Transmission Rights and Market Power on Electric Power Networks. II: Physical Rights

Paul Joskow * and Jean Tirole†

December 2, 1998

1 Introduction

In a companion paper (Joskow-Tirole 1998), we examined whether and how the allocation of financial transmission rights may enhance the market power of sellers or buyers of electricity in a deregulated electricity auction market with locational pricing that reflects congestion on the transmission network. In such a market, prices may differ from one location to another when the network becomes congested. Generators located on the "exporting" side of the congested link will receive a lower price than generators located on the "importing" side of the congested link. Moreover, the price consumers on the importing side of the congested link pay is higher than the price "exporting" generators receive for their supplies. The difference between the delivered price in the importing region and the price generators in the exporting region receive represents the cost of congestion. Generators located in the exporting region implicitly pay a transmission charge equal to this cost of congestion to sell their output to consumers in the importing region. A *financial* transmission right associated with a particular transmission link entitles the holder to a share of these congestion rents.

Focusing our analysis on a simple two-node electricity network, we found that when generators and consumers behave competitively at all locations, financial transmission rights do not affect the allocation of production or the prices paid by consumers for electricity compared

*Department of Economics, and Center for Energy and Environmental Policy Research, MIT.
†IDEI and GREMAQ (UMR 5603 CNRS), Toulouse, CERAS (URA 2036 CNRS), Paris, and MIT.

to the “no rights” competitive case. However, when electricity suppliers or consumers have market power we found that the allocation of financial rights could enhance that market power and reduce welfare. Specifically, we found that allocating financial transmission rights to a generator with market power located in the importing region enhances its market power. The more rights that the generator with market power in the importing region is allocated, the more its market power is enhanced and the higher is the delivered price of energy paid by consumers. The number of rights that will be allocated to the generator with market power through the rights market depends on the microstructure of the market for financial rights. In particular, it depends on whether the rights market is organized in a way that mitigates free riding by others on the increased congestion rents that a generator with market power can earn by increasing energy prices in the importing region to levels above their “no rights” level. Allocating rights to a generator with market power in the exporting region does not enhance its market power or affect delivered prices paid by consumers. This is the case because a generator with market power in the exporting region can already capture the scarcity rents associated with transmission congestion. Indeed, the presence of a generator with market power in the exporting region mitigates the market power enhancing effects of financial rights allocated to a generators with market power in the importing region.

We also found that financial rights can affect the behavior of electricity consumers as well. Financial rights allocated to a buyer of energy with market power (a monopsony) located in the importing region will reduce the buyer’s incentives to exercise such market power. On the other hand, allocating financial rights to a monopsony buyer of energy located in the exporting region would enhance its market power. The ultimate allocation of rights continues to depend on the microstructure of the rights market and the relative valuations of the rights by generators and consumers with market power in the energy market.

Our companion paper focused primarily on a simple two-node network with a single link connecting low cost generators located in the North with high cost generators and electricity consumers located in the South. However, we also provided a brief analysis of a three-node network. This allowed us to examine whether and how the introduction of the kind of “loop flows" that characterize many electric power networks affected these results. We found that our results are robust to loop flow considerations, although loop flow introduced some interesting twists to the analysis.

The purpose of this paper is to extend the analysis in our companion paper to consider *physical* transmission rights. Under a physical transmission rights system, the capacity of each of the potentially congested interfaces is defined and rights to use this capacity are created and allocated in some way to suppliers and consumers. A supplier of electricity that uses a congested transmission link must possess a physical right to have its supplies accepted for scheduling or "transportation" by the network operator. Once an electricity supplier has such a physical right, there is no additional charge for using the associated congested transmission interface. That is, a generator located in the exporting region which has acquired the necessary physical rights would be paid a price equal to the delivered price in the importing region for any net supplies it sells to consumers located in the importing region; there is no implicit or explicit congestion charge assessed on such a generator. Of course, the generator would have to cover the cost of acquiring its rights from the gross proceeds from such electricity sales. The physical rights are fully tradable and it is the market for these rights that determines the market clearing price of congestion.

The focus of this paper is identical to our companion paper. We focus on (a) how the allocation of physical rights may affect competition or enhance seller and buyer market power in the markets for electric generation when a transmission network is congested and (b) how rights markets with different microstructures allocate physical rights among generators and consumers and determine rights prices. As in the companion paper, most of our analysis examines a simple two-node network. However, we also provide a brief analysis of a three-node network which allows us to examine whether and how the introduction of the kind of "loop flows" that characterize many electric power networks affects our results. We find that our results are robust to loop flow considerations, although loop flow introduces some interesting twists to the analysis.

The results for physical rights are very similar to those that we obtained in our analysis of financial rights. There are two major differences. The first difference is that a physical rights system introduces the possibility that owners of transmission rights can withhold these rights from the market, effectively reducing the capacity of the congested interface. When this type of behavior is profitable, the withholding of rights leads to production inefficiency that does not arise in a financial rights system. Potential withholding problems leads naturally to the consideration of regulatory rules requiring "capacity release" by physical rights holders when they do not use their rights to schedule deliveries of electricity supplies. Accordingly, the paper contains a discussion of capacity release rules as an additional regulatory response to market power problems caused by the allocation of transmission rights. The second difference is that the assumed timing of the rights and power markets has implications for the market power-enhancing effects of rights initially allocated to generators or consumers with market power in the electricity market. Specifically, a physical rights system can create a Coasian durable goods situation that may limit the market-power enhancing effect of these rights on the behavior of some agents. From a welfare perspective (prices and supply efficiency), and ignoring other considerations affecting the costs and benefits of physical vs. financial rights, the choice between the two rights systems depends on both the nature of generation supplier and buyer market power and the microstructure of the rights market.

The paper proceeds in the following way. For the convenience of the reader, in Section 2 we review the basic attributes of the competitive electricity market and the two-node network that we worked with in our companion paper. The economic attributes of the system that we work with here are the same as in the companion paper, but we alter the institutional arrangements slightly to make them more compatible with a physical rights system. If the generation market is competitive, there is no buyer market power in the energy market, and the rights markets are competitive, then the allocation of resources is the same for a financial rights system and a physical rights system. (This result ignores any differences in transactions costs associated with operating the different types of transmission rights systems, something which is not the subject of these papers.) We also begin to develop the intuition regarding potential differences in their effects on market power between financial and physical rights systems. Section 3 contains an integrated analysis of the allocation of physical rights and their effects on market power under the same three rights market microstructures we examined in our companion paper. The analysis focuses on situations where the generator in the importing region is a monopoly and the generators in the exporting region are competitive, but includes some discussion of other market power configurations as well. Section 4 extend the analysis to a three-node network to see if loop flow considerations change our results in any significant ways. Section 5 discusses "capacity release" rules as a method for mitigating capacity withholding problems that may arise with a physical rights system. Section 6 presents a summary of our conclusions.

2 An electricity market with physical transmission rights in the absence of capacity release rules

We examine a restructured electricity sector that has on the supply side a number of unintegrated and unregulated generating companies that supply electricity. On the demand side there are consumers which can either be final end-use retail consumers or distribution companies that resell the electricity they acquire to end-use customers. There is an independent system operator (ISO) that is responsible for operating the transmission network reliably. Unlike the situation in our companion paper, the ISO does not operate an auction market that determines which generators will be scheduled and what the purchase and sale prices will be at different locations. Instead, generators enter into bilateral contracts with specific wholesale or retail consumers (including intermediary brokers and marketers) to supply their electricity needs. The generators in turn are required to submit "balanced schedules" to the ISO. Balanced schedules are schedules which match exactly the generation that is supplied to the network with what the customers on the other side of the bilateral contracts consume from the network. Generators are also required to possess *physical rights* to use the transmission network that match the supply schedules that they submit to the ISO. The ISO accepts all balanced schedules that come with matching physical rights attached to them and rejects any schedules that do not have these rights. There are no additional charges for congestion that the generators are required to pay when they schedule electricity supplies over a congested interface (the ISO earns no merchandising surplus from these transactions). Finally, there are markets for physical rights, but the ISO plays no role in these markets.

