Strengths and weaknesses of S-curves

Theodore Modis

Growth-Dynamics, Via Selva 8, Massagno, 6900 Lugano, Switzerland

For the last 22 years I have been fitting logistic S-curves to data points of historical time series at an average rate of about 2–3 per day. This amounts to something between 15,000 and 20,000 fits. Combined with the 40,000 fits of the Monte Carlo study we did with Alain Debecker to quantify the uncertainties in logistic fits [1], probably qualifies me for an entry in the *Guinness Book of Records* as the man who carried out the greatest number of logistic fits.

It hasn't all been fun and games. There have also been blood and tears and not only from human errors. There have been what I came to recognize as “misbehaviors” of reality. I have seen cases where an excellent fit and ensuing forecast were invalidated by later data. But well-established logistic growth reflects the action of a natural law. A disproved forecast is tantamount to violating this law. A law that becomes violated is not much of a law. What is going on? There is something here that needs to be sorted out.

1. Sources of difficulty

Logistics have been used extensively in the widest range of applications. When I confronted Cesare Marchetti – celebrated veteran of logistics – with the idea that *anything* that grows will fit a logistic, he replied, “Yes, of course. Anything that begins and ends an existence will fit a logistic.”

Marchetti was being frolicsome and provocative as usually. A straight line can be fitted to an S-curve if we ignore – with some justification – the beginning and end periods.[1] But to establish the action of a natural law such as logistic growth, it requires a little more than that.

E-mail address: tmodis@yahoo.com.

[1] Deviations are expected before and after logistic growth [2].

1.1. When can we use S-curves?

The indispensable ingredients for natural growth are the ability of a "species" to multiply and a finite niche capacity. Of course it is understood that we are dealing with a well-defined species, which is not something obvious. In generalizing the concept of a species we may run into difficulties deciding what constitutes one. For example, the Soviet Union constituted a species more defensibly than the USA. After their revolution the Bolsheviks inherited a multinational empire that was created by Tsarist expansion over four centuries. But this mosaic of cultures did not change appreciably during the lifetime of the Soviet Union, which was isolated from the outside world. The USSR was a rather stable "species" that came into existence at a well-defined moment and occupied a well-defined niche in absence of excessive mutations. In contrast, the USA took 150 years to take shape, and different-culture people never ceased to pour in it and continue to do so even today. As a consequence, the evolution of the USSR can be described better by an S-curve than that of the USA [3].

In general, a forecaster must have good reasons to believe that an S-curve fit is appropriate. Logistic growth is natural growth in competition. Therefore besides a species capable of multiplying there should also be competition for a limited resource. Alternatively, an *a posteriori* argument for using an S-curve can be made if a good fit has been achieved. If the data do fit well a large section of an S-curve, then one can argue for logistic growth just because "if it walks like a duck and quacks like a duck, it must be a duck." Still, one will gain confidence if one can identify a well-defined species "genetically" stable that is competing for a limited resource.

1.2. When should we fit cumulative data?

A frequent point of misunderstanding and confusion is whether a forecaster should fit an S-curve to the raw data or to the cumulative number. Here the forecaster must exercise wise judgment. What is the species and what is the niche that it is growing into? To the frustration of business-school teachers there is no universal answer. When forecasting the sales of a new product it is rather obvious that one should fit the cumulative sales because the product's market niche is expected to eventually fill up. It is also obvious that one should *not* fit the cumulative number concerning forecasts on the evolution of the Internet or of the Earth's population. But if we are trying to forecast the evolution of Microsoft, should we fit the evolution of its revenue or of its cumulative revenue from the beginning of Microsoft's existence?

The question that needs to be answered is whether Microsoft's existence is a transient phenomenon (like a product or a fad, if on a larger timeframe) or whether Microsoft is here to stay with us "forever" like the Internet. In the former case, the forecaster should fit an S-curve to the cumulative revenue; in the latter he or she should fit it to Microsoft's raw revenue numbers. The forecasts obtained are dramatically different.

My view is that Microsoft has been filling a niche, which could well be half full at present. Unless Microsoft undergoes a major change (tantamount to becoming a different species) I would favor the forecast on the right of Fig. 1.

