Report 1.1: Energy Pricing

by

Alan Ransil

23 August 2018

Protocol Labs, MIT

Thank you to Sandhya Prabhu, MITei, for help with LMP methodology

Contents

1 Introduction: overview of GTDR system design **3**

2 Pricing in the bulk power system **5**

2.1 The simple case: no transmission constraints or losses 7

2.2 Locational marginal pricing 9

2.2.1 LMP Step 1: Solve Optimal Power Flow 11

2.2.2 LMP Step 2: Determine binding constraints 14

2.2.3 LMP Step 3: Price calculation 16

2.3 Multisettlement Systems and the rationale behind auctions 18

2.4 Ancillary services 19

3 Pricing in retail markets **21**

3.1 Retail rate setting in regulated markets 22

3.2 Other retail market arrangements 23

3.3 Why don't retail customers pay time of use rates? 25

1. Introduction: overview of GTDR system design

This project has two overall goals, as described in the work plan. Firstly, it will identify value propositions that an open-source energy trading protocol can provide as as the power grid is restructured to accommodate increased penetration of renewable energy. Second, it will outline the design of such a protocol.

This first report focuses on the mechanisms by which electricity is priced in today's power markets. Existing energy markets govern the infrastructure that any widely-used trading protocol must interface with in the short and medium terms. They also suggest the operational requirements that a protocol must satisfy in the long-term, if it is to eventually replace existing systems or become a ubiquitous platform upon which future wholesale markets are built.

Power grids throughout the world are based on a Generation, Transmission, Distribution and Retail (GTDR) model as shown in figure 1.

The system is highly centralized, with radial power flows from large power stations to small consumers. In the US in 2016, there were 8,084 power plants above 1MW nameplate generation capacity and over 150M retail customers, implying that each generating facility serves on average 18.6 thousand retail nodes.

This radial, centralized structure is the result of historical economies of scale in power plants. Because of the increased energy conversion efficiencies that could be achieved in larger plants, the ability to spread fixed costs over a larger generation capacity, and other economic efficiency gains, there was a strong incentive to increase the size of generation plants from the inception of the electricity industry until the 1970's when these economies of scale were exhausted.

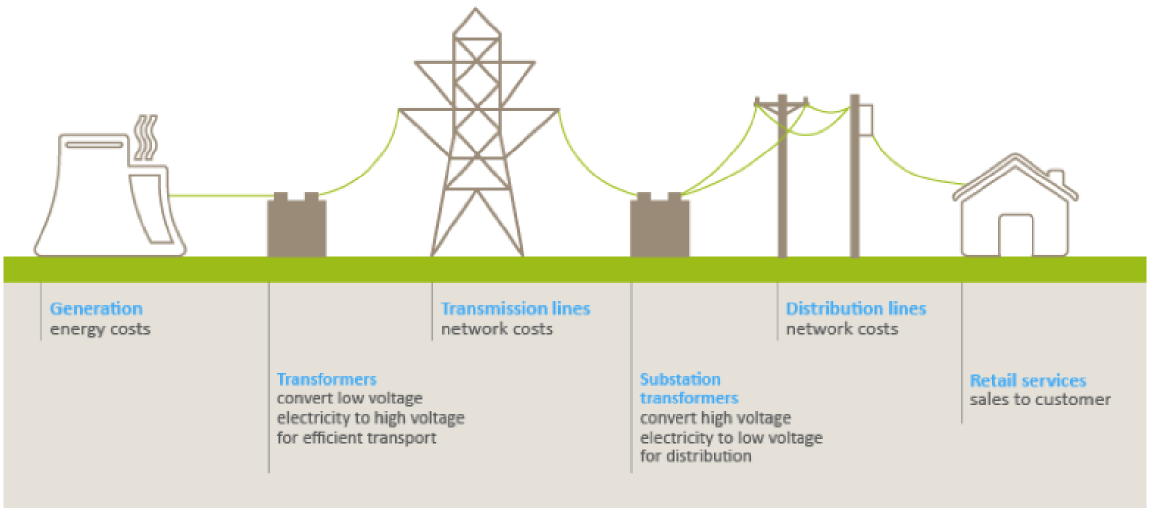

Figure 1: Schematic of the Generation, Transmission, Distribution and Retail (GTDR) model. From [2].

