

Some Costs of Reliability Limits

Douglas R. Hale and Thomas Leckey

DHALE@eia.doe.gov

tleckey@eia.doe.gov

U.S. Energy Information Administration
Washington, DC 20585 USA

Thomas J. Overbye

Overbye@ece.uiuc.edu

University of Illinois at Urbana-Champaign
Urbana, IL 61801 USA

Abstract

To ensure that the electrical grid will continue to operate even when generators and lines fail, electricity system managers restrict power flows between areas to levels below the physical capability of the lines. Electricity traders have complained that these safety margins, which are formalized in area interchange limits, can preclude profitable trades. Line limits can also protect local markets from outside competition. This paper quantifies some of the costs of area interchange limits constraining power flows between areas in the Eastern Interconnect in the summer of 2000. Recognizing the hidden costs and competitive effects of line limits is a prerequisite to balancing reliability goals against their cost.

1. Introduction

Utilities, Independent System Operators (ISOs) and National Electric Reliability Council (NERC) reliability areas rely on excess generator capacity and redundant line capability to ensure electricity demand, including losses, does not exceed generation. If it does, voltages and frequency may drop, possibly causing blackouts. System managers have traditionally required generation capacity that exceeds expected peak demand. The “safety margin” in the Northeast Power Coordinating Council (NPCC)¹ is currently about 15%; suppliers (load serving entities) in the PJM Interconnect² are required to have a 17% excess of generation over their peak demand.

Similarly, system operators, usually control areas, enforce power flow limits on lines with their neighbors that include safety margins derived from contingency analyses. In a contingency analysis, engineers present a model of the electrical system with hypothetical demand conditions and a base

¹ NPCC consists of New York State, the New England States, and the Ontario, Quebec and Maritime Provinces.

² PJM is an acronym for Pennsylvania, (New) Jersey, Maryland, and the system also includes Washington, DC and the Delmarva Peninsula.

case of operating generators and lines. Large generators and major lines are then taken off line one at a time to mimic unplanned outages. This is called an n-1 contingency analysis: all but 1 of the n pieces of major equipment in the electrical system is assumed to operate normally. The analysts note those operating regimes that cause failure of other large lines, potentially resulting in cascading blackouts. Through a planning procedure, they preclude catastrophic failures, essentially “outlawing” failed operating regimes, by de-rating vulnerable power lines. The line limits that are imposed to ensure that the system continues to operate after a failure are called area interchange limits, n-1 limits, contingency limits, reliability limits or some similar term. System operators treat these area interchange limits as if they were actual physical limits when they dispatch the system, thus reducing the chance of outage induced blackouts.³

The n-1 approach to achieving reliable electricity supply has been effective: Large area blackouts have been rare in the United States since the North American Reliability Council was established in 1968. But in recent years, some control areas have canceled large volumes of otherwise profitable trade to protect the lines against hypothetical events.⁴ Unfortunately since the probability of blackouts is not calculated as part of the n-1 approach,⁵ it is not possible to know whether canceling these trades actually reduced the risk of blackout. Another problem is once limits are set they are not dynamically updated to reflect the real time evolution of demand, supply, equipment failures and weather. Consequently trades could be disallowed when there is no realistic chance of blackout and allowed when there is. The emergence of competition has also led to new complaints that reliability limits are being manipulated to protect incumbent suppliers from outside competition.

This paper quantifies the short term operating costs of the area interchange limits reported by NERC for the Eastern Interconnect in the summer of 2000. The Eastern Interconnect is the US east of the Rocky Mountains, exclusive of Texas, and several Canadian Provinces. The paper also illustrates how more realistic line limits can act to protect local markets from high prices.

The paper is organized as follows. Section 2 presents the model used to make the cost and price calculations. Section 3 shows that managing the Eastern Interconnect at its physical potential would bring significant benefits. Under peak summer demand conditions across the Interconnect, system operating costs could be reduced almost 4 percent by loosening area interchange (reliability) limits. This is a large gain considering that to meet peak demand simultaneously requires almost all of the available generator capacity. It is on the order of the savings, also about 4 percent, estimated for the restructuring of the United Kingdom’s electricity industry.⁶ When demand is less and more relatively cheap capacity is available the potential cost reductions would be larger. At lower demand levels, the savings increase to 6.6%.

³ Note that if a major piece of equipment has already failed, the n-2 limits become the relevant constraints.

⁴ The Federal Energy Regulatory Commission has endorsed the use of transmission loading relief orders to individual generators to keep line flows below area interchange limits. These orders are based on a transaction's "priority" and not on its economic value. In addition, those generators who have a priority cannot sell it to others who are willing to pay. A summary of recent curtailments appears in Transmission Constraint Study, Presentation of FERC Staff to the Commission, December 19, 2001.

⁵ The probability of system failure depends not only on what equipment fails but also on what information is available to operators and what they do with it. These considerations are currently resolved by operator judgment.

⁶ D.M. Newbery & M.G. Pollitt "The Restructuring and Privatisation of the CEGB: Was It Worth It?" *Journal of Industrial Economics*, 1997, 45(3): 269-303.

Section 4 illustrates how the choice of area interchange limit can dominate pricing. As shown by the examples of New England and Florida, transmission managers such as Regional Transmission Organizations (RTOs) will have the power to eliminate the gains from trade and competition, which we estimate as 24% for New England and 5% for Florida, by their choice of contingency (trade) limits. Section 5 discusses congestion problems around a load pocket, New York City, and Section 6 is the conclusion.

