Chapter 4.
Real-time Energy Market

4.1 Introduction

4.1.1 Chapter Overview

This chapter reviews the performance of the ISO's real-time imbalance energy market during the first year of operation. Section 4.1 provides a brief description of the real-time market; a more detailed description is provided in Section 2.5. Section 4.2 summarizes key successes, failures and ongoing issues in the real-time markets. Section 4.3 presents a detailed look at demand in the real-time market, while Section 4.4 looks at supply, and Section 4.5 discusses prices.

4.1.2 Real-time Market Description

Because there is always some deviation in real time of actual generation and load from what was scheduled, one of the key functions of the ISO is to perform real-time balancing of loads and generation. The ISO performs this through the real-time imbalance energy market, which is the mechanism whereby supply resources are selected to be increased (*incremented*) or decreased (*decremented*) in order to maintain system balance. Real-time balance is maintained through the combined use of units providing Regulation Reserves, which are on automatic generation control (AGC), and units providing other ancillary services and/or supplemental energy bids. The non-AGC units are dispatched every 10 minutes through the Balancing Energy and Ex-Post Pricing system (BEEP), while the Regulation units are used only to respond instantaneously to system imbalances. Resources dispatched through the BEEP system are incremented or decremented as needed to allow Regulation units to return to their Preferred Operating Point (POP).

Resources available in the BEEP system for incrementing generation include Spinning, Non-spinning and Replacement Reserves, as well as resources that submitted supplemental bids for incremental energy. The only resources available in the BEEP system for decrementing generation are those that submitted supplemental bids for decremental energy. Decremental energy bids are submitted by units that are, in effect, willing to curtail their generation schedules by buying back a portion of those schedules in real time. All of these resources are pooled in the BEEP system and arranged in merit order based on their bid prices. In the case of the A/S units, the bids used in the BEEP are the energy components of their two-part A/S bids.

Operation of the real-time market is closely connected to the ISO's responsibility for controlling the actual dispatch of generation and managing the reliability of the transmission grid. Although generation and load are in balance when the ISO issues the final hour-ahead schedules, the market design does not impose explicit penalties on generators or loads for real-time deviations from the final schedules. All deviations from final schedules are settled at the hourly real-time imbalance energy price. Those SCs providing extra supply (or having less than scheduled load) earn this price, while those having extra demand (or providing less than scheduled supply) pay this price. The ISO's real-time imbalance energy market is, in practice, the only truly physical market for energy. In the California market structure the real-time market is the *spot market* for energy, upon which all financial settlements for unscheduled generation and load are ultimately based.

4.2 Key Market Issues, Successes and Failures

4.2.1 Real-time Price Cap

Just prior to the beginning of market operation on March 31, 1998, the ISO imposed a cap of $125/MWh on energy bids into the ISO's real-time imbalance energy market. This cap became known as "the BEEP cap." It was put in place as an interim measure to prevent predicted run-ups in the market clearing price resulting from problems with the BEEP software, and was later raised to $250/MW.[1] Subsequently, based on operating experience in the summer of 1998, a number of additional problems with the ISO's BEEP software and the current market design were identified as inhibiting the efficient operation of the real-time market. As a result, the BEEP cap has remained in effect, although it is explicitly recognized as a temporary measure to be used only until design imperfections and operational problems have been resolved. In fact, the timing and rate at which the price cap will be raised have been linked to the implementation of several market redesign elements scheduled for later this year.[2] It is important to note that the real-time price cap serves as a *de facto* cap on the PX day-ahead price, which is currently capped by the PX at $2,500. Buyers would have no reason to purchase PX energy at a price above $250, since they can rely on the ISO to meet any unscheduled load at a real-time price not exceeding $250. Thus, due to inter-relationships among the markets, the ISO's real-time imbalance price cap can be expected to have significant impacts on PX prices and on the ISO's own A/S markets.

4.2.2 Price Spikes

During the first year of operation, the price of real-time energy hit $250 in 48 hours, and exceeded $200 in 72 hours. Review of market performance during peak months since the beginning of market operation indicates that market power has frequently existed in the real-time energy market, particularly when high overall system loads have coincided with relatively high demand for imbalance energy.

4.2.3 Impacts of Operational and Software Constraints on Real-time Market Efficiency

During the ISO's first year of operation, a variety of operational constraints were identified that have had a significant impact on the efficiency of the real-time market by preventing the ISO from obtaining the resources necessary to meet demand at the least cost. While some of these constraints may be resolved with the software modifications scheduled for implementation in

[1] The initial cap of $125/MWh imposed prior to market open was raised to $250 on May 27, 1998, after the $125 cap was hit on May 18.
[2] In its Order of May 26, 1999, the FERC removed the ISO's authority to impose price caps effective November 15, 1999. Should events in the markets between now and then indicate a need to continue the use of price caps, the ISO will have to apply to FERC for an extension of this authority.

summer 1999, the efficiency of the real-time market can continue to be improved as operational protocols and software are further enhanced and refined.

4.2.4 Under-scheduling of Load

The basic design of California's energy markets calls for the bulk of demand to be met either through bilateral contracts or in the forward (day-ahead or hour-ahead) energy markets. Either way, the bulk of demand and the supply to meet that demand would be scheduled in advance. This principle is reflected in ISO protocols, which require that Scheduling Coordinators (SCs) submit balanced schedules to the ISO. These schedules represent only financial commitments by the SCs, however. As noted above, the ISO's imbalance market is, in practice, the only real physical market for electricity in California and represents a spot market for energy in the ISO's control area. During the first months of operation buyers quickly recognized this fact. They often under-scheduled load in the forward markets and relied on the ISO to meet the balance of their demand in the real-time market. During numerous hours during the peak summer months, total system loads exceeded final hour-ahead schedules by over 5,000 MW, or more than 10 percent of total system load. Certain incentives existed which reinforced this practice. First, such under-scheduling provided a source of demand elasticity which could limit prices in the PX markets. Second, the ISO was billing A/S costs to SCs based on their *scheduled* rather than their *metered* loads. During periods of the highest demand, and the greatest under-scheduling of demand, A/S costs billed on scheduled demand equaled about 15 percent of overall energy costs. Looking forward, the ISO and other market participants are exploring ways to increase the flexibility of load to respond to market conditions without simply shifting it to the real-time market. With respect to the second incentive, the ISO will begin billing SCs for A/S based on their metered load as an element of the market redesign being implemented in summer 1999.

4.2.5 Lack of Demand Elasticity

As noted above, the ability of buyers to shift load into the ISO's real-time market currently represents the primary source of demand elasticity in the forward markets. In competitive markets, demand elasticity – the ability of customers to respond to prices by modifying their demands or finding other ways to meet them – is the primary protection consumers have against market power. Thus, introducing additional sources of demand elasticity into the forward and real-time markets has been identified as a key necessity for increasing the competitiveness of California's energy markets and for reducing and eventually eliminating the need for price caps. Two approaches presently under development are: (1) a participating load agreement (PLA) through which loads can contract directly with the ISO to participate in the A/S markets by bidding load reduction to offset some of the ISO's A/S requirements, and (2) new interruptible retail tariffs proposed by the investor owned utilities (IOUs), to provide incentives for their larger retail customers to curtail their loads quickly in response to instructions from the IOUs.

4.3 Demand for Real-time Imbalance Energy

Imbalances between system loads and generation occur in real time due to differences between *scheduled* and *actual* demand and supply. Figure 4-1 shows how the real-time imbalance energy requirements for the ISO system have varied at different levels of total system load over the first year of operation. System imbalances have varied widely at any level of total system load. At system loads of up to about 35,000 MW, imbalances have been evenly distributed between negative and positive, representing conditions of over- and under-generation, respectively. At system loads above 35,000 MW, however, imbalance energy requirements are predominantly positive. The ISO has needed to increment generation in approximately 85 percent of the hours in which loads have exceeded 35,000 MW.

Since ISO protocols require SCs to submit balanced schedules, all load imbalances in the ISO system ultimately result from deviations between scheduled and actual loads and generation. Two additional factors affecting system imbalances are: (1) dispatch of RMR generation that is not balanced against scheduled load, and (2) inaccuracies in how transmission losses are accounted for in schedules submitted to the ISO. Figure 4-2 depicts the various sources of real-time imbalances, which are discussed in more detail below.

