Evolutionary Model of Moore’s Law

Kaldasch, Joachim

EBC Hochschule Berlin

9 March 2014
Evolutionary Model of Moore’s Law

Joachim Kaldasch

EBC Hochschule Berlin, Alexanderplatz 1, 10178 Berlin, Germany

Correspondence should be addressed to Joachim Kaldasch; joachim.kaldasch@international-business-school.de

Received 6 January 2014; Accepted 6 February 2014; Published 9 March 2014

Copyright © 2014 Joachim Kaldasch. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Moore suggested an exponential growth of the number of transistors in integrated electronic circuits. In this paper, Moore’s law is derived from a preferential growth model of successive production technology generations. The theory suggests that products manufactured with a new production technology generating lower costs per unit have a competitive advantage on the market. Therefore, previous technology generations are replaced according to a Fisher-Pry law. Discussed is the case that a production technology is governed by a cost relevant characteristic. If this characteristic is bounded by a technological or physical boundary, the presented evolutionary model predicts an asymptotic approach to this limit. The model discusses the wafer size evolution and the long term evolution of Moore’s law for the case of a physical boundary of the lithographic production technology. It predicts that the miniaturization process of electronic devices will slow down considerably in the next two decades.

1. Introduction

The exponential growth of the number of transistors in integrated electronic circuits is known as Moore’s Law [1]. In order to derive Moore’s law, the paper develops an evolutionary model of successive production technologies. A large part of the evolutionary economic literature is devoted to the study of the economic consequences of innovation and technical change [2]. Initiated by Nelson and Winter [3], the Neo-Schumpeterian literature focuses on production technologies and routines, while the technological evolution is suggested to be governed by a fitness related to some sort of cost functions. It studies in particular technological and organizational opportunities [4, 5], while the impact of the market dynamics plays only a secondary role [6]. It is mediated by assuming a productivity-driven exogenous demand function or by postulating a population dynamics in evolutionary game theory. An attempt to take the market dynamics explicitly into account was based on the time evolution of firms [7].

In the presented evolutionary model, however, not firms but product variants manufactured with different production technologies are regarded to suffer from a preferential growth process. These product variants are the elementary units in a preferential growth process recognized independently by a number of researchers [8–13]. It is based on the idea that the demand side of a market is determined by the selection behaviour of potential buyers, while a key variable is the price. During the reproduction process, in order to make profit, the supply side responds on purchase decisions of the buyers and preferentially reproduces the best-selling products. Which means that manufacturers increase the output of the best-selling product variants and reduce, respectively, and terminate the production of the slow sellers. As a result, preferentially selected products increase their market shares in time, while less favoured variants are replaced.

Similar to the Neo(Post)-Schumpeterian literature, the presented evolutionary model suggests that as a result of this evolutionary adaptation process of products, also the production technologies suffer from a preferential growth process [14]. The presented model relates the market dynamics of products manufactured with different production technologies with the replacement evolution of successive technology generations. The theory further establishes the time evolution of a cost relevant characteristic of the production technology for the case that this characteristic is constrained by a technological limit. It is shown explicitly that in the technological evolution this limit is approached asymptotically.

The model is exemplified at empirical results of memory chips used in electronic consumer goods, the so-called
Dynamic Random Access Memory (DRAM). They find their main application in computers primarily in PCs and servers. There is an extensive literature available on the technological evolution of DRAMs [10, 15–19]. The production technology to create memory chips is quite complex. Moore discovered that the number of transistors in integrated electronic circuits had doubled approximately every two years. However, as already realized by Moore and other researchers, the miniaturization process must have a physical limit [20, 21]. Since Moore’s law is taken as an empirical fact and is regarded by a physical limit.

The paper is organized as follows. In the next section, a preferential growth model of production technology generations is established. It takes the case of a technological limit of a cost relevant characteristic into account. In order to test the theory, two examples of the technological evolution of the DRAM semiconductor industry are studied: the wafers size and Moore’s law. The paper ends with a conclusion.

