Rethinking Gas Markets – and Capacity

Larry E. Ruff[1]

Independent Consultant
Senior Advisor, Market Reform

ABSTRACT

The "US Model" of natural gas markets is based on long-term, point-to-point commercial capacity rights (MDQ_{XY}) that reflect the physical capacity of the pipeline and are traded frequently among system users (shippers) in markets independent of the transmission system operator (TSO). When physical capacity is complex and scarce and the gas market is dynamic there are many MDQ_{XY} that must be continually reallocated and reconfigured, making trading difficult/illiquid and market outcomes suboptimal.

Entry and exit capacities defined separately for a few large zones make trading easy and liquid but operationally problematic. Shipper-only trading would result in such a large gap between market and optimal (or even just feasible) outcomes that the TSO must engage in active capacity and gas trading itself to offset unconstructive/dangerous shipper trades.

These conclusions suggest that, at least in some/many complex situations, commercial capacity should be eliminated altogether and replaced with a TSO-operated on-the-day market that prices and allocates physical capacity directly, with financial hedging as an equivalent (or better) substitute for commercial capacity. A simplified version of such a market has been operating successfully in Victoria, Australia, since 1999.

Keywords: gas market; capacity rights; configure; point-to-point; entry/exit; spot market; Victoria

1. INTRODUCTION

New production technology is making natural gas a fast-rising star in the energy firmament, but the theory and practice of pipeline gas markets seem mired in the past. It may be time to rethink the conventional wisdom about pipeline gas markets, particularly the idea that daily market and operational outcomes should be determined by decentralized trading of some form of capacity right.

1.1. A Brief History

The US gas pipeline industry began in the early twentieth century in the form of independent, point-to-point pipelines project-financed on the back of long-term take-or-pay contracts on both ends – a beginning that shaped the future by making long-term, point-to-point transport seem like the natural or even only function of a pipeline. By the end of the century, and after several major government-mandated restructurings, the industry had (in the words of Jeff Makholm) been "transformed into an

[1] leruff@gmail.com; 8017 Oak Way, Windsor, CA, 95492, USA: 1-707-836-0437

industry that exhibits true Coasian bargaining in [long-term, point-to-point] transport entitlements and supports the world's only vigorously competitive and openly transparent gas market with an equally vigorous futures market." [2]

The "US Model" is often viewed as the best or even only logical form of gas market, but has not been fully or successfully applied outside the United States. In particular, the European Commission (EC) has decided that point-to-point capacity is not suitable for Europe, and that transmission system operators (TSOs) there should adopt the entry/exit approach used by the UK TSO – even though the economic and operational logic and efficiency of entry/exit capacity are questionable.

1.2. Focus and Approach

This paper analyzes market design issues on gas pipelines with complex and scarce physical capacity and frequently changing gas supplies and demands. These characteristics are linked by the facts that complexity and market volatility have no effect on a pipeline that has so much physical capacity it never constraints market outcomes, and that scarce capacity is easy to manage if the pipeline is simple enough or market conditions seldom change. The focus here is on situations where scarcity cannot be managed efficiently with simple commercial capacity or eliminated at acceptable cost with excess physical capacity.

The effects of pipeline complexity are usually ignored in discussions of gas markets, largely because pipelines are assumed to be simple. For example: "There is no ambiguity about where an oil or gas pipeline runs or what it does: transporting fuel from *one* location along a defined path to another [i.e., *one* other] location."[3] (emphasis added). This paper deals with more complex pipelines, beginning with one that transports gas from one location to *two* other locations – a seemingly trivial complication that, in fact, fundamentally changes the problem of defining and trading capacity. The paper concludes by discussing a complex pipeline system – essentially the one in Victoria, Australia – that takes gas from several, widely-separated injection points to more than 100 withdrawal points, with storage facilities and interconnections that can be injection points one day and withdrawal points the next, multiple laterals interconnected by a large ring, gas flows that can reverse direction from day to day or within a day, volatile weather that can cause the mostly-residential demand to change significantly and unpredictably from day to day and during the day, and little linepack that must be managed carefully to deal with the unpredictable swings in demand from day to day and within days. Whether or not such a system is a "pipeline" at all by the one-point-to-one-other-point definition above, it seems worthy of discussion.

2 Makholm, p. 119. Makholm provides a comprehensive review of how and why the US Model developed as it did, and suggests that many/most other gas systems would benefit from adopting this model.