Figure 4-1. Real-time Energy Imbalances and Total System Loads

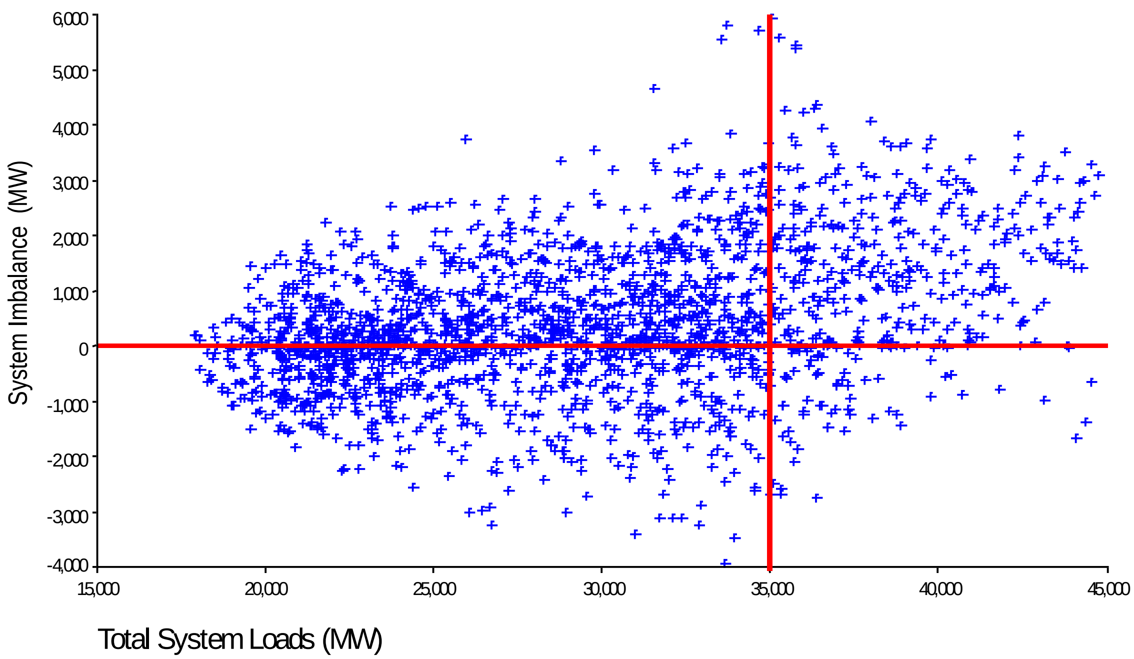
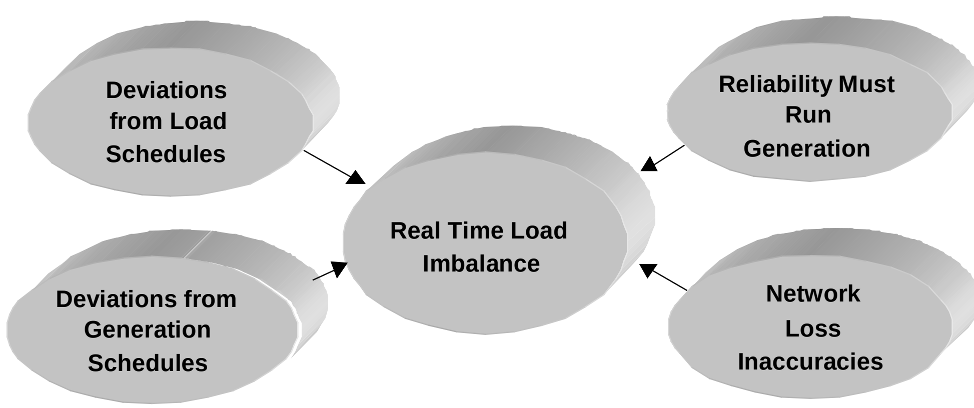


Figure 4-2. Sources of Real-time Energy Imbalances



4.3.1 Deviations from Scheduled Loads

Differences between scheduled and actual demand may result from both unintentional and intentional under- or over-scheduling of loads:

- **Load forecast errors, or unintentional mis-scheduling of demand.** Even the best forecasting models in use by SCs typically have average errors of at least ± 2 percent. When extreme or unanticipated weather conditions occur, loads tend to deviate by even greater margins from the forecasts upon which day-ahead schedules are based.

- **Intentional mis-scheduling.** ISO protocols require that SCs submit balanced schedules at two times, i.e., after the close of the day-ahead and hour-ahead PX markets. As noted above, however, in practice the ISO's imbalance energy market serves as a spot market for energy and buyers use it to their advantage. If buyers believe that real-time imbalance prices will be lower than day-ahead prices, they may intentionally under-schedule loads by submitting a balanced schedule in which expected loads are underestimated. If buyers expect the real-time prices to be higher than forward prices, they may over-schedule loads and then sell any extra energy procured in the forward markets back to the ISO at the real-time price. The ability of buyers to shift a portion of their demand from the forward market to real time provides one of the only sources of demand elasticity in the forward market, and may help to keep prices lower in the forward markets at times when market power exists. At the same time, this behavior also increases price volatility, the probability of real-time price spikes, and overall average prices in the real-time market.

Large differences between scheduled loads and actual system loads were prevalent during the ISO's first six months of operation. During these months, total system loads frequently exceeded final hour-ahead schedules by over 5,000 MW, or more than 10 percent of total system loads. Figures 4-3 through 4-6 show average hourly schedules and system loads for the ISO's first year of operation. During the first three months, final hour-ahead schedules were an average of about 9 percent below actual loads during peak hours, and about 6 percent lower than actual loads during off-peak hours (see Figure 4-3). During the peak months of July to September, when real-time prices rose significantly above day-ahead PX prices, final-hour ahead schedules were an average of almost 3 percent below actual loads during peak hours, and about 1.5 percent lower than actual loads during off-peak hours (see Figure 4-4).

Since October 1998, final hour ahead schedules have closely tracked actual system loads during off-peak hours, while final schedules continue to fall systematically short of actual loads during peak hours. From October to December, system loads exceeded final hour-ahead schedules by an average of about 2,000 MW, or 1.6 percent of total loads (see Figure 4-5). From January to March 1999, system loads exceeded final hour ahead schedules by an average of about 3,000 MW, or over 3 percent of total system loads. As shown in Figures 4-5 and 4-6, the highest loads and the largest amounts of under-scheduling in the fall and winter months occurred during the evening hours (hours 18 to 22).

Figure 4-3. Average Hourly ISO System Loads and Schedules (April–June 1998)

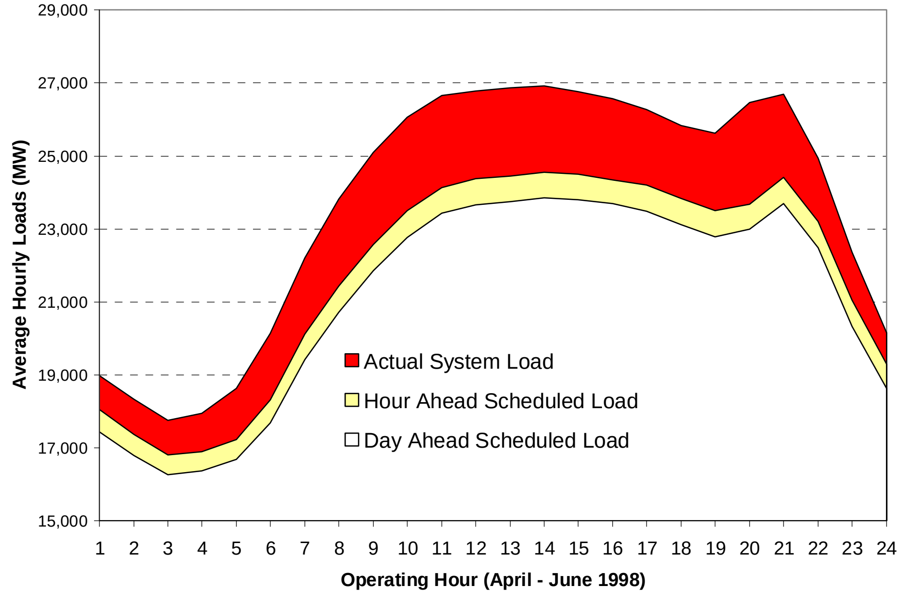


Figure 4-4. Average Hourly ISO System Loads and Schedules (July–Sept. 1998)