The Eastern Interconnect Model

The short term operating cost of area interchange (reliability) limits is the difference between the least cost of meeting demand while enforcing the limits and the least cost of meeting demand consistent with the transmission system's physical capabilities. Specifically we ask, if the Eastern Interconnect were managed to meet demand at least cost while respecting all electrical constraints, but not necessarily all of the area interchange limits, would there be significant short-term benefits? To estimate the costs of interchange limits we assume that the existing utility-based systems minimize costs within their borders but restrict trade with their neighbors to honor the limits reported by NERC. The current situation is called the "restricted trade" regime. By comparison, the least cost alternative allows any economic (lower cost) trade that is physically feasible. This is called "unrestricted trade." Since regional markets are not perfectly competitive, their actual costs are understated by this estimate. Consequently the actual costs of area interchange limits are higher than reported here.

Trade reduces the cost of meeting electricity demand by replacing generation from high cost, but local, generators with generation from lower cost, distant suppliers. How much cost is reduced and how much electricity can be transmitted from low cost to higher cost areas depends upon the relative costs of hundreds of generators and the details of the transmission system (lines, capacitors, transformers and so forth) connecting generators to electricity users. The opportunities for trade are also limited by constraints placed on line flows in order to promote reliability.

Computing the cost of meeting demand under "restricted" and "unrestricted" trade in the face of this complexity requires a mathematical model. The model used in this study consists of four components: a description of the electrical properties of essentially all of the generators and higher voltage power lines in the Eastern Interconnect; estimated operating costs of most of the generator capacity; the contingency limits; and, an algorithm for calculating the least cost of meeting demand consistent with the physics of the transmission system. The Eastern Interconnect includes the National Electrical Reliability Council's (NERC's) regions shown in Figure 1 except for WSSC and ERCOT. It is roughly the United States east of the Rocky Mountains and excluding Texas. WSSC and ERCOT are electrically separate from the East and so very little electricity crosses their boundaries.

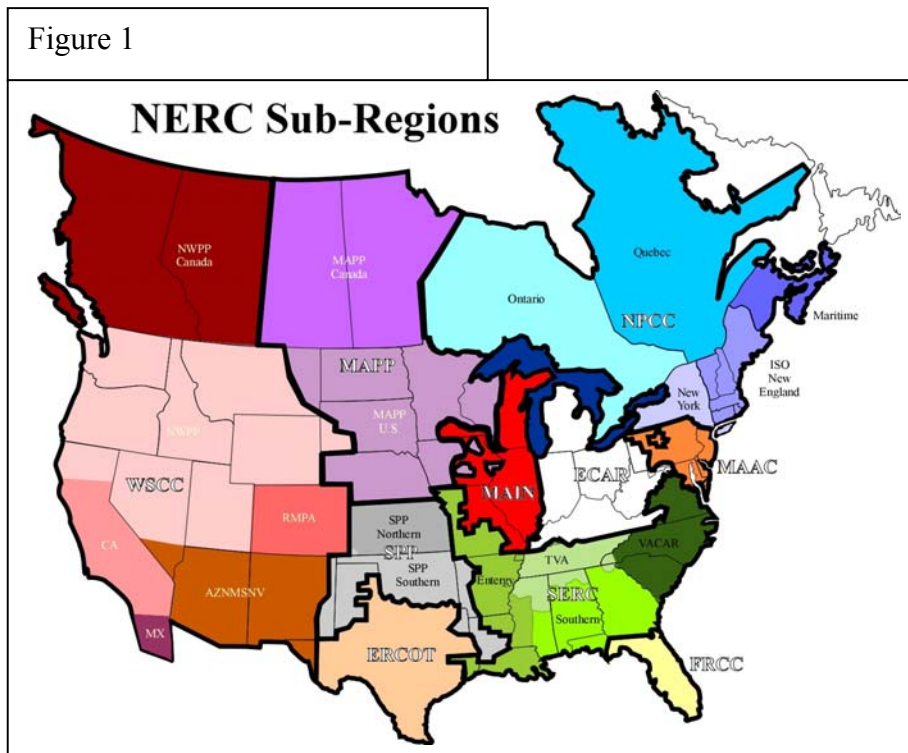
The electrical model for the study is based on the NERC's summer 2000 Base Case model of the Eastern Interconnect. NERC's Multiregional Modeling Working Group built the Summer 2000 Base Case from regional power flow models provided by NERC's members. This model is intended to "...realistically simulate bulk electric system behavior." The Committee's guidance for component models ensures that the integrated model is realistic:

Of paramount importance in this effort is the detail in which the various systems are modeled. The detail included in each system model must be adequate for all inter and intra-regional study activities but not necessarily as detailed as required for internal studies. This means that each system model should include sufficient detail to ensure that power transfers or contingencies can be realistically simulated.⁷