We also assume that the ISO runs a real time spot "balancing market" that it relies on to deal with imbalances between individual suppliers' commitments to supply their customers' demand under bilateral contracts and the actual generation supplies that they deliver to the network. "Involuntary" imbalances may arise as a consequence of unanticipated outages of generating or transmission capacity or deviations from anticipated demand. "Voluntary" imbalances may arise as a result of a conscious decision to under or overschedule supply or demand. In this paper, the balancing market and the potential for voluntary schedule imbalances are relevant to our discussion of capacity release rules in Section 5 and we discuss both in more detail there. (Involuntary imbalances play no role here since we ignore uncertainty.)

As in our companion paper, we work initially with a simple two-node network (no loop flow) where there are a set of low-cost generators (G_1) in the North which produce output q_1 and have an aggregate cost function $C_1(q_1)$ with $C_1' > 0$ and $C_1'' > 0$. We focus here on situations where the generators in the North behave competitively. There is no (net) demand in the North and we refer to the North as being either the upstream location or the exporting region. In the South, there are electricity consumers and a set of generators (G_2) that have higher production costs (within the relevant range) than do the generators in the North and produce q_2. We refer to the South as the downstream location or the importing region. We initially assume that the generators in the South behave competitively as well, but most of the paper focuses on cases where G_2 is a monopoly. Consumers in the South have a demand function $Q = q_1 + q_2 = D(p_2)$ with $D' < 0$ and where p_2 is the delivered price in the South. Since this system operates under bilateral contracts between generators and consumers, generators located in both the North and the South are paid the same price for the electricity they supply, which is the delivered price p_2 in the South.

Finally, there is a transmission line linking the North and the South which has a fixed capacity equal to K. We assume that the fixed costs of this line and any ongoing operation and maintenance costs that do not vary with utilization are recovered separately from consumers in lump sum charges and we do not consider these costs further in our analysis. We ignore thermal losses on the network[1] and focus on situations where demand is sufficiently high that it cannot all be fully served by generators in the North because the transmission capacity constraint (K) is binding. That is, some supplies from the less efficient generators in the South are required to balance supply and demand at the competitive prices. Thus, the marginal cost of generation in the North must be lower than the marginal cost of generation in the South when the transmission capacity constraint is binding.

Generators in the North must have physical rights to schedule their generation pursuant to their bilateral contracts with consumers in the South. Since the transmission capacity is a binding constraint, rights to use it have a market value (η) that is greater than zero. The net price p_1 (net of the cost of physical rights) generators in the North receive for their generation supplies is then simply the difference between the price they are paid by their customers in the South (p_2) minus the market value of the physical transmission rights they need to deliver it (η). (Alternatively, we can think of p_1 as the price of generation produced and delivered in the North. Retail marketing intermediaries would acquire the generation in the North, acquire the necessary transmission rights in the rights market at a market price η, and schedule the supplies with the ISO for delivery to their customers in the South who pay a delivered price p_2.) Generators in the South do not need physical transmission rights since they do not use the transmission line as a result of their proximity to consumers and receive both a gross and net price equal to the delivered price in the South p_2.

When the energy and rights markets are perfectly competitive the equilibrium conditions are as follows:

[1]Thermal losses increase with the usage of the line in a quadratic fashion. Thus if thermal losses are not negligible, a charge must ex post be added for the usage of the line, that depends on total usage of the line.

$$p_2^* = C_2'(D(p_2^*) - K),$$

$$p_1^* = p_2^* - \eta^*,$$

$$p_1^* = C_1'(K),$$

$$p_2^* > p_1^*,$$

$$\eta^* = p_2^* - p_1^*,$$

$$D(p_2^*) = q_1 + q_2 = K + q_2.$$

These are the same equilibrium conditions that emerged under perfect competition with a financial rights system in our companion paper. The price of physical and financial rights are the same, the delivered price in the South is the same and the price received by generators in the North net of the cost of physical rights is equivalent to the nodal price in the North derived under bid-based dispatch and nodal pricing in our companion paper. This verifies for our model the more general result (due to Chao-Peck 1996) concerning the equivalence of financial and physical rights when the energy and rights markets are perfectly competitive.

Why might there be *any* differences between a financial rights system and a physical rights system?[2] One potential difference is that unlike the case with financial rights, the operation of the physical rights market might lead to a reduction in the effective capacity of the transmission link. Since generators in the North must have rights to use the transmission link, the rights that they acquire effectively defines the capacity of the link (up to K). If the market leads to an allocation where generators at the cheap node (the North) do not end up holding all of the rights (K) and cannot (or do not) use all of the capacity (K) available on the link, then the supply of "cheap" power from the North available to meet demand in the South would be reduced. In this case, supply from the cheap generators in the North (q_1) will be restricted ($q_1 < K$) and more demand than is necessary will be satisfied with expensive power from the South. Thus, "withholding" of rights from generators in the North results in production inefficiency since expensive power from the South is substituted for cheaper power in the North.[3] Moreover, if a

[2] We ignore here additional complications that may arise when there are more than two nodes and uncertainty about supply and demand conditions.

[3] We note that production inefficiency in this sense is not possible in Bushnell (1998)'s model since he assumes

generator in the South with market power controlled the physical rights, it would potentially have two instruments available to exercise market power: contracting supply in the South to increase the value of rights and withholding physical rights in order to contract the supply from the North. Thus, the combination of market power over generation in the South and a physical rights system could lead to both production inefficiency and downstream prices that exceed the marginal cost of the downstream supplier.

Whether or not any difference between financial and physical rights will occur depends on both the willingness to pay of different stakeholders and the microstructure of the physical rights market. We now proceed to analyze these issues in more detail.

3 Physical transmission rights and market power

3.1 Physical rights initially held by a single nonstakeholder owner (no free riding)

Consider first the situation in which a non-stakeholder owner initially owns all of the physical rights and there is a single monopoly generator G_2 in the South. It is optimal for G_2 to acquire the rights in order to avoid non-internalized externalities between the two players. In essence, G_2 then produces electricity in two ways: first, by selling rights to generators in the North or by purchasing power from them and then keeping the rights to dispatch the power produced in the North, and, second, by producing power in the South.

G_2 obtains the maximum profit by importing power from the North and reselling this power together with its own power to the consumers. In contrast, if G_2 first sells $q_1 \leq K$ rights to generators in the North and then chooses its own production q_2 in the South, G_2 does not internalize in the latter decision the change in value of the rights sold earlier in the rights market. Because G_2 then sells power in two stages, it tends to overproduce in the electricity market, in the same way Coase's durable good monopolist floods the market after having previously sold. The standard solution to Coase's durable good problem is leasing or vertical integration, which here corresponds to G_2's purchasing power in the North and thus keeping

that the marginal costs of the generators are the same at each node. We examine additional differences with Bushnell's model below.

an exclusive relationship with the consumers in the South.

We are thus led to consider two cases:

Commitment: G_2 imports power $q_1 \leq K$ from the North and sells $q_1 + q_2$ to consumers in the South.

Noncommitment: G_2 cannot resell power produced in the North (say, because competition policy prohibits it). It sells $q_1 \leq K$ rights to producers in the North, who contract with consumers; G_2 cannot commit to a level of production q_2 in the South when selling rights to generators in the North.

