Do We Need to Implement Multi-Interval Real-Time Markets?

Darryl R. Biggar[a] *and Mohammad Reza Hesamzadeh*[b]

ABSTRACT

Many market-based power systems have implemented a form of 'look-ahead dispatch' which simultaneously solves for the optimal dispatch and prices over several intervals into the future. A few papers have pointed out that the dispatch outcomes which emerge from look-ahead dispatch may not be time consistent. We emphasise that this time inconsistency is not inherent in look-ahead dispatch but is a consequence of the assumption of linear cost and utility functions, which is arguably a special case. Various augmentations to the dispatch process to resolve the time inconsistency problem have been proposed, but these augmentations suffer from the drawback that they do not allow the power system to efficiently adjust to new information. We query whether it is necessary to implement multi-interval real-time markets. We show how under certain assumptions, a sequence of one-shot dispatch processes will achieve the efficient outcome.

Keywords: Look-ahead dispatch, Pricing, Real-time market, Dual degeneracy

https://doi.org/10.5547/01956574.43.2.dbig

1. INTRODUCTION

Around the world, as the penetration of intermittent generation and distributed energy resources increases, concerns are being raised about the ability of wholesale electricity spot markets to deliver adequate reliability and efficient dispatch outcomes.[1] In particular, faced with the risk of increasingly large swings in the supply/demand balance there are growing concerns about the ability of 'one-shot' economic dispatch processes to efficiently dispatch the available generation and load resources in the face of ramp rate constraints and other short-term inter-temporal constraints.[2]

For example, the famous California 'duck curve'[3] highlights how, with increasing penetration of solar in a power system, conventional sources of generation will likely be required to ramp up quickly as solar generation declines in the late afternoon. But some generation resources may be limited in how quickly they can increase their output. If these ramp rate constraints are forecast to bind in the future, it may be efficient to pre-position some resources in the power system ex ante to

1. See, for example, Moarefdoost et al. (2016); Herrero et al. (2018).

2. See Hogan (2020). In some markets the response has been to limit the rate at which intermittent generators are allowed to increase or decrease their output Hittinger et al. (2014). The issues discussed here could, in principle arise in any real-time wholesale electricity market, with or without locational marginal pricing, whether in the U.S., Europe, Australia or New Zealand, particularly markets with a high level of penetration of intermittent generation. In practice, look-ahead dispatch has primarily been an issue for consideration in the U.S. (see the summary in the online appendix), although with increasing penetration of intermittent generation, related concerns have arisen in Australia, as noted in footnote 5.

3. See Wikipedia: Duck Curve.

a Australian Competition and Consumer Commission, Australia. E-mail: darryl.biggar@accc.gov.au.

b Corresponding author. KTH Royal Institute of Technology, Sweden. E-mail: mrhesamzadeh@ee.kth.se.

reduce the cost of adjustment to a new supply/demand balance ex post.[4] In fact, such pre-positioning may be essential if power system reliability is to be maintained. This has led some markets to consider introducing separate procurement of ramping capability or flexible operating reserves.[5]

But how should such 'pre-positioning' be achieved? Many authors have argued that, in the face of binding ramp rate constraints, the wholesale market should not compute the efficient dispatch outcome for a single period at a time (which we will refer to as 'one-shot' dispatch). Instead, the dispatch process should compute the optimal dispatch over several dispatch intervals at a time, taking into account potential short-term inter-temporal constraints such as ramp rate constraints or short-term energy limits.[6] In fact, as set out in the online appendix, many wholesale power markets have already implemented such practices. These policies are known by various names in the literature including "dynamic economic dispatch"[7], "multi-interval real-time markets" (MIRTM)[8] or "look-ahead dispatch" (LAD)[9]. Hua et al. (2019) explain the justification for MIRTM as follows:

> MIRTM allows for more efficient dispatch of generation to meet projected system conditions by pre-positioning resources to cope, for example, with large ramps in net load. In particular, a satisfactory short-term forecast for renewable generation can allow LAD to meet high ramping requirements due to generator variability in a reliable and economic manner. Consequently, the New York Independent System Operator (NYISO) and California ISO (CAISO) have already implemented MIRTM, while the Electricity Reliability Council of Texas (ERCOT) has proposed the approach.

In practice, in those jurisdictions which have implemented look-ahead dispatch, the dispatch process forecasts the market state (demand, supply, and network conditions) several dispatch intervals into the future and then computes the optimal dispatch and prices for the current and forecast market states. The dispatch and prices from the first interval in the sequence are used for dispatch and settlement purposes. All the future prices are merely advisory. In this paper we will refer to the first price in the sequence of prices from the dispatch process as the spot price. All the other prices (which, to repeat, are merely advisory) we will refer to as forecast prices.

However, as conventionally implemented, such look-ahead dispatch processes suffer from two problems:

1. First, as a few authors have pointed out, the prices that emerge in the look-ahead dispatch may not be 'time consistent': Even if the forecasts of the future market state are

4. A 'dispatch' (that is, a set of production and consumption instructions for all controllable generators and loads) is efficient if it maximises the total economic welfare subject to the physical constraints of the power system. In the context of inter-temporal constraints (such as ramp rate constraints) the efficient dispatch must maximise the total economic welfare over a period of time subject to the physical constraints of the power system.

5. See, for example, Daraeepour et al. (2019); Cornelius et al. (2018). The system operator in Australia, the Australian Energy Market Operator, has argued for the introduction of 'ahead' markets on the basis that this will allow for "pre-positioning the system to deal with expected system conditions ... to help improve system efficiency and reliability". AEMO (2018). Many ISOs in the U.S. are considering the procurement of 'ramping products' (CAISO, 2015; Parker, 2015; Schiro, 2018) to ensure that sufficient ramping capability will be available in the near future. A related concept is the procurement of operating reserve over a time period that extends beyond the current dispatch interval, which is only necessary if inter-temporal constraints are forecast to bind, as currently being considered in the Australian NEM (AEMC, 2020).

6. See, for example, Xie et al. (2013). According to Ela and O'Malley (2015), "increased variability and uncertainty can cause the need for advanced scheduling strategies in order to improve or maintain the level of reliability, while obeying system constraints at least production cost".

7. Ross and Kim (1980), Han et al. (2001), Xia and Elaiw (2010), Raithel et al. (1981).

8. Hua et al. (2019), Zhao et al. (2018), Ela and O'Malley (2015).

9. Wu et al. (2014), Xie et al. (2011).

accurate, the profile of spot and forecast prices which are forecast at the outset may not be the same as the sequence of spot prices that arise as we step forward in time. As a consequence, the generation and load resources may not have an incentive to comply with the dispatch instructions. This is a potentially serious problem as resources may not respond to dispatch instructions (including pre-positioning), potentially undermining system security and reliability.

2. Second, the conventional look-ahead dispatch task only forecasts a single market state in each near-term future dispatch interval. This approach does not lead to efficient dispatch outcomes (including pre-positioning) in the real-world context in which new information can arrive over time about market demand, supply or network conditions in the future. If, for example, there is uncertainty about future wind generation conditions, an efficient dispatch in the face of ramp-rate constraints should take that into account. This is not possible in the conventional formulation of look-ahead dispatch.

In this paper we explore the significance of these problems and suggest a solution. We conclude that, while the first problem has attracted attention in the literature, it only arises in what is arguably a special case (the case of linear cost and utility functions). The second problem, however, has no practical solution. But, we point out that, under reasonable assumptions, implementing look-ahead dispatch is not necessary to achieve efficient dispatch (including pre-positioning) of the power system. Instead, a single one-shot dispatch process (with no requirement to forecast future market states) can achieve the efficient outcome, provided market participants can effectively forecast the future prices in different scenarios.

Ramp-rate constraints and their impact on electricity market design have long been discussed in the electricity market literature. For example, Han et al. (2001) model ramp rate constraints in the formulation of what they call 'dynamic economic dispatch'. Several papers have explored the application of look-ahead dispatch in the context of specific wholesale markets. For example, Yu et al. (2005) discuss the implementation of a multi-interval economic dispatch model for the Ontario real-time electricity market; Price and Rothleder (2011) discuss extending the dispatch horizon and look-ahead unit commitment in California's energy market; Xu and Howard (2013) consider a proposal for look ahead security-constrained economic dispatch in ERCOT.

Several papers assess the theoretical benefits of look-ahead dispatch. Xie and Ilic (2008) discuss the potential benefits of applying model predictive control in the context of a multi-interval real-time economic dispatch problem. Using a model that includes stored energy resources, Zhu and Chen (2012) show that the cost of energy dispatch and the cost of procuring regulation reserve are reduced in their look-ahead multi-interval formulation as compared to the traditional single-interval formulation. Thatte et al. (2014) highlight the cost-saving benefits arising from pre-positioning some generators in response to changes in future dispatch intervals.

A few papers have highlighted some theoretical issues. For example, Choi and Xie (2016) perturb the Karush-Kuhn-Tucker conditions of the look-ahead economic dispatch to calculate the impact of data corruption on look-ahead economic dispatch results, concluding that the look-ahead dispatch is more vulnerable to data corruption than the static one-shot economic dispatch. Wang and Hobbs (2013) compare the look-ahead real-time economic dispatch model under deterministic and stochastic settings.

