Table of Contents

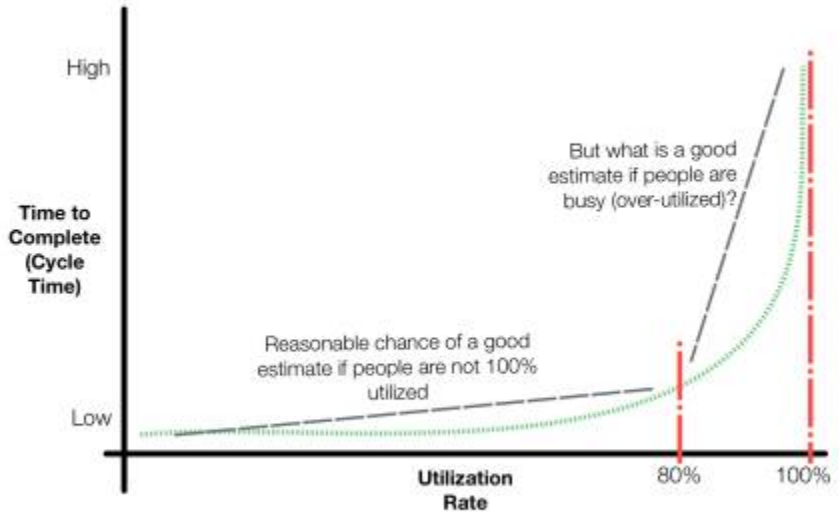
Why Does a High Utilization Rate Make It Impossible to Create Good Estimates?

Or “Why do we have to match capacity with demand in order to have predictable results?”

People often wonder why we are so poor at estimating knowledge work. There are many reasons, but one effect that is often neglected is the effect of high utilization rates. Many product development organizations and IT shops pride themselves on their high utilization rate, thinking that the more busy people are the more we are getting from the dollars we spend on those people, and so the more efficient we will be. Little's Law tells us that while this is true up to a point, there is a point where this is not the case, and the more people and Teams are utilized, the (significantly) longer things will take on average.

> What this means is that if you have full utilized people or Teams you will probably be seriously wrong with estimates you make.

To understand this impact, have a look at the following chart:



This a well known result for all kinds of queues of work, including knowledge work. The chart shows the effect of increasing average utilization on the average time it takes to get something done. As you can see at the beginning of the chart, as you increase utilization, it may take a little longer to do something, but generally not a lot. But as you can see, as you increase the utilization rate of people, the amount of time it takes to complete something on average increases, and after about 80% increases dramatically, exponentially.

Now, instead of understanding the effect of utilization on time it takes to get things done, lets say we want to understand how long it will take to get something done. If we have reasonable utilization rates then we can make a pretty good guess at how long something will take, based on what has occurred in the past.

But if you are trying to estimate work and you have a high utilization rate, then which number would you use - the one represented at the lower end of the curve, or the one at the high. The difference is “orders of magnitude” on the estimate.

What this means is that before you can begin to expect reasonable estimates (discounting all the other things that could go “wrong”) you have to match the demand (what you are asking the people or Team to do) with the supply (the capacity of the people or Team to do the work) and then allow them to step back from 100% utilization. Anything else will just be a guess.

FAQ, estimation, utilization, basics