3 Makholm, p. 17.

Because the objective here is to discuss problems caused by complex and scarce capacity and dynamic gas supplies/demands, other potential problems, such as equipment failures, overruns of contractual capacity limits, profile issues and trading inefficiencies, are ignored; these problems are often important in practice, but assuming them away makes it clear that the problems identified here are due to something more fundamental than (e.g.) failure of a compressor or inadequate trading platforms. The analysis here also focuses on short-term market and operational issues, with only a few comments about longer-term capacity expansion and financing.

1.3. Outline and Summary

Section 2 demonstrates the principal analytic proposition in this paper: except in the very simplest situations, a gas market based on point-to-point commercial capacity requires a lot of decentralized capacity trading among shippers and yet virtually always results in a “gap” between any market outcome constrained by the commercial capacity and an optimal market outcome constrained only by the physical capacity. To reduce this gap efficiently, the TSO can/should operate (but not itself trade in) short-term capacity and/or gas markets – and if it does this efficiently enough, the point-to-point capacity becomes essentially a financial hedge against the prices in the TSO-operated markets, creating the kind of market described in Section 4.

Section 3 discusses the entry/exit capacity concept supported by the EC as applied in the UK. Entry/exit capacity is easy for shippers to trade among themselves, but the results of shipper-only trading would be very large gaps between market outcomes and efficient (or even feasible) outcomes. To prevent such large gaps from emerging, the TSO must not only operate markets, but must actively buy and sell capacity and gas itself to counter the effects of shipper trades of capacity that poorly reflects pipeline realities. TSO trading is inherently problematic, but if it becomes extensive and efficient enough to reduce the gap significantly it is not clear what role entry/exit capacity is playing (other than raising TSO revenue). A fully-efficient entry/exit capacity market could look much like the market described in Section 4.

Section 4 discusses the development of the Victorian gas market and outlines a model market that is based closely on that market, i.e., it does away with commercial capacity and the costly and often unproductive trading it requires. In this market, a TSO-operated spot market prices and allocates physical capacity (and gas) efficiently on the day, and Financial Transport Rights (FTRs) are used to provide the kind of hedge against that congestion that point-to-point capacity is intended to provide, but in a much more flexible and efficient way.

Section 5 offers some concluding thoughts, and comments briefly on some longer-term issues that are otherwise ignored here.

2. POINT-TO-POINT CAPACITY – AND THE GAP

The basic assumption underlying the US Model is that point-to-point capacities can be defined in advance based on the physical capabilities of the pipeline, and then trading of these point-to-point capacities independent of the TSO – what Makholm calls "Coasian bargaining" in the quote in Section 1.1 above – will result in efficient use of the physical pipeline under any gas market conditions. Unfortunately, this assumption is true as a logical matter only for a pipeline that takes gas from one point "A" to one other point "B" (although a "point" here may include several operationally equivalent physical locations). This section explains why this is so and discusses some of the implications.

2.1. Simplicity, Complexity and the Capacity of a Pipeline

Consider the simplest-possible pipeline, one with a single injection point at A, a single withdrawal point at B, and physical capabilities that never change except in predictable ways. The on such a simple pipeline can use the physical capabilities of the pipeline, and reasonable assumptions about linepack and injection/withdrawal profiles, to estimate the maximum sustainable daily flow (by displacement) from A to B, apply a safety margin, and define this as the daily A-to-B capacity or "maximum daily quantity" MDQ_{AB} (in, say, GJ/day); MDQ_{AB} may be profiled over the year to take account of known changes in pipeline capacity, such as maintenance schedules, but does not depend on market conditions. Once allocated among shippers somehow, this MDQ_{AB} can be easily and efficiently traded among shippers so that it is always held by, and hence the physical capacity is always used by, those with the highest-value uses for it on every day under any market conditions.

Now consider what happens when this simplest-possible, A-to-B pipeline is complicated in the simplest-possible way – by adding a second withdrawal point C downstream from B, so there are now two point-to-point capacities to determine, MDQ_{AY} for Y = B and C. This seemingly trivial complication fundamentally changes the problem of defining and trading commercial capacity, because the two MDQ_{AY} are joint products produced by shared physical capacity, and (as always with joint products) the optimal quantity and price of each depend on, and hence change with, market conditions. Physical pipeline characteristics determine the trade-off between the two MDQ_{AY} , but to determine the economically optimal mix of the MDQ_{AY} the TSO must know their relative values in the market, which depend on the values of gas at different points on the pipeline, which depend on the two MDQ_{AY} the TSO is trying to set. This sort of logical circularity – or, more accurately, simultaneity – arises in much more serious and complex forms on electricity networks, but appears here on a very simple gas pipeline that has nothing like AC loop flow (although gas pipelines can have multiple paths between points and hence a form of loop flow).