For the purposes of this paper, we are particularly interested in the strand of the literature that argues that look-ahead dispatch or multi-interval real-time markets gives rise to a *time inconsistency problem*—that the sequence of 'ex post' prices arising from the first dispatch interval in each of sequential runs of the look-ahead dispatch process may not coincide with the sequence of prices forecast by the dispatch process at the outset—even if nothing changes in the power system. These issues are raised by Hogan (2012); Ela and O'Malley (2015); Hua et al. (2019); Guo et al. (2019).

Such a time inconsistency would be a serious problem. In normal circumstances, wholesale power markets rely on market participants *voluntarily* complying with their dispatch instructions.[10] If participants expect to face a different sequence of prices in the future than originally forecast, they will not necessarily have an incentive to comply with the dispatch instructions at the outset, undermining the dispatch process, leading to inefficient outcomes and threats to system security. To address this problem Hogan (2016) and Hua et al. (2019) propose extensions or augmentations to the look-ahead dispatch problem. These extensions tie future or subsequent price outcomes to prices that were determined at an earlier time. These mechanisms resolve the time inconsistency problem in the special case of perfect foresight.

The contribution of this paper is as follows:

1. First, we clarify that the problem of time inconsistency is not a fundamental problem of look-ahead dispatch, *per se*. Rather it is a consequence of the assumption that the dispatch task is linear.[11] We suggest that the assumption that the dispatch task is linear is a special case. When the generators' supply curves are continuous and upward sloping (or, if demand curves are continuous and downward sloping), the time inconsistency problem disappears. In such cases, the mechanisms proposed in Hua et al. (2019) and Hogan (2016) are not necessary to achieve efficient dispatch.
2. Second, we point out that the look-ahead dispatch task, as formulated in theory and practice, does not allow for efficient handling of new information as it arrives over time and so does not yield efficient forecasts of dispatch outcomes (including pre-positioning). The same critique applied to the solutions to the time-inconsistency problem proposed in the literature (by Hua et al. (2019) etc.)—those solutions do not result in efficient dispatch and pricing outcomes in a context in which the power system must adjust to changing demand, supply and network conditions all the time.
3. Third, we emphasise that, under reasonable assumptions, implementing a multi-interval real-time market is not necessary to achieve efficient dispatch. If we assume that all market participants are price takers, with continuous upward-sloping supply curves and/or continuous downward-sloping demand curves, a simple, one-shot bid-based security-constrained economic dispatch will achieve the overall efficient outcome. Provided the market participants have sufficient flexibility in the bids and offers they can make to the central dispatch process, and have accurate forecasts of future prices in difference scenarios, all of the inter-temporal concerns can be delegated to the market participants themselves—including responsibility for 'pre-positioning' in advance of binding ramp rate constraints if necessary. We show that the resulting sequence of one-shot dispatch outcomes is efficient overall.

10. That is, for market participants voluntarily participating in the wholesale market process, the dispatch instructions received must maximise their profit given the price they are paid. If it did not, they would have an incentive to generate at levels inconsistent with their dispatch instructions and/or to distort their offers to change their dispatch instructions.

11. If generators' cost functions and the utility function of loads are both assumed to be linear (or piece-wise linear) there may arise a situation where the spot price (i.e., the marginal value on the energy balance constraint) is not uniquely determined. This is the problem of *dual degeneracy*, as mentioned in Hua et al. (2019); Ela and O'Malley (2015); Hogan (2012). Moreover, the range of optimal prices tends to increase over time. In this context, even with perfect foresight, there is no guarantee that later runs of the dispatch process will choose prices from within the set of prices which were (forecast) optimal in earlier runs.

This paper has two substantive parts. Section 2 examines the time-inconsistency problem under look-ahead dispatch, highlighting the problem of dual degeneracy. We show that under the assumption of strict convexity the time inconsistency problem disappears. We also briefly review possible mechanisms for resolving the time inconsistency problem under perfect foresight, and point out that these models are unsatisfactory in that they do not correctly handle the new information that arrives over time. Section 3 demonstrates that implementing look-ahead dispatch mechanisms in the real-time market are not necessary for efficient dispatch: under certain assumptions, efficient dispatch outcomes can be achieved with a simple one-shot dispatch. This arises because all of the inter-temporal constraints are resolved through the profit-maximising decisions of the market participants themselves. Section 4 concludes.

2. LOOK-AHEAD DISPATCH AND TIME-INCONSISTENCY

2.1 Introduction to look-ahead dispatch

Let's start by considering a very simple model of look-ahead dispatch, with N generators (indexed by i). There is a single, perfectly inelastic load. All generators and loads are located at the same pricing node. For simplicity all generators and loads are assumed to be price-takers (i.e., have no market power).

We have tried to keep the notation as simple as possible. We will use g_i, L, and p to refer to production (MW) of generator i, consumption (MW) of inelastic load L, and spot price p ($/$MWh$), respectively. However, we must also keep track of which time period we are referring to. We will use superscripts to denote the time at which the dispatch process is run, and subscripts to denote the time in the future to which the dispatch or pricing refers. For example $g^t_{i,s}$ will refer to the dispatch of generation i at future time s, in the economic dispatch which takes place at time t. Similarly, p^t_s refers to the spot price at future time s when viewed from the dispatch which takes place at time t.

In a simple one-shot market, at some time t, all of the generators submit offers to the system operator reflecting their willingness to produce electricity at time $t+1$. In a competitive market these offers reflect each generator's cost function $c_i(\cdot)$. The cost function is assumed to be increasing and convex: $c_i'(\cdot) \geq 0$ and $c_i''(\cdot) \geq 0$. Similarly, the load forecasts the demand at time $t+1$, L^t_{t+1}. The system operator determines the optimal dispatch $(g^t_{i,t+1})_{i=1,\ldots,N}$ which minimises the total cost of dispatch, and publishes the resulting spot price p^t_{t+1}.[12]

The standard extension of this model to accommodate ramp rate constraints is as follows: We assume that in period t all of the generators submit offers reflecting their willingness to produce for periods $t+1, t+2, \ldots, T$. T is assumed to be a point in time in the future at which no inter-temporal constraints are binding.[13] Similarly, the load submits bids for consumption of electricity in each of the same periods $L^t_{t+1}, L^t_{t+2}, \ldots, L^t_T$. Finally, generators are assumed to announce the maximum rate at which they are capable of ramping up or down. The maximum rate at which generator i can change its output per dispatch interval is R_i. The dispatch process is assumed to minimise the total cost of

12. This paper focuses on the question of efficient dispatch on scales longer than the dispatch interval. We are ignoring the question of efficient dispatch on scales shorter than the dispatch interval. Specifically we are implicitly assuming that, during each dispatch interval all generators and loads ramp linearly towards their new dispatch targets.

13. The choice of T goes beyond the scope of this paper. We will merely assume that T is sufficiently far in the future that the power system has enough time to fully adjust to any changes to the supply/demand balance, with no future binding inter-temporal constraints.

producing sufficient electricity to meet demand (summed over the period up to time T), subject to any ramp rate constraints.

As discussed further below in section 2.4, we consider this formulation (as set out in equations 1–3) is flawed in that it only allows for one possible future state of the world in each future dispatch interval. In practice, in the real world, future load may depend on factors such as the ambient temperature, which cannot be forecast with certainty several dispatch intervals in advance. Similarly, the cost function of some generators (particularly intermittent generators) will depend on factors such as wind speed which is not known with certainty in advance. As discussed in 2.4 the only way to efficiently take these factors into account in the ex ante dispatch process is through a full state-contingent dispatch. In practice this is impractical. For the moment we will put this aside and follow the standard approach in which there is only one possible state of the world in each future dispatch interval.

At time t, the power system is assumed to be in some initial state. Since the only inter-temporal constraints we are handling are ramp rate constraints, the initial state of the power system is reflected in the current rate of production of the power system g_t^t.[14] Given this initial state, the dispatch process solves the look-ahead dispatch problem set out in equations 1–3. We will refer to this problem as $LAD(t,T \mid g_t^t)$.[15] The outcome of the problem is a set of forecast dispatch outcomes $(g_{t+1}^t, g_{t+2}^t, \ldots, g_T^t)$ and forecast prices $(p_{t+1}^t, p_{t+2}^t, \ldots, p_T^t)$ extending out into the near future.

$$\min_{g_{i,s}^t} \sum_{i,s=t+1,\ldots,T} c_i(g_{i,s}^t) \text{ s.t.} \tag{1}$$

$$(p_s^t): \forall s \in [t+1,T], \sum_i g_{i,s}^t = L_s^t \tag{2}$$

$$(\gamma_{i,s}^{t,U}, \gamma_{i,s}^{t,D}): \forall i,s \in [t+1,T], -R_i \le g_{i,s}^t - g_{i,s-1}^t \le R_i \tag{3}$$

Here equation 1 is the simple objective function (the total dispatch cost), which is convex by assumption. Equation 2 is the energy balance equation; equations 3 reflect the ramp rate limits. The variables in brackets are the dual variables for each constraint. In general, the optimal dispatch $(g_s^t)_{s=t+1,\ldots,T}$ and prices $(p_s^t)_{s=t+1,\ldots,T}$ are not necessarily unique.

As we have noted, the look-ahead dispatch problem $LAD(t,T)$ relies on a single forecast of future supply and demand conditions. We will say that the power system has *perfect foresight* at time t if at some later time $\hat{t}$, the forecasts of the demand and supply conditions over the remaining periods remain unchanged. Loosely speaking, we would expect that under perfect foresight, due to Bellman's Principle, a set of forecast dispatch and prices which are a solution to $LAD(t,T)$ at the outset would remain a solution to $LAD(\hat{t},T)$ as we step forward in time. This intuition is confirmed in theorem 1 below.