2. The Model

We want to consider a competitive commodity market. The market is characterized by product variants manufactured with production technologies. The commodity demand at time step \( t \) is determined by the total demand rate \( d(t) \). It is the result of first and repurchases of buyers. We assume that the total unit sales \( y(t) \) are (nearly) equal to the total demand rate:

\[
y(t) \equiv d(t). \tag{1}
\]

We want to distinguish the unit sales of products of the commodity manufactured with different production technologies. For the \( i \)th generation of the production technology, the unit sales are indicated by \( y_i(t) \). The total unit sales are determined by

\[
y(t) = \sum_{i=1}^{N(t)} y_i(t), \tag{2}
\]

where the number of different technological generations at a given time step is \( N(t) \). The unit sales of products manufactured with the \( i \)th generation grows with

\[
\frac{dy_i}{dt} = \theta_i y_i, \tag{3}
\]

where \( \theta_i \) is the growth rate of the corresponding generation. The growth process of products in a competitive market is not further discussed here (for details see, e.g., [13]). It is assumed that products manufactured with the \( i \)th production technology have costs per unit \( c_i \). They decrease with increasing technological generation such that \( c_i > c_{i+1} \). Since suppliers can offer the corresponding products for lower prices, the growth rates \( \theta_i \) of the unit sales are different. Applying (3), the evolution of the total unit sales is given by

\[
\frac{dy}{dt} = \sum_{i} \frac{dy_i}{dt} = \sum_{i} \theta_i y_i = \langle \theta \rangle y, \tag{4}
\]

where \( \langle \theta \rangle \) is the mean unit sales growth rate of the commodity market.

The market share \( m_i \) of sold units manufactured with the \( i \)th technological generation is defined as

\[
m_i(t) = \frac{y_i(t)}{y(t)}. \tag{5}
\]

The time derivative of this relation yields

\[
\frac{dy_i}{dt} = \frac{dm_i}{dt} y + \frac{dy}{dt} m_i(t). \tag{6}
\]

Using (3) and (4), this relation can be rearranged with (5):

\[
\frac{dm_i}{dt} = (\theta_i(t) - \langle \theta \rangle) m_i(t), \tag{7}
\]

which is a replicator equation for the market shares of the technological generations. Since products manufactured with different technological generations have different growth rates, they suffer from a preferential growth process, while the growth rate \( \theta_i \) represents the fitness of the \( i \)th generation. The sales dynamics of products are therefore governed by the fitness function \( \theta_i \) which is regarded to be a function of the costs per unit \( \theta_i = \theta_i(c_i) \).

Note that (7) can be rewritten as

\[
\frac{d \ln(m_i(t))}{dt} = \theta_i - \langle \theta \rangle. \tag{8}
\]

In order to derive the market share evolution of different generations, we diminish a second generation with index \( j \) from (8) and obtain

\[
\frac{d \ln(m_i(t))}{dt} - \frac{d \ln(m_j(t))}{dt} = \theta_i(t) - \theta_j(t) = \Delta \theta_{ij}(t), \tag{9}
\]

where \( \Delta \theta_{ij}(t) \) is termed fitness (competitive) advantage. The relation between two market shares becomes as follows:

\[
m_i(t) / m_j(t) = m_i(t_0) / m_j(t_0) \exp \left( \int_{t_0}^{t} \Delta \theta_{ij}(t') dt' \right). \tag{10}
\]

In order to simplify this relation, we rewrite the fitness advantage as a mean fitness advantage and time dependent fluctuations:

\[
\Delta \theta_{ij}(t) = \langle \Delta \theta_{ij} \rangle_{\Delta t} + \delta \theta_{ij}(t), \tag{11}
\]

where the mean fitness advantage is the average over a time period \( \Delta t \):

\[
\langle \Delta \theta_{ij} \rangle_{\Delta t} = \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} (\theta_i(t) - \theta_j(t)) dt. \tag{12}
\]
The integral in (12) yields $\Delta t = t - t_0$:

$$
\int_{t_0}^{t} \Delta \theta_{ij} (t') \, dt' = \langle \Delta \theta_{ij} \rangle_{\Delta t} (t - t_0) + \Xi_{ij} (t),
$$

(13)

while the time dependent function $\Xi_{ij} (t)$ is obtained from the integration over the fluctuations $\delta \theta_{ij} (t)$. With identity [24],

$$
m_i (t) = \frac{m_i (t)}{\sum_j m_j (t)} = \frac{1}{1 + \sum_{j \neq i} \left( m_j (t) / m_i (t) \right)},
$$