To set the two MDQ_{AY}, the TSO must know or assume something about gas market conditions on the high-demand days when the MDQ_{AY} will matter, i.e., when they may constrain market outcomes. In principle, the TSO should select a few high-demand "design days", each representing a different season of the year and/or day of the week; forecast/assume the daily supply/demand curves for gas at A, B and C for each design day; use constrained-optimisation techniques to determine the highest-value/competitive-market-clearing set of gas injections and withdrawals subject to a model of the pipeline constraints on that day; and set the MDQ_{AY} for each season/day of the year at the calculated optimal withdrawals at Y (presumably reduced by a safety factor to reflect uncertainties) for the relevant design day.

Alternatively, the TSO can use a market to shift the forecasting burden to shippers, by inviting shippers to bid for the MDQ_{AY} based on their own gas market forecasts. The TSO can then choose a mix of the MDQ_{AY} that maximizes the total value of the cleared bids subject to the pipeline constraints. TSOs use simple versions of such a process when they hold "open seasons" to determine the most valuable mix of long-term capacity contracts for financing a new pipeline, or to reconfigure and reallocate released long-term capacity.

The essential common features of either of these (or any other reasonably logical) processes for setting/configuring the optimal MDQ_{AY} is that they consider simultaneously all gas flows on, and the physical characteristics of, the pipeline, and that the optimal MDQ_{AY} both affect and are affected by the gas market. Any such process is inherently a centralized, and hence monopoly, market, that is most naturally operated by the TSO, although in principle some entity could operate it using a TSO-provided model of pipeline characteristics and operations.

2.2. Reconfiguration and the Gap

The critical point here is that, no matter how, by whom or at what levels the MDQ_{AY} are set in advance for each day, the optimal mix or configuration of the MDQ_{AY} for any day will depend on actual gas market conditions on that day, so any MDQ_{AY} configured in advance will be different from, and hence suboptimal compared to, the optimal MDQ_{AY} virtually every day. Trading among shippers can, in principle, allocate the preconfigured MDQ_{AY} to those shippers with the highest-value use for it on every day, but cannot, even in principle, reconfigure it as market conditions from day to day. The market outcome will be inefficient virtually whenever the MDQ_{AY} affect/constrain the market.

For example, consider an A-to-B-to-C pipeline with a large, high-value load at B that was assumed to be on-line when one of the above processes configured the MDQ_{AY} for the (say) year. If this load goes off-line but all other (actual or potential) loads are unchanged during a high-demand period of the year, it would (probably) be optimal to decrease MDQ_{AB} and increase MDQ_{AC} so that more load can be served at C. But trading of the existing MDQ_{AY} among shippers cannot accomplish such a reconfiguration, so the MDQ_{AB} not used by the industrial load at B will at best be sold at a low price to low-value load at B and may even be unsold at a price of zero – even if the physical capacity could easily supply higher-value load at C. Even with perfectly efficient trading of the preconfigured

MDQ_{AY}, there may be a large gap between the market outcome based on point-to-point commercial capacity and a market outcome constrained only by the actual physical capacity.

If the TSO had no alternative to setting the MDQ_{AY} for the year in advance, the suboptimal market outcomes would properly be attributed to the cruel reality that any year-ahead forecast will often be significantly wrong, not to any flaw in the market design. But there is no fundamental reason the MDQ_{AY} should be based on year-ahead gas market forecasts. The processes suggested above for determining the MDQ_{AY} could operate every month based on month-ahead gas market forecasts or shipper bids, or every day based on day-ahead forecasts or bids, or even during the day based on observed current conditions. The gap resulting from inefficiently preconfigured MDQ_{AY} is attributable entirely to the fact that the market is based on preconfigured MDQ_{AY}.

The materiality of the avoidable inefficiencies inherent in point-to-point capacity will depend on the situation. On a long-distance pipeline with few injection points X and withdrawal points Y, stable demands and a lot of excess capacity, the optimal configuration of the point-to-point MDQ_{XY} may be stable or unimportant enough that the inefficiencies are small enough to be acceptable. But on a complex, potentially congested pipelines with dynamic gas supplies and demands, there may be many high-demand days when the market outcome based on preset MDQ_{XY} would result in some scarce physical capacity being used inefficiently or not at all – particularly if a large safety factor must be included to assure that the market solution is always feasible, which will result in unused physical capacity on virtually every high-demand day.