3.1.1 Commitment

As we discussed, G_2's preferred outcome is obtained when it imports q_1 units (at price $C_1'(q_1)$ each) from the North, or, equivalently, when G_2 simultaneously sets a price p_2 for power in the South and a price η for rights (which it acquired earlier from the non-stakeholder owner). These two prices determine (in the relevant range) a quantity $q_1 \in [0, K]$ flowing through the congested interface, with

$$p_2 - C_1'(q_1) = \eta.$$

G_2's profit is

$$\max_{\{p_2\,,\,q_1\}} \left\{ p_2 \left[D(p_2) - q_1\right] - C_2\left(D(p_2) - q_1\right) + \left[p_2 - C_1'(q_1)\right] q_1 \right\}.$$

Note that G_2 is a "gatekeeper" for production in the North when it controls all of the physical rights. It is both a monopsonist and a monopolist. It sells its own power and then it "outsources" to G_1 as well. So, G_2 faces a "make or buy" decision:

$$\begin{aligned} &\text{either} \quad q_1 = K \\ &\text{or} \quad q_1 < K \text{ and } C_2'(D(p_2) - q_1) = C_1'(q_1) + q_1 C_1''(q_1). \end{aligned}$$

The term on the left of the latter equality is the marginal cost of (internal) production in the South and the term on the right is the " virtual marginal cost" (external) production in the North or the "perceived marginal cost" of G_2. In this case, G_2 finds it optimal to substitute expensive supplies from the South for cheaper supplies from the North in order to extract some inframarginal rents from the cheap generators in the North.

Accordingly, when $q_1 < K$, p_2 will be higher than in the case where there is no generator market power at either node both as a result of withholding rights and as a result of the contraction of output in the South given q_1.

3.1.2 Noncommitment

Now assume that G_2 sells rights to generators in the North and cannot commit on its own production when selling these rights. One may have in mind that the rights market operates first and then the power market (day- ahead or hour-ahead) operates given the distribution of physical rights arrived at the first stage. But the two markets can be simultaneous as long as G_2 is not able to demonstrate its own level of production when selling rights to generators in the North.

Power market: In the electricity market, G_2 takes q_1 as given and sets $p_2 = \hat{p}_2(q_1)$, where

$$\hat{p}_2(q_1) \text{ maximizes } p_2[D(p_2) - q_1)] - C_2(D(p_2) - q_1).$$

Rights market: In the first stage G_2 sells $q_1 \leq K$ rights so as to maximize:

$$\max_{q_1} \{\hat{p}_2(q_1)[D(\hat{p}_2(q_1)) - q_1)] - C_2(D(\hat{p}_2(q_1)) - q_1) + [\hat{p}_2(q_1) - C_1'(q_1)]q_1)\}$$

given the function $\hat{p}_2(q_1)$.

Using the envelope theorem, the derivative of the latter objective function is equal to:

$$C_2'(q_2) - [C_1'(q_1) + q_1 C_1''(q_1)] + [d\hat{p}_2(q_1)/dq_1]\, q_1.$$

The third term, which is nonpositive, does not appear in the equivalent condition for the commitment case. It equals the change in value of the physical rights as downstream prices change.

3.1.3 Comparison of the commitment and noncommitment cases

The easiest case to examine is where there is constant returns to scale in the South (C_2' is constant). In this case, G_2 withholds "weakly" more physical rights in the noncommitment case.[4] This is a standard conclusion of "Coasian dynamics". Yet we cannot conclude that the commitment case dominates from a welfare point of view. It dominates from a production efficiency point of view, in that production from the cheap node is weakly greater than it is in the noncommitment case. However, it does not dominate from a "market power" or downstream pricing point of view. In the noncommitment case, in the energy market G_2 ignores the effects of its production on the value of rights and "floods the market" to maximize profit given the output in the North which is defined by the rights that G_2 has sold in the first stage.

Example: Assume constant returns to scale in the North and the South ($C_1'' = C_2'' = 0$, $C_1' < C_2'$) and linear demand [$D(p) = 1 - p$]. Our analysis shows that there are no withholding under commitment. If $C_2' - C_1' > K/2$, then there are no withholding under noncommitment either.

Thus, if $C_2' - C_1' > K/2$, then noncommitment dominates since there is no withholding in either case and downstream prices p_2 would be lower under noncommitment. Under these conditions physical rights would dominate financial rights since the physical rights do not provide an additional incentive to G_2 to contract output in the energy market and, here, like financial rights do not generate withholdings and production inefficiency. We return to the comparison of the two systems shortly.

Remark: The motivation for and implications of physical rights withholding here are different from those in Bushnell's (1998) recent paper. Bushnell assumes that production at both nodes is equally efficient and that both exhibit constant returns to scale. The only role that the transmission link plays in this model is to mitigate the generator's local market power at one of the nodes. This assumption seems less realistic than those made here. Historically, major interregional transmission lines were built to bring electricity from areas where it is cheap to produce to areas where it is more expensive to produce. They were not built to mitigate local

[4] "Weakly" comes from the fact that there may be corner solutions at $q_1 = K$. If $q_1 < K$ in the noncommitment case, then "strictly" is correct.

market power problems. In the future, in restructured electric power sectors, transmission lines are expected to be built for the same reasons, unless regulators do not have other instruments available to mitigate local market power problems. Also, the mechanism here is different from that in Bushnell (1998). Bushnell assumes that there is a lot of capacity K on the link and that it is competition from generators using this link that keeps G_2 from exercising market power. Here we have assumed that capacity K is limited even when there is no generation market power and does not prevent the exercise of market power by G_2. Withholding of physical rights is motivated by the desire to extract rents from G_1.

Physical vs financial rights: What can we say about whether financial or physical rights dominate from a welfare perspective in the case where there is a monopolist at the expensive node and no free riding in the rights market? Clearly, financial rights dominate from a production efficiency perspective, since it is only with physical rights that the allocation of rights can reduce output from generators at the cheap node and substitute output from the expensive node for it. Financial rights also dominate physical rights with commitment from an overall welfare perspective: G_2 maximizes $p_2D(p_2) - C_2(D(p_2) - q_1)$, whereas under financial rights G_2 maximizes $p_2D(p_2) - C_2(D(p_2) - K)$. This implies that p_2 is higher with physical rights unless $q_1 = K$ or $C_2'' = 0$. So, if $q_1 < K$, financial rights dominate physical rights with commitment from a total welfare perspective. Financial and physical rights with commitment are equivalent when $q_1 = K$.

The welfare comparison of financial and physical rights under noncommitment is less clear. There is a potential tradeoff between production inefficiency in the North and market power in the South associated with the choice between physical and financial rights. On the one hand, physical rights can lead to production inefficiency by reducing supply from the North while financial rights do not. Reducing the supply from the North also increases prices in the South compared to the competitive case. On the other hand, the effect of Coasian dynamics with physical rights leads to less market power and lower prices in the South given a value for q_1. So, unlike the commitment case, physical rights *may* dominate financial rights from an overall welfare perspective.

3.1.4 Other types of market power in the energy market

a) When there is a *monopolist at the cheap node*, competitive suppliers in the South and there is no buyer market power, the monopolist in the North acquires all rights from the non-stakeholder owner. G_1 may withhold some rights and use the others to dispatch its supplies. Note, though, that in the case of constant marginal cost in the South there is no withholding by G_1 at all ($q_1 = K$); for, if G_1 signs q_1 contracts with consumers in the South, then competitive behavior of generators in the South yields p_2 such that

$$p_2 = C_2'(D(p_2) - q_1), \tag{1}$$

and G_1 chooses q_1 so as to maximize

$$q_1 p_2 - C_1(q_1),$$

where p_2 is given by (1). So, *if* $q_1 < K$ then

$$C_2' - C_1' = q_1[C_2''/(1 - C_2''D')].$$

This is impossible if $C_2'' = 0$. In contrast, if marginal cost is upward sloping in the South it may be profitable for G_1 to withhold output ($q_1 < K$) in order to raise the price in the South. Note, though, that physical rights are not needed by G_1 to implement this withholding strategy. Under financial rights, G_1 captures congestion rents ($p_1 = p_2$) and schedules the profit maximizing value of q_1 with the ISO.