1.3. Fitting techniques and uncertainties

The best technique for fitting an S-curve on data points involves the minimization of a rigorously constructed Chi-Square function. This procedure permits the estimation of confidence levels. Unfortunately, the construction of such a function requires knowledge of the complete and correct uncertainty for each data

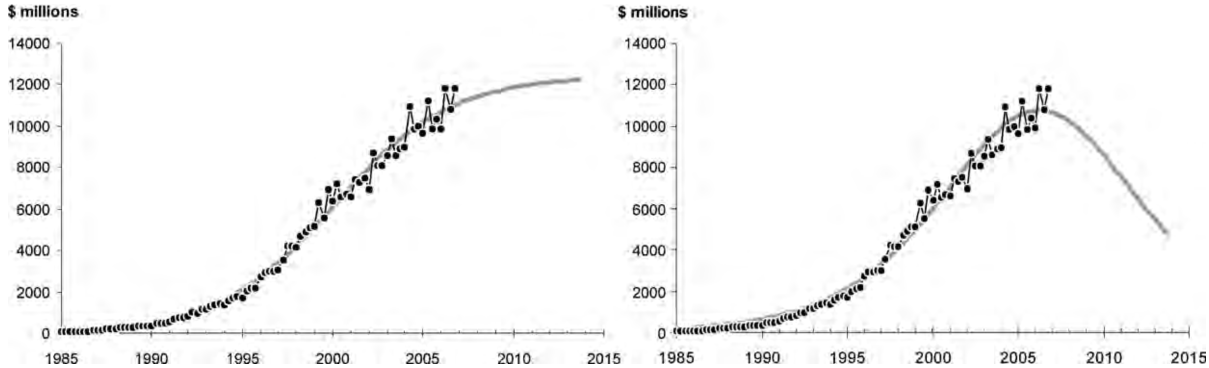

Fig. 1. The corrected-for-inflation quarterly revenue of Microsoft has been fitted by S-curves. On the left we see raw revenue data. On the right we see the life cycle of an S-curve fitted on cumulative revenue.

point. Had we been studying outcomes in casino roulettes, this uncertainty would simply be the statistical error. But for such cases as product sales, annual revenues, discovery of oil reserves, and the productivity of artists as measured by their cumulative artistic achievement, it is unrealistic to rigorously determine an uncertainty for each data point. When I went back to make an update of the AIDS-victims curve in the US ten years later, I found my old data restated with values that differed up to 80%!

One may think that there exist data sets with no uncertainties whatsoever, for example, the data on Nobel laureates. But the uncertainties needed for the Chi-Square do not concern documentation and reporting but the evolution of a natural law. During World War II there were no Nobel prizes given at all. There is no uncertainty about this fact; nevertheless it constitutes a large fluctuation below the trend in the evolution of the number of Nobel laureates. Also sometimes there were many laureates who share one prize and other times few. It is the deviations (fluctuations) from the natural-growth pattern that reflect the size of the uncertainty on the data points. For the calculation of the Chi-Square one can generally make only rough estimates for such uncertainties.

Another aspect overlooked by those forecasters who take the trouble to quote an error on their forecasts is the confidence level. A ± 10% error on a prediction may sound good until we realize that it usually refers to a confidence level of 68%. In other words, our result is 10% accurate in less than 7 out of 10 times. In physics research a one-standard-deviation “discovery” is considered worthless.

In our Monte Carlo study we give tables that link the expected forecasting error to the uncertainties of the data points, to the confidence levels, and to the range of the S-curve that the data point cover. For example, if the data cover about half of the fitted S-curve and the uncertainty on each data point is of the order of ±10%, then the error on the forecasted ceiling will be ±20% with a confidence level of 95% [1].

1.4. Beware of a bias toward a low ceiling

No matter what program one uses, the fitted S-curve will flatten toward a ceiling *as early and as low as* it is allowed by the constraints of the fitting procedure. All fitting programs yield logistic fits that are generally biased toward a low ceiling. Uncertainties on the data points accentuate this bias by permitting larger margins for the determination of the S-curve parameters.

To compensate for this bias I usually make several fits with different weights on the data points.[2] I then keep the answer that gives the highest ceiling for the S-curve (most often obtained by weighting heavily the recent historical data points). All this within reason! Here again the forecaster must exercise wise judgment.

Despite my careful compensation for the low-ceiling bias, I obtained a forecast for US Nobel laureates that proved too low by 18% ten years later. There was no violation of the natural law; I had simply been unlucky. Golden and Zantek pointed out the inaccuracy of my forecast with fanfares [4]. They failed to notice, however, that 18% was smaller than the expected error for 95% confidence level and 10% uncertainty per data point.

1.5. Appreciation of the forecasts

No niche in nature was ever left partially completed under *natural circumstances* and that is why logistics possess forecasting power. The catchall phrase natural circumstances has been misunderstood more often than not.

> "You mean that is what's going to happen if we do nothing?" was the typical reaction of marketers when I presented to them my logistic forecasts.
>
> "No," I would reply, "that is what is going to happen if you keep doing the same type of actions that you have been doing."
>
> "What if the competition comes up with a powerful new product?"
>
> "How many times has the competition come up with powerful new products during our historical window?" I would challenge them.
>
> "Many times," was the invariable answer.
>
> "Than the launching of powerful new products by competitors is a *natural circumstance*."

