Electricity Markets: Summer Semester 2016, Lecture 7

Tom Brown, Mirko Schäfer

30th May 2016

Frankfurt Institute of Advanced Studies (FIAS), Goethe-Universität Frankfurt
FIAS Renewable Energy System and Network Analysis (FRESNA)

{brown,schaefer}@fias.uni-frankfurt.de

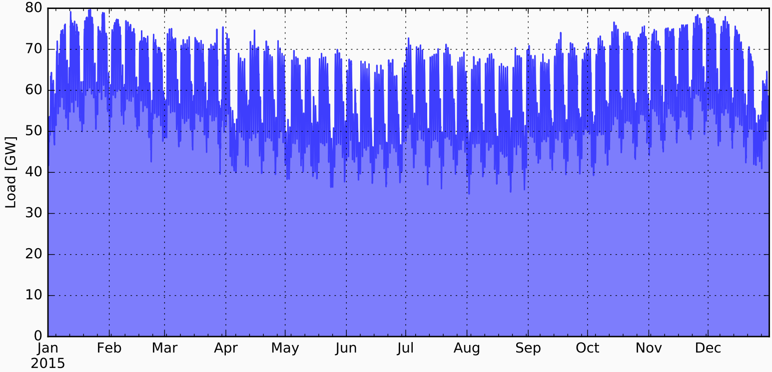
FIAS Frankfurt Institute for Advanced Studies

Table of Contents

Duration Curves and Capacity Factors: Examples from Germany in 2015

Load curve

Here's the electrical demand (load) in Germany in 2015:

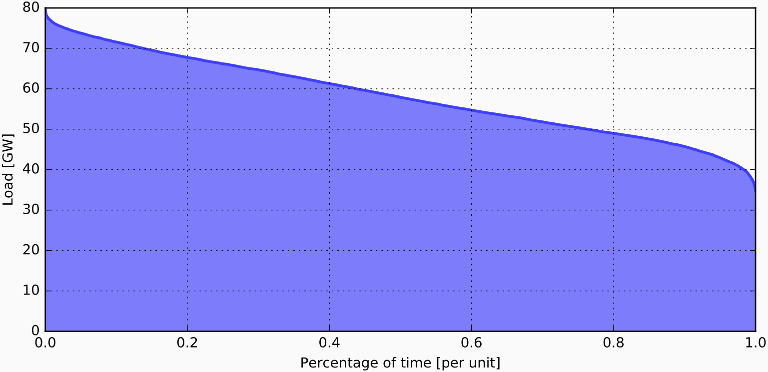


To understand this curve better and its implications for the market, it's useful to stack the hours of the year from left to right in order of the amount of load.

Load duration curve

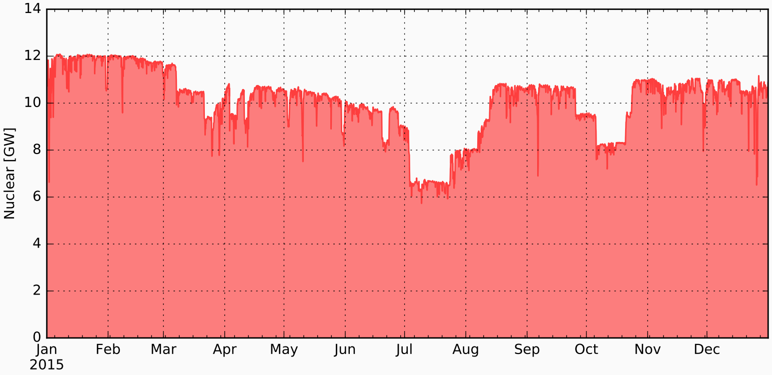
This re-ordering is called a duration curve.

For the load it's the load duration curve.



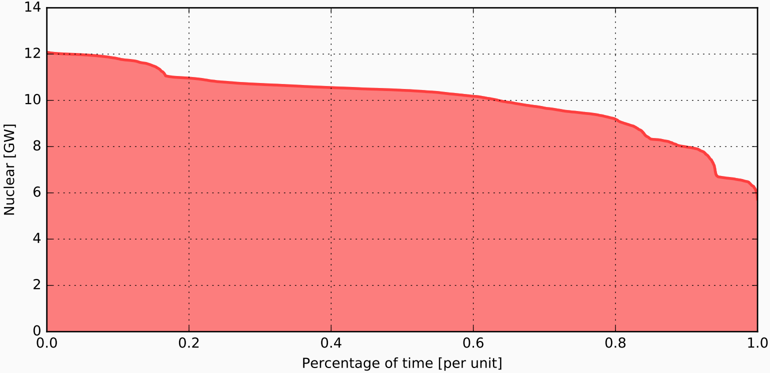
Nuclear curve

Can do the same for nuclear output:



Nuclear duration curve

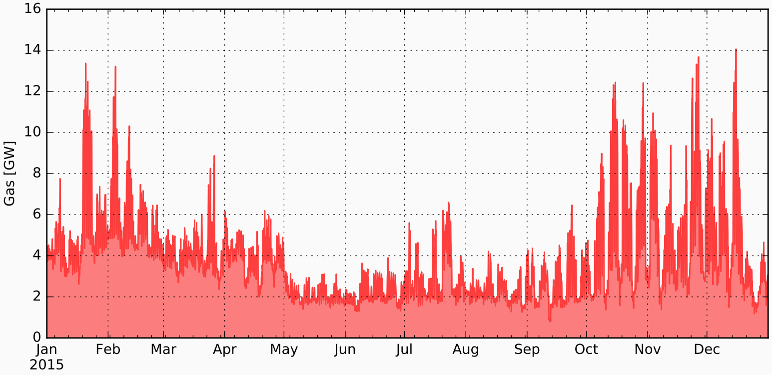
Duration curve is pretty flat, because it is economic to run nuclear almost all the time as baseload plant:



The equivalent fraction of time that the plants run at full capacity over the year is the capacity factor - nuclear has a high capacity factor, usually around 70-90%.