b) Consider now the case where there is *buyer market power* (a monopsony) *in the South*. Recall that in the case of financial rights, holding financial rights reduces the monopsony power of the buyer in the South with market power. This is the case because the value of the rights declines as p_2 is reduced. Let us briefly analyze this situation in the case of physical rights. The monopsonist in the South purchases all rights from the nonstakeholder owner, and then purchases q_1 units of power in the North at price $p_1 = C_1'(q_1)$, and q_2 units of power in the South at price $p_2 = C_2'(q_2)$. Denoting by $S(q_1 + q_2)$ the monopsonist's gross surplus, the latter maximizes $\{S(q_1 + q_2) - q_1 C_1'(q_1) - q_2 C_2'(q_2)\}$ over input purchases $\{q_1, q_2\}$. If returns in the

North are constant or do not decrease fast (C_1'' small), then there are no withholdings ($q_1 = K$) and the outcome is the same as under financial rights.[5] If C_1'' is large, the monopsonist withholds rights ($q_1 < K$) to extract rents from the generators in the North through higher rights prices which more than compensate for the effects of reduced supplies from the North on the price in the South. The monopsonist and the generators in the South are better off than under financial rights, and the generators in the North are worse off.[6]

c) When there is a *monopsony buyer in the North* and competitive behavior in the South, we saw that in the case of financial rights the behavior of the monopsony in the North could be affected if it held financial rights. In that case, reducing the price in the North would increase the value of the financial rights and, as a result, increase incentives to further distort demand in the upstream market. In the case of physical rights, the same incentives appear to operate if the markets operate simultaneously. However, if they operate sequentially, the monopsonist in the North gets no additional rights value by reducing prices further in the North ex post. Again, the potentially interesting twist associated with physical rights worth exploring further involves the potential for the allocation of physical rights to further restrict exports from the North. Under what if any conditions would a monopsony buyer in the North benefit (net of the cost of the rights) by acquiring and then withholding physical rights from the suppliers in the North to further reduce the nodal price in the North by restricting exports?[7] If the marginal cost curve in the North is upward sloping, consumers in the North can indeed be shown to have

[5]Under financial rights, the generators in the North are dispatched first, and there is no way for the monopsonist in the South to capture their inframarginal rents. The monopsonist solves:

$$\max_{q_1} \{S(K+q_2) - (q_2+K)C_2'(q_2) + [C_2'(q_2) - C_1'(K)]K\}$$
$$= \max_{q_2} \{S(K+q_2) - q_2 C_2'(q_2) - KC_1'(K)\}.$$

[6]When $C_1(q_1) = c_1 q_1 + b\frac{q_1^2}{2}$, $C_2(q_2) = c_2 q_2$, and $c_2 - c_1 < bK$ (so there are withholdings), then $q_1 = (c_2 - c_1)/b$ and $S'(q_1+q_2) = c_1$ under physical rights. So total output is the same as under financial rights; production inefficiency under physical rights implies that financial rights dominate physical rights from a social welfare perspective.

[7]Obviously, buyers in the North would be very interested in convincing the government to restrict exports in order to reduce local nodal prices. So, we should not be surprised to find consumer groups in exporting areas like Oregon or Washington state to be cautious about deregulation and increased exports from their low-cost suppliers. Since there are gains from trade, it would make more sense for regulators to give local consumers an entitlement to a share of the additional profits earned from price deregulation and unrestricted exports (e.g. regulatory entitlements to export profits), rather than restricting exports of cheap power.

an interest in withholding rights (over and beyond the incentive of a monopoly rights owner to withhold rights to raise price in the South if the marginal cost curve is upward sloping in the South). Accordingly, we would expect to find buyers located in an exporting region to try to exploit a physical rights system by engaging in collective action to withhold export rights in order to drive the local price for power down below competitive levels. There is a long history in the U.S. of consumer interest groups in states with cheap power supplies (typically hydroelectric production) trying to restrict exports of wholesale power to other states. The use of physical rights in this way would simply be another manifestation of such behavior.

3.2 Tender offer by G_2 (full free riding)

Suppose now that generator G_2 makes an unconditional tender offer at some price η to dispersed owners of rights who do not have market power. For simplicity, suppose that physical rights are initially held by producers in the North.

Stage 1: G_2 offers a price η and a fraction α_2 are tendered. Rights are then registered with the ISO.

Stage 2: The electricity market operates given the allocation of rights in stage 1.

At stage 2, G_2 selects p_2, knowing that $q_1 = (1 - \alpha_2)K$. G_2 selects $p_2 = \hat{p}_2(q_1)$, that solves

$$\max_{p_2} \{p_2[D(p_2) - q_2] - C_2(D(p_2) - q_1)\}.$$

At stage 1, selecting η is equivalent to selecting α_2 and q_1, knowing that

$$\eta = \hat{p}_2(q_1) - C_1'(q_1)$$

from rational expectations. G_2 therefore solves:

$$\max_{q_1} \{[\hat{p}_2(q_1)[D(\hat{p}_2(q_1)) - q_1] - C_2(D(\hat{p}_2(q_1)) - q_1)] - (K - q_1)[(\hat{p}_2(q_1) - C_1'(q_1)]\}.$$

The derivative of the maximand with respect to q_1 is

$$-p_2 + C_2' + p_2 - C_2' - (K - q_1)[\hat{p}_2' - C_1''] = (C_2' - C_1') + (K - q_1)(C_1'' - \hat{p}_2') > 0.$$

Increasing q_1 involves substituting cheap for expensive power ($C_2' - C_1' > 0$) and lowering the cost of purchasing the financial rights ($C_1'' > 0$, $\hat{p}_2' < 0$). Accordingly, G_2 has no incentive to acquire physical rights to enhance its market power. Since G_2 buys no rights, it maximizes its profits on the residual demand curve and physical rights do not enhance the market power of the monopoly generator in the South. This is the same result that we got for financial rights with this microstructure.

3.3 Auctioning of the rights by the ISO (partial free riding)

We now turn to the case where the ISO auctions off the physical rights. For simplicity let us assume that there are constant returns to scale in the North ($C_1(q_1) = c_1 q_1$).

Let us first explore the possibility that G_2 buys no rights. We are then simply back to the case where G_2 maximizes profit on its "no rights" residual demand curve:

$$\max_{p_2} \{p_2[D(p_2) - K] - C_2(D(p_2) - K)\}$$

and the resulting nodal price in the South is $p_2^m \equiv \hat{p}_2(K)$. The market price for physical rights, assuming constant returns to scale in the North, is then

$$p_2^m - c_1.$$

Suppose G_2 were to deviate and purchase some physical rights. If G_2 purchases some rights then it can control supply in the North (q_1) by withholding some rights from the market. G_2's profit is then:

$$\max_{p_2} \{p_2 [D(p_2) - q_1] - C_2(D(p_2) - q_1) - (p_2^m - c_1)(K - q_1)\}.$$

It is clear that a small purchase by G_2 is never profitable for G_2. The derivative of the profit function at $q_1 = K$ is equal to $(C_2' - c_1) < \eta$. That is, the increase in G_2's profits is less than the market value of the rights. This is the case because by reducing the availability of rights to generators in the North, expensive power in the South is substituted for cheaper power in the North. In contrast recall that with financial rights, if investors anticipate that G_2 will not buy rights, then it is optimal for G_2 to buy any amount of the rights at the low price and make them more valuable.

At this presumed no-withholdings equilibrium, G_2's decision problem is "bang-bang": G_2's optimal decision is to purchase either no rights or all of the rights. The intuition is that the marginal purchase of rights raises the value of other rights, and so the generator's profit is convex in the quantity purchased.

