more than double the peak in the north, but there was also much more generation capacity owned by smaller firms located south of this constraint.

In Section 4, we noted that the presence of price-taking fringe firms in either or both markets would not alter the analysis. This is also true if those fringe firms are in fact in other regional markets that are connected to one of the dominant firm's markets through a transmission path, as is the case in California. The state is connected with Mexico, the desert southwest, and the northwest states, though there is very little of the northwest's hydro-electric power available during the peak demand months of August through October. Essentially, the models of the previous sections are directly applicable to markets where the dominant-firm regions are connected to competitive regions, and to each other, via a radial or linear network.

Almost half of PG&E's capacity is hydro generation, which is most abundant in spring and early summer, and offers very little power in the fall and early winter, which is the same time that imports from the northwest (also mostly hydro generation) become much less available. For this reason, northern California imports substantial quantities of power from the south during the fall

[25]Since 1998, the incumbent utilities have been forced to divest most of their generation assets to several firms of approximately equal size in the state, which has reduced the concentration of ownership. Despite the divestitures, market power remains a concern in these markets (See Borenstein, Bushnell and Wolak, 1999).

[26]PG&E owned approximately 2,700 MW of generation capacity that is located south of the Path 15, most of which was from the Diablo Canyon nuclear plant, which is under a separate regulatory agreement. Borenstein and Bushnell (1999) explains how the output from these plants are treated as fringe production in the south.

and winter. As a result, Path 15 can become congested with flows from southern California into the north.

Most analysts agree that PG&E, in its 1998 form, had significant potential market power when the northern California region is isolated from its neighbors. The interesting question, however, is how often would this isolation arise? It is in addressing this question that the analysis of the previous sections comes into play. While it would be tempting to simply calculate the flows that would occur in the absence of the Path 15 constraint and then see if the constraint is violated, the model we have presented indicates that this would fail to capture the potential for strategic congestion of the line.[27]

Borenstein and Bushnell (1999) develop a Cournot firms-fringe firms simulation of the California and neighboring electricity markets using data on the production costs and capacities of all generators in the region. Here we adapt that model to examine the potential for congestion of Path 15. We consider only the two largest suppliers as of 1998, PG&E and SCE, to be strategic players, and treat the other electricity producers in California, as well as any out-of-state firms exporting power to California, as price-taking fringe firms.

We use a constant elasticity demand curve of the form $q(p) = xp^{-\varepsilon}$ with an elasticity of $\varepsilon = .1$. The constant x is set such that the demand curve in the peak hour intersects the relevant forecast price-quantity points.[28] We focus on two months, September and December. For each of these months, we simulated six demand levels starting with the peak hour, then the 150th highest demand level, and so on by increments of 150 down to the minimum demand level of the month. Each simulation finds a static Cournot equilibrium that is assumed to be independent of the outcomes of the other hours simulated.[29]

We first find the Cournot equilibrium production levels of the two large firms under the assumption that Path 15 has unlimited transmission capacity. Transmission capacity over this path is in fact about 3,000 MW in the south-to-north direction. An examination of the net flows over this path (Table 1) reveals that flows exceed 3,000 MW in only one hour of September, but surpass the limit in three hours of December. These are shown in italics in Table 1. If one ignored strategic incentives to congest the line, one would assume that the northern and southern markets are seldom isolated from each other in the month of September. In December, one would find that the line is congested during the three highest demand hours simulated.

As we have shown, however, analysis of the potential for congestion between these two markets must account for the incentives of the suppliers when they incorporate the transmission constraint into their profit calculations. We next calculated the profit made by PG&E if it produced its optimal passive output (causing south to north congestion along Path 15) and compared that profit to the profit that arises in the state-wide unconstrained Cournot equilibrium. This calculation reveals that in many of the hours examined, starting from the uncongested-line Cournot equilibrium, it is

[27]Some studies have relied upon historical information on the power flows over the various transmission paths (see Pace, 1996). However, these data were generated by firms operating in a tightly regulated environment. The incentives faced by these firms after the transition period to a deregulated industry are significantly different. It appears from the brief transition period that has occurred in California that transmission patterns have indeed changed substantially with deregulation.

[28]Borenstein and Bushnell (1999) includes a detailed description of the assumptions used in this analysis.

[29]Hydro releases were "scheduled" outside of the Cournot calculation according to a demand peak-shaving heuristic as described in Borenstein and Bushnell (1999). Bushnell (1998) examines the strategic use of hydro in this market.

profitable for PG&E to reduce its output and induce congestion over Path 15. Of the eight hours in Table 1 in which Path 15 would be uncongested without strategic behavior (those with flows below 3,000 MW), five exhibit congestion (shown in bold) when PG&E accounts for the possible strategic gains from congesting the line.

For this response by PG&E to lead to a pure-strategy passive/aggressive equilibrium, however, it is necessary that PG&E continue to want to produce its optimal passive output even after SCE (and all fringe firms) optimally respond to PG&E's passive output, and that the line remain congested. For this to be the case, (a) the resulting price in the north must be greater than the price in the south and (b) PG&E, when faced with SCE's optimal response to a passive, congestion-inducing output by PG&E, must not find it more profitable to increase output and decongest the path. Empirically, it is straightforward to test for these conditions. If they fail to hold, then only mixed-strategy equilibria are possible for a path of the particular capacity assumed. If they do hold, then a pure strategy equilibrium in which the path is congested from south-to-north does exist.

We found that only at the lowest demand level in December is it more profitable for PG&E to increase output and decongest the line in response to the "aggressive" output level by SCE than to continue to produce its optimal passive output. This indicates that, for a line capacity of 3,000 MW, there is no pure-strategy equilibrium for that hour.[30] In the other four hours in which PG&E prefers the optimal passive output to remaining in the unconstrained Cournot equilibrium, we found there to be a pure-strategy passive/aggressive equilibrium.

We now focus on one specific demand level, the 600th highest demand hour in December, in order to examine the impact of increased transmission capacity on this market. If there is unlimited transmission capacity along Path 15, south-to-north flows are 2,264 MW. With the actual Path 15 capacity of 3,000MW, PG&E can make higher profit in this hour if it reduces its output below its unconstrained Cournot levels and induces congestion.

For a range of transmission capacities of Path 15, we examined the conditions for a pure strategy passive/aggressive equilibrium. We found that not only does a pure-strategy passive/aggressive equilibrium exist when the line capacity is 3,000 MW, this result held for all capacities up to 3,835 MW, the threshold line size (k^*) that yields the unconstrained Cournot equilibrium during this hour. We can confirm that the overlap case (Case 2 in Theorem 5) exists for this demand level by comparing the unconstrained Cournot output of SCE to the optimal aggressive output of SCE when the line has a capacity of 3,835. In the uncongested case, SCE's optimal output is 2,953 MW, while its optimal aggressive output is 4,057 MW. Since $q_m^+ > q_c$ for SCE in this case, by Theorem 6, the asymmetric pure strategy equilibrium must exist for $k^* = 3,835$ and for all smaller capacity lines.

Table 2 contrasts the market outcomes for the unconstrained Cournot equilibrium that can obtain when the line capacity is 3,835 MW with those for the actual path capacity of 3,000 MW.[31] Total output in the state (including imports) increases from 27,114 MW under the passive/aggressive

[30]This therefore must fall into the non-overlap case, Case 1, in Theorem 5. We did not attempt to find the mixed strategy equilibria in these cases.