The division between Transmission and Distribution results from a need to minimize losses between Generation and Retail nodes. As power plants are in general far from retail customers, the electricity they produce must be transmitted over long distances. Because power line losses are proportional to the square of current, $P = I^2R$, lower currents will minimize losses.* As dictated by Ohm's law, $V/I = R$, high voltage may be used to minimize current and so high voltage lines (roughly 69kV - 765kV) are used in the Transmission

*Because T and D use AC power, there are non-ohmic losses associated with complex impedance and inductance as well.

system to transmit power over long distances. Substations, where high voltage power is stepped down for local distribution (ranging from about 2kV to 35kV) form the interface between the bulk power system (GT) and the distribution system. Distribution transformers, often inside of grey cylinders mounted on telephone poles and in metal boxes hiding in plain sight on street corners, then step down the distribution voltages to those used in residential, commercial, and industrial buildings.

The bulk power system (GT) is divided from the distribution and retail system (DR) not only by the technical details of power flow but also by the fundamentals of its market structure. Wholesale power markets governing the bulk power system, designed to achieve an economically efficient solution, are more or less free. Regulation at this level exists largely to level the field between players. On the other hand, the retail system is more or less socialized. Because power is considered a public good, customers pay the same average rate determined by regulators. This rate generally doesn't vary based on when or where power is used, despite the fact that this can greatly influence the cost of providing it. This report will explain the dominant pricing mechanisms in both of these markets, and then highlight some of the variations and exceptions to these general principles.

2. Pricing in the bulk power system

Power grids are grouped into ínterconnections' that tend to cover broad areas. The Unified Power System of Russia is an interconnected grid spanning nine time zones. All of Europe and most of North Africa form a single interconnection. China's grid has been fully interconnected since 2011. Brazil, Paraguy, Uraguay and Argentina share an interconnected grid, with the major Garabi interconnection linking Brazil's 60 Hz system to Argentina's 50 Hz grid. The United States has three interconnections. One covers its western half and another covers its east. Texas (of course) has its own.

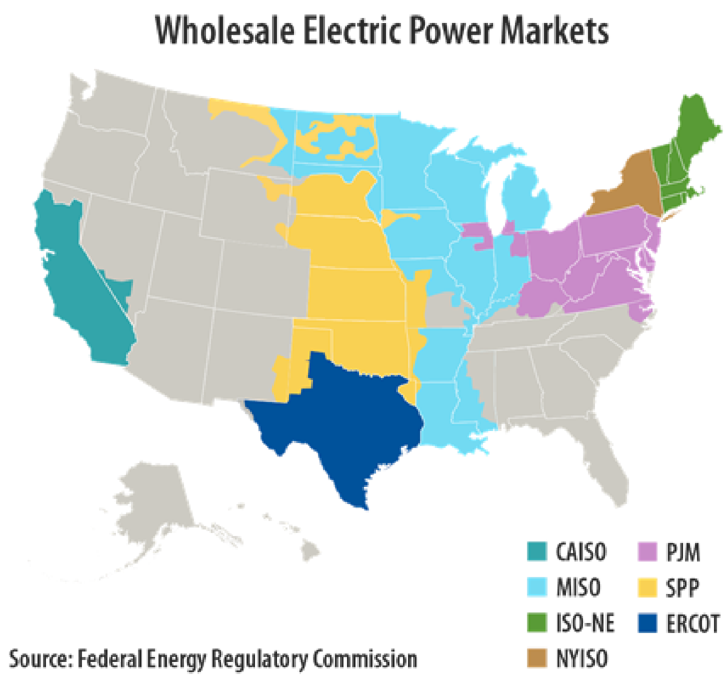

Figure 2: A map of bulk power system operators in the US. These operators run wholesale power markets. Grey regions are served by regional balancing authorities. From [1].

Large sections of all these power grids are run synchronously, meaning that they operate at the same frequency in phase across the entire grid. Because all major power grids have very little storage capability, it is necessary for a balancing authority to coordinate injections into and monitor withdrawals from the system to ensure that these are balanced on timescales of about one second. This role is accomplished in many regions by a system operator (known variously as an ISO, RTO or TSO). System operators in the US are shown in figure 2.

