The dynamics of technological substitutions

Brice Dattee
Imperial College London, Tanaka Business School
London, SW7 2AZ , UK
Tel: +44(0)2075941864
b.dattee@imperial.ac.uk

David FitzPatrick
University College Dublin, College of Engineering
Dublin 4, Ireland
Tel: +353(0)17161829
david.fitzpatrick@ucd.ie

Henry Birdseye Weil
Massachusetts Institute of Technology, Sloan School of Management
Cambridge, MA, US
Tel: +1(617)258-6101
weilco@compuserve.com

Abstract

Technological substitution is the process by which disruptive technologies replace the dominant ones in an industry. The formulation of classical models of diffusion and substitution impose simplification constraints to reach analytical solvability. We use the system dynamics methodology to build upon existing models by integrating dynamic aspects derived from a broad theoretical framework and to explore the links between social dynamics, technological developments and substitution patterns. Our simulation model generates a substitutive drop in the life cycle which is not replicated by classical models but substantiated by empirical data from various industries. The more general theory embodied in the model allows to better understand the underlying dynamics of technological substitutions. The generic structure can generate the dynamics of a sailing ship effect and account for the non-uniformity of interpersonal communications.

Keywords: *Technological substitutions, diffusion models, expected utility, heterogeneity of markets.*

1 Introduction

The diffusion process is a very well-ploughed academic ground; a widely researched and extensively documented social phenomenon. Yet, classical models of diffusion typically make oversimplifying assumptions to describe it. The classical models usually are analytical refinements of the Bass model (Bass, 1969) and generate bell-shape life cycles; hence the smooth logistic shape of cumulative adoption. Despite their good fit to historical data, their epidemic structure – based on external and internal communications – lack explanatory power by oversimplifying the adopter's decision making process and do not fully account for market heterogeneity. Moreover, the innovation under scrutiny is almost always considered as static; classical models ignore the technological evolution over the life cycle. Some models focus on substitution and account for successive generations of technologies. These are usually descriptive models limited to two competing technologies following the tradition of the Fisher and Pry model (Fisher and Pry, 1971).

A review of the literature shows that several diffusion and substitution models attempt to account for non-linear influence, for a multi-stage adoption process, for heterogeneity at the individual level, for a dynamic potential market, for technological evolution and finally for multi-innovations substitutions dynamics.

Nevertheless, there is an opportunity to broaden the boundaries of existing models by integrating these different stream of works. We show that it is possible to develop a model which accounts for technological evolution, market adoption and the dynamics induced by market heterogeneity and social networks (Dattee and Weil, 2005). Thus, we present a simulation model developed with the system dynamics simulation methodology. We discuss its theoretical underpinnings, describe the model's structure and present how its dynamic behavior replicate both diffusion and substitution effects. The model generates an asymmetrical life cycle whereby there is a sudden drop in the sales of the current technology when it is confronted by the take off of a new generation. This dynamic behavior is not replicated by classical analytical models of diffusion; yet historical data from different industries (DRAM, VHS/DVD, etc…) clearly corroborate this substitutive drop.

With the broader theoretical framework embedded in its structure, the model allows the exploration of more complex dynamics. We illustrate how a defensive surge from the threatened technology can induce a delay in the substitution time-path. This is the classical "sailing ship effect" described in the literature on technology and innovation management. These are broader dynamics that cannot be captured by classical diffusion models. We then discuss how the model's structure could be modified to account for some of the social dynamics occurring during technological substitutions.

2 Existing models of diffusion

The underlying assumption of diffusion research is that an innovation is communicated and absorbed over time into a social system in stages, corresponding to the psychological and social profiles of various segments within that population. Diffusion models have been developed to represent the spread of an innovation in terms of a simple mathematical function of the time that has elapsed from the introduction of the innovation. Thus, the need for a simpler structure leads to several simplifying assumptions which could seem unrealistic.

Classical models do not account for important interdependencies and structural fundamentals. They make strong assumptions on the process of innovation diffusion by considering that adoption is a binary process, the potential market size is constant, there is no repeat purchase, there are uniform probability of dyadic interactions between prior and potential adopters, and that the innovation itself does not change over the diffusion process. This latter assumptions implies that for a technological innovation, further developments in price and performance are not accounted for in the modeling of the diffusion process. Figure 1 illustrates the formulation and structure of the classical Bass model (Bass, 1969).