[31]In order to illustrate the *possible* welfare gains, this comparison focuses only on the unconstrained Cournot equilibrium when $k = 3,835$. There is also still a passive/aggressive equilibrium at this line capacity and a mixed-strategy equilibrium.

equilibrium with $k = 3,000$ to 28,121 MW under the unconstrained Cournot equilibrium with $k = 3,835$. Thus, the additional line capacity is about 83% of the resulting increase in output. Market clearing prices drop by 80% in the north when transmission capacity is increased and the unconstrained Cournot outcome obtains. The addition of 835 MW of capacity to Path 15 in this model can result in an increase in consumer surplus of over one million dollars, but most of this is a transfer from generation and transmission line owners. The change in total surplus is much smaller.[32]

It is also interesting to examine the market outcomes under the assumption that the market in southern California is perfectly competitive. If this were the case, even if Path 15 had a capacity of 3,835 MW, PG&E would reduce output and congest the path since the unconstrained Cournot outcome is no longer an alternative. If Path 15 were congested, the northern California market would have a price of $\$64.7/MWh$ as opposed to a price of $\$25.9/MWh$ when all firms but PG&E act as a price-taking fringe in a statewide market at this demand level. PG&E's profits are higher when the line is congested, even at a capacity of 3,835 MW. Thus, for Path 15 capacities in this range,[33] consumers in northern California may be better off if the southern California market is *less* competitive.

These results indicate that strategic congestion of transmission lines may play an important role in the forthcoming deregulated electricity markets. These effects would not be captured by the most widely used methods for estimating the scope of geographic markets in this industry that rely upon historical flow data or the simulation of a perfectly competitive regional market. Such methods could fail to recognize some of the most tangible benefits that might result from expansion of transmission capacity.[34]

6. Conclusions

Transmission constraints will be at the heart of market power issues in a restructured electricity market. In the absence of transmission constraints in the western U.S., it is less likely that any firm would have sufficient market power to significantly elevate prices. Transmission congestion, however, is likely to occur at some times. We have made a preliminary exploration into the market power issues raised by the central role of transmission and the resulting possible benefits to increasing transmission capacity. We have shown that the social value of transmission capacity may not be closely related to the actual flow that occurs on the line. We also have demonstrated how to analyze the impact of transmission capacity on competition. Our results indicate that expanding

[32]Consumer Surplus and Total Surplus are unbounded due to the use of a constant elasticity demand function, so the table shows only changes in these measures. PG&E's operations include hydroelectric power which we treated as having zero marginal cost. Transmission Rents are the difference between the north price and the south price multiplied by the flow on the line. We do not attribute the transmission rents to any of the generating companies.

[33]For a Path 15 capacity of less than 3,835 MW, additional competition in the south would have no effect on prices in the north since PG&E prefers to congest Path 15 anyway.

[34]The full analysis of the social value of a transmission path upgrade would require carrying out the exercise we have done here for every hour of the year. If the upgrade would not eliminate congestion in all hours, as would almost certainly be the case in the surplus-maximizing outcome, then it would also require calculating the outcomes in cases in which only mixed-strategy equilibria exist. Finally, after doing this benefits calculation, one would have to compare it with the cost of such an upgrade at various different possible upgraded capacities to see which generates the greatest net surplus. Also, the cost of such upgrades is idiosyncratic and often controversial. See Baldick and Kahn (1992).

transmission capacity between markets that suffer from market power problems may have very high payoffs in terms of reduced prices, increased consumption, and lower deadweight loss.

Our analysis, however, has considered only one-shot Nash equilibria with a single dominant firm in each regional market. In reality, the firms that compete in electricity markets will do so repeatedly and, thus, may be able to reduce rivalry through the threat of retaliation. To the extent that firms can reach more cooperative outcomes through such supergame strategies, the competitive effects of transmission lines, as well as most other remedies for the exercise of market power, are likely to be dampened.

Entry of new producers could also mitigate the effects of transmission that we identify here. If entry in the two markets is easy, then sufficient entry in each market could produce as great or greater competitive effects as increasing transmission capacity. It is important to note that entry would have to occur in both markets for it to replicate the competitive effect of a transmission line that we have demonstrated here. If, however, there are sunk costs of entry or fixed costs of operation, then the transmission capacity could be a less expensive route to achieving the benefits of increased competition.

Appendix

Proofs of Lemma 4, Theorem 3, and Theorem 5 follow.

Proof of Lemma 4. First note that for production by a rival of $q \leq q_m^+(k) - 2k$, the best response must be $q_m^+(k)$. The output of a rival for which the best response is the optimal passive output must therefore be greater than $q_m^+(k) - 2k$.

To show that q_{sw} exists and is unique in the range $[q_m^+(k) - 2k, \infty]$, we will show that $\Pi\left(\min(UBR[q_{sw}], q_{sw} + 2k), q_{sw}\right)$ (the left hand side of (3)) is a continuous decreasing function formed by the intersection of two functions that are tangent at the point $\tilde{q}$ where $UBR[\tilde{q}] = \tilde{q} + 2k$. This function is monotonically decreasing in q over the relevant range, has a value greater than $\pi(q_m^-(k), k)$ (the right hand side of (3)) at $q = q_m^+(k) - 2k$, and has a value less than $\pi(q_m^-(k), k)$ for some $q < \infty$.

By Assumption 3, $\Pi\left(UBR(q), q\right)$ is decreasing in q. Also note that $\Pi\left(q + 2k, q\right) = \pi(q + 2k, -k)$, which is a concave function with a maximum at the optimal aggressive output, where $q + 2k = q_m^+(k)$. The function $\Pi\left(q + 2k, q\right)$ is therefore decreasing in q for $q > q_m^+(k) - 2k$. Lastly, note that

$$\Pi\left(UBR(q), q\right) \geq \Pi\left(q + 2k, q\right)$$

for all q and that these two values are equal at $\tilde{q}$ where $UBR(\tilde{q}) = \tilde{q} + 2k$. These two functions are therefore tangent at this point, with the left-hand side of (3) equal to $\Pi\left(q + 2k, q\right)$ for $q \leq \tilde{q}$ and equal to $\Pi\left(UBR(q), q\right)$ for $q \geq \tilde{q}$. Since this tangency point must occur where both functions are decreasing, $\tilde{q} > q_m^+(k) - 2k$.

Lastly note that $\Pi\left(q + 2k, q\right) = \pi(q_m^+(k), -k)$ for $q = q_m^+(k) - 2k$. Therefore $\Pi\left(\min(UBR[q], q + 2k), q\right) > \pi(q_m^-(k), k)$ for $q = q_m^+(k) - 2k$ because the optimal aggressive output, $q_m^+(k)$, is always more profitable than the optimal passive output, $q_m^-(k)$. By assumption 3, $\Pi\left(\min(UBR[q], q + 2k), q\right) \leq \pi(q_m^+(k), -k)$ for a sufficiently large q. The left-hand side of (3), which by the above arguments is continuous and decreasing in q, is therefore greater than the right-hand side of (3) at $q = q_m^+(k) - 2k$ and less than or equal to the right-hand side for some $q < \infty$.