In fact, natural circumstances are all the types of events that have taken place during the historical window covered by the data. In contrast, if there is a nuclear war, or other never-seen-before incident, then one has a valid reason to doubt the forecast because the natural law is indeed being violated.

2. Intuitive use

Do the arguments presented in Section 1 diminish the usefulness of S-curves?

Not at all! S-curve forecasts supplemented by correct error estimates are neither magical nor useless; they constitute a statement of fact: how a natural law evolves. But S-curves can also be used qualitatively to obtain rare insights and intuitive understanding.

---

[2] In the fitting procedure the weights of the data points are inversely proportional to their uncertainties.

2.1. Cascading S-curves and the four phases of a growth cycle

S-curves are known to cascade with a new one beginning where the last one leaves off. New products replace old products just as new technologies replace old technologies. But even in careers and personal relationships one may find successions and replacements as a new one begins when the old one ends. Each S-curve has its own life cycle, undergoing good and bad seasons. Successive life cycles give rise to a landscape reminiscent of the wavy sinusoidal pattern of the harmonic oscillator. I have studied the conditions for a "harmonious" succession of S-curves, where the sum of their life cycles perfectly matches a sinusoidal wave, see Fig. 2. It turns out that optimum introduction of change occurs when 3% completion of the new process coincides with about 90% completion of the old process [5]. This then can serve as a quantitative rule for the perennial quest of *just-in time* replacement.

A long wave can be seen as a succession of S-shaped steps reflecting a sequence of niches with a new one opening up as soon as the old one has been filled. Every period of the wave will feature four "seasons" of growth.

I have tabulated a large set of behaviors, each one best suited for a particular season. Becoming conservative – seeking no change – is appropriate in the summer when things work well; this is the time to strive for excellence and total quality management. But excellence drops in second place during the difficult time of the winter when fundamental change must take place; one should now become entrepreneurial and innovative.

Learning and investing are appropriate for spring, but teaching, tightening the belt, and sowing the seeds for the next season's crop belong in the fall.

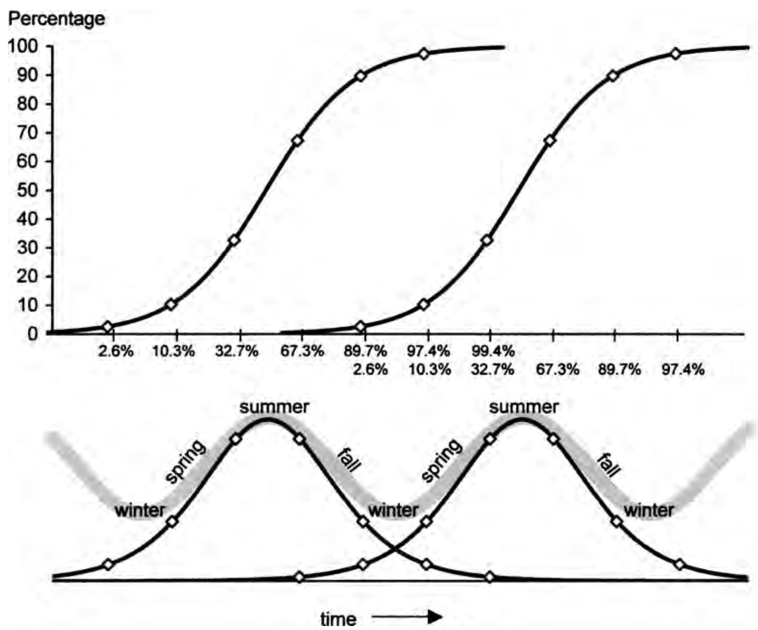

Fig. 2. The two S-curves represent two consecutive natural-growth processes. The corresponding life cycles below show how their rates of growth go through "seasons." The sum of a series of such life cycles matches perfectly the thick gray line, which is a sinusoidal (harmonic) wave.

The seasons metaphor can illuminate Mozart's career. When Mozart's productivity began waning in the early 1780s he found himself in a “fall” season. He appropriately – if subconsciously – tried to sow the seed for a different crop. His *Dissonant Quartet in C Major* K465 (1785) has been cited as evidence for Mozart's possible evolution, had he lived. I consider this an unlikely scenario. The learning curve of music lovers of that time could not accommodate the kind of music that became acceptable more than 150 years later. Mozart would have soon stopped exploring musical directions that provoked public rejection.

S-curves can shed special light on the overall evolution of music through the eons.