(14)

we formally obtain the time evolution of the market share:

$$
m_i (t) = \frac{1}{1 + e^{-\langle \Delta \theta_{ij} \rangle_{\Delta t} + \kappa_i}}, \quad m_j (t) = \frac{1}{1 + e^{-\langle \Delta \theta_{ij} \rangle_{\Delta t} + \kappa_j}},
$$

(15)

with the coefficients $\kappa_{ij} = t_0 \langle \Delta \theta_{ij} \rangle_{\Delta t}$. For the case where the impact of the time-dependent fluctuations $\Xi_{ij} (t)$ is sufficiently small to be neglected, we obtain for the two generation case the well-known Fisher-Pry substitution relation [25]:

$$
m_i (t) = \frac{1}{1 + e^{-\langle \Delta \theta_{ij} \rangle_{\Delta t} + \kappa_i}}, \quad m_j (t) = \frac{1}{1 + e^{-\langle \Delta \theta_{ij} \rangle_{\Delta t} + \kappa_j}}.
$$

(16)

Applying (1) and (5), the unit sales of products of successive technological generations become

$$
y_i (t) = \frac{1}{1 + \sum_{j \neq i} e^{-\langle \Delta \theta_{ij} \rangle_{\Delta t} + \kappa_i}}.
$$

(17)

The sales evolution of products manufactured with different production technologies is essentially determined by their fitness advantage. Products with lower unit costs have a fitness advantage, since they can be sold for a lower price. Those product generations generating higher costs per unit are therefore replaced in time by those with lower costs per unit according to a logistic law of the unit sales given by (17).

### 2.1. Constrained Technological Evolution

Implementing a new production technology with lower costs per unit gives manufacturers a competitive advantage. We want to discuss the technological evolution for the case that this advantage is related to a cost relevant characteristic $\Gamma$ of the production technology. That means that the fitness is regarded to be as a function of a technological or physical characteristic $\theta_i = \theta_i (\Gamma_i)$. The unit costs of a product of the $i$th generation depend on the magnitude of $\Gamma_i$ such that its increase (or decrease) reduces the costs per unit significantly. The mean magnitude of the characteristic at a given time step can be obtained by taking the average over the sold number of products manufactured with current technologies:

$$
\langle \Gamma \rangle = \frac{1}{y} \sum_i \Gamma_i y_i.
$$

(18)

In the run of time successive generations of the production technology are replaced as described by (17). This leads to a slow shift of the characteristic $\langle \Gamma \rangle$ and simultaneously to a lowering of the mean costs per unit. In order to quantify this slow technological evolution, we introduce a separation of the time scales. On the short time scale $t$ the change of the mean characteristic can be neglected as follows:

$$
\frac{d}{dt} \langle \Gamma \rangle \sim \varepsilon \approx 0,
$$

(19)

where $\varepsilon \ll 1$. The long time scale $T$, however, is chosen such that the replacement of successive production technology generations can be regarded as a continuous process with

$$
\frac{d}{\varepsilon dt} = \frac{d}{dT} \neq 0.
$$

(20)

Hence, the time scales are related by $T = \varepsilon t$. In order to establish this evolution, we take advantage from the definition of $\langle \Gamma \rangle$ in a continuous form:

$$
\langle \Gamma \rangle = \int_0^\infty P (\Gamma) \Gamma d\Gamma,
$$

(21)

where $P (\Gamma)$ is the probability density distribution characterizing the relative abundance of products manufactured with different production technologies. The time derivative of the mean characteristic becomes

$$
\frac{d}{dt} \langle \Gamma \rangle = \int_0^\infty \frac{dP (\Gamma)}{dt} \Gamma d\Gamma.
$$

(22)

For the time evolution of the relative abundance $P (\Gamma)$, we take advantage from the evolution of the unit sales market shares. They are governed by the replicator equation (7). We can therefore write

$$
\frac{d}{dt} \langle \Gamma \rangle = \int_0^\infty P (\Gamma) \theta (\Gamma) - \langle \theta (\Gamma) \rangle \Gamma d\Gamma
$$