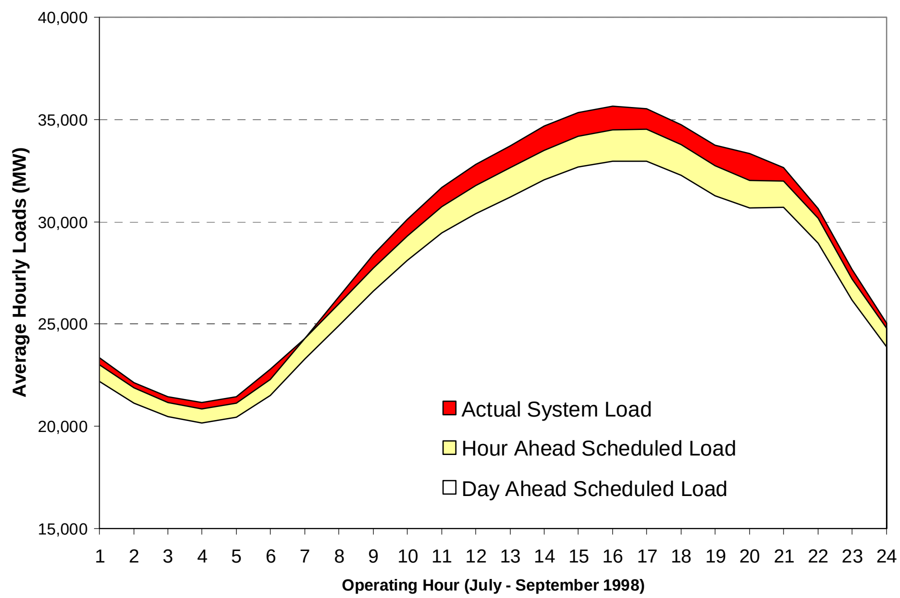


Figure 4-5. Average Hourly ISO System Loads and Schedules (Oct.–Dec. 1998)

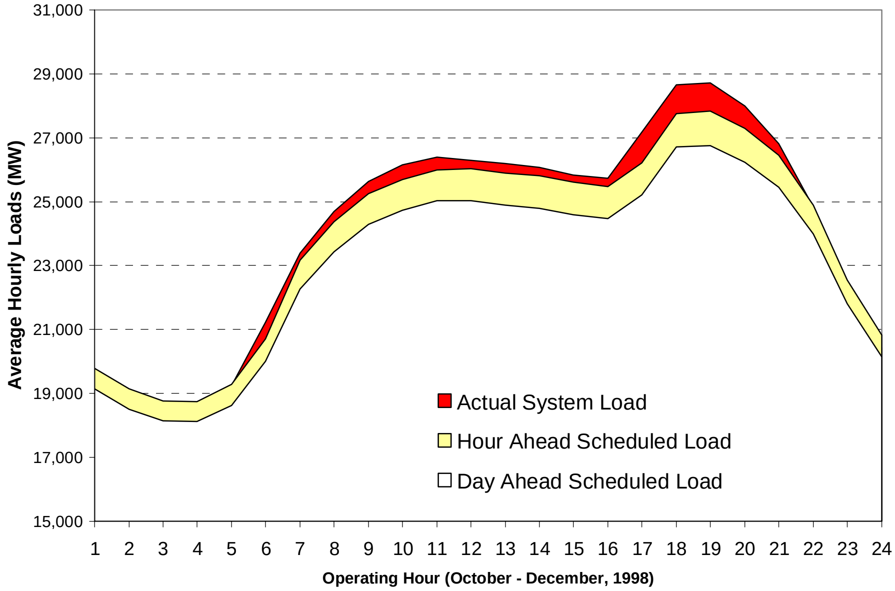
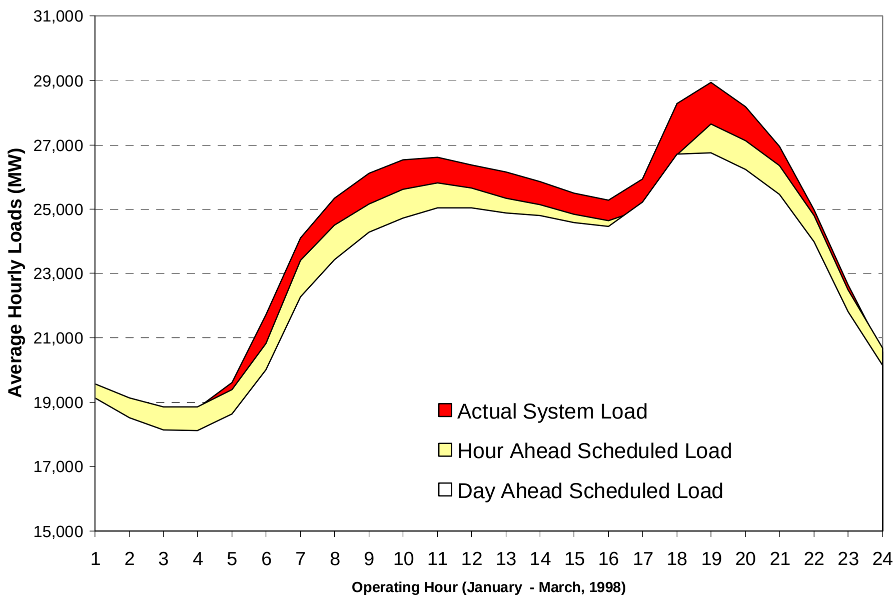


Figure 4-6. Average Hourly ISO System Loads and Schedules (Jan.–March 1999)



4.3.2 Deviations from Generation Schedules

Differences between scheduled and actual supply occur for a variety of reasons, including:

- **Unplanned outages and deviations from generation schedules.** Actual generation may deviate significantly from scheduled generation due to unit tripping and other sources of unplanned outages. Even if a unit is operating, generation may deviate from schedule due to ramping and unit operating constraints, or natural variations in output from hydro, wind and solar resources.

- **Scheduling inaccuracies.** Forward market energy schedules may not be achievable given the generators' unit commitment constraints (start-up time, shut-down time, minimum up time, minimum down time, and ramp rate). They could fall short on some of these constraints and rely on the real-time market to make up the difference.

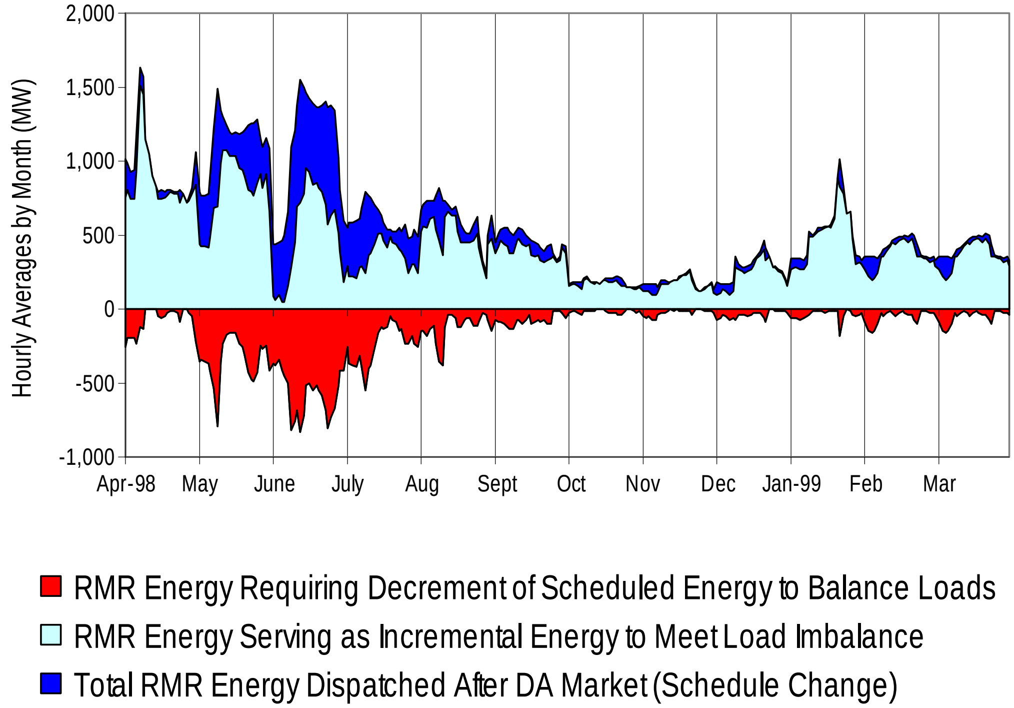
- **Intentional deviations from generation schedules.** Under current ISO protocols, suppliers may intentionally engage in *uninstructed deviations* from generation schedules for several reasons. Suppliers may run units above their generation schedules to earn additional profits when real-time prices exceed the units' variable operating costs. Even if real-time prices are lower than a unit's operating costs, units without generation scheduled may be kept on at minimum operating levels due to the costs and constraints associated with shutting down and restarting. The "self-dispatch" feature of California's energy markets, combined with the frequent under-scheduling of load mentioned above, facilitates strategic bidding at high prices, or withholding of capacity, by suppliers who may seek to exercise market power. The reason is that suppliers who intentionally engage in uninstructed deviations suffer reduced opportunity costs in those instances when some of their capacity fails to clear in any of the energy or A/S markets.