2.1. The simple case: no transmission constraints or losses

System operators (for example, ISOs) typically run wholesale markets using an auction system. The ISO accepts bids from generating facilities willing to produce power at a given price, and from retail utilities expecting to purchase a given amount power based on their anticipated demand. In a competitive market (meaning that no one entity controls enough generating capacity to give it market power, and wholesale markets are audited to assure that this is the case), every producer should bid exactly their marginal cost of production. If a producer bids a lower price they may produce power at a loss, whereas if they bid in a higher price they risk not being dispatched and forgoing revenue. In a typical fossil fuel power station, this marginal cost is determined by the cost of fuel.

The ISO will then construct a supply curve in which generation capacity is ordered according to its bid price, as shown in figure 3. The market clearing price resulting from the auction is then the price at which this supply curve intersects with demand.

In the case of an idealized grid, with no transmission constraints and no losses, the market clearing price defines the cost of power on the grid. All generators that bid prices below or equal to this price are dispatched and are paid this price.

This means that producers with a marginal cost lower than the clearing price, such as solar and hydroelectric generators with no fuel costs, are producing at a profit. Meanwhile the marginal producer, or the producer generating the most expensive unit of electricity, defines the clearing price and makes no profit. Generators that bid higher prices into the system are not dispatched by the ISO and sit idle.

On the other side of the market, all distribution utilities purchasing power from the grid pay the market clearing price for each unit of power.

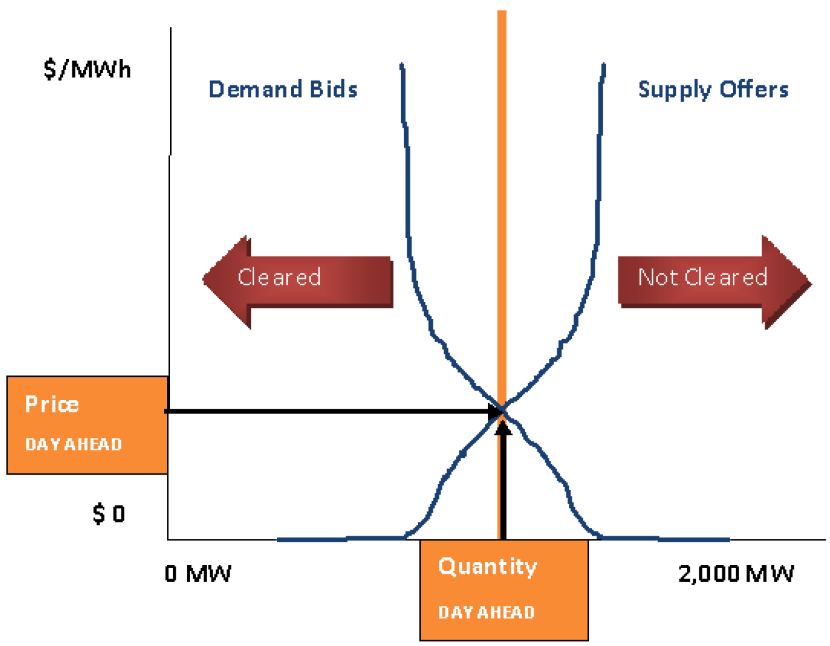

Figure 3: An example of a power auction in a day-ahead wholesale market. The ISO accepts bids from generation capacity and demand in the grid. The point at which supply meets demand sets the market clearing price, at which all wholesale transactions occur. All producers with marginal cost below this market clearing price are dispatched and operate at a profit, while the marginal producer is dispatched and makes no profit. From [3].

2.2. Locational marginal pricing

In reality power transmission lines do exhibit losses and transmission constraints, which lead to nonuniform costs throughout the system. Locational Marginal Prices (LMPs) take these factors into account, and generate price signals that vary across the grid based on location.

Losses are the result of electrical resistance through the cables: because the cables have a finite resistance (superconducting grid cables are regrettably rare), ohmic heating implies that some electricity will be lost as heat. Meanwhile, cables themselves have a finite capacity. This is determined by the wire temperature: wires will undergo ohmic heating, causing degradation and dangerous sagging, if their capacity limits are exceeded. Transmission constraints result from the limit on the current that can be run safely through a cable, and the number of cables that have been installed along a given path.

To take transmission losses and constraints into account, the ISO maintains a grid model of all of the injection (or electricity generation) and withdrawal (substation) nodes in the transmission system. Frequently, the system managed by one ISO will contain several thousand nodes.

Terms used in this section include:

- **Branch**: the equivalent of a transmission line in a grid model

- **Bus**: a node connecting branches, frequently a power station

Bids received from generators and retailers then specify supply and demand at specific buses in this model.