To show that $q_{sw}(k)$ is increasing in k, note that if $UBR(q_{sw}) \leq q_{sw} + 2k$, q_{sw} is defined by the equality

$$\Pi\left(UBR[q_{sw}], q_{sw}\right) - \pi(q_m^-(k), k) = 0$$

using the implicit function theorem, we can express the effect of line capacity on the switch point as

$$\frac{\partial q_{sw}}{\partial k} = -\frac{\frac{\partial}{\partial k} \quad \Pi\left(UBR[q_{sw}], q_{sw}\right) - \pi(q_m^-(k), k)}{\frac{\partial}{\partial q_{sw}} \quad \Pi\left(UBR[q_{sw}], q_{sw}\right) - \pi(q_m^-(k), k)} = -\frac{-\pi_2(q_m^-(k), k)}{\Pi_2\left(UBR[q_{sw}], q_{sw}\right)} > 0 \qquad (A1)$$

Note that both π_1 and Π_1 are zero at the best-response outputs given above and that π_2 and Π_2 are always negative.

If $UBR(q_{sw}) > q_{sw} + 2k$ then q_{sw} is defined by the equality

$$\Pi(q_{sw} + 2k, q_{sw}) - \pi(q_m^-(k), k) = 0.$$

Again applying the implicit function theorem, we have

$$\frac{\partial q_{sw}}{\partial k} = -\frac{2\Pi_1(q_{sw}+2k, q_{sw}) - \pi_2(q_m^-(k), k)}{\Pi_1(q_{sw}+2k, q_{sw}) + \Pi_2(q_{sw}+2k, q_{sw})} \tag{A2}$$

Note that the first term in the numerator represents the change in profit from producing the optimal aggressive output when the line capacity is increased. This must be positive since $q_{sw}+2k < UBR(q_{sw})$. An increase in k therefore allows this firm to increase output towards the optimal level of $UBR(q_{sw})$. The second term in the numerator, π_2, is, as before, everywhere negative. The numerator is therefore positive. To evaluate the sign of the denominator, first note that $\Pi(q, q) = \pi(q, 0)$, which is concave by assumption with a maximum at $q = q_m$. Because of this concavity, $\frac{\partial}{\partial q_{sw}}\Pi_1(q_{sw}+2k, q_{sw})$ must be negative for $q_{sw} > q_m$. For any non-zero k, q_{sw} must be greater than q_m since by Lemma 1 each firm would prefer to increase its output if the other firm produces q_m. Therefore the denominator is negative.

q_{sw} is therefore always increasing in k, whether $UBR(q_{sw})$ is greater or less than $q_{sw}+2k$. *Q.E.D.*

Proof of Theorem 3. In the presence of losses, the profit of firm s when the line is uncongested would be[35]

$$P(q^s - f - rf^2)\cdot q^s - C(q^s) \quad \text{for} \quad q^s < q^n \tag{A3a}$$

$$P(q^s - f)\cdot q^s - C(q^s) \quad \text{for} \quad q^s > q^n \tag{A3b}$$

and its condition for profit maximization would be:

$$[1+(1+2rf)\frac{\partial f}{\partial q^s}]P'(q^s - f - rf^2)q^s + P(q^s - f - rf^2) - C'(q^s) = 0 \text{ for } q^s < q^n \tag{A4a}$$

and

$$[1-\frac{\partial f}{\partial q^s}]P'(q^s - f)q^s + P(q^s - f) - C'(q^s) = 0 \text{ for } q^s > q^n \tag{A4b}$$

From equation (5) we can infer the derivative of the flow with respect to a change in q^s by applying the implicit function theorem:
For $q^s < q^n$,

$$\frac{\partial f}{\partial q^s} = -\frac{P'(q^s - f - rf^2)(1+2rf)}{P'(q^s - f - rf^2)(1+2rf)^2 - 2rP(q^s - f - rf^2) + P'(q^n + f)} \tag{A5a}$$

and for $q^s > q^n$,

$$\frac{\partial f}{\partial q^s} = -\frac{P'(q^s - f)}{-P'(q^s - f) - (1-2rf)^2 P'(q^n + f - rf^2) + 2rP(q^n + f - rf^2)}. \tag{A5b}$$

[35] We assume throughout this discussion that the condition $rf^2 \leq |f|$ is not binding. Losses are always less than the flow quantity.

Note that both profit and marginal profit are continuous: when $q^s = q^n$, (A3a) and (A3b) are equivalent. When $r = 0$, *i.e.,* there are no losses, at the point of a symmetric equilibrium ($q^n = q^s$) $\frac{\partial f}{\partial q^s} = 1/2$, so

$$\frac{1}{2}P'(q^s)q^s + P(q^s) - C'(q^s) = 0. \qquad (A6)$$

When $r > 0$, however, at the point of a symmetric equilibrium ($q^n = q^s$), $\frac{\partial f}{\partial q^s} = -\frac{P'(q^s)}{2P'(q^s)-2rP(q^s)} > -1/2$. In this case, (A3b) and (A4b) become

$$\phi P'(q^s)q^s + P(q^s) - C'(q^s) = 0 \qquad 1/2 < \phi < 1. \qquad (A7)$$

The presence of losses lowers the marginal revenue of each firm and, thus, at the no-losses symmetric Cournot equilibrium outputs, each firm would now, with losses, want to reduce its output. Thus, if a symmetric Cournot equilibrium exists, it would be at a lower quantity than in the absence of losses.

For sufficiently small r an equilibrium will, in fact, exist. To see this, note that, by Assumption 3, each firm's profits are concave in its own output when $r = 0$. By differentiating (A5) with respect to q^s, it can be verified that the second order impact of output on profits is continuous in r. Since this second order impact is negative when $r = 0$, for sufficiently small r, profits are concave for both $q^s < q^n$ and $q^s > q^n$. Since the derivative of the profit function is continuous, even at $q^s = q^n$, the profit function (A3) must be the inner envelope of two concave functions that define the two segments of (A3) and will therefore be concave. *Q.E.D.*

Proof of Theorem 5.

Below we prove by example the existence of each of these two cases. To prove that these two cases are exhaustive, we must show that (1) If no passive/aggressive equilibrium exists for a line of capacity $\tilde{k}$, then no passive/aggressive equilibrium exists for any line of capacity greater than $\tilde{k}$, and (2) if the unconstrained Cournot equilibrium exists for a line of capacity $\tilde{k}$ then the unconstrained Cournot equilibrium exists for any line of capacity greater than $\tilde{k}$. To prove (1), note that the passive/aggressive equilibrium ceases to exist when the passive firm (n is the discussion above) finds it less profitable to produce its optimal passive output than to act more aggressively, decongesting the line or congesting it in the opposite direction. Increases in k reduce the profit from producing the optimal passive quantity and weakly increase the profit from acting more aggressively, since acting more aggressively does not require the firm to congest the line in the opposite direction. Thus, if no passive/aggressive equilibrium exists for a line of capacity $\tilde{k}$, then no passive/aggressive equilibrium exists for any line of capacity greater than $\tilde{k}$. To prove (2), simply note that at k^*, at least one firm is indifferent between the unconstrained Cournot equilibrium and producing its optimal passive quantity in response to the other firm producing its Cournot equilibrium quantity. Increases in k beyond k^* have no effect on profit in the Cournot equilibrium and strictly decrease the profit of a firm from producing its optimal passive quantity. Thus, if the unconstrained Cournot equilibrium exists for a line of capacity $\tilde{k}$ then the unconstrained Cournot equilibrium exists for any line of capacity greater than $\tilde{k}$.