Decentralized trading can reconfigure commercial capacity if (staying with the example) the A-to-B-to-C pipeline can be regarded as separate A-to-B and B-to-C pipelines, with MDQ_{AB} and MDQ_{BC} determined purely from pipeline capabilities, not market conditions. In this case, trading among shippers could reallocate some MDQ_{AB} to supply load at C instead of at B (if there is enough MDQ_{BC}) if demand shifts from B to C. Such segmentation of capacity can and does work – albeit with a lot of complex trading – when/where each segment is operationally independent enough. But on complex pipelines with many short segments, operational interactions across segments create externalities that make such segment-specific capacity rights impractical or (if large safety margins are required) highly inefficient.

2.3. Reducing the Gap Inherent in Point-to-Point Capacity

On a complex pipeline with dynamic gas supplies and demands, even perfectly efficient trading of preconfigured point-to-point capacity (MDQ_{XY}) will leave a gap, sometimes a large gap, between any market outcome constrained by the MDQ_{XY} and an optimal outcome constrained only by the actual physical capacity. When this gap is too large to tolerate, the TSO should or even must (if reliability is threatened) take action to reduce it, either by reconfiguring the MDQ_{XY} *ex ante* so the independent markets will clear with a smaller gap, or by intervening *ex post* to reduce the gap remaining after the independent markets have cleared, or both.

Most TSOs with long-term MDQ_{XY} reconfigure it *ex ante* from time to time by buying low-value MDQ_{XY} and selling higher-value MDQ_{XY} consistent with the physical capacity. Such reconfiguration increases the value of the MDQ_{XY} portfolio for both the TSO and shippers, and could/should be routine in a complex, dynamic gas market: a TSO-operated periodic or as-needed reconfiguration market should solicit shippers' buy/sell bids for MDQ_{XY} and use these to determine a new mix of MDQ_{XY} that maximizes the bid-defined value of the total portfolio subject to the modelled pipeline constraints. Such a process could maintain the commercial benefits of long-term contracts while providing some of the flexibility needed to respond to changing market conditions; for example, multi-year MDQs could be reconfigured annually to deal with secular market trends, and/or seasonally to deal with weather patterns.

To deal with day-to-day changes in market conditions, the TSO's capacity reconfiguration market should operate every day, at least during periods of high demand. This is technically feasible with modern information technology; but if the TSO reconfigures (and hence reprices and reallocates) the MDQ_{XY} every day to reflect market conditions on the day, it is the TSO's on-the day market, not the MDQ_{XY} set in advance, that determines market and operational outcomes. Every day, yesterday's MDQ_{XY} are sold and the proceeds are applied to purchase today's MDQ_{XY}, making the MDQ_{XY} essentially financial hedges against the daily price of X-to-Y capacity. Taking this concept to its logical and efficient conclusion suggests that the TSO should operate a daily spot market that prices and allocates physical capacity directly each day and provides financial hedges against congestion prices. Such a market is described in Section 4.

Most TSOs that use MDQ_{XY} also intervene in or override market outcomes *ex post* to reduce the most obvious gaps remaining after the independent capacity markets have closed; for example, interruptible capacity is called when the market outcome threatens to be unfeasible, or non-firm capacity is scheduled when the market is likely to leave valuable physical capacity unused. Such *ex post* actions are best thought of as gas, not capacity, transactions; the fact that they are typically taken under ISO/shipper contracts that are usually regarded as special forms of capacity contract means only that they are very restrictive and inefficient ways to buy and sell spot gas, not that they have anything to do with capacity in the MDQ_{XY} sense.

To reduce the gap efficiently *ex post*, the TSO should operate an incremental on-the-day gas market, in which any shipper/load could offer to take, if the price is right, more or less gas at a specific location than allowed by the MDQ_{XY} it holds, in effect buying or selling the gas increment. Again, this is technically feasible with modern IT; but if shippers can buy and sell gas at efficient locational market prices on the day without regard to the MDQ_{XY} they hold, the final outcome will be determined by the TSO's spot gas market. MDQ_{XY} will simply be the right to buy gas at X and sell it at Y, which is effectively a financial hedge against the Y-minus-X spot gas price differential or (the same thing) the price of one-day X-to-Y capacity. Again, making a market based on MDQ_{XY} fully efficient is equivalent to doing away with MDQ_{XY} altogether and providing financial hedges against spot price differentials, as discussed Section 4.