The summer 2000 Base Case consists of 33,528 buses, 22,812 loads, 5,312 generators, 2,361 switched shunts, 45,421 AC lines/transformers (lines), 10 DC lines, and 107 control areas. Total load at peak is 536.4 GW. A significant portion of the modeled case lies outside the United States portion of the Eastern Interconnect, which is the primary focus of this analysis. Modeled peak load in the United States is about 510 GW, around 30 GW less than modeled in the entire case. Similarly, about 4,740 generators and 31,000 of the buses are located in the United States. Of the 107 areas represented, 11 are in Canada, and another 7 represent equivalenced power flows to areas in the Western Systems Coordinating Council (WSCC) and the Electric Reliability Council of Texas (ERCOT). These 89

areas of interest comprise 9 NERC regions in the Eastern Interconnect (Figure 1). The NERC flowgates (line limits) described on the NERC Market Re-dispatch website were also included in the electrical model.⁸

Short run operating costs for generators are essential in analyzing optimal power flows. Unfortunately, the NERC 2000 case does not report operating costs for generators, nor is it straightforward to match the generators listed in the NERC 2000 case with those reported on EIA's inventory survey forms.⁹ Table 1 reports the NERC regional



distribution of generators modeled and identified. Overall, 86 percent of the generators were identified which enabled the technological character to be determined for 96 percent of the capacity.

⁷ NERC, *Multiregional Modeling Working Group Procedural Manual*, June 2000, page 2

⁸ The original NERC case contained a large number of electrical violations, including overloaded lines and substandard voltage in some areas. Typically the violations occurred in remote locations and involved small amounts of energy. Most of these violations were fixed by correcting apparently inaccurate data. Details are available from Thomas Overbye.

⁹ Utility units are detailed on the EIA-860A, "Annual Electric Generator Report – Utility" and non-utility units are reported on the EIA-860B, "Annual Electric Generator Report – Nonutility."

Identification was most difficult in New England and New York where small hydroelectric, cogeneration, and non-utility generators are most prevalent.

Our approach to estimating operating

Region	Units			Capacity (GW)		
	Modeled	Identified	Percent	Modeled	Identified	Percent
ECAR	551	466	84.6%	107.1	102.4	95.6%
FRCC	309	289	93.5%	38.3	37.6	98.2%
MAIN	435	361	83.0%	56.1	53.1	94.5%
APP	512	441	86.1%	36.9	35.5	96.1%
NEPOOL	475	334	70.3%	26.0	23.7	90.9%
NYPP	387	286	73.9%	34.4	32.0	93.1%
PJM	484	422	87.2%	59.0	57.0	96.6%
SERC	1209	1117	92.4%	170.8	164.3	96.2%
P	362	329	90.9%	45.6	44.1	96.9%
East Interconnect	4724	4045	85.6%	574.2	549.7	95.7%

costs per kilowatt-hour for fossil units was to use the cost of fuel as reported on the FERC Form 423¹⁰ for the month of June 2000, and then to adjust the fuel cost by multiplying by the heat rate.¹¹ Representative heat rates were developed from a variety of sources, most often the EIA-767¹² for large steam turbines, and the FERC Form 423 in combination with the EIA-759.¹³ These costs were developed at the plant level, and then assigned to the appropriate units. About 69% of the capacity was assigned costs using this method. Where it was possible to identify individual steam units burning coal and another fossil fuel, unit level estimates of costs associated with the various fuels were developed. Four percent of the capacity was assigned these unit level costs.¹⁴ Because many of the generating units modeled either fell below the reporting threshold or were outside the survey frame (non-utilities), a costing alternative was developed. Fossil generators without reported plant level costs were assigned state-level averages derived from the FERC Form 423, and adjusted using plant-specific heat rates. Another four percent of capacity was assigned using this alternative methodology. Costs for nuclear units (15 percent) were developed from the FERC Form I, for 1999, adjusted to 2000 dollars.¹⁵ These costs were uniformly low. Hydroelectric and pumped storage plants, about eight percent of the capacity, were assigned the same nominal cost, low enough to make them inframarginal.

Table 2 reports the capacity modeled¹⁶ in the NERC 2000 case, as well as the forecasted demand (load) by NERC region. The ratio of regional capacity to regional load is generally indicative of the reserve capacity, the ability of a region to serve its load, but it may also indicate the ability of a region to serve

Region	Capacity (GW)	Load (GW)	Ratio
ECAR	107.1	93.0	1.15
FRCC	38.3	36.5	1.05
MAIN	56.1	48.8	1.15
MAPP	36.9	35.5	1.04
NEPOOL	26.0	22.9	1.14
NYPP	34.4	28.5	1.21
PJM	59.0	49.5	1.19
SERC	170.8	158.0	1.08
SPP	45.6	37.4	1.22
East. Int.	574.2	510.1	1.13

¹⁰ “Monthly Cost and Quality of Fuels for Electric Plants.” Reporting threshold is 50 MW at the plant.

¹¹ \$/MMBtu times thousand BTU per kwh, yielding either mills/kWh or \$/MWh.

¹² “Steam-Electric Plant Operation and Design Report.”

¹³ “Monthly Power Plant Report.”

¹⁴ This differentiation was especially important in PJM, which has about a dozen plants where some of the units burn coal, while associated units burn gas or petroleum.

¹⁵ Inflated about two percent.

¹⁶ More than twenty GW of this capacity was originally modeled as off-line, or uncommitted, so in a dispatch the ratios are more stringent than shown in Table 2.

load elsewhere. By this measure, SERC and ECAR, significant areas of base load capacity appear to be “tight” power pools, with relatively slim operating reserves above peak load requirements. Conversely, New England and New York, two regions characterized by higher than average electricity prices and which presumably would benefit from trade, appear to have adequate reserves. Florida has little spare capacity in the NERC 2000 case.