At time t, let's suppose that generator i faces future prices $(p_s^t)_{s=t+1,\ldots,T}$. The expected future profit of generator i when dispatched in the amount $(g_{i,s})_{s=t+1,\ldots,T}$ is the sum of the one-period profit, extending over the future time horizon:

$$\Pi_i^t(g_i, p^t) = \sum_{s \in [t+1,T]} (p_{i,s}^t g_{i,s} - c_i(g_{i,s})) \tag{4}$$

14. Here g_t^t is interpreted as a vector with one value for each generator $g_t^t = (g_{1,t}^t, g_{2,t}^t, \ldots, g_{N,t}^t)$.

15. Here, t in $LAD(t,T \mid g_t^t)$ refers to the time at which the dispatch process operates, T refers to the final time (the 'time horizon') of the dispatch process, and g_t^t reflects the current state of the power system (that is, the current rate of production of each generator). At times we will shorten this to $LAD(t,T)$ where the state of the power system g_t^t is clear from the context.

A key issue in dispatch of electric power systems is ensuring that generators have no incentive to deviate from the dispatch instructions they are given. A sequence of forecast dispatch and prices $(g_{i,s}^t, p_s^t)_{s=t+1,\ldots,T}$ will be said to be *profit maximising*[16] for generator i at time t if, given the set of spot and forecast prices, the dispatch maximises the expected future profits of the generator over the remaining timeframe, given any binding inter-temporal constraints, i.e., if and only if:

$$\Pi_i^t(g_i^t, p^t) \geq \max_{g_{i,s}} \Pi_i^t(g_i, p^t) \tag{5}$$

$$\text{s.t.} \forall s \in [t+1, T], -R_i \leq g_{i,s} - g_{i,s-1} \leq R_i \tag{6}$$

It is a well-known result that, provided each of the generators is a price-taker, the dispatch and pricing outcomes which emerge from the dispatch process $LAD(t,T)$ are profit maximising for each generator (in other words, provided the generators were actually exposed to those prices they would have no incentive not to comply with the dispatch instructions). This is stated formally in theorem 1 below.

In addition, we will say that a dispatch and pricing outcome (g_i^t, p^t) is **time consistent** if, under perfect foresight, for all subsequent times $\hat{t}$, the subsequence $(g_s^t, p_s^t)_{s=\hat{t}+1,\ldots,T}$ is a solution to $LAD(\hat{t}, T \mid g_{\hat{t}}^t)$.

In a result which is useful for the discussion that follows, theorem 1 shows that the dispatch and pricing outcomes which emerge from the dispatch process $LAD(t,T)$ are time consistent. The proof is in the online appendix of this paper.

> **Theorem 1** *If, at time t,* $(g_i^t, p^t)_{i,t+1,\ldots,T}$ *solves* $LAD(t,T)$e*, then for each generator i, the dispatch and prices* $(g_i^t, p^t)_{t+1,\ldots,T}$ *are profit maximising for generator i and, under perfect foresight, for each value of* $\hat{t}$ *in the range* $t < \hat{t} \leq T$ *the sub-sequence of dispatch and prices* $(g^t, p^t)_{\hat{t}+1,\ldots,T}$*:*
>
> *1. is a solution to* $LAD(\hat{t}, T \mid g_{\hat{t}}^t)$*; and*
> *2. is profit maximising for generator i at time* $\hat{t}$ *(for all i).*
>
> *It follows that the sequence of dispatch and prices* $(g_i^t, p^t)_{t+1,\ldots,T}$ *is time consistent.*

2.2 Exploration of time inconsistency under look-ahead dispatch

In recent years, a number of papers have emerged which suggest that the dispatch and prices which emerge from a sequence of look-ahead dispatch runs are *not* time consistent (in apparent contradiction to theorem 1). Moreover, these papers have raised concerns that the prices emerging from a sequence of look-ahead dispatch processes may not be profit maximising: Faced with expectations of those prices emerging, the individual generators would not find it profit-maximising to comply with the dispatch instructions. This point is noted, for example by Ela and O'Malley (2015):

> The crucial issue is that when future intervals affect the schedules of current intervals, the pricing must incentivize resources to perform as directed. If this is not done or guaranteed, resources will not be incentivized to assist during ramping periods or they may be incentivized to provide false offer information (e.g., offer costs or ramp rate information).

16. Guo et al. (2019) use the equivalent term 'individual rationality' to express the same condition.

To illustrate the time-inconsistency problem, let's consider a simple network model, with three generators, each with constant marginal cost up to a capacity limit. One of the generators will be assumed to be limited in its ramp rate. The parameters for this simple model are set out in Table 1.

Table 1: Generator parameters for example 1

Gen	MC ($/MWh)	Capacity MW	Ramp Limit MW/DI	Initial state MW ($t = 0$)
G1	$10	55	–	55
G2	$80	200	10	45
G3	$100	100	–	0

(MC: marginal cost, DI: dispatch interval)

We will consider the efficient response of this power system to a perfectly-anticipated step down in demand. We will assume that demand, which starts at 100 MW is forecast to step down to 60 MW at time $t = 6$. Figure 1 sets out the efficient dispatch and forecast prices resulting from the dispatch $LAD(0,12)$.[17]

In the absence of any ramp rate constraints, the efficient response to the drop in demand would simply be to reduce the output of G2 from 45 MW to 5 MW. However, due to the ramp rate constraint on G2, this dispatch is not feasible. Instead, the dispatch process must trade off between (a) reducing the output of G1 ex post (until G2 ramps down to its long-run level) and (b) increasing the output of G3 ex ante (i.e., 'pre-positioning' G3, thereby reducing the length of the adjustment to the new equilibrium ex post). Which of these is preferred depends on whether it is more costly to draw on more expensive (G3) generation ex ante, or to reduce the output of cheaper generation (G1) ex post.

In general, a mix of ex ante and ex post adjustment will be optimal, depending on the difference in the marginal cost of G2 and G3, and the difference in the marginal cost of G2 and G1. If the marginal cost of G2 is closer to the marginal cost of G1 than G3, it is relatively cheaper to reduce G1 ex post, and relatively more expensive to increase the output of G3 ex ante. In this case the dispatch process will tend to choose more ex post adjustment and less ex ante pre-positioning. Conversely, if the marginal cost of G2 is closer to the marginal cost of G3 than G1, the dispatch process will tend to prefer ex ante pre-positioning. In the example we are using, where the marginal cost of G2 is closer to the marginal cost of G3 than G1, the efficient dispatch outcome turns out to rely exclusively on pre-positioning: it is efficient to 'pre-position' G2 down to 15 MW (offset by an increase in G3) before the anticipated drop in demand.[18]

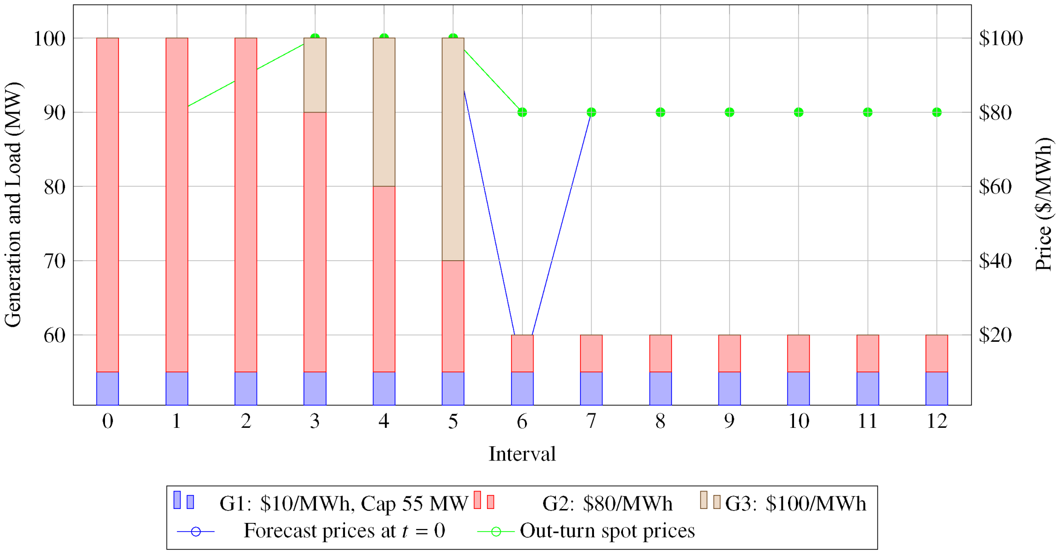
As set out in Table 2, at time $t = 0$ the sequence of future forecast prices are $p_s^0 = \$80, \$90, \$100, \$100, \$100, \$10, \$80, \$80, \ldots, \$80$.[19] However, the sequence of prices which emerges as we

17. Intuitively, what gives rise to this range of optimal prices? This effect can be thought of as arising from an asymmetry in the most efficient way to address an increment or decrement to demand at time $t = 6$. A small increment in demand at time $t = 6$ is most efficiently addressed by increasing the output of G2 by the same amount in times $t = 3$ to $t = 6$ while decreasing the output of G3 by the same amount in times $t = 3$ to $t = 5$. This has the effect of reducing the generation cost by $100-$80=$20 per unit in times $t = 3$ to $t = 5$ and increasing the generation cost in $t = 6$ by $80 per unit. The overall benefit is $20/unit of increase in demand, giving rise to the upper forecast price of $20/MWh. Conversely, a reduction in demand by a small amount at time $t = 6$ is most efficiently addressed by reducing output of G1 by that amount at time $t = 6$, reducing the total generation cost by $10/unit of demand, giving rise to the lower forecast price of $10/MWh.

