5. THE ECONOMIES OF SPEED

Chen Zhou

Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology

Contents

We often hear or say “time is money”. In supply chain, it is literally true. If you can provide products with shorter leadtime, or provide faster service, you can charge more or increase revenue. In fact, if everything can be instantaneous, there will be no worry about variations, buffers, inventories, etc. However, when you communicate your ideas of shorter leadtime and faster service to upper level management, you often need to convert these into financial terms. As a supply chain engineer or manager, you have to learn to quantify the economic value of the speed.

1. THE ECONOMIES OF SPEED OF PRODUCTION

Faster production reduces the production throughput time and requites less resources. Faster transportation reduces the time it takes to delivery goods, and requires less transporters. These are direct benefits of speed.

2. THE ECONOMIES OF SPEED WITH RESPECT TO CYCLE AND LEAD

A shorter leadtime can lead to higher prices or competitiveness of the product or services. A shorter leadtime requires shorter production cycles, which lead to lower cycle stocks and safety stocks.

2.1 Cycle time and speed

Most production or service systems produce multiple product or services. This include manufacturing, wave order picking in distribution centers, paper process in offices. The different jobs in batches are produced alternatively. The cycle time determines the WIP, cycle stock and safety stocks. If the cycle times can be reduced, there will be benefits in many aspects.

The leadtime is a function of cycle time in a multiple product system.

$$CT \sim T^* = \sqrt{\frac{2\sum_{j=1}^{n} K_j}{\sum_{j=1}^{n} h_j^{'} D_j}}$$

The lot sizes of each product, which is approximately the cycle stock are

$$Q_j^* = D_j T^*$$

For single product, the average cycle time is how long an EOQ can last before you run below the safety stock level. In industry, a term called "days of supply".

$$CT = \frac{EOQ}{D} = \sqrt{\frac{2K}{hD}}$$

Cycle time is proportional to the square root of setup cost K, $CT \propto \sqrt{K}$. The setup cost is also inversely proportional to unit holding cost and demand. You can reduce the CT by reducing K, or increasing h or D. Increasing h is crazy. However, to capture true cost of h may be practical. Many business only account the opportunity cost of money tied in the inventory as holding cost. In practice, there are cost of storage, obsolescent, price drop (consider electronics). A HP study show that in electronics industry, the opportunity cost can be as low as a quarter of the total holding cost. By capturing the real holding cost, one can reduce the cycle time. Interestingly, increasing the demand can also reduce the cycle time, even when the lot size increases. Therefore, increasing demand has benefits of higher revenue, larger scale as well as shorter cycle times, and lead times.

2.2 Setup costs and setup time

Setup time costs have different components. In manufacturing, there is the cost of changeover. The easiest way to relate is the change of color. One must clean out the current color, add the new color, etc. When process start, the color of the initial product may not be uniform. Therefore, there is waste of materials, in addition to the waste of time.

During setup, more attention is needed. An operator can attend multiple machines in modern manufacturing because they are automatic or semi-automatic. However, the setup would require someone's full attention.

During setup, the resource for processing, such as the machine, is not producing. The cost can be very high if the machine is the bottleneck and is expensive. Typically, the bottleneck machine is expensive. Otherwise, one can always add another machine.

For a new work order, time is needed to process the order. If approval is involved, the wait for approval can be long, and uncertain.

2.3 Lead time and speed

The lead time is the time from when the order is placed to the time when the order is received. In manufacturing, the manufacturer must finish the jobs already planned before it can process the new order, without making exceptions. There is always a cost associated with exceptions, or rush orders. The larger the cycle times for each planned job, the longer the leadtime. Therefore, the leadtime depends on the cycle times of all jobs that have been planned. Therefore, the cycle time is at the main obstacle enemy for speed. Since cycle time $CT \propto \sqrt{K}$, the root cost of slow response time is the setup costs.

In supply chain, transportation time can be significant, from hours to days, to weeks to months. The speed of transportation depends on distance and transport mode: airplane, truck, ships, pipelines have certain speed.

Example: A company orders products from overseas. The demand is 200 unit per month. The ordering and transportation cost is \$400 within the quantity range. The holding cost is \$1 per unit per month. The variance is 60^2 for each type.

1. What are the cycle time and leadtime?
2. If the company can apply various automation to reduce the setup and fixed part of the transportation cost down to \$25 per setup, what should be the cycle time before and after?

Solution

Assumptions:
EOQ assumptions approximately holds.