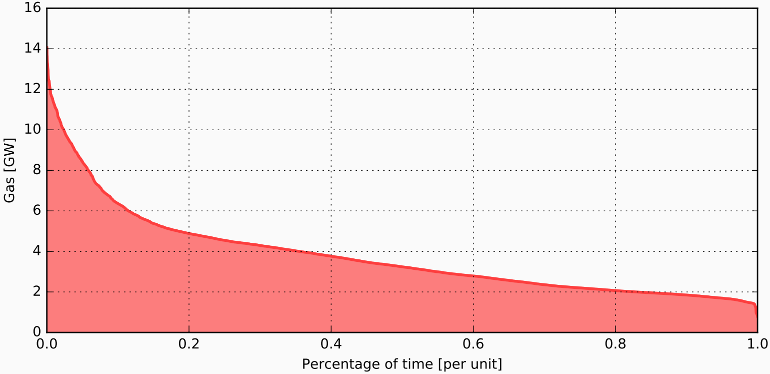
Gas curve

Can do the same for gas output:



Gas duration curve

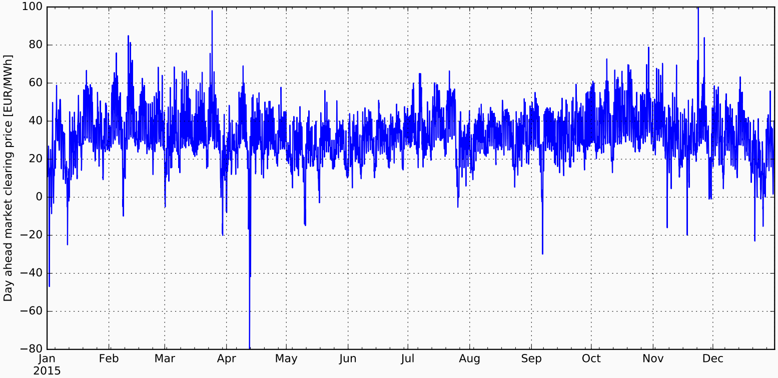
Duration curve is partially flat (for heat-driven CHP) and partially peaked (for peaking plant):



The capacity factor for gas is much lower - more like 20%.

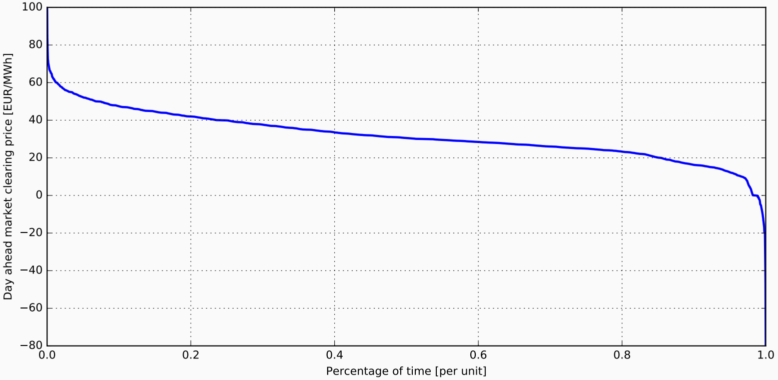
Price curve

Can do the same for price during the year:



Price duration curve

Price duration curve:



Question

Now we are in a position to consider the questions:

- What determines the distribution of investment in different generation technologies?
- How is it connected to variable costs, capital costs and capacity factors?

We will find the price and load duration curves very useful.

Long-run Efficiency: Optimal Investment in Generation with a Single Technology

Definition of long-run efficiency

Up until now we have considered short-run equilibria that ensure short-run efficiency (static), i.e. they make the best use of presently available productive resources.

Long-run efficiency (dynamic) requires in addition the optimal investment in productive capacity.

Concretely: given a set of options, costs and constraints for different generators (nuclear/gas/wind/solar) what is the optimal generation portfolio for maximising long-run welfare?

From an indivdual generators' perspective: how best should I invest in extra capacity?

We will show again that with perfect competition and no barriers to entry, the system-optimal situation can be reached by individuals following their own profit.

Simple example: Single generator type with downward sloping demand

Consider the long-run efficiency from the total system perspective with a single generator type with linear cost function and downward-sloping demand (taken from Biggar-Hesamzadeh pages 21 and 183).

We have to consider marginal costs arising from each unit of production Q and capital costs that arise from fixed costs regardless of the rate of production (such as the investment in building capacity K).