Table 3 displays the regional breakdown and technology of those generators for which cost estimates were developed and included in the model. About 65% of the capacity that could be assigned a cost is fossil steam, burning coal, natural gas, residual fuel, or petroleum coke. On average, steam units fired by coal had costs roughly one-fourth of those fired by either natural gas or petroleum.¹⁷ Short run

Table 3. Capacity and Number of Units by Technology Type and Region

Region	Combined Cycle		Gas Turbine		Fossil Steam		Nuclear		Hydroelectric and Pumped Storage		Other		Total	
	Units	MW	Units	MW	Units	MW	Units	MW	Units	MW	Units	MW	Units	MW
ECAR	4	1,620	104	5,404	299	86,362	6	5,732	38	2,901	15	378	466	102,397
FRCC	37	4,524	128	5,672	98	23,035	5	3,954	4	47	17	383	289	37,615
MAIN	0	0	110	6,173	170	32,540	17	13,305	61	975	3	58	361	53,051
MAPP	5	359	107	4,146	159	23,569	6	3,580	69	2,770	95	1,065	441	35,489
NEPOOL	42	3,057	53	2,225	76	10,671	5	4,324	118	2,986	40	416	334	23,678
NYPP	54	3,527	60	2,520	94	15,862	6	4,922	70	5,115	2	50	286	31,996
PJM	41	3,231	157	8,633	161	28,862	13	13,035	40	2,887	10	351	422	56,999
SERC	58	5,410	279	14,224	374	91,742	32	31,063	373	21,762	1	77	1,117	164,278
SPP	11	1,277	63	3,612	155	34,994	1	1,164	76	2,488	23	615	329	44,150
East Interconnect	252	23,004	1,061	52,609	1,586	347,637	91	81,079	849	41,930	206	3,393	4,045	549,652

marginal costs for nuclear units were also low. Most of the combined cycle units burned relatively costly natural gas in June 2000, but the efficiency of these units worked to reduce the estimated operating costs overall. Generally, the gas turbines were the most costly units dispatched in the model.¹⁸ Hydroelectric and pumped storage units were dispatched at a uniformly low cost on the assumption that they would run during the peak hour. Where reasonable estimates of costs could not be developed, unit output was fixed at the level of the original NERC 2000 filing, and the generator was not applied in the optimization algorithm.

Because regions have different mixes of generators, each region ends up with a different estimated cost profile (Table 4). ECAR, SERC, MAIN, and MAPP possess large amounts of coal and nuclear baseload capacity, making those regions the lowest cost generation resources in the aggregate. NYPP and PJM, with larger shares of gas and petroleum steam units, show higher cost profiles. Florida, with very little coal capacity has the highest short run operating costs on average. The cost profile for both NEPOOL and NYPP includes significant hydroelectric resources, which are made available at relatively low cost: under more realistic unit commitment, costs in these areas would be higher than indicated here. In general, as the model optimizes to reduce cost over large areas, we would expect to see costs fall in the higher cost areas, though localized transmission constraints could prevent that outcome in some areas.

¹⁷ <http://www.eia.doe.gov/cneaf/electricity/epav1/ta20p1.html>

¹⁸ “Other” units are mostly internal combustion, consuming gas. A few are MSW (relatively expensive) and wind (relatively cheap).

In order to calculate the costs of contingency limits it is necessary to calculate the least cost in each trading area when those limits are honored. NERC's Summer 2000 case reports the relevant contingency flow limits. Table 5 reports a few of the specific flow limits across the trading areas.¹⁹ Negative values indicate imports.

The physical description of the electrical network and the generator cost data are inputs to an algorithm that calculates the least cost of meeting demand. The algorithm determines the output of each and every generator in such a way that overall cost of meeting demand is as low as possible consistent with the laws of physics, the capabilities of the transmission system, generator availability and capacity and, in the restricted trade regime, administrative limits on each region's imports. In the free trade regime there are no administrative limits on imports, just physical limits.

Large non-linear minimum cost problems are hard to solve. Our approach was to solve the non-linear alternating current equations (i.e., the power flow equations) for an initial trial solution, note the line and voltage violations, and make a linear approximation to the power flow equations. We then solved for the minimum cost of a linear version of the non-linear electrical model in such a way as to remove the violations. This resulted in another "solution" for generator outputs. That solution was given to the electrical equations and the process started again. These iterations continued until the solution coming from the cost minimization model was consistent with all of the electrical equations.

Least cost operation is an attribute of a perfectly competitive market.²⁰ Schweppe and his colleagues showed precisely how least cost operation (net benefit maximization) defines competitive equilibrium in an electricity market.²¹ They showed the competitive price at any location is equal to the system cost of one more unit of production, plus the value of the additional losses incurred in transmitting power to the location and the cost of the increased line congestion brought about by increased consumption. In other words, price at any location is just the system marginal cost of production plus the

Table 4. Weighted Average Cost of Generation by Region (nominal dollars/MWh)

Region	Capacity (GW)	\$/MWh
FRCC	37.6	38.22
NEPOOL	23.7	34.33
SPP	44.1	32.11
PJM	57.0	30.95
NYPP	32.0	29.37
MAIN	53.1	24.48
MAPP	35.5	22.79
SERC	164.3	21.84
ECAR	102.4	16.45
East Interconnect	549.7	24.96

Table 5. Selected Flow Limits, Summer 2000 Peak

Area	Net Scheduled Transactions MW
PJM	124
Southern	598
FRCC	-2268
TVA	-177
NY ISO	-1585
ntergy	-111
ISO NE	-2526
AEP	-543
Commonwealth Edison	-435
Consumers (MI)	-508
Duke	144
VA Power	352
First Energy	-1815
Cinergy	239
Ameren	-711

¹⁹ FERC Form 715, planning contingencies for 2000.