(23)

and further evaluate

$$
\frac{d}{dt} \langle \Gamma \rangle = \int_0^\infty \theta (\Gamma) P (\Gamma) \Gamma d\Gamma - \langle \Gamma \rangle \langle \theta \rangle .
$$

(24)

For the case that the distribution $P (\Gamma)$ is located around the mean value, the fitness function can be expanded around $\langle \Gamma \rangle$ as

$$
\theta (\Gamma) = \theta (\langle \Gamma \rangle + \delta \Gamma) = \theta (\langle \Gamma \rangle) + \frac{\partial \theta (\Gamma)}{\partial \Gamma} \bigg|_{\langle \Gamma \rangle} \delta \Gamma,
$$

(25)

with the small deviation $\delta \Gamma = \Gamma - \langle \Gamma \rangle$. Inserting this relation in (24), we obtain for the evolution of the characteristic on the long time scale,

$$
\frac{d}{dT} \langle \Gamma \rangle = \frac{\partial \theta (\langle \Gamma \rangle)}{\partial \Gamma} \bigg|_{\langle \Gamma \rangle} \text{Var} (P (\Gamma)),
$$

(26)

where $\langle \theta \rangle = \theta (\langle \Gamma \rangle)$ and the variance is defined as

$$
\text{Var} (P (\Gamma)) = \int P (\Gamma) \Gamma^2 d\Gamma - \left( \int P (\Gamma) \Gamma d\Gamma \right)^2.
$$

(27)

This result suggests that the mean characteristic will change in time as long as the fitness can be increased by a variation.
of $\Gamma$, since in this case the variance is nonzero. The velocity by which the mean technological characteristic varies depends on the selection pressure given by $d\theta(\Gamma)/dT$. In order to determine the selection pressure, the function $\theta(\Gamma)$ has to be known.

We want to confine the model to the special case that $\Gamma$ has a technological limit at $\Gamma^*$, such that $\Gamma$ cannot have a magnitude beyond this limit. Since a variation of $\Gamma$ reduces the unit costs, the function $c(\Gamma)$ must have a minimum at $\Gamma^*$. Because the costs per unit have an overall minimum at $\Gamma^*$, the long term fitness function $\theta(\Gamma)$ must have a maximum there. Expanding the fitness function sufficiently close to $\Gamma^*$, we can write

$$\theta(\Gamma) = \theta^\text{max} - \theta_0 (\Gamma - \Gamma^*)^2,$$  \hfill (28)

where we consider the case that the technological characteristic $\Gamma$ must decrease in order to decrease the costs per unit. In this relation, $\theta^\text{max}$ indicates the maximum fitness at minimum costs per unit and $\theta_0$ is a free parameter. Taking advantage form (26), we obtain, with this approximation,

$$\frac{d (\Gamma)}{dT} = \frac{2}{\varepsilon} \text{Var} (P (\Gamma)) \theta_0 (\langle \Gamma \rangle - \Gamma^*) = -a^\Gamma (\langle \Gamma \rangle - \Gamma^*),$$  \hfill (29)

where the parameter $a^\Gamma$ is denoted as decline rate of $\Gamma$. Treating the decline rate in a first approximation as constant, we finally get

$$\langle \Gamma(T) \rangle \equiv \Gamma_0 \exp (-a^\Gamma T) + \Gamma^*,$$  \hfill (30)

where $\Gamma_0$ is as a free parameter related to $\Gamma$ at introduction of the technology. This result suggests that after sufficient time, the mean magnitude $\langle \Gamma(T) \rangle$ approaches the technological limit at $\Gamma^*$ asymptotically in time. It is a result of the replacement process of successive technologies with different production costs.

2.2. Moore’s Law. We want to apply the model describing the evolution of successive production technologies to the semiconductor industry in order to establish Moore’s law. As mentioned above, Moore’s Law is related to the number of transistors in integrated electronic circuits. The production technology to create transistors on a wafer is essentially determined by the lithographic method. The number of transistors and hence the memory size depends on the minimum feature size $f_s$. The increase of the density of electronic devices on a microelectronic chip leads to faster switching times and to lower cost per bit [10]. The presented evolutionary model can be applied by regarding the minimum feature size $f_s$ as the cost relevant technological characteristic $\Gamma$.