2.2. The music curve

The seasons metaphor dictates that in spring the focus is on *what* to do, whereas in fall the emphasis shifts to the *how*. The former appears early in the growth process, the latter late. The evolution of classical music can be visualized as a large-timeframe S-curve beginning its development sometime in the fifteenth century and reaching a ceiling in the twentieth century. In Bach's time composers were concerned with what to say. The value of their music is on its architecture and as a consequence it can be well interpreted by any instrument, even by simple whistling. But two hundred years later composers such as Debussy wrote music that depends crucially on the interpretation, the how. Classical music was still “young” in Bach's time but was getting “old” by Debussy's time (when you hear people say that they need to focus on the how, you can understand that they are referring to something that is getting old,) see Figs. 2 and 3. One may wonder why Chopin is more popular than Bartók. Chopin composed during the “summer” of music's S-curve when public preoccupation with music went over a maximum. Around that time composers' efforts were rewarded more handsomely than today. The innovations they made in music were assimilated by the public within a short period of time because the curve rose steeply and would rapidly catch up with each innovation. But today the curve has flattened and gifted composers are given very limited space, see Fig. 3. Even if they make only small innovations they find themselves above the

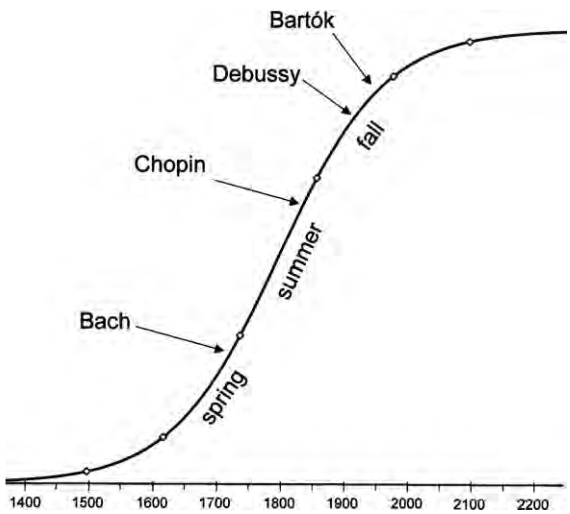

Fig. 3. This S-curve is constructed by only qualitative arguments and yet it seems accurate and informative. The vertical axis could be something like “importance”, “public appreciation”, or “public preoccupation with music”, (always cumulative).

curve and there won't be any time in the future when the public will appreciate their work. On the other hand, if they don't innovate, they will not be saying anything new. In either case today's composers will not be credited with an achievement.

3. S-curves as building blocks

S-curves enter as modular components in many intricate natural patterns. One may find S-curves inside S-curves because logistics portray a fractal aspect. A large S-curve can be decomposed in a cascade of smaller ones [6]. One may also find chaos by rendering the logistic equation discrete [7]. Finally, S-curves sit in the heart of the Volterra–Lotka equations, which describe the predator–prey relations and other forms of competition.

In its full generality, the logistic equation, in a discrete form, with cross terms to account for all interrelations between competing species, would give a complete picture, in which growth in competition, chaos, self-organization, complex adaptive systems, and other such academic formulations ensue as special cases [8].

4. Conclusions

The original reading line on the cover of my first book was, “Society ticks like a clock leaving a telltale signature.” The signature I had in mind was the logistic S-curve. It has a ubiquitous way to enter just about every aspect of our lives. As for using it in forecasting, it can be very useful when done judiciously. The forecaster's involvement is crucial because he or she plays a decisive role. The forlorn hope of marketers and artificial-intelligence designers for a push-button solution whereby everyone will be able to trivially obtain the same forecast without the need to exercise judgment is a utopia. The forecaster's ultimate test is the goodness of his or her forecasts not the elegance or the easiness of the method used.

References

[1] A. Debecker, T. Modis, Determination of the uncertainties in S-curve logistic fits, Technol. Forecast. Soc. Change 46 (1994).

[2] T. Modis, A. Debecker, Chaoslike states can be expected before and after logistic growth, Technol. Forecast. Soc. Change 41 (2) (1992).

[3] T. Modis, Life cycles — forecasting the rise and fall of almost anything, The Futurist (September–October 1994) 20–25.

[4] B. Golden, P. Zantek, Inaccurate forecasts of logistic model for Nobel Prizes, Technol. Forecast. Soc. Change 71 (4) (2004).

[5] T. Modis, The normal, the natural, and the harmonic, Technol. Forecast. Soc. Change 74 (3) (2007) 391–399.

[6] T. Modis, Fractal aspects of natural growth, Technol. Forecast. Soc. Change 47 (1) (1994) 63–73.

[7] T. Modis, A. Debecker, Chaoslike states can be expected before and after logistic growth, Technol. Forecast. Soc. Change 41 (2) (1992) 111–120.

[8] T. Modis, Genetic re-engineering of corporations, Technol. Forecast. Soc. Change 56 (2) (1998) 107–118 (This article received the journal's Outstanding-Paper-of-the-Year Award).