Gas markets based on point-to-point capacity often (at least appear to) to work "well enough" in practice for various reasons: many of the inefficiencies and their costs are not obvious; TSOs use occasional *ex ante* reconfigurations and frequent (if inefficient) *ex post* gas market interventions to reduce the most obvious inefficiencies; and (most importantly) point-to-point capacity is used only where pipeline and market conditions are simple and stable enough that such *ad hoc* fixes are good enough. Where physical capacity is complex and scarce and gas supply/demand is dynamic, the gap created by MDQ_{XY} is too large to ignore or for the TSO to reduce much; in these situations, point-to-point capacity is unlikely to be adopted in the first place or to be retained where it is tried – which helps explain why Victoria did not adopt, and Europe is moving away from, point-to-point capacity.

3. ENTRY/EXIT CAPACITY – AND A LARGER GAP

Experience with point-to-point capacity in the early years of European gas market liberalization (the 1990s) quickly persuaded the European Commission (EC) that point-to-point capacity cannot support effective and efficient competition on Europe's complex pipeline system; since 2002 the EC has been encouraging European TSOs to move to the kind of entry/exit capacity used in the UK. Unfortunately, the EC (and others frustrated with point-to-point capacity on complex systems) have misdiagnosed the problem: the flaw in point-to-point capacity is less the "point-to-point" part than the "capacity" part. Replacing point-to-point capacity with entry/exit capacity is not necessarily an improvement – as this section illustrates, using the UK entry/exit system as an example.

3.1. The UK Entry/Exit Capacity System

In the UK, National Grid Gas (NGG) owns and is the TSO of the National [gas] Transmission System (NTS). NGG aggregates the actual injection and withdrawal points on the NTS into several entry zones and many exit zones. Monthly entry capacity for zone A is defined as the right to inject in A up to a maximum daily quantity (MDQ_A) without regard to where the gas will be withdrawn or how it might get there; the commercial fiction is that all injected gas flows to a notional National Balancing Point (NBP). There is no real exit capacity or other limit on withdrawal locations or amounts, just exit-zone-specific charges on actual peak-day withdrawals within a month; the commercial fiction is that monthly exit capacity, along with the obligation to pay a regulated capacity charge for it, is "assigned" *ex post* to the shippers who actually "used" the capacity during the month.

NGG auctions monthly MDQ_X , for X = A, B, C, ..., every six months (or so). Shippers can trade monthly entry capacity for any entry zone among themselves bilaterally (NGG provides an on-line bulletin board to help) or in independent markets or exchanges without NGG involvement. NGG operates a series of auctions in which shippers (and NGG) can buy and sell daily MDQ_X for each operational day starting a week or so prior to and continuing into the day; thus, on every day there will be auctions for daily MDQ_X for each of several future and the current day(s). There is an On-the-Day Commodity [gas] Market (OCM) that trades gas at the NBP and at various physical locations on the system; shippers can use the NBP to trade their individual daily balances and as an alternative to trading entry capacity: a shipper with excess MDQ_X can use it to put gas into the system and then sell that gas in the OCM (notionally at the NBP) to shippers that do not have enough MDQ_X.

Shippers must provide NGG with information on their planned operations. When NGG sees a potential problem for an operating day within the week or so ahead, it responds by buying and selling MDQ_X in its daily capacity auctions. To manage system balance and congestion on the operational day, NGG buys and sells OCM gas at specific locations as well as daily capacity in its within-day capacity auctions.

There is much more to the UK entry/exit system in detail, but this summary includes enough to illustrate the points made next.

3.2. The Dubious Economic Logic of Entry/Exit Capacity

Entry/exit capacity makes capacity trading much easier and (particularly if there are only a few, large zones) more liquid than it is with point-to-point capacity. But entry/exit capacity is such an oversimplified/inaccurate representation of physical capacity that trading purely among shippers would often produce a very large gap between the market outcome and an efficient or even feasible outcome, sometimes even when the entry/exit capacity is not constraining the market – unless NGG uses large safety margins in setting the MDQ_X, in which case the gap will take the form of unused capacity on virtually every day, even high-demand days. The gap will occur more frequently and be larger than the analogous gap with point-to-point capacity, which occurs only when the market is capacity-constrained (assuming perfectly efficient trading of the MDQ_{XY} among shippers) and is limited by the fact that the MDQ_{XY} reasonably reflect the ability of the physical capacity to move gas from X to Y for all X,Y pairs simultaneously, even when the configuration is economically suboptimal.

The gap with entry/exit capacity is potentially so large that the TSO cannot manage it by just operating markets in which shippers traded capacity or gas, as is possible with point-to-point capacity. As illustrated by the NGG example above, the TSO must continually manage the gap by buying and selling capacity and gas itself, using multiple markets that it and others operate. Any TSO trading gains or losses are presumably shared among the TSO and shippers in regulator-approved ways, but this can result in a lot of money flowing around the system to mute price signals and distort incentives.