$$\frac{n(t)}{m - N(t)} = p + \frac{q}{m} N(t)$$

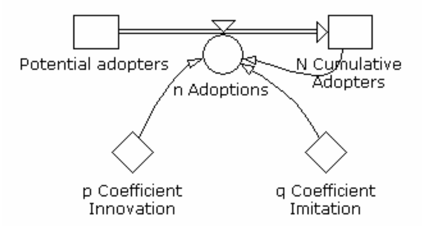


Figure 1: The classical Bass diffusion model: analytical formulation and system dynamics structure

Parker offers a review of analytical refinements nested in the fundamental model formulation (Parker, 1991). Easingwood et al. proposed flexible versions of the internal-influence and mixed-influence models, the Non-symmetric Responding Logistic (NSRL) and Non-Uniform Influence (NUI) respectively (Easingwood, Mahajan et al., 1983). The purpose of these models is to overcome the inherent assumption that the “word-of-mouth effect remains constant over the entire diffusion span”.

Some attempts have also been made to extend the two-stage model to incorporate the multi-stage nature of the diffusion process. These models hypothesize that social members first become potential adopters and then current adopters. Dekimpe et al. in their study of the diffusion of technological innovations at the national policy level (Dekimpe, Parker et al., 1998) adopted an hazard-rate structure applied to a two-stage process. Similarly, there are examples of multi-stage models structure with an untapped market, potential adopters and current adopters (Milling, 1996; Maier, 1998; Milling, 2002). However, these models do not offer a theoretical rationale for the growth of the potential adopters group.

The Fisher-Pry model is an analytical formulation used to project the market share evolution of an emerging technology. The model is based upon three explicit assumptions: that “many technological advances can be considered as competitive substitutions of one method of satisfying a need for another”; that new technologies often completely supplant older ones; and “the rate of fractional substitution of new for old is proportional to the remaining amount of the old left to be substituted”. They assert that the rate constant of substitution, once begun, does not change. The market share model is expressed as:

$$\frac{ds(t)}{dt} = ks(t)[1 - s(t)] \qquad 1$$

where *s(t)* is the fractional market share of the innovation at time t, and *k* is a constant of proportionality (Fisher and Pry, 1971). Using the assumption that there are only two competing technologies, Fisher and Pry derive a more convenient form for purposes of estimation. The result is that the log of the ratio of the market share of the succeeding technology to that of the first is a linear function of time:

$$\ln\left(\frac{s}{(1 - s)}\right) = kt \qquad 2$$

Nevertheless, by normalizing the market potential with a market share formulation for only two competing technologies, the system appears static. As an innovation invades a market, it starts interacting with the technologies already established. Pistorius and Utterback argue that the interactions between technologies should be viewed in a broad sense (i.e. not just pure competition) and suggest a multi-mode technological interaction framework (Pistorius and Utterback, 1995; 1997). Their formulation through a Lotka-Volterra formulation is a general model for multi-technology, multi-mode interaction. It can be used to model the interaction of any finite number of technologies where the interaction among any pair can either be pure competition, predator-prey, or symbiosis. However, this Lotka-Volterra modelling formulation remain descriptive and do not account for technological development.

Classical models characteristically make the following assumptions; that adoption is a binary process; the size of the potential market is constant; there are no repeat or upgrading purchases; dyadic interactions between prior and potential adopters; and the innovation itself is static over the diffusion process. Hence, developments in price and performance are ignored. Several attempts have been made to release one or two of these limitations based on either analytical refinements (Easingwood, Mahajan et al., 1983; Parker, 1991), a multi-stage structure (Dekimpe, Parker et al., 1998), a multi-innovations model (Fisher and Pry, 1971; Kabir, Sharif et al., 1981; Norton, 1986; Norton and Bass, 1987, 1992; Pistorius and Utterback, 1995; 1997), individual level parameters (Roberts and Urban, 1988; Chatterjee and Eliashberg, 1990; Lattin and Roberts, 2000; Adner and Levinthal, 2001) or a dynamic potential market (Homer and Finkelstein, 1981; Kabir, Sharif et al., 1981; Norton and Bass, 1987; Lyneis, 1993; Maier, 1996; Milling, 1996). The model developed by Weil and Utterback has dynamic market size, repeat purchases of two generations of technology, and dynamic product price/performance (Weil and Utterback, 2005). It does not, however, contain the social dynamics and other refinements in the digital music version later described by Weil (Weil, 2007). There is still an opportunity to develop a model that integrates these various attempts in order to fully link social dynamics, technological developments, and market adoption. We can see that there is an opportunity to improve the descriptive approach with a model that combines the underlying dynamics of market adoption with technological developments. Moreover, a better handling of the social dynamics could lead to a powerful structure explaining diverse patterns of technological substitutions.

