

An options view of designing capacity -- An introductory tutorial

Stefan Scholtes

We use a simplistic example to lead the reader's intuition. Suppose you need decide on the capacity for an infrastructure system. Projected demand for the system is 60,000 units, the operational margin is \$800 per unit demanded, and the capacity costs are \$700 per unit capacity. What capacity should you install to optimise the value of the system?

The naïve view is that for every unit demanded, you will have a net gain of \$100, therefore the system can be operated economically and the optimal decision is to install capacity for the full demand of 60,000 units. The value of the system would then be \$ 6 Million. The problem with this back-of-the-envelope calculation is obvious to every planner: We need to install capacity before demand is observed. What if demand is lower or higher than projected? In the case of higher demand the system will be over utilized, with the consequence of service level deterioration, long waiting times and unsatisfied customers. In the case of low demand, we have installed a "white elephant". A lot of money may be wasted for a system that will never earn its set-up costs.

We need to be able to deal with uncertainties in demand when we make capacity decisions. Suppose, for example, that there is a 50/50 chance of a demand of 40,000 or 80,000 units, respectively. We have still got an expected demand of 60,000. But what is the expected profit? Suppose we install 60,000 units capacity. In the case of an actual demand of 40,000 units we have:

$$\begin{aligned} \text{operating profits} &= \$800 \text{ per unit} * 40,000 \text{ units demanded} = \$32 \text{ Million} \\ \text{set-up costs} &= \$700 \text{ per unit} * 60,000 \text{ units capacity} = \$ 42 \text{ Million.} \end{aligned}$$

This results in a loss of \$10 Million. On the upside, if we have a demand of 80,000 units but can only satisfy 60,000 units, the result is a profit of \$ 6 Million. Given the 50/50 chance of the two scenarios, the expected value of the system is negative \$ 2 Million.

We should therefore choose the capacity of the system so that the expected value is maximised. This is different from maximising the value of the system for the expected demand level, which is naïve approach. Optimal capacity for the example equals the low-scenario demand, and the corresponding expected value is \$ 4 Million,¹ see Figure 1.

Low capacity has obviously the disadvantage of losing on economies of scale. Scale economies will increase the optimal capacity level, without bringing it to the level of the naïve capacity level.

Low capacity, however, has an enormous advantage: It can be seen as a first stage, enabling future expansion if the demand is high. In many infrastructure systems, low-

¹ Notice that the optimal value is lower than the naïve value. This is not a coincidence but a typical phenomenon: Capacity induces a constraint on the system; the system value suffers in case of downside demand and this is not fully balanced out by the possibility of upside demand since upsides are capped at the capacity limit. This can be formalised mathematically through the use of Jensen's inequality.

cost expansions are only possible if the initial design takes account of this possibility. In other words, an often small amount of money needs to be invested early on, when the low-capacity is installed, to enable a cost-efficient later expansion of the system. This is the options view of designing capacity: Build small but invest in the option to expand if needed.

To see how this effects our example, suppose we install some initial capacity now, then observe demand, and can then make an expansion decision which will allow us to capture 50% of the residual demand, i.e. of the demand that we have not satisfied with our initial capacity. It is easy to see that the optimal decision is to install the low-scenario capacity first and then install 50% of the residual capacity later, if the high-demand scenario has occurred. This increases the expected value of the system from \$ 4 to \$ 5 Million. The difference between these two figures is the expected value of the option to expand. Some of this option value may have to be invested now, before the scenario is observed, to enable later expansion. Such initial investment to “buy the option” is often considerably lower than the expected gain from the option.

To see the effect of uncertainty on option value, let us assume the upside and downside scenarios are 60,000 units plus or minus some uncertainty level, measured in percentage deviation; the uncertainty level in the example is thus 33%, leading to a downside of 40,000 and an upside of 80,000 units. Figure 2 shows how the value of the staged system changes as the uncertainty increases. As uncertainty increases, the total value of the staged system decreases but the value of the expansion option, i.e. the difference between the value of the staged system and the value of the non-staged system, increases.²