Examples of Asymmetric Market Equilibria

For an example of the first case in Theorem 5, where there exists a range of k for which no pure strategy equilibrium exists, consider two markets, N and S, where demands are $q^S = 10 - p$, and

$q^N = 12 - p$, and each market has one producer with cost function, $C^s(q) = C^n(q) = 0$. With no transmission line, the prices in S and N are 5 and 6, respectively, and the outputs are 5 and 6, respectively, for a total industry output of 11. The Cournot equilibrium quantities are $\frac{22}{3}$ for each firm with a total industry output of $\frac{44}{3}$ and price of $\frac{11}{3}$, and a net flow on the line of 1 from S to N. The profit of each firm in this equilibrium is $\frac{22}{3} \cdot \frac{11}{3} = \frac{242}{9}$ or about 26.89.

The northern firm's profit if a line of size k is congested into its market is

$$\pi^n(q^n(k)) = P^N(q^n + k)q^n = (12 - q^n - k)\, q^n.$$

and the first order condition implies that $q_m^{n-} = 6 - \frac{k}{2}$, where q_m^{n-} is q_m^- for firm n, giving profit from producing the optimal passive output of

$$\pi^n(q_m^{n-}(k)) = \left(6 - \frac{k}{2}\right)^2.$$

The line capacity, k, that leaves the northern firm indifferent between the unconstrained Cournot and the optimal passive output is therefore $(6 - \frac{k}{2})^2 = 26.89 \quad \Longrightarrow \quad k = 1.6291$.[36] For a line size slightly less than this, $k = 1.6$ for instance, the unconstrained Cournot equilibrium is not achievable; firm n would (just barely) prefer to produce the optimal passive output than to play its Cournot best response to s producing its unconstrained Cournot quantity. But, if n produces its optimal passive output, s will revert to selling its profit-maximizing quantity that congests the line, $q_m^{s+} = 5.8146$. This is less than s's Cournot quantity, 7.3333. As s reduces its output, producing its optimal passive output becomes less attractive to n. In this case, n will jump to producing its Cournot best response to 5.8146, which is 8.0927. With the line uncongested, however, s will then respond with its Cournot best response of 6.9536, and the process will once again iterate towards the unconstrained Cournot equilibrium. Since, the line is just slightly below the level that can support the unconstrained Cournot, as s's output approaches its Cournot equilibrium quantity, and strictly before it equals that quantity, n will once again revert to producing its optimal passive output. No pure strategy equilibrium exists.

For an example of the second case, consider markets with the following demands: $q^N = 15 - p, q^S = 15 - 4p$ and again one producer in each market with $C^n(q) = C^s(q) = 0$. If there is no transmission line, the monopoly quantities in the two markets are each 7.5 and the monopoly prices are $P^N = 7.5$ and $P^S = 1.875$. If there were a transmission line with no capacity constraint, then the Cournot problems for the two firms would be identical. In the resulting Cournot duopoly equilibrium, $q^n = q^s = 10$, $P = 2$ and each firm earns profit of $\pi = 20$.

We now examine the market outcomes when the transmission constraint is binding. If firm n produces its optimal passive output, its profit as a function of the line capacity k is

$$\pi^n(q_m^{n-}(k)) = \left[P^N(q^n + k)\right] q^n = (15 - q^n - k)\, q^n. \tag{A8}$$

The first order condition implies that $q_m^{n-} = 7.5 - \frac{k}{2}$, giving profit for firm n when it produces its optimal passive output of

$$\pi^n(q_m^{n-}(k)) = \left(7.5 - \frac{k}{2}\right)^2. \tag{A9}$$

[36] Note that the line of $k = 1.6291$ yields additional output of only 3.6666, so then line capacity is 44% of the additional output, in contrast to about 17% with linear demand and constant marginal cost in the symmetric case.

The line capacity, k, that leaves the northern firm indifferent between the unconstrained Cournot duopoly equilibrium and producing the optimal passive output is therefore $(7.5 - \frac{k}{2})^2 = 20 \implies k = 6.057$. [37] For a line size slightly less than this, $k = 6$ for instance, the unconstrained Cournot duopoly equilibrium is not achievable; firm n would (just barely) prefer to produce its optimal passive output than to play its Cournot best response in response to s producing its unconstrained Cournot quantity. At $k = 6$, however, there is a passive/aggressive equilibrium in which s produces $q^s = 10.5$ of which 6 is exported to N and firm n responds by producing $q^n = 4.5$.

At $k = 6.057$, the unconstrained Cournot duopoly equilibrium is feasible. From that equilibrium, if the north firm were to reduce its output from its Cournot quantity of 10 to its optimal passive output of $7.5 - \frac{6.057}{2} = 4.472$, the line would become congested with flow from S to N. Given that the line is congested, the optimum output for s would be $q_m^{s+} = \frac{15+k}{2} = 10.528$.

Unlike in Case 1, if firm n jumps from the unconstrained Cournot duopoly equilibrium to producing its optimal passive output, firm s's best response in this case is to *increase* its output from the quantity it produces in the unconstrained Cournot duopoly equilibrium. This reinforces n's incentive to produce its optimal passive output. Thus, when $k = 6.057$, the critical line size to support the unconstrained Cournot duopoly equilibrium, there is a also a passive/aggressive equilibrium. In this equilibrium, firm n produces $q^n = 4.472$, and firm s produces $q^s = 10.528$, the line is congested with flows from S to N of 6.057. The result is a price in the north of $P^N = 4.471$ and in the south of $P^S = 2.632$.[38] This sort of outcome with two pure-strategy equilibria is illustrated in Figure 7.

As the line size increases beyond $k = 6.057$, both the uncongested and the passive/aggressive equilibria are possible until the line expands to about $\bar{k} = 6.36$, the line size at which the passive/aggressive equilibrium disappears. At any line size above $\bar{k}$, firm n's profit from producing its optimal passive output, which congests the incoming line, are smaller than the profit of playing the unconstrained Cournot best response.[39]

Two pure-strategy equilibria exist for some line sizes in Case 2, but not in Case 1, because in Case 2, when n switches from the unconstrained Cournot equilibrium to producing its optimal passive output, firm s responds by increasing its own output. An increase in s's output decreases the profitability for n of playing its unconstrained Cournot best response, reinforcing its decision to produce its optimal passive output. In contrast, in Case 1, s's output fell in response to n producing its optimal passive output, thereby inducing n to deviate from producing its optimal passive output. This illustrates the necessary and sufficient condition expressed in Theorem 5. *Q.E.D.*

[37] Note that the line of $k = 6.057$ yields additional output of only 5, so the line capacity is 121% of the additional output it elicits.

[38] If the passive/aggressive pure-strategy equilibrium obtains, then total output is unchanged from the outcome with no transmission line (with linear demand and constant marginal cost). There would, however, still be a positive total surplus effect from building the line: output would be reallocated from s, where it has lower value on the margin, to n. This is reflected in the fact that the line causes P^S to increase and P^N to decline.

[39] The profit of n playing its unconstrained Cournot best response to the aggressive firm playing q_m^{s+} is $\pi = 18.66$. Comparing this to the profit from producing its optimal passive output $(7.5 - \frac{k}{2})^2$ and solving for k gives $\bar{k}$.

References

Baldick, R. and Kahn, E. "Transmission Planning in the Era of Integrated Resource Planning: A survey of recent cases." Working Paper LBL-32231, Lawrence Berkeley Laboratory, 1992.

Borenstein, S. and Bushnell, J.B. "An Empirical Analysis of the Potential for Market Power in California's Electricity Industry." *Journal of Industrial Economics*, Vol. 47 (1999), pp. 285-324.

—, —, and Stoft, S. "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry." POWER Working paper PWP-040, University of California Energy Institute, 1998.

—, —, and Knittel, C.R. "Market Power in Electricity Markets: Beyond Concentration Measures." *Energy Journal*, Vol. 20 (1999), pp. 65-88.