²⁰ Least cost operation is also consistent with operating philosophy of regulated utilities. The difference is that regulated utilities price at average cost and the competitive price is marginal cost. When demand is responsive to price, the levels and patterns of demand will change. Also, competitors will invest differently from regulated utilities whose returns on investment are protected by its regulators.

²¹ Schweppe, Fred, Caramanis, Tabors and Bohn, *Spot Pricing of Electricity*, Klur Academic Publishers, 1988

value of induced losses plus the cost of induced congestion.

When congestion and losses are significant, some low cost generators cannot be fully used. Delivering their full output levels to the grid would cause serious physical damage, perhaps blackouts. In those circumstances, operators replace generation from cheaper facilities with higher cost generation. This causes the system cost of supplying an additional unit of electricity to increase, contributing to a system wide increase in electricity's price.

Prices typically vary by location because the losses and congestion induced by transmission depend on where electricity is being demanded. These location dependent prices are called Locational Marginal Prices (LMPs). Losses and congestion themselves depend on how others are using the transmission system. For example, when demand is low, lines will usually have excess capacity, and congestion costs would likely be negligible or absent entirely. If an increase in one customer's demand is enough to congest a line, all customers may have to bear congestion charges.²² These external effects of individuals' production and consumption decisions are included in the results reported below. But, in the real world properly pricing and charging for externalities are significant problems: Everyone attempts to push external costs on others.

To summarize, the model calculates the least cost of meeting demand and the associated competitive prices. The model contains an electrical component because losses and congestion can only be calculated with a detailed electrical model. The model contains a detailed description of generator costs and capabilities because it is necessary to tradeoff costs, congestion and losses across all potential suppliers to calculate system marginal cost and least cost.

3. The Cost and Price Consequences of Area Interchange Limits

This section reports the estimated impact on system operating costs and prices of area interchange limits. That amounts to calculating the differences in prices and system costs between the "restricted trade" regime and the "unrestricted trade" regime. In the restricted trade regime, the relevant demands and supplies are primarily those within each region. There are some imports and exports but they cannot exceed the area interchange limits specified in the NERC 2000 case.²³ Each region is minimizing its own costs using its own resources but not fully exploiting further cost reductions through trade.

In the unrestricted trade regime, the relevant demands and supplies are those across the Eastern Interconnect and there are no administratively imposed trade restrictions. Trade is, however, restricted by the physical capabilities of the transmission system. The prices that are calculated in both cases are those appropriate to a competitive market.

We made cost estimates for summer peak and for a "shoulder" period. The summer peak is NERC's estimate of demand when all areas are simultaneously in a heat wave. At such times, essentially all available generators are on line and there is little room for cost reduction. The shoulder period is nominally defined as 80% of peak demand. During a shoulder period generation capacity exceeds

²² Calculation and enforcement of congestion charges varies across the regulated jurisdictions in the Eastern Interconnect.

²³ The FERC 715 files specifies the total imports or exports for each area, and in rare instances bilateral transactions are modeled.

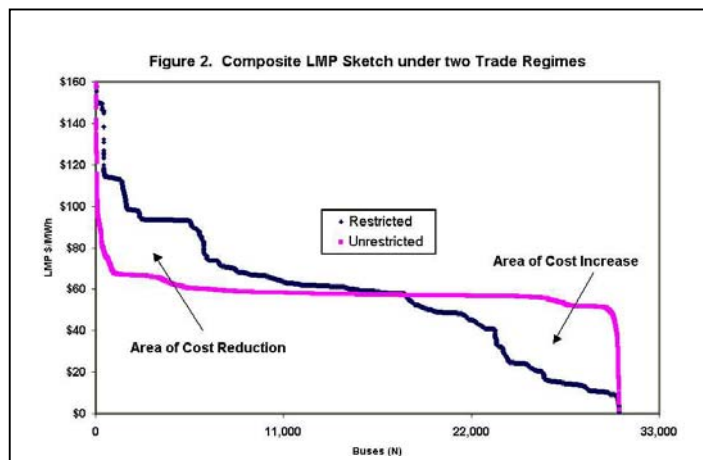
demand so there is more potential for reducing costs by calling on relatively lower cost generators. Since demand is reduced in all areas at the same time, and almost all areas have some low cost generators, there may be limited need for any area to import cheaper power.

The restricted trade case has total operating cost of \$10,968,547/hour.²⁴ LMPs were computed at slightly more than 31,000 buses. The LMPs ranged from a high of \$ 987/MWh to a low of \$ -16/MWh with less than 6% having values above \$ 100/MWh and only eight buses having negative values.²⁵ The average LMP was \$ 57.96/MWh while the standard deviation was \$33.71/MWh. The area interchange constraints are what permit the area average LMPs to vary considerably. They range from a high of \$ 149.82/MWh to a low of \$ 6.49/MWh. At the solution the flows on 14 lines were being enforced as binding constraints; this represents only about 0.05% of the 41,000 lines modeled in OPF areas.