18. Again, we emphasise that this is just one possible outcome. If the marginal cost of G2 was $75/MWh, the efficient dispatch outcome would involve less pre-positioning (G3 increases to only 20 MW ex ante) and more adjustment ex post (G2 drops to 45 MW for one dispatch interval ex post).

19. It is worth noting that although the dispatch of G1 does not change over this time period, it is not the case that G1 is irrelevant to the determination of the efficient dispatch outcome. The marginal cost of G1 affects the range of efficient spot

Figure 1: Pricing and dispatch outcomes for example 1



step forward in time is not the same as these prices that are forecast at the outset. Let's assume perfect foresight and that generators perfectly comply with their dispatch instructions. Table 2 shows the pricing outcomes for $LAD(t,12)$ for $t = 0,\ldots,6$. As noted earlier, in each dispatch interval, the spot price is defined to be the first in the sequence of forecast prices—which is highlighted in a box in Table 2. As can be seen, the sequence of spot prices that emerges from consecutive look-ahead dispatch is $80, $90, $100, $100, $100, $80, $80, …. This sequence of prices is not profit maximising. Faced with this sequence of future prices, generator G2 can increase its profit by choosing a different dispatch.

Table 2: Forecast prices for $LAD(t,12)$, $t = 0,\ldots,6$ for example 1.

Time	p_1^t	p_2^t	p_3^t	p_4^t	p_5^t	p_6^t	p_7^t
$t = 0$	$80	$90 80...90	$100	$100	$100	$10 10...20	$80
$t = 1$		$90 80...90	$100	$100	$100	$10 10...20	$80
$t = 2$			$100	$100	$100	$20 10...20	$80
$t = 3$				$100	$100	$40 10...40	$80
$t = 4$					$100	$60 10...60	$80
$t = 5$						$80 10...80	$80
$t = 6$							$80

(Where the price is not uniquely determined the numbers in smaller font underneath show the range of possible values)

What is driving this result? In power system dispatch problems, if we assume linear (or piecewise-linear) cost and utility functions, situations can arise where the wholesale price—which

and forecast prices.

is defined as the marginal value of the energy balance constraint—is not uniquely determined (see Hogan et al., 1996). As Hogan (2012) observes:

> The concern is not that there is no market-clearing price; rather the situation is that there are many market-clearing prices. For these supply and demand conditions, any price in the degenerate range would suffice. . . . Selecting any price in the range would support the economic dispatch solution. . . . And any price in this range have a claim to be an efficient price because it supports this surplus maximizing outcome.

The observation that the ex post spot price outcomes are not time consistent appears, at first sight, inconsistent with theorem 1. But theorem 1 only asserts that *a* pricing outcome which is optimal in an earlier dispatch process will, under perfect foresight, remain an optimal price in a later dispatch process. But there could be many possible, acceptable spot price outcomes. There is no guarantee that the dispatch algorithm will, at a later time, choose the same price (from the set of acceptable prices) that it chose at an earlier time.

This effect is illustrated in Table 2. As noted earlier, the first row of Table 2 shows the forecast prices which emerge from the original problem *LAD*(0,12). Where there is some ambiguity in the possible spot or forecast price, the range of possible values is shown in smaller font below the value chosen by the optimization algorithm. As can be seen, even the original problem has some ambiguity in the possible forecast prices. As we step through the subsequent dispatch tasks *LAD*(1,12), *LAD*(2,12) and so on, we find that this ambiguity in the price increases. But the spot and forecast prices which were a possible solution to *LAD*(0,12) remain an optimal solution for the subsequent dispatch tasks. In this sense there is no time inconsistency. But there is no guarantee that subsequent runs of the dispatch process will choose a spot price within the range of the forecast prices which were efficient at the outset.[20]

In short, under perfect foresight, the multi-interval dispatch problem is time consistent—a path of prices which is optimal for a given run of the dispatch process, remains optimal for later runs of the dispatch process. The time inconsistency issue highlighted by Ela and O'Malley (2015), Hua et al. (2019) and Zhao et al. (2018) is a consequence of the assumption that the dispatch task is (piecewise) linear. In this case, there can be a range of optimal prices consistent with any given dispatch. Furthermore, the range of prices which is optimal for a later run of the dispatch process may be larger than the range of prices for that time which is optimal in an earlier run of the dispatch process. Even under perfect foresight, there is no guarantee that the dispatch algorithm will happen to choose a spot price which falls within the range of values which was optimal at an earlier time.

A diagram may help in developing intuition on this point. As illustrated in Figure 2, the dispatch task can be viewed as finding the maximal value of a hyperplane on a convex set representing the set of feasible points. The slope of the hyperplane is determined by the dual variables (in this case the price). When the problem is linear, the feasible set is a polygon. The maximal point (i.e., the efficient dispatch) may be at a vertex of that polygon, in which case the slope of the tangent hyperplane is not uniquely determined. A range of different prices will yield the same maximal point. In contrast, when the optimisation problem is strictly convex (as illustrated in the second row of

20. Hua et al. (2019) note their belief that a commercial solution algorithm is likely to choose the price at the lower end of the feasible range. A similar observation is made by Ela and O'Malley (2015) but they observe that 'this is not proven'. We suggest that a commercial solver is likely to choose outcomes at the end of the range (either the upper or lower end), rather than the interior, as we observe in the simple examples used in this paper.

Figure 2: Linear dispatch models and price ambiguity

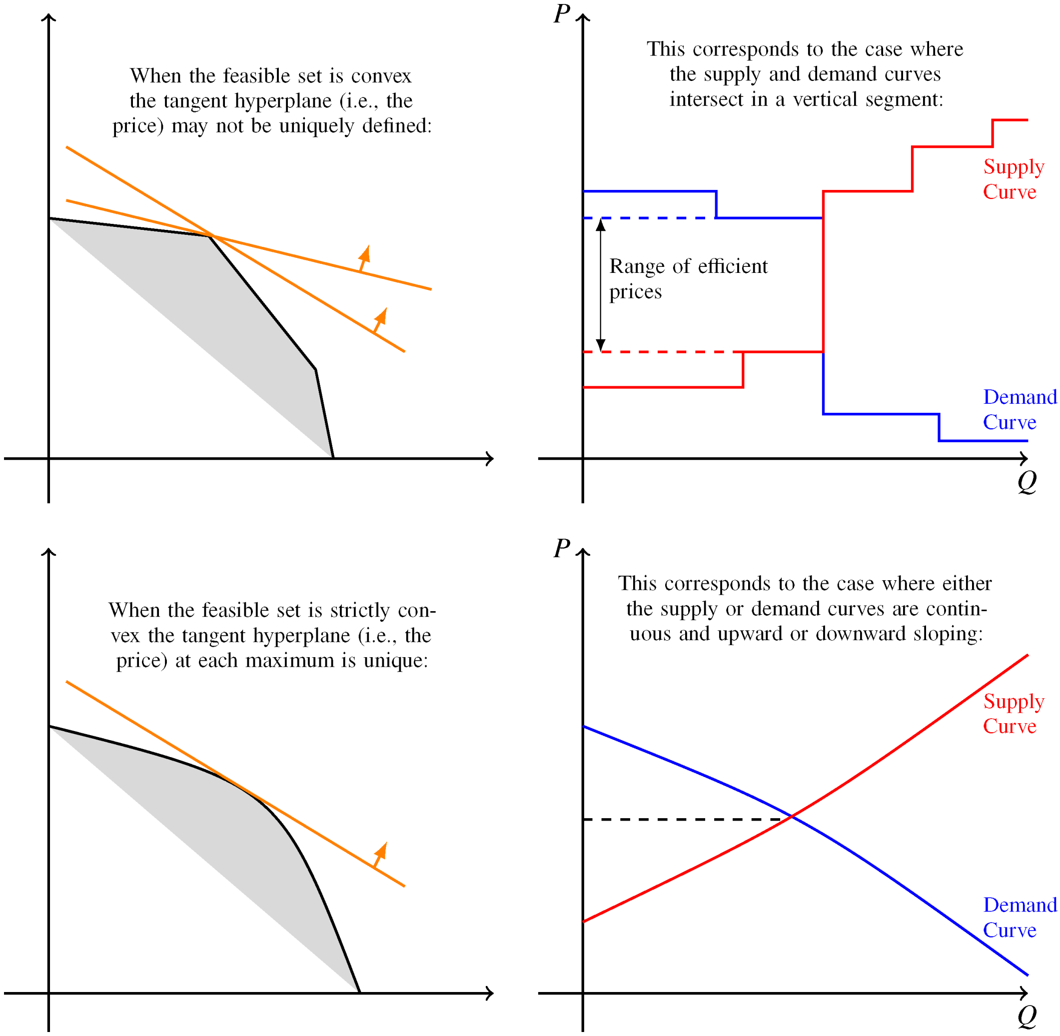


Figure 2), the tangent hyperplane is unique—there is only a single set of prices associated with each optimal dispatch. The time inconsistency problem disappears.[21]

In our view, the assumption that both the cost and utility functions are linear is quite strong. If, instead, the supply curves are continuous and upward sloping or the demand curves continuous and downward sloping, the dispatch task is strictly convex and the time inconsistency problem no longer arises.