Because the grid is ultimately a physical system governed by physical power flow laws, changing the amount of power injected or withdrawn at any given location will change power flows throughout the entire system. Locational Marginal Prices (LMPs) take this into account, and are calculated based on the effect that individual bus behavior has on total system-level generation costs.

LMPs are found using a three-step process:

1. A physical model is solved in order to determine the power flow that minimizes total system-level generation cost
2. This model is used to determine:
    - The effect that changing the amount of power injected at any given bus would have on the total system-level power loss
    - The effect that adjusting each individual transmission constraint would have on the total has on the total cost of power generation in the system
    - The effect that changing the amount of power injected at any given bus would have on the amount of power flowing through each branch
3. The marginal cost at each location is calculated based on the system-wide market

clearing price and the effect that each bus has on system-level losses and congestion costs

Additionally, while the prices are determined at each node in the transmission system the LMP prices are not used directly to set prices. Instead, nodes are grouped into hubs and the weighted average of the LMPs at these hubs is used to set prices for all nodes in these hubs. This simplifies accounting and auditing.

The following description of LMP pricing calculations heavily references [4].

2.2.1. LMP Step 1: Solve Optimal Power Flow

The first step in determining LMPs is to solve a physical power flow model. Given the bus connectivity, the resistance of each branch, and the power withdrawals that will occur at each substation based on retail utility bids, a unique system state can be defined by the power P_G injected at each generator: $StateA = \{P^A_{Gi}\}$ for the set of generators i. Because setting the injection and withdrawal at each bus arbitrarily will overspecify the system state, injection at one bus must be allowed to vary. This is designated the ‘slack bus’. Computationally, injections and withdrawals at other buss are specified and power flows through each branch as well as injection at the slack bus are solved using Kirchoff’s laws.

A full, physically accurate transmission system model takes into account that transmission grids use AC power. Because the impedance in transmission lines is complex, power flow through a physical line induces a phase shift. Power injected into a transmission line with a zero phase shift (conventionally measured relative to the slack bus) will arrive at the other end with some finite shift. Accurate AC models are frequently used in order to calculate system-level power flows in day-ahead markets. Proprietary software, such as PowerWorld, PROMOD, and UPLAN, is generally used to do this, implementing more sophisticated algorithms than the linear model described below. However, while the physical model differs, the process (solving the physical model, using this to determine sensitivities, then using those sensitivities to calculate LMPs) is unchanged.

In a DC model, the power system is modeled as a linear set of equations. These models simply assign a real-valued resistance R to each branch and apply kirchoff's laws.

In order to determine the most economically efficient configuration of power flows, the ISO finds the solution to the set of physical power flow equations that minimizes the total cost of generation in the system. For generators i each producing power P_{Gi} at cost C_i this total cost of generation is:

$$Min \sum_{i=1}^{N} P_{Gi} \cdot C_i \tag{1. Minimize Generation Cost}$$

The system of physical equations is found such that this cost is minimized. The first equation is that the power injected into the system minus the power withdrawn and lost due to ohmic heating must equal zero. In addition to generated capacity this includes loads j each withdrawing power P_{Lj}, as well as the system loss L:

$$\sum_{i=1}^{N} P_{Gi} - \sum_{j=1}^{M} P_{Lj} - L = 0 \tag{2. Energy Balance}$$

Next, transmission constraints must be taken into account. It must be assured that for each branch k the power along the branch is lower than or equal to the transmission limit of this branch. In a linear model, the shift factor S_{ki} is defined as the increase in power running through branch k corresponding to an increase in injection at generator i. Then there exists a constraint on the system for every branch k such that:

$$\sum_{i=1}^{N} P_{Gi} \cdot S_{ki} \leq T_k \qquad \text{(3. Transmission Constraints)}$$

Lastly, each generator G has a set of generator capacity constraints. The minimum capacity is the lowest that the generator can safely be set to, which may be nonzero (particularly in the case of some hydroelectric and nuclear plants).

$$P_{Gi}^{M}in \leq P_{Gi} \leq P_{Gi}^{Max} \qquad \text{(4. Generation Constraints)}$$

This model determines the cheapest way to supply all demand in the system. Next, the solution must be optimized in order to ensure system reliability. This is done through contingency simulations. In these tests, beginning with the previously determined optimal system state the effect of removing any one element (such as a branch or generator) in the system is simulated. The resulting power flow is then calculated (allowing the injection state to vary at the slack bus, as before). Thermal and stability limits, such as transmission line tolerances, are then checked in this new contingent state. Adjustments are made to the optimal generation configuration such that thermal and stability limits would not be exceeded by these contingencies. This process can involve tens of thousands of simulations, in order to ensure system reliability in case of component failure.