3 A system dynamics model of technological substitutions

We believe that a multi-stage, multi-innovation, dynamic potential diffusion model based on the aggregation of restricted individual level parameters will allow exploring the links between technological evolution, social dynamics and substitution patterns. We use the system dynamics simulation methodology to integrate all these processes, linking them to the trajectories of successive technologies. The use of system dynamics is particularly interesting for the study of social factors in technological substitution because it considers system causation as endogenous, i.e. not brought by external variations or shocks, but by the way feedback structures process external events. Moreover, our approach integrates several models that taken together cannot be solved analytically. The system dynamics methodology enables us to simulate their behavior. Figure 2 shows a synoptic view of the structure of the simulation model.

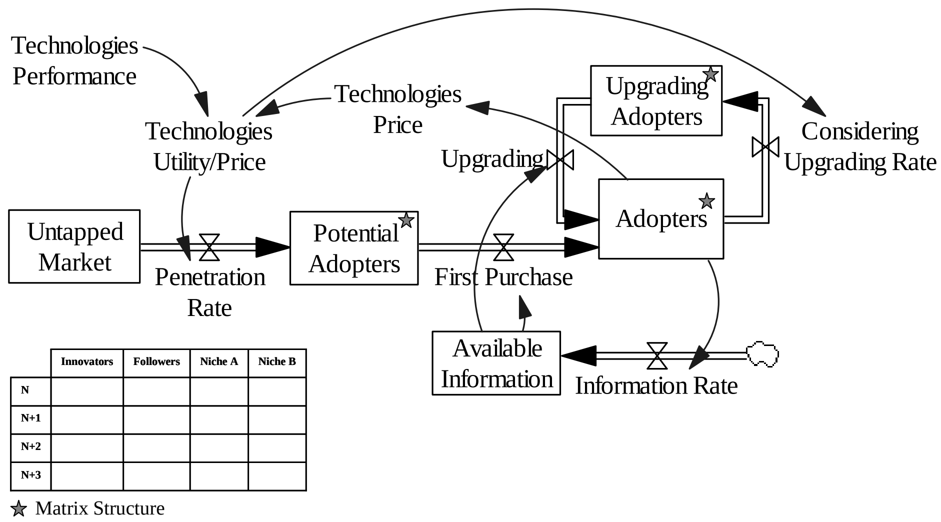


Figure 2: Synoptic view of the model structure

Technology assessment

Consumers in the untapped market are assumed to evaluate the innovation along a performance index and price. We operationalize the evaluation process through a von Neumann-Morgenstern framework via the uniattribute negative exponential utility function and risk aversion r (Chatterjee and Eliashberg, 1990). Performance is modeled as an aggregated index of attributes moving along a technological trajectory and price is decreased as a function of cumulative volumes. The value of a technological innovation n is thus represented by:

$$X_n = X_n(y_{n1}, y_{n2}, \ldots, y_{nK}) = \sum_{k=1}^{K} w_k y_{nk} \qquad 3$$

where y_{nk} is the level of attribute k for technology n and w_k is the relative importance weight of attribute k in the targeted market segment. The mean value X_n changes over time due to both technological developments and word of mouth communication which influences the estimation. DeGroot's formulation of Bayesian estimation theory is used to reflect the updating of prior perceptions of mean and uncertainty (DeGroot, 1970; Roberts, 1984). As more "units" of information are received, expectations move towards the true value.

Given the technological innovation's dynamic price, Ψ_n, consumers in the untapped market perceive a utility that may be represented by the additive utility function:

$$U(\tilde{X}_n, P_n) = k_x u_\chi(\tilde{X}_n) + k_p u_p(\Psi_n) \qquad 4$$

where $\tilde{\chi}_n$ denotes the potential adopter's uncertain perception of performance after receiving τ_n "units" of information, and where k_x and k_p are the scaling constants which may be interpreted as importance weights associated with the two uniattribute utility functions for performance and price respectively (Chatterjee and Eliashberg, 1990). We assume that the consumers' uncertainty about the measurable value of technology n, $\tilde{X}_n$ is characterized by a normal distribution (mean X_n, variance σ_n^2) and that a consumer's utility for price is linear in its argument. Consistent with the microeconomic model of choice based on maximization of expected utility subject to a budget constraint, the expected utility of technology *n* given the risk aversion *r* and the adoption decision rule are (Roberts and Urban, 1988; Chatterjee and Eliashberg, 1990):

$$E(U(\tilde{X}_n, p)) = 1 - \exp(-r(X_n - \frac{r}{2}\sigma_n^2)) - k_p \Psi_n \qquad 5$$