—, —, and Wolak, F.A. "Diagnosing Market Power in California's Deregulated Wholesale Electricity Market." POWER Working Paper PWP-064, University of California Energy Institute, 1999.

Brander, J. and Krugman, P. "A 'Reciprocal Dumping' Model of International Trade." *Journal of International Economics*, Vol. 15 (1983), pp. 313-321.

Bushnell, J.B. "Water and Power: Hydroelectric Resources in the Era of Competition in the Western U.S." POWER Working Paper PWP-056r, University of California Energy Institute, 1998.

California Independent System Operator. 1998 Transmission Reliability Report. Folsom, CA: California Independent System Operator, 1998.

Cardell, J.B., Hitt, C.C., and Hogan, W.W. "Market Power and Strategic Interaction in Electricity Networks." *Resources and Energy Economics*, Vol. 19 (1997), pp. 109-137.

Chao, H.P. and Peck, S. "A Market Mechanism for Electric Power Transmission." *Journal of Regulatory Economics*, Vol. 6 (1996), pp. 25-60.

Fudenberg, D. and Tirole, J. *Game Theory.* Cambridge, MA: MIT Press, 1991.

Green, R.J. and Newbery, D.M. "Competition in the British Electricity Spot Market." *Journal of Political Economy*, Vol. 100 (1992), pp. 929-953.

Joskow, P.L. and Frame, R.. "Report on Horizontal Market Power Analysis." PJM Supporting Companies Request for Market Based Rates. Federal Energy Regulatory Commission. Docket No. ER97-3729-000, 1997.

Klemperer, P.D. and Meyer, M.A. "Supply Function Equilibria in Oligopoly under Uncertainty." *Econometrica*, Vol. 57 (1989), pp. 1243-1277.

Krishna, K. "Trade Restrictions as Facilitating Practices." *Journal of International Economics*, Vol. 26 (1989), pp. 251-270.

Oren, S.S. "Economic Inefficiency of Passive Transmission Rights in Congested Electricity Systems with Competitive Generation." *The Energy Journal*, Vol. 18 (1997), pp. 63-83.

Pace, J."Market Power Analysis of Pacific Gas and Electric Company In Support of Joint Application", Federal Energy Regulatory Commission, Docket No. ER96-1663-000, 1996.

Schweppe, F., Caramanis, M., Tabors, R. and Bohn, R. *Spot Pricing of Electricity.* Boston: Kluwer Academic Publishers, 1988.

Tirole, J. "The Theory of Industrial Organization." Cambridge, MA: MIT Press, 1988.

von der Fehr, N. and D. Harbord. "Spot Market Competition in the UK Electricity Market," *Economic Journal*, Vol. 103 (1993), pp. 531-546.

Western Systems Coordinating Council. WSCC Transmission Bottleneck Study Report. Salt Lake City, UT: Publisher, 1997.

Wolak, F.A. and R.H. Patrick. "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market." POWER Working Paper PWP-047, University of California Energy Institute, 1997.

Wolfram, C.D. "Measuring Duopoly Power in the British Electricity Spot Market," *American Economic Review*, Vol. 89 (1999), pp. 805-826.

Table 1: Flows over Path 15 assuming unlimited capacity

Demand Level	Peak	150th	300th	450th	600th	744th
September Path 15 flow (S-N)	1,116	**2,736**	**2,841**	*3,198*	895	-1,167
December Path 15 flow (S-N)	*5,016*	*5,263*	*4,697*	**2,705**	**2,264**	**2,310**

Note: Italics indicate violation of transmission capacity when firms do not consider these limits. Bold indicates violation of transmission capacity only when firms incorporate these limits into their strategies.

Table 2: The impact of increased transmission capacity
(600th Highest Demand Hour in December)

Path 15 Capacity	3,000 MW	3,835 MW
Flow on Path 15	3,000 MW	2,264 MW
PG&E output	683 MW	3,448 MW
N. California Price ($\$/MWh$)	169.3	27.8
S. California Price ($\$/MWh$)	27.8	27.8
N. California Consumption (MWh)	8,443	9,450
S. California Consumption (MWh)	18,671	18,671
PG&E Profit	\$117,378	\$36,713
Industry Profit	\$1,138,776	\$390,164
Change in Consumer Surplus	—	+\$1,251,517
Transmission Rents	\$424,380	\$0
Change in Total Surplus	—	+\$78,525