To illustrate this, suppose that the demand curve and production cost in the South are linear:

$$D(p_2) = 1 - p_2$$

and

$$C_2(q_2) = c_2 q_2.$$

The derivative of G_2's profit with respect to q_1, when the market expects no purchase by G_2 is then equal to:

$$c_2 - c_1 - (K - q_1)/2.$$

In particular, if $(c_2 - c_1) > K/2$ then G_2 strictly prefers not to buy physical rights. Basically, in this case, physical rights ownership by G_2 destroys value and the auction forces G_2 to bear the costs of this reduction in value if it buys physical rights at their fair market value in the auction. Since G_2 buys no physical rights, the availability of physical rights does not enhance its market power. In this sense, physical rights dominate financial rights.

More generally, there exists $\Delta < K/2$ such that:

i) if $c_2 - c_1 > \Delta$ then G_2 purchases no physical rights and the price of physical rights in the auction is $(p_2^m - c_1)$.

ii) if $c_2 - c_1 < \Delta$ then the equilibrium is in mixed strategies as in section 4.3 of our companion paper.

3.4 Physical rights on a two-node network: summing up

Our analysis of physical rights focused mainly on the case where there is a monopoly generator G_2 in the South, competitive generators in the North and no buyer market power in the North or in the South. Physical rights exhibit two behavioral factors which potentially have implications for the effects of the allocation of transmission rights on generator market power. First, physical rights make "rights withholding" a possibility. The effect of rights withholding is to reduce the effective capacity of the North-South transmission link and to reduce the supply of cheap energy from the North compared to the competitive case. This results in both production inefficiency and higher prices in the South compared to the no-withholding case. Second, G_2 is unable to commit to a level of production when selling rights, the effect of Coasian dynamics is to reduce the market power of the monopoly generator in the South given the supply from the North. However, physical rights are more likely to cause production inefficiency than if G_2 can commit.

Whether or not the availability of physical rights impacts these dimensions depends on the microstructure of the physical rights market. When the market is organized so that there is no free riding, in the benchmark case the monopoly generator in the South will buy all of the rights. This in turn triggers the behavior that potentially leads to production inefficiency and enhances G_2's market power. When the rights market is characterized by full free riding, G_2 does not buy any physical rights. Physical and financial rights are equivalent from this perspective. When the physical rights market is characterized by partial free-riding, G_2 either buys no rights or randomizes its bids as in section 4.3 of our companion paper. Last, it appears that the most interesting cases among alternative market power configurations will arise when buyers in the North can use the rights to restrict exports from the North so as to reduce the nodal prices at the cheap node.

4 Loop flows

As in our companion paper we extend our analysis to a three-node network to take the existence of loop flows into account. For conciseness, we will focus on the standard loop flow problem described in Figure 1. Production occurs at nodes 1 and 2 and consumption at node 3. Only the line between nodes 1 and 2 is constrained. Production at node 1 is cheaper, but production at node 2 is required for increasing production at node 1 beyond some threshold so that supply and demand can be balanced. Production is competitive at node 1, and monopolized at node 2. We refer to our companion paper for a fuller description of this network.

Let us remind the reader that because one-third (two-thirds) of the power produced at one

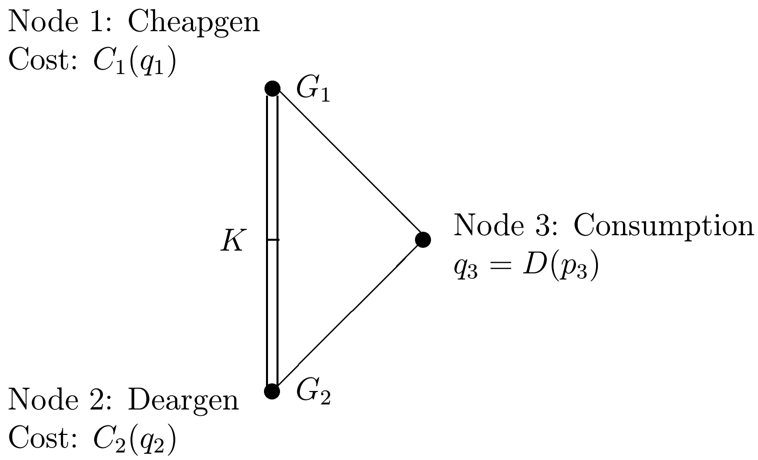

Figure 1

generation node flows through the indirect (direct) route to the consumer on a network with this configuration, the constraint on the North-South line can be written as

$$\frac{1}{3}(q_1 - q_2) \leq K.$$

The power produced by G_2 thus unloads the congested link by creating a "counterflow". The nodal prices, namely the prices that would prevail under bid-based dispatch (no bilateral trades), would "tax" and "subsidize" productions at the two generating nodes to reflect the marginal cost of congestion associated with increased production at each node.

$$p_1 = p_3 - \frac{\eta}{3}$$

and

$$p_2 = p_3 + \frac{\eta}{3}.$$

And so:

$$p_3 = \frac{p_1 + p_2}{2}.$$

We make three observations regarding physical transmission rights on a network with loop flow. Our first two observations restate for our simple network more general points made by Chao and Peck (1996) in a perfectly competitive environment; these points have not always been well understood and certainly haven't yet been fully incorporated into current reform proposals, and therefore are worth belaboring. Even in the absence of market power associated with the production or purchasing of electricity, the efficient implementation of a physical rights system on a network with loop flows must confront a number of significant challenges. These challenges must be understood to talk intelligently about physical rights systems for managing congestion on electric power networks.

Observation #1: Imputing capacity usage to a bilateral contract under loop flows.

Because an injection at one node of the network and an equal withdrawal at another node affect the flows through all links, the Independent System Operator must verify that the players scheduling a bilateral trade also possess the relevant physical rights on the network's links. The quantities of the "relevant physical rights" required in turn depend on the physical attributes of the network under different supply and demand conditions. For example, for our simple three-node network, a generator in the North (node 1) selling 1MW to a consumer at node 3 must own two thirds of a physical right on the line from node 1 to node 3, and one third of a physical right on the lines from node 1 to node 2 *and* from node 2 to node 3.

The designer of a physical rights system a priori can choose between two types of rights accounting systems: *a system with an exhaustive set of bidirectional rights or a system with a parsimonious set of unidirectional rights.* In the former case, the designer creates six rights, that is one per line in each direction. In the latter case, the designer contents herself with three directed rights (one per line), *and* allows for negative capacity usage. For example, when selecting directed rights from 1 to 3, 1 to 2, and 2 to 3, then a bilateral unit trade between G_2 and a consumer at node 3 consumes two thirds of a unit of transmission capacity on the 2-3 line (direct path), *minus* one third on line 1-2 and plus one third on line 1-3 (indirect path). We will discuss shortly the feasibility of either approach.

Observation #2: Unloading a link: creation of rights vs netting.

Ignoring for the moment market power, a fundamental issue in a physical rights system with loop flows relates to the provision of incentives for a generator located in the South to unload the congested link.

In the *exhaustive* set case, 5 out of the 6 types of rights are valueless provided that the corresponding directed flows do not congest their respective lines. Only physical rights for capacity for transferring power from node 1 to node 2 have positive value, η say. Then, G_2 receives no direct financial incentive (or "subsidy") for unloading the line. A bilateral trade by G_2 with consumers yields G_2 price $p_2 = p_3$ per unit. G_2 should receive $p_2 = p_3 + \eta/3$ have the proper incentives to produce. In contrast, a bilateral trade between a generator in the North and a consumer yields the generator $p_1 = p_3 - \frac{\eta}{3}$, as it should be. The basic problem here is that the value to generators in the North (G_1) of the generator in the South (G_2) producing some additional output is greater than the cost to G_2 of producing that additional output (note that we continue to assume that G_2 behaves competitively). If G_2 produced more then the G_1 generators could profitably produce more as well. Thus, there is an opportunity for G_2 to enter into mutually beneficial production and sales agreements with the generators at G_1 that would result in G_2 producing more and getting paid more for what it produces. For example, G_2 could contract with generators in the North offering to supply q_2 overall (recall $q_2 < q_1$) and bundle its own output q_2 with theirs to sell $2q_2$ to consumers at node 3. G_2 would then get credit for the value of its unloading the congested line by $q_2/3$. Netting[8] would occur as long as the Independent System Operator recognizes that there is no net flow created by the bundled outputs along the congested line, and so no physical rights would be demanded for dispatching them. The generators would then receive $2p_3q_2 = \left(p_3 - \frac{\eta}{3}\right) q_2 + \left(p_3 + \frac{\eta}{3}\right) q_2$, as it should be. Of course, in general, such agreements among producers might raise concerns about collusive behavior, and this consideration may make bundling an unattractive policy option. It must also be the case that the ISO and the stakeholders share a common physical model of the network, so there is a match between what the ISO recognizes at "nets" and what the stakeholders can agree to do.