$$\frac{E\left[u_\chi(\tilde{X}_n)\right]}{\Psi_n} > k_p \qquad 6$$

Market penetration

At time *t*, consumers in the untapped market who believe that the technology *n* offers a utility-price ratio $UPR_n(t)$ that exceeds their threshold requirement k_p will become potential adopters of that technology. This condition assumes that the penetration of the untapped market is driven by the technology perceived to offer the best *UPR* in the set *S* of available technologies. The model assumes that parameter k_p is distributed across the entire population with the density function $f_k(\cdot)$ and the cumulative function $F_k(\cdot)$. The cumulative penetration of the market, driven by the successive generations *n* in the technology landscape *S*, is thus:

$$C(t) = M.F_k(Max(UPR_n(t))), \forall n \in S \qquad 7$$

where *M* is the total size of the targeted market. This means a common market for all competing technologies as opposed to unrelated individual markets as in other existing models (Norton and Bass, 1987). The penetration rate which depends on technological evolution and market heterogeneity is thus given by the first derivative:

$$\dot{C}(t) = M \frac{\partial Max(UPR_n(t))}{\partial t} f_k(Max(UPR_n(t))), t > 0, \forall n \in S \qquad 8$$

If a normal distribution (μ_k, σ_k^2) is used for the individual threshold k_p and assuming that technology *n* drives market growth, then the rate of potential adopters of technology *n* is given by:

$$\dot{C}(t) = M(t) * \frac{\partial(UPR_n(t))}{\partial t} * \frac{1}{\sigma_k\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{UPR_n(t)-\mu_k}{\sigma_k})^2} \qquad 9$$

Rogers identifies the innovators adopters as the first 2.5% of the social system (Rogers, 2003, p. 281). Therefore, up to this percentile of $F_k(.)$, those potential adopters will be considered as innovators independently of the technology generation; past this value the market penetration flow will be distributed into the followers subgroup of potential adopters of technology *n*. If a new technology *n+1* subsequently takes the lead, then these followers become potential adopters of technology *n+1*. Figure 3 shows the system dynamics structure of technology assessment and market penetration, inclusive of equations (3) to (9).

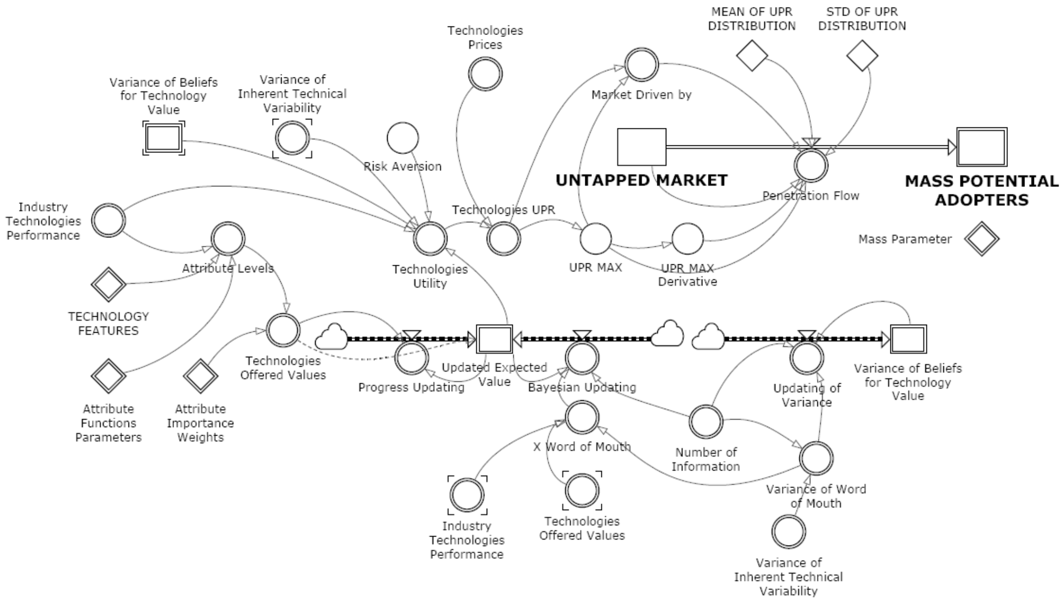


Figure 3: System Dynamics structure of consideration rate

First time purchase

Chatterjee and Eliashberg introduced the notion of 'critical amount of information', i^*, which is the cumulative amount of information about a technology that an individual requires so that the degree of uncertainty passes under the risk hurdle. A consumer receives p_n units of information from external sources and q_n units from word of mouth communication. We consider that a potential adopter receives units of word of mouth information from a constant proportion, λ of $A_n(t)$, the cumulative number of adopters of technology n at time t. Discounting factors are introduced to account for the moderating effect of the social topology, i.e. the credibility of information and relevance of opinion in inter-segments communication. If technology n is introduced at time t_{n0} , the amount of information available is:

$$i_n(t) = \int_{t_{n0}}^{t} \eta_n(\tau) d\tau \qquad 10$$

Thus, the pattern of first purchases is a function of cumulative information received and the distribution of i^* in the population. Let the distribution of i^* be the density function $f_{i^*}(\cdot)$. The fractional rate of purchase by potential adopters is obtained with the first derivative:

$$\dot{A}_n(t) = \eta_n(t) f_{i^*}(i_n(t)) \qquad 11$$

In classical diffusion models the potential adopters all have the same hazard rate, i.e. likelihood of adoption at time t. In contrast, the density $f_{i^*}(i_n(t))$ captures only those consumers who are "ready" to adopt. As shown in figure 4, the structure of this information feedback replicates the behavior of classical diffusion model when i^* has a negative exponential distribution[1] across the population:

[1] Chatterjee and Eliashberg give the relationships between parameters of the Bass model (p, q) and this critical information framework (Chatterjee and Eliashberg, 1990)

$$F_{i^*}(i^*) = 1 - \exp(-a_{i^*} i^*) \qquad 12$$

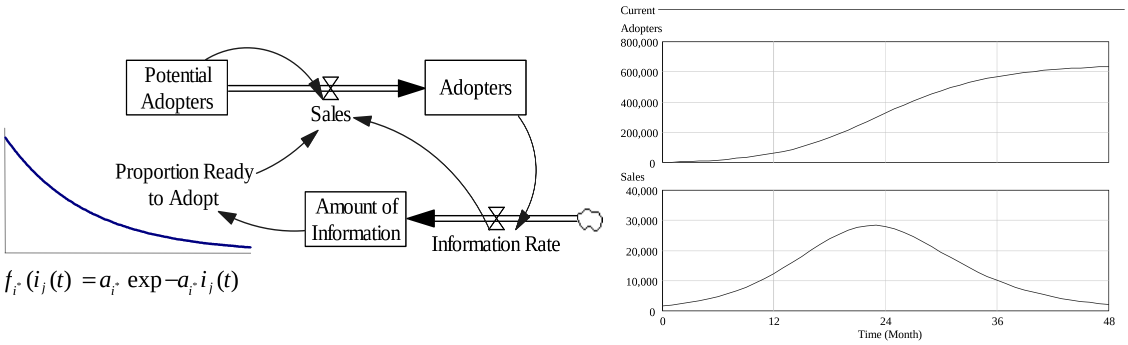


Figure 4: Diffusion pattern based on distribution of i^*

Niche markets

When a technology emerges, it is often crude, expensive, and does not appear to offer significant improvement on any of the usual dimensions. However, it can often perform a new functionality unrecognized in the mass market, but highly valued in some niche segments (Christensen, Suarez et al., 1998). We consider two niches: a niche A in which consumers apply a disjunctive rule by requiring a technology to perform, either a function F1, or another function F2; and a niche B in which consumers apply a conjunctive rule and require a combination of both. If a technology emerges which satisfies such a decision rule, then it initiates diffusion through the concerned segments and the feedback information above applies to the given subgroup. The word of mouth from niche adopters is integrated into the mass market by applying an inter-segment relevance factor. These decision rules and flow structure allow technologies with lower performance and higher price to still start diffusing by offering new functionalities. Figure 5 shows the adoption flow from both the niche and the mass markets potential adopters into the adopters stock matrix (technologies * categories).