Next the system was modeled assuming unrestricted trade, i.e., the area interchange contingency constraints between the U.S. operating areas were relaxed (those with Canadian areas were still enforced). Thus the entire U.S. portion of the Eastern Interconnect was treated as though it were a single operating area.²⁶

As expected, relaxing the area contingency constraints at peak demand resulted in a lower operating cost, with the value dropping by about 3.7% from \$11 million/hour to \$10.6 million/hour. With the area constraints relaxed, power was free to flow from the low cost areas to those with higher costs. The largest absolute changes occurred in the Virginia Power control area, which increased its exports by over 1,400 MW, while TVA and the Florida Control area saw their imports increase by 1,161 and 1,674 MW respectively.

This resulted in the average area LMPs converging as well, with the highest value now just \$76.87/MWh and the lowest \$ 52.46/MWh.²⁷ Compared to the administered trade model, the average LMP²⁸ remained about the same, \$57.95/MWh. The standard deviation of the LMPs, however, decreased substantially from \$33.71/MWh to \$8.44/MWh. The price leveling would have been greater except that increased trade led to the congestion of 17 lines compared to 14 in the restricted trade case. Figure 2, which



²⁴ NewIND100d010402.pwb.

²⁵ Negative LMPs simply indicate that buses on one side of a constraint should reduce generation or increase their load. Negative LMPs occurred on the PJM system during each of the three summer months in 1999, reaching \$-199.33MWh in July. One area, Consumers Energy in Michigan has an additional 15 LMPs with negative values, stemming from large reactive power loads.

²⁶ NewSA100d010402.pwb.

²⁷ One nonutility generator, modeled as an area with negligible load, solves with an average LMP of \$31.20/MWh.

²⁸Not load weighted. Median LMP cost drops slightly.

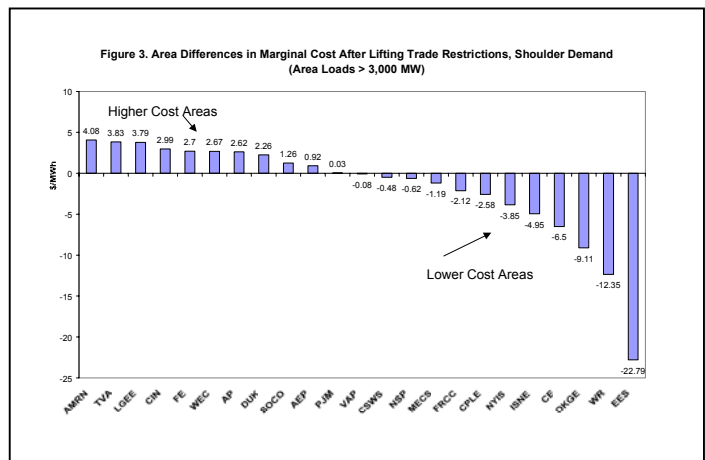
graphs LMPs at about 31,000 locations under both trade regimes, illustrates the degree to which prices are flattened by free trade.

Area interchange limits at peak demand account for about 4% of operating costs. They also permit high price pockets throughout the Eastern Interconnect. Recall that the short term operating cost of area interchange limits is the difference between the least cost of meeting demand while enforcing the limits and the least cost of meeting demand consistent with the transmission system’s physical capabilities. In reality the current balkanized transmission system is unlikely to minimize costs within regions. Consequently this estimate of the cost of area interchange limits understates actual costs.

Regime	Demand Levels		Cost Reductions	
	Peak	Shoulder	(\$)	%
Restricted	\$ 11.0	\$ 7.6	\$ 3.4	31%
Unrestricted	\$ 10.6	\$ 7.1	\$ 3.5	33%
Reductions				
(\$)	\$ 0.4	\$ 0.5		
Savings (%)	3.6%	6.6%		

Not surprisingly the costs of interchange limits increase during periods of shoulder demand (Table 6). When peak load is reduced by 20% (to about 404 GW), the cost of meeting demand falls over 30% in both trade regimes. System operating cost under restricted trade is about \$7.6 million per hour and under free trade is about \$7.1 million. The cost of restricting trade is about \$500 thousand/hour, about a 25% increase over the costs of restrictions at peak. Because the cost of restricting trade increases at the same time total costs decrease the cost of interchange limits as a percent of system operating costs increases from 3.6% to 6.6%.

Power flows generally proceed west to east, and north to south as the model seeks to minimize cost differences. The largest net exporters are TVA (4.4 GW), Duke (3 GW), and Southern Company (2 GW). Large net importers are Entergy (5.3 GW), CPLE (4 GW), FRCC (2.9 GW), Commonwealth Edison (2.6 GW), and ISNE (1.3 GW). Marginal cost in Entergy falls about \$23/MWh, or 43%, while costs in Ameren increases by about \$4/MWh (Figure 3). Note that, in this case, load in Ameren is only half that served in Entergy, indicating the value of the optimal power transactions.



4. Competitive Consequences of Area Interchange Limits: Examples from New England and Florida

This section shows that the choice of interchange limits can have a major impact on prices by limiting competition. New England and Florida are used as hypothetical examples. They were chosen in part because they are in established NERC regions (NPCC and FRCC respectively) that operate under NERC interchange limits. They differ in their physical ability to import power and in the mix of generators within their borders, leading to major differences in the impact of interchange (contingency) limits on competition.