In order to bring electronic units on a chip, the circuits need to be delineated with a finer and finer brush. Applying optical lithography, the size of this brush, respectively, the resolution, is defined by [26]:

$$f_s = k_1 \frac{\lambda_{\text{light}}}{\text{NA}},$$  \hfill (31)

where $\lambda_{\text{light}}$ is the wavelength of the light source used, NA is the numerical aperture of the lens applied to image the circuit patterns, and $k_1$ (known as $k$-factor) is a factor describing the ability of the recording process to resolve small features.

In order to increase the memory size, more and more transistors are placed on a given surface. Applying the lithographic method, the number of transistors $N_t$ must increase proportional to the inverse of the minimum feature size:

$$N_t (T) \sim \frac{1}{f_s (T)^d},$$  \hfill (32)

where $d$ is a dimension. Since the lithographic method is based on the exposure of light on a surface, it could be expected that the dimension is that of a surface $d = 2$. However, lithographic methods take advantage from several stacked electronic layers. Therefore, the dimension will deviate from that of a surface. Instead, the dimension must be higher than a surface but less than the spatial dimension: $2 < d < 3$. Since $f_s$ has the dimension of a length, we introduce the length $f_s^*$ as a proportionality factor and establish a relation between the number of transistors and the minimum feature size of the form:

$$N_t (T) \equiv \left( \frac{f_{s^*}}{f_s (T)} \right)^d.$$  \hfill (33)

Note that such relations are known in statistical physics as scaling relations [27]. It is suggested that transitions are arranged in some sort of a self-similar structure.

Increasing the number of transitions per chip decreases the costs per memory unit (bit). In order to increase $N_t$, the minimum feature size must be decreased. In the evolution of the lithographic method this takes place mainly by decreasing the wavelength of the applied light source [26]. However, the minimum feature size is bounded by the shortest wavelength of light. It determines the technological limit of the minimum feature size $f_s^*$. The technological evolution of the number of transistors can be understood in terms of the presented evolutionary model as an adaptation process towards this limit. Applying (30) with $\Gamma = f_s$, the mean minimum feature size evolution is expected to have the form:

$$f_s (T) = f_{s^*} e^{-a^s T} + f_s^*,$$  \hfill (34)

where $f_{s^*}$ is related to the minimum feature size at introduction, $f_s^*$ is the technological limit, and $a^s$ is the corresponding decline rate. Note that this relation is only valid on the long time scale sufficiently close to $f_s^*$. It expresses the fact that the costs per bit cannot decrease below $f_s^*$ with the standard lithographic technology.

Applying this time evolution of the feature size in (33), we obtain Moore’s law in the form:

$$N_t (T) = \left( \frac{f_{s^* e^{-a^s T}}}{f_{s^* e^{-a^s T} + f_s^*}} \right)^d.$$  \hfill (35)

The miniaturization process increases the number $N_t$ of electronic devices on a chip. For small $T$ and sufficiently far
from the limit $f_s^*$, the relation suggests that the number of transistors increases exponentially with $N_t(T) \sim \exp(a f_s T)$ in agreement to Moore’s law, while Moore estimated the growth rate $a f_s$ to be equivalent to an eighteen month doubling of $N_t$. However, the number of transistors cannot increase to infinity but must be limited by the minimum feature size that can be physically achieved. This effect is included in (35). It describes therefore the long term technological evolution of Moore’s law even close to the technological limit.

3. Comparison with Empirical Results

The presented evolutionary model suggests a logistic replacement of successive production technologies when they can be related to a reduction of the production costs per unit. In the case that the main competition is between neighbouring generations, the replacement takes place according to a Fisher-Pry-law in terms of the unit sales market shares. If the production technology is constraint by a physical or technological limit, then the technological evolution evolves such that the mean value of a characteristic parameter ($\langle \Gamma \rangle$) related to the costs per unit approaches the technological limit $\Gamma^*$ asymptotically as given by (30). We want to compare the theory with empirical results of two characteristics of the production technology of DRAM chips. We first study the wafer size as a cost relevant characteristic. Then, the evolution of Moore’s law is predicted under the condition of a technological limit.