The resulting power flow configuration that minimizes generation cost, satisfies transmission constraints and has passed the contingency tests is specified by some set of generation targets which are communicated to power generators as dispatch orders. This is known as Security Constrained Economic Dispatch.

2.2.2. LMP Step 2: Determine binding constraints

The next step is to assess the sensitivities of the model to changes in power injected at a bus and to changes in the transmission constraints imposed by branches in the system. For this analysis, a linear model can be used and sensitivities can be calculated numerically.

Changes in injection at each bus are modeled in order to determine both their effect on power transmitted through each branch, and on the overall system loss. The effect on power in each branch is quantified through the shift factor S_{ik}, which is used in the linear model (above). This can be evaluated numerically by beginning with the optimal dispatch configuration, adjusting the current injected at bus i by a small amount, and calculating the resulting overall system state (again, with injection at the slack bus being altered to maintain system balance). Dividing the resulting change in power through branch k, P_k, by the change in injected power at bus i, P_{Gi}, results in a ratio:

$$S_{ik} = \frac{\Delta P_k}{\Delta P_{Gi}} \qquad \text{(Shift Factor)}$$

Similarly, the overall system loss will change as injections at various buses are varied. This is because different transmission lines have different resistances, so altering the distribution of power flow will change thermalization losses. The change in system-level transmission loss as a result of changing the injection at bus i, the loss factor LF_i, can be calculated similarly to the shift factor:

$$LF_i = \frac{\Delta Loss_{System}}{\Delta P_{Gi}} \tag{Loss Factor}$$

Finally, congestion due to transmission line constraints will also affect costs. This is because such constraints require power flows to be adjusted on the grid in order to avoid overloading branches. Adjusting power flow is accomplished by changing generation levels. For example, when a transmission constraint forces a low marginal cost generator to inject less power because of its location on the grid a higher cost generator at a different location may be dispatched to maintain system balance. This will raise the marginal cost of power generation. Additionally, adjusting power flow will alter system losses which may require additional injection.

Sensitivity to transmission constraints is determined numerically by lessening the transmission constraint T_k^{Max} by one unit, and calculating the corresponding optimal power flow. That optimal power flow will correspond to a new generation cost $C_{G,Total}$ This is known as the shadow price, μ_k, corresponding to the transmission constraint on branch k:

$$\mu_k = \frac{\Delta C_G}{\Delta T_k^{Max}} \qquad \text{(Shadow Price of Transmission Constraint)}$$

For transmission lines that are operating below their maximum capacity in the optimal power flow configuration, adjusting the transmission limit will have no effect on power flow and the resulting derivative μ_k will be zero. These are called 'non-binding' constraints. By contrast, transmission lines that are operating at their limit for a given system configuration are known as binding constraints and will affect marginal prices.

The final sensitivity to be calculated is the overall system sensitivity to changes in net power demand. This quantity, λ, is the marginal cost of energy based on generation bids in the system (as explained in the simple case involving no transmission losses above).

2.2.3. LMP Step 3: Price calculation

LMPs, λ_i, are the result of the marginal energy cost set by the marginal producer, and the binding constraints on the system. At a given bus i:

$$\lambda_i = \lambda - LF_i \cdot \lambda + \sum_{k=1}^{K} S_{ik}\mu_k \qquad \text{(LMP at bus i)}$$

This equation has three distinct parts. The first component, λ, is the marginal cost of energy set by the marginal producer. This corresponds to the system balance constraint.

The second component, $LF_i \cdot \lambda$, accounts for the cost of losses caused by injection at bus i. This cost scales both with the effect of bus i on overall system losses, and with the cost λ of supplying a marginal unit of energy to the system. This term is negative because high losses resulting from power injection at bus i will cause power at that location to be less valuable (less highly compensated) than power injected elsewhere in the system.