For a given production rate Q and capacity K we have in this simple example a cost

$$C(Q, K) = cQ + fK$$

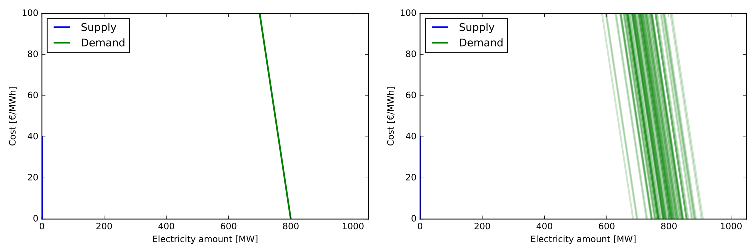
with $0 \leq Q \leq K$, where $C(Q, K)$ has units €/h, c has units €/MWh, Q and K have units MW and f has units €/MW/h ('hourised' capital cost).

Note again: the term fK is constant regardless of production rate Q.

Can't just consider just one load situation

Up until now, in our considerations of short-run efficiency, we've considered just a single demand situation.

Now that we're considering long-term investment, we have to consider many or even all demand situations.

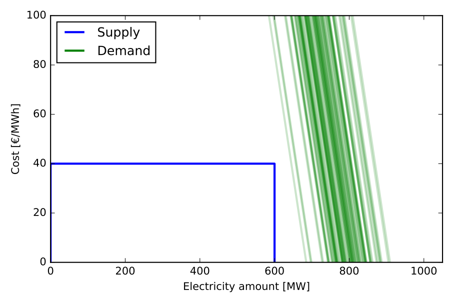


We consider many different utility curves $U_t(Q)$ for different times t, each of which occurs with probability $p_t > 0$, $\sum p_t = 1$.

Simple example: Consumer with downward sloping demand

Suppose the generators have a marginal cost of $c =$40 €/MWh and the downward-sloping demand fluctuates over time.

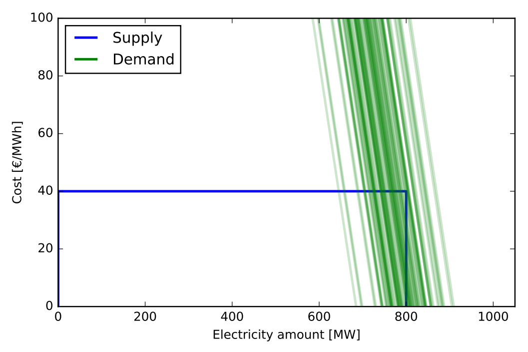
If total generation capacity is always below demand, the demand will set the price at MCB (Marginal Consumer Benefit) and the generators will always earn above their Marginal Generation Cost (MGC):



But then why don't they build more capacity to make even more profit?

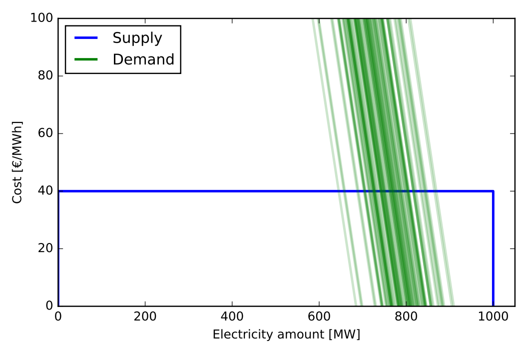
Simple example: Consumer with downward sloping demand

If sometimes the price is set by MCB and sometimes by the MGC then the generators might still earn enough to cover their capital costs:



Simple example: Consumer with downward sloping demand

If generation capacity is so large that it can always cover the demand, regardless of the MCB, then generators will never earn enough money to regain their capital costs, because the price will always be set by the marginal generation cost:



Simple example: optimisation problem

Now consider the maximisation of long-run welfare, including the capital costs:

$$\max_{\{Q^B_t\},\{Q^S_t\},K} \sum_t p_t \left[U_t(Q^B_t) - C(Q^S_t, K)\right]$$

i.e. with cost $C(Q, K) = cQ + fK$ we optimise

$$\max_{\{Q^B_t\},\{Q^S_t\},K} \sum_t p_t \left[U_t(Q^B_t) - (cQ^S_t + fK)\right]$$

given

$$\begin{aligned}
Q^B_t - Q^S_t = 0 \quad &\leftrightarrow \quad p_t\lambda_t \quad && \forall t \\
-Q^S_t \leq 0 \quad &\leftrightarrow \quad p_t\underline{\mu}_t \quad && \forall t \\
Q^S_t \leq K \quad &\leftrightarrow \quad p_t\bar{\mu}_t \quad && \forall t
\end{aligned}$$

(We have taken the liberty to multiply the KKT multipliers by a constant $p_t > 0$, to make the resulting equations easier to read.)

