**Predicting Semiconductor Business Trends**

Chapter 1 Understanding Learning Curves

**1952** AT&T sold licenses to patents and basic know-how for their newly developed solid-state transistor (STT) technology to anyone willing to pay $50K.

**1956** IBM developed the Winchester disk drive but did not have a formal licensing process.

Companies choosing to commercialize AT&T’s technology competed on a level playing field with no initial competitive barriers such as patents or market share pressures. Furthermore, regulations were non-existent and as a result these early companies were a hotbed for new ideas, business approaches, and financial growth. Today, the SCI industry provides the most significant example in recent history of global **free market** economics. Similar to SSTs, the Winchester disk drive market also encompassed businesses operating in nearly ideal conditions of supply, demand, and free market economics. The SCI exhibited life cycles that were longer than the disk drive industry but had the same free market characteristics.

Over time, SSTs and disk drives followed trends in global markets that could be **quantified and used to predict the future**.

**Key Takeaway** The economics and technology of the SCI can be used to predict its future direction and magnitude with a high degree of consistency.

The **transistor learning curve** (**TLC**), which plots transistor unit revenue versus cumulative transistors shipped, **is the most predictable tool in the SCI**. Transistor unit revenue is a proxy for transistor unit cost.[[1]](#endnote-1) The TLC is so prescriptive that SCCs can use it for SCO and 4Ps strategic planning[[2]](#endnote-2).

In the days of germanium and silicon discrete transistors, Texas Instruments could use TLCs to predict the marginal cost on the 100,000th transistor after the 1,000th transistor was produced. With this insight TI could initially price a particular transistor product at a loss to gain market share and parlay this advantage into higher profitability and market influence when they reached future high unit volume sales.

LCs can be **used for any product**. The vertical axis plots the log of unit cost while the horizontal axis plots the log of cumulative units produced over the product’s entire history.[[3]](#endnote-3) The catch is that it must be the **same product being produced over and over again**. For all LCs, unit cost decreases monotonically with more experience or “learning”. Doubling cumulative volume results in a fixed percentage decrease in unit cost. This **percentage decrease is different for different products** but tends to be similar across a broad range of products in an industry like semiconductors.

Many efficiencies contribute to unit cost reduction as volume increases such as spreading depreciation and development costs over larger chunks of volume per unit time. There are a few more considerations when using LCs:

* They must be adjusted for inflation.
* They only apply in free markets characterized by minimal tariffs, trade barriers, taxes, regulations, freight, and other non-direct production costs.

LC have utility for as long as a product is produced. Thus, **LCs never become irrelevant or obsolete**. However, as an industry matures, the cumulative number of products stops moving as quickly to the right on a log scale resulting in prices decreasing less rapidly and diminishing observable effects of improved learning. At some point, monetary inflation surpasses manufacturing cost reduction and unit prices may start increasing in absolute dollars terms.

In 2019, transistor unit revenue decreased at a rate of 32% per year. This rate was applicable to all SC components produced. However, when you partition all SC categorically – memory, logic, analog, etc. – you find that rates by partition vary from the general groups’ average. Specifically, looking at the NAND Flash and DRAM segments, you discover that these memory products’ transistor unit cost and cumulative volume is way ahead of the non-memory segment.[[4]](#endnote-4)

LCs do not specify how decreases in unit cost are achieved, only that it will happen as cumulative volume increases.[[5]](#endnote-5)

In ideal market conditions, as cited earlier, LCs are strictly monotonic but deviations from monotonicity will result from non-ideal market scenarios, such as the 2016 – 2018 DRAM memory shortages. This shortage created price inflation stemming from the supply-demand imbalance.[[6]](#endnote-6) Whenever supply-demand imbalances occur, unit cost **temporarily** moves above or below the LC’s long-term trend line. When supply-demand comes back in balance, unit cost will move to the other side of the LC leading to a “mean reversion” back to the long-term trend.

While LC are primarily used for predicting unit costs, they can also be used to predict other metrics such as improvements in performance, reliability, power dissipation, and any other parameters that benefit from the experience gained as cumulative volume increases.

LCs are also useful for predicting “tipping points” for new technology adoption. The introduction of **compression technology** or **embedded deterministic tests**, in the SC test industry in 2001 illustrates effect. If compression technology had never been introduced the automated testing equipment (ATE) LC would have eventually intersected the TLC. If this had happened, economically it would have meant that it would cost more to test a transistor than to manufacture it. However, with compression technology, the ATE LC steepened and over time made the eventual intersection of ATE-LC and TLC unlikely. Note that the steepening of the ATE-LC reduced the cost for customers buying transistors and at the same time reduced the revenue earned by the ATE industry.[[7]](#endnote-7)

Questions

1. How do LCs behave in relation to product complexity – transistors, toasters, cars, and airplanes?

1. Does swapping revenue with cost per transistor require assumptions about transistor variation margins? [↑](#endnote-ref-1)
2. See BCG report “Perspectives on Experience”. [↑](#endnote-ref-2)
3. How does time factor into an LC? [↑](#endnote-ref-3)
4. Fully understand the visualization and underlying mathematics of LCs in terms of position on the LC, varying slope, and vertical shifts. [↑](#endnote-ref-4)
5. What scenarios would result in this statement being violated? Can violation of the statement be modeled and correlated back to what aspects of business and engineering would cause the violation? [↑](#endnote-ref-5)
6. What are other causes for non-monotonic moves in LCs? [↑](#endnote-ref-6)
7. This last point is a bit tricky to understanding. Think about it more. [↑](#endnote-ref-7)