1. $CT = \sqrt{\frac{2*400}{1*200}} = 2$ months, LT = 1 month.
2. $CT = \sqrt{\frac{2*25}{1*200}} = 0.5$ months, LT = 1 month.

3. SAFETY STOCK REDUCTION DUE TO SPEED

Shorter cycle time and shorter leadtimes reduces the window of uncertainty and therefore the safety stock.

$$SS = \sigma_W z = \sigma\sqrt{W}z = \sigma\sqrt{LT + CT}z$$

Where σ is the standard deviation in a unit time, such as per week. If you can reduce the length of the window of uncertainty, you can reduce the safety stock without additional risk. Consider that the same example above.

Example The company in the example above sets the probability of no shortage in a production cycle to be 98%. The variance in the demand is 60^2. What should be the safety stock before and after.

Solution

Before the setup reduction, the window for the risk is about 2 months because you can produce another lot in 2 months.

$$z = \Phi^{-1}(0.98) = 2.05$$

$$SS \geq \sigma_L z = \sqrt{W}\sigma z = \sqrt{2+1} * 60 * 2.05 = 213.04 = 214$$

Theoretically, you should round it up to 214. However, the cost at 213 will be nearly identical to that in 214.

If K is cut down to \$25/setup, The window of uncertainty is 0.5 month or about 2 business days,

$$SS \geq \sigma_L z = \sqrt{0.5+1}\sigma z = 150.64 = 151$$

4. SAFETY STOCK REDUCTION DUE TO BETTER SHORTER TERM FORECAST

The savings from the variability in leadtimes reduction is based on the assumption that the demand variability per unit time from forecasting remain the same over time. If the leadtime is L week, shown in the figure

The square root came from the independent and identical assumption, i.et.

$$\sigma_L = \sqrt{\sigma_1^2 + \cdots \sigma_l^2} = \sqrt{W}\sigma$$

However, the forecasting model can take additional information to make it more accurate as the forecasting horizon decreases. For weather, there is 10-day, 5-day, daily, and hourly forecasts. The shorter forecasting horizon lead to higher forecasting accuracy They are based on different models with different information. For the hourly and daily forecast, with the modern super computer, the accuracy is so high, many times it is near certainty. The 3 day and 5 day forecasts are much less accurate. The 10-day, monthly or long term forecast are almost speculative.

Shorter term forecast is more accurate is also true in customer demand. There are lot of indicators available for short term forecasting not available for long term forecast, such as market fluctuation, weather influence, disruptions in the global supply chain. Smith S., Sincich conducted an empirical analysis of the effect of length of forecast horizon on population, and concluded, that forecast error grows nearly linearly with the length of forecasting horizon.

If demand also follow the same linear relationship, we have, $\sigma_i = i\sigma_1,\ 1 \le i \le l$, and σ_l is the forecast for *l* periods. If we can shrink the window of uncertainty from *l* to 1, we have

$$\sigma_l = \sqrt{\sum_{i=1}^{l}\left(i\sigma_1\right)^2} = \sqrt{\sum_{i=1}^{l}\left(i\right)^2}\sigma_1 = \sqrt{\frac{2l^3+3l^3+l}{6}}\sigma_1$$

While *l* = 2, 3, or 4, the value in the square root are 5, 14 and 30. Without the adoption of the short term higher forecast accuracy, the standard deviation in each period is the same, we have

$$\sigma_l = \sqrt{\sum_{i=1}^{l}\left(l\sigma_1\right)^2} = \sqrt{\sum_{i=1}^{l}\left(l\right)^2}\sigma_1 = \sqrt{l^3}\sigma_1$$

When *l* = 2, 3, or 4, the value in the square root are 8, 27, 64. Very different from 5, 14 and 30.

Let's revisit the example in the previous section. Before the setup reduction, the window of uncertainty is cut in half. The safety stock was 214. This should be cut by a factor of $\frac{5}{8}$, or

$SS \ge \frac{5}{8} * \ 151 = 94.4 = 95$, which is much lower than 214!

This is a more practical setting of the safety stocks and the resulting savings when more accurate of the shorter term forecast are considered.

Stanley K. Smith, Terry Sincich, "An Empirical Analysis of the Effect of Length of Forecast Horizon on Population Forecast Errors", Demography, Volume 28, Issue 2, pp 261-274, May 1991.

5. The economies of speed at macro level

At macro-level, we consider the impact of faster supply for satisfy human needs.