Consider now the *parsimonious* set. The number of physical rights from node 1 to node 2 is no longer a fixed number equal to K unlike in the case of an exhaustive set, but rather is determined *endogenously* by G_2's production. Because each unit of production in the South unloads the congested link by one third the total number of rights available for bilateral trades

[8] "Netting" is called "counterscheduling" in the policy debate in California.

between G_1 and the consumers should be equal to

$$K + \frac{q_2}{3},$$

resulting in the following constraint on production in the North:

$$\frac{q_1}{3} \leq K + \frac{q_2}{3}.$$

Furthermore the newly-created rights should be turned over to G_2 who then resells them at price η each to producers in the North. The total revenue for a unit production in the South is therefore $p_3 + \frac{\eta}{3}$, as it should be.

Note three potential difficulties with this arrangement: First, it would seem that bilateral trades between G_2 and consumers and the associated production in the South must be scheduled ahead of those in the North, so as to allow G_2 to resell the newly-created permits to generators in the North. This unfortunate sequentiality, which may disturb the price discovery process, might be circumvented by allowing G_2 to sell short (that is, to sell in advance) physical rights that it anticipates receiving at the scheduling date, with clearing and settlements occurring at that date.

Second, the use of a parsimonious set may face difficulties in situations in which a link may be constrained in opposite directions at different times of the day or seasons.

Third, one might worry about G_2 possessing market power in the physical rights market (besides that on the energy market). For the same reason as in the two-node network, G_2 may want to withhold some of the newly-created rights. To see this, let us distinguish between the number of rights, $q_2/3$, held by G_2 as a result of producing q_2, and the number of rights, $\hat{q}_2/3$, sold to generators in the North, where

$$\hat{q}_2 \leq q_2.$$

Production in the North is then

$$q_1 = \hat{q}_2 + 3K;$$

and because $p_3 = p_1 + \frac{\eta}{3}$, G_2's profit can be written as

$$\begin{aligned} & p_3 q_2 - C_2(q_2) + \eta \frac{\hat{q}_2}{3} \\ & \qquad = P\left(3K + q_2 + \hat{q}_2\right) q_2 - C_2(q_2) + \hat{q}_2 \left[P(3K + q_2 + \hat{q}_2) - C_1'(3K + \hat{q}_2)\right]. \end{aligned}$$

G_2 withholds none of the newly-created rights if and only if the derivative of its profit function with respect to $\hat{q}_2$ at $\hat{q}_2 = q_2$ is nonnegative, that is if and only if (using the first-order condition with respect to q_2)

$$C_2'(q_2) - C_1'(q_1) - q_2 C_1''(q_1) \geq 0.$$

As in the two-node network, G_2 trades off the need for substituting expensive for cheap power (which argues in favor of no withholding) and the desire to extract G_1's inframarginal rents (if any).

Finally, we note that an identical "withholding" strategy for G_2 is feasible under exhaustive rights and netting, as long as G_2 can choose to schedule some of its production in the South without netting it with an equal production in the North. Thus, the two institutions do not differ with respect to their scope for withholding transmission capacity. [Similarly, prohibition of unmatched production by G_2 under exhaustive rights, or of withholding newly-created rights under parsimonious rights would be the counterpart to the capacity release program.]

Observation #3: Closed-end physical rights portfolios.

Whichever way one proceeds, the thrust of the introduction of markets for physical rights is to have such rights traded among stakeholders. Efficiency requires that the rights corresponding to links with excess capacity be traded at zero price. But if such rights were indeed worthless, an investor or a stakeholder could costlessly create a spurious scarcity by purchasing a sufficient fraction of them and withholding some of them. The parties engaged in bilateral trades would then have to pay for more than one link.

Thus, it does not seem reasonable to organize separate markets for physical rights on the different links. Indeed stakeholders value *bundles* of rights, rather than individual rights (which per se are useless). In our context, this suggests that one could for example offer two bundles of rights. The first bundle, with K such rights, tailored for dispatching Northern production *on a stand-alone basis*, would give the rights to two units of capacity between nodes 1 and 3, and one unit between nodes 1 and 2 and between 2 and 3. The second bundle, tailored to *joint dispatching* of equal (netted) quantities at the two generation nodes gives no rights on the line from 1 to 2, and a unit right on lines 1 to 3 and 2 to 3. This approach has the benefit of preventing anyone from creating a spurious scarcity of rights on noncongested lines; more thought however should be devoted to the design of this portfolio of bundles in situations in which the location and the direction of the binding constraints is uncertain.

5 Capacity release rules

One of the primary differences between a financial transmission rights system and a physical transmission rights system arises as a result of withholding of physical rights from the market which leads to an artificial contraction of the capacity of the transmission system. The potential for transmission capacity withholding naturally leads to the question of whether regulatory rules can be crafted which restrict the ability of stakeholders to withhold physical rights from the market. The transportation of natural gas on the interstate natural gas pipeline system in the U.S. is governed by a physical rights system.[9] Pipelines are required to offer to enter into transportation contracts with gas shippers and gas consumers that give them the physical right to transport gas from one point to another on their pipeline networks. These physical rights are tradable, subject to regulatory price caps. Rights holders who do not use their rights to support the transport of gas by a certain time period prior to any particular transportation date are required to "release" those unused rights for sale to other shippers and consumers in the gas transportation market.[10]

Let's consider how a capacity release program might be implemented for electricity. We will ignore all of the problems associated with dealing with loop flow in a physical transmission rights regime and return to the two-node model. The most interesting cases are when G_2 is a monopoly and G_1 is competitive and withholding occurs in either the simultaneous or the sequential cases. Several issues need to be addressed. First, at what time in the generation scheduling process are physical rights deemed to be "unused" and available for release for use by other generators? Second, when an unused right is used by another user what, if anything,

[9] Nothing like the "loop flow" phenomenon is observed to any significant degree on natural gas pipeline networks.

[10] See U.S. Energy Information Administration, *Natural Gas 1996: Issues and Trends*, DOE/EIA-0560(96), December 1996, Washington, D.C.; U.S. Federal Energy Regulatory Commission, Notice of Proposed Rulemaking, *Regulation of Short-term Natural Gas Transportation Services*, July 29, 1998.

is the initial owner of the right paid for its use? Third, how does the system respond to an ex post realization that some rights that were designated for use in the scheduling process, and not made available under the capacity release program, are found not to have been used either due to conscious overscheduling by generators or due to unanticipated plant outages or reductions in consumer demand served under bilateral requirements contracts?

Counteracting physical transmission capacity withholding behavior that is a component of G_2's strategy to exercise market power in the electricity market, requires that the unused capacity be released for sale to competing generators in sufficient time that they can use the capacity effectively. In a regime governed solely by bilateral contracts between generators and consumers and a requirement that generators submit balanced schedules to the ISO, the value of the physical rights to competitors and the effects of their release on market power could be heavily influenced by how far in advance of the formal scheduling periods the rights are released and made available to others. It is difficult however to conceive of a pure bilateral contract system with a release program because there is then no natural date at which the bilateral market closes and the leftover capacity is released to allow...further bilateral trades. A realistic release program therefore seems to require a sequence of a bilateral market followed by a centralized auction market similar to those set up in the restructured electricity systems. Thus, we will consider a two-stage timing in which the bilateral market closes, say, a day ahead, and is then followed by bid-based dispatch for the remaining capacity:

Stage 1: *Bilateral market.* Bilateral contracts between buyers and sellers can be negotiated at any point of time (five years ahead, a year ahead, a week ahead...) before the date, say a day ahead, at which the balanced trades together with the associated physical rights must be registered with the ISO. Let q_1 denote the amount of power injected in the North as an outcome of the bilateral market.