4.3.3 Reliability Must Run (RMR) Generation

Under current protocols, the ISO dispatches additional generation from RMR units after the day-ahead market closes, when it needs such generation to meet local reliability criteria. Since no load is scheduled corresponding to the RMR energy dispatched, this generation could have the effect of increasing or decreasing whatever net load imbalance exists due to other factors. Since RMR generation is, in effect, *must-take* energy in the real-time market, other generating units must be decremented if the RMR generation creates or exacerbates an excess of supply. Of course, when a shortage of supply already exists due to other factors, RMR energy serves to reduce this imbalance.

As shown in Figure 4-7, a significant amount of RMR energy is dispatched into the real-time market during all hours and months. Such RMR generation, against which no load is scheduled, has averaged over 500 MW per hour, or about 2.3 percent of total system loads. During the first three months of operation, RMR energy typically exacerbated over-generation that occurred due to other factors. In contrast, since July virtually all RMR energy has offset significant negative imbalances due to other factors, and has required that other units be decremented only during off-peak winter hours. During peak hours, when system imbalances are highest, RMR generation dispatched by the ISO after the day-ahead market has offset about 500 MW of the real-time load imbalances on average.

Figure 4-7. Impact of Reliability Must Run Generation on Energy Imbalance



4.3.4 Inaccurate Accounting for Transmission Losses

ISO protocols and scheduling software require that SCs submit schedules in which generation and loads are balanced *after accounting for transmission losses*. Transmission losses are accounted for by applying Generation Meter Multipliers (GMMs) to scheduled generation. At present, day-ahead and hour-ahead energy schedules do not account for transmission losses due to a limitation in the PX scheduling software (although GMMs are applied to metered PX generation during the ISO's settlement process). This creates a systematic scheduling imbalance – a deficit of actual supply available to meet scheduled load. To date, a significant portion of this imbalance has been offset by additional unscheduled RMR generation. In the past year such unscheduled RMR generation has averaged over 2 percent of total system loads.

4.4 Supply of Imbalance Energy

The demand for real-time imbalance energy is met by four major categories of resources, as depicted in Figure 4-8. These resources are:

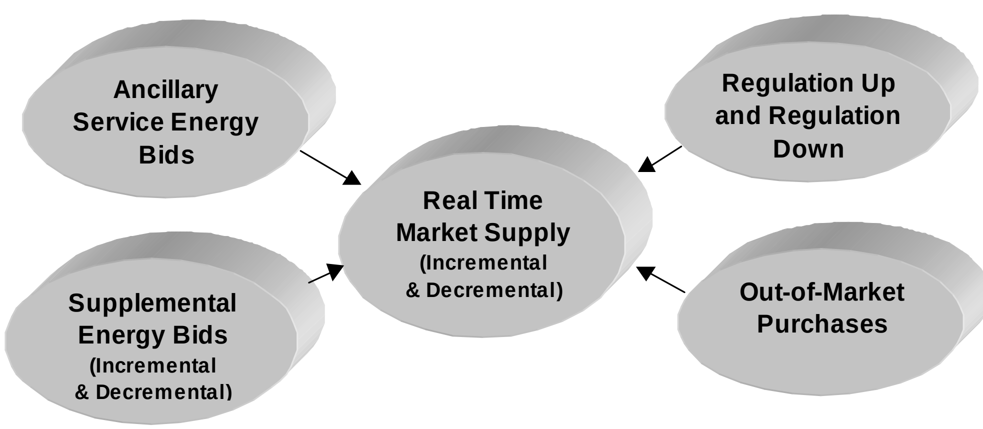
- **Ancillary Services Energy Bids**. Units providing Spinning, Non-spinning and Replacement Reserve capacity submit separate energy bids (along with the capacity bids they submit for the A/S auctions) for any energy they might be called upon to provide in real time. These energy bids may be as high as the ISO's $250 cap on real-time energy prices.

- **Supplemental Energy Bids**. Generators can submit supplemental energy bids for real-time energy up until 45 minutes before the beginning of each operating hour. Supplemental energy bids can be for *incremental energy,* which if called upon would increase their scheduled generation, or for *decremental energy,* which represents energy already included in their generation schedules which they are willing to buy back and not generate if the ISO needs to reduce supply.

- **Regulation**. Although Regulation Reserve is different from the other A/S by virtue of its use in conjunction with automatic generation control (AGC) to balance random fluctuations in system balance, during the first summer of operation the ISO frequently utilized Upward Regulation as a significant source of short-term real-time energy.

- **Out-of-Market Purchases**. During the first summer of operation, the ISO purchased energy "out-of-market" in order to meet actual or expected loads. This occasionally required pre-arranged bilateral purchases of blocks of energy to be imported from adjacent control areas at agreed upon prices. In other instances, it involved calling resources within the ISO control area to provide additional energy that was not bid into the real-time market. In these cases, resource owners received the hourly ex-post price.

Figure 4-8. Sources of Real-time Imbalance Energy



4.4.1 The BEEP Stack

Real-time energy bids from the units providing A/S are combined with supplemental energy bids for incremental and decremental energy to form the “BEEP Stack,” which represents the merit order ranking of bids based on price. Figures 4-9 and 4-10 depict the BEEP stack during a sample of typical off-peak and peak hours. The BEEP stack is typically composed of two distinct categories of real-time supply. One category is bid at prices which increase gradually in a way that closely mirrors the short-term variable operating costs of less efficient units in California, while the other category is bid at much higher prices, at or near the ISO’s $250 cap for real-time energy.

As shown in Figure 4-9, during off-peak hours, over 5,000 MWh of real-time energy are often offered at prices up to about $40/MWh to $60/MWh. This price range corresponds closely to the variable operating costs of the less efficient plants in California, and suggests that during off-peak hours many bids are based on marginal operating costs. Beyond this point in the BEEP stack, however, bid prices for additional supply jump significantly above marginal operating costs for even the least efficient generating units, to prices at or near the $250 price cap for real-time energy.

During peak hours, a much smaller amount of energy is typically offered at prices below the $50 to $60 range, so that the real-time supply curve is effectively “shifted” to the left relative to off-peak hours, as depicted in Figure 4-10. Thus, during peak hours the $250 price cap may be hit if demand for imbalance energy reaches levels of 2,000 to 4,000 MWh. Additional discussion and analysis of real-time prices at different levels of demand is provided in Section 4.5 of this chapter.

Figure 4-11 depicts the month-to-month variation in the amount of real-time energy bid at the ISO’s $250 price cap during peak hours, for each of the A/S and for supplemental energy. For units providing A/S capacity, the portion of energy bids at the ISO’s $250 price cap during peak operating hours (hours 7 to 22) increased significantly over the first few months of operation. During the peak months of July to September, nearly 20 percent of the energy bids for units providing Non-spinning and Replacement Reserves were at or just below the $250 price cap. During this period, as much as 40 percent of the energy bids by units providing Spinning Reserve were at or just below the $250 price cap. At the same time, only a very small portion of supplemental energy bids were submitted at the $250 price cap. As discussed in the following section, most of the demand for imbalance energy has been met through supplemental energy rather than A/S energy bids, due to the more competitive pricing of supplemental energy bids.