Simple example: KKT

From stationarity we get:

$$\frac{\partial \mathcal{L}}{\partial Q_t^B} \Rightarrow p_t U_t'(Q_t^B) - p_t \lambda_t = 0$$
$$\frac{\partial \mathcal{L}}{\partial Q_t^S} \Rightarrow -p_t c + p_t \lambda_t + p_t \underline{\mu}_t - p_t \bar{\mu}_t = 0$$
$$\frac{\partial \mathcal{L}}{\partial K} \Rightarrow -f + \sum_t p_t \bar{\mu}_t = 0$$

From primal feasibility we get $Q_t^B = Q_t^S = Q_t^*$ and from complementary slackness we have $\underline{\mu}_t^* = 0$, assuming the demand is always positive, and $\bar{\mu}_t^* \geq 0$. Thus we get

$$\lambda_t^* = U_t'(Q_t^*)$$
$$\lambda_t^* = c + \bar{\mu}_t^*$$
$$f = \sum_t p_t \bar{\mu}_t^*$$

Simple example: KKT interpretation

We have

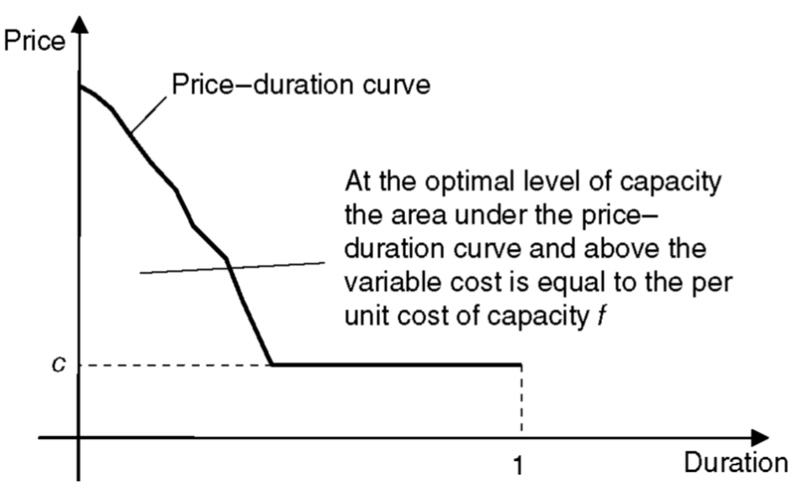
$$
\begin{aligned}
\lambda_t^* &= U_t'(Q_t^*) \\
\lambda_t^* &= c + \bar{\mu}_t^* \\
f &= \sum_t p_t \bar{\mu}_t^*
\end{aligned}
$$

So $\bar{\mu}_t^*$ is the difference between the Marginal Generation Cost (MGC) c and the Marginal Consumer Benefit (MCB) $U_t'(Q_t^*)$.

If the constraint $Q_t \leq K$ is binding, then $\bar{\mu}_t^* \geq 0$.

The optimal investment level happens when the average value of $\bar{\mu}_t^*$, $\sum_t p_t \bar{\mu}_t^*$, is equal to the capital cost f.

Downward Price Duration



Source: Biggar and Hesamzadeh, 2014

Simplified case: inelastic demand

Now consider a simplified case where the demand is inelastic with a very high Marginal Consumer Benefit of V, sometimes called the Value of Lost Load (VoLL) i.e.

$$U_t(Q_t) = \begin{cases} VQ_t & \text{for } Q_t \leq \hat{Q}_t \\ \frac{1}{2} V \hat{Q}_t^2 & \text{for } Q_t > \hat{Q}_t \end{cases}$$

i.e.

$$U'_t(Q_t) = \begin{cases} V & \text{for } Q_t \leq \hat{Q}_t \\ 0 & \text{for } Q_t > \hat{Q}_t \end{cases}$$

In this case the clearing price is binary: it is either set by the marginal generation cost c for the case that demand is lower than generation capacity $\hat{Q}_t < K$ (where $\bar{\mu} = 0$) or by the marginal consumer benefit V if demand exceeds the generation capacity $\hat{Q}_t \geq K$ (where $\bar{\mu} = V - c$).

In this case

$$f = P(\hat{Q}_t > K)(V - c)$$

Example of VoLL from New South Wales, Australia

VoLL depends on consumer:

	Value of Customer Reliability ($/MWh)
Residential	20,710
Small business	413,120
Large business	53,300
Average (volume weighted)	94,990

(1 Australian $ = 0.65 €)

The average value of customer reliability is estimated at close to A$95,000/MWh.

Source: AEMC, 2012, via Jenny Riesz

Simplified case: inelastic demand

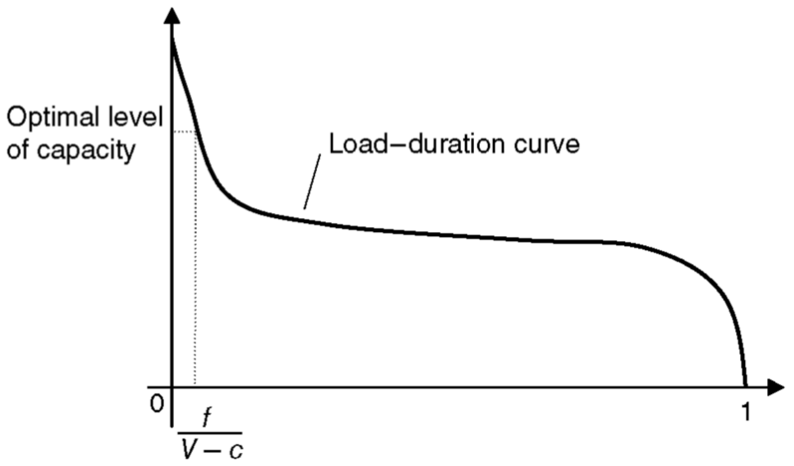
We can turn this around and define the loss of load probability (LOLP) as the probability that voluntary load shedding will be required:

$$LOLP(K) = P(\hat{Q}_t > K) = \frac{f}{V - c}$$

In other words, it will not be efficient to choose a level of capacity that guarantees 100% reliability (that is, the probability that load will be left unserved is not zero).

