

LMPS FOR (TECHNICALLY-INCLINED) DUMMIES

David W. Savitski[*]

Synopsis: This article presents an elementary yet rigorous derivation of locational marginal prices (LMPs) for energy using a model with three injection or withdrawal points (a three-bus model), intended for technically-inclined readers wanting to learn the rudiments of LMP. While rigorous, the mathematics and intuition are presented in small steps over a series of increasingly complex examples. The problems encountered in setting up the model and solving for the LMPs are subtle, bearing such repetition. The basic model explores how transmission congestion raises generation costs, and two extensions explore how transmission congestion affects market power. While the examples are elementary compared to real-world systems, the model offers a rich set of insights. These concern congestion costs, constrained dispatch, deadweight loss, demand response, hockey-stick offers, inframarginal rents, LMP determination, market power, reference buses, shadow prices, transmission expansion, and transmission rents, among others. Analysts, lawyers, and judges involved with energy markets, especially those new to these markets, would thus benefit from a deeper understanding of LMP pricing, which is at the heart of wholesale energy markets.

* DavidSavitski42@gmail.com. The author previously worked at FERC, in the Office of Administrative Litigation. This work has benefited from comments from James Ballard, Joseph Cholka, Charlton Clark, Edward Gross, David Mead, Edward Mills, Guych Nuryyev, Ronald Ratty, Aaron Siskind, and two anonymous referees, none of whom are responsible for any errors.

SYMBOL KEY

AC = alternating current
DC = direct current
Δ = change in variable (*e.g.*, ΔQ = a change in Q)
df(x)/dx = the derivative (slope) of the function f(x)
$\partial f(x,z)/\partial x$ = the partial derivative (slope) of the function f(x,z)
GEN1, GEN2, GEN3 = generator 1, 2, and 3
$\int f(x)dx$ = integral, the area under the function f(x)
K_1, K_2, *etc*. = the capacity of generators 1, 2, *etc*.
L = Load
LMP = locational marginal price
$\mathscr{L}(\bullet)$ = the Lagrangian (the constrained objective function)
λ, λ_t, λ_1, λ_2, and λ_3 = the Lagrangian multipliers (shadow prices) for energy at the reference bus, transmission, and generators 1, 2, and 3
MC_1, MC_2, MC_3 = the marginal (incremental) cost of one MWh of energy supplied by generators 1, 2, and 3
MW = megaWatt
MWh = megaWatt hour
P_1, P_2, P_3 = price of energy (the LMPs) at buses 1, 2, and 3
π_1, π_2, π_3 = the profit of generators 1, 2, and 3
Q = Total energy demanded (and supplied)
Q_1, Q_2, Q_3 = the energy supplied by generators 1, 2, and 3
S_1, S_2 = flows of GEN1 and GEN2 over the constrained line *to the reference bus*
$\sum$ = summation
T = the maximum flow on the constrained transmission line
z = transmission line resistance

I. INTRODUCTION

Markets commonly use price to equate supply and demand, perhaps best illustrated by the stock market.[1] Many wholesale electricity markets similarly use price to equate the supply and demand for energy (transmission prices are still

1. INVESTOPEDIA, HOW DOES THE LAW OF SUPPLY AND DEMAND AFFECT THE STOCK MARKET? (Aug. 12, 2019), https://www.investopedia.com/ask/answers/040215/how-does-law-supply-and-demand-affect-stock-market.asp.

regulated), including CAISO, MISO, and PJM, among others.[2] Because the transmission of electricity is sometimes constrained, however, there is often not a single market-clearing price as in the stock market. Instead, prices often depend on location, hence "locational marginal prices" (LMPs). LMPs result from an optimization process whereby wholesale electricity prices in organized markets are set locally (at the bus level) based on least-cost dispatch, subject to generation and transmission constraints.[3] These constraints play a central role in dispatch and greatly complicate analyses of LMPs.[4] In such markets, load (electricity demand) pays the LMP at the bus where energy is taken and generators receive the LMP at the bus where energy is delivered.[5]

The Federal Energy Regulatory Commission (FERC) first approved LMPs for PJM in its November 25, 1997, order in Docket No. OA97-261.[6] This seminal case involved debate over the merits of LMP, which FERC evaluated and summarized in this order (at pages 37-51).[7] Proponents argued that LMPs send correct price signals for efficiency because the marginal benefit to load equals the marginal cost of delivering energy to each location, with transmission congestion costs reflected in LMP differences, unlike for average-cost pricing.[8] The

---

2. *See generally* Paul L. Joskow, *Challenges for Wholesale Electricity Markets with Intermittent Renewable Generation at Scale: The U.S. Experience*, MIT CTR. FOR ENERGY AND ENVTL. POL'Y RES. 1, 32 (2019). There are seven transmission system operators, classified as Independent System Operators (ISOs) and Regional Transmission Organizations (RTOs) in the U.S: California ISO (CAISO), Electric Reliability Council of Texas (ERCOT), ISO New England (ISO-NE), Midcontinent Independent System Operator, Inc. (MISO), New York ISO (NYISO), PJM Interconnection (PJM), and Southwest Power Pool (SPP). There is also a large non-RTO bilateral market in the Western Electricity Coordinating Council, comprised of Arizona, New Mexico, Southern Nevada Power Area, the Northwest Power Pool, and the Rocky Mountain Power Area. Centralized and bilateral markets differ in how prices are set. Centralized markets produce market-level prices by equating supply and demand. Bilateral markets, in contrast, produce individual-level prices as a result of decentralized trades negotiated between buyers and sellers. (Given trade between market participants across these organizational forms, these sets of prices will be correlated.) The Western Systems Power Pool, for example, facilitates trading by offering members a standardized default contract (for example, limiting details to specifying such terms as location, price, quantity, firmness, time). Both centralized and bilateral energy markets, however, must have system operators to schedule flows to respect transmission constraints, deal with imbalance energy, reserves, and so on.

3. *Price Formation in Organized Wholesale Electricity Markets*, Docket No. AD14-14-000 at 1 (Dec. 2014), https://www.ferc.gov/legal/staff-reports/2014/AD14-14-operator-actions.pdf (citing *Price Formation in Energy and Ancillary Services Markets Operated by Regional Transmission Organizations and Independent System Operators*, Notice, Docket No. AD14-14-000 (June 19, 2014)).

4. *See generally* Collin Cain & Jonathan Lesser, *A Common Sense Guide to Wholesale Electric Markets*, BATES WHITE ECON. CONSULTING, 20 (Apr. 2007).

5. Ezra Hausman et al., *LMP Electricity Markets: Market Operations, Market Power, and Value for Consumers*, SYNAPSE ENERGY ECON., INC. 4 (Feb. 5, 2006), http://www.synapse-energy.com/project/review-lmp-markets.

6. *Pennsylvania-New Jersey-Maryland Interconnection, et al.*, 81 F.E.R.C. ¶ 61,379, at p. 62,785 (1997) *citing* 81 F.E.R.C. ¶ 61,257 (1997), *order on clarification*, 82 F.E.R.C. ¶ 61,068 (1998), *order on reh'g and clarification*, 92 F.E.R.C. ¶ 61,282 (2000), *remanded on other grounds sub nom. Atlantic City Elec. Co. v. FERC*, 295 F.3d 1 (D.C. Cir. 2002). These orders, other issuances, and submissions can be found on FERC's website by searching eLibrary (http://elibrary.ferc.gov/idmws/docket_search.asp) by docket number.

7. FERC Docket No. OA97-261-001 (Nov. 25, 1997).

8. *Id.* at 14 (explaining that average-cost pricing is based on the total cost of the service, averaged over the total quantity).

proponents also argued that LMPs send correct signals for investment in the industry (*e.g.*, producers prefer to build generators in locations where LMP is high and large energy consumers prefer to locate their load where LMP is low) and for transmission use and planning, among other things.[9] Opponents argued that LMPs overcompensate inframarginal baseload units,[10] are based on offer prices rather than cost data, will not lower prices, are not known *ex ante*, and are overly complex.[11] While the merits of these points are not explored here, an understanding of LMP derivation sheds light on them, as noted periodically below.[12]

The LMP literature varies greatly based on approach, depending on audience and application. A relatively simple approach describes LMPs in general as consisting of three cost components (energy, transmission congestion, and transmission losses) of delivering a MWh of electricity to specific locations.[13] A more complicated approach explores efficiency, LMPs, market power, welfare effects, and so on using graphs and tables.[14] A more abstract approach explores similar issues using advanced mathematics, often vector calculus.[15]

The approach taken here is that of a textbook, combining elements of these three approaches, but with an elementary yet rigorous derivation of LMPs. This simplified approach emphasizes explanation and intuition, using a simple three-bus model with numeric examples.[16] It is intended for readers who have limited knowledge of LMPs and mathematics, yet nevertheless want an in-depth under-

9. FERC Docket No. OA97-261, at pp. 4-9 (Dec. 31, 1996).

10. FERC Docket No. AD08-4-000 (Apr. 30, 2008). If generators are arranged from low to high offer prices, the marginal unit is the most expensive unit dispatched; inframarginal units are those with lower offer prices and do not set the market price. *Appendix B: Overview of the U.S. Electric System*, EPA (Mar. 2016), https://www.epa.gov/sites/production/files/2016-03/documents/overview_of_the_electric_system.pdf.

11. FERC Docket No. OA97-261, *supra* note 9, at pp. 2-6.

12. For a subsequent evaluation of these points, *see* Hausman, *supra* note 5.

13. *See, e.g.*, CAISO, LOCATIONAL MARGINAL PRICING (LMP): BASICS OF NODAL PRICE CALCULATION (Dec. 6, 2005), https://www.caiso.com/Documents/02LMPOverview.pdf; MISO, PRICE FORMATION AT MISO MARKETS (Apr. 22, 2014), www.caiso.com/Documents/6_MidwestISO_MarketOverview.pdf. PJM's LMP training manual defines LMP as the sum of three components: system energy, congestion, and losses. PJM INTERCONNECTION, LLC, LOCATIONAL MARGINAL PRICING COMPONENTS (July 13, 2017), https://www.pjm.com/-/media/training/nerc-certifications/markets-exam-materials/mkt-optimization-wkshp/locational-marginal-pricing-components.ashx?la=en. These and the next sets of examples are to provide a flavor rather than a survey of the vast literature.

14. *See, e.g.*, Judith B. Cardell et al., *Market power and strategic interaction in electricity networks*, 19 RES. AND ENERGY ECON. 109-137 (1997); William W. Hogan, *Competitive Electricity Market Design: A Wholesale Primer* (Dec. 17, 1998), https://sites.hks.harvard.edu/fs/whogan/empr1298.pdf; Carolyn A. Berry et al., *Understanding How Market Power can Arise in Network Competition: A Game Theoretic Approach,* UTIL. POL'Y 8, 139-158 (Oct. 26, 1999); and Hausman, *supra* note 5.

15. *See, e.g.*, J. D. Weber & J. T. Overbye, *A two-level optimization problem for analysis of market bidding strategies*, 1999 IEEE POWER ENG'RING SOC'Y SUMMER MEETING CONFERENCE PROCEEDINGS (July 18–22, 1999), vol. 2, pp. 682–687, https://ieeexplore.ieee.org/document/787399; Yong Fu & Zuyi Li, *Different models and properties on LMP calculations*, 2006 IEEE POWER ENG'G SOC'Y GEN. MEETING (June 18-22 2006), https://pdfs.semanticscholar.org/01c7/31c5011c51893bb4769ef559f07834df0ad1.pdf?_ga=2.19311163.2082278785.1568057263-1753666943.1568057263; and H. Liu, L. Tesfatsion & A. Chowdhury, *Locational marginal pricing basics for restructured wholesale power markets*, 2009 IEEE POWER & ENERGY SOC'Y GEN. MEETING 1-8 (July 26-30, 2009), https://ieeexplore.ieee.org/document/5275503.

16. The models presented here have been analyzed in various forms by others; the contribution here is to explain the process in intuitive detail.

standing of LMPs. While rigorous, the mathematics and logic are presented in small steps, though simple calculus is used. The appendix presents the rudiments of differential and integral calculus and works out some of the examples presented here. The goal is to provide a detailed explanation of the model, including many of the subtle aspects often ignored in general presentations or easily overlooked in abstract presentations. Analysts, lawyers, and judges involved with energy markets, especially those new to these markets, would benefit from a deeper understanding of LMP pricing, which is at the heart of wholesale energy markets. Such an understanding allows for more accurate and insightful analysis in testimony, briefs, and decisions, and would enable greater access to the more theoretical analysis of energy markets found in journal articles and filed testimony.

The problems encountered in setting up the model and solving for the LMPs are complicated, subtle, and bear repetition. Hence, the model is developed over a series of increasingly complex examples. The initial examples involve competitive generators and fixed load. The first involves an unconstrained (benchmark) transmission system with generators dispatched in merit (least-cost) order to meet demand, producing uniform LMPs across buses, set by the marginal cost of the last generator needed to meet load. The second adds a binding transmission constraint, a situation where the unconstrained flow would exceed the line's capacity and so must be constrained by redispatching generators away from this least-cost benchmark solution. The redispatch raises production cost, illustrating congestion costs, and produces non-uniform LMPs across buses within an RTO.[17] The second example is extensively discussed, as many of the LMP basics are presented there: least-cost (constrained) dispatch, LMPs and shadow prices,[18] and the RTO choice of reference bus.

The remaining examples explore how transmission congestion affects market power. Specifically, the third introduces downward sloping demand, shifting the focus from minimizing cost to maximizing welfare (social surplus) – the difference between what load is willing to pay and the cost of the energy[19] – and sets up two market power examples. The fourth introduces market power for one generator, exercised by *decreasing* output to limit competitor access to transmission.[20] The fifth, involving a reconfigured system, illustrates generator market power exercised by *increasing* output, again to limit competitor access to

17. Bernard C. Lesieutre & Joseph H. Eto, *Electricity Transmission Congestion Costs: A Review of Recent Reports*, ERNEST ORLANDO LAURENCE BERKELEY NAT'L LAB. (Oct. 2003), https://www.energy.gov/sites/prod/files/oeprod/DocumentsandMedia/review_of_congestion_costs_october_03.pdf.

18. Shmuel S. Oren, *Capacity Payments & Supply Adequacy in Competitive Electricity Markets*, VII SYMP. SPECIALISTS IN ELEC. OPERATIONAL & EXPANSION PLAN., (May 21-26, 2000), https://oren.ieor.berkeley.edu/workingp/sepope.pdf. In general, a shadow price measures the benefit of having another unit of a constraining resource.

19. *Id*.

20. It does this by decreasing its output that flows counter to congestion. Since the counter-flow, in effect, creates additional capacity on the constrained transmission line over which the competitor ships energy, decreasing that counter-flow forces a decrease in the competitor's output.

transmission.[21] That market power might be exercised in some circumstances by decreasing output and in other circumstances by increasing output illustrates a challenge for detection.[22]

While the examples are simple compared to real-world systems, many model insights apply to more complex systems. These insights concern congestion costs, constrained dispatch, deadweight loss, demand response, hockey-stick offers, inframarginal rents, LMP determination, market power, reference buses, shadow prices, transmission expansion, and transmission rents, among others.[23] An understanding of the basic LMP model and these market concepts also makes the economics literature more accessible, allowing further exploration of these and other ideas.

II. THE MODEL

A three-bus model is used to explore LMPs, based on a standard DC approximation of the more complex AC system.[24] Such elementary models are common in the economics literature.[25] Some extensions of the elementary models are discussed in the Extensions section below.

The core grid model has three injection or withdrawal points, or buses.[26] Three buses are necessary to explore the problems caused by loop flow – the flow of electricity over multiple paths.[27] A one-bus model has generation and load adjacent to one another, making analysis of transmission simple.[28] A two-bus model allows generation and load to be adjacent to one another or to be separated.[29] Transmission congestion may now emerge, for example, if load is at one bus, generation is at the other, and the transmission line is too small during peak hours.[30] Market power may also be examined (as is the case with generation and load located at a single bus).[31]

A two-bus model, however, excludes loop flow, which is the property that electricity flows over all available paths, and which allows for more complicated

21. Cardell, *supra* note 14, at 122.

22. *See generally id.*

23. The definitions of these terms will be made clear as they are introduced in the examples below.

24. *See generally* Haifeng Lie, Leigh Tesfatsion & Ali A. Chowdhury, *Derivation of Locational Marginal Prices for Restructured Wholesale Power Markets*, 2 J. ENERGY MKTS. 3 (2009) (discussing DC and AC systems).

25. *See e.g.*, Cardell, *supra* note 14; and Berry, *supra* note 14 (exploring market power within such models, allowing different forms of competition among sellers).

26. Berry, *supra* note 14.

27. POWER WORLD CORP., *Solving the Power Flow*, https://www.powerworld.com/WebHelp/Content/MainDocumentation_HTML/Solving_the_Power_Flow.htm.

28. Berry, *supra* note 14, at 148.

29. Fu, *supra* note 15, at 1.

30. *Id.*

31. *See generally* William W. Hogan, *A Market Power Model with Strategic Interaction in Electricity Networks*, INT'L PROGRAM PRIVATIZATION & REG. REFORM HARV. INST. INT'L DEV. (July 15, 1997), https://sites.hks.harvard.edu/fs/whogan/hiid797b.pdf.

forms of market power. Loop flow also allows for cascading blackouts, such as the one that affected approximately 50 million people in the northeastern U.S. and Ontario, Canada on August 14, 2003.[32] Changes in loop flow occur when a generator or a transmission line trips (shuts down), forcing the electricity to flow over remaining parallel lines.[33] If the additional flows overload other lines they, too, could trip, leading to yet higher flows over remaining lines, and so on, resulting in a cascading blackout.[34] Such cascades are not possible in one- or two-bus models, as the outage is system-wide – either the line is open or tripped.[35]

A three-bus model is thus the simplest model to allow for analysis of congestion cost, loop flow, and market power in ways that are reasonable analytical approximations of reality.[36] And while more buses allow for more insights, they come at considerable analytical or computational expense (and so are beyond the scope of this article, though some additional insights are noted below).