Figure 4-9. Real-time Energy Bids (BEEP Stack) for Typical Off-Peak Hours

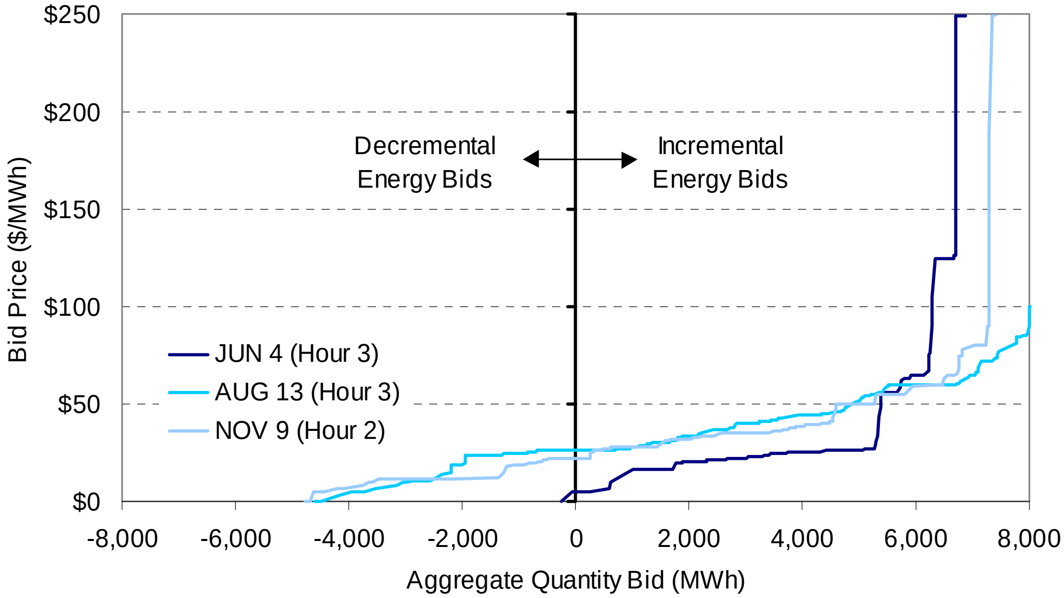


Figure 4-10. Real-time Energy Bids (BEEP Stack) for Typical Peak Hours

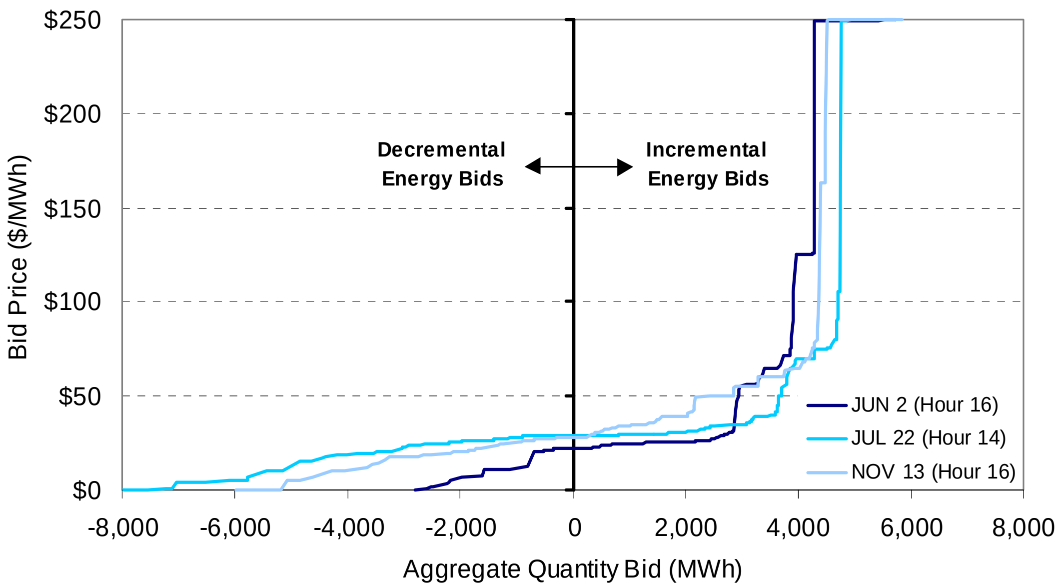
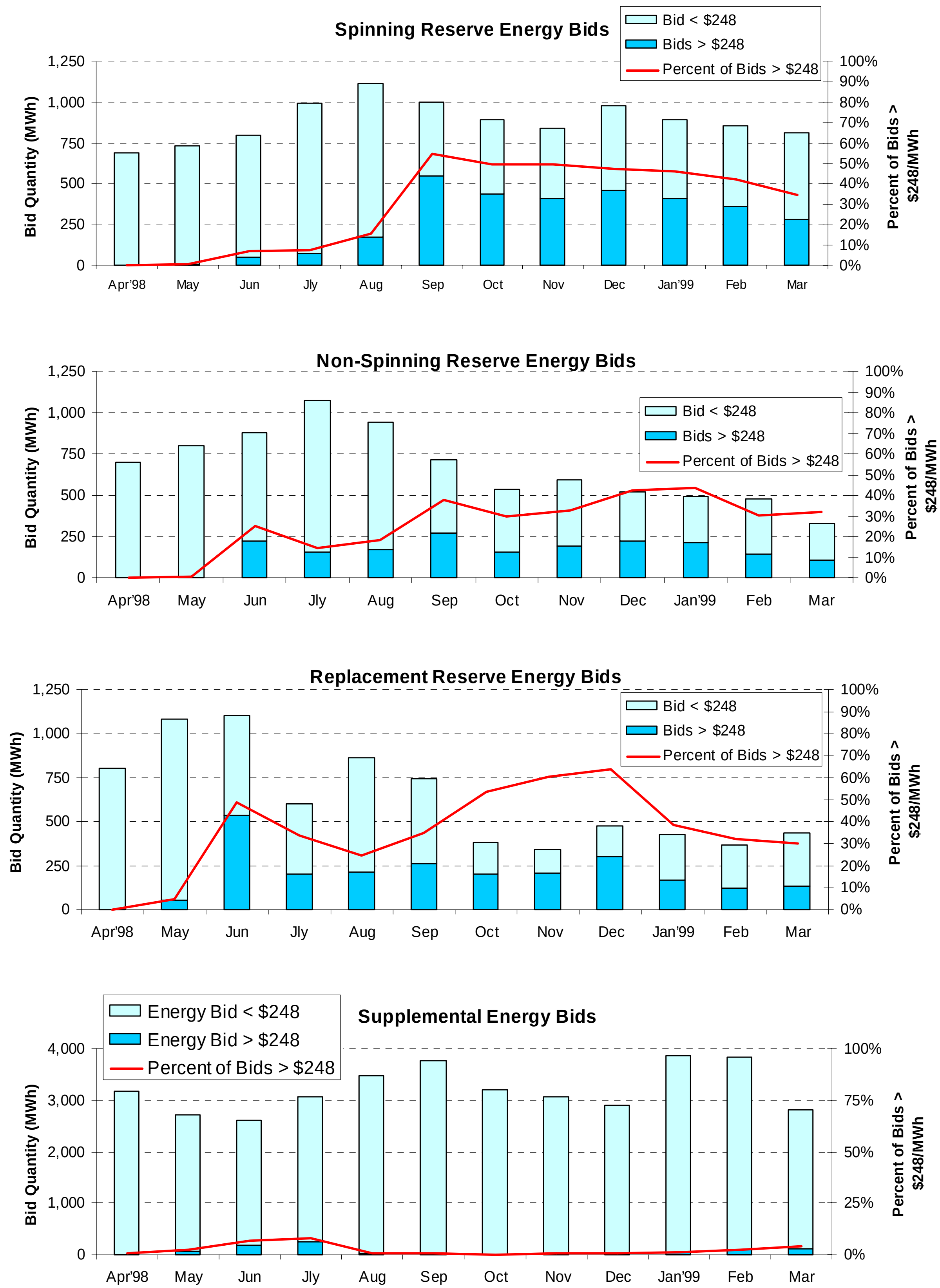


Figure 4-11. Real-time Energy Bids at $250 Price Cap (Peak Hours 7-22)



4.4.2 Real Time Energy Dispatched

As shown in Figure 4-12 and 4-13, supplemental energy bids have consistently provided the major source of real-time imbalance energy dispatched by the ISO since the market opened. During the first year of operation, over 80 percent of real-time imbalance energy has been provided by supplemental energy bids, with imported supplemental energy accounting for 38 percent.

The ISO's reliance on supplemental energy bids to supply the bulk of real-time imbalance energy reflects the bid prices of the different sources of supply. Only occasionally has the ISO skipped some of the energy bids submitted by units providing A/S in order to maintain adequate levels of Operating Reserves, or for other operational reasons. A more detailed discussion and analysis of merit order dispatch of energy bids in the real-time market is provided below.