3.1. Wafer Size Evolution. We want to consider the wafer size as a cost relevant technological characteristic $\Gamma$ of the DRAM production technology. A wafer serves as the substrate for microelectronic devices characterized by its diameter. The increase of the wafer diameter increases the number of units that can be obtained from a single wafer. Therefore, the costs per unit decrease with increasing wafer size. The model suggests that the replacement process of different wafer sizes must be governed by the Fisher-Pry law for the case that competition takes place essentially between neighbouring generations. Displayed in Figure 1 is the evolution of the wafer size market shares $m_i$ of the production technology with different diameters in a Fisher-Pry plot. The solid lines represent a fit with the evolutionary model.

3.2. Moore’s Law. In order to test Moore’s law, we take advantage form empirical data of the minimum feature size. They are displayed in Figure 2 [28, 29]. The empirical data (squares) show the expected exponential decline of the minimum feature size $f_s$ in nm [30]. The solid line is a fit of (34) with $f_{s_0} = 8100$ nm and $a f_s = 0.16$ per year, and the technological limit can be estimated to be of the order of the minimum wavelength of UV light $f_s^* \approx 10$ nm.

From the evolution of the minimum feature size follows Moore’s law given by the number of transistors per chip $N_t$ (squares) for DRAM’s as a function of time. The solid lines represent a fit with the evolutionary model.

4. Conclusion

Since the costs per unit of a good are governed by the production technology, the presented evolutionary model suggests that manufacturers have a competitive advantage when they...
apply new generations of the production technology. If the main competition is confined to neighbouring generations, the unit sales market shares of sold products are expected to evolve according to a Fisher-Pry-plot law. Also, derived is the case that a process technology is governed by a cost relevant characteristic that is constrained by a technological or physical boundary. The model suggests that in this case the limit is approached asymptotically in time.

In order to test the model two characteristics of the DRAM semiconductor production technology are studied. The wafer size is a cost relevant characteristic of the production technology because the costs per unit decrease with an increasing wafer size. The model suggests that different wafer sizes replace each other according to a Fisher-Pry law. Empirical data confirm this replacement process of successive generations of the wafer size. Note that similar replacement processes are known from other technologies [9, 32–34].

Another cost relevant characteristic of the DRAM production technology is related to the minimum feature size of electronic elements on a chip. It determines the number of transistors per chip and governs therefore Moore's law. Applying the lithographic method, the minimum feature size is bounded by the minimum wavelength that can be applied. This limit restricts the density of transistors. While Moore's law suggests an exponential increase of the number of transistors per chip, the model agrees with this statement far from the technological limit but suggests a deviation from Moore's law in the run of time. It predicts that the miniaturization process will slow down considerably in the next two decades (Figure 2).

Note that the evolution of the production technology can be regarded as a learning process. It is a search for decreasing costs per unit in the production process. However, we have to distinguish this “evolutionary learning process” from the “learning by doing”. The technological learning process described here is driven by the commodity market. It is in fact an evolutionary adaptation process of the suppliers to the demands of the buyers. Since, the demand side prefers a lower price for a unit of the commodity, production technologies with lower costs per unit have an evolutionary advantage. As a consequence, cost reducing technologies replace each other in time leading to a decrease of the mean price for a unit of this commodity. Moore's law is a prominent example of this process. The ability to increase the number of transistors per chip by improvements of the lithographic method not only increases the performance of a computer chip but also decreases its mean price considerably in the run of time [10].

The two learning processes can be distinguished by the costs for variations of the production technology. If the costs for a change of the production process are low, the variation can be applied easily within short time periods. This effect can be regarded as learning by doing process. In the semiconductor industry, the fast decrease of the lead time in the production process of computer chips can be regarded as learning by doing [35]. However, if a technology variation requires considerably investments, manufacturers will try to apply the old technology as long as possible. In this case, the replacement process of successive production technologies takes place by a preferential growth process and is governed by the costs per unit, respectively, the product price. The evolutionary learning process discussed here is therefore confined to production technologies associated with high investments.

Production technologies are bounded by physical constraints. Learning processes aim at approaching these limits. The prediction of the time evolution of Moore's law in Figure 2 is based on the application of the lithographic method. The learning process implies, however, that manufacturers may apply alternative techniques in order to circumvent this limit. But even so, new physical boundaries will appear. As soon as the corresponding production technology will lead to an increase of the costs per unit, the model suggests that the miniaturization process will come to an end.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References