Figure 1 below provides the model details for the initial examples. The three buses are represented by the three corners of the triangular grid, labeled 1, 2, and 3. The 1-2 line is arbitrarily constrained to a maximum flow of T = 100 MW; the other lines are assumed to have excess capacity at any load considered. The transmission constraint plays a central role in the analysis, just as in real-world dispatch.

Parenthetically, the *instantaneous* flow of electricity, as though time is frozen, is measured in MW. The *cumulative* flow of electricity over time, as discussed next, is measured in MWh. For example, one MW flowing for one hour or two MW flowing for half an hour equals one MWh.

32. U.S.-CANADA POWER SYSTEM OUTAGE TASK FORCE, FINAL REPORT ON THE AUGUST 14, 2003 BLACKOUT IN THE UNITED STATES AND CANADA: CAUSES AND RECOMMENDATIONS (Apr. 2004), www.ferc.gov/industries/electric/indus-act/reliability/blackout/ch1-3.pdf.

33. *Id.* at 6-7.

34. AM. SUPERCONDUCTOR CORP., 2004 ANNUAL REPORT 5 (2004), https://ir.amsc.com/static-files/5e1fc832-c4b6-40f6-9d31-4049746c07b3.

35. *Id.*

36. Fu, *supra* note 15, at 6.

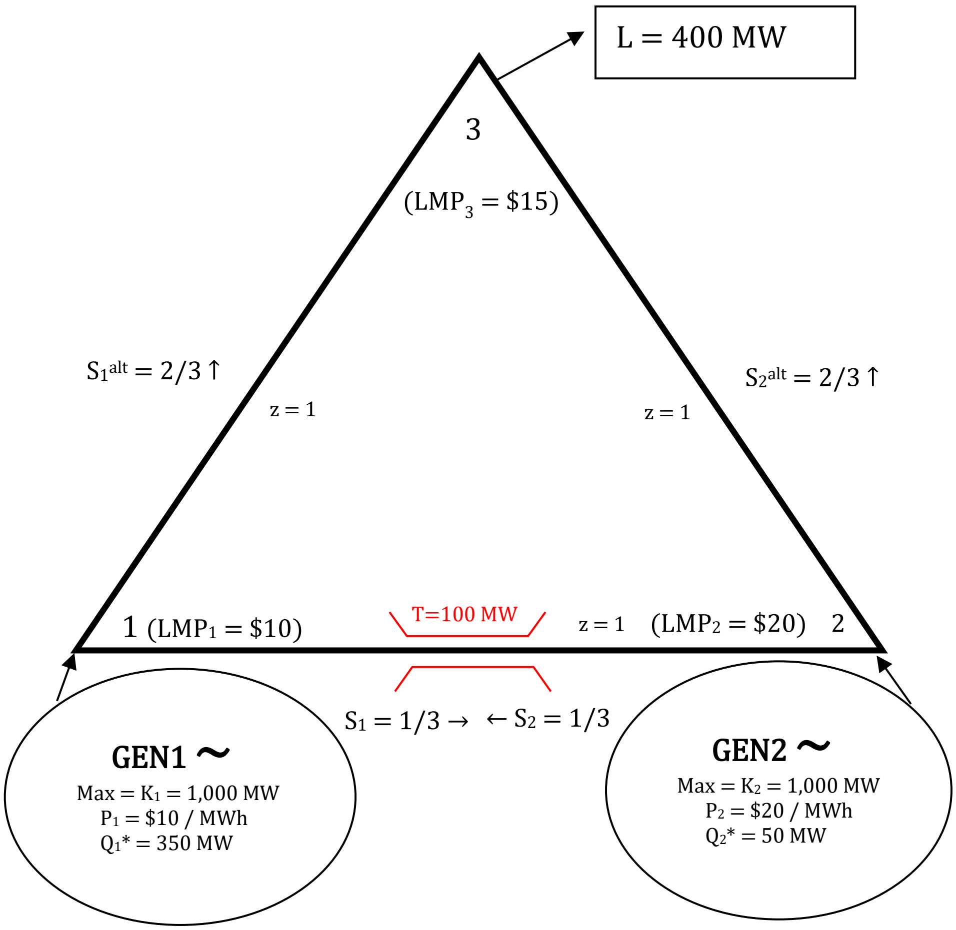

Figure 1. LMP in a 3-Bus, 1-Load Example.

Load ("L"), or electricity demand, is located at bus 3. It is initially fixed at L = 400 MW, and later varies according to the downward sloping demand curve P(Q) = 100 - Q/10.[37] Load is served by two (later, three) generators. The gener-

37. This demand curve is linear, with a vertical intercept (when quantity Q is zero) of price P(0) = 100. This is approximately the most buyers would pay for the first unit. The horizontal intercept (when Q is 1,000) represents how much buyers would take at a zero price. The constant slope of -1/10 means that, to encourage buyers to demand another unit, the price must be decreased by $1/10 per MWh. The negative slope forces generators with market power to consider the price-quantity trade-off: profit falls to zero when price is increased to $100/MWh, as quantity decreases to zero.

ator at bus 1, GEN1, has a capacity ("K") of K_1 = 1,000 MW and a constant marginal cost[38] ("MC") of MC_1 = \$10/MWh. The generator at bus 2, GEN2, has a capacity of K_2 = 1,000 MW and a constant marginal cost of MC_2 = \$20/MWh. Later, the generators are assumed to have increasing marginal costs. Generators are also initially assumed to competitively offer energy into the market at marginal cost.[39] In the last two examples, market power is introduced which produces offer prices that diverge from marginal cost.

As noted, central to the analysis is the constrained flow of energy over the transmission grid from generation to load.[40] Kirchhoff's Current Law implies that energy flows such that flow times resistance is equal over all paths.[41] Each transmission line in the three-bus model has equal resistance, indicated by z = 1.[42] For example, suppose GEN1 injects 1 MW at bus 1 for sale at bus 3. The energy flows over both paths. Since each segment is assumed to have equal resistance, the long path (1-2-3) has two segments and, therefore, twice the resistance of the short path (1-2). Hence, half as much energy will flow over the long path (x) as over the short path. Given that 1MW is delivered to bus 3, we have x + 2x = 1 MW, and so x = 1/3 MW flows over the long path and 2/3 MW flows over the short path.[43] These are indicated by the shares ("S") S_1^{alt} = 2/3 and S_1 = 1/3 in Figure 1, with the arrows indicating flow direction.[44] The subscript indicates GEN1's (S_1) or GEN2's (S_2) flow.[45] Since the transmission constraint is central, the focus is on the shares flowing over the constrained line (the shift factors), S_1 and S_2, to reference bus 3. It is thus easy to see how, with parallel loop flows combined with binding transmission constraints, one generator's output can limit another generator's output, an analysis that helps demonstrate how market power might be exercised.

38. Puneet Chitkara & Jin Zhong, *The IEEE/PES Transmission & Distribution Conf. & Exposition, Equilibrium Analysis in Imperfect Traders' and GenCos' Market* 4-5 (Apr. 19-22, 2010), http://hdl.handle.net/10722/126137.

39. A noted criticism of LMP is that dispatch is based on offers rather than on marginal cost. One justification is that competition pressures generators to make offers close to marginal cost. If there is no capacity market, however, as in ERCOT, then the competitive price must at times exceed marginal cost if the marginal generator is to cover capacity cost from the energy market. Later, strategies such as hockey-stick offers, where energy is offered at low prices for most of the units offered, but at high prices for the last units offered, are explored. *See* Berry, *supra* note 14, at 140-144.

40. Berry, *supra* note 14, at 143-144.

41. *See, e.g.*, ELECTRONICS TUTORIALS, KIRCHHOFF'S CURRENT LAW, https://www.physics.uoguelph.ca/tutorials/ohm/Q.ohm.KCL.html. Kirchhoff's Current Law describes how electricity flows over all available transmission lines from generator to load. *Id.*

42. PJM, BASICS OF ELECTRICITY: POWER FLOW ON AC TRANSMISSION LINES (July 11, 2013), https://www.pjm.com/-/media/training/nerc-certifications/trans-exam-materials/bet/bet-lesson5-power-flow-on-ac-transmission-lines.ashx?la=en.

43. Another way to see this is to equate the flow over the short path (S) times the resistance (z) to the flow over the long path (L) times the resistance (2z). This yields Sz = L2z, implying S = 2L. Then, since one MW is to be delivered to bus 3, S + L = 1. Solving the two equations yields S = 2/3 MW and L = 1/3 MW.

44. *See* Figure 1.

45. *Id.*

The analysis focuses on optimal, least-cost constrained dispatch under various system conditions; it excludes the *adjustment process* whereby generators move from one dispatch to another.[46] Hence, adjustment complications, such as generator start-up and shut-down costs, minimum-run times or quantities, ramp rates, and so on, are ignored.[47] Line losses are also ignored.[48]

Finally, the RTO (or ISO) is the central coordinator of organized wholesale energy markets.[49] It accepts generator offers to sell energy, constructs a supply curve, and dispatches generators based on the supply curve to meet demand in real time.[50] This serves load at the least cost, subject to generator and transmission constraints, and determines the resulting LMPs.[51] Loads are charged the LMP at the bus where they consume energy and generators are paid the LMP at the bus where they deliver energy.[52]

III. EXAMPLES

The three-bus model is explored below over several examples of increasing complexity. The first involves an unconstrained system with fixed (vertical) demand.[53] Without transmission congestion, generators are dispatched in merit order, *i.e.*, in order of increasing marginal cost.[54] This forms the least-cost benchmark for generating a given amount of energy.[55] Then, a binding transmission constraint is imposed, necessitating out-of-merit-order dispatch to prevent line overload.[56] This increase in production cost defines the congestion cost.[57]

46. *See generally* Fu, *supra* note 15.

47. *Id*.

48. INMR, USING ARRESTERS TO REDUCE TRANSMISSION LINE LOSSES (June 11, 2016), https://www.inmr.com/using-arresters-reduce-transmission-line-losses/.

49. Patrick McGarry, *Why We Need RTO Markets*, ENERGY CENT. (Aug. 5, 2019), https://www.energycentral.com/c/pip/why-we-need-rto-markets.

50. RTOs, of course, do more than dispatch energy in real time, involving other products (*e.g.*, capacity, financial transmission rights, reactive power, and reserves) and functions (*e.g.*, administration, market monitoring, seams issues, and transmission planning).

51. PJM, ENERGY MARKETS LOCATIONAL MARGINAL PRICING (2011), https://www.pjm.com/~/media/training/nerc-certifications/em2-LMP.ashx.

52. AEP ENERGY, TRANSMISSION CONGESTION: HOW DOES THIS AFFECT YOUR ENERGY PRICE? (June 5, 2018), https://www.aepenergy.com/2018/06/05/transmission-congestion-affect-energy-price/.

53. T. Nireekshana et al., *Locational Marginal Pricing Calculation with Rescheduling of Generation in Deregulation,* 1 INT'L J. OF ENG'G RES. & DEV. 61 (2012). In general, references to a constraint will concern transmission, as the bulk of the analysis focuses on transmission. When generator constraints are discussed, generator capacity will be explicitly noted.

54. Lesieutre, *supra* note 17, at 31. Merit order can be quite complicated to determine when the offers themselves are complicated and because of physical constraints on generators. For example, offers might have steps that start high and gradually increase or start low and rapidly increase. Similarly, there may be minimum run times, minimum shutdown times, and so on for generators. For a given load, all combinations of generators must be considered to determine the least-cost dispatch. But, as noted, the (dynamic) adjustment process is ignored to focus on the (static) equilibrium dispatch, so many of these complications are also ignored.

55. *Id.* at 6, 10.

56. *Id.* at 12 n.9.

57. *Id.* at 12.

Next, downward sloping demand is added. This changes optimization from minimizing cost to maximizing the net gain from energy production and consumption (maximizing "social surplus"), and allows market power to be examined in the last two examples.[58] The subtlety of the constraints is also explored in how they impact the dispatch, the LMPs, and the exercise of market power. For example, one configuration allows market power to be exercised by decreasing output while another configuration allows market power to be exercised by increasing output, illustrating a challenge for detecting the exercise of market power.

A. The Unconstrained Benchmark

Without a binding transmission constraint, generators are dispatched in merit order until load (400 MW) is served.[59] This defines the least-cost benchmark.[60] Since GEN1 is the cheapest generator, it is dispatched first.[61] And since its capacity exceeds load, it is dispatched to 400 MW and GEN2 is not dispatched.[62] The LMPs are uniform without congestion, and determined by the marginal generator, GEN1, the last (most expensive) generator dispatched.[63] Since another MWh consumed at any bus would require the additional dispatch of GEN1 at $10/MWh, the LMP at each bus is $10/MWh. Economically, the model collapses to a single bus with a single LMP. Total production cost is minimized at 400 MWh x $10/MWh = $4,000. This is also the amount billed to load.

Load follows a standard profile over the day, with a peak period midafternoon, surrounded by shoulder periods, and an off-peak period at night.[64] As load exceeds 1,000 MW in the example, GEN2 must be dispatched. For example, suppose that load is 1,500 MW. GEN1 is dispatched to 1,000 MW (its capacity), and GEN2 is dispatched to 500 MW. The LMPs are again determined by the marginal generator, now GEN2. Another MWh consumed at any bus requires the additional dispatch of GEN2, at $20/MWh, so all LMPs are now $20/MWh. Total production cost is minimized at 1,000 MWh x $10/MWh + 500 MWh x $20/MWh = $20,000.

A difference now emerges between the amount billed to load and paid to generators ($30,000/hour) and the cost of production ($20,000/hour). The difference arises because, as soon as GEN2 is needed, GEN1 is paid more than its cost – it receives "rents." Rents are a net gain, an excess of revenue over cost.

---

58. *Id.* at 34.

59. Hausman, *supra* note 5, at 35.

60. *Id.*

61. Fu, *supra* note 15, at 2; Sumit Kumar & Harinder Singh Sandhu, *Economic Dispatch & Its Various Impacts on Power System Generation,* 3 INT'L J. OF ENHANCED RES. IN SCI. TECH. & ENG'G 275 (Sept. 2019).

62. Kumar, *supra* note 61.

63. *Id*.

64. ENERGY INSIDER, ENERGY USE PATTERNS: UNDERSTANDING YOUR LOAD PROFILE, http://members.questline.com/Article.aspx?articleID=779&accountID=296&nl=10006.

The significance of rents is that such payments are not necessary to induce generators to operate.[65] Once GEN2 is dispatched, all LMPs increase from \$10/MWh to \$20/MWh, resulting in a profit for GEN1 of \$10,000 per hour. Note that GEN1 is not exercising market power, as it competitively offers energy at marginal cost.

A related insight illustrates why the introduction of LMPs has stimulated demand response.[66] Suppose that load is initially 1,000 MW. The total amount billed to load per hour is 1,000 MWh x \$10/MWh = \$10,000. Now suppose that load increases to 1,001 MW. Since GEN1 is at capacity, GEN2 must generate the 1,001st MW, raising the LMP to \$20/MWh for all units. Hence, the total amount billed to load per hour is now 1,001 MWh x \$20/MWh = \$20,020. The 1,001st MW thus adds \$10,020 to load's bill.[67] This justifies paying customers to decrease load (demand response), to avoid dispatching an expensive generator or, more generally, from jumping to the next step on the supply curve as load would otherwise increase.

B. Competition Under Fixed Demand

Now consider the 100 MW transmission constraint on the 1-2 segment. Since S_1 = 1/3 of GEN1's output flows over the line, Q_1 x S_1 = 400 x (1/3) = 133 1/3 > 100, the constraint binds. Hence, the constraint is violated if GEN1 serves the entire 400 MW load. To prevent overload, GEN2 must displace some of GEN1's output, as it flows counter to GEN1's, decreasing the net flow on the 1-2 line. Redispatch (an adjustment to a prior dispatch) away from the unconstrained, least-cost dispatch gives rise to congestion costs (the redispatch-induced cost increase) and to non-uniform LMPs.[68]

65. Hence, the criticism by LMP opponents, noted above, who argued that inframarginal base-load units are over-compensated. Because their costs are low, they have an incentive to offer energy at marginal cost, as that ensures dispatch, while the marginal generator offers energy at a higher price, setting a higher LMP from which base-load units benefit.

66. Demand response involves paying load to decrease consumption relative to some base-load quantity. FERC mandated that RTOs and ISOs treat demand response on par with dispatchable generation resources. Final Rule, *Wholesale Competition in Regions with Organized Elec. Mkt.*, 125 F.E.R.C. ¶ 61,071 (2008) (to be codified at 18 C.F.R. pt. 35); *Order on reh'g.*, 128 F.E.R.C. ¶ 61,059 (2009); *Order on reh'*g., 129 F.E.R.C. ¶ 61,252 (2009). FERC also set the price for compensating demand response at LMP. Final Rule, *Demand Response Compensation in Organized Wholesale Energy Markets*, 134 F.E.R.C. ¶ 61,187 (2011) (to be codified at 18 C.F.R. pt. 35); *Order on reh'g and Clarification*, 137 F.E.R.C. ¶ 61,215 (2011); *Reh'g denied*, 138 F.E.R.C. ¶ 61,148 (2012). Order No. 745 was ultimately upheld by the Supreme Court. FERC v. Elec. Power Supply Ass'n, 136 S. Ct. 760, 784 (2016).