In deciding when and where to buy and sell capacity and gas to solve a problem, the TSO must use models of the physical capacity, and presumably considers the market prices of capacity and gas in different zones to find (somehow) a set of capacity and gas transactions that solve the problem while advancing some economic or commercial objective. Such a process is analogous to the *ex ante* capacity reconfiguration and *ex post* gas markets suggested above for reducing the gap inherent in point-to-point capacity, but is much less logical, transparent and efficient – and has a much larger gap to close. Entry/exit capacity has such a weak link to actual gas flows that the TSO cannot really know how buying or selling it somewhere will affect gas flows; the TSO sees only the market prices resulting from individual transactions, not the full set of possible transactions that could be used to solve a problem; the economic objectives are confused by the fact that the TSO keeps at least some of its trading profits. The overall process is *ad hoc*, opaque and inefficient, even if it is implemented through markets.

In principle, the processes for managing the large gap in an entry/exit system could be made more transparent and efficient by using the kinds of centralized *ex ante* capacity and *ex post* energy trading processes suggested above for use with point-to-point capacity. But, just as in that case, if the TSO-operated markets are truly efficient, they will allocate and price physical capacity (and gas) directly with no need for the entry/exit capacity, which will at best become a financial hedge against the prices in the TSO markets. It would be easier, more logical and more efficient just to eliminate the entry/exit capacity altogether and use the kind of market process described in Section 4.

4. No Commercial Capacity – and No Gap

This section uses recent experience in Victoria, Austalia, to illustrate the proposition, suggested by the analysis above, that on a complex pipeline with scarce capacity and dynamic gas supplies and demands, the gas market should not use any form of commercial capacity, but should use a spot market to price and allocate physical capacity (and gas) directly on the day, with financial instruments available as hedges against congestion prices. This approach eliminates both the costly and sometimes unproductive trading activity required with commercial capacity and the possibility of a gap between the outcome of the actual market and the outcome of market constrained only by physical capacity (because the actual market ***is*** constrained only by physical capacity). The situation and process in Victoria are summarized first, and then a model market based close on the actual Victorian market is outlined.

4.1. The Victorian Experience

This paper, and its focus on complex pipelines with dynamic market, was motivated primarily by the author's experience with the Victorian natural gas system. In the late 1990s, Victoria restructured and privatized its then-state-owned natural gas monopoly, the high-pressure portion of which was essentially the complex, dynamic, and operationally challenging system described in Section 1.2, and needed a market that could manage operations and competitive commercial operations on such a system. There were vague suggestions that the US Model could and should be used, but nobody ever provided even an in-principle description of how the hundreds of required long-term point-to-point capacity rights might be defined, used to manage system operations, and reallocated and reconfigured as market conditions changed from day to day and often within the day, in a small market with few players and virtually no experience with or infrastructure for sophisticated trading. So the market designers turned to economic and market first-principles and lessons from modern electricity markets.

The Victorian gas market began operations in 1999, and has been operating and evolving successfully since then. It has accommodated, and perhaps even helped stimulate in some cases, new private sources of gas, power generators, underground storage facilities and interconnectors. The market has produced price-driven responses to many significant and several very large shocks without ever requiring market suspension or even operator intervention (with one exception). It is generally regarded as the most competitive and dynamic gas market in Australia, with extensive bilateral contracting, multiple retailers and active customer switching. And it has accomplished all this even though – or, more likely, because – it does not use commercial capacity of any kind.

4.2. A Gas Market with No Commercial Capacity

This section outlines, at a very high level that ignores many details and complications, a model gas market with no commercial capacity. This model market is essentially the Victorian market, but with one important difference: the model market uses congestion pricing and Financial Transport Rights (FTRs) but the Victorian market does not. The Victorian market uses a spot market to schedule operations and manage congestion efficiently, and provides a form of hedge against the allocated out-of-market costs of congestion, but does not actually price congestion in the market and hence does not need the FTRs described here.[4] Sophisticated congestion pricing and FTRs are now routine in some large and complex US electricity markets, but have not been proven for pipeline gas markets. Significant creative work would be required to implement the congestion price/FTR arrangements outlined here, but this has been true of every advance in modern network markets, including the initial implementation and subsequent evolution of the Victorian market.[5]

[4] Pepper, *et al*, describes the market-clearing and pricing process and model used in the Victorian market.

[5] Read, *et al*., discusses issues involved in using linear programming (LP) to clear a gas market, including calculating congestion prices.