Figure 4-12. Sources of Real-time Energy Dispatched by ISO

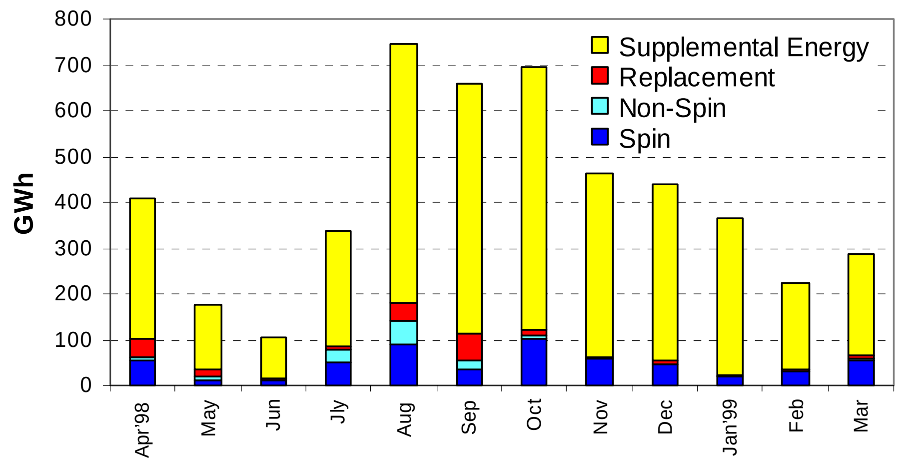
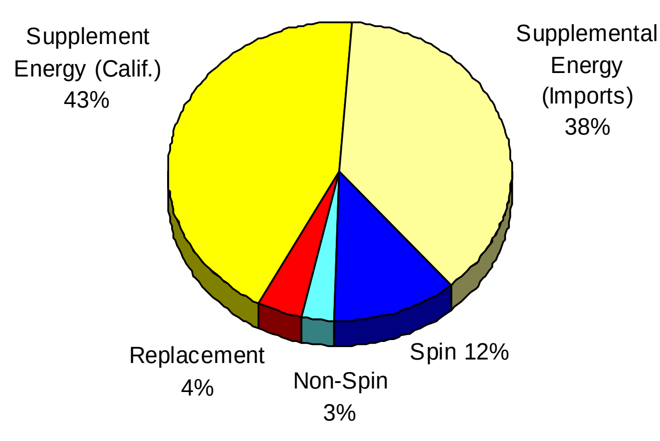


Figure 4-13. Sources of Real-time Energy Dispatched by ISO (12 months)



4.4.3 Merit Order Dispatch of Real Time Energy Bids

Supplemental energy bids and A/S energy bids are dispatched in merit order based on bid price, with assistance from the BEEP software, subject to a variety of system and unit constraints taken into consideration by the ISO Generation Dispatcher. Bids in the BEEP merit order stack may be skipped due to system conditions and resource constraints that exist or are expected to exist during the operating hour, such as:

- **Operating Reserve (O/R) deficiencies**. When demand for energy is high and the ISO determines that a deficit in O/R may develop, it will skip energy bids from O/R (Spinning and Non-spinning Reserve) capacity in order to maintain proper O/R levels.
- **Unplanned forced outages**. A unit may become unexpectedly unavailable after its supplemental energy bids are submitted.
- **Decline of dispatch orders**. SCs sometimes decline to follow an instruction to dispatch a supplemental energy bid accepted by the ISO.
- **Inability to provide energy due to uninstructed deviations**. In some cases, units may be unable to provide real-time energy because they have deviated from their accepted operating schedules. This may tend to occur precisely when real-time energy prices are already high, since generators receive the imbalance price for uninstructed deviations.

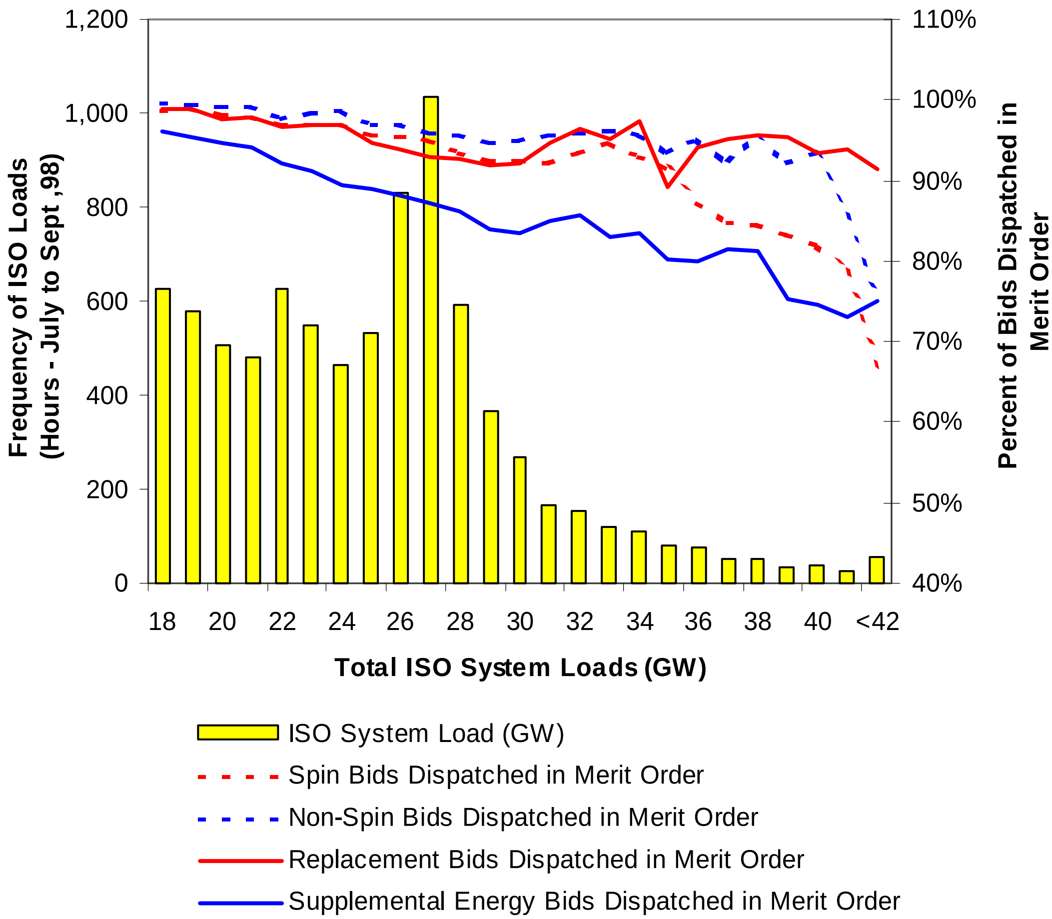
While skipping of bids in the BEEP stack is sometimes inevitable due to special reliability and operational constraints, skipping bids can have significant impacts on the prices of real-time energy and A/S capacity, and should therefore be minimized.

- **Real-time energy prices.** When a bid is skipped, the BEEP interval price is raised (or lowered in the case of a decrement) to the bid of the next feasible resource selected. During periods of peak loads or significant load imbalances, this can have a significant effect on market prices, due to the steep slope of the supply curve under such conditions. At these times the decision to skip over bids in the BEEP stack can represent a major trade-off between reliability and real-time energy costs.

- **Ancillary service capacity prices.** The expected revenues earned by A/S units whose energy bids are accepted in the real-time market represent a key incentive for suppliers to provide A/S, rather than only providing energy through the forward or supplemental energy markets. In fact, the *probability* of being dispatched to provide real-time energy, based on a bid price lower than the market clearing price, is likely to affect prices at which suppliers bid their A/S capacity. For instance, if energy bids submitted by O/R units are frequently skipped over for reasons of system reliability, this will likely translate into higher market clearing prices for A/S capacity. Likewise, the more likely O/R units are called to provide real-time energy, the lower their O/R capacity prices should be.

Figure 4-14 shows the portion of real-time incremental energy bids (supplemental and A/S) submitted at prices below the *ex post* real-time price which were accepted and dispatched by the ISO at different levels of total system load. The difference between the amount of bids submitted at prices below the *ex post* price (100 percent on the graph) and the amount of bids dispatched by the ISO suggests the magnitude market impacts associated with the skipping of bids in the BEEP stack. Included in the market impacts is the effect on suppliers' perceptions of the probability that they may be called to provide real-time energy if the *ex post* price reaches their bid price.[3] The data have been categorized based on the level of total ISO system loads, since this is the single factor most likely to affect both the real-time price and the probability that bids may be skipped due to various reliability and operational constraints. Some of the specifics for each category of energy bids are discussed below.