67. FERC Docket No. ER03-1086-001 et al. (July 9, 2004). The marginal cost is the \$20/MWh incurred by GEN2 for the unit it produced. But since load pays LMP, its total bill will more than double in this example, as the LMP on all units double. This also illustrates a complexity of merit order – the offer steps can create significant jumps in total cost for small changes in output. Under average-cost pricing, in contrast, the marginal cost of \$20/MWh gets averaged over 1,001 units, with little effect on rates. A second impetus for demand response is the increasing integration of renewables, as their non-dispatchability puts additional burden on thermal reserves. Demand response can substitute, albeit imperfectly, for reserves.

68. Mathew Morey, *Power Market Auction Design*, EDISON ELEC. INST. (2001), http://web.mit.edu/esd.126/www/MktsAuctions/EEI.pdf.

The RTO dispatches generators to minimize generation costs, as before, but now must do so respecting the transmission constraint.[69] Mathematically, the optimization gets more complicated, however, as congestion introduces a second dimension to the problem, and so direct application of the prior optimization technique to the constrained problem will yield incorrect results. A new technique is thus needed that incorporates the constraint into the optimization problem. Mathematics has a tool that accomplishes this task - minimizing cost subject to a transmission (or other) constraint – called the Lagrangian.[70] (It naturally extends to multiple constraints.)

Minimizing the Lagrangian thus minimizes total energy cost, as before, but now subject to the constraints. The Lagrangian has an additional benefit: it provides an implicit price for each constrained resource, measuring the benefit of having another unit of the resource in question on energy cost.[71] This price (benefit) can then be compared with the cost of the resource to guide investment.

1. Objective Function

In terms of notation, let $\mathcal{L}(\bullet)$ denote the Lagrangian function.[72] Also, let the λs denote the Lagrangian multipliers, the implicit (shadow) prices of the constraining resources.[73] (They are discussed in detail in the next section.) The bracketed comments describe the function and the constraints. Solving for the optimal dispatch now means minimizing the cost of serving the fixed load, as before, but subject to the constraints:

$$
\begin{aligned}
(1)\quad & \mathcal{L}(Q_1, Q_2, \lambda, \lambda_t, \lambda_1\, \lambda_2) & \\
& = 10Q_1 + 20Q_2 & \text{[total (variable) cost]} \\
& + \lambda(L - Q_1 - Q_2) & \text{[ensures supply } (Q_1 + Q_2) \text{ equals demand (L)]} \\
& + \lambda_t(T - S_1Q_1 - S_2Q_2) & \text{[ensures that the transmission constraint holds]} \\
& + \lambda_1(K_1 - Q_1) & \text{[ensures GEN1 does not exceed its capacity, } K_1\text{]} \\
& + \lambda_2(K_2 - Q_2) & \text{[ensures GEN2 does not exceed its capacity, } K_2\text{].}
\end{aligned}
$$

69. U.S. DEP'T OF ENERGY, THE VALUE OF ECONOMIC DISPATCH 28 (2005), https://www.energy.gov/sites/prod/files/oeprod/DocumentsandMedia/value.pdf.

70. *See, e.g.*, MALCOLM PEMBERTON & NICHOLAS RAU, MATHEMATICS FOR ECONOMISTS: AN INTRODUCTORY TEXTBOOK (4th ed. 2015); ALPHA C. CHIANG & KEVIN WAINWRIGHT, FUNDAMENTAL METHODS OF MATHEMATICAL ECONOMICS (4th ed. 2005) (referencing the Lagrangian method, as well as much of the other mathematics used here). The appendix below presents a primer on differential and integral calculus, including the basics of the Lagrangian and the optimization. Several of the examples discussed here are also worked out in detail.

71. E. INTERCONNECTION STATES PLANNING COUNCIL, CO-OPTIMIZATION OF TRANSMISSION AND OTHER SUPPLY RESOURCES 39 (Sept. 2003), https://pubs.naruc.org/pub.cfm?id=536D834A-2354-D714-51D6-AE55F431E2AA.

72. KHAN ACADEMY, LAGRANGE MULTIPLIERS, https://www.khanacademy.org/math/multivariable-calculus/applications-of-multivariable-derivatives/constrained-optimization/a/lagrange-multipliers-single-constraint.

73. *Id.*

The first term on the right-hand side, $10Q_1 + 20Q_2$, is the total (variable) cost, where Q_1 costs \$10/MWh and Q_2 costs \$20/MWh.[74] The Lagrangian has the property that the remaining terms are all zero at the optimum, as discussed next. Hence, minimizing $\mathcal{L}(\bullet)$ means minimizing variable cost but, crucially, subject to the constraints.

The second term, $\lambda(L - Q_1 - Q_2)$, ensures that total supply $(Q_1 + Q_2)$ equals total demand (L = 400). Since supply must equal demand, $L - Q_1 - Q_2 = 0$, and so the product $\lambda(L - Q_1 - Q_2)$ is zero.

The third term, $\lambda_t(T – S_1Q_1 - S_2Q_2)$, ensures that the flow across the 1-2 line respects the constraint (T = 100 MW). The output of GEN1, which is dispatched first and to the extent possible (since it is the cheapest), flows across the line from left to right. This flow consumes $S_1 = 1/3$ MW of the transmission capacity per MW generated, and hence S_1Q_1 (= $Q_1/3$) is subtracted from capacity, T. The output of GEN2 flows counter to that, reducing the net flow by 1/3 MW ($S_2 = -1/3$) for each MW generated, in effect adding capacity. Hence, $Q_2/3$ is added to T. The net flow across the line (from left to right) is $S_1Q_1 + S_2Q_2 = Q_1/3 - Q_2/3$. Since the constraint binds, based on the prior unconstrained example, $T = Q_1/3 - Q_2/3$. Rearranging gives the term in parentheses, $T - S_1Q_1 - S_2Q_2$ (= $T - Q_1/3 + Q_2/3 = 0$). If the transmission constraint did not bind, so that $T - S_1Q_1 - S_2Q_2 > 0$, then $\lambda_t = 0$ (as shown below). Hence, the product $\lambda_t(T – S_1Q_1 - S_2Q_2)$ is always zero.

The fourth and fifth terms are symmetric, so consider just the fourth, $\lambda_1(K_1 - Q_1)$. This ensures that GEN1's output Q_1 does not exceed its capacity (K_1 = 1,000). When it is at capacity, $K_1 - Q_1 = 0$. If the generator was not at capacity, so that $K_1 - Q_1 > 0$, then $\lambda_1 = 0$. Again, the product $\lambda_1(K_1 - Q_1)$ is always zero at the optimum.

2. Shadow Prices

The Lagrangian multipliers, the λs, are the implicit or shadow prices of the constraining resources, determined by generator offers and optimal dispatch.[75] Shadow prices indicate the impact on variable cost (the objective function) of loosening the constraints by one unit.[76] Binding generation constraints imply that more expensive generators are dispatched and binding transmission constraints imply out of merit-order dispatch, both of which increase generation cost

74. Fixed costs could be added as constants in the generators' cost functions. Under the optimization, however, they drop out, and so do not affect the least-cost dispatch. Since the goal is to explain optimal dispatch, rather than the adjustment from one optimal dispatch to another (which can involve costs unrelated to energy), they are ignored. Young-Beom Jung et al., *A Study of Optimization Modeling Generation Cost under the Stable Operation of Power System* INT'L COUNCIL ON ELEC. ENG'G. 208, 210 (2012), https://doi.org/10.5370/JICEE.2012.2.2.208.

75. David Albouy, Berkley, *Constrained Optimization*, SHADOW PRICES 4 https://eml.berkeley.edu/~webfac/saez/e131_s04/shadow.pdf.

76. *Id.*

(variable cost),[77] as this example shows. Loosening a generation or transmission constraint therefore lowers cost, and so the shadow prices of generation and transmission are negative (when the constraints bind) or zero (otherwise).[78]

Mathematically, this is seen by differentiating with respect to a constraining variable; *i.e.*, by asking what happens to total cost when the constraining resource is made more abundant.[79] Consider, *e.g.*, an increase in load. Differentiating with respect to load yields $\partial\mathcal{L}(\bullet)/\partial L = \lambda$. Without going into the details of derivatives, and focusing instead on their intuition, note that $\partial\mathcal{L}(\bullet)/\partial L \approx \Delta\mathcal{L}(\bullet)/\Delta L$, where "Δ" means "change in." That is, the derivative ($\partial\mathcal{L}(\bullet)/\partial L$) tells us how variable cost changes as load (alone) increases by one MWh ($\Delta L = 1$): $\Delta\mathcal{L}(\bullet)/\Delta L = \Delta VC/\Delta L = \Delta VC/1$. The derivative measures the relationship for "very small" changes in load whereas the delta form does so for larger, discrete changes in load. The two are reasonably close if delta is small. This indicates the cost of delivering a MWh to bus 3. Note again that, at the optimum, all terms except variable cost (VC) are zero: $\mathcal{L}(\bullet) = VC = 10Q_1 + 20Q_2$ (plus zero terms), so a change in the Lagrangian is a change in variable cost. Hence, λ indicates the cost (shadow price) of delivering another MWh to bus 3, given system constraints and generator offers. It is thus positive.

Similarly, consider an increase in the constrained transmission resource, T. Differentiating with respect to T yields $\partial\mathcal{L}(\bullet)/\partial T = \lambda_t$. This measures the impact on variable cost when the transmission constraint is loosened (T is increased) by 1MW. Since loosening the binding constraint implies more efficient dispatch, variable cost falls, and so $\lambda_t < 0$.[80] If the line is not constrained, additional capacity would not lead to redispatch, meaning there is no impact on variable cost, and so $\lambda_t = 0$.

Finally, consider an increase in generator capacity. For GEN1, $\partial\mathcal{L}(\bullet)/\partial K_1 = \lambda_1$ measures the impact on variable cost when the generator constraint is loosened (K_1 is increased) by 1MW. If GEN1 is at capacity, then $1{,}000 - Q_1 = 0$, and its shadow price is negative, $\lambda_1 < 0$. Another, more expensive, generator is being dispatched. An increase in capacity would thus allow for more output from the (constrained) efficient generator, and less from the more expensive generator, decreasing variable cost. If it has excess capacity, so that $1{,}000 - Q_1 > 0$, then its

77. *See generally* Avinash Vijay et al., *The value of electricity and reserve services in low carbon electricity systems*, 201 APPLIED ENERGY 111 (2017).

78. *See generally* Richard Green, *Electricity Transmission Pricing: How much does it cost to get it wrong?* U. CAL. ENERGY INST. 6 (1998). Later, when optimization involves maximizing social surplus (the value load places on the energy less the cost of the energy to the generators), the shadow prices of generation and transmission are positive. These are not inconsistent results. Here, since cost is being minimized, a looser constraint means lower cost, which implies a positive gain in social surplus.

79. Differentiating a function yields a function that indicates the slope of the original one. It is commonly set to zero to solve for the value of the variable in question that maximizes or minimizes the objective function. The appendix discusses differentiation in some detail.

80. Once this example is complete, the solution is compared to the unconstrained benchmark case. The comparison will show how the binding constraint increases generation costs. Increasing T thus reverses this, decreasing generation costs.

shadow price is zero, $\lambda_1 = 0$. Since the generator is below capacity, adding to its capacity would not change dispatch, and so would not change variable cost.

3. Optimization

The Lagrangian, with the parameter values inserted, is

$$(2)\quad \mathcal{L}(Q_1, Q_2, \lambda, \lambda_t, \lambda_1\, \lambda_2) = 10Q_1 + 20Q_2 + \lambda(400 - Q_1 - Q_2) + \lambda_t(100 - Q_1/3 + Q_2/3) + \lambda_1(1{,}000 - Q_1) + \lambda_2(1{,}000 - Q_2).$$

In general, optimization is akin to getting to the top of a hill (*e.g.*, when maximizing profit) or getting to the bottom of a valley (*e.g.*, when minimizing cost), which implies that the slope is zero there. Hence, to solve the optimization problem, the slope with respect to each choice variable – here, the generator dispatches – and the Lagrangian multipliers (this keeps the constraints as part of the solution equations) are all set to zero. For example, GEN1's output (alone) is varied until cost is at a minimum (the zero-slope condition for GEN1), given the output of GEN2, demand, and so on.[81] The same is done for GEN2. Then these conditions, along with the constraints, are solved simultaneously for the optimal outputs. Mathematically, then, to minimize the Lagrangian, differentiate equation (1) with respect to the generator output variables and the Lagrangian multipliers. The derivatives (slopes) are set to zero.

A constrained solution is more complicated than an unconstrained one, however, as the constraints might not bind, depending on system conditions. For example, during off-peak periods, transmission lines are generally not constrained.[82] But during peak periods, they might be.[83] To allow for that possibility, the zero-slope conditions are modified to have three parts: (1) the slope with respect to each variable might be zero (or not), (2) the variable itself might be zero (or not), and (3) the product of the first two conditions, which must be zero (because at least (1) or (2) must be zero).[84] These three conditions are known as the Kuhn-Tucker conditions.[85] Hence, the zero-slope conditions for a constrained optimum that reflect these three conditions are

$$(3a)\quad \partial\mathcal{L}(\bullet)/\partial Q_1 = 10 - \lambda - \lambda_t/3 - \lambda_1 \geq 0; \qquad Q_1 \geq 0;\ \text{and} \qquad Q_1 \text{ x } \partial\mathcal{L}(\bullet)/\partial Q_1 = 0$$

$$(3b)\quad \partial\mathcal{L}(\bullet)/\partial Q_2 = 20 - \lambda + \lambda_t/3 - \lambda_2 \geq 0; \qquad Q_2 \geq 0;\ \text{and} \qquad Q_2 \text{ x } \partial\mathcal{L}(\bullet)/\partial Q_2 = 0$$

$$\mathbf{(3c)\quad \partial\mathcal{L}(\bullet)/\partial\lambda = 400 - Q_1 - Q_2 = 0; \qquad \lambda \geq 0;\ \text{and} \qquad \lambda \text{ x } \partial\mathcal{L}(\bullet)/\partial\lambda = 0}$$

81. *See* Brent Eldridge et al., *Pricing in Day-Ahead Electricity Markets with Near-Optimal Unit Commitment*, U. CAMBRIDGE ENERGY POL'Y RES. GRP., at 2 (2018). Strictly speaking, the intuition does not hold when the objective function (here, total cost) and the constraints are linear. In such cases, the zero-slope conditions produce an inconsistent set of equations, as will be seen below. Nevertheless, the intuition is useful and holds in many cases, including those considered below.

82. *Id.* at 5.

83. *Id*.

84. *Id.* at 2-9, 17.

85. Albouy, *supra* note 75, at 3.

(3d) $\partial\mathcal{L}(\bullet)/\partial\lambda_t = 100 - Q_1/3 + Q_2/3 \geq 0$;[86] **$\lambda_t \leq 0$; and $\lambda_t \times \partial\mathcal{L}(\bullet)/\partial\lambda_t = 0$**
(3e) $\partial\mathcal{L}(\bullet)/\partial\lambda_1 = 1{,}000 - Q_1 \geq 0$; $\lambda_1 \leq 0$; and $\lambda_1 \times \partial\mathcal{L}(\bullet)/\partial\lambda_1 = 0$
(3f) $\partial\mathcal{L}(\bullet)/\partial\lambda_2 = 1{,}000 - Q_2 \geq 0$; $\lambda_2 \leq 0$; and $\lambda_2 \times \partial\mathcal{L}(\bullet)/\partial\lambda_2 = 0$.[87]

These zero-slope conditions involve six equations and six unknowns (Q_1, Q_2; λ, λ_t, λ_1, λ_2). One technique (that works here) to solve these equations is to focus on subsets of equations. For example, equations (3c) and (3d) (boldfaced) involve two equations and two unknowns. Equation (3c) holds with equality, $400 - Q_1 - Q_2 = 0$, since demand must equal supply. Equation (3d) also holds with equality, $100 - Q_1/3 + Q_2/3 = 0$, as the transmission constraint binds. Rearranging yields $Q_1 + Q_2 = 400$ and $Q_1 - Q_2 = 300$. Adding these two equations eliminates Q_2, so $2Q_1^* = 700$, or $Q_1^* = 350$. Equation (3c) then implies $Q_2^* = 400 - 350 = 50$.

With the generator outputs determined, solving for the multipliers is feasible, beginning with (3e) and (3f). Since the analysis of the two generators is symmetric, as neither GEN1 nor GEN2 is at capacity, consider again GEN1. Since the generator is not constrained, the slope is positive. That is, $Q_1 < 1{,}000$ implies $\partial\mathcal{L}(\bullet)/\partial\lambda_1 = 1{,}000 - Q_1 = 1{,}000 - 350 = 650 > 0$. Since the product is zero, $\lambda_1 \times \partial\mathcal{L}(\bullet)/\partial\lambda_1 = 0$, then $\lambda_1^* = 0$: having another unit of capacity for GEN1 means more idle capacity, with no change in cost, so the shadow price (the benefit in decreased cost) is zero. Symmetry implies that $\lambda_2^* = 0$.

Finally, with $\lambda_1^* = \lambda_2^* = 0$, equations (3a) and (3b) reduce to two equations and two unknowns (λ and λ_t). Since Q_1 and Q_2 are positive, the third (product) terms in equations (3a) and (3b) mean that the slopes are zero:

$\partial\mathcal{L}(\bullet)/\partial Q_1 = 10 - \lambda - \lambda_t/3 = 0$; and
$\partial\mathcal{L}(\bullet)/\partial Q_2 = 20 - \lambda + \lambda_t/3 = 0$.[88]

Adding equations yields $2\lambda^* = 30$, or $\lambda^* = 15$. Plugging $\lambda^* = 15$ into the first equation yields $15 = 10 - \lambda_t^*/3$, or $\lambda_t^* = -15$.