In the case of New England, the n-1 contingencies limit the flow of imports to the region off-peak (Table 7). When imports are capped at the contingency level, 2,526 MW, generation of 16,200 MW is necessary to meet the ISO demand of 18,299 MW. Losses total 410 MW. These constraints result in a marginal cost of \$36.17/MWh. The minimum LMP in New England is \$9.17, and the maximum LMP is \$53.65.

Case Description	Marginal Cost	Imports	Revenue
	ISNE	(MW)	(000)
Off-peak demand, restricted trade	36.17	2,526	588.8
Off-peak, restricted, NE offers up 20%	43.52	2,526	707.5
Off-peak demand, optimized trade	27.39	4,016	401.5
Off-peak, with trade, NE offers up 20%	29.69	4,355	423.3

Suppose generator owners, confident that the contingency constraints would limit imports to about 2500 MW, were to independently increase their bids by 20% over (marginal) cost.²⁹ The average price would increase to \$43.52/MWh, the standard deviation of prices would increase to \$2.26/MWh, the maximum local price increases to \$65.04/MWh and the minimum is \$11.08/MWh. Imports remain fixed at 2,526 MW, and generator revenue increases to \$707,477 per hour. Since lines are at their area interchange limit imports cannot increase, and since demand does not react to price, generators could continue to increase their bids, consequently raising prices and profits.

Under unrestricted trade and competition, New England (Connecticut, Massachusetts, Maine, Vermont and New Hampshire) generators produce only 14,706 MWh, a 9 percent decrease.³⁰ Examination of the supply curve shows this level of production has a marginal cost of \$27.39 per MWh, which is the average competitive price.

Eliminating the line constraints reduces the average price (marginal cost) 24% compared to the corresponding restricted case. Competition and trade reduces price 37%, from \$43.52 to \$27.39, compared to the non-competitive case. Differences in the impact of local demand on congestion and line loss cause local prices to be dispersed around the average. The maximum local price is \$39.59, the minimum is \$9.61, and the standard deviation, a measure of spread of local prices about the average, is \$2.02. Generator revenue is \$401,492 per hour. The supply curve implies that a price of \$27.39 is not sufficient to induce generation within New England such that demand is met and line losses accounted for, so the model optimizes by importing 1,516 MW over the planning contingency of 2,526 MWh. Imports at these levels reduce marginal cost in New England by about 24%.

If generators attempt to raise prices by bidding 20% over (marginal) cost, imports would increase to 4,355 MWh and average price would drop to \$29.69.³¹ Imports are not at their physical limit-more could actually be imported- because at the margin it would cost more to increase output outside of New England than would be saved within New England. Absent the contingency constraint, generator revenues would fall to \$423,328 per hour and costs would fall because output declines to 14,391 MWh. Prices would still exceed the competitive standard of \$27.39 but not by as much as when the contingency constraints are applied.

²⁹ Contact Thomas Leckey for details of this case (NewIND80NE20d010402.pwb).

³⁰ Contact Thomas Leckey for details of this case (NewSA80d010402-a.pwb).

³¹Contact Thomas Leckey for details of this case (NewSA80NE20d010402-a.pwb).

Noncompetitive bidding would increase total profits for New England generators above the competitive level, but some higher cost generators that operated profitably under restricted trade conditions would be idle. Since the highest bidders are the ones that pay the price in lost profit, they would have substantial reason to bid true marginal cost. If they “break ranks,” prices will fall. These pressures will remain until all of the generators threatened with loss of output bid their true (marginal) costs. In this sense, free trade helps enforce competitive bidding on generators.

Like New England, Florida imports to the level of its contingency limit. Consequently, were generators to increase bids 20 percent above marginal cost, prices and profits would increase by nearly the full amount (Table 8). Florida’s consumers get less relief from imports than New England consumers. Floridians face higher cost generators and have less recourse to cheap imports. Under unrestricted trade, Florida generators produce 24,805 MWh,³² resulting in an average competitive price of about \$44.97 MWh with a standard deviation of \$7.32. Imports total 5,186 MWh.

Case Description	Marginal Cost ERCC	Imports (MW)	Revenue (000)
Off-peak demand, restricted trade	47.52	2,268	1,308
Off-peak, restricted, FL offers up 20%	57.09	2,268	1,572
Off-peak demand, optimized trade	44.97	5,186	1,088
Off-peak, with trade, FL offers up 20%	54.31	5,188	1,312

Were Florida generators to raise their offers by 20%, prices would increase to \$54.31 MWh.³³ Prices essentially increase as much as offers, but Florida generation is almost unchanged because the transmission system does not permit additional imports. Given that additional imports are not available and demand does not respond to price increases, generators can raise their prices and increase profits at will.