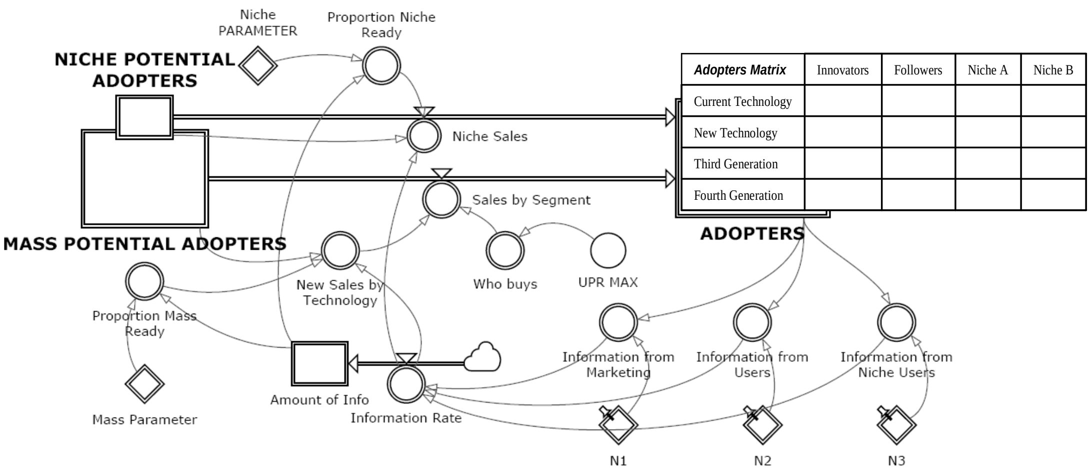


Figure 5: System dynamics structure of adoption rate

Upgrading and renewal

Lattin and Roberts found that market heterogeneity variables have a greater explanatory power than that of individual dynamic variables (Lattin and Roberts, 2000, p.22). The model assumes that the upgrading process of innovators is driven by the best available performance. The relative performance of a technology *n* in the set of technologies is:

$$RP_n = \frac{MAX(P_1,..,P_N)}{P_n}, \forall n \in S \tag{13}$$

Lattin and Roberts have considered the distribution of the requirement for relative utility to be uniform over [L;U] (Lattin and Roberts, 2000). Because the innovators represent a small fraction of the market, we also assume a uniform distribution with regards to the requirement for relative performance RP_n among the innovators category. The fractional rate of innovators upgrading from technology *n* and to the leading-edge can thus be written:

$$\text{If } RP_n \geq 1,\ UPI_n(t) = A_{n,Innov}(t) * \frac{\partial RP_n(t)}{\partial t} * \frac{1}{U-L}, \forall n \in S \tag{14}$$

with $A_{n,Innov}(t)$ the number of innovators currently possessing technology *n*. Due to the upgrading distribution, the innovators can be spread over several generations of technology. The total rate of upgrading to the leading-edge technology *N* is the sum of all innovators upgrading from their current technology $n < N$.

The pragmatic followers are assumed to be more price sensitive; we assume that they consider upgrading based on perceived relative *UPR*. Because of their larger number, we assume that the requirement for upgrading with regard to RU is normally distributed (μ_{RU}, σ^2_{RU}) across the followers population. Thus, the rate of followers upgrading from their technology *n* to the technology offering the best UPR can be expressed:

$$\text{If } RU_n \geq 1,\ UPF_n(t) = A_{n,Follow}(t) * \frac{\partial (RU_n(t))}{\partial t} * \frac{1}{\sigma_{RU}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{RU_n(t)-\mu_{RU}}{\sigma_{RU}}\right)^2}, \forall n \in S \tag{15}$$

with $A_{n,Follow}(t)$ the number of followers possessing technology *n*. These followers considering upgrading become potential adopters of the technology with the best *UPR*; their effective adoption is thus controlled by the information feedback described previously.

Finally, we account for sales derived from the renewal purchases after the physical life of an artifact. With a total number of adopters *A* and an average physical life of *x* years, then the average renewal rate is *A/x* per year. We further assume that adopters renewing will always choose to buy the best technology according to the decision criteria of their adopters category. Figure 6 shows the upgrading and renewal structure; table 1 highlights the key points of our system dynamics model.