Figure 4-14. Percent of Bids Dispatched in Merit Order, by System Load



[3] Data in Figure 4-14 are based on real-time energy bids at prices equal to or lower than the final *ex post* imbalance price for each hour, compared to actual real-time energy provided by each unit on an hour-by-hour basis. As previously noted, a bid equal to or lower than the final *ex post* price may fail to be dispatched for operational or reliability reasons, including cases in which units are unable or decline to provide energy for which a bid had been submitted to the ISO. Since the reasons for skipping individual bids are not formally tracked or categorized for all hours, this analysis does not attempt to describe or quantify the reasons why particular bids were skipped.

- **Replacement Reserve.** The portion of energy bids below the real-time *ex post* price that were dispatched for real-time energy ranged from 90 to 100 percent for various levels of system load. This reflects the fact that bids from Replacement Reserve units would rarely be skipped for reliability reasons. It should be noted that such units were sometimes called by the ISO but were unable to provide energy, for a variety of reasons. During peak periods, this was sometimes due to the fact that units were already operating above their scheduled generation level and therefore could not ramp up to provide additional energy. The ISO is implementing a number of steps as part of its A/S redesign to make the supply of energy from A/S units more reliable, as discussed in Chapter 3.

- **Spinning Reserve.** The portion of energy bids below the real-time *ex p*ost price that were dispatched for real-time energy ranged from 90 to 100 percent for load levels up to 35,000 MW, but dropped at higher load levels, falling to about 70 percent for hours where loads exceeded 42,000 MW. This pattern may reflect the fact that bids from units providing Spinning Reserve are those most likely to be skipped for reliability reasons, particularly at very high load levels. In other cases, bids may also tend to be skipped at high load levels due to their inability to provide energy as a result of uninstructed deviations, which tends to occur most when real-time energy prices are highest.

- **Non-spinning Reserve**. The portion of energy bids below the real-time *ex post* price that were dispatched for real-time energy ranged from 90 to 100 percent for loads levels up to 40,000 MW, but dropped to about 80 percent during the hours that loads exceeded 42,000 MW. This trend reflects the fact that bids from units providing Non-spinning Reserve, like those providing Spinning Reserve, are also likely to be skipped for reliability reasons, particularly at very high load levels.

- **Supplemental Energy**. The portion of bids submitted below the real-time *ex post* price that were dispatched by the ISO declined gradually from nearly 100 percent at very low levels of system load, to about 70 percent at very high loads above 40,000 MW. Unlike energy bids from A/S units, supplemental energy bids do not represent a firm commitment of energy at the bid price. In many cases, units submitting supplemental energy bids may respond to high real-time prices by running uninstructed in real-time, rather than waiting to be called by the ISO. Supplemental energy bids from outside the service territory may also be unavailable at such times due to congestion.

4.5 Real-time Energy Prices

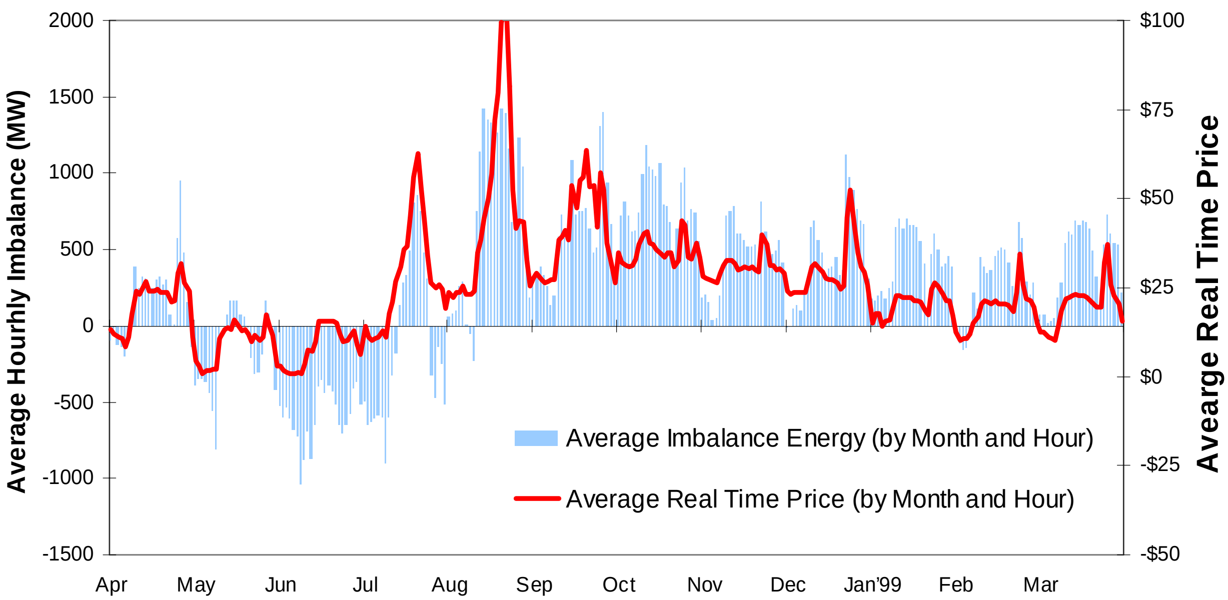
Despite a high level of volatility, real-time energy prices have closely tracked the demand for imbalances over the ISO's first year of operation, as depicted in Figure 4-15.

- From April to June, an abundance of hydro resources in northern California and low loads led to over-generation conditions for much of the period. As a consequence, prices were very moderate and the ISO frequently had to decrement generation. Hourly decrements of 1,000 MWh or more frequently resulted in real-time prices at or near $0/MWh. Prices typically increased linearly at higher levels of real-time demand up to 2,500 MWh, with prices rarely exceeding $50/MWh.

- During the peak summer season (July to September) the ISO often had to make very large real-time energy adjustments to maintain system balance. Prices were typically much higher than during the spring months, for the same level of incremental energy. This upward shift in prices was largely due to higher system loads requiring dispatch of higher cost thermal units. Prices typically rose in linearly with the amount of imbalance energy needed up to about 1,000 MWh, but increased exponentially when demand for imbalance energy went to higher levels. During this period a series of price spikes occurred, with the ISO's price cap for-real time energy being hit in more than 40 hours. A more detailed discussion of price spikes is provided in the next section.

- During the fall season (October to December) real-time energy imbalances moderated considerably compared to the summer months, and generally ranged between -1,000 MWh to +2,000 MWh. Prices were similar to those experienced during the summer up to 1,000 MWh of incremental energy demand, but at levels of real-time demand above 1,000 MWh prices were generally much lower than during the summer months. This result is mainly due to lower system load conditions, which allowed real-time demand to be met from lower cost generation. Although the relationship between real-time prices and energy adjustments was highly linear, prices in the fall months were on average significantly higher than in spring. This result is largely attributable to the fact that hydro resources are scarce in the fall months and prices are predominately set by higher cost thermal generation.

Figure 4-15. Real-time Energy Imbalances and Total System Loads



4.5.1 Real Time Price Spikes During the Summer Peak Season

The positive correlation between prices and the demand for imbalance energy is consistent with a market which is well-functioning under most conditions. During the peak system load conditions from July to September, however, extreme price increases have frequently occurred which are consistent with the exercise of market power that is created when tight supply conditions exist.

As shown in Figure 4-16, average prices in the PX and real-time market from July to September were virtually equal, and rose linearly up to total system loads of about 32,000 MW. Above this point, prices in both markets rose exponentially in relation to total system loads. Then, above 42,000 MW the average prices in the real-time market rose dramatically above average prices in the PX, with average real-time prices about $100 higher than average PX prices during system peak hours. As shown in Figure 4-17, the frequency of hitting the $250 price cap in the real-time market also rose exponentially as total system loads increased.