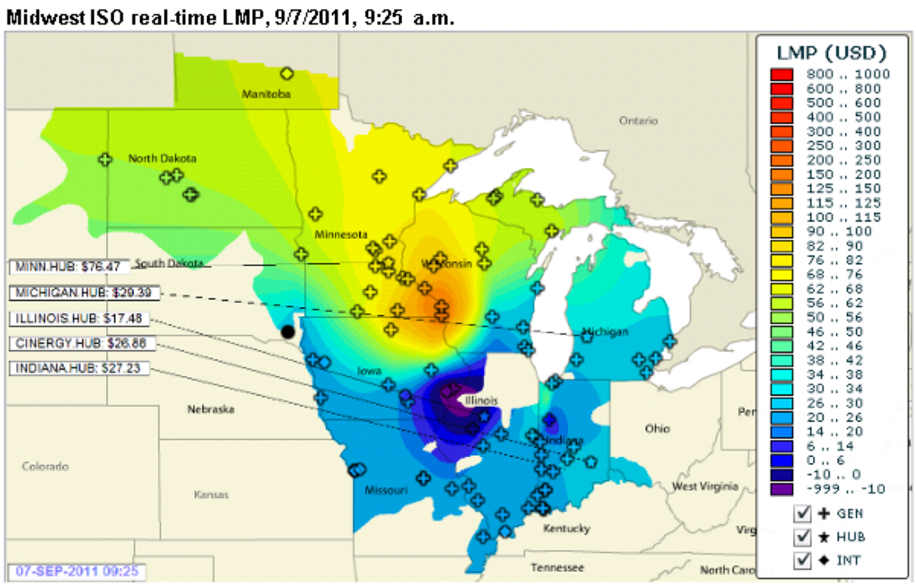

Figure 4: A heat map of locational marginal prices (LMPs) across the great lakes region at 9:25 AM on September 7th, 2011. As shown, the marginal price varies from below zero in regions with excess generation to \$200/MWh in areas with high demand. These differences are due to transmission constraints and losses. From [5].

The third component accounts for binding congestion constraints. The sum is taken over all branches k. This quantifies the effect that injection at i has on congestion. If a generator i is downstream of congestion through branch k, then increasing its generation would lower congestion through k. In this case, $S_{ik}\mu_k$ would be positive. Generation at this location is considered to be more valuable, so the congestion component of λ_i is positive which increases revenue to the generator.

In the case of zero losses (all $LF_i = 0$) and no transmission constraints (all $\mu_k = 0$), all LMPs λ_i reduce to the system-wide marginal cost λ. Losses and congestion constraints are the factors causing price to vary at different buses across the transmission system.

Figure 4 shows a snapshot of LMPs in the great lakes region at a moment in time. As shown, prices can vary dramatically across an interconnection due to differences in the local marginal cost of production (ex: the placement of renewable generation) coupled with binding system constraints.

2.3. Multisettlement Systems and the rationale behind auctions

The auctions described above are not run as a continuous free and open market. Rather, power markets are 'multisettlement' systems. In such a system, power trades are committed at distinct time points prior to the time the power is actually produced and consumed (consumption happen milliseconds after production, due to the lack of storage in the grid).

On long timescales, multiyear contracts commit some generation and consumption in the system. These are frequently bilateral agreements between power generators and consumers. Generators may be large utilities or smaller installations (such as the Independent Power Producers introduced by the Public Utility Regulatory Policies Act, or PURPA). Consumers may be retail utilities aiming to ensure price stability or individual entities such as universities or companies purchasing renewable energy to meet voluntary energy portfolio targets. The fraction of power that is traded in such agreements prior to wholesale market auctions varies, but can be as high as 80%.

Wholesale market auctions as described above are split into a day-ahead market which is run every hour as well as 'real-time' auctions run fifteen minutes and five minutes ahead of the time of use. Exact time points vary based on the ISO's market design decisions. This allows generators to plan ahead on different timescales according to their technical and economic requirements and ensures that supply will meet demand based on increasingly accurate demand forecasts as the appointed minute draws nearer.

Auctions at specific time points were chosen as opposed to a continuous free market for two reasons. Firstly, the ISO needs to simulate the grid's physical state in order to maintain balance in the system. This technical problem is more tractable when run at discrete time points rather than as a continuous market, which was especially relevant given the limited IT systems of past decades. Secondly, the absolute number of generators selling power and distribution retailers purchasing it is low compared to those in other markets (such as those for commodities or stocks). Without some structure imposed on the market to coordinate transactions between sellers and buyers in time, the relatively small number of players could lead to illiquidity. Running auctions at specified time points is a way of bringing market participants together to ensure market liquidity.