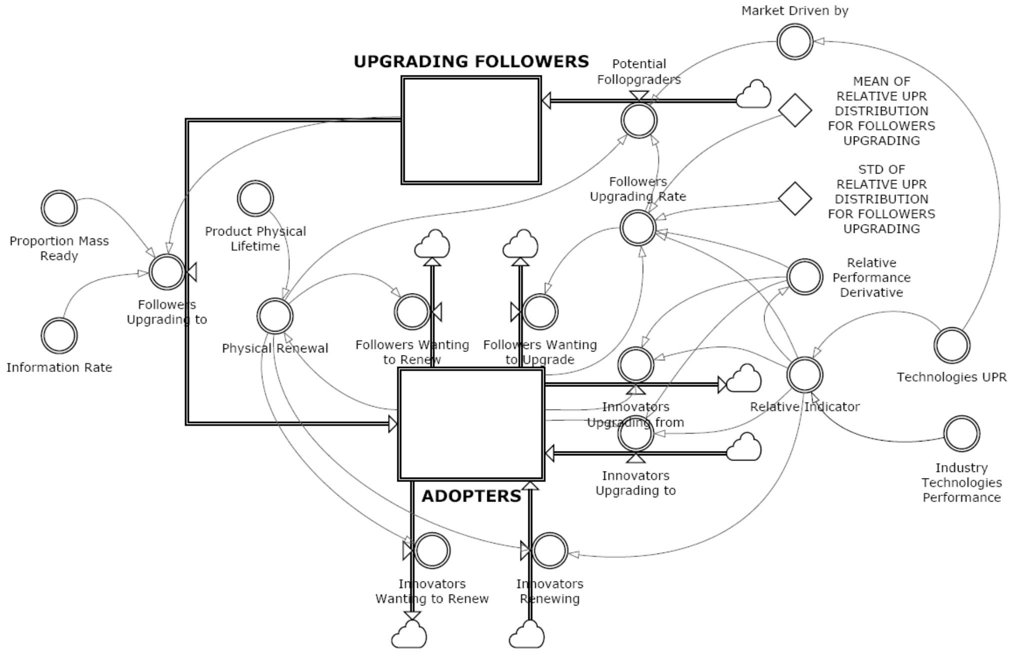


Figure 6: System dynamics structure for upgrading and renewal

Property / Dynamics	Model assumption
Decision process	Three stages: untapped, potential, adopters.
Market structure	Innovators (2.5%), Followers (97.5%), and Niche segments
Penetration of main market	Utility Price Ratio
Decision to adopt in niche	Conjunctive or disjunctive rule over given functionalities
Heterogeneous UPR thresholds	Normally distributed across entire population
Communication structure	Word of Mouth (lead users / followers / niche) ; relevance factor.
Uncertainty threshold	Negative exponential distribution of critical amount of information
Innovators upgrading	Relative Performance (upgrade to newer from same or next generation)
Followers upgrading	Relative UPR (upgrade to newer from same or next generation)
Physical renewing	Average annual flow given life cycle

Table 1: Summary of modeling key points

4 Simulation behavior

One technology

When the current technology (CT) is introduced, the market penetration is driven by improvements along the UPR_{CT} trajectory presented in figure 7.

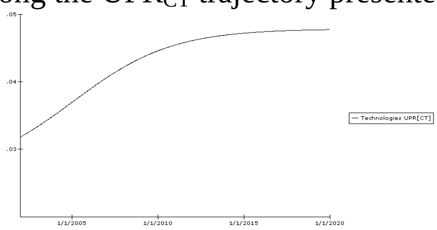


Figure 7: Current Technology UPR evolution

Accounting for uncertainty about the expected UPR and the normal distribution (μ_k, σ_k^2) of UPR requirements across the population, the market penetration flow is shown in figure 8. This figure also shows that the flow variability is reduced as $\hat{\mu}_n$ the expectations of the mean level of value of technology n converge towards μ_n the mean of the technology true average value through Bayesian updating.

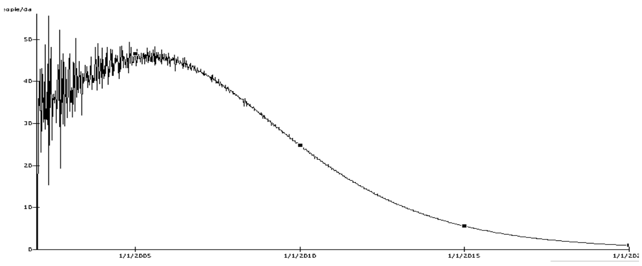


Figure 8: Market penetration flow for the Current Technology alone

This market penetration flow will flow into the potential adopters stock for technology CT. The potential adopters adopt on receiving their critical amount i^* of information about technology CT. Figure 9 illustrates the time delay that exists between consideration and the effective adoption for CT. This is an important feature of our multi-stage adoption model. Figure 10 shows the life cycle of one technology when accounting for physical renewal.

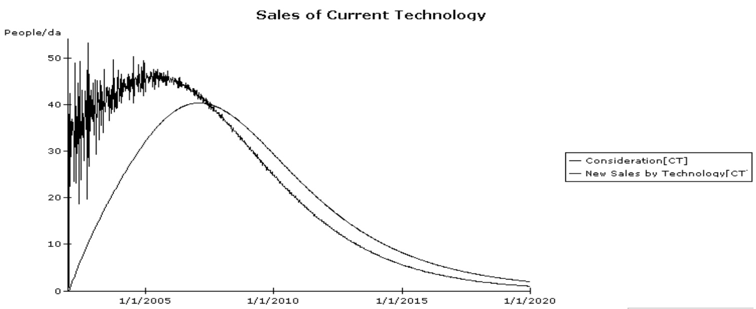


Figure 9: New Sales Rate for CT alone

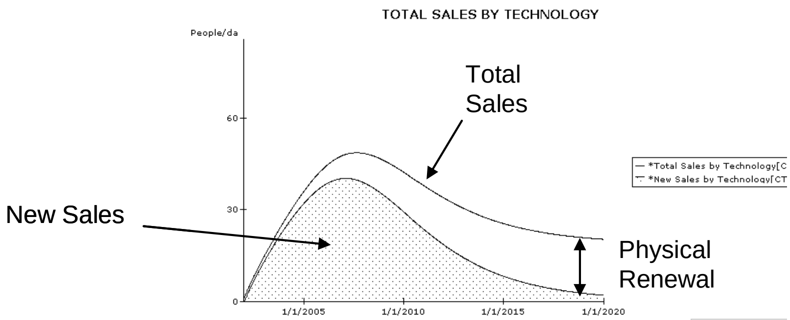


Figure 10: Total Sales of CT (New Sales + Physical Renewal)