---

86. *See* Eldridge, *supra* note 81, at 2-3. Note that the transmission constraint in the Lagrangian cannot be simplified by multiplying by three to remove the fractions ($300 - Q_1 + Q_2$), as this changes the zero-slope conditions. Multiplying the constraint in the Lagrangian by three would change (3a) and (3b), giving the wrong shadow price of transmission capacity, whereas doing so in the zero-slope conditions, multiplying (3d) by three, changes nothing.

87. Albuoy, *supra* note 75. A second-order test distinguishes between a minimum and a maximum value of the objective function. This is necessary, as a zero slope is consistent with being at a minimum or at a maximum. Essentially, this requires that, when passing through an output with a zero slope, the slope itself is increasing (for a minimum) or decreasing (for a maximum). Note, too, that these tests are for local minimums and maximums, and so other points, including corner solutions, must be checked.

88. To develop the intuition for these equations, reconsider the prior unconstrained example, where $\lambda_t^* = 0$. The solution was $\{Q_1^*, Q_2^*\} = \{400, 0\}$ and $\lambda^* = \$10/\text{MWh}$. The first term of (3a) is thus zero; the first term of (3b) is $20 - 10 > 0$. Hence, the first terms of (3a) and (3b) are either zero or positive, as indicated. *Id.* at 2-3.

4. Shadow Prices Revisited

Reconsider the shadow prices in light of the constrained solution. Again, a shadow price indicates the change in the objective function (variable cost) from having one more unit of the constraining resource in question.[89]

The first shadow price is for load: $\lambda^* = \$15$; if load increases by 1MWh at bus 3, then variable cost increases by \$15. To see this, note that redispatch requires $\Delta Q_1 + \Delta Q_2 = 1$. Since the transmission constraint must still hold, the *additional* flow over the line ($\Delta Q_1/3$) must balance the *additional* counter flow over the line ($\Delta Q_2/3$). This implies $\Delta Q_1 = \Delta Q_2$. Substituting this into the required redispatch equation yields $\Delta Q_1 = \Delta Q_2 = ½$. Hence, the cost of serving another MWh of load at bus 3 is $MC_1 \times \Delta Q_1 + MC_2 \times \Delta Q_2$ = \$10/MWh x (1/2)MWh + \$20/MWh x (1/2)MWh = \$15.

The second shadow price is for transmission capacity: $\lambda_t^* = -\$15$; if transmission capacity increases by 1 MW, variable cost *decreases* by \$15. The constraint implies out of merit order dispatch compared to unconstrained dispatch, and so having greater transmission capacity implies having more efficient dispatch. Specifically, since load is fixed, the changes in generator outputs, made possible by the loosened constraint, must sum to zero: $\Delta Q_1 + \Delta Q_2 = 0$. Similarly, since the constraint binds before and after redispatch, $\Delta Q_1/3 - \Delta Q_2/3 = 1$, or $\Delta Q_1 - \Delta Q_2 = 3$.[90] Adding these two equations yields $2\Delta Q_1 = 3$, or $\Delta Q_1 = 3/2$. Since supply and demand must balance, Q_2 decreases by the same amount: $\Delta Q_2 = -3/2$ MW. Now, the increase in Q_1 changes cost by \$10/MWh x (3/2) MWh = \$15, and the decrease in Q_2 changes cost by \$20/MWh x (-3/2) MWh = -\$30. Adding the two yields a net change in cost of -\$15 when the constraint is loosened by 1 MW. Hence, λ_t^* prices congestion relief.[91]

Finally, the last two shadow prices are for generator capacity: $\lambda_1^* = \lambda_2^* = 0$. These indicate the impact on variable cost of an increase in generator capacity. Since neither generator is at capacity, having another unit of capacity means more idle capacity. With no change in dispatch, there is no impact on variable cost and, hence, zero shadow prices.

To further illustrate, reconsider the unconstrained benchmark example, with a load of 1,500 MW. Without a binding transmission constraint, generators are dispatched in merit order, starting with GEN_1. As noted, the dispatch is $\{Q_1^*, Q_2^*\} = \{1{,}000, 500\}$, at a cost of \$20,000. Now that GEN_1's capacity constraint binds, the shadow price is no longer zero. If GEN1's capacity increased by 1

89. *Id.* at 4.

90. Samir Sayah & Khaled Zehar, *Economic Lead Dispatch with Security Constraints of the Algerian Power System using Successive Linear Programming Method*, LEONARDO J. SCI., 3 (2006). Before redispatch the constraint is $Q_1/3 - Q_2/3 = 100$. After loosening the constraint and the redispatch, the constraint becomes $(Q_1+\Delta Q_1)/3 - (Q_2 +\Delta Q_2)/3 = 101$. Subtracting the first equation from the second implies $\Delta Q_1 - \Delta Q_2 = 3$.

91. *See* Kameshwar Poolla, *The Flow of Money: Electricity Market Tutorial*, U. CAL. BERKELEY, 29 (Jan. 26, 2018), https://simons.berkeley.edu/sites/default/files/docs/9080/simonsbootcamppoolla.pdf. It would not be efficient to build the transmission system such that it was never constrained (so that $\lambda_t^* = 0$). That would imply an inefficient (excessive) investment in transmission. Instead, efficiency requires that the cost of investment in transmission balance the decrease in congestion costs at the margin.

MW, for example, the new dispatch would be $\{Q_1*, Q_2*\} = \{1{,}001, 499\}$, for a variable cost of \$19,990. The \$10 decrease in variable cost implies $\lambda_1* = -\$10/MW$.

5. LMPs

The final step in the optimization is to determine the LMPs, perhaps the subtlest part of the analysis. These indicate the marginal cost of delivering one MWh (*e.g.*, one MW for one hour) to each bus.[92] The LMP formula is given by

(4) $LMP = \lambda + [S_1\Delta Q_1 + S_2\Delta Q_2]\ \lambda_t$.[93]

The subtlety arises because this formula is oriented towards a reference bus, here where load is located (bus 3).[94] Specifically, λ is the price of one MWh delivered to the reference bus.[95] Furthermore, the shares of an additional MW from the generators (S_1 and S_2) are defined as the flow from the generators over the constrained line to the reference bus.[96] If GEN1 injects a MW at bus 1 delivered to bus 3, 1/3 flows over the constrained line in the congested direction: $S_1 = 1/3$. Similarly, if GEN2 injects a MW at bus 2 delivered to bus 3, 1/3 flows over the constrained line against the congestion: $S_2 = -1/3$. The value of the shadow price of the constrained transmission resource, $\lambda_t = -15$, however, does not depend on the identity of the reference bus.

The orientation of the LMP formula towards a reference bus requires that these constraint-impact shares be used *even when examining sales to other buses*, as λ must be adjusted to determine the LMPs at other buses based on congestion costs (the second term in (4)).[97] Thus, while counter-intuitive, these flow shares apply regardless of where a MW is delivered. The ΔQ_1 and ΔQ_2, on the other hand, are determined by the least-cost dispatch to the bus under consideration. There is thus an unfortunate mix of fixed flow shares combined with flexible dispatch when considering deliveries to various buses. The LMP formula with the fixed parameter values for the example is

92. *See* MONITORING ANALYTICS, LLC, 2017 QUARTERLY STATE OF THE MARKET REPORT FOR PJM: JANUARY THROUGH SEPTEMBER, CONGESTION AND MARGINAL LOSSES 489 (2017), http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2017/2017q3-som-pjm.pdf [hereinafter 2017 MARKET REPORT].

93. *See* Fu, *supra* note 15. PJM's LMP training manual lists three components: LMP = system energy price (λ) + transmission congestion cost ($(S_1\Delta Q_1 + S_2\Delta Q_2)\ \lambda_t$) + cost of marginal losses (https://www.pjm.com/-/media/training/nerc-certifications/markets-exam-materials/mkt-optimization-wkshp/locational-marginal-pricing-components.ashx?la=en). Marginal transmission losses are ignored, as noted above. The LMP formula is derived below, as it is more intuitive to do so with downward sloping demand than with fixed demand.

94. *Id.* at 4. The choice of reference bus is examined next.

95. *Id.*

96. *See* 2017 MARKET REPORT, *supra* note 92, at 491.

97. *See* Fu, *supra* note 15, at 2.

(5) $LMP = \lambda + [(1/3)\Delta Q_1 + (-1/3)\Delta Q_2] (-15)$.

To deliver a MW to bus 3 (for one hour), the RTO cannot simply increase Q_1 (the cheapest generator) by 1MW, as $S_1 = 1/3$ will flow over the constrained line, overloading it.[98] Since the impact of each generator on the constraint is equal but in opposite directions, their outputs must be increased equally, as shown above: $\Delta Q_1 = \Delta Q_2 = ½$. Plugging these values into the LMP formula yields

$$LMP_3 = \lambda + [(1/3) \text{ x } (1/2) - (1/3) \text{ x } (1/2)] (-15) = \lambda = \$15.$$

Hence, as noted, λ equals the LMP (per MWh) *at the reference bus*.

Determining the LMP at buses 1 and 2 requires adjusting λ by the congestion costs. At bus 1, the cheapest generation is by GEN1: $\Delta Q_1 = 1$. This does not violate the transmission constraint, since the MW is consumed locally (at bus 1). Also, $\Delta Q_2 = 0$, since total additional output must sum to one. Hence,

$$LMP_1 = 15 + [(1/3) \text{ x } (1) - (1/3) \text{ x } (0)] (-15) = \$10.$$

Hence, to deliver another MW to bus 1, the RTO increases Q_1^* by 1MW at the cost of $10 per MWh, the local (marginal) generator's marginal cost.

Finally, to deliver a MW to bus 2, the RTO cannot rely on GEN1, as the constraint precludes additional output of GEN1 alone. Hence, $\Delta Q_1 = 0$ and $\Delta Q_2 = 1$. Substituting these into the formula yields

$$LMP_2 = 15 + [(1/3) \text{ x } (0) - (1/3) \text{ x } (1)] (-15) = \$20.$$

Again, this makes sense, as $20 is the marginal cost of GEN2, the local (marginal) generator at bus 2.

Summarizing, the generator outputs are $\{Q_1^*, Q_2^*\} = \{350, 50\}$ and the shadow prices are $\{\lambda^*, \lambda_t^*, \lambda_1^*, \lambda_2^*\} = \{15, -15, 0, 0\}$. Total variable cost is $VC = 10Q_1^* + 20Q_2^* = 10 \text{ x } 350 + 20 \text{ x } 50 = \$4{,}500$. Load pays $LMP_3 \text{ x } Q^* = \$15/\text{MWh x } 400 \text{ MWh} = \$6{,}000$. The flow over the transmission line is $Q_1^*/3 - Q_2^*/3 = 350/3 - 50/3 = 100$, which is at capacity. Since the RTO charges load $6,000 but pays generators $4,500, it nets $1,500 due to congestion. Without congestion, the uniform LMPs imply that the price paid to generators equals that received from load, so there is no RTO net revenue.[99] Because the RTO net revenue arises from congestion, efficiency implies that it should accrue to the own-

98. *See* 2017 MARKET REPORT, *supra* note 92, at 502.

99. *Id.* at 496. The focus on RTO net revenue here is limited to the difference between what the RTO receives from load for energy consumed and what it pays to generators for energy produced.

ers of the scarce transmission resource.[100] The transmission constraint means that each generator is paid its constant marginal cost,[101] so generator profits are zero.

Finally, comparing the unconstrained (benchmark) and constrained examples illustrates the cost of congestion. The unconstrained solution, $\{Q_1^*, Q_2^*\} = \{400, 0\}$, had a cost of $4,000; the constrained solution, $\{Q_1^*, Q_2^*\} = \{350, 50\}$, had a cost of $4,500. Hence, the cost of congestion is $500 per hour.[102] Such costs may influence transmission investments.

6. Choice of Reference Bus

The last major issue is the choice of reference bus. Implicit in the above analysis is the choice of bus 3 as the reference bus, a natural choice since load is there.[103] Other buses, however, may be chosen.[104] While the choice does not affect the LMPs, it does affect the shadow price (λ) and the shift factors (the S_i). Recall the definition of LMP:

$$(4) \quad LMP = \lambda + [S_1\Delta Q_1 + S_2\Delta Q_2]\,\lambda_t,$$

where λ is the cost (shadow price) of delivering one MWh to bus 3. Implicit in this is the choice of bus 3 as the reference bus. More generally, therefore, λ is the cost to deliver one MWh to the reference bus. The shift factors, S_1 and S_2, represent flows from GEN1 and GEN2 over the constrained line *to the reference bus*. The remaining variables, ΔQ_1, ΔQ_2, λ_t, and the LMPs, do not depend on the reference bus, but derive from the constrained, least-cost dispatch.

a. Reference Bus 1

Suppose, instead, that bus 1 is the reference bus. Hence, if GEN1, located at bus 1, produces one MWh that also sinks at reference bus 1, nothing flows over the constrained line.[105] To distinguish among reference buses, let the superscript (1, 2, or 3) indicate the reference bus, so $S_1^1 = 0$. Similarly, if GEN2, located at bus 2, produces one MWh that sinks at reference bus 1, 2/3 MWh flows

100. *See also* Seth Blumsack, *Introduction to Electricity Markets: Auction Revenue Rights*, PENN ST. COLL. EARTH AND MIN. SCI., https://www.e-education.psu.edu/ebf483/node/719 (last visited Sept. 18, 2019).

101. *See* 2017 MARKET REPORT, *supra* note 92, at 496.

102. The bill to load increases from $4,000 to $6,000 when the constraint is imposed. But $1,500 of that is a transfer from load to the RTO. Hence, only the $500 increase in variable cost represents the cost of congestion.

103. *Id.* at 489.

104. PJM, for example, uses a (dynamic) load-weighted reference bus.

105. *See* ELEC. TUTORIALS, KIRCHHOFF'S CURRENT LAW, https://www.electronics-tutorials.ws/dccircuits/kirchhoffs-current-law.html. Even though no load is modeled at bus 1, it is possible to consider the cost of delivering a MWh there. Indeed, load there would have first call on the energy produced, as energy follows the path of least resistance (Kirchhoff's Current Law), and so GEN1's output can be thought of as net of implicit load there.

over the constrained line. But since it flows against the congestion, its sign is negative: $S_2^1 = -2/3$. Equation (4), with these substitutions, becomes

(6) $\text{LMP} = \lambda^1 + [0 \text{ x } \Delta Q_1 + (-2/3) \text{ x } \Delta Q_2] (-15).$

The revised formula nevertheless yields the same set of LMPs. Consider delivering one MWh to bus 1 first (it is easiest to start with the reference bus). The least cost way of doing that, from above, is $\{\Delta Q_1, \Delta Q_2\} = \{1, 0\}$. Plugging these values in to equation (6) yields the LMP at bus 1 (known to be $10):

$\text{LMP}_1 = \lambda^1 + [0 \text{ x } 1 + (-2/3) \text{ x } 0] (-15) = \lambda^1 = 10.$

Similarly, the least cost way of delivering one MWh to bus 2 is $\{\Delta Q_1, \Delta Q_2\} = \{0, 1\}$. Plugging these values (including $\lambda^1 = 10$) into equation (6) yields the LMP at bus 2 ($20):

$\text{LMP}_2 = 10 + [0 \text{ x } 0 + (-2/3) \text{ x } 1] (-15) = 20.$

Finally, the least cost way of delivering one MWh to bus 3 is $\{\Delta Q_1, \Delta Q_2\} = \{1/2, 1/2\}$. Plugging these values into equation (6) yields the LMP at bus 3 ($15):

$\text{LMP}_3 = 10 + [0 \text{ x } (1/2) + (-2/3) \text{ x } (1/2)] \text{ x } (-15) = 15.$

b. Reference Bus 2

Finally, suppose bus 2 is the reference bus. If GEN1 produces one MWh that sinks at reference bus 2, 2/3 MWh flows over the constrained line (in the direction of congestion). Hence, $S_1^2 = 2/3$. Similarly, if GEN2 produces a MWh that sinks at reference bus 2, 0 MWh flows over the constrained line, so $S_2^2 = 0$. Equation (4) becomes

(7) $\text{LMP} = \lambda^2 + [(2/3) \text{ x } \Delta Q_1 + 0 \text{ x } \Delta Q_2] (-15).$

Again, the formula yields the same set of LMPs. The least cost way of delivering one MWh to reference bus 2 is $\{\Delta Q_1, \Delta Q_2\} = \{0, 1\}$, so the LMP at bus 2 is

$\text{LMP}_2 = \lambda^2 + [(2/3) \text{ x } 0 + 0 \text{ x } 1] (-15) = \lambda^2 = 20.$

The least cost way of delivering one MWh to bus 1 is $\{\Delta Q_1, \Delta Q_2\} = \{1, 0\}$, so the LMP at bus 1 is

$\text{LMP}_1 = 20 + [(2/3) \text{ x } 1 + 0 \text{ x } 1] (-15) = 10.$

Finally, the least cost way of delivering one MWh to bus 3 is $\{\Delta Q_1, \Delta Q_2\} = \{1/2, 1/2\}$, so the LMP at bus 3 is

$$LMP_3 = 20 + [(2/3) \times (1/2) + 0 \times (1/2)] \times (-15) = 15.$$

c. Summary of the Reference Bus Results

Figure 2 summarizes the results of the three-bus model for each choice of reference bus. The least-cost constrained dispatch necessary to deliver one MWh to a given bus, $\{\Delta Q_1, \Delta Q_2\}$, the corresponding LMPs, and λ_t, are shown in the top panel. These do not depend on the reference bus,[106] and are shown once. The choice of reference bus, however, affects the shadow price, λ^i (the LMP at the reference bus), and the shift factors, $\{S_1^i, S_2^i\}$.[107] These are shown in the bottom panels for each case. The general formula for the LMP is also repeated, tailored for the reference bus, as are the LMP calculations.