In an optimal dispatch with inelastic demand up to some valuation V, at the optimal level of generation capacity, there will remain a finite probability that some load will be left unserved.

Inelastic load duration curve



Source: Biggar and Hesamzadeh, 2014

Long-run Efficiency: Optimal Investment in Generation with Multiple Technologies

Extension to multiple technologies

Now consider several different generation technologies T each with different marginal and capital costs, with downward sloping demand.

For a given production rate Q_T and capacity K_T we have for the linear case

$$C(Q_T, K_T) = c_T Q_T + f_T K_T \text{ for } 0 \leq Q_T \leq K_T$$

The maximisation of long-run welfare, including the capital costs, is:

$$\max_{\{Q_t^B\},\{Q_{t,T}^S\},\{K_T\}} \sum_t p_t \left[U_t(Q_t^B) - \sum_T C_T(Q_{t,T}^S, K_T) \right]$$

given

$$\begin{aligned}
Q_t^B - \sum_T Q_{t,T}^S = 0 &\qquad \leftrightarrow \qquad p_t \lambda_t \qquad \forall t \\
-Q_{t,T}^S \leq 0 &\qquad \leftrightarrow \qquad p_t \underline{\mu}_{t,T} \qquad \forall t, T \\
Q_{t,T}^S \leq K_T &\qquad \leftrightarrow \qquad p_t \bar{\mu}_{t,T} \qquad \forall t, T
\end{aligned}$$

Multiple example: KKT

From KKT we have for $Q_t^* = \sum_T Q_{t,T}^*$:

$$\frac{\partial \mathcal{L}}{\partial Q_t^B} \Rightarrow U_t'(Q_t^*) = \lambda_t^*$$
$$\frac{\partial \mathcal{L}}{\partial Q_{t,T}^S} \Rightarrow \lambda_t^* - c_T = \bar{\mu}_{t,T}^* - \underline{\mu}_{t,T}^*$$
$$\frac{\partial \mathcal{L}}{\partial K_T} \Rightarrow f_T = \sum_t p_t \bar{\mu}_{t,T}^*$$

If $c_1 \leq c_2 \leq \cdots \leq c_N$ and $S_n = \sum_{i=1}^n K_i$ then:

If $\lambda_t^* \geq c_T$ we have $Q_{t,T}^* = K_T$, $\bar{\mu}_{t,T}^* = \lambda_t^* - c_T$ and $\underline{\mu}_{t,T}^* = 0$.

If $\lambda_t^* < c_T$ then $Q_{t,T}^* = 0$, $\bar{\mu}_{t,T}^* = 0$ and $\underline{\mu}_{t,T}^* = c_T - \lambda_t^*$.

Multiple example: KKT

$f_T = \sum_t p_t \bar{\mu}^*_{t,T}$ becomes

$$f_T = \sum_{t \text{ s.t. } \lambda^*_t \geq c_T} p_t (\lambda^*_t - c_T)$$

i.e.

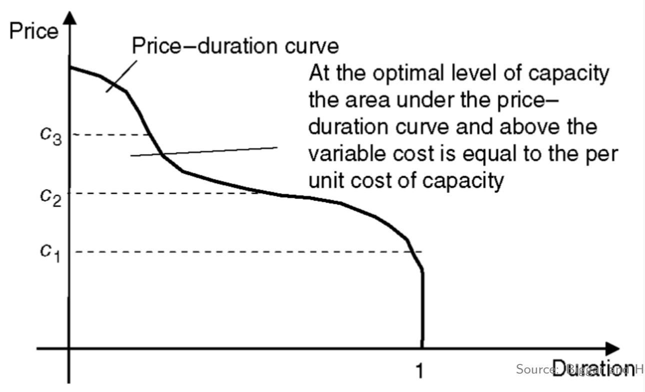
$$\mathbb{E}(\lambda | \lambda \geq c_T) = c_T + \frac{f_T}{P(\lambda \geq c_T)}$$

At the efficient level of capacity, the expected price given that the price is above variable cost is equal to that variable cost plus the fixed cost of capacity discounted by the probability the price is above the variable cost.

The expression on the right-hand side is often known as the long-run marginal cost (LRMC) of a generator of type T. The LRMC is equal to the variable cost plus the fixed cost (per unit of capacity) discounted by the probability that the generator is operating.

Multiple price duration

The optimal mix of generation is where, for each generation type, the area under the price–duration curve and above the variable cost of that generation type is equal to the fixed cost of adding capacity of that generation type.