Stage 2: *Bid-based dispatch.* The unused transmission capacity, namely $K - q_1$, is released. An auction market, run as described in our companion paper, opens with transmission capacity $K - q_1$. That is, the stage-2 market is the standard auction market except that the transmission capacity is reduced to the leftover capacity.

Compensation for the released capacity: If there is congestion for the released capacity $K - q_1$, the ISO accrues some merchandising surplus from its operation of the stage-2 market. The ISO could return the merchandising surplus to the owners of the physical rights that it has taken possession of, using the difference in nodal prices in the stage-2 market to value these rights. This effectively turns any released physical rights into financial rights. We call this the *use-it-or-get-paid-for-it* rule. Alternatively, the ISO could give the merchandising surplus produced in the stage-2 market to charity or use to to help to defray the ISO's fixed costs. In this case, the holders of the released rights get nothing for them. We call this the *use-or-lose-it* rule.

a) *Use-or-lose-it rule.* This rule appears to provide the most powerful incentives for physical rights holders not to withhold rights from the market. The release of any rights they withhold to the ISO undermines the profitability of a withholding strategy and they lose entirely the value of any rights withheld from the market that they might otherwise earn if they sold (or used) the rights before the close of the day-ahead market. So, even if G_2 holds all the physical rights initially (at the start of stage 1), G_2 does not withhold any. The bid-based dispatch market is inactive. As in section 3.1, it makes a difference whether G_2 can centralize sales to consumers by purchasing power in the North, or whether G_2 sells electricity to consumers without internalizing the value of the rights sold to generators in the North. We thus conclude that, under the use-or-lose-it rule, G_2 obtains the commitment or noncommitment profit corresponding to $q_1 = K$.

Remark: The absence of stage-2 (last day) uncertainty in our model may conceal a potential cost of the use-or-lose-it rule if interpreted too rigidly. It may be the case that an a priori efficient plant in the North is registered at the end of stage 1 but becomes incapacitated or more generally becomes a high-cost unit at stage 2. Some flexibility should then be created so as to allow substitution possibilities for power at stage 2; the challenge is then how to provide stakeholders with incentives to reallocate production efficiently without altering the spirit of the use-or-lose-it rule. We leave this issue (which does not arise in our model) for future research.

b) *Use-it-or-get-paid-for-it rule*: This rule undermines the direct value to G_2 of withholding physical rights from the market since the withheld rights must be released, but imposes no penalties for doing so. Indeed, G_2 in equilibrium withholds all rights and so the bilateral market is inactive. Given that all transmission capacity will be used under any strategy, G_2's total profit (from generation, from the sale of physical rights, and from the dividends received for the financial rights resulting from withheld physical rights) is bounded above by

$$\max_{p_2} \{p_2 D(p_2) - C_2(D(p_2) - K) - K C_1'(K)\}.$$

But G_2 can get exactly this upper bound by withholding all rights and transforming them into financial rights (see our companion paper). We conclude that the use-it-or-get-paid-for-it rule, while preventing production inefficiency, allows G_2 to optimize against the full demand curve and leads to a high price in the South.

6 Summing up

Based on the analysis in this paper and in our companion paper, it is clear that when there is seller and/or buyer market power in an unregulated electricity market, the allocations of firm transmission rights can enhance market power and induce production inefficiency. Whether and how transmission rights can have such effects depends upon numerous factors, including the configuration of the underlying market power problems (location, buyer vs. seller), whether the transmission rights are physical or financial, the microstructure of the market for transmission rights, and in the case of physical transmission rights the timing of the rights market and the power market. For both physical and financial rights, their allocation is most likely to have adverse welfare effects when rights are initially allocated to a generator with market power at the expensive node (the importing region) or to a buyer with market power at the cheap node (the exporting region). We have identified hazards associated with either system and have discussed remedies.

While most of our analysis has focused on a simple two-node network without loop flow, we have also provided some analysis of a three-node network with loop flow. In the case of financial rights, the extension to the three-node network is reasonably straightforward and the results vis-à-vis market power do not change in any significant way. In the case of physical rights, designing a workable physical rights system in the presence of loop flow, even on a simple three-node network without market power, is a significant challenge. However, assuming that these design and implementation challenges can be overcome, the basic nature of the market-power enhancement problems identified for the two-node case does not change in important ways when loop flow is introduced.

In order to sum up our two papers in a concise way, let us focus on the benchmark case of a two-node network with generator market power at the expensive node. As we showed, a good understanding of this case almost effortlessly provides the deciphering key for the other situations.

6.1 The underlying issue

Generator G_2 can attempt to capture *three rents* corresponding to the three markets (two local electricity markets and rights market). The first rent is the consumer net surplus in the South; in all our variants, G_2 has local monopoly power in the South and so the same ability to extract consumer surplus. Indeed, in the two papers, the price in the South always exceeds the price that maximizes generation profit in the South when G_2 faces the residual demand curve when the link is congested ($p_2 \geq \hat{p}_2(K)$). Thus in the two papers the action is with respect to G_2's impact on the other two rents markets: value of rights and inframarginal rents of the competitive generators in the North.

The study of financial rights (paper I) centers around G_2's impact on the value of rights. Financial rights do not enable G_2 to reduce the power flow from North to South and thus to reduce the inframarginal rents of generators in the North. In contrast, under physical rights (paper II), G_2 can withhold transmission capacity and thereby capture some of the inframarginal rents in the North; on the other hand, physical rights receive no dividend, and G_2 therefore does not affect the value of associated dividends. Physical and financial rights therefore do not allow G_2 to impact the same rent. It is remarkable then that the two systems can be compared so readily.

6.2 Comparison

To save on notation (this is not essential), let us assume that $C_2(q_2) = c_2 q_2$, that is, production in the South exhibits constant returns to scale. This assumption allows us to compare G_2's optimal price function when $K - q_1$ physical rights are withheld,

$$\hat{p}_2(q_1) \equiv \arg\max_{p_2} \{p_2 [D(p_2) - q_1] - C_2 (D(p_2) - q_1)\},$$

with the price function that prevails when G_2 holds a fraction α_2 of financial rights (see our companion paper),

$$p_2(\alpha_2) \equiv \arg\max_{p_2} \{p_2 [D(p_2) - (1 - \alpha_2)K] - C_2 (D(p_2) - K)\}.$$

Under constant returns in the South,

$$\hat{p}_2(q_1) = p_2 \left(1 - \frac{q_1}{K}\right).$$

Social welfare in all our variants is a simple, decreasing function of the price p_2 in the South and of the level of production, $K - q_1$, withheld in the North:

$$W(p_2, K - q_1) \equiv S(D(p_2)) - C_2 (D(p_2) - q_1) - C_1(q_1),$$

where $S(\cdot)$ is the consumer gross surplus. Given local market power in the South, the optimum is obtained when the price in the South is $\hat{p}_2(K) = p_2(0)$, and when there is full production in the North ($q_1 = K$).