	Bus 1	Bus 2	Bus 3
$\{\Delta Q_1, \Delta Q_2\}$	**{1, 0}**	**{0, 1}**	**{1/2, 1/2}**
LMP	\$10	\$20	\$15
λ_t	-\$15	-\$15	-\$15

	Bus 1	**Bus 2**	**Bus 3**
Reference Bus: 3			
Shadow Price: $\lambda^3 = 15$			
Shift Factors: $\{S_1^3, S_2^3\} = \{1/3, -1/3\}$			
$LMP = \lambda^3 + [S_1^3 \times \Delta Q_1 + S_2^3 \times \Delta Q_2]\, \lambda_t$	$15 + [(1/3) \times (\mathbf{1}) + (-1/3) \times (\mathbf{0})](-15)$ $= \mathbf{\$10}$	$15 + [(1/3) \times (\mathbf{0}) + (-1/3) \times (\mathbf{1})](-15)$ $= \mathbf{\$20}$	$15 + [(1/3) \times (\mathbf{1/2}) + (-1/3) \times (\mathbf{1/2})](-15)$ $= \mathbf{\$15}$

	Bus 1	**Bus 2**	**Bus 3**
Reference Bus: 1			
Shadow Price: $\lambda^1 = 10$			
Shift Factors: $\{S_1^1, S_2^1\} = \{0, -2/3\}$			
$LMP = \lambda^1 + [S_1^1 \times \Delta Q_1 + S_2^1 \times \Delta Q_2]\, \lambda_t$	$10 + [(0) \times (\mathbf{1}) + (-2/3) \times (\mathbf{0})](-15)$ $= \mathbf{\$10}$	$10 + [(0) \times (\mathbf{0}) + (-2/3) \times (\mathbf{1})](-15)$ $= \mathbf{\$20}$	$10 + [(0) \times (\mathbf{1/2}) + (-2/3) \times (\mathbf{1/2})](-15)$ $= \mathbf{\$15}$

	Bus 1	**Bus 2**	**Bus 3**
Reference Bus: 2			
Shadow Price: $\lambda^2 = 20$			
Shift Factors: $\{S_1^2, S_2^2\} = \{2/3, 0\}$			
$LMP = \lambda^2 + [S_1^2 \times \Delta Q_1 + S_2^2 \times \Delta Q_2]\, \lambda_t$	$20 + [(2/3) \times (\mathbf{1}) + (0) \times (\mathbf{0})](-15)$ $= \mathbf{\$10}$	$20 + [(2/3) \times (\mathbf{0}) + (0) \times (\mathbf{1})](-15)$ $= \mathbf{\$20}$	$20 + [(2/3) \times (\mathbf{1/2}) + (0) \times (\mathbf{1/2})](-15)$ $= \mathbf{\$15}$

Figure 2. Choice of Reference Bus in the Three-Bus Model.

106. 2017 MARKET REPORT, *supra* note 92, at 489.

107. *Id.* Table notes: The shadow prices for generator capacity have been left out, as they are not relevant to the LMP calculations. They are zero for each reference case in this example.

C. Competition Under Downward Sloping Demand

The two models developed above have a fixed load (demand). A fixed load precludes analysis of market power, however, as buyers are implicitly modeled as willing to pay any amount to get the fixed quantity (modeled with a vertical demand). A generator with market power could thus charge an indefinitely high price. To examine market power, therefore, price sensitive (downward sloping) demand is necessary, as this forces a generator with market power to trade off a lower quantity for a higher price.[108] Competition continues to be assumed for this example, however, to limit the introduction of new ideas. Also, many of the details of optimization, reference buses, and shadow prices are left behind, with the emphasis shifting to the new features (downward sloping demand and market power).

With fixed (vertical) demand, optimization means cost minimization. Implicitly, however, this maximizes social surplus, defined as total willingness to pay minus production cost.[109] With demand explicitly modeled, the objective becomes maximizing social surplus.[110] To that end, let demand be given by P(Q) = 100 - Q/10, where P(Q) is the unit price and Q is the corresponding quantity demanded (load) at bus 3. Analytically, the marginal willingness to pay for *the* Q^{th} *unit* is given by P(Q). Similarly, the total willingness to pay for *Q units* is the marginal willingness to pay for each unit, summed over the units consumed. This corresponds to the area under the demand curve over the Q units, given by the integral

$$\int_0^Q P(q)dq\,.$$[111]

The Lagrangian is set up as before, except for the change in objective function from total cost to social surplus, and that the output constraint that ensures

108. Chitkara, *supra* note 38, at 2-3.

109. Shmuel S. Oren, Stephen A. Smith, & Robert B. Wilson, *Competitive Nonlinear Tariffs*, 20 J. ECON. THEORY 49, 51-52 (1983).

110. *Id.* at 52.

111. *Id.*; Fredrik Carlsson, Peter Martinsson, & Alpaslan Akay, *The Effect of Power Outages and Cheap Talk on Willingness to Pay to Reduce Outages*, INST. STUDY LABOR ECON., July 2009, at 7. The integral is the sum (suggested by the elongated "S") of the maximum price buyers are willing to pay for each of the Q units, and corresponds to the area under the demand curve, summed from zero to Q. The appendix explores integrals in more detail. While abstract, gross total willingness to pay (total value) is related to the value of lost load (VoLL), a net measure of how much load is willing to pay to avoid an outage. Note that there is a time dimension that is excluded – namely, the duration of an outage. The time dimension parallels the distinction between MW (an instantaneous measure of energy) and MWh (a cumulative measure of energy over time). The integral here is adding willingness to pay over units of energy, not time. Carlsson, Martinsson, and Akay (2011), for example, used surveys to estimate VoLL around the time of a severe storm in Sweden, where they asked people how much they would pay to avoid an outage, characterized as planned or unplanned, and of varying durations.

demand (Q) equals supply ($Q_1 + Q_2$) now allows for price sensitive demand.[112] Also, negative one multiplies λ, so that $\lambda \geq 0$, to retain its interpretation as the LMP at the reference bus.[113] The Lagrangian is thus defined as total value minus total cost, subject to the constraints:[114]

$$(8)\ \mathcal{L}(Q, Q_1, Q_2, \lambda, \lambda_t, \lambda_1, \lambda_2) = \int_0^Q P(q)dq - 10Q_1 - 20Q_2 - \lambda(Q - Q_1 - Q_2) + \lambda_t(100 - Q_1/3 + Q_2/3) + \lambda_1(1{,}000 - Q_1) + \lambda_2(1{,}000 - Q_2).$$

The zero-slope conditions are

(8a) $\partial\mathcal{L}(\bullet)/\partial Q = P(Q) - \lambda \leq 0$;[115]	$Q \geq 0$; and	$Q \times \partial\mathcal{L}(\bullet)/\partial Q = 0$;
(8b) $\partial\mathcal{L}(\bullet)/\partial Q_1 = -10 + \lambda - \lambda_t/3 - \lambda_1 \leq 0$;	$Q_1 \geq 0$; and	$Q_1 \times \partial\mathcal{L}(\bullet)/\partial Q_1 = 0$;
(8c) $\partial\mathcal{L}(\bullet)/\partial Q_2 = -20 + \lambda + \lambda_t/3 - \lambda_2 \leq 0$;	$Q_2 \geq 0$; and	$Q_2 \times \partial\mathcal{L}(\bullet)/\partial Q_2 = 0$;
(8d) $\partial\mathcal{L}(\bullet)/\partial\lambda = Q - Q_1 - Q_2 = 0$;	$\lambda \geq 0$; and	$\lambda \times \partial\mathcal{L}(\bullet)/\partial\lambda = 0$;
(8e) $\partial\mathcal{L}(\bullet)/\partial\lambda_t = 100 - Q_1/3 + Q_2/3 \geq 0$;	$\lambda_t \geq 0$; and	$\lambda_t \times \partial\mathcal{L}(\bullet)/\partial\lambda_t = 0$;
(8f) $\partial\mathcal{L}(\bullet)/\partial\lambda_1 = 1{,}000 - Q_1 \geq 0$;	$\lambda_1 \geq 0$; and	$\lambda_1 \times \partial\mathcal{L}(\bullet)/\partial\lambda_1 = 0$; and
(8g) $\partial\mathcal{L}(\bullet)/\partial\lambda_2 = 1{,}000 - Q_2 \geq 0$;	$\lambda_2 \geq 0$; and	$\lambda_2 \times \partial\mathcal{L}(\bullet)/\partial\lambda_2 = 0$.

These zero-slope conditions involve seven equations and seven unknowns ($Q, Q_1, Q_2; \lambda, \lambda_t, \lambda_1, \lambda_2$). This is more complicated to solve than equations (3a-f), as there is no subset of two equations the solution to which is $\{Q_1^*, Q_2^*\}$, from which the remaining variables can be determined.[116] Another approach is to solve the unconstrained problem, assuming both outputs are positive at the optimum (implying zero slopes), and then check the constraints for violations.

Hence, ignoring the transmission and generator constraints yields the simplified Lagrangian

$$(9)\ \mathcal{L}(\bullet) = \int_0^Q P(q)dq - 10Q_1 - 20Q_2 - \lambda(Q - Q_1 - Q_2).$$

The corresponding zero-slope conditions are

$$(9a)\ \partial\mathcal{L}(\bullet)/\partial Q = P(Q) - \lambda = 0;$$

112. Fu, *supra* note 15, at 4.

113. *Id.* Multiplying by negative one changes only the sign of λ, as $\lambda(Q - Q_1 - Q_2) = 0$ at the optimum.

114. *Id.* at 9.

115. Oren, *supra* note 109, at 52. The derivative of the integral $\int_0^Q P(q)dq$ with respect to Q is P(Q), the price. If we add a *very* small increment to output, the additional value that consumers get (the additional area under the demand curve as Q increases a bit), *measured on a full unit basis*, is price. The weak inequality implies that, if $P(Q) < \lambda$, then buyers are unwilling to pay λ (the marginal cost) even for the first unit, and so Q = 0.

116. *See generally id.* Since equations (8d) and (8e) involve three variables (Q, Q_1, Q_2), it is not possible to solve for Q_1 and Q_2 as was done for equations (3c) and (3d) above.

(9b) $\partial\mathcal{L}(\bullet)/\partial Q_1 = - 10 + \lambda = 0$;
(9c) $\partial\mathcal{L}(\bullet)/\partial Q_2 = - 20 + \lambda = 0$; and
(9d) $\partial\mathcal{L}(\bullet)/\partial\lambda = Q - Q_1 - Q_2 = 0$.[117]

These zero-slope conditions are inconsistent: (9b) yields $\lambda = 10$ and (9c) yields $\lambda = 20$, with outputs Q = 900 and 800 from (9a). But since MC_1 = \$10 is the lowest of the generators' marginal costs, GEN1 can produce either output, and a lower price implies a higher social surplus, Q = 900 is the efficient solution. Hence, the unconstrained solution is $\{Q_1^*, Q_2^*\} = \{900, 0\}$, with all LMPs at \$10/MWh.

This tentative unconstrained solution must then be checked for constraint violations. For example, does it overload the transmission line? Neither generator output constraint is violated, and so $\lambda_1 = \lambda_2 = 0$ (having more of either capacity would not change the dispatch so there is no change in the objective function, social surplus). The transmission constraint (8e) is, however, violated: $100 - Q_1/3 + Q_2/3 < 0$. Hence, $100 - Q_1/3 + Q_2/3 = 0$ holds, which requires decreasing Q1 and increasing Q2, and so $\lambda_t \neq 0$.

Returning to the original zero-slope conditions, (8a-g), the 7x7 system then simplifies to a 5x5 system, given that both outputs are positive and less than generator capacity and that the transmission constraint binds. Hence, the first terms in the zero-slope conditions again hold with equality:

(10a) $\partial\mathcal{L}(\bullet)/\partial Q = P(Q) - \lambda = 0$;
(10b) $\partial\mathcal{L}(\bullet)/\partial Q_1 = - 10 + \lambda - \lambda_t/3 = 0$;
(10c) $\partial\mathcal{L}(\bullet)/\partial Q_2 = - 20 + \lambda + \lambda_t/3 = 0$;
(10d) $\partial\mathcal{L}(\bullet)/\partial\lambda = Q - Q_1 - Q_2 = 0$; and
(10e) $\partial\mathcal{L}(\bullet)/\partial\lambda_t = 100 - Q_1/3 + Q_2/3 = 0$.

This is easily solved by adding (10b) and (10c), eliminating λ_t, so $\lambda^* = 15$. Plugging this into (10a) yields $Q^* = 850$. With total output determined, equations (10d) and (10e) form the 2x2 system

$850 = Q_1+Q_2$ and
$300 = Q_1 - Q_2$,

the solution to which is $\{Q^*, Q_1^*, Q_2^*\} = \{850, 575, 275\}$. Now all generator and transmission constraints hold.

The shadow prices are $\{\lambda, \lambda_t, \lambda_1, \lambda_2\}$. From above, λ^* = \$15, the price of another MWh at bus 3. Since the objective function has changed from total cost to social surplus, the shadow prices of transmission and generation now have different interpretations. In the previous example, cost minimization, given demand, meant that loosening the transmission constraint allowed for more efficient dispatch, lowering cost. Hence, λ_t (the value of another unit of the binding

117. Assuming that both outputs are positive implies equalities in each of the first terms.

transmission capacity) was negative. In this example, social surplus (the net benefit of the energy) is maximized. Since the cost of delivering another MWh to bus 3 remains \$15 even though the constraint is loosened by one MW, the quantity demanded (and total value) remains fixed. Loosening the transmission constraint and the more optimal dispatch decreases cost by the same \$15 as with fixed demand. Hence, social surplus increases by \$15, and so $\lambda_t^* = \$15$. Finally, since neither generator's output is constrained, as noted, the shadow prices for additional capacity are zero (increasing each generator's capacity will not affect output or social surplus). (If, alternatively, GEN1's capacity was binding, then $\lambda_1^* > 0$.) Hence, $\{\lambda^*, \lambda_t^*, \lambda_1^*, \lambda_2^*\} = \{\$15, \$15, 0, 0\}$.

The LMPs are the same as before, as the generators' constant marginal costs have not changed: the cost of delivering one MWh to bus 1 and bus 2 are \$10 and \$20. With downward sloping demand, load (Q) adjusts so that the demand price equals the \$15 cost of delivering the marginal MWh to bus 3. Hence, $\{LMP_1, LMP_2, LMP_3\} = \{\$10, \$20, \$15\}$.

Finally, social surplus (total value less cost) is given by $\int_0^Q (100 - q/10)dq$ $-10Q_1 - 20Q_2$, evaluated at the optimal output: $\{Q^*, Q_1^*, Q_2^*\} = \{850, 575, 275\}$.[118] This yields $(100 \times 850 - 850^2/20) - 10 \times 575 - 20 \times 275 = \$37{,}625$.[119] This surplus is allocated only to load and the RTO. The net gain to load (consumer surplus) is the area under the demand curve above the price, or $\int_0^{850} (100 - q/10)dq - 15 \times 850 = \$36{,}125$.[120] Since load is billed \$12,750 (= \$15/MWh x 850 MWh), while generators are paid \$11,250 (= \$10/MWh x 575 MWh + \$20/MWh x 275 MWh), the RTO nets \$1,500. Each generator's profit is zero, as the price received per MWh equals their unit costs.

Before continuing, reconsider the LMP formula (4) in light of the downward sloping demand curve. The formula is implicit in the zero-slope conditions. To make it explicit, rewrite the zero-slope conditions (10a)-(10c) by substituting back in the general designations for marginal costs ($MC_1 = 10$ and $MC_2 = 20$) and the shift factors ($S_1 = 1/3$ and $S_2 = -1/3$), yielding

(11a) $\partial\mathcal{L}(\bullet)/\partial Q = P(Q) - \lambda = 0$;
(11b) $\partial\mathcal{L}(\bullet)/\partial Q_1 = -MC_1 + \lambda + S_1\lambda_t = 0$; and
(11c) $\partial\mathcal{L}(\bullet)/\partial Q_2 = -MC_2 + \lambda + S_2\lambda_t = 0$.

The role of the reference bus (bus 3) is seen directly in equation (11a), which defines λ as $P(Q)$, the price at bus 3. Adding equations (11b) and (11c) yields $\lambda = (MC_1 + MC_2)/2 - (S_1/2 + S_2/2)\lambda_t$. Plugging in the shares, $S_1 = 1/3$ and $S_2 = -1/3$, yields $\lambda = (MC_1 + MC_2)/2$. Since λ is the price at the reference bus,

118. *Id.* at 54.

119. *Id.* The integral of the demand curve is $\int_0^Q (100 - q/10)dq = 100q - q^2/20$, which is evaluated at Q = 850. See the appendix for details.