The $250 price cap for real-time energy was reached in over 48 hours in the first year of operation.[4] All but five of these hours were during the peak summer period from July 7 to September 3, 1998. During this period, price spikes typically occurred in the real-time market when high system loads coincided with relatively high demand for imbalance energy, as shown in Figure 4-18.[5] During these hours, load imbalances ranged from 1,000 to 3,800 MWh, with average demand for imbalance energy of over 2,400 MWh. Total system loads when the price cap was hit always exceeded 35,000 MW, and the cap was reached in 90 percent of the hours when system loads exceeded 40,000 MWh. As depicted in Figure 4-18, when total system loads were lower, imbalances in the range of 1,000 to 4,000 MWh occurred frequently without prices hitting the cap.

[4] The final *ex post* price, which represents the average of six 10-minute interval prices, was $250 only 45 times. The price cap was reached during other hours, but not for all six 10-minute intervals. Statistics in the report are based on 48 hours in which the final average *ex post* price was $249 or higher.

[5] During one hour, a price of $250 was reached despite real-time imbalance energy demand of only 360 MWh. This hour was preceded by an hour in which the price cap was hit due to demand for over 1500 MWh of imbalance energy. This event illustrates how the real-time price can be very high when units bidding very high prices are accepted for the hour following a high-demand hour, but may not ultimately be needed for the second hour.

Figure 4-16. Total System Loads and Energy Prices (July–Sept. 1998)

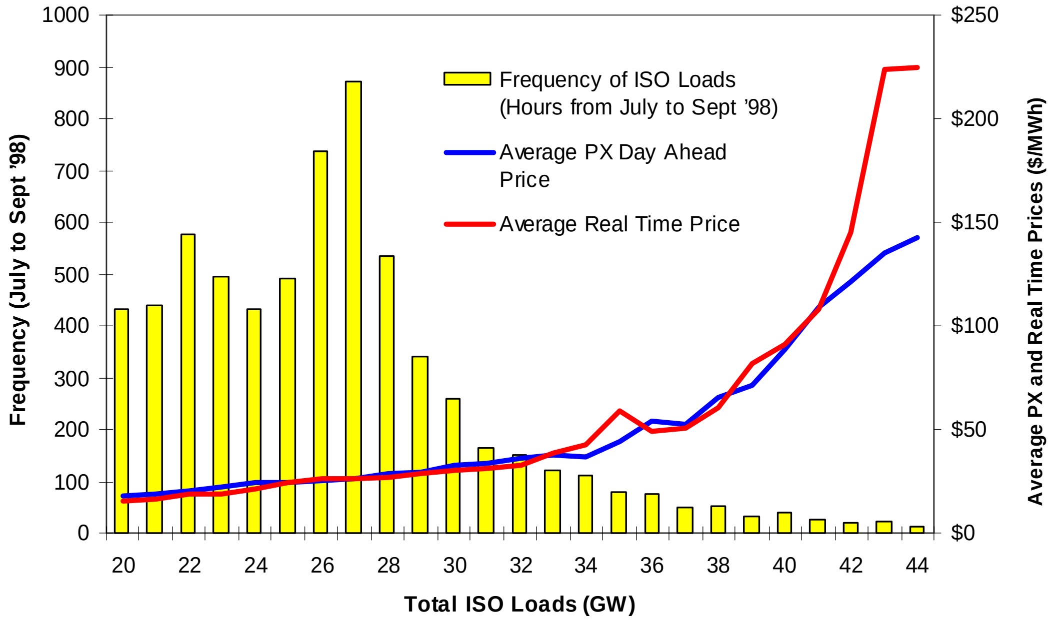


Figure 4-17. Total System Loads and Real-time Price Spikes (July–Sept. 1998)

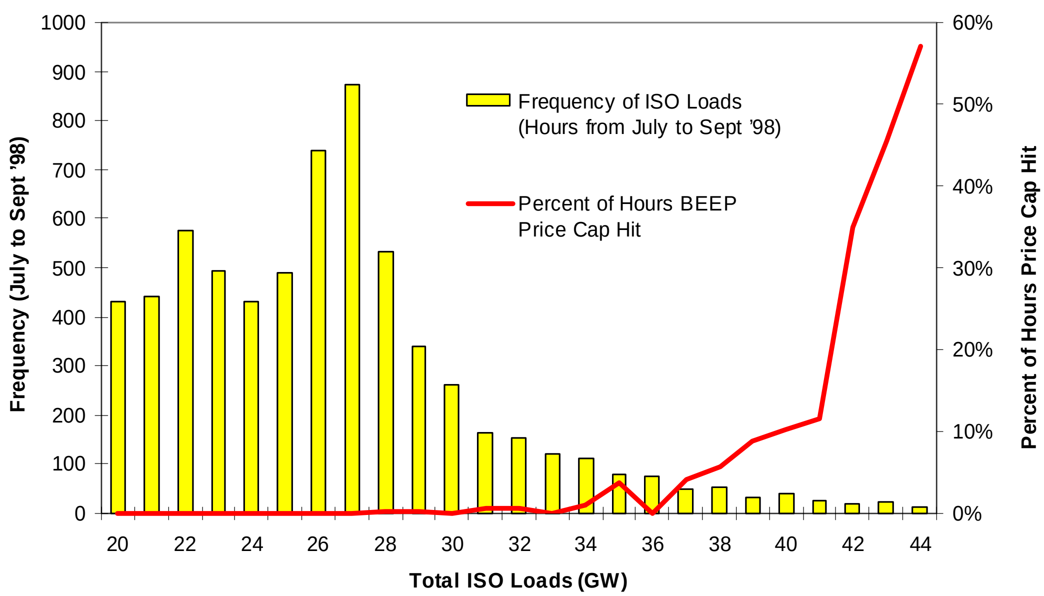
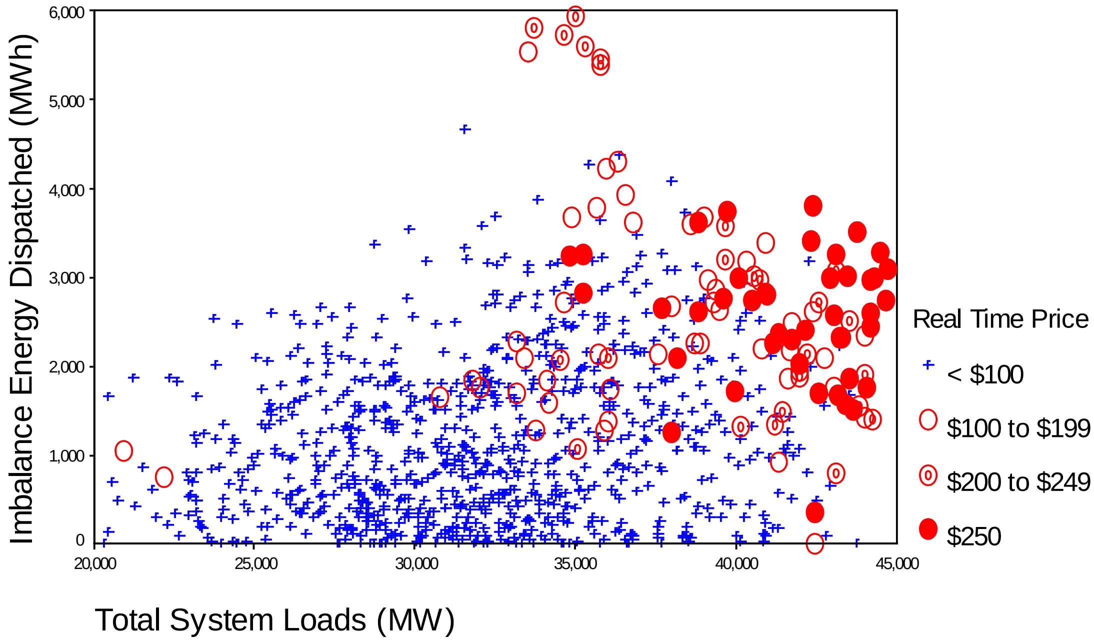


Figure 4-18. Total System Loads and Real-time Imbalances During Real-time Price Spikes (July–Sept. 1998)



4.5.2 Real Time Price Spikes During the Winter and Spring Seasons

The price cap was hit during four hours in December during the week before Christmas. These price spikes coincided with very cold weather in northern California and the Pacific Northwest, which created severe gas shortages. With very high prices for gas and in some cases physical shortages, the supply of thermal generation to the real-time market was significantly reduced and bids that were provided were at very high prices.

The price cap was hit for one hour on March 15, due to a combination of relatively high demand for imbalance energy during the winter evening peak hours, and transmission constraints on COI which prevented supplemental energy import bids from being available.