Source: Biggar and Hesamzadeh, 2014

Multiple generators with inelastic demand

Assume again we have $c_1 \leq c_2 \leq \cdots \leq c_N$ and $S_n = \sum_{i=1}^{n} K_i$ then:

$$\lambda_t = \begin{cases} V & \text{for } Q_t > S_N \\ c_i & \text{if } S_{i-1} < Q_t \leq S_i, \qquad\qquad \text{for } i = 1, \dots N \end{cases}$$

Looking at the area under the price duration curve but above the variable cost, we then find:

$$f_i = (V - c_i)P(Q > S_N) + \sum_{j=i+1}^{N} (c_j - c_i)P(S_{j-1} < Q \leq S_j)$$

Screening curve

These equations can be rewritten recursively using the substitution $\theta_i = P(Q > S_i)$ (see Exercise Sheet 4):

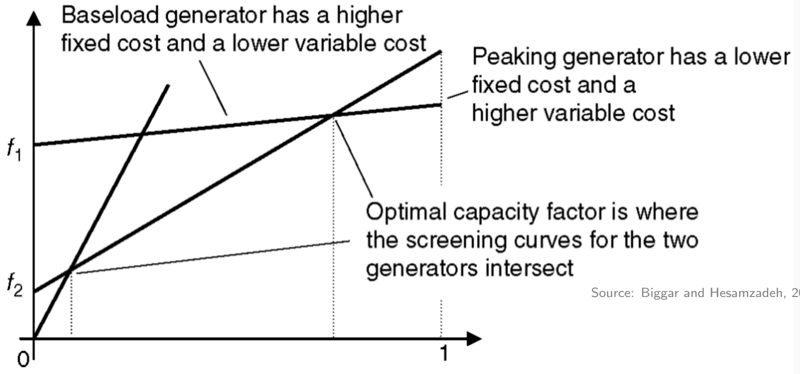
$$\begin{aligned} f_N + \theta_N c_N &= V\theta_N \\ f_i + \theta_i c_i &= f_{i+1} + \theta_i c_{i+1} \qquad \forall i = 1, \ldots N-1 \end{aligned}$$

The first equation can be solved to find θ_N, then the other equations can be solved recursively to find the remaining θ_i. The θ_i correspond to the optimal capacity factors of each type of generator, which correspond to the fraction of time the generator runs at full power.

Screening curve

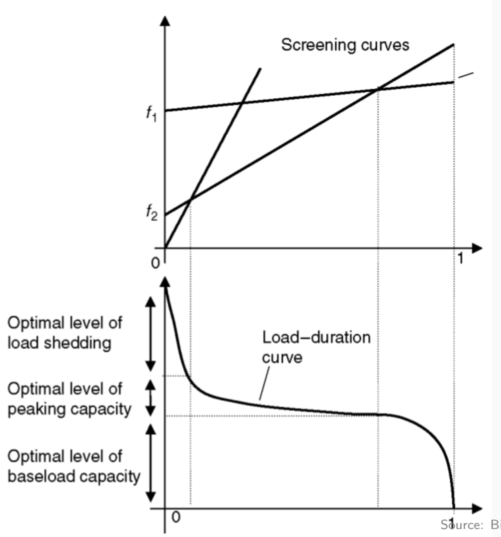
The costs as a function of the capacity factors can be drawn together as a screening curve (more expensive options are *screened* from the optimal inner polygon).

The intersection points determine the optimal capacity factors and hence, using the load duration curve, the optimal capacities of each generator type.



Source: Biggar and Hesamzadeh, 2014

Screening curve versus Load duration



Source: Biggar and Hesamzadeh, 2014

Market-Based Investment in Electricity Generation

Relation to individual investment decisions

Consider a generator who is a price taker for the price λ_t at each time and who can invest even in small amounts of capacity. The expected value of his profit π is then

$$\mathbb{E}(\pi) = \sum_{t:\lambda_t \geq c} p_t(\lambda_t - c) - f$$

The generation entrepreneur will add a small amount of capacity of this generation type if and only if this expression is positive. Similarly, the generation entrepreneur will withdraw a small amount of capacity of this generation type if this expression is negative. Therefore, we can conclude that in a free-entry-and-exit equilibrium, this expression will be zero.

Relation to individual investment decisions

Turning this around we get the same expression for f as before:

$$f = \sum_{t:\lambda_t \geq c} p_t(\lambda_t - c)$$

So: Under the assumptions we just set out, in a free-entry-and-exit equilibrium, with a high level of competition between generators, the equilibrium level of capacity chosen for each type of generator is optimal.

But how does this work in practice?

Grit in the machine 1/2

Several factors make this theoretical picture quite different in reality:

- Generation investment is lumpy i.e. you can often only build power stations in e.g. 500 MW blocks, not in continuous chunks.
- Some older generators have sunk costs, i.e. costs which have been incurred once and cannot be recovered, which alters their behaviour (i.e. the f term is not evenly distributed across all hours)
- Returns on scale in building plant are not taken into account (we did everything linear)
- Site-specific concerns ignored (e.g. lignite might need to be near a mine and have limited capacity)
- Variability of production for wind/solar ignored
- There is considerable uncertainty given load/weather conditions during a year, which makes investment risky; economic downturns reduce electricity demand

Grit in the machine

Several factors make this theoretical picture quite different in reality:

- Fuel cost fluctuations, building delays which cost money
- Risks from third-parties: Changing costs of other generators, political risks (CO_2 taxes, Atomausstieg, subsidies for renewables, price caps)
- Political or administrative constraints on wholesale energy prices may prevent prices from rising high enough for long enough to justify generation investment ("Missing Money Problem")
- Lead-in time for planning and building, behaviour of others, boom-and-bust investment cycles resulting from periods of under- and over-investment in capacity
- Exercise of market power

Episodes of High Prices are an Essential Part of an Energy-Only Market

In an energy-only market (in which generators are only compensated for the energy they produce), the wholesale spot price must at times be higher than the variable cost of the highest-variable-cost generating unit in the market. Episodes of high prices and/ or price spikes are not in themselves evidence of market power or evidence of market failure.