120. *Id.* at 54.

replace λ on the left-hand side by "LMP" and replace $(MC_1 + MC_2)/2$ on the right-hand side with λ. Note that the 1/2 terms multiplying the flow shares S_1 and S_2 are the ΔQ_1 and ΔQ_2 necessary to deliver 1 MW to bus 3. Since those will depend on where the 1MW is delivered, substitute in ΔQ_1 and ΔQ_2 for the 1/2s. Hence, the general formula for LMP at any bus is $LMP = \lambda - (\Delta Q_1 S_1 + \Delta Q_2 S_2)\lambda_t$.[121]

D. Market Power via Decreasing Output

The properties of electricity create opportunities for the exercise of market power, especially when coupled with significant regulatory change (*e.g.*, the introduction of markets) or technological change (*e.g.*, the introduction of significant amount of renewables) where regulators struggle to keep up with market evolution, as supply must constantly match demand,[122] and where failure to do so can lead to load shedding or blackouts.[123] And when reserves are low, say at peak demand, the ability to exercise market power is enhanced.[124] Hence, significant resources are devoted to monitoring markets.[125] FERC has its Office of Enforcement[126], CAISO has its Department of Market Monitoring,[127] PJM has its Independent Market Monitor (Monitoring Analytics),[128] and SPP has its Market Monitor.[129] Potomac Economics, a consultant, also monitors several markets, including ERCOT, ISO-NE, MISO, and NYISO.[130]

Generators are aware that their actions can impact prices and profits, and the western energy markets during the California energy crisis provide a rich set of examples.[131] Simple examples include economic (energy priced very high) or

121. The negative sign is absent in (4), as the objective function there was cost, implying $\lambda_t < 0$. Here, the objective function is social surplus (net benefits), so $\lambda_t > 0$. Multiplying λ_t by negative one, as for λ, would then yield equation (4) exactly, but would then give the wrong sign for λ_t.

122. Storage technologies exist, such as batteries, but for now are insufficient to decouple generation from load. *See*, *e.g.*, U.S. ENERGY INFO. ADMIN, BATTERIES PERFORM MANY DIFFERENT FUNCTIONS ON THE POWER GRID (Jan. 8, 2018), https://www.eia.gov/todayinenergy/detail.php?id=34432.

123. *Id*.

124. FED. ENERGY REG. COMM'N, WHY WERE PRICES HIGH THIS SUMMER?, https://www.ferc.gov/legal/maj-ord-reg/land-docs/section5.PDF (last visited Sept. 19, 2019).

125. RAP ONLINE, MARKET POWER AND MARKET MONITORING – CRITICAL ISSUES FOR SERC AD COMPETITIVE WHOLESALE MARKETS, https://www.raponline.org/wp-content/uploads/2016/05/rap-sercissuesandwholesalemkts.pdf.

126. FED. ENERGY REG. COMM'N, ENFORCEMENT, https://ferc.gov/enforcement/enforcement.asp (last visited Oct. 26, 2019).

127. CAL. ISO, WE KEEP A CLOSE WATCH ON OUR MARKETS, http://www.caiso.com/market/Pages/Market Monitoring/Default.aspx.

128. MONITORING ANALYTICS, MONITORING ANALYTICS IS THE INDEPENDENT MARKET MONITOR FOR PJM INTERCONNECTION, http://www.monitoringanalytics.com/company/role.shtml.

129. S.W. POWER POOL, MARKET MONITORING, https://www.spp.org/markets-operations/market-monitoring/.

130. POTOMAC ECON., LEADING INDEPENDENT MARKET MONITORING, ANALYSIS, & LITIGATION SUPPORT SERVICES, https://www.potomaceconomics.com.

131. *See generally* FERC Docket No. PA02-2-000, *Final Report on Price Manipulation in Western Markets Fact-Finding Investigation Of Potential Manipulation Of Electric And Natural Gas Prices* (Mar. 2003).

physical (a generator taken offline) withholding or hockey-stick offers (energy offered at low prices for most units, but at high prices for the last units).[132] Egregious examples of the exercise of market power involved the trading at Enron (among others).[133] It employed various techniques to exploit weaknesses in the young market and its regulation.[134]

In a stylistic "Death Star," Enron would schedule energy over a "circular" path comprised of multiple segments.[135] One segment would have energy flowing against congestion in the CAISO to free transmission capacity, just as in Figure 1 above, where additional flow from GEN2 from bus 2 to bus 1 would alleviate congestion on the 1-2 line. The second segment would have energy loop back to the origin outside of CAISO (and so unseen by CAISO). In terms of Figure 1, imagine a neighboring grid (not shown) that was connected to buses 1 and 2, but had a third bus (4) that was the mirror image of bus 3 below line 1-2. The second flow would be 1 to 4 to 2. The net effect of the circular schedule is that no energy flowed over the grid, as it is injected and withdrawn at bus 2. But CAISO nevertheless made payments for congestion relief on line 1-2.[136] Another circular schedule, "Ricochet," involved Enron buying energy in CAISO's day-ahead market, selling it outside of the state to a second party, and buying the energy back for sale in California thereby evading price caps on "local" energy.[137] In terms of Figure 1, the export might be from bus 2 to bus 4 (not shown) and the import from bus 4 to bus 2. The net effect is that energy produced by GEN2 could be sold locally for prices in excess of the cap.

The next two examples illustrate how such market power can be used to limit the output of the competing generator by exploiting the transmission constraint. Limitations of the transmission grid can enhance generator market power and can, in some instances, make detection challenging. In this example, strategic use of the constraint (decreasing output, and so counter-flow) limits the competing generator's transmission access by decreasing effective capacity. In the next example, with an alternative configuration, strategic use of the transmission constraint (increasing output, and so competing flow) limits the competing generator's transmission access by displacement. Detecting market power is thus challenging, as it can be consistent with decreasing or increasing output. For both examples, GEN2 has market power, whereas GEN1 continues to competitively offer energy at marginal cost.[138]

Specifically, Chapter VI provides details on manipulation strategies, including links with the natural gas industry. *Id.* at VI-1-59.

132. *Id.* at VI-47.

133. *Id.* at VI-1, VI-3.

134. *Id.* at VI-3, VI-6, VI-12, VI-17-17, and VI-20.

135. FERC Docket No. PA02-2-000 at VI-26.

136. *Id.* at VI-30.

137. *Id.* at VI-17.

138. Whether GEN1 or GEN2 is assumed to have market power makes little difference. To assume GEN1 has market power would result in even higher prices, though, as it is the cheapest generator. This would

In this example, the transmission constraint allows GEN2 to limit GEN1's output by restricting its own. Since GEN1's flow dominates (being the cheapest generator), the line is constrained at 100 MW from bus 1 to bus 2, with 1/3 of each MW flowing over the line. GEN2's flow is in the opposite direction, in effect adding capacity, also with 1/3 of each MW flowing over the line. The binding constraint, $100 = Q_1/3 - Q_2/3$, implies $Q_1 = 300 + Q_2$. When GEN2 exercises market power by decreasing Q_2, the constraint forces GEN1 to decrease its output one-for-one, amplifying GEN2's market power.

GEN2's profit depends on LMP_3, which increases as GEN2 decreases its output, given the downward sloping demand. The relationship among LMPs stems from the constraint: as noted, an additional MW delivered to bus 3 requires 1/2 MW from each of GEN1 and GEN2, at a delivered cost of the average of their offer prices. Hence, $LMP_3 = (LMP_1 + LMP_2)/2$. Since GEN1 behaves competitively, $LMP_1 = \$10$/MWh, and so $LMP_2 = 2 \text{ x } LMP_3 - 10$.

GEN2's profit is revenue minus cost: $\pi(Q_2) = LMP_2 \text{ x } Q_2 - 20Q_2$. Substituting for LMP_2 yields $\pi(Q_2) = (2 \text{ x } LMP_3 - 10) \text{ x } Q_2 - 20Q_2$. The demand curve links LMP_3 to Q_2: $P(Q) = 100 - Q/10 = 100 - (Q_1 + Q_2)/10$. Lastly, the transmission constraint determines GEN1's output based on GEN2's output: $Q_1 = 300 + Q_2$. Making these two substitutions (sequentially, in boldface) into the profit function yields:

$$
\begin{aligned}
\pi(Q_2) &= \{2 \text{ x } [\mathbf{100 - (Q_1 + Q_2)/10}] - 10\} \text{ x } Q_2 - 20Q_2 \\
&= \{2 \text{ x } [100 - (\mathbf{300 + Q_2} + Q_2)/10] - 10\} \text{ x } Q_2 - 20Q_2 \\
&= 110Q_2 - (2/5)Q_2^2.
\end{aligned}
$$

Again, optimization (profit maximization) is akin to getting to the top of the profit hill, and is found by examining profit at different rates of output. For example, at zero output, profit is zero. Then the generator would consider one unit of output – does profit increase? If yes, another unit is considered in the same way. At some point adding another unit to output will decrease profit, implying that the slope is zero (at the top of the hill).

Mathematically, this involves setting the derivative (slope) equal to zero and solving for Q_2: $d\pi(Q_2)/dQ_2 = 110 - 2(2/5)Q_2 = 0$, and so the profit maximizing quantity is $Q_2^* = 137.5$.[139] The transmission constraint implies that $Q_1^* = 300 + Q_2^* = 437.5$. Total output then determines $LMP_3 = P(Q^*) = 100 - (437.5 + 137.5)/10 = \$42.5$/MWh. Finally, the LMP relationship implies $LMP_2 = 2 \text{ x } LMP_3 - 10 = 2 \text{ x } 42.5 - 10 = \$75$. GEN2 thus maximizes profit by offering en-

also reverse the net flow across the constrained line, as GEN2 would become the cheapest generator, and so dispatched to the extent possible.

139. Paul Joskow, *Transmission rights and market power on electric power networks,* 31 RAND J. ECON. 3 (2000). As noted, the RTO is also optimizing in the background to maximize social surplus (net benefits), given the generator offers (inclusive of market power, itself the object of optimization). Hence, optimization may benefit or harm customers, depending on who is doing the optimization.

ergy at \$75/MWh. The RTO, facing offer prices of \$10 and \$75 per MWh, then dispatches the corresponding quantities.

Now consider the welfare effects of market power. Output fell from the competitive quantities, $\{Q^*, Q_1^*, Q_2^*\} = \{850, 575, 275\}$, to $\{575, 437.5, 137.5\}$. Globally, the social surplus is given by $\int_0^Q (100 - q/10)dq - 10Q_1 - 20Q_2$, evaluated at $\{Q^*, Q_1^*, Q_2^*\} = \{575, 437.5, 137.5\}$. This yields $(100 \times 575 - 575^2/20) - 10 \times 437.5 - 20 \times 137.5 =$ \$33,843.75, \$3,781.25 below the competitive amount. The lost \$3,781.25 in social surplus is a deadweight loss (a loss not offset by gains elsewhere in the market) from the exercise of market power. Load suffers a considerable drop in consumer surplus, from \$36,125.00 to \$16,531.25. GEN1's profit remains zero, as its unit cost and price are the same (\$10/MWh). Two parties gain. GEN2 gains from its market power; its profit is $LMP_2 \times Q_2^* - 20 \times Q_2^* = 75 \times 137.5 - 20 \times 137.5 =$ \$7,562.5 (versus zero under competition). The RTO also gains; its net increases from \$1,500 to \$9,750. Their gains come at the expense of load.

E. Market Power via Increasing Output

The last example illustrates another way in which transmission congestion interacts with market power. In the previous example, GEN2's output flows against congestion, so the generators did not compete for transmission capacity. But when GEN2 decreased its output to exercise market power, the decreased counter flow forced a decrease in GEN1's output. If the grid is reconfigured so that GEN2's output flows with congestion, then the generators compete for transmission capacity, akin to a zero-sum game. GEN2 then exercises market power by increasing output, forcing a decrease in GEN1's output by using transmission capacity otherwise used by GEN1.

To illustrate, the grid is reconfigured in several ways. First, the transmission constraint is moved to the 2-3 line. Since GEN1 and GEN2 are now on the same side of the constraint relative to load, 1/3 of Q_1 and 2/3 of Q_2 flow over the 2-3 line in the same direction, forcing them to compete for transmission capacity. Since the flows are in the same direction, transmission capacity is also increased to 300 MW. Hence, the constraint becomes $(1/3)Q_1 + (2/3)Q_2 \leq 300$. Second, GEN1 remains competitive, but with increasing marginal cost: $MC_1 = 2 + Q_1/200$. This change offers an additional insight into hockey-stick offers (the initial units offered at low prices and the last units offered at high prices). Third, GEN2 continues to have market power, but now also with increasing marginal cost: $MC_2 = 4 + Q_2/100$. To expand GEN2's output to limit GEN1's access to transmission capacity, GEN2 must decrease its offer price to undercut GEN1, incurring a loss. To make the output expansion profitable, GEN2 has an affiliated, relatively expensive, generator at bus 3, with increasing marginal cost of $MC_3 = 6 + Q_3/100$. All generators have capacity constraints of 1,000 MW. The reconfigured grid is illustrated in Figure 3, with prices and quantities based on the constrained market power example.

The optimization problem is solved as before, in steps, starting with the unconstrained competitive case, then adding the binding transmission constraint, and finally adding market power. Because the grid configuration has changed, however, this example is not directly comparable to the previous examples in terms of output, prices, and surpluses.

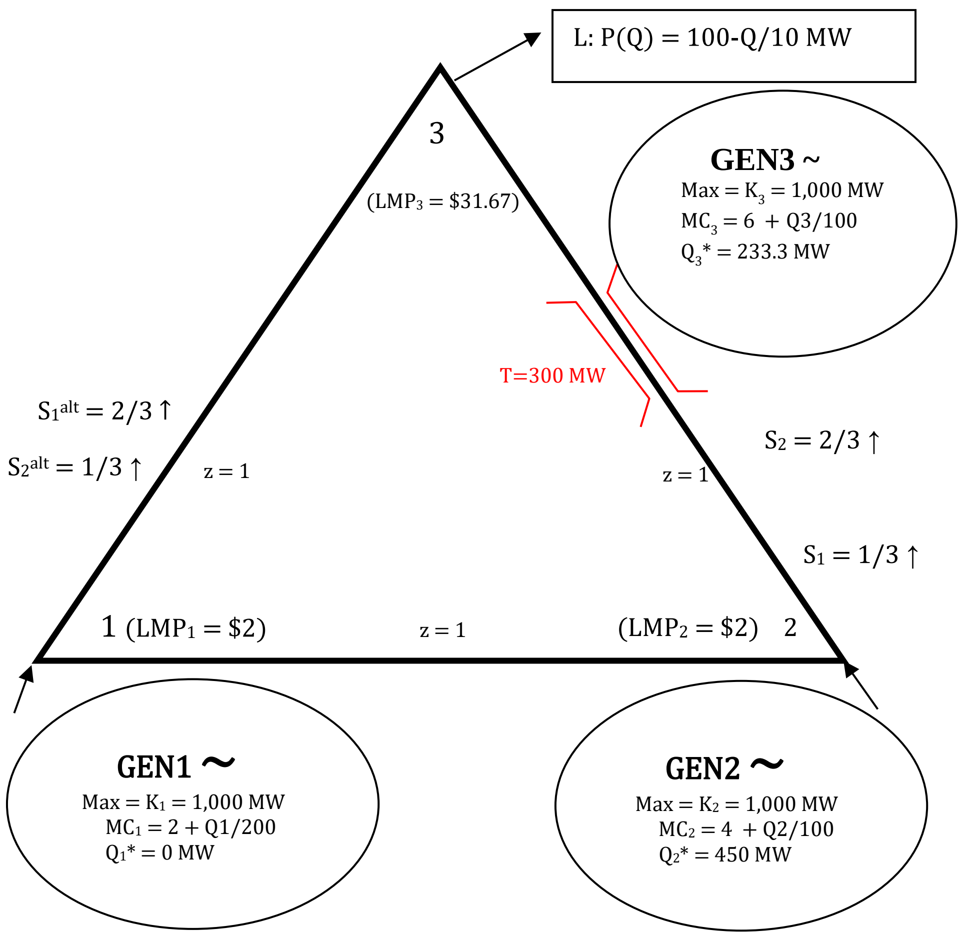

Figure 3. Market Power in a 3-Bus, 1-Load Example.

The competitive, unconstrained case involves maximizing social surplus as in the previous example. The Lagrangian is given by

$$(12)\quad \mathcal{L}(\bullet) = \int_0^{Q*} P(q)dq - \int_0^{Q1} MC_1(q)dq - \int_0^{Q2} MC_2(q)dq - \int_0^{Q3} MC_3(q)dq - \lambda(Q - Q_1 - Q_2 - Q_3).$$[140]

The zero-slope conditions are

(12a) $\partial\mathcal{L}(\bullet)/\partial Q = P(Q) - \lambda \leq 0$; $Q \geq 0$; and $Q \text{ x } \partial\mathcal{L}(\bullet)/\partial Q = 0$;
(12a) $\partial\mathcal{L}(\bullet)/\partial Q_1 = \lambda - MC_1(Q_1) \leq 0$; $Q_1 \geq 0$; and $Q_1 \text{ x } \partial\mathcal{L}(\bullet)/\partial Q_1 = 0$;
(12b) $\partial\mathcal{L}(\bullet)/\partial Q_2 = \lambda - MC_2(Q_2) \leq 0$; $Q_2 \geq 0$; and $Q_2 \text{ x } \partial\mathcal{L}(\bullet)/\partial Q_2 = 0$;
(12c) $\partial\mathcal{L}(\bullet)/\partial Q_3 = \lambda - MC_3(Q_3) \leq 0$; $Q_3 \geq 0$; and $Q_3 \text{ x } \partial\mathcal{L}(\bullet)/\partial Q_3 = 0$; and
(12d) $\partial\mathcal{L}(\bullet)/\partial\lambda = Q - Q_1 - Q_2 - Q_3 = 0$; $\lambda \geq 0$; and $\lambda \text{ x } \partial\mathcal{L}(\bullet)/\partial\lambda = 0$.