5.1 The benefits of speed

For the physiological needs, faster supply reduces the wait, reduces the storage requirements, reduces the fund tied up in the storage. In emergency situations, faster supply can deliver air, water, food faster to save lives, reduces suffering. At the safety level, faster can increase the response to assault, shelter needs such as in hurricane.

The speed at any node in a loop, can also enable the speed at other nodes. Imagine a supplier who can change deliveries from every week to every day. This will enable the client to serve its customer faster.

5.2 The drawbacks of speed

However, ever faster motion can potentially increase the severity of accidents. The faster pace in life adds anxiety to some. The food from faster production of certain food are lower in nutrition and taste, lead to more environmental degradation. Examples can be chicken farms, CAFOs, etc. because some of the negative impact are not part of financial accounting system, such as health, pollution to air and water, degradation of land.

In March 2019, FDA lifted an important restriction on genetically engineered salmon to allow AquaBounty, a biotech company with facilities in Cannada and Panama, to start marketing first GMO seafood. They grow twice as fast. Anyone had tried farm raised salmon and wild salmon knows speed matters to the quality of the fish for consumption.

6. Summary on economies of speed

Time is money is true. It can reduce the cost due to wait, inventories, uncertainty. It can also increase revenue. However, the cost of wait, inventories, uncertainty are not as tangible money spent in investment and setup. Once setup, managers love to produce large batches over long cycle times. The root cause of the batching or longer cycle time is the setup cost K, which also impact leadtimes. The cycle time plus leadtime yield the window of uncertainty. The window of uncertainty only have impact if there is uncertainty. Therefore, a successful engineer should also work to reduce the level of uncertainty. One of the major objectives of current drive in big data, business analytics are all effort to reduce the uncertainty.

Exercises

1. A company produces tools in a production line with 10 stations. The takt time in each station is 1 minutes per part on each station.
   1.1. (10) The parts are individually transferred on a conveyor from station to station (no batching, batching, delaying or interruptions) with conveying time of 2 seconds. What is the flow time of a part from when the part enters the first station to when it is completed in the last station.
   Flow time of one part: 10*(60 sec) + 9*(2 sec) = 618 sec
   1.2. (10) The parts are batched into 20 parts totes, and moved together from station to station when a tote is full. The moving time from station to station is 20 seconds. What is the flow time from the time when a full tote arrive at the first station to the time when the last part is completed in the last station?
   Flow time of all totes: 10*(20*(60 sec)) + 9*(20 sec) = 12,180 sec

2. (10*) You are pooling among two sources. The means and variances of the two sources are μ_i and VAR_i, for $i = 1,2$.

$$CV_{pool} = \frac{\sqrt{VAR_1 + VAR_2 + 2COV_{1,2}}}{\mu_1 + \mu_2}$$

   Now, from the correlation coefficient definition, we can find bounds on the value of the covariance of a sum of two random variables $-\sigma_1\sigma_2 \le COV_{1,2} \le \sigma_1\sigma_2$
   2.1. What is the minimum value of the CV_{pool}? Please show the condition for this minimum quantitatively.

   The value of the covariance is minimized when the two variables have a perfect negative correlation: $-\sigma_1\sigma_2 = COV_{1,2}$

   In this case,

$$CV_{pool} = \frac{\sqrt{VAR_1 + VAR_2 + 2COV_{1,2}}}{\mu_1 + \mu_2} = \frac{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2}}{\mu_1 + \mu_2} = \frac{|\sigma_1 - \sigma_2|}{\mu_1 + \mu_2}$$

   The minimum occurs when $\sigma_1 = \sigma_2$, and $CV_{pool} = 0$

   2.2. What is the maximum value of the CV_{pool}? Please show the conditions for this maximum quantitatively.
   The coefficient of variation is maximized when there is a perfect positive correlation between the two sources: $+\sigma_1\sigma_2 = COV_{1,2}$
   In this case,

$$CV_{pool} = \frac{\sqrt{VAR_1 + VAR_2 + 2COV_{1,2}}}{\mu_1 + \mu_2} = \frac{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2}}{\mu_1 + \mu_2} = \frac{\sigma_1 + \sigma_2}{\mu_1 + \mu_2}$$

The maximum occurs when $\sigma_1 = \sigma_2$ and $\mu_1 = \mu_2$, and $CV_{pool} = CV_1 = CV_2$. Since perfect correlation would mean certainty for one of the variables, pooling effect almost always exist, to some extent, even when two random variables are positively correlated.