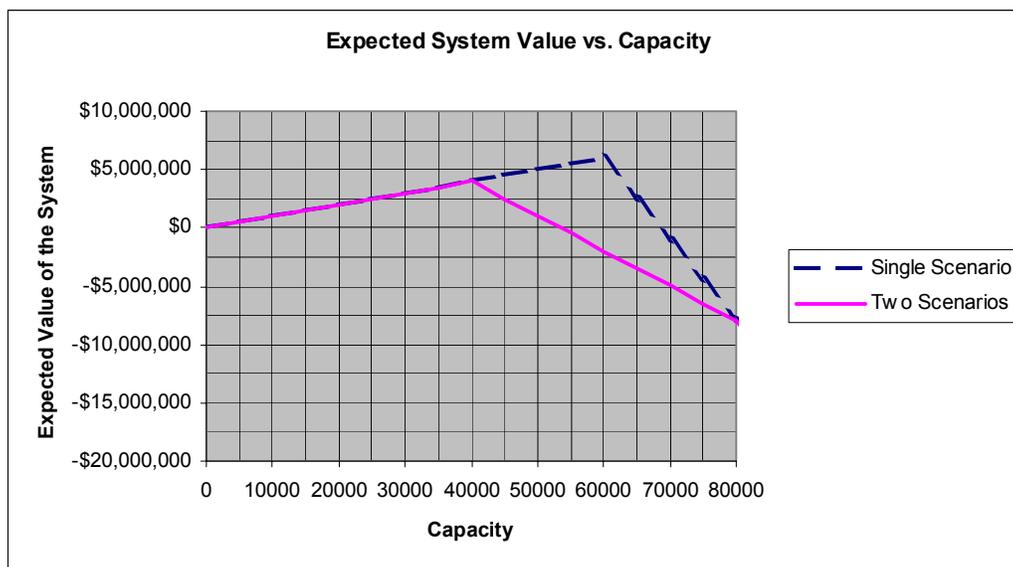


Figure 1

² This is a general phenomenon, formally provable using Jensen’s inequality. The value of an option increases as the underlying uncertainty increases. This balances with the effect of high capacity, which becomes more detrimental to the system value as demand uncertainty increases. In other words, the higher the uncertainty, the lower the initial capacity and the higher the expansion capacity should be.

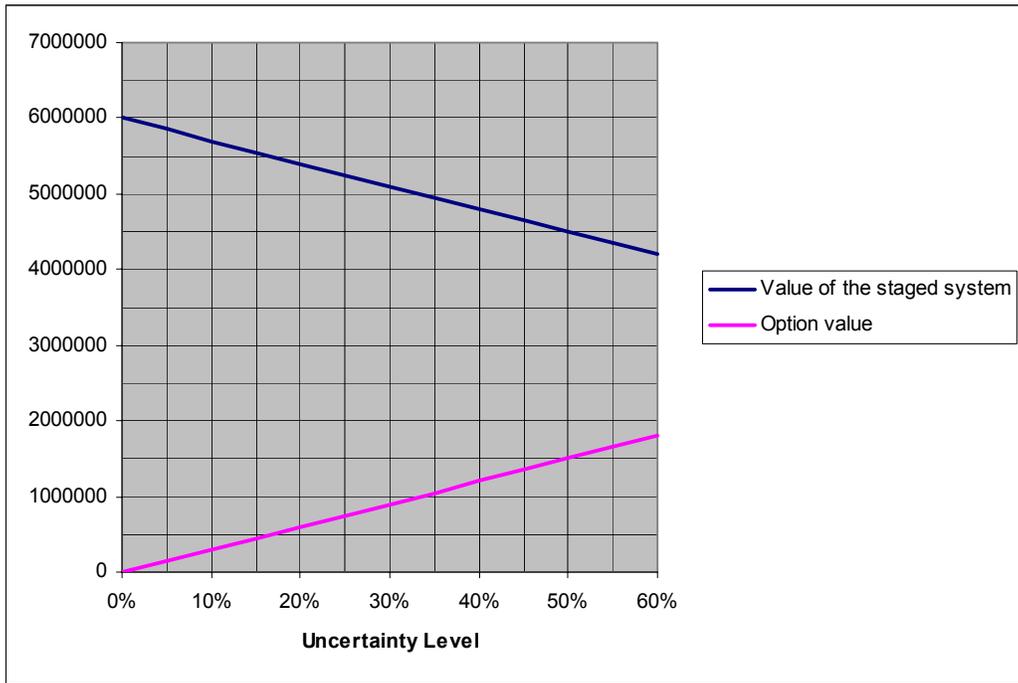


Figure 2