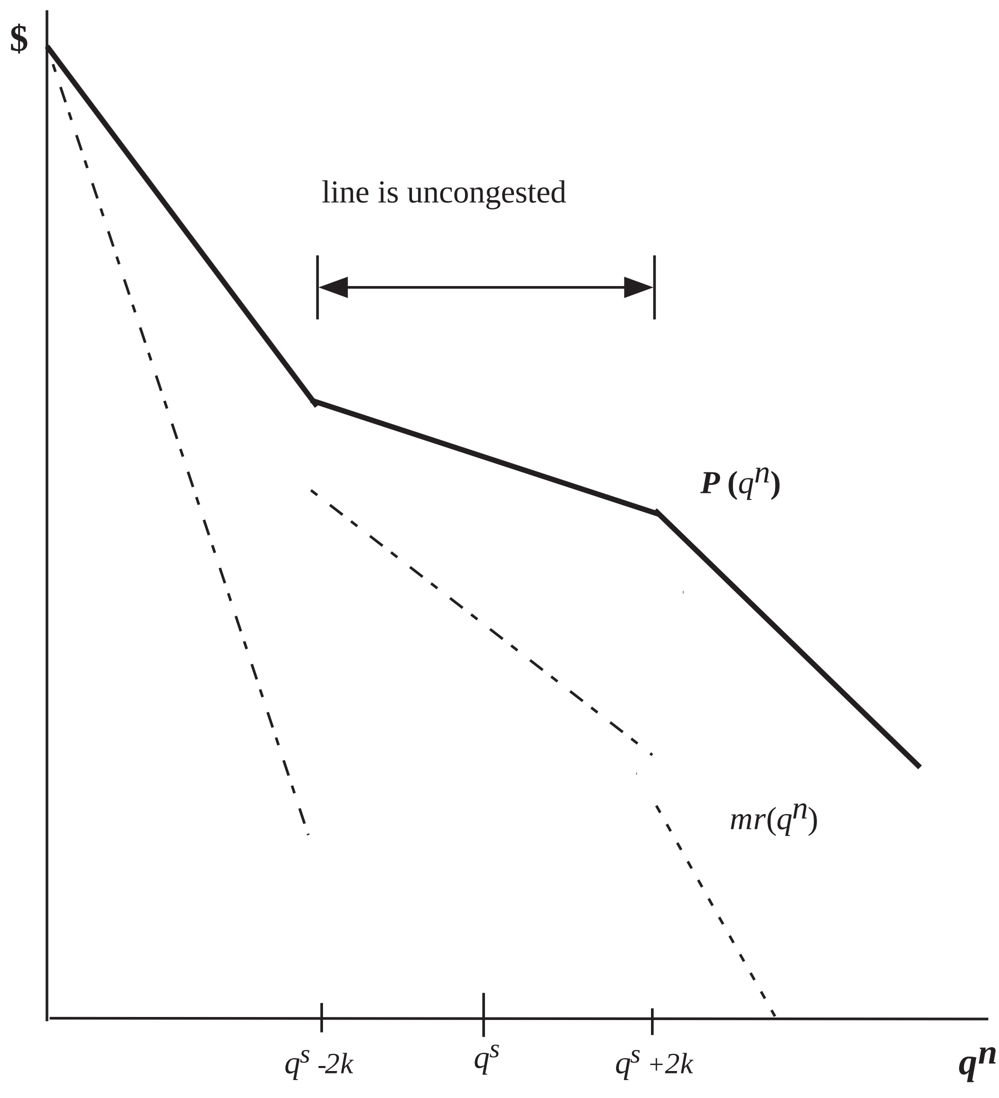

Figure 1: Demand curve faced by firm *n* when firm *s* produces q^S

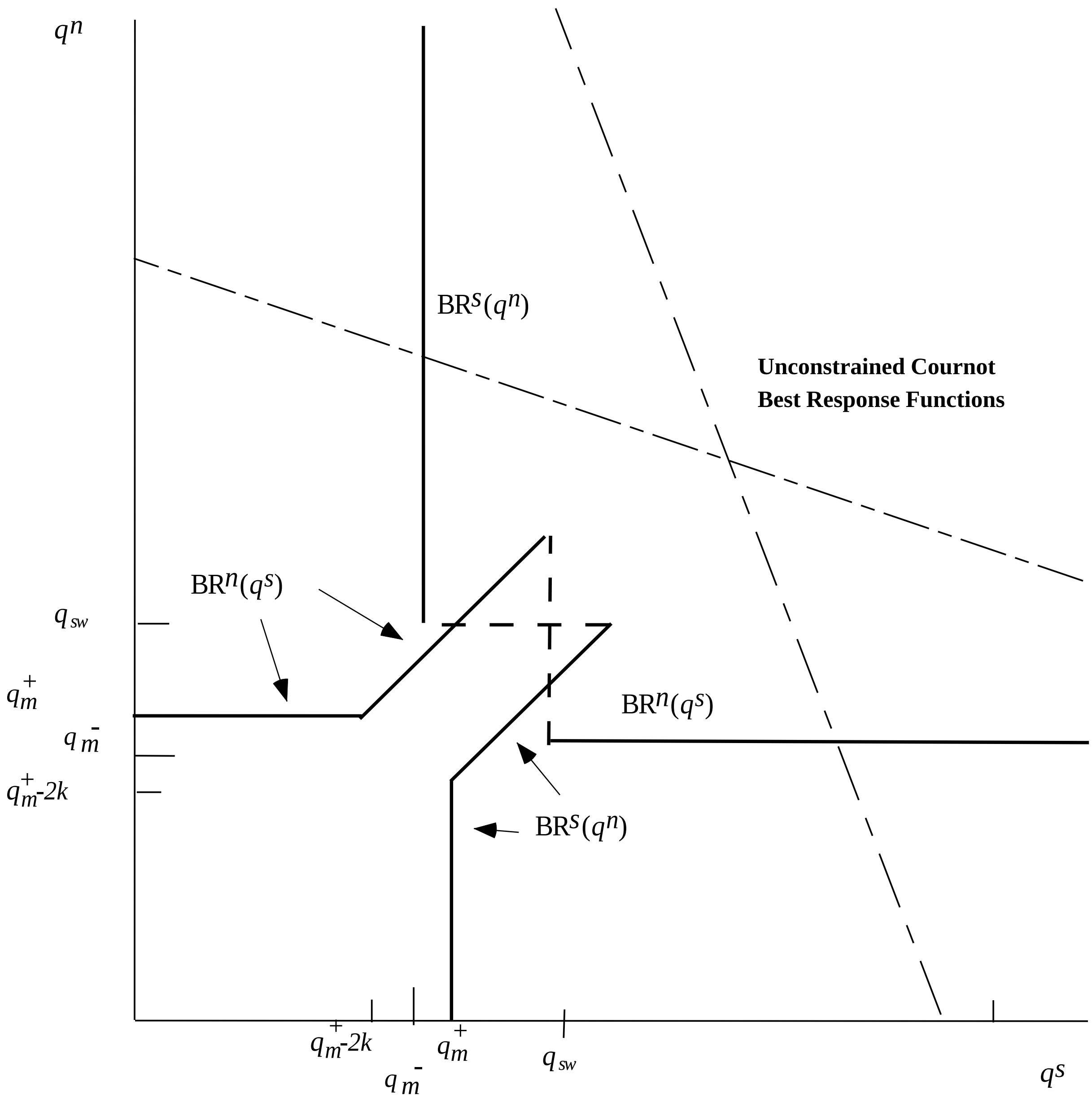

Figure 2: Best Response Function for Very Small Line

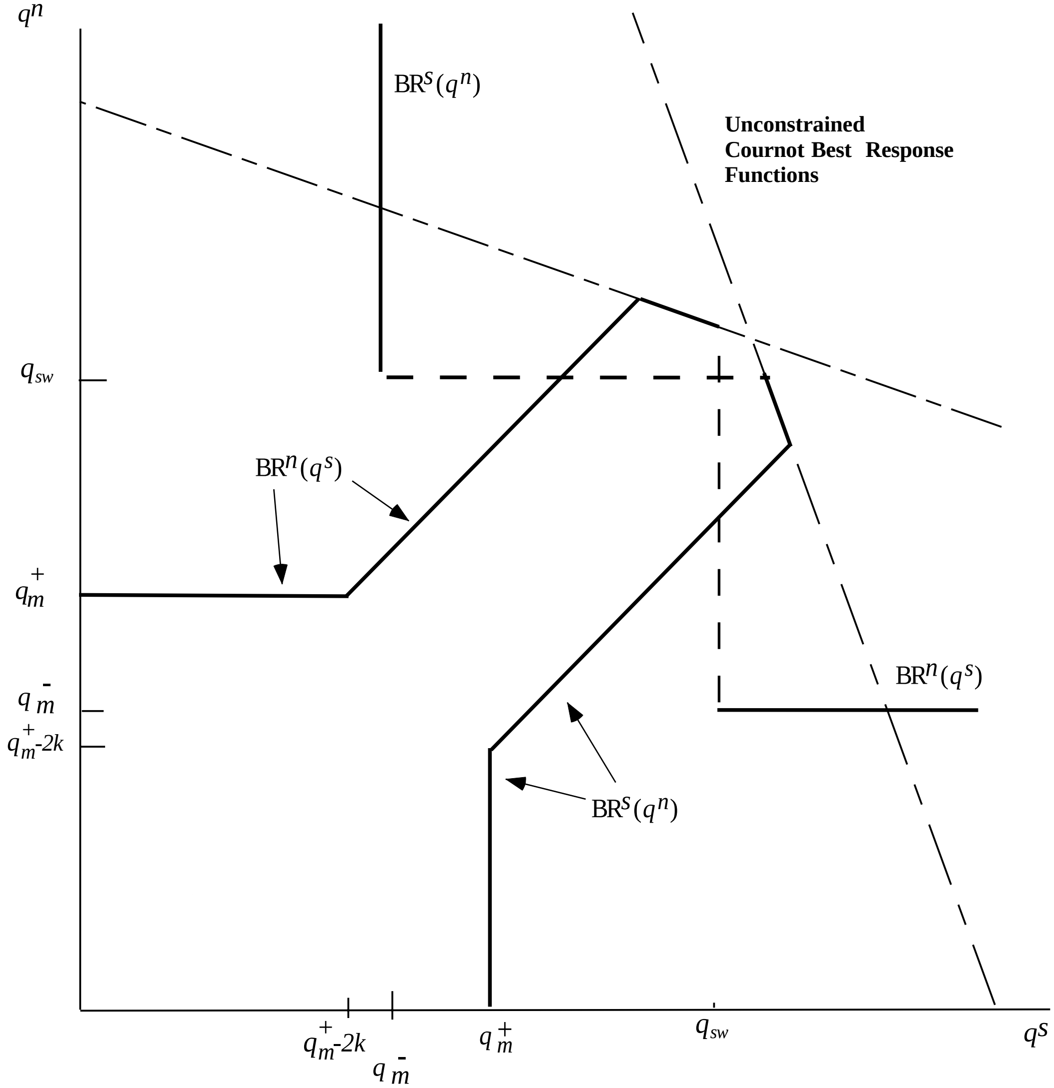