The tentative unconstrained solution is $\{Q^*, Q_1^*, Q_2^*, Q_3^*\} = \{941.5, 770.7, 185.4, -14.6\}$. Since output must be non-negative, however, Q_3^* is set to zero. The revised solution, found by removing GEN3 from the problem, is $\{Q^*, Q_1^*, Q_2^*, Q_3^*\} = \{941.9, 761.3, 180.6, 0\}$, with the generator constraints holding. Without a binding transmission constraint, LMPs are uniform as before, at \$5.81, even though the generators now have different marginal cost curves. This is because the generators are dispatched to equate marginal costs (for each generator dispatched) to the uniform LMP. (Equations (12a-c) set the LMP at the reference bus (λ) equal to each generator's marginal cost.) The shadow prices are $\{\lambda^*, \lambda_1^*, \lambda_2^*, \lambda_3^*\} = \{\$5.81, \$0, \$0, \$0\}$, which (except for λ) equal zero, as the generator constraints are nonbinding.

Consider now the welfare effects under the new configuration with competition. Globally, the social surplus is

$$\int_0^{941.9}\left(100 - \frac{q}{10}\right)dq - \int_0^{761.3}\left(2 + \frac{q}{200}\right)dq - \int_0^{180.6}\left(4 + \frac{q}{100}\right)dq = \$45{,}974.19.$$

This surplus is allocated largely to load. The consumer surplus is \$44,362.12, and the generator profits are \$1,448.91; \$163.16; and \$0. The RTO net revenue is \$0, as the transmission constraint does not bind.[141]

Note that, even though GEN1 behaves competitively, offering energy at marginal cost, its profit increases from zero (in the previous example with a constant marginal cost) to \$1,448.91. This results from having increasing marginal

140. Just as the integral involving the demand curve (the first term on the right hand side) sums the price that load is willing to pay for each unit consumed (the area under the demand curve), the integrals involving the marginal cost curves sum the marginal cost for each unit produced for each generator (the area under the marginal cost curve).

141. Again, without a binding transmission constraint, all generators are paid the common LMP (set by the marginal generator, though all are marginal with rising marginal cost), and load pays the same common LMP for each unit. Rents accrue to inframarginal generators (when marginal cost is constant) or to inframarginal units for all generators (when marginal cost is increasing).

cost. While profit on the last unit is approximately zero, as LMP equals marginal cost, it earns LMP on the inframarginal units, which have lower marginal cost. This illustrates another way in which hockey-stick offers can be profitable.[142]

Checking the transmission constraint indicates a violation: $(1/3)Q_1^* + (2/3)Q_2^* = (1/3) \times 761.3 + (2/3) \times 180.6 = 374.2 > 300$. Imposing the constraint may result in GEN3 being dispatched, so it is added back in the model. The transmission constraint (boldface) is thus added to the Lagrangian:

$$(13)\quad \mathcal{L}(\bullet) = \int_0^{Q^*} P(q)dq - \int_0^{Q1} MC_1(q)dq - \int_0^{Q2} MC_2(q)dq - \int_0^{Q3} MC_3(q)dq - \lambda(Q - Q_1 - Q_2 - Q_3) + \boldsymbol{\lambda_t(300 - Q_1/3 - 2Q_2/3)}.$$

The zero-slope conditions are

(13a) $\partial\mathcal{L}(\bullet)/\partial Q = P(Q) - \lambda \le 0$; $Q \ge 0$; and $Q \times \partial\mathcal{L}(\bullet)/\partial Q = 0$;
(13b) $\partial\mathcal{L}(\bullet)/\partial Q_1 = \lambda - MC_1(Q_1) - \lambda_t/3 \le 0$; $Q_1 \ge 0$; and $Q_1 \times \partial\mathcal{L}(\bullet)/\partial Q_1 = 0$;
(13c) $\partial\mathcal{L}(\bullet)/\partial Q_2 = \lambda - MC_2(Q_2) - 2\lambda_t/3 \le 0$; $Q_2 \ge 0$; and $Q_2 \times \partial\mathcal{L}(\bullet)/\partial Q_2 = 0$;
(13d) $\partial\mathcal{L}(\bullet)/\partial Q_3 = \lambda - MC_3(Q_3) \le 0$; $Q_3 \ge 0$; and $Q_3 \times \partial\mathcal{L}(\bullet)/\partial Q_3 = 0$;
(13e) $\partial\mathcal{L}(\bullet)/\partial\lambda = Q - Q_1 - Q_2 - Q_3 = 0$; $\lambda \ge 0$; and $\lambda \times \partial\mathcal{L}(\bullet)/\partial\lambda = 0$; and
(13f) $\partial\mathcal{L}(\bullet)/\partial\lambda_t = 300 - Q_1/3 - 2Q_2/3 \ge 0$; $\lambda_t \le 0$; and $\lambda_t \times \partial\mathcal{L}(\bullet)/\partial\lambda_t = 0$.

The constrained, competitive solution is $\{Q^*, Q_1^*, Q_2^*, Q_3^*\} = \{930.2, 765.1, 67.4, 97.7\}$, so the generator constraints hold. The LMPs are $\{LMP_1, LMP_2, LMP_3\}$ = {\$5.83, \$4.67, \$6.98}. The shadow prices are $\{\lambda^*, \lambda_t^*, \lambda_1^*, \lambda_2^*, \lambda_3^*\}$ = {\$6.98, \$3.45, 0, 0, 0}. The social surplus is \$45,836.63, again mostly allocated to load. Consumer surplus is \$43,266.63, and the generator profits are \$1,466.89; \$22.44; and \$47.70. Now that the transmission constraint binds, the RTO net revenue is \$1,036.05.

Given this constrained competitive benchmark, consider now the exercise of market power. GEN2 exercises market power by increasing its output to preclude GEN1's access to transmission. It ensures dispatch (use of transmission) by offering energy below its marginal cost and, in particular, below GEN1's marginal cost. Selling energy below marginal cost would not be profitable, and so to allow GEN2 to exercise market power, the affiliated generator at bus 3 is needed to profit from the higher price there.

The optimization is as follows. Since GEN1 is the most efficient generator (lower and flatter marginal cost), it has the largest output of the three generators under competitive dispatch. Because GEN1 and GEN2 compete for transmission

142. This relates to a point raised by LMP opponents noted in the introduction, who argued that inframarginal base-load units were over compensated. Because their costs are low, they have an incentive to offer energy at marginal cost, as that likely ensures dispatch, while the marginal generator offers energy at a higher price, from which they benefit. Here, all generators (with positive output) benefit on their inframarginal units.

capacity, GEN2 exercises market power by expanding output to limit GEN1's output.[143] GEN2 excludes GEN1 from the market by producing at the limit defined by the transmission constraint, where two thirds of its output flows over the line: $(2/3)Q_2^* = 300$, or $Q_2^* = 450$. If it offered 450 MWh at \$2, it would undercut GEN1 and be dispatched to 450 MW, binding the constraint.[144] Hence, LMP_2 = \$2/MWh. Since GEN1 has not been dispatched, LMP_1 = \$2/MWh as well (GEN1's marginal cost at $Q_1 = 0$). Once GEN1 has been precluded from the market, and GEN2 has maximized its output, subject to the transmission constraint, the problem simplifies to setting the affiliate's output (Q_3) to maximize GEN3's profit.

GEN3's profit is $\pi(Q_3) = P_3 \text{ x } Q_3 - \int_0^{Q3} MC_3(q)dq$, where P_3 depends on total output: $P_3(Q) = 100 - Q/10 = 100 - (Q_1 + Q_2 + Q_3)/10$. Since $Q_1^* = 0$ and $Q_2^* = 450$, $P_3(Q_3) = 100 - (0 + 450 + Q_3)/10 = 55 - Q_3/10$. GEN3's profit function thus reduces to

$$\pi(Q_3) = (55 - Q_3/10) \text{ x } Q_3 - \int_0^{Q3}(6 + \frac{q}{100})dq.$$

The zero-slope condition is $d\pi(Q_3)/dQ_3 = 55 - 2Q_3/10 - (6 + Q_3/100) = 0$, implying $Q_3^* = 233.3$. The constrained, market power solution is thus $\{Q^*, Q_1^*, Q_2^*, Q_3^*\} = \{683.3, 0, 450, 233.3\}$, the LMPs are $\{LMP_1, LMP_2, LMP_3\}$ = {\$2, \$2, \$31.67}, and shadow prices are $\{\lambda^*, \lambda_t^*, \lambda_1^*, \lambda_2^*, \lambda_3^*\}$ = {\$31.67, \$50.51, 0, 0, 0}.[145]

Now consider the welfare effects of granting GEN2 market power. Social surplus falls from \$45,836.63 to \$27,149.11. Load suffers the most: consumer surplus falls from \$43,266.63 to \$23,341.50. GEN1 also suffers, as it is excluded from the market; its profit falls from \$1,463.51 to \$0. Two parties gain from the exercise of market power. GEN2 (obviously) gains from having market power. GEN2's profit is thus $\pi_2(450) = 2 \text{ x } 450 - (4\text{x}450+450^2/(2\text{x}100))$ = -\$1,912.50; the profit of its affiliate GEN3 is $\pi_3(233.3) = (55 - 233.3/10)\text{x}233.3$ –

143. *Id.* GEN2 cannot exercise market power by decreasing output, as in the prior example, as both generators now flow over the constrained line in the same direction. Since GEN1 is cheaper, it would be dispatched until the transmission constraint binds, with total output remaining constant. An exception would be if GEN1's output became constrained. It might then be possible (depending on how limited GEN1's output is) to exercise market power by increasing price, as in the previous example.

144. Leslie Liu & Assef Zobian, *The Importance of Marginal Loss Pricing in an RTO Environment,* 15 ELEC. J. 8 (2002). GEN2 cannot offer, for example, 449 MWh at \$2 and additional MWh at marginal cost (or higher price), to eliminate losses by pushing up LMP_2 at the margin. At that price, it would be dispatched to 449, and GEN1 would be dispatched to 3 MWh (1/3 of GEN1's output flows over the line), which then binds the transmission line. The problem for GEN2 is that any attempt to push LMP_2 above \$2/MWh results in GEN1's dispatch up to the point of binding the transmission line. This is also true if there is load at bus 2. If GEN2 offered units in excess of 450 at, for example, \$10, hoping that LMP2 would be \$10, it would again be undercut by GEN1. Sales at bus 2 would not violate the constraint (indeed, would increase effective capacity) as 1/3 of GEN1's sales to bus 2 flows against the congestion (bus 1 to bus 3 to bus 2), so GEN1 can serve load at bus 2. Hence, \$2/MWh is the highest offer price possible which still precludes dispatch of GEN1.

145. Fu, *supra* note 15. Once Q_3 is determined, LMP_3 is $55 - Q_3^*/10 = 55 - 23.33 = 31.67$.

$(6x233.3 + 233.3^2/(2x100)) = \$5,716.67$. The combined profit of GEN2 and GEN3 is thus $3,804.17. The corresponding profit under competition was $70.44. The RTO (less obviously) also gains when GEN2 exercises market power: the RTO net increases from $1,036.05 under competition to $13,351.50 under market power.

A final insight emerges regarding the shadow price of transmission, λ_t*, in the presence of market power. Social surplus depends on the marginal cost curves rather than on the offer curves tainted with market power,[146] even though GEN2 offers energy below its marginal cost (the loss for GEN2 is a gain for load). If the model is solved assuming "competition," but with prices determined by the exercise of market power (*i.e.*, GEN1 offering energy at $MC_1 = 2+Q_1/200$ at bus 1 and GEN2 offering energy at $2/MWh at bus 2 and at $31.67/MWh at bus 3), then $\lambda_t^* = \$50.51$ as indicated. That is, suppose that the RTO invested in transmission to loosen the constraint by 1 MW, to 301 MWs. Assuming that generators did not adjust their offers in response, the RTO will then be able to use more output from the low-cost generator and less from the high-cost generator. The redispatch increases social surplus by $50.51.

The exercise of market power, however, prevents such redispatch: the additional output of GEN2 necessary to exclude GEN1 by using the additional available transmission capacity would be exactly offset by decreased output of GEN3, keeping the total output fixed. There would be a small decrease in social surplus, as the marginal cost of GEN2 is $0.17/MWh higher than that of GEN3. Essentially, expanding the transmission line to lessen the constraint is offset by the redispatch of GEN2 to again congest the line and of GEN3 to keep total output fixed. Whether social surplus increases or decreases depends on whether the marginal cost of GEN2 (whose output would increase) exceeds that of GEN3 (whose output would decrease). But λ_t* no longer measures the gains (at the margin) from loosening the transmission constraint by one MW in the presence of market power.

IV. EXTENSIONS

Once the basic LMP model is understood, many extensions are relatively straightforward. In general, extensions involve the design of the transmission grid, the economics of competition, and policy considerations.[147] While a survey of extensions is beyond the scope of this article, some are outlined here.

Technical extensions involve changing the model topology to more accurately reflect the grid.[148] Simple transmission extensions involve adding buses, lines, and constraints.[149] Similarly, generators and loads can be added at various

146. *Id.*

147. Joskow, *supra* note 139, at 10-13.

148. *Id.* at 46-47.

149. Fu, *supra* note 15.

buses.[150] This greatly increases the number of combinations to consider, however, and so greatly increases computational demands.[151] Also, additional transmission constraints imply that additional generators become marginal.[152] For example, without congestion, generators are dispatched in merit order, resulting in a single marginal generator that responds to load changes (Example III.A).[153] Adding a binding constraint results in a second generator being marginal – two are required to serve additional load while respecting the constraint (Example III.B).[154] Each additional binding constraint results in an additional marginal generator.[155]

Some extensions quickly leave the analysis presented here behind, however, as the electrical engineering becomes more realistic. For example, by replacing the simplified DC approximation with an AC model.[156] In addition, the modeling gets more abstract, for example, by using vector calculus.[157]

Economic extensions involve changing the form of competition.[158] And with more complexity, more complex forms of market power may emerge, such as GEN1 and GEN2 each exercising market power (duopoly). For example, Cardell *et al.* (1997) explore different forms of competition among generators, including duopoly.[159] They also explore how transmission rights impact incentives to exercise market power, as tradable transmission rights represent another revenue stream over which to optimize.[160] Coordination can also be extended across other products, such as energy, reserves, capacity, and reactive power, and related industries, such as natural gas.

Another extension allows generators to compete over time. The above analysis focused on a single period. When time is modeled, how the market participants learn and react to one another must also be modeled, as that also plays a

150. *Id.*

151. *Id.*

152. *See generally id.*

153. Joskow, *supra* note 139, at 14.

154. MONITORING ANALYTICS, LLC, 2016 QUARTERLY STATE OF THE MARKET REPORT FOR PJM: JANUARY THROUGH MARCH, CONGESTION AND MARGINAL LOSSES (2016), https://www.monitoringanalytics.com/reports/ PJM_State_of_the_Market/2016/2016-som-pjm-sec11.pdf.

155. *See, e.g.* PJM INTERCONNECTION, L.L.C., LOCATIONAL MARGINAL PRICING COMPONENTS (2017), http://pjm.com/~/media/training/nerc-certifications/markets-exam-materials/mkt-optimization-wkshp/locational-marginal-pricing-components.ashx.

156. *See, e.g.*, Bautista, G. et al., *Modeling Market Power in Electricity Markets: Is the Devil Only in the Details?*, 20 ELEC. J. 82-92. DC models have linear equations, resulting from several simplifying assumptions, such as ignoring reactive power, and are relatively simple to solve. AC models, in contrast, have nonlinear equations involving sines and cosines, consider active and reactive power, and are much more difficult to solve.

157. *See, e.g.*, Tina Orfanogianni & George Gross, *A General Formulation for LMP Evaluation*, 22 IEEE TRANSACTIONS ON POWER SYS. 1163 (Aug. 2007); Richard M. Benjamin, *An Electricity Primer for Energy Economists: Basic EE to LMP Calculation*, SSRN (Feb. 13, 2013), https://papers.ssrn.com/ sol3/papers.cfm?abstract_id=2216338.

158. *See generally* Cardell et al., *supra* note 14.

159. *Id.*

160. *Id.*

role in establishing equilibrium and in exercising market power.[161] Market power could also be introduced on the buyer side. Demand response is a way to introduce price sensitive demand to limit seller market power.[162] Similarly, large buyers or buyer groups limit seller market power.[163]

Finally, policy considerations are an important part of LMP modeling exercises. This article focuses on market power, but several extensions naturally arise. For example, Berry *et al.* (1999) find that the effects of localized competition (adding generation at a given bus) depends on how electrically close the bus is to binding transmission constraints, with effects more pronounced closer to the constraints.[164] Hence, they note that measures of market concentration, such as the HHI, that ignore location or grid topology, might be an unreliable guide to market power.[165]

Reliability, a central concern, can be modeled by adding generator and transmission derate or outage probabilities and solving on an expected value basis.[166] A related reliability concern arises from the integration of large amounts of non-dispatchable solar and wind capacity.[167] As the share of non-dispatchable resources grows, the burden on the dispatchable resources needed to balance load and respond in emergencies increases.[168] Growing solar and wind capacity also raise potential problems in energy-only markets, such as ERCOT, as they have no fuel cost, pressuring the energy margins needed to recover capacity costs.[169]

V. CONCLUSION

This article explained the basics of LMP in a three-bus model through a series of examples. Two examples, the Unconstrained Benchmark[170] and Competition Under Fixed Demand,[171] illustrated unconstrained and constrained dispatch

161. *See, e.g.,* Andrew Lu Liu, Repeated-Game Models of Competitive Electricity Markets: Formulations and Algorithms, (July, 2019) (unpublished Ph.D. dissertation, Johns Hopkins University); Michael H. Rothkopf, *Daily Repetition: A Neglected Factor in the Analysis of Electricity Auctions,* ELEC. J. 60 (Apr. 1999).

162. Cherelle Eid et al., *Time-Based Pricing and Electricity Demand Response: Existing Barriers and Next Steps*, UTIL. POL'Y 15, 17 (Apr. 20, 2016).