The upper bound, Π_1, for G_2's and the rights owners' joint profit under any institution is

$$\Pi_1 \equiv \max_{\{p_2, q_1 \leq K\}} \{p_2 D(p_2) - C_2 (D(p_2) - q_1) - C_1(q_1)\}.$$

This upper bound is obtained for $p_2 = \hat{p}_2(0)$ and $q_1 \leq K$ (with $q_1 < K$ if and only if $c_2 - C_1'(K) < C_1''(K)$). Letting q_1^c ("c" for "commitment") denote the optimal q_1 in this program, let

$$W_1 \equiv W(\hat{p}_2(0), K - q_1^c).$$

Let us also define

$$\Pi_2 \equiv \max_{p_2} \left\{ p_2 D(p_2) - C_2 \left(D(p_2) - K \right) - K C_1'(K) \right\},$$

$$W_2 \equiv W \left(\hat{p}_2(0) \, , \, 0 \right),$$

$$\Pi_3 \equiv \max_{q_1} \left\{ \hat{p}_2(q_1) D \left(\hat{p}_2(q_1) \right) - C_2 \left(D(\hat{p}_2(q_1)) - q_1 C_1'(q_1) \right\},$$

and letting q_1^{nc} ("nc" for "noncommitment") denote the optimal q_1 in this program,

$$W_3 \equiv W \left(\hat{p}_2(q_1^{nc}) \, , \, K - q_1^{nc} \right).$$

Last, let

$$\Pi_4 \equiv \max_{p_2} \left\{ p_2 \left[D(p_2) - K \right] - C_2 \left(D(p_2) - K \right) + \left[\hat{p}_2(K) - C_1'(K) \right] K \right\}.$$

and

$$W_4 \equiv W \left(\hat{p}_2(K) \, , \, 0 \right).$$

We have

$$\Pi_1 \geq \Pi_2 > \Pi_3 \geq \Pi_4,$$

$$W_4 > W_2 \geq W_1 \quad \text{and} \, W_4 > W_3.$$

We summarize the analyses of the two papers in Figure 2. Figure 2 assumes away free riding and therefore posits that gains from trade between the generator with market power and the rights owners are realized. In Figure 2, welfare decreases when moving east (increase in local market power) or north (increased withholdings).

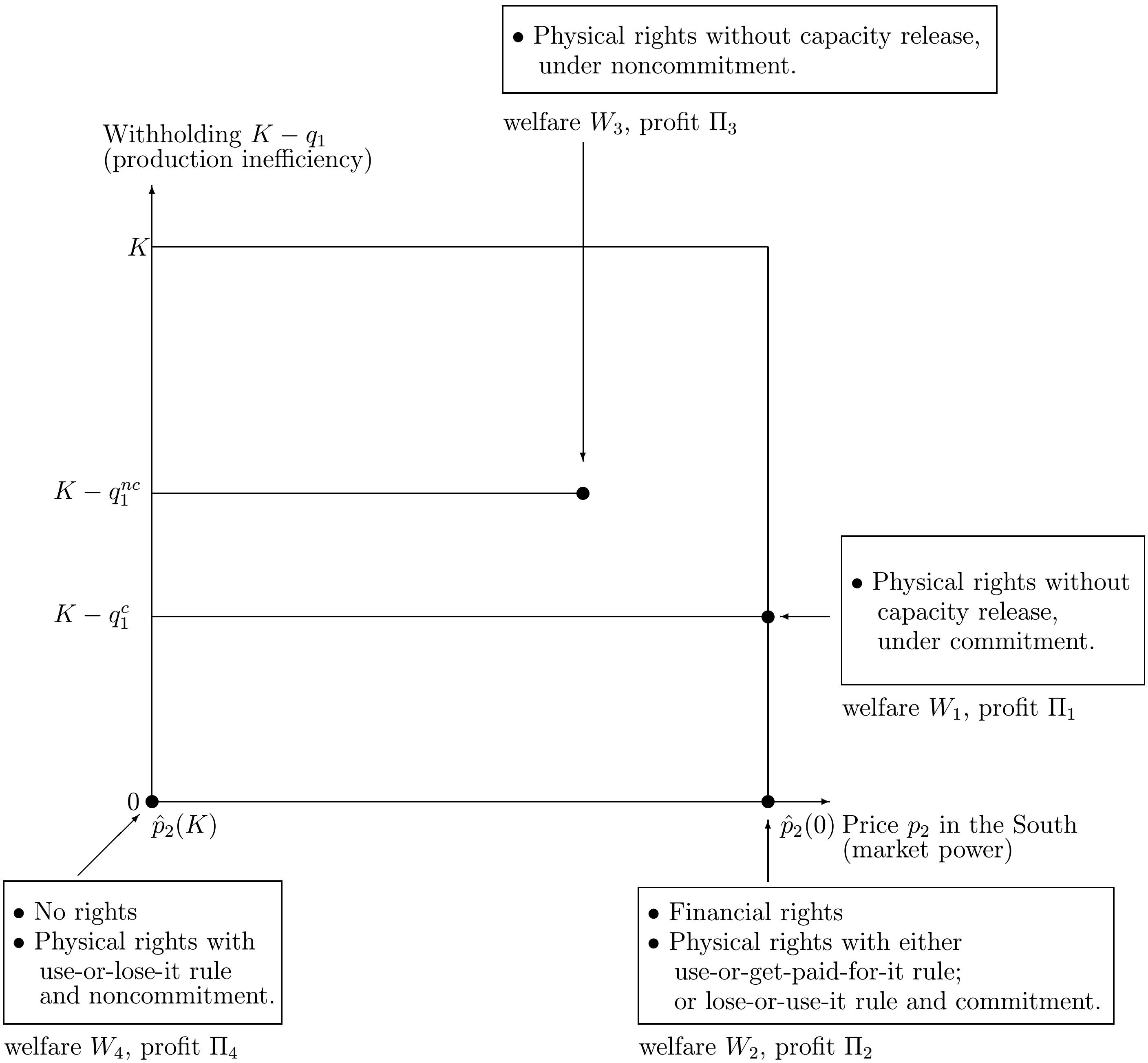

Figure 2

6.3 Future research

A striking implication of our policy analysis is that the absence of rights (the "zero net supply solution") does as well as and in general better than either system of rights. This leads naturally to the question of why is the ISO creating financial or physical rights if insurance opportunities can be created " synthetically" through ordinary insurance markets. There may be two reasons why a positive net supply may be unavoidable; we have not explored either reason and think this topic is a central area of potential research in view of the fact that all current policy proposals emphasize institutions with positive net supply of rights.

First, it may be the case that zero net supply (pure insurance markets) is not a feasible option. The ISO's merchandizing surplus must go to someone. To the extent that it goes to nonstakeholders or to stakehoders with no local market power, how can we prevent side deals between these investors and large stakeholders, that is stakeholders like G_2 who through their local market power can affect the value of the rights? Avoiding such side deals requires some form of " insider trading regulation", in which stakeholders with market power are not allowed to engage in side deals. The question then is: If one can prevent such side deals under zero net supply, can't one also prevent perverse holdings of financial rights by large stakeholders under positive net supply (see the discussion of the prohibition of "gambling behaviors" in our companion paper)? We leave this issue for future research.

Another argument may be that the creation by the ISO of transmission rights is required for the provision of transmission investment incentives. According to Hogan (1992), when new transmission investments are made, the ISO is supposed to create new financial rights to match the additional network capacity that has been created by the new transmission investments. The dividends from these financial rights then are supposed to become the (sole) source of the transmission investors' revenue. A similar investment motivation is associated with physical rights. The study of long-term incentives for investments in transmission is still in its infancy, and much work will be required in order to understand the articulation between these incentives and the design of transmission rights.

References

Bushnell, J. (1998) "Transmission Rights and Market Power," mimeo, University of California Energy Institute.

Chao, H.P., and S. Peck (1996) "A Market Mechanism for Electric Power Transmission," *Journal of Regulatory Economics*, 10(1): 25–60.

Hogan, W. (1992) "Contract Networks for Electric Power Transmission," *Journal of Regulatory Economics* , 4(3): 211–42.

Joskow, P. and J. Tirole (1998) "Transmission Rights and Market Power on Electric Power Networks. I: Financial Rights," mimeo, MIT and IDEI.