1-2016

Bell-Shaped Curve for Productivity Growth: An Explanation

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-16-04

Bell-Shaped Curve for Productivity Growth: An Explanation

Olga Kosheleva and Vladik Kreinovich
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

A recent analysis of the productivity growth data shows, somewhat surprisingly, that the dependence of the 20-century productivity growth on time can be reasonably well described by a Gaussian formula. In this paper, we provide a possible theoretical explanation for this observation.

1 Formulation of the Problem

An empirical fact. A recent book [2] shows that, when averaged over decades, the productivity growth $n(t)$ in the US from 1900 until now follows a bell-shaped Gaussian curve $n(t) = c_0 \cdot \exp(-k_0 \cdot (t - t_0)^2)$, for appropriate values c_0, k_0, and t_0 (see also [1]).

This fact is somewhat surprising. Such curves normally describe how the probability of a certain value x depend on this value. It is somewhat surprising to find a similar curve in the description of how productivity growth depends on time.

What we do in this paper. In this paper, we provide a possible explanation for this surprising phenomenon.

2 Our Explanation

Main idea. Eventually, productivity growth can be traced to new inventions. However, the appearance of a new invention does not immediately boost the productivity growth:

- inventions are usually formulated in somewhat abstract theoretical terms, and therefore

- it takes quite some time and effort to adopt and modify the original invention so that it would start boosting up productivity.

From the main idea to precise formulas: first approximation. Let $c(t)$ be the number of inventions per time unit. An invention made at time t_1 leads to a productivity boost at some later time $t > t_1$. The corresponding time delay $\Delta t = t - t_1$ is, in general, different for different inventions. The exact value of this delay is unpredictable, so it makes sense to consider this delay as a random variable.

Let $\rho(\Delta t)$ be the probability density that describes the probability distribution of different delays. The boost of productivity at moment t_1 can be caused by inventions made at different past moments of time $t = t_1 - (t_1 - t)$. At each moment t, $c(t)$ inventions were made, and the probability of each of these inventions leading to a productivity boost at moment t_1 – i.e., the fraction of those inventions that lead to a productivity boost at moment t_1 – is proportional to $\rho(t_1 - t)$. Thus, at moment t_1, the increase in productivity caused by these inventions is proportional to the product $\rho(t_1 - t) \cdot c(t)$. The overall productivity increase at moment t_1 can be obtained if we add up all the increases corresponding to all moments t. Thus, the productivity growth $n(t_1)$ at moment t_1 is proportional to the sum $\int \rho(t_1 - t) \cdot c(t)\, dt$.

A more detailed analysis. In the above description, we considered a transition from an invention to productivity growth as a single stage. In reality, this transition is very complex, it contains many stages.

First, we need to transform the original raw ideas into solid science. This may also involve several steps. For example:

- first, we test the idea on a small sample, to provide a proof of concept;
- once this testing confirms the idea, we get the funding to test it on a larger sample, etc.

Often, during this testing, it becomes necessary to modify and update the original idea.

All this constitutes research. Once the research is done, we need to work on development, to think of how the original research ideas can be best implemented in an efficient way. This may also take several steps:

- first we implement it on a small scale (as computers and additive manufacturing were),
- then we find the way to make it more widely spread, etc.

How this informal analysis changes the corresponding mathematical model. According to the above analysis, instead of a *single* large delay Δt, it is more appropriate to consider it as the sum of *several* (much smaller) delays corresponding to different stages of the transition from the original invention to the increase in productivity: $\Delta t = \Delta t_1 + \ldots + \Delta t_m$.

Different stages are reasonably independent. Thus, if we denote by $\rho_i(\Delta t_i)$ the probability distribution corresponding to the i-th stage, then:

- the average number $c_1(t_1)$ of inventions that finished the first stage at moment t_1 will be proportional to $\int \rho_1(t_1 - t) \cdot c(t)\,dt$;
- the average number $c_2(t_2)$ of inventions that finished the second stage at moment t_2 is proportional to $\int \rho_2(t_2 - t_1) \cdot c_1(t_1)\,dt_1$, i.e., to

$$\int \rho_2(t_2 - t_1) \cdot \rho_1(t_1 - t) \cdot c(t)\,dt\,dt_1;$$

- the average number $c_3(t_3)$ of inventions that finished the third stage at moment t_3 is proportional to $\int \rho_3(t_3 - t_2) \cdot c_2(t_2)\,dt_2$, i.e., to

$$\int \rho_3(t_3 - t_2) \cdot \rho_2(t_2 - t_1) \cdot \rho_1(t_1 - t) \cdot c(t)\,dt\,dt_1\,dt_2;$$

- etc.
- finally, the productivity growth $n(t_m)$ at moment t_m is proportional to the average number of inventions that finished all m stages at this moment, i.e., to

$$\int \rho_m(t_m - t_{m-1}) \cdot \ldots \cdot \rho_1(t_1 - t) \cdot c(t)\,dt\,dt_1\,\ldots\,dt_{m-1}.$$

This formula helps explain the Gaussian shape. From the mathematical viewpoint, this formula means that the desired function $n(t)$ is proportional to the convolution of a large number of functions $c(t)$, $\rho_1(\Delta t_1)$, $\rho_2(\Delta t_2)$, ..., and $\rho_m(\Delta t_m)$. This is exactly the same formula that describes how the probability density function (pdf) of the sum of many independent random variables depends on the pdfs of the components of this sum.

For random variables, there is a Central Limit Theorem, according to which, under some reasonable conditions, the distribution of the sum of many relatively small random variables is close to Gaussian; see, e.g., [3]. In terms of convolution, this means that, under the corresponding conditions, the convolution of a large number of non-negative functions is close to the Gaussian bell-shaped curve.

Since, according to our argument, the productivity growth function $n(t)$ can be described as such a convolution, it then follows that this function is indeed close to Gaussian. In other words, we have the desired explanation.

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

References

[1] E. Glaeser "Those were the days", *The Wall Street Journal*, January 16–17, 2016, pp. C5 and C7.

[2] R. J. Gordon, *The Rise and Fall of American Growth*, Princeton University Press, Princeton, New Jersey, 2016.

[3] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2011.