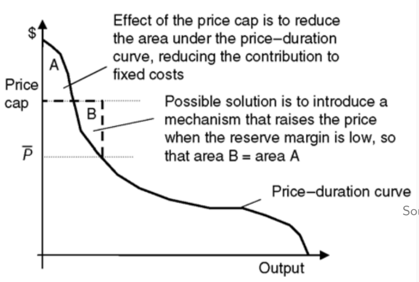
However, there may be political or administrative restrictions on prices going to very high levels (i.e. consumer protection, concerns about market abuse).

Price cap

Some markets implement a maximum market price cap (MPC), which may be below the Value of Lost Load (VoLL) (V for the inelastic case).

In the Australian market, a MPC of A$13,800/MWh for the 2015-2016 financial year is set, corresponding to the price automatically triggered when AEMO directs network service providers to interrupt customer supply in order to keep supply and demand in the system in balance.

This can introduce distortions which make it difficult for some generators to recover costs.



Source: Biggar and Hesamzadeh, 2014

Australian Market in Future

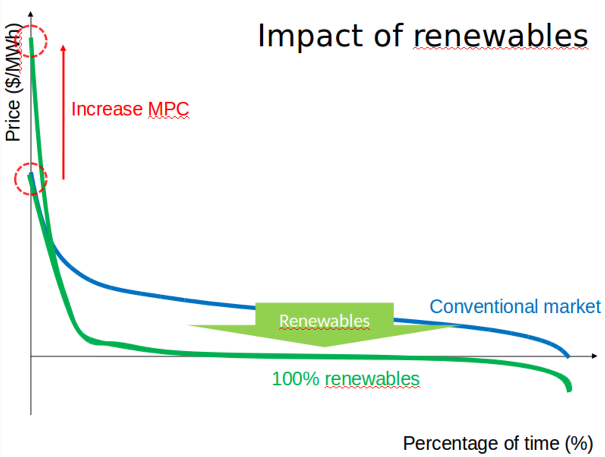
A recent paper on the Australian market: "100% Renewables in Australia: Will a Capacity Market be Required?" by Jenny Riesz, Iain MacGill, 2013, link, analysed the effect high shares of renewables would have on the market and the MPC.

"However, most renewables have very low SRMCs, which in a competitive market is likely to lead to an increasing proportion of low priced periods. This has led to suggestions that a capacity market may be required in the Australian National Electricity Market (NEM). This analysis suggests that existing energy- only market mechanisms in the NEM have the potential to operate effectively in a 100% renewables scenario, but success will rely upon two critical factors: (1) further increase to the already high Market Price Cap (MPC) of A$12,900/MWh. Initial analysis suggests this may need to increase by a factor of six to eight."

(Australia is also interesting for having a market with 5 min dispatch intervals, compared to hourly for Day Ahead in Germany and quarter-hourly for Intraday Market)

Australian Market in Future

MPC may have to increase 6-8 fold to close to VOLL with high shares of RE so that conventional backup generators can recover their costs:



Source: Jenny Riesz

UK price spikes

Price spikes are also caused by reliability problems with other generators, as has happened recently in the UK:

NATIONAL GRID PLC

5:45 PM Tuesday

UK electricity prices spike after power stations break down

By Kiran Stacey, Energy Correspondent

Electricity prices spiked on Monday night, it has emerged, after several old power stations broke down, forcing National Grid to issue an emergency request for more supplies.

Power prices jumped to £1,250 per megawatt hour at one point as the company that runs the UK's electricity network rushed to make sure there was a big enough gap between demand and supply. The normal price per MWh in the summer is about £50.

National Grid said on Tuesday it had issued a so-called "notification of inadequate system margin [Nism]" at 7pm on Monday — only the third time it has done so since 2009.

The last time the company took such action was after similar plant breakdowns last November, when it also paid heavy users for the first time to turn down their equipment.

Source: Financial Times

Comparison with Uber surge pricing

Note the same, since no long-term investments that need to be covered.

Similar in terms of political concerns.

Capacity Markets

A cap on the price in the wholesale market that is binding at times reduces the revenue that generators can earn from the market thereby reducing their incentives to invest. This is known as the ‘missing money’ problem and results in an inefficient mix of generation. The incentives for investment can be restored by making additional payments to generators based on their available capacity. These payments are often determined through a market process known as a ‘capacity market’. Capacity markets represent a response to an existing market defect (the price cap) and are not necessary where the price cap has been removed.

Flexibilise demand

High price spikes can also be ameliorated by adjust demand, which was here assumed to be fairly inelastic.

Flexibilise demand by making it price-responsive.

The technology required to make a sufficient portion of the demand responsive to short-term price signals is not yet available, although some large loads (cement works, etc.) may already implement demand-side management (DSM).