In our view, the assumption that demand curves are continuous and downward sloping is more likely in practice. Typically the supply curve of most generators is flat or slightly upward sloping up to the capacity of the generator. But if different types of generators have very different cost curves, the aggregate or market supply curve may have vertical segments when all generators

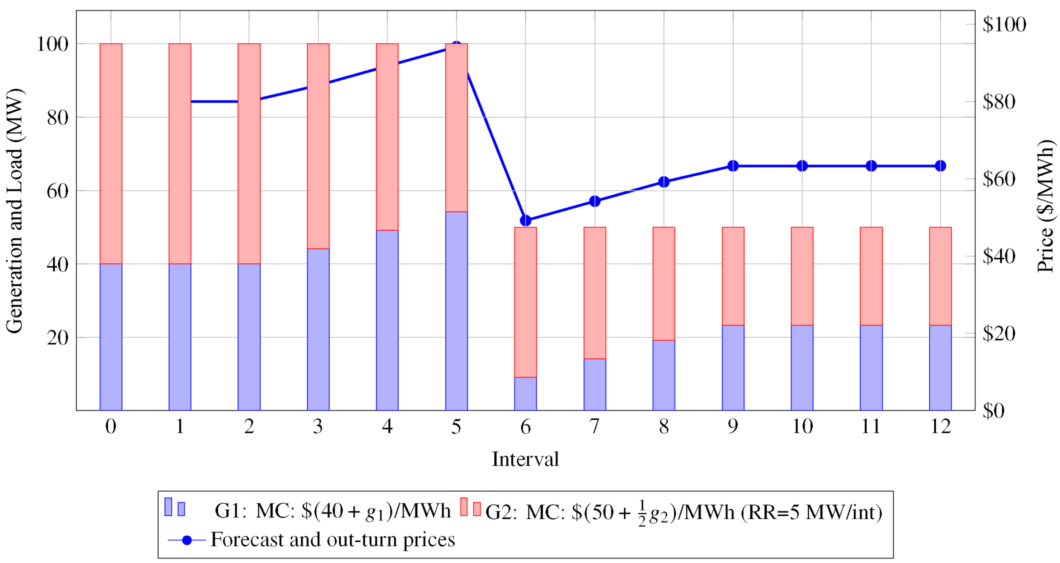
21. Prof. Ross Baldick points out to us that the link between time inconsistency and dual degeneracy, and the fact that this problem disappears when we assume strict convexity, was pointed out to him by Prof. Ben Hobbs from Johns Hopkins University in an email exchange in December 2012.

of one type are operating at their capacity, and generators which are the next highest in the merit order have not yet been activated. In contrast, a typical power system has a large number of differentiated electricity consumers. The aggregate demand curve is therefore expected to be continuous and downward sloping.[22]

To confirm this intuition, let's consider a couple of further examples. Example 2a involves generators with a linear upward-sloping supply curve, while retaining perfectly inelastic demand. Specifically, in example 2a there are two generators with cost functions of the form: $c_i(g_i) = a_i g_i + \frac{1}{2} b_i g_i^2$ where G1 has parameters $a_1 = 40$ and $b_1 = 1.0$ and G2 has parameters $a_2 = 50, b2 = 0.5$. Generator G2 is assumed to be limited to ramp up or down no more than 5 MW per dispatch interval. The demand is assumed to start at 100 MW, but is anticipated to drop to 50 MW at time $t = 6$. The optimal dispatch outcome $LAD(0,12)$ is set out in Figure 3. As can be seen, the optimal response of the power system is to ramp down G2 (while ramping up G1) in anticipation of the fall in demand. The drop in demand at time $t = 6$ is met by a decrease in production in G1, which is subsequently ramped up as G2 continues to ramp down to the long-run steady state.

Figure 3 also illustrates the corresponding forecast profile of prices. As can be seen, the price starts at the initial steady state of $80/MWh. The future price is forecast to increase for several intervals *before* the drop in demand, reaching a peak of $94.17/MWh, dropping suddenly to $49.17 with the drop in demand, and then increasing over several intervals back to the new long-run level of $63.33/MWh. Importantly, these prices are unique. As we expect, these prices also emerge as the spot price in subsequent runs of the dispatch process $LAD(t,12)$, $t = 1,\ldots,12$.

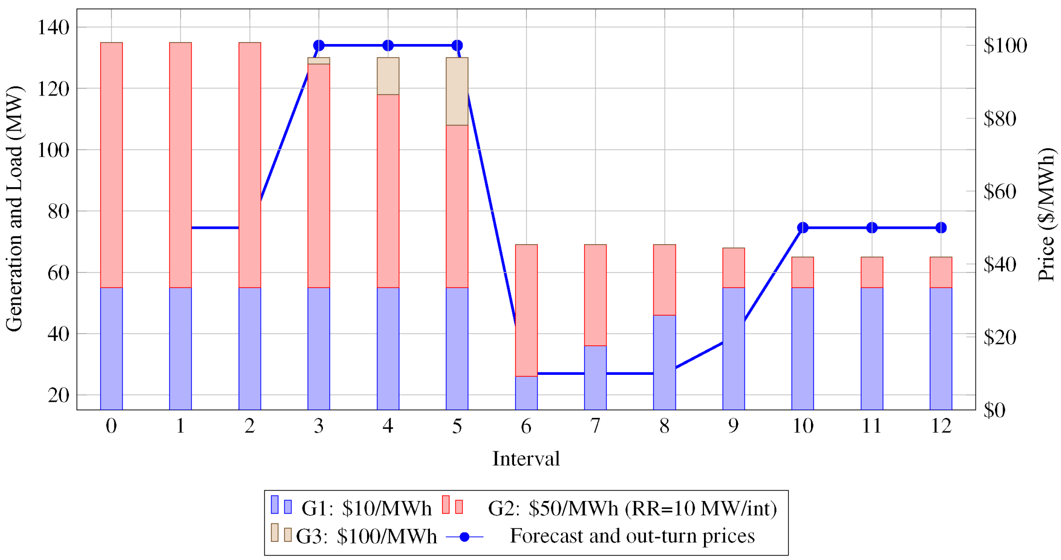
Figure 3: Pricing and dispatch outcomes for example 2a (upward sloping supply, perfectly inelastic demand)



To complete this story, let's consider one further example. Example 2b retains the generators G1-G3 (with constant marginal cost, and ramp rate constraint on G2) from example 1, but

22. Historically, many power systems have modelled demand as being perfectly inelastic—as in example 1 and 2a. With increasing penetration of smart meters it is increasingly feasible for electricity consumers to reflect their demand curve to the wholesale market.

Figure 4: Pricing and dispatch outcomes for example 2b (constant MC generation, elastic demand)



introduces a downward sloping demand curve of the form $a - 10L$, where L is the load (MW) and a is a parameter. The demand parameter a starts at 1400, but drops to 700 at time $t = 6$. The optimal dispatch outcome $LAD(0,12)$ is set out in Figure 4. The spot price starts at the initial steady state of $50/MWh, but is forecast to increase to $100/MWh for several intervals before the drop in demand, dropping to $10/MWh with the drop in demand, and then increasing over several intervals back to the long-run level of $50/MWh (but at a lower level of demand). Again, these prices are time consistent: They also emerge as the spot price in subsequent runs of the dispatch process $LAD(t,12)$, $t = 1,\ldots,12$.

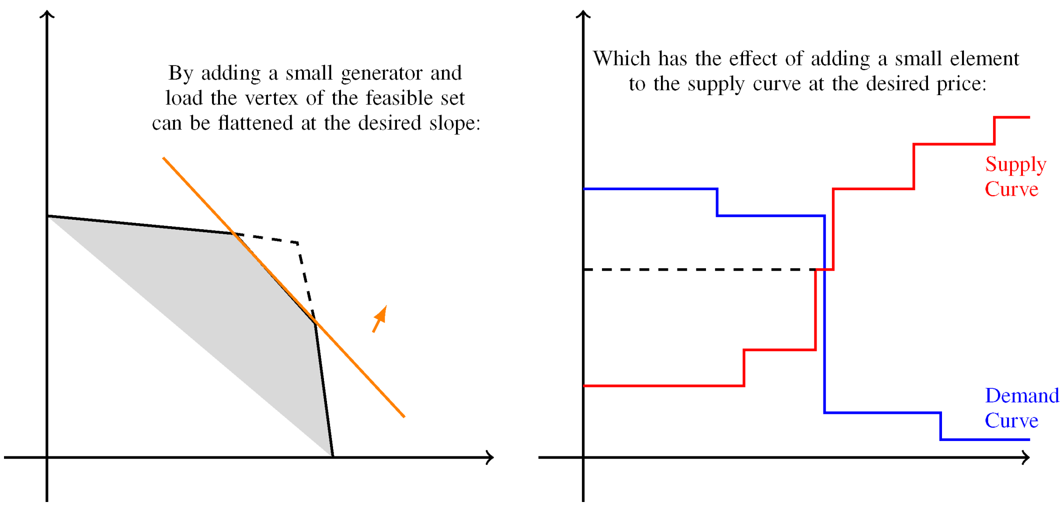
2.3 Proposals for augmenting the dispatch process to resolve time inconsistency

A concern with time inconsistency has prompted several authors to propose augmentations or enhancements to the simple dispatch process. For example, Hua et al. (2019) propose an approach which recognises that, at any point in time, the value of the constraint marginal value on the ramp rate constraint reflects the impact of that constraint on the economic cost of dispatch. Therefore, at each subsequent point in time, if we incorporate the earlier value of the ramp constraint marginal value into the objective function we can force the dispatch engine to take into account the past costs. Hua et al. (2019) refer to this approach as ‘constraint-preserving multi-interval pricing’ or CMP. The marginal values of the ramp constraint can themselves be derived from past spot and forecast prices. Therefore it is not surprising that we can also achieve the same outcome using historic prices. Hua et al. (2019) refer to this as ‘price-preserving multi-interval pricing’ or PMP.