The central feature of any gas market that has no commercial capacity must be a daily spot market operated by the TSO (or some other entity coordinating closely with the TSO) that allocates and prices physical capacity and gas. The principal features of such a spot market can be illustrated by outlining what happens on an operational day, beginning on a day when no pipeline constraints affect the market outcome, i.e., when there is no congestion.

- Some hours prior to the start of the day, each shipper decides how much gas it wants to inject (sell to the spot market) at each point A and withdraw (buy from the spot market) at each point B to deliver to its contract customers (or to storage or a connected pipeline) there, and at what prices, if any, it is willing to increase or decrease these amounts. This information is submitted to the daily market/scheduling process in the form of daily nominations and "inc/dec" bids.

- The nominations and inc/dec bids are used in a market-clearing optimisation to determine a set of gas transactions/schedules and prices that clears the market (i.e., that maximises the total bid-defined value of all transactions) subject to the constraints in a model of the pipeline constraints. On a no-congestion day no constraints are binding, so there is a single morning price (P^M) everywhere on the system.

- Shippers are notified before the start of the day of their cleared transactions/schedules (say, sale of Q_A at A and purchase of Q_B at B) at the morning price P^M, and the implied net payments to the settlement system, $P^M \times (Q_B - Q_A)$, are booked for later settlement, i.e., the schedules are financially firm commitments. There will be no transactions or payments in the initial schedule unless the nominations and bids indicate that some shippers want (in effect) to schedule and trade imbalances ($Q_B - Q_A$) in advance.

- If conditions and expectations can change during the day enough to matter, the market process can be repeated during the day to determine and price incremental quantities of daily gas; the Victorian market has four incremental markets during the day. (For simplicity in describing settlements, this example assumes only a single, morning run of the market.)

- At the end of the day, aggregate market outcomes and the latest market bids determine an end-of-day price (P^E) that each shipper will (eventually) pay or be paid for any differences between the morning quantities and end-of-day actual quantities. P^E will be about the same as (higher/lower than) P^M if total demand turns out to be about the same as (higher/lower than) projected in the morning run of the market.

- Any aggregate gas imbalance for the day is bought/sold by the market at the end-of-day price and carried forward in linepack, where it is sold/bought at tomorrow's morning price. Any net monetary gains or losses from these linepack transactions are periodically allocated to shippers.

- An *ex post* settlement process computes the net payments due to/from each shipper and manages payments and (an important issue) credit.

Now consider a day in which congestion, in the simple form of a binding transport constraint between upstream zone A and downstream zone B, appears in the morning run of the market. The focus here is on the effects of congestion, so it is assumed that the morning forecasts are accurate and all schedules are followed, i.e., there are no incremental end-of-day prices and quantities to worry about. The submission and processing of bids, etc., are the same as before, but now:

- The market-clearing/scheduling optimization encounters A-to-B congestion – i.e., not enough gas can flow from A to B to meet total demand in B at the upstream price P_A – so it automatically finds and schedules a combination of demand decs and supply incs (if there are any) in B that satisfies the constraint at least cost; the downstream price P_B is set equal to the highest inc/dec bid needed to equate supply and demand there. The price difference (P_B - P_A) due to the A-to-B congestion is the "congestion price".

- Shippers implement their schedules, and the high price in B gives them all incentives to reduce their loads and, if possible, increase supply there, whether or not they have real-time meters, submitted dec bids to the morning market or have FTRs (discussed next).

- At settlement, a shipper who sold Q_A in A and purchased Q_B in B will pay $P_B \times Q_B$ for the downstream gas it bought, will be paid $P_A \times Q_A$ for the upstream gas it sold, and – the new wrinkle here – will be paid the congestion price (P_B - P_A) for the quantity of A-to-B Financial Transport Rights it holds (FTR_{AB}); some algebraic manipulation shows that the shipper's net payment to the settlement system will be

  $$P_A \times (Q_B - Q_A) + (P_B - P_A) \times (Q_B - FTR_{AB}),$$

  i.e., the shipper buys (or sells, if negative) its net imbalance (Q_B - Q_A) at the market gas price in A, P_A, and pays (or is paid) the congestion price (P_B - P_A) on the difference between the amount of gas it flows across the constraint (Q_B) and its FTR_{AB} holding.

- If the total FTR_{AB} held by shippers equals the capacity of the A-to-B constraint, total FTR payments by the settlement system will equal the total "congestion rent" collected by the settlement system; if FTR_{AB} is higher (lower) than the physical capacity of the constraint, the settlement system will have a deficit (surplus) proportional to the difference.