Figure 3: Best Response Function with Unconstrained Cournot Region

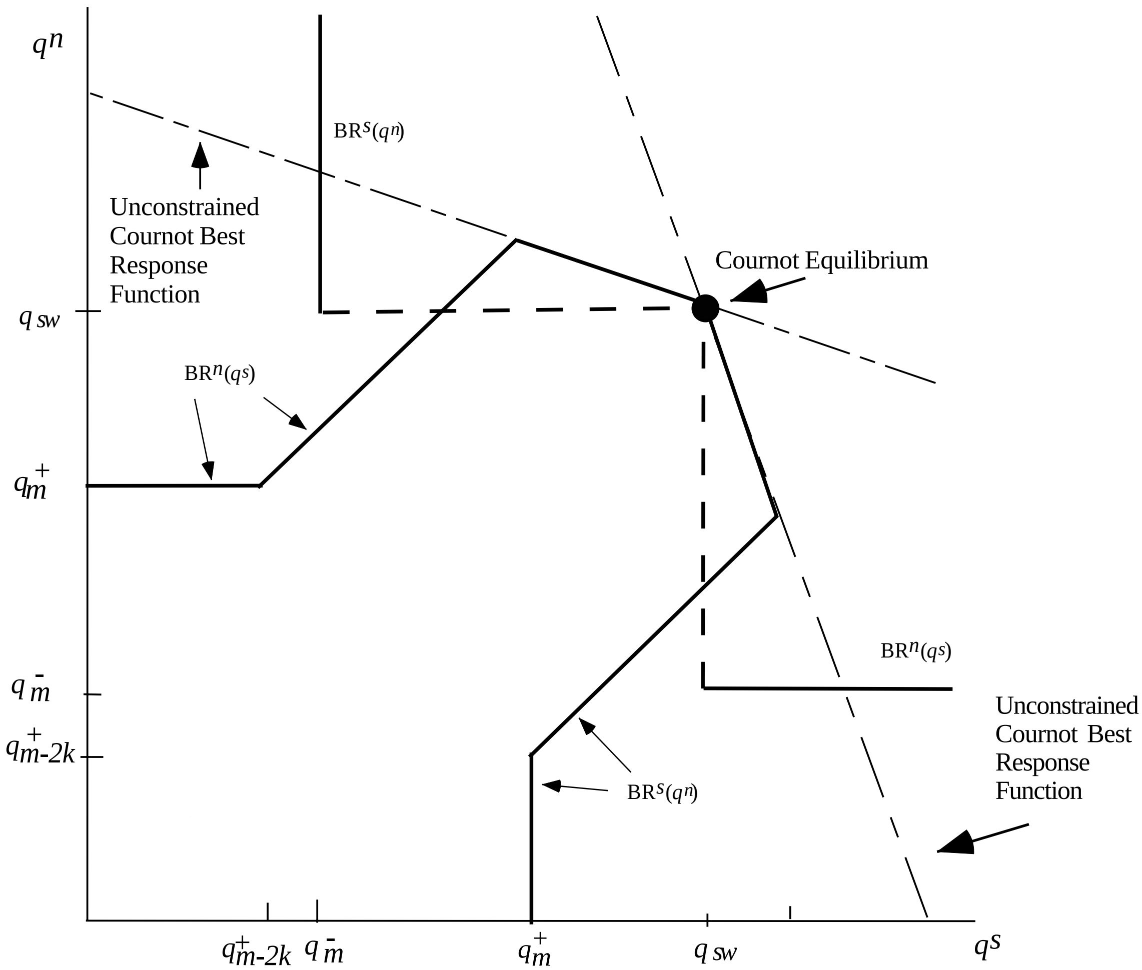

Figure 4: Line Capacity Allows for Unconstrained Cournot Equilibrium

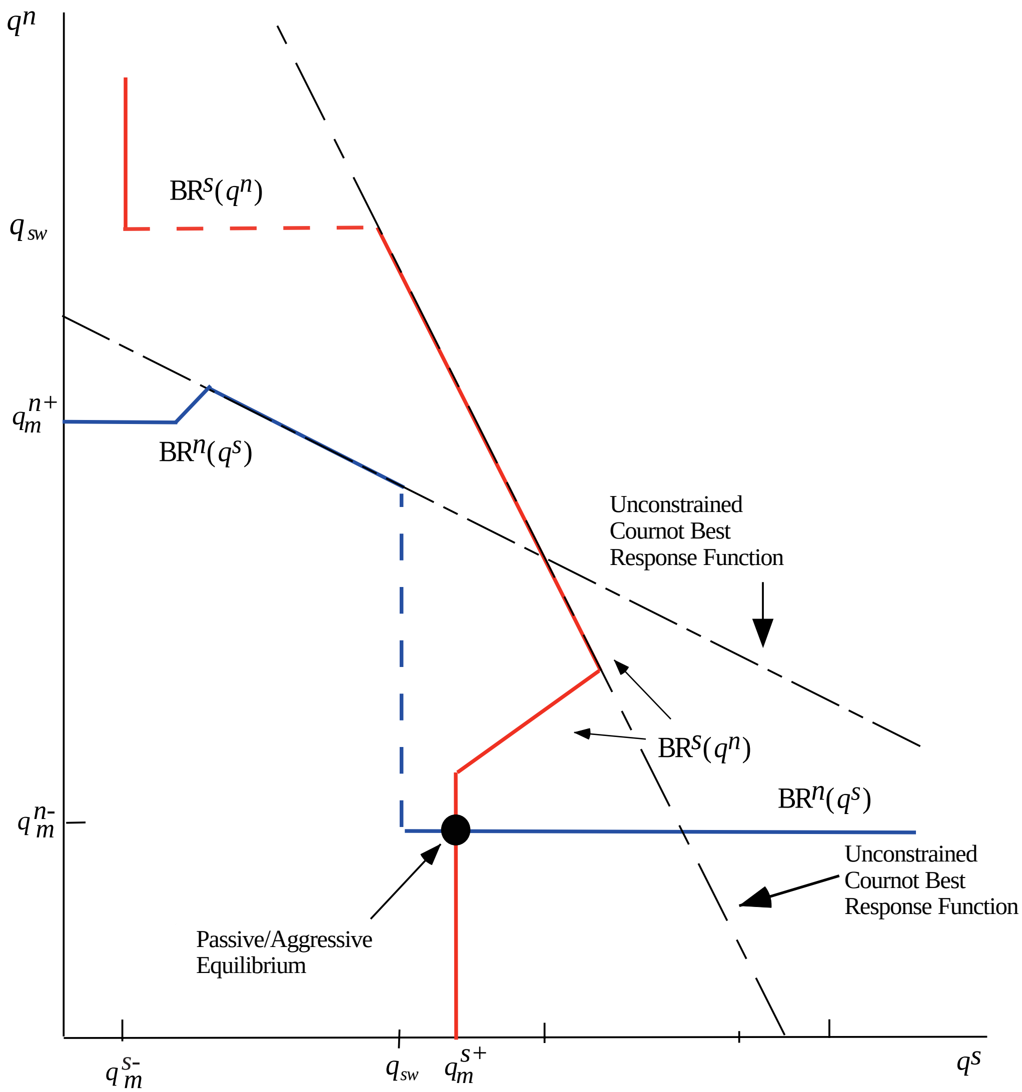