Perhaps the simplest and most intuitive way to resolve the time inconsistency problem in the face of dual degeneracy is to introduce a ‘phantom generator’ and/or a ‘phantom load’ into the dispatch process. A phantom generator (or phantom load) is a resource with a very small capacity and with a marginal cost (or willingness to pay) equal to the desired price outcome. As long as the capacity of the generator or load is very small, the overall dispatch outcome will not be affected; but the presence of these very small resources adds a small ‘kink’ in the supply or demand curves at the required location, which causes the dispatch process to yield the desired price in the special case of perfect foresight.

This can be explained further as follows: Let's suppose that, at time t the system operator solves the dispatch task $LAD(t,T\,|\,g_t^t)$ resulting in a sequence of forecast prices $(p_s^t)_{s=t+1,\ldots,T}$. At some later time $\hat{t}$ $(t+1 \le \hat{t} \le T)$ we then augment the dispatch task $LAD(\hat{t},T\,|\,g_{\hat{t}}^{\hat{t}})$ by adding a generator and a load with capacity ε and a marginal cost p_s^t at times $s = \hat{t}+1,\ldots,T$. Here ε is a small number. This approach ensures that, under perfect foresight, where there is ambiguity in the spot price, the dispatch process chooses a forecast price $p_s^{\hat{t}}$ equal to the original forecast p_s^t. Intuitively, the effect of this change is to 'flatten' the vertex of the feasible set by a small amount, by adding a small section with a slope equal to the desired price. This is illustrated in Figure 5.

Figure 5: Intuition behind the 'phantom generator' approach



Formally, the augmented dispatch process in this 'phantom resource' approach $LAD^{PR}(\hat{t},T\,|\,p^t, g_{\hat{t}}^{\hat{t}})$ is defined as follows:

$$\min_{g_{i,s}^{\hat{t}},\, g_{+s}^{\hat{t}},\, g_{-s}^{\hat{t}}} \sum_{s=\hat{t}+1,\ldots,T} \sum_i c_i g_{i,s}^{\hat{t}} + p_s^t (g_{+s}^{\hat{t}} - g_{-s}^{\hat{t}}) \tag{7}$$

s.t.

$$\forall s \in [\hat{t}+1,T],\ \sum_i g_{i,s}^{\hat{t}} + g_{+s}^{\hat{t}} - g_{-s}^{\hat{t}} = L_s^{\hat{t}} \tag{8}$$

$$\forall i,s \in [\hat{t}+1,T],\ -R_i \le g_{i,s}^{\hat{t}} - g_{i,s-1}^{\hat{t}} \le R_i \tag{9}$$

$$\forall s \in [\hat{t}+1,T],\ g_{+s}^{\hat{t}} \le \varepsilon, g_{-s}^{\hat{t}} \le \varepsilon \tag{10}$$

Here $g_{+s}^{\hat{t}}$ and $g_{-s}^{\hat{t}}$ are the dispatch of the phantom generator and load respectively.

Applying this approach to the look-ahead dispatch task of example 1, the first run of the dispatch process $LAD(0,12)$ yields a sequence of forecast prices $p^0 = \$80, \$90, \$100, \$100, \$100, \$10, \$80, \ldots$. Using these prices as a phantom generator in the subsequent runs of the dispatch, and assuming perfect foresight, we find that the sequence of spot prices is equal to the original forecast prices; time consistency has been restored.

2.4 Handling new information in economic dispatch

But there remains a further problem. As we have seen, it is possible to modify the simple look-ahead dispatch process to solve the time-inconsistency problem. But the look-ahead dispatch formulation considered above only allows for a single state of the world in each future dispatch interval. In real-world markets, new information arrives all the time. This could be new information about demand conditions (due to, e.g., changes in the weather), network conditions (such as network outages), or supply conditions (e.g., unplanned outages, changes in fuel costs, changes in wind or solar production). The power system must be able to efficiently take that new information into account and reflect it in both current and future prices and dispatch.[23]

But all of the mechanisms discussed above tie future spot prices to past forecast price outcomes. The only way this can be made consistent with efficient dispatch in a context in which new information arrives over time, is if these mechanisms are applied in a context of full **state-contingent dispatch**.

Formally, let's suppose that σ_s is a random variable which completely describes the state of the power system at some time s. A state-contingent dispatch is a set of prices $p_s^t(\sigma_s)$ and a dispatch $g_{i,s}^t(\sigma_s)$ which depend on the state of the power system in the future.

As we have observed earlier, in a context in which both cost and utility are linear, there can be ambiguity in the optimal price at a given time and state of the world. The mechanisms discussed above can ensure, at a given time and state of the world, the dispatch process will choose a price from within that range which is consistent with the prices forecast at the outset. This resolves the time consistency problem, in theory.

To make this point concrete, let's extend example 1 above to illustrate a simple state-contingent dispatch. In example 1, demand in the power system starts out at 100 MW, and is forecast to drop to 60 MW at time $t = 6$. Let's now assume that there is some uncertainty in that drop in demand. Specifically, let's now assume that there is a 50% chance that the demand will drop to 60 MW at time $t = 6$, and a 50% chance that demand will only drop to 90 MW. We will assume that this uncertainty is realised at time $t = 5$ (i.e., although the change in demand is forecast, the precise amount by which it drops is not known until the start of the dispatch interval before the drop in demand occurs). The efficient dispatch in this case is to pre-position the power system by ramping down G2 to 25 MW, while ramping up G3 to 20 MW, before the drop in demand. Following the drop in demand the efficient dispatch involves backing off G1 for one dispatch interval in the case when the out-turn demand is 60 MW, and increasing the output of G2 to 35 MW in the case when the out-turn demand is 90 MW.

Table 3 shows the efficient pricing outcomes in this simple model as a function of the state of the world (in this case reflected in the out-turn level of demand). Until time $t = 5$ there is a single set of prices p_s^t, reflecting the known path of demand up until this time. From time $t = 6$ there are two sets of prices $p_s^{t,LO}$ and $p_s^{t,HI}$, corresponding to the outcome that demand is low (60 MW) or high (90 MW). As before, the numbers in smaller font underneath each price outcome represents the range of possible efficient price outcomes in the event of dual degeneracy.

23. To repeat the point noted earlier, in this paper we are not focusing on the question of efficient response to new information on a timescale shorter than the dispatch interval. Therefore, we will make the assumption that the new information arrives at discrete intervals, at the end of each dispatch interval (in time to be taken into account in the subsequent dispatch process).

Table 3: Forecast prices for $LAD^{SCD}(t,12)$, $t = 0,\dots,6$ for example 3.

Time	p_1^t	p_2^t	p_3^t	p_4^t	p_5^t	$p_6^{t,LO}$ / $p_6^{t,HI}$	$p_7^{t,LO}$ / $p_7^{t,HI}$	$p_8^{t,LO}$ / $p_8^{t,HI}$
$t = 0$	$80	$80	$100	$100	$100	$10	$30 (10...70)	$80
						$100 (80...100)	$80	$80
$t = 1$		$80	$100	$100	$100	$10	$10 (10...70)	$80
						$100 (80...100)	$80	$80
$t = 2$			$100	$100	$100	$10	$10 (10...70)	$80
						$100 (80...100)	$80	$80
$t = 3$				$100	$100	$10	$10 (10...70)	$80
						$100 (80...100)	$80	$80
$t = 4$					$100	$10	$10 (10...80)	$80
						$100 (80...100)	$80	$80
$t = 5$						$10	$10 (10...80)	$80
						$100 (80...100)	$80	$80
$t = 6$							$80 (10...80)	$80
							$80	$80

(Where the price is not uniquely determined the numbers in smaller font underneath show the range of possible values)

As can be seen, there remains a time inconsistency problem: The forecast prices for the low-demand case are $80, $80, $100, $100, $100, $10, $30, $80, but the out-turn prices are $80, $80, $100, $100, $100, $10, *$80*, $80. In principle, this time-inconsistency could be resolved by implementing one of the mechanisms identified above.

However, we emphasise that in practice such state-contingent dispatch is entirely impractical. In any real world market there are a very large number of independent contingencies possible in each time period. Combining these possibilities into sequences of possible paths of outcomes quickly gives rise to an exponentially large number of combinations. Estimating the probabilities of all such events is likely to be difficult if not impossible. Computing the optimal dispatch in all such contingencies is almost certainly impossible.[24]

State-contingent dispatch remains a useful theoretical idea but is infeasible in practice. If state-contingent dispatch were feasible we would not need a real-time market at all—we could run the dispatch process say, once per day (perhaps at the same time as the 'day-ahead market'), evaluating the optimal dispatch and pricing outcomes once for each of the very large number of contingencies that might evolve over the remainder of the day. This is not feasible. It is precisely for this reason that in practice the real-time dispatch is carried out as close as possible to real-time—so that all the latest information about the state of the market can be fully taken into account at the time of dispatch.

24. State contingent dispatch or 'stochastic market clearing' is modelled in Daraeepour et al. (2019).

3. DO WE NEED TO IMPLEMENT MULTI-INTERVAL REAL-TIME MARKETS?

How then, can we achieve efficient dispatch in the face of binding ramp rate constraints and new information arriving all the time?