Widespread load disconnections are extremely unpopular and often have disastrous social consequences (accidents, vandalism). They are also economically very inefficient. Their impact can be estimated using the value of lost load (VOLL), which is several orders of magnitude larger than the cost of the energy not supplied. Consumers are not used to such disruptions and it is unlikely that their political representatives would tolerate them for any length of time.

Integrating Renewables in Power Markets

Why do renewables need support at all?

NB: In some markets, they do not.

- Variability (unpredictable revenue leads to investor risk)
- Technology immaturity; needs support in early stage of learning curve
- External benefits not seen by market: lower pollution, CO_2 emissions
- High ratio of capital costs to marginal costs require measures to reduce investment risk

Types of support schemes for renewables

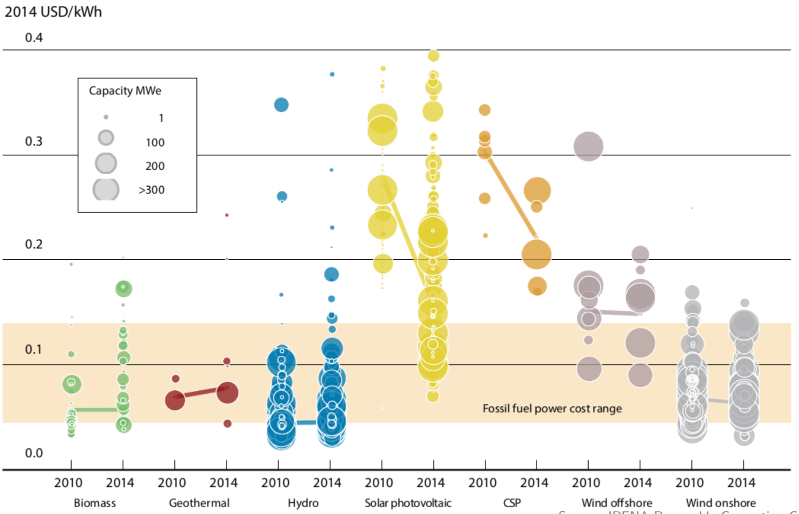
- No support at all (in some markets with good wind speeds, wind can survive on the spot market without any subsidy at all)
- Feed-In Tarriff (FIT, e.g. EEG - fixed payment per kWh for lifetime of project, often technology-specific, can also be region-specific, most popular form of RE subsidy worldwide)
- Contract for Difference (variation on the FIT, where market price is topped up to a set strike price)
- Production Tax Credit (essentially a tax discount corresponding to a direct subsidy per kWh)
- Quotas (E.g. Renewable Obligation Certificates in UK)
- Carbon taxes (increase relative cost of fossil fuels, but doesn't remove investor uncertainty)

Levelised Cost of Electricity (LCOE)

LCOE are full lifecycle costs per MWh including all fixed and variable costs.

However it is FLAWED, doesn't take into account grid, balancing costs from uncertainty and backup costs.

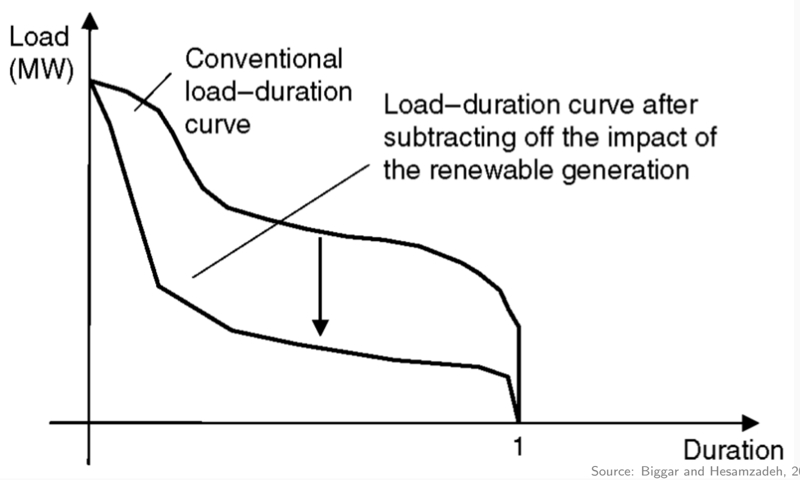
RE costs



Source: IRENA Renewable Generation Costs

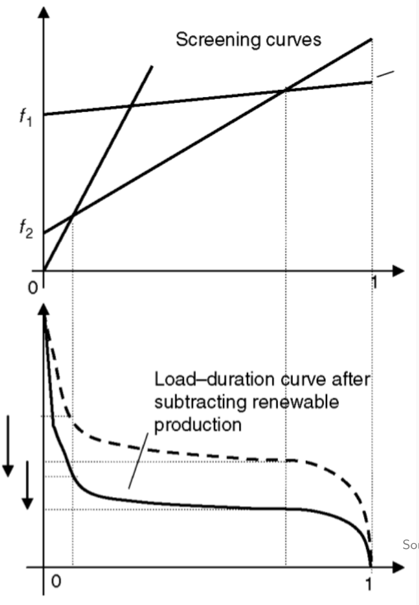
Residual load curve

Simplest: subtract from load.



Source: Biggar and Hesamzadeh, 2014

Residual load curve and screening curve



Source: Biggar and Hesamzadeh, 2014

Copyright