Figure 5: Asymmetric pure strategy equilibrium

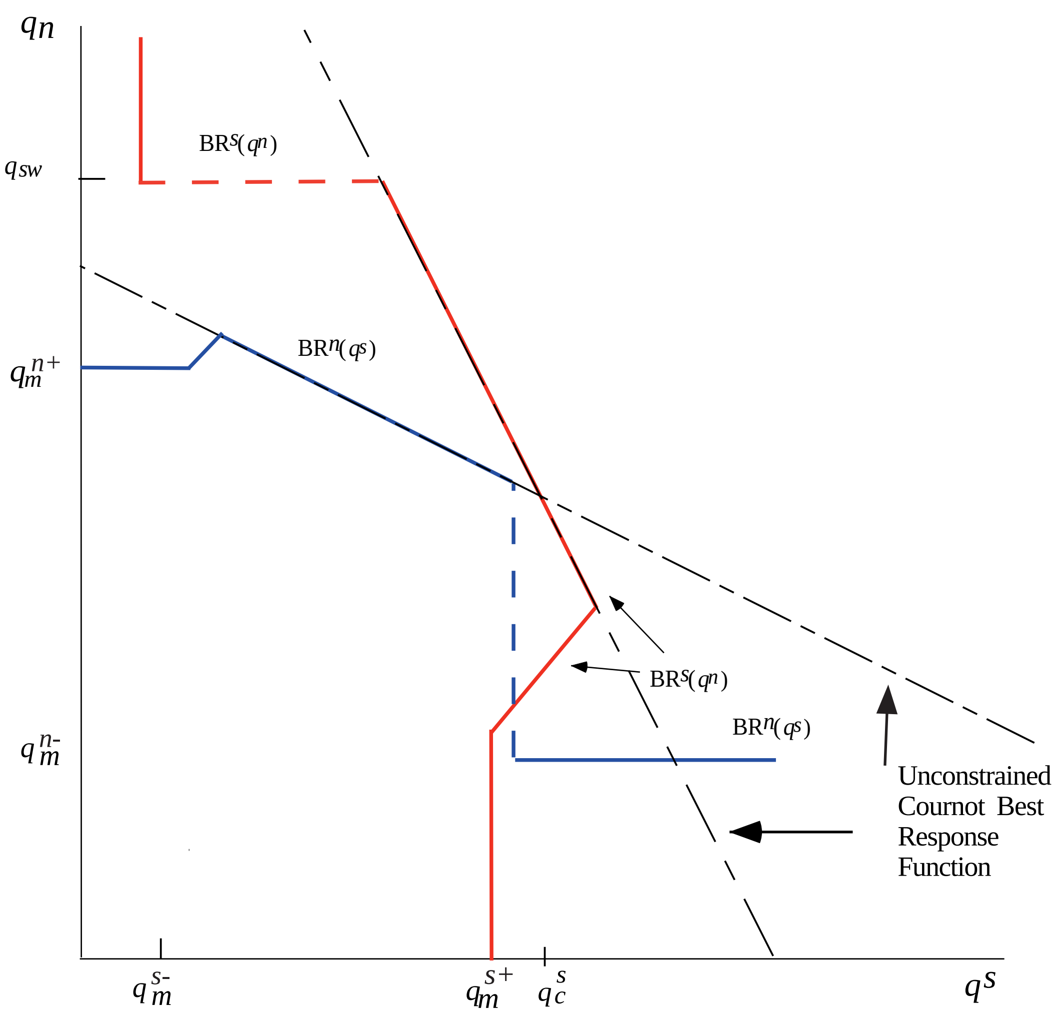

Figure 6: No pure-strategy equilibrium

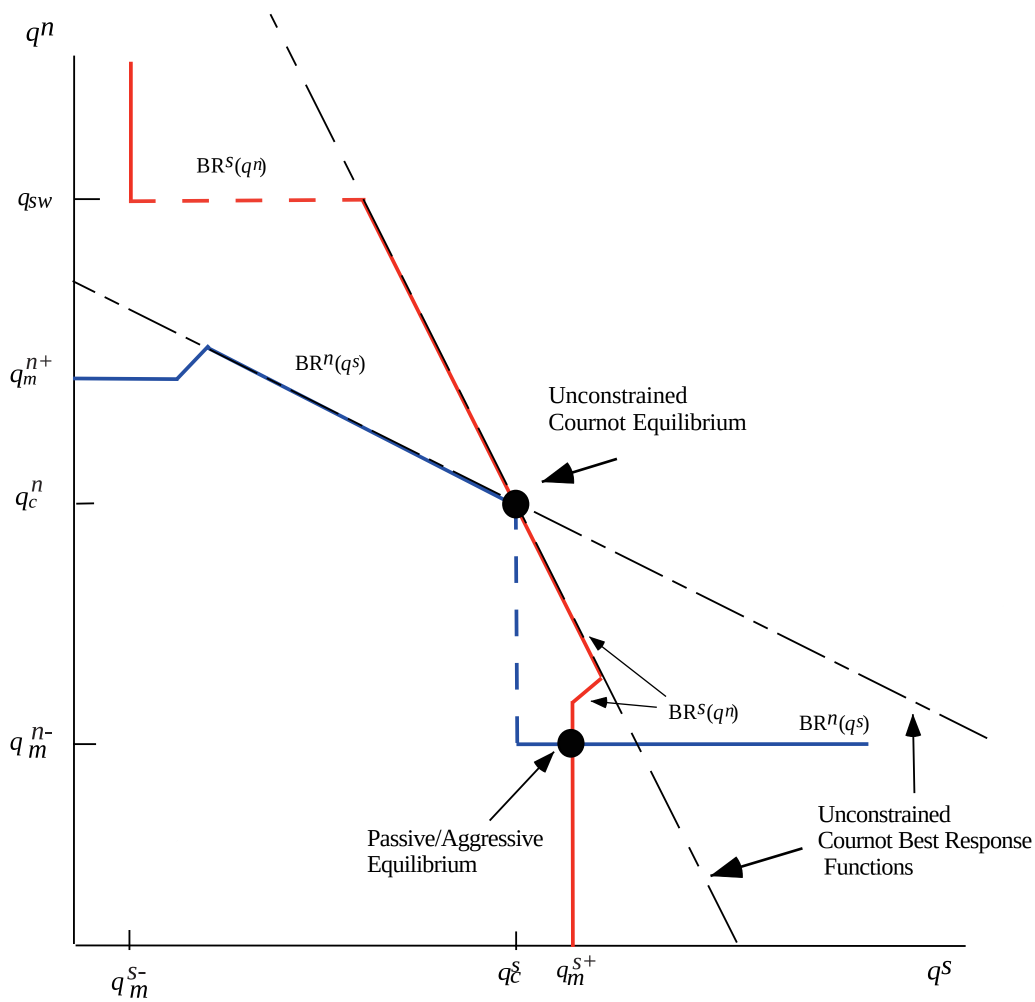

Figure 7: Case 2, 'overlapping' equilibria