We suggest that it may not be necessary to implement 'look-ahead dispatch' in the real-time dispatch process at all. In fact, it may not be necessary for the central dispatch process to take into account *any* inter-temporal constraints.

To see this, let's consider a power system in which:

1. All generators and loads are price takers;
2. All generators and loads have continuous upward-sloping supply curves or continuous downward-sloping demand curves;
3. There are no limits on the supply and demand curves which generators and loads can submit to the dispatch process each dispatch interval; and
4. There are no inter-temporal constraints in the network (the only inter-temporal constraints are in the physical operation of the generators and loads).

Under these conditions, it is straightforward to demonstrate that a one-shot or myopic dispatch process can achieve the fully efficient outcome, even in the presence of ramp-rate constraints and other inter-temporal constraints.

Intuitively, the process operates as follows: Let's suppose that, each period, we have a set of forecast future prices[25]. Each generator (or load) then carries out its own inter-temporal profit-maximisation (or utility-maximisation) problem, based on the forecast prices, to determine their own desired offer (or bid) curve, which they submit to the dispatch process. The dispatch process then carries out a conventional centralised one-shot dispatch process, based on the bids and offers submitted, to determine the optimal short-run price and dispatch outcomes. Provided there are no restrictions on the offer (or bid) curve that each generator or load can communicate to the central dispatch process each dispatch interval, the sequence of one-shot processes matches the efficient inter-temporal dispatch outcome. Also, under perfect foresight, the resulting prices match the forecast prices.

The following theorem states the main result (the proof is in the online appendix of this paper).

> **Theorem 2** *Let's suppose that, at time t, we have a set of forecast prices* $(p_s^t)_{s=t+1,\dots,T}$ *which is an unbiased estimate of the actual future prices. Let's suppose that, at each time s, each generator i submits an offer curve which for each possible spot price p, maximises the generator's expected profit given any ramp rate constraints. In addition, let's suppose that the system operator uses these offer curves in a simple, one-shot dispatch process to find the optimal dispatch and pricing outcome for that time period.*
>
> *The resulting sequence of dispatch and pricing outcomes are efficient (in other words, are a solution to* $LAD(t,T)$*). In addition, if the problem is strictly convex the sequence of dispatch and pricing outcomes is the same as was forecast at the outset.*

To see how this might work in practice, let's return to example 2a above. Starting at time $t = 0$ the market participants forecast the sequence of spot prices given in Figure 3, that is:

$$p^0 = (\$80, \$80, \$84.17, \$89.17, \$94.17, \$49.17, \$54.17, \$59.17, \$63.33, \dots)$$

25. These prices should, in theory, include forecasts of prices in all future contingencies.

In each subsequent period, the market participants submit an offer curve to the system operator which maximises their profit given any inter-temporal constraints and price forecasts. Since G1 is not ramp constrained, and since we are assuming price-taking behaviour, it has an incentive to submit an offer curve which matches its cost curve. In the case of G2, however, we find that the profit-maximising offer curve varies from interval to interval, as set out in Table 4. Table 4 sets out the parameters for the quadratic cost function within the range that G2 can feasibly ramp. Outside this feasible set, the cost function is assumed to have a large negative slope (below the range) and a large positive slope (above the range), as in equation 17 in the online appendix.

Table 4: Profit-maximising offer curves for G2 when facing prices $p^0 = (\$80, \$80, \$84.17, \$89.17, \$94.17, \$49.17, \$54.17, \$59.17, \$63.33, ...)$

Dispatch Interval	Initial Output	a_i (\$/MWh)	b_i (\$/MW^2h)	Range
$s = 1$	60MW	\$50	0.5	$55 \le g \le 65$
$s = 2$	60MW	\$50	0.5	$55 \le g \le 65$
$s = 3$	60MW	\$-84.76	3.03	$55 \le g \le 65$
$s = 4$	55.83MW	\$-65.29	3.04	$50.83 \le g \le 60.83$
$s = 5$	50.83MW	\$-41.79	2.97	$45.83 \le g \le 55.83$
$s = 6$	45.83MW	\$-72.71	2.99	$50.83 \le g \le 50.83$
$s = 7$	40.83MW	\$-52.79	2.99	$35.83 \le g \le 45.83$
$s = 8$	35.83MW	\$49.02	0.33	$30.83 \le g \le 40.83$
$s = 9$	30.83MW	\$50	0.5	$25.83 \le g \le 35.83$
$s = 10$	26.67MW	\$50	0.5	$21.67 \le g \le 31.67$

When these offer curves are submitted to a one-shot dispatch process we find that the dispatch and pricing outcomes are efficient (the same as illustrated in Figure 3). In other words, generator G2 engages in efficient pre-positioning—without any need for look-ahead dispatch. Furthermore, the pricing outcomes are time consistent—under perfect foresight, the same sequence of prices emerges as forecast at the outset. Since this is a one-shot dispatch process it can easily be applied in a context in which new information arrives all the time, although in this case, the price forecasts must be contingent on new information as it arrives.

It is worth noting that, even though there is no exercise of market power in this model, the offer curves that emerge from this process are not necessarily 'cost related' in the sense that they do not directly reflect the cost of production in any given time period. Instead they reflect a combination of the direct production costs and the inter-temporal opportunity costs. In some markets, there are rules requiring generator offers to be cost-related as a way of controlling for the exercise of market power. Those rules may need to be adjusted to accommodate the types of offers set out here.

4. CONCLUSION

As the California 'duck curve' highlights, with increasing penetration of intermittent generation, there is concern that inter-temporal constraints, such as ramp rate constraints, will be binding with increasing frequency—potentially leading to inefficient or even infeasible dispatch outcomes. As a consequence, there is increasing pressure to incorporate inter-temporal considerations in the real-time dispatch process. This might take the form of separate procurement of ramping capability, or procurement of additional flexible operating reserve. Several markets, especially in the U.S., have adopted a form of 'look-ahead dispatch' or 'multi-interval real-time market', and use the output of these processes to commit generators or to pre-position the power system in advance of anticipated binding constraints.

Several papers in the literature have highlighted a potential problem with look-ahead dispatch: The prices that emerge from a sequence of look-ahead dispatch processes may not be 'time consistent' in the sense that, if those prices were anticipated, the market participants would not have an incentive to comply with the dispatch instructions. This is a potentially serious concern. We emphasise, however, that this problem of time inconsistency only arises in what is, arguably a special case: The case where both cost and utility functions of market participants are (piece-wise) linear. In this case there arises a problem of dual degeneracy and consequent ambiguity in the definition of the wholesale spot price. Under perfect foresight, the forecast prices emerging from an earlier run of the dispatch process remain within the set of optimal outcomes for subsequent runs of the dispatch process, but there is no guarantee that subsequent runs of the dispatch process will choose those earlier prices. As a result, the sequence of prices that emerges may not be time consistent.

In response, several authors have proposed mechanisms to resolve the time inconsistency problem—that is, to induce the dispatch process to choose prices which fall within a range defined ex ante. However, as proposed, these mechanisms assume there is only one possible state of the world in each future dispatch interval. This is only possible where no new information comes along over time. In all real world markets, new information about demand, supply, and network conditions arrives all the time and must be reflected in the dispatch process. This could be achieved by applying these mechanisms in the context of 'state contingent dispatch'. This is a theoretical dispatch process in which, at a given time, future efficient dispatch outcomes are determined as a function of the possible future states of the world. When applied in the context of state-contingent dispatch, the time-inconsistency problem could, in principle, be resolved, while allowing for efficient response to new information as it arrives over time. In practice, however, due to the large number of possible contingencies, such state-contingent dispatch is a practical impossibility.

We emphasise that, even in a context in which ramp rate constraints are binding, it may not be necessary to implement look-ahead dispatch to achieve efficient generator commitment or pre-positioning decisions at all. We show that under certain assumptions, a sequence of one-shot dispatch processes, combined with individual profit-maximising (or utility-maximising) decisions by generators and loads achieves the efficient overall outcome, including efficiently anticipating and 'pre-positioning' the power system in the face of anticipated swings in the supply/demand balance. These arguments lend weight to those who argue for a simple, one-shot, central dispatch process.[26] In fact, we would predict that we should already observe generators choosing to make non-cost-based offers in one-shot markets such as the Australian NEM in order to pre-position themselves efficiently in advance of binding ramp rates. This could be tested empirically.

In this context, the role of a centralised look-ahead dispatch process is, at most, to provide market participants improved forecasts of short-term prices. However, in our view, if a centralised look-ahead dispatch process is to be retained, consideration should be given to not just computing the optimal dispatch and prices in a single state of the world in the future, but instead calculating the optimal dispatch-and-pricing outcomes over a range of future credible contingencies. Such an approach could be a tool for improving short-term price forecasts thereby encouraging market participants to efficiently pre-position themselves for potential future market outcomes.

26. For example, increasing concerns about variability in the supply/demand balance in the Australian National Electricity Market (NEM) have led the system operator to consider introducing new markets to commit generators, or procure additional ramping capability, in advance of potentially binding ramp rate constraints, as noted in footnote 1. The analysis in this paper lends weight to the view that there is no need to augment the existing one-shot dispatch process in the NEM.

ACKNOWLEDGMENTS

The views expressed in this article do not reflect the views of the Australian Competition and Consumer Commission (ACCC) or the Australian Energy Regulator (AER). The authors would like to thank Prof. Bill Hogan from Harvard Electricity Policy Group (HEPG) and Prof. Ross Baldick from University of Texas at Austin, and the anonyous reviewers for their valuable comments on the earlier versions of this paper.