These arrangements for managing congestion risks are different from, but no more complex and in many ways better than, those in a market based on point-to-point commercial capacity. A shipper with long-term A-to-B capacity rights can be given the same amount (in GJ/day) of A-to-B FTRs, which will allow it to flow that amount of gas from A to B without paying anything other than the price of the right itself, just as its point-to-point capacity rights would do. But FTRs accomplish this basic objective much more flexibly and efficiently than do point-to-point capacity rights.

The simplicity and efficiency of a spot market/FTR system can be illustrated with the example in Section 2.2. In that example, an A-to-B-to-C pipeline configures its point-to-point capacity on the assumption that a large, high-value load at B is on-line; when that load goes off-line, the inability of shippers to reconfigure point-to-point capacity by trading among themselves results in physical capacity being used to supply low-value load at B or even being unused, even though it could be used to supply higher-value load at C. If this pipeline used a spot-market/FTR system, with the point-to-point FTRs configured the same way, the result would be very different: when the high-value load at B shuts down, the spot market would automatically reprice physical capacity, reallocate it to serve additional load at C (if that load outbid the low-value load at B), charge the newly-served load at C (which has no FTRs) the market value of the physical capacity it used, and pay that value to the now-off-line load at B under its FTRs. The new situation could last indefinitely, with the new load at C simply using the physical capacity on a spot basis and paying the off-line load at B the spot market value of the physical capacity every day; but if the load at B is never going to come back, the FTRs should eventually be reconfigured in an FTR auction that the TSO could/should operate (say) monthly or as-needed.

In effect, a spot market/FTR system automatically finds, prices, executes and settles the kinds of short-term capacity trades and reconfigurations that are essential for efficiency in a complex and dynamic gas market but that are difficult or impossible with point-to-point capacity rights – and probably with entry/exit capacity as well, although the TSO processes there are so *ad hoc* and opaque that it is hard to know.

5. CONCLUSIONS

Gas markets operate on many different kinds of pipelines with many different market situations. A Victoria-type market, based on a spot market and FTRs rather than commercial capacity, is surely not the best solution for every gas system – just as the US Model that evolved over the twentieth century in the unique US environment is surely not the best solution for every non-US gas system in the twenty-first century. Large gas systems consist of many different types of subsystems interconnected in various ways, and there is no reason every subsystem should or even could use the same form of gas market. It is quite conceivable, for example, that a continental gas market could include several subsystems with Victoria-like markets interconnected and fed by long-haul, point-to-point pipelines using point-to-point capacity. If all these markets must use the same form of market, it should be one based on a spot market and FTRs; such a market would be easy to design and operate on a long point-to-point pipeline, while a point-to-point capacity market would be unworkable on a Victoria-like system.

The analysis here shows that, on a complex pipeline with scarce capacity and dynamic gas supplies/demands, a market based on point-to-point (or any other) commercial capacity defined and configured in advance will produce inefficient outcomes virtually whenever the commercial capacity affects the market. Point-to-point capacity may have longer-run benefits that exceed the costs of the short-run inefficiencies, and if the longer-run benefits cannot be obtained any other way and exceed the short-run costs, point-to-point capacity is the better solution. But if not, not. Point-to-point capacity should not be used in the misguided belief that efficient trading of it among shippers will produce efficient market outcomes or that there are no other options.

The most-frequently-cited long-run benefit of long-term point-to-point capacity is its alleged value in underwriting the financing of new capacity. This was a real and large benefit in the early days, when long-haul pipelines were being built to supply monopoly distribution systems that could safely sign long-term, take-or-pay contracts. But those days are long gone in many/most places. On an existing complex pipeline with a dynamic and competitive gas market, such as the Victorian system, a new point-to-point pipeline within the system is unlikely to be economic or to find many customers willing to contract long-term for its point-to-point capacity (except perhaps as a bypass threat during tariff negotiations). In such situations, both existing and new capacity will have to paid for through some combination of throughput charges and the (say) annual auction of capacity rights or FTRs

References

Makholm, Jeff (2012). *The Political Economy of Pipelines: A Century of Comparative Institutional Development*, University of Chicago Press.

Pepper W., Ring B.J., Read E.G., and Starkey S.R. (2012). "Implementation of a Scheduling and Pricing Model for Natural Gas", in M. V. Pereira and P. M. Pardalos (editors), *Handbook of Networks in Power Systems*, Vol II, p3-36, Springer.

Read E.G., Ring B.J., Starkey S.R., and Pepper W. (2012). "An LP Based Market Design for Natural Gas", in M. V. Pereira and P. M. Pardalos (editors), *Handbook of Networks in Power Systems*, Vol II , p77-114, Springer.