163. *Id.*

164. Carolyn A. Berry et al., *Understanding How Much Market Power Can Arise in Network Competition: A Game Theoretic Approach*, UTIL. POL'Y 139 (Sept. 1999).

165. *Id.* Conversely, standard concentration analysis could be used for regions without transmission constraints, as the grid effectively collapses to a single bus. The Herfindahl-Hirschman Index (HHI), a common measure of concentration, is defined as the sum of the squared market shares (as percentages). For example, if a market has three firms, with shares of 20%, 20%, and 60%, the HHI is $20^2 + 20^2 + 60^2 = 4{,}400$.

166. Michael Hogan, *Follow the missing money: Ensuring reliability at least cost to consumers in the transition to a low-carbon power system,* 30 ELEC. J. 55 (2017).

167. *Id.*

168. *See, e.g., Id.*; Thure Trabera & Claudia Kemfert, *Gone with the wind? - Electricity market prices and incentives to invest in thermal power plants under increasing wind energy supply*, 33 ENERGY ECON. 249 (Mar. 2011).

169. *See, e.g.,* William W. Hogan, *On An "Energy Only" Electricity Market Design for Resource Adequacy* (Sept. 23, 2005), https://sites.hks.harvard.edu/fs/whogan/Hogan_Energy_Only_092305.pdf.

170. *See* discussion *supra* Part III.A.

171. *See* discussion *supra* Part III.B.

to explore how transmission congestion increases production cost. Two additional examples, Market Power via Decreasing Output[172] and Market Power via Increasing Output,[173] explored how transmission constraints can affect market power. While the models are simple, they yield a rich set of insights into the operation of wholesale electricity markets. These insights include congestion costs, constrained dispatch, deadweight loss, demand response, hockey-stick offers, inframarginal rents, LMP determination, market power, reference buses, shadow prices, transmission expansion, and transmission rents.

There are, as noted, many extensions of the model left unexplored. But once the basics are understood, extensions in the economics literature become more approachable. Even when the mathematics used is more advanced than that presented here, important insights are still possible with an understanding of the models presented here.

VI. APPENDIX

This appendix presents a short calculus primer, stating results without proof. Very generally, calculus may be broken into two related parts: differential calculus (concerning slope) and integral calculus (concerning area and volume).[174] After presenting basic results, several examples presented above are discussed.

Differential calculus focuses on slope, the heart of optimization: maximizing profit or minimizing cost.[175] This means locating the top or bottom of the function, where the slope is zero (for well-behaved functions).[176] The slope of a function is given by its derivative.[177] Calculating derivatives is fairly mechanical, and for simple functions, quite easy.

To motivate the derivative, consider the linear function $y = a + bx$, where a and b are constants. If x changes by Δx, y will change by $b\Delta x$: $\Delta y = b\Delta x$. Dividing by Δx yields the average slope over the interval Δx which, for a linear function, is constant: $\Delta y/\Delta x = b$. Letting Δx get vanishingly small, replacing Δ with d, yields the derivative of y with respect to x: $dy/dx = b$. The rules for finding derivatives of simple functions are also simple: the exponent on the independent variable is brought out front to multiply the function and then the exponent is decreased by one. For example, if $y = bx^1$, then $dy/dx = 1(bx^{1-1}) = b$, as just shown. Similarly, if $y = bx^2$, then $dy/dx = 2(bx^{2-1}) = 2bx$. The general rule is that, if $y = bx^n$, then $dy/dx = n(bx^{n-1}) = nbx^{n-1}$.

172. *See* discussion *supra* Part III.D.

173. *See* discussion *supra* Part III.E.

174. ENCYCLOPAEDIA BRITTANNICA, ANALYSIS CALCULUS, https://www.britannica.com/science/analysis-mathematics/Calculus#ref731795.

175. Jeff Cruzan, *The Derivative*, http://xaktly.com/TheDerivative.html.

176. *Id.*

177. *Id.*

This rule extends to polynomials: the derivative of a sum is the sum of the derivatives (of the individual terms).[178] The original example can be written as $y = ax^0 + bx^1$, so $dy/dx = 0(ax^{0-1}) + 1(bx^{1-1}) = b$. The constant does not alter the slope, but rather just raises or lowers the function parallel to itself. More generally, if $y = a + bx + cx^2$, then $dy/dx = b + 2cx$.

To then find the maximum (*e.g.*, profit maximizing) or minimum (*e.g.*, cost minimizing) values of a function, set the slope equal to zero and solve for the corresponding x*.[179] For example, let $y = 10 + x^2$. The slope is $dy/dx = 2x$. Setting the slope equal to zero yields the solution for a local minimum at $x^* = 0$. The minimum value of the function is thus $y(0) = 10$.

An extension involves taking derivatives of derivatives, call second derivatives ($\mathbf{d}(dy/dx)/\mathbf{dx} = d^2y/dx^2$), third derivatives ($\mathbf{d}(d^2y/dx^2)/\mathbf{dx} = d^3y/dx^3$), and so on. Second derivatives are used to distinguish between a local maximum and a local minimum, both of which have a zero slope (first derivative).[180] To see how, note that, at a maximum, the slope is decreasing as x increases.[181] That is, slope is positive and decreasing as x approaches x* from the left, is zero at the maximum, and is negative and decreasing as x continues to increase. Hence, the second derivative test for a maximum is that, when dy/dx is zero (which yields x*), d^2y/dx^2 (the change in the slope) is negative at x*. Similarly, the second derivative test for a minimum is that d^2y/dx^2 is positive at x*.[182]

For example, let $y = 10 - 8x - x^2 + ⅓x^3$. The first derivative is $dy/dx = - 8 - 2x + x^2$, implying potential extrema of $x^* = \{-2,4\}$. To distinguish between a minimum and a maximum, d^2y/dx^2 is evaluated at these candidates. The second derivative is $d^2y/dx^2 = -2 + 2x$, and so $d^2y(-2)/dx^2 = -2 + 2(-2) = -6 < 0$, implying that the function has a local maximum of $y^*(-2) = -22/3$ at $x^* = -2$. Similarly, $d^2y(4)/dx^2 = -2 + 2(4) = 6 > 0$, implying that the function has a local minimum of $y^*(4) = -50/3$ at $x^* = 4$.[183]

Consider the profit maximization problem in Example E. GEN2's profit function is given by $\pi_2(Q_2) = 110Q_2-(2/5)Q_2^2$. The first derivative is given by $d\pi_2(Q_2)/dQ_2 = 110 - (4/5)Q_2$. Equating the slope to zero, $d\pi_2(Q_2)/dQ_2 = 110 - (4/5)Q_2^* = 0$ implies that $Q_2^* = 137.5$. To show that this maximizes profit, note that $d^2\pi_2(137.5)/dQ_2^2 = -4/5 < 0$. Since the slope always decreases as Q_2 increases, the profit function is always concave, and so the local maximum in this example is also the global maximum.

---

178. *Id.*

179. *Id.*

180. Cruzan, *supra* note 175.

181. *Id.*

182. Sometimes the test fails, as when the second derivative is zero. This occurs when, *e.g.*, a function reaches a plateau and then continues to increase. For example, $y = x^3$, where $dy/dx = d^2y/dx^2 = 0$ at $x = 0$. With the exception of $x = 0$, y always increases as x increases.

183. The qualifier "local" is necessary because the conditions for a maximum and a minimum only hold in a neighborhood around x*. Since the cubic term in this example enters positively, as x increases (decreases), y increases (decreases) without bound.

Yet another extension occurs where y depends on two independent variables: x and z. Now y changes as x alone changes ($\partial y/\partial x$, with z treated as a constant) or as z alone changes ($\partial y/\partial z$, with x treated as a constant), keeping other variables fixed.[184] The notation changes from dy/dx to $\partial y/\partial x$, with the derivatives now referred to as partial derivatives. The intuition is still that of the slope, but now along the dimension of the variable in question. Hence, $\partial y/\partial x$ measures the slope in the x dimension and $\partial y/\partial z$ measure the slope in the z dimension. In general, y may depend on many variables.

For example, let $y = ax^2z$. Then $\partial y/\partial x = 2(ax^{2-1}z) = 2axz$, with z treated as a constant. Similarly, $\partial y/\partial z = 1(ax^2z^{1-1}) = ax^2$, with x now treated as a constant. The rule applies to sums of functions, too. For example, let $y = ax^2z + bz^4$. Then $\partial y/\partial x = 2(ax^{2-1}z) = 2axz$, as bz^4 is treated as a constant and so drops off. Similarly, $\partial y/\partial z = 1(ax^2z^{1-1}) + 4bz^{4-1} = ax^2 + 4bz^3$.

Optimization is more complicated beyond two-dimensional functions. When functions are well-behaved, the minimum or maximum is characterized by simultaneous zero slopes in every dimension.[185] Mathematically, this boils down to a set of slope equations (first order conditions) that are set to zero and jointly solved for the independent variables.[186] (The test to distinguish between a minimum or a maximum is also more complicated, involving a matrix of second partials.)[187]

For example, suppose $y = a - bx^2 - cz^2 + xz$. The zero-slope conditions are the two-by-two system of partial derivatives set to zero:

(A1) $\partial y/\partial x = - 2bx + z = 0$; and
(A2) $\partial y/\partial z = - 2cz + x = 0$.

This yields $(x^*, z^*) = (0, 0)$. In general, for n independent variables, there would be n equations defining the solution.[188] Solving the system can be complicated, and there may not be a unique solution or even a solution.

A natural extension is to add constraints, which are common in electricity markets: supply must always equal demand, generator output cannot exceed capacity, energy flows cannot exceed transmission capacity, and so on.[189] This extension involves adding the constraints to the objective function. The constrained objective function is called a Lagrangian.

---

184. Of course, y will change as both x and z change. The partials measure the marginal effects of x ($\partial y/\partial x$) and of z ($\partial y/\partial z$), holding the other variables constant. It is also possible to find the slope along paths other than those parallel to the x and z axes.

185. Cruzan, *supra* note 175.

186. *Id.*

187. *Id.*

188. STATISTICS SOLS., INDEPENDENT AND DEPENDENT VARIABLES, https://www.statisticssolutions.com/independent-and-dependent-variables/ (last visited Sept. 15, 2019).

189. Bautista, *supra* note 156.

Consider the Lagrangian, equation (1) from section B, where load and the constraining variables have been replaced by their generic variable names (in place of the constraints):

$$(1) \quad \mathcal{L}(Q_1, Q_2, \lambda, \lambda_t, \lambda_1\, \lambda_2) = 10Q_1 + 20Q_2 + \lambda(L - Q_1 - Q_2) + \lambda_t(T - S_1Q_1 - S_2Q_2) + \lambda_1(K_1 - Q_1) + \lambda_2(K_2 - Q_2).$$

The constraint terms (with the λs) are all zero at the optimum: either $\lambda = 0$ (when the constraint does not bind) or the constraint binds. The coefficients on the outputs, Q_1 and Q_2, are the unit (incremental) costs. Hence, this specification of the Lagrangian indicates the total cost of production, subject to the constraints.

Partial derivatives can be taken with respect to any variable in the Lagrangian, treating the others as constants. To find the optimal values of the independent variables that maximize the Lagrangian, the partial derivatives in each dimension are set to zero.[190] Hence, for Q_1, all other variables are treated as constants, and so the partial is given by $\partial\mathcal{L}(\bullet)/\partial Q_1 = 10 - \lambda - \lambda_t /3 - \lambda_1$, which is set to zero. A similar result holds for the other variables. The full set of partials yields the first terms in equations (3a-f) above. The weak inequalities arise as the constraints sometimes preclude a solution with a zero slope for some variables. For example, not all generators will operate at capacity. To find the optimum, the set of n equations and n unknowns is solved for the n unknowns.

A related element of the Lagrangian involves the shadow prices, the implicit prices of the constraining resources, indicated by the lambdas. For example, the partial derivatives with respect to load, L, is $\partial\mathcal{L}(\bullet)/\partial L = \lambda$. Since $\mathcal{L}(\bullet)$ is the cost of producing Q_1 and Q_2 units subject to the constraints, $\partial\mathcal{L}(\bullet)/\partial L$ indicates how much cost increases if load increases by one unit. And since the load bus was designated the reference bus, lambda is thus the shadow price of another MWh delivered to the reference bus. The same intuition holds for the other shadow prices: $\partial\mathcal{L}(\bullet)/\partial T = \lambda_t$, the shadow price of the constraining transmission resource; and $\partial\mathcal{L}(\bullet)/\partial K_1 = \lambda_1$, the shadow price of the constraining generator capacity resource. The examples in the text walk through solving several n x n systems of equations in some detail.

Integral calculus, in contrast to differential calculus, concerns area or volume.[191] For example, consider the problem of finding the area under f(x) from x = a to x = b. The area may be approximated by breaking the range of x considered, b-a, into a set of equal segments, Δx. For each segment, evaluate the height by f(x) where x is the midpoint of the segment. The area then may be approximated by summing the areas of a series of rectangles whose height is f(x) and

190. U. CAL., RIVERSIDE, FUNCTIONS WITH VANISHING PARTIAL DERIVATIVES, http://math.ucr.edu/~res/math10A/zeropartials.pdf (last visited Sept. 15, 2019).

191. EUR. MATHEMATICAL SOC'Y ENCYCLOPEDIA MATHEMATICS, INTEGRAL CALCULUS, https://www.encyclopediaofmath.org/index.php/Integral_calculus (last visited Sept. 15, 2019).

whose base is Δx: $\Sigma f(x)\Delta x$. The approximation gets better as the width of the rectangles, Δx, gets smaller. As with derivatives, letting $\Delta x \rightarrow 0$ gives the exact area in the limit. The notation similarly changes: the summation sign Σ becomes the integral sign $\int$, and delta Δ becomes d. Hence, $\Sigma f(x)\Delta x \approx \int f(x)dx$.

Differential and integral calculus are, however, intimately related. The relationship, the Fundamental Theorem of Calculus, says that, for a continuous function on an interval [a,b], $\int f(x)dx = F(b) – F(a)$, where f(x) is defined as dF(x)/dx, the derivative of F(x).[192] Integral calculus is thus based on a process (integration) that reverses, in some sense, that of differentiation.[193] It is more difficult (sometimes impossible), however, to integrate than to differentiate a function.

Consider several examples. A simple example is f(x) = 2. To illustrate the basics, note that the area under f(x) from 0 to 2 is 4. To evaluate the area under the curve using the integral, note that F(x) is 2x, verified by taking the first derivative of F(x): dF(x)/dx = 2 = f(x).[194] Then the integral is $\int_0^2 2dx$ = 2(2) – 2(0) = 4, as expected.

A more complicated example is given by the demand curve of Section C: P(Q) = 100-Q/10. Total value (total willingness to pay) is the area under the demand curve from 0 out to Q units: $\int_0^Q P(q)dq$. This was applied to find consumer surplus, defined as total willingness to pay minus the required payment.[195] The unconstrained solution to the problem is Q* = 850 at a price of \$15, with a consumer surplus of \$36,125. To see this, first find total value at 850 units. This is given by $\int_0^{850}(100 - q/10)dq$. The antiderivative of 100 – q/10 is F(q) = $100q – q^2/20$.[196] This is evaluated at 850 and at 0, the difference being the integral value: $\int_0^{850}(100 - q/10)dq = [100(850) – (850)^2/20] – [100(0) – (0)^2/20]$ = \$48,875. This area under the demand curve is the most load is willing to pay for the 850 MWh of energy. Load only has to pay \$15 per unit, so after netting out \$15/MWh x 850 MWh = \$12,750, the net gain to load (consumer surplus) is \$48,875 – 12,750 = \$36,125.

---

192. This corresponds to the (signed) area under f(x) from x = a to x = b. Areas below the x axis have a negative sign attached to them.

193. EUR. MATHEMATICAL SOC'Y ENCYCLOPEDIA MATHEMATICS, *supra* note 191.

194. There is a constant of integration that appears. Recall that, when differentiating a function, the constant drops off. Reversing the process, integrating, means that it must be added. But when finding areas, the constant will again disappear, as F(b) – F(a) adds and subtracts the arbitrary constant of integration.

195. CORP. FIN. INST., CONSUMER SURPLUS FORMULA-GUIDE, EXAMPLES: HOW TO CALCULATE, https://corporatefinanceinstitute.com/resources/knowledge/economics/consumer-surplus-formula/ (last visited Sept. 15, 2019).

196. This can be verified by taking the derivative of $F(q) = 100q – q^2/20$: dF(q)/dq = 100 – q/10 = P(q). Again, the constant of integration is ignored.

A related example involves finding total cost when marginal cost is increasing, as in Section E. GEN1 had a marginal cost of $MC_1 = 2 + Q_1/200$. The total cost is the sum of the marginal costs for each of the units produced and corresponds to the area under the marginal cost curve over those units. For example, GEN1 produced 761.3 MWh in the unconstrained example. Its total cost is thus given by $\int_0^{761.3} MC1(q)dq = \int_0^{761.3}(2 + q/200)dq = 2q + q^2/400\Big|_0^{761.3} = 2(761.3) + (761.3)^2/400 = \$2{,}971.54$.[197]

A final problem relates to the link between the derivative and the integral. Going back to the total value example, it is natural to ask what happens to total value as Q increases marginally. The additional (marginal) value is given by the derivative of the integral used to measure total value. In general, the derivative of the integral measuring total value is price: $d/dq \int_0^Q P(q)dq = P(Q)$ - if we add a *very* small increment to output, the additional value that consumers get (the additional area under the demand curve as Q increases a bit), *measured on a full unit basis*, is price.

197. Again, to verify the integration: $(d/dq)\,(2q + q^2/400) = 2 + 2q/400 = 2 + q/200$.