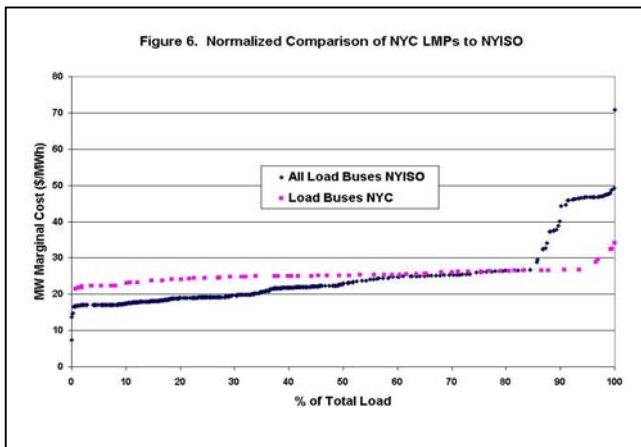
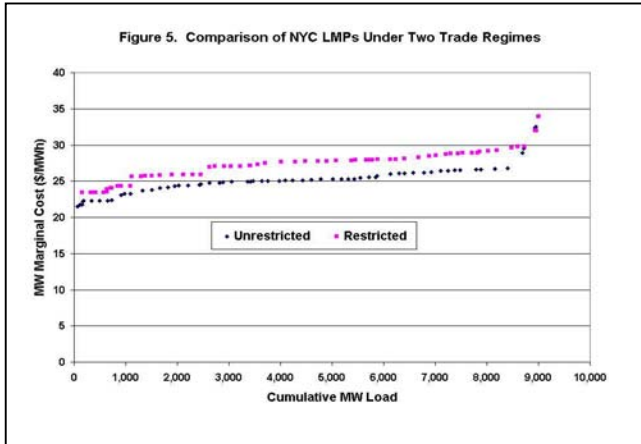
Florida’s electricity market is vulnerable to opportunistic pricing in the short run even with unrestricted trade. The summer limits exacerbate a difficult problem by constraining imports even during non-peak times. In the long term electrical connections could be strengthened to allow more imports. Perhaps lower cost generators could be built. But, in the short run competitive trade is not enough to force competitive pricing.

5. Implications of Load Pockets: New York City

Even if area interchange limits were eased, some local areas would continue to face persistently high costs by virtue of their location in the existing transmission grid. New York City, with just under 11 GW of load at peak, constitutes about 40% of peak load in the New York ISO. In 2000 the city was served by four large fossil plants sited within the city totaling about 5 GW generating capacity, and by three utility transmission interfaces which are capable of supplying the remainder of peak demand.

³² Contact Thomas Leckey for details of this case (NewSA80d010402-a.pwb).

³³ Contact Thomas Leckey for details of this case (NewSA80FL20d010402-a.pwb).



When trade restrictions are in place, LMPs in New York City run about 9% higher compared to an open trade regime (Figure 5).³⁴ The graph depicts 64 load bearing buses in the city under restricted and free trade, where about 9 GW of demand is present. When trade is introduced, the average cost at these 64 buses declines from \$27.64 MWh to \$25.28 MWh. Loadings on the thirteen main lines to the city approach capacity, a significant portion of which is dedicated to reactive power. Despite the cost reduction that comes along with increased trade, costs within the city remain significantly higher than in the NYISO as a whole (Figure 6). By means of a load normalized curve, the city's 9 GW of load is compared to the nearly 23 GW of load in the ISO.³⁵ Costs at the city load buses exceed those in the ISO for the first 80% of load, after which the high cost buses elsewhere in the State outstrip the costs in the city. Because access to the city through the existing interfaces is constrained, the area would remain vulnerable to market power, even with broad interregional trade. This situation is similar to the physical constraints observed in Florida.

7. Conclusion

Area interchange limits based on n-1 contingency studies have been an effective weapon for 40 years in limiting area blackouts in the Eastern Interconnection. However, area interchange limits are blunt tools, insensitive to the actual risks and opportunities facing the system in real time and ignorant of economic consequences.

This paper has demonstrated that the hidden costs of these limits can also be significant. Operating costs could be reduced on the order of 4-7% if the network were used to its physical potential. Subtler is the finding that area interchange constraints can shield noncompetitive pricing and inefficient operations. In the case of New England this shield is decisive. Even were New England's market perfectly competitive, the average price under area interchange limits would be 30% above those prevailing under free trade. Florida's consumers are disadvantaged by area interchange limits but removing them is not enough to compensate for a weak inter-regional transmission system and high cost generation. Florida generators are well protected either from pressures to economize or to price at

³⁴ NewSA80d010402-a.pwb. About 80% of peak load.

³⁵ The x-axis arranges the approximately 700 load-bearing buses in the ISO according to their share of the 23 GW. Similarly, each of the 64 city load buses constitute a share of the 9 GW load, so that both curves extend the full length of the x-axis. The 64 NYC buses are also represented in the ISO curve.

marginal cost. New York City's consumers face a constrained transmission system, resulting in prices higher than those obtained in the remainder of the ISO.

The tools for detecting, quantifying, and communicating risk are far advanced from what they were 40 years ago. And, the cost and competitive consequences of interchange limits are significant. It is probable that significant savings and efficiencies can be realized by re-examining reliability tools in the light of changed technology and market realities.

8. Acknowledgement and References

Research for the generator cost database was conducted by Ms. Marilyn Walker of the Department of Justice. A table matching NERC Names and reference numbers to those used by the EIA and U.S. Environmental protection agency are available from the authors on request.

9. Biographical Information

Douglas Hale (202.287.1723) is a senior economist at the Energy Information Administration (EIA) of the U.S. Department of Energy; Thomas Overbye is a professor of electrical engineering, University of Illinois; and Thomas Leckey (202.586.9413) is an industry specialist at EIA.