2.4. Ancillary services

In addition to wholesale markets and long-term procurement contracts, markets for Ancillary Services (AS) are used to ensure grid reliability over the long and short term. The design of these markets varies greatly by region. A (non-exhaustive) list of common AS are listed below.

- **Forward capacity market**: a market to incentivize investment in generation infrastructure to ensure adequate capacity over longer timescales than those served by wholesale markets. Annual auctions are held in which capacity is purchased three years ahead. Payments are made whether or not the capacity is dispatched, as the goal is to ensure that capacity will be adequate whether or not it is actually needed in practice. Some systems (such as Texas' ERCOT system and the Southwestern Power Pool) do not run capacity markets and rely on intentionally high short-term scarcity prices instead to incentivize investment.

- **Forward Reserve Market and Real-Time Reserve Pricing**: Reserve markets pay generation resources to 'stand by' and not produce power, but to be ready to in case demand isn't met by other sources. Like forward capacity markets, these payments are made whether or not this capacity is dispatched.

- **Frequency and Voltage Regulation Markets**: Frequency and Voltage regulation must be dispatched on a second by second basis. In the US, FERC mandates that frequency on the wholesale system must not deviate from tolerances set by the North American Electric Reliability Corporation. Generally, Automatic Generation Control assets are owned by utilities but controlled by balancing authorities (ie ISOs) and both capacity and actual service mileage are compensated as AS. It is worth noting that the value of regulation is highly locationally dependent.

- **Blackstart Service**: Generation able to start the grid following a blackout is compensated in a separate AS market.

3. Pricing in retail markets

As shown in the GTDR model above, the process of actually delivering and selling electricity to consumers is carried out in the Distribution and Retail (DR) system. Physically, these are the medium and low voltage power lines that run throughout our communities and bring power into our homes, stores, offices and factories.

The distribution system is separated from the bulk power system not only physically (by substations) but from an economic and regulatory standpoint. A major difference between wholesale and retail markets is that while generators are paid a time-dependent rate, these price signals are generally not passed on to retail customers. Instead, electricity retailers purchase power both on the wholesale market and through long-term contracts. They then sell this power to customers at some predetermined rate approved by their regulators. As a result of this system, the distribution utility is exposed to price fluctuations in wholesale markets but must get approval in order to pass these costs on to its customers. This situation can result in utilities losing money or going bankrupt if wholesale market costs rise quickly (PG&E's 2001 bankruptcy being a famous example of this)

One consequence of this market design is that customers are not incentivized to adjust their power use based on grid demand, because time-of-use pricing signals are generally not passed to them. If customers were able to benefit from adjusting their behavior in ways that stabilize the grid, they may be more conducive to grid stability and higher renewables penetration. (See section below, 'Why don't retail customers pay time of use rates?')

3.1. Retail rate setting in regulated markets

Investor Owned Utilities (IOUs) historically have been operated as monopolies in their territories. In exchange for this privilege, they are required to service all paying customers in that territory and to submit to a high degree of regulation. In these markets, the local regulator such as the state Public Utility Commission (PUC) must approve the rate schedule charged to retail IOU customers. This schedule sets rates based on classes of customer, frequently charging different rates to Residential, Commercial and Industrial customers.

The bill will have several components, including:

- Supply charge: covers the cost of procuring electricity
- Transmission charge: covers cost of infrastructure in the bulk power system
- Distribution charge: for local distribution system infrastructure
- Service Charge: covers the operations of the utility
- Other miscellaneous charges to support renewables, recover the costs of stranded assets, etc

For a utility to adjust these rates, they must undergo a rate review with the PUC in order to seek approval. As a first step in this process, the utility must provide supporting data justifying the rate change. This will include information such as costs the IOU has spent or proposes to spend on infrastructure, the rate base (ie assets for which the IOU believes it should earn a rate of return), corporate structure and proposed profit margin. The PUC then evaluates these data and holds a series of public hearings in order to make the rate setting process transparent. It will then approve or modify the proposed rate schedule.

3.2. Other retail market arrangements

In many cases, the distribution market structure differs somewhat from that described above. While in regulated markets retail customers are required to purchase both delivery and supply from one local utility, there has been a trend in the US towards deregulation of retail markets since the 1990s. Retail market deregulation allows customers to purchase supply (but not delivery, which includes Transmission and Distribution) from competing retailers. This allows competitive non-utility suppliers to buy electricity in either wholesale markets or from producers and sell to customers. These markets also allow brokers to arrange sales (called bilateral trades) directly between producers and consumers. While large consumers such as factories and large companies have most frequently availed themselves of these arrangements, smaller consumers such as residences have increasingly purchased power from non-utility suppliers in deregulated markets.