REFERENCES

AEMC (2020). "System Services Rule Changes" https://www.aemc.gov.au/sites/default/files/2020-07/ System%20services%20rule%20changes%20-%20Consultation%20paper%20%E2%80%93%20%202%20July% 202020.pdf accessed: 2020-07-27.

AEMO (2018). "AEMO Observations: Operational and market challenges to reliability and security in the NEM."

CAISO (2015). "Flexible ramping product: Revised Draft Final Proposal." http://www.caiso.com/Documents/Revised-DraftFinalProposal-FlexibleRampingProduct-2015.pdf accessed: 2020-07-27.

CAISO (2019). "Price Performance in the CAISO's Energy Market." http://www.caiso.com/Documents/Report-PricePerformanceAnalysis.pdf accessed: 2020-07-27.

Choi, D.-H. and L. Xie (2016). "Data perturbation-based sensitivity analysis of real-time look-ahead economic dispatch." *IEEE Transactions on Power Systems* 32(3): 2072–2082. https://doi.org/10.1109/TPWRS.2016.2598874.

Cornelius, A., R. Bandyopadhyay, and D. Patiño-Echeverri (2018). "Assessing environmental, economic, and reliability impacts of flexible ramp products in MISO's electricity market." *Renewable and Sustainable Energy Reviews* 81: 2291–2298. https://doi.org/10.1016/j.rser.2017.06.037.

Daraeepour, A., D. Patiño-Echeverri, and A.J. Conejo (2019). "Economic and environmental implications of different approaches to hedge against wind production uncertainty in two-settlement electricity markets: A PJM case study." *Energy Economics* 80: 336–354. https://doi.org/10.1016/j.eneco.2019.01.015.

Ela, E. and M. O'Malley (2015). "Scheduling and pricing for expected ramp capability in real-time power markets." *IEEE Transactions on Power Systems* 31(3): 1681–1691. https://doi.org/10.1109/TPWRS.2015.2461535.

Guo, Y., C. Chen, L. Tong, et al. (2019). "Pricing Multi-Interval Dispatch under Uncertainty Part I: Dispatch-Following Incentives." *arXiv preprint arXiv:1911.05784* .

Han, X., H. Gooi, and D.S. Kirschen (2001). "Dynamic economic dispatch: feasible and optimal solutions." *IEEE Transactions on power systems* 16(1): 22–28. https://doi.org/10.1109/59.910777.

Herrero, I., P. Rodilla, and C. Batlle (2018). "Enhancing intraday price signals in us iso markets for a better integration of variable energy resources." *The Energy Journal* 39(3). https://doi.org/10.5547/01956574.39.3.iher.

Hittinger, E., J. Apt, and J. Whitacre (2014). "The effect of variability-mitigating market rules on the operation of wind power plants." *Energy Systems* 5(4): 737–766. https://doi.org/10.1007/s12667-014-0130-8.

Hogan, W.W. (2012). "Multiple market-clearing prices, electricity market design and price manipulation." *The Electricity Journal* 25(4): 18–32. https://doi.org/10.1016/j.tej.2012.04.014.

Hogan, W.W. (2016). "Electricity market design: Optimization and market equilibrium."

Hogan, W.W. (2020). "Electricity Market Design: Multi-interval Pricing Models." https://scholar.harvard.edu/ files/whogan/files/hogan_hepg_multi_period_062220.pdf accessed: 2020-08-12.

Hogan, W.W., E.G. Read, and B.J. Ring (1996). "Using mathematical programming for electricity spot pricing." *International Transactions in Operational Research* 3(3-4): 209–221. https://doi.org/10.1111/j.1475-3995.1996.tb00048.x.

Hua, B., D. Schiro, T. Zheng, R. Baldick, and E. Litvinov (2019). "Pricing in Multi-Interval Real-Time Markets." *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2019.2891541.

Mickey, J. (2015). "Multi-interval real-time market overview: ERCOT Board of Directors Meeting." http://www.ercot.com/content/wcm/key_documents_lists/76342/5_Multi_Interval_Real_Time_Market_Overview.pdf accessed: 2020-07-27.

Moarefdoost, M.M., A.J. Lamadrid, and L.F. Zuluaga (2016). "A robust model for the ramp-constrained economic dispatch problem with uncertain renewable energy." *Energy Economics* 56: 310–325. https://doi.org/10.1016/j.eneco.2015.12.019.

NYISO (2019). "Reliability and Market Considerations for a Grid in Transition: A Report from the New York Independent System Operator." https://www.nyiso.com/documents/20142/9869531/Reliability%20and% 20Market%20 Considerations%20for%20a%20Grid%20in%20Transition%20-%2020191220%20Final.pdf/ 7846db9c-9113-a85c-8abf-1a0ffe971967 accessed: 2020-07-27.

Parker, N. (2015). "Ramp Product Design." https://www.spp.org/documents/29342/ramp%20product%20design. pdf accessed: 2020-07-27.

PJM (2019). "PJM Manual 11: Energy and Ancillary Services Market Operations: Revision 108." https://www.pjm.com/~/media/documents/manuals/m11.ashx accessed: 2020-07-27.

Price, J.E. and M. Rothleder (2011). "Recognition of extended dispatch horizons in California's energy markets." In "2011 IEEE Power and Energy Society General Meeting," IEEE 1–5. https://doi.org/10.1109/PES.2011.6039484.

Raithel, R., S. Virmani, S. Kim, and D. Ross (1981). "Improved allocation of generation through dynamic economic dispatch." In Seventh Power Sys. Comput. Conf. 273–280.

Ross, D.W. and S. Kim (1980). "Dynamic economic dispatch of generation." *IEEE transactions on power apparatus and systems* (6): 2060–2068. https://doi.org/10.1109/TPAS.1980.319847.

Schiro, D.A. (2018). "Procurement and Pricing of Ramping Capability." https://www.iso-ne.com/static-assets/ documents/2018/03/price_information_technical_session3.pdf accessed: 2020-07-27.

Thatte, A.A., D.-H. Choi, and L. Xie (2014). "Analysis of locational marginal prices in look-ahead economic dispatch." In "2014 Power Systems Computation Conference," IEEE 1–7. https://doi.org/10.1109/PSCC.2014.7038489.

Wang, B. and B.F. Hobbs (2013). "Flexiramp market design for real-time operations: Can it approach the stochastic optimization ideal?" In "2013 IEEE Power & Energy Society General Meeting," IEEE 1–5.

Wu, W., J. Chen, B. Zhang, and H. Sun (2014). "A robust wind power optimization method for look-ahead power dispatch." *IEEE transactions on sustainable energy* 5(2): 507–515. https://doi.org/10.1109/TSTE.2013.2294467.

Xia, X. and A. Elaiw (2010). "Optimal dynamic economic dispatch of generation: A review." *Electric power systems research* 80(8): 975–986. https://doi.org/10.1016/j.epsr.2009.12.012.

Xie, L. and M.D. Ilic (2008). "Model predictive dispatch in electric energy systems with intermittent resources." In "2008IEEE International Conference on Systems, Man and Cybernetics," IEEE 42–47. https://doi.org/10.1109/ICSMC.2008.4811248.

Xie, L., X. Luo, and O. Obadina (2011). "Look-ahead dispatch in ERCOT: Case study." In "2011 IEEE Power and Energy Society General Meeting," IEEE 1–3. https://doi.org/10.1109/PES.2011.6039515.

Xie, L., S. Puller, M. Ilic, and S. Oren (2013). "Quantifying benefits of demand response and look-ahead dispatch in systems with variable resources." *PSERC Final Report M* 26.

Xu, X. and R. Howard (2013). "Ramp rate modeling for ERCOT look ahead SCED." In "2013 IEEE Power & Energy Society General Meeting," IEEE 1–5. https://doi.org/10.1109/PESMG.2013.6672741.

Yu, C.-N., A.I. Cohen, and B. Danai (2005). "Multi-interval optimization for real-time power system scheduling in the Ontario electricity market." In "IEEE Power Engineering Society General Meeting, 2005," IEEE 2717–2723.

Zhao, J., T. Zheng, and E. Litvinov (2018). "A Multi-Period Market Design for Markets with Intertemporal Constraints." *arXiv preprint arXiv:1812.07034* .

Zhu, Y. and Y. Chen (2012). "Modeling short-term stored energy resources under real time look-ahead dispatch on energy and ancillary service market." In "2012 IEEE Energytech," IEEE 1–6. https://doi.org/10.1109/EnergyTech.2012.6304630.

IAEE

International Association for
ENERGY ECONOMICS

Membership in the International Association for Energy Economics is open to anyone worldwide who has an interest in the fields of energy or energy economics. Our membership consists of those working in both the public and private sectors including government, academic and commercial. Our current member base consists of 3900+ members in over 110 nations, with 28 nations having local affiliate organization.

We are an independent, non-profit, global membership organization for business, government, academic and other professionals concerned with energy and related issues in the international community. We advance the knowledge, understanding and application of economics across all aspects of energy and foster communication amongst energy concerned professionals.

We are proud of our membership benefit offerings, which include access to a rich library of energy economics related publications and proceedings as well as a robust line-up of webinars, podcasts and conferences. Learn more about the benefits of membership at: https://www.iaee.org/en/membership/benefits.aspx

In addition to traditional membership, we offer student and institutional memberships.