The reasons that customers have for entering deals with non-utility suppliers vary. They may buy energy from a supplier to contractually assure price stability, to purchase renewable energy directly, or to undercut the market rate by making a deal with an inexpensive provider. In any case, it should be noted that deregulated suppliers are generally (though not always) competing with a regulated utility. This would prevent them from gaining market power and raising rates, as could happen in a truly free market.

Another trend is towards 'decoupling' of utility revenues from sales. The motivation to do this is to incentivize utilities to promote energy efficiency among their customers. If revenues depend on sales and PUC-approved schedules change only rarely, then utilities have a financial incentive to encourage customers to purchase more. In a decoupled system, PUCs allow the utility the right to collect a given amount of revenue regardless of the amount of electricity sold. In this case, lower sales allow the utility to spend less procuring power without sacrificing revenue.

A panoply of other arrangements exist, often heavily dependent on local history and regulation. For example, municipal utilities are government-owned and set their own rates rather than requiring PUC approval. Community choice aggregation is another alternative arrangement, in which a group of residential customers form a single entity able to negotiate supply contracts with providers.

3.3. Why don't retail customers pay time of use rates?

Some do (as will be discussed in subsequent reports), but this is rare. As described above, there is a theoretical economic benefit to passing price signals on to customers in that they might then adjust their power use in ways that could help stabilize the grid. There are three reasons that retail markets are not generally designed this way today.

The first reason is social and economic. Because electricity is a public good and a crucial element of our economy, we want to protect residential customers and businesses from large price swings. We have decided as a society that everyone should have access to an inexpensive, reliable electricity supply. This is a major reason that even when some price signals are passed to the end customer, there will likely need to be a ceiling imposed on the price in order to prevent sudden price spikes.

The second reason is technical. Historically, it has not been possible to charge TOU rates to retail customers because the measurement, reporting and computing infrastructure did not exist. Before the advent of the smart meter, price signals could not be measured by utilities in real time (and are even now frequently only measured once every month for billing). Furthermore, because the number of nodes involved in the retail system is two to three orders of magnitude greater than those involved in the distribution system, the computational resources to take distribution nodes into account when running the market did not exist. This is an argument for designing future systems that shift computing to the grid edge using distributed algorithms where possible. This is particularly true if IOT devices will play a role in markets and increase the number of relevant nodes even further.

The third reason that consumers aren't exposed to price signals is path-dependant and economic. The GTDR system is the prototypical example of a 'natural monopoly,' because of the economies of scale in generation and the fact that wires have to physically go everywhere served by the utility. In the twentieth century, the prevailing logic was that it therefore made sense to grant monopolies to a given utility in a given geographical area. This necessitated regulation in order to prevent that company from raising prices charged to a captive consumer base, and paying no penalty in market share.

Such received wisdom, that the electric system is a natural monopoly, is unchallenged when it comes to transmission and distribution infrastructure. (Power lines are expensive and must geographically connect all nodes on the grid.) However, in contemporary design of wholesale power markets and deregulation of retail markets is has been generally accepted that the GTDR system is not monolithic. While the T&D sections are natural monopolies, the end members G and R can be supplied by competing providers albeit within the context of a heavily controlled and regulated framework.

Additional variations on all of the markets discussed such as passing price signals to end users, regulations that affect priced in order to introduce clean energy, alternative resources such as aggregated demand response and the operation of solar home systems and microgrids as an alternative to or supplement to the grid, will be discussed in further reports.

References

[1] United States Environmental Protection Agency. U.s. electricity grid and markets. https://www.epa.gov/greenpower/us-electricity-grid-markets. (Accessed on 7/16/2018).

[2] Queensland Competition Authority. Electricity facts. http://www.qca.org.au/Electricity/Further-Info/Electricity-facts. (Accessed on 7/29/2018).

[3] ISO New England. Overview of new england's wholesale electricity markets and market oversight. 2014.

[4] ISO New England Litvinov, Eugene. Power system and lmp fundamentals. 2008.

[5] Midwest Independent System Operator. Midwest iso real-time lmp map. 2011.