Two competing technologies

We now consider that a new technology *NT* emerges while the current technology *CT* is diffusing into the market. Multiple modes of interaction could be considered, but we restrict our analysis of technological substitutions to a pure competition mode between successive generations of technology. How will the introduction of this second technology interfere with the diffusion process of current one? In our framework, the substitution dynamics depend on the development path of each technology and the heterogeneous requirements in the population as given by (μ_k, σ_k^2).

Figure 11 gives an example of the utility per price trajectory of the two competing technologies. As long as CT offers the best UPR, its development is driving market penetration. It should create a market whose total size is a fraction of the untapped market, as given by the position of its UPR upper limit on the cumulative distribution. However, once $UPR_{NT} > UPR_{CT}$, the penetration of the untapped market is thereafter driven by the evolution of NT and the overall market size for the product category can grow because the penetration of the untapped market has been driven further to the right of the requirement distribution by a higher UPR_{NT}.

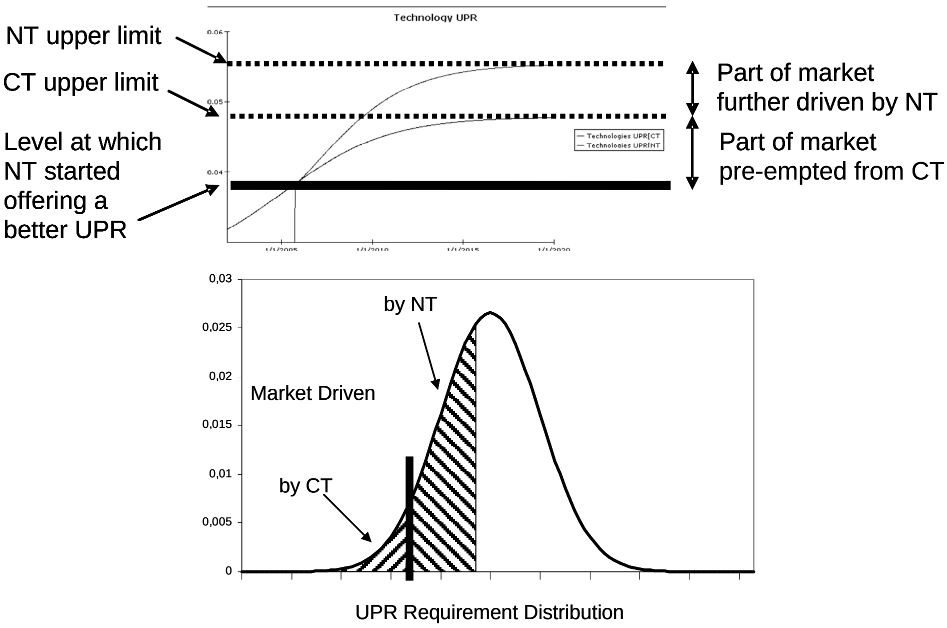


Figure 11: Technology evolution and UPR requirement distribution

After the discontinuity, the number of potential adopters for technology *N* can no longer increase. Nevertheless, the stock of consumers already engaged in the decision process for this technology does not disappear instantly; it gets depleted as they continue their information gathering. Figure 12a shows that the simulation of these dynamics creates a large pre-emption effect by technology *N+1* and this clearly results in what we call a “substitutive drop” in the sales of *N*. If unforeseen, this could have a devastating effect on a firm’s expected return on investment. Indeed, we can see on figure 12b that by using the data up to the discontinuity point, a classical Bass model can be satisfactorily fitted to the life cycle of technology *N*. However, it would completely miss the substitutive drop; any investment based on this expected profile could be seriously threatened. To overcome this structural mismatch, classical diffusion models are calibrated a posteriori with a smaller market size parameter and they anticipate the peak of sales as shown in figure 12c. This behavior is extremely clear in the application of Norton and Bass’ multi-generations model to DRAM devices (Norton, 1986; Norton and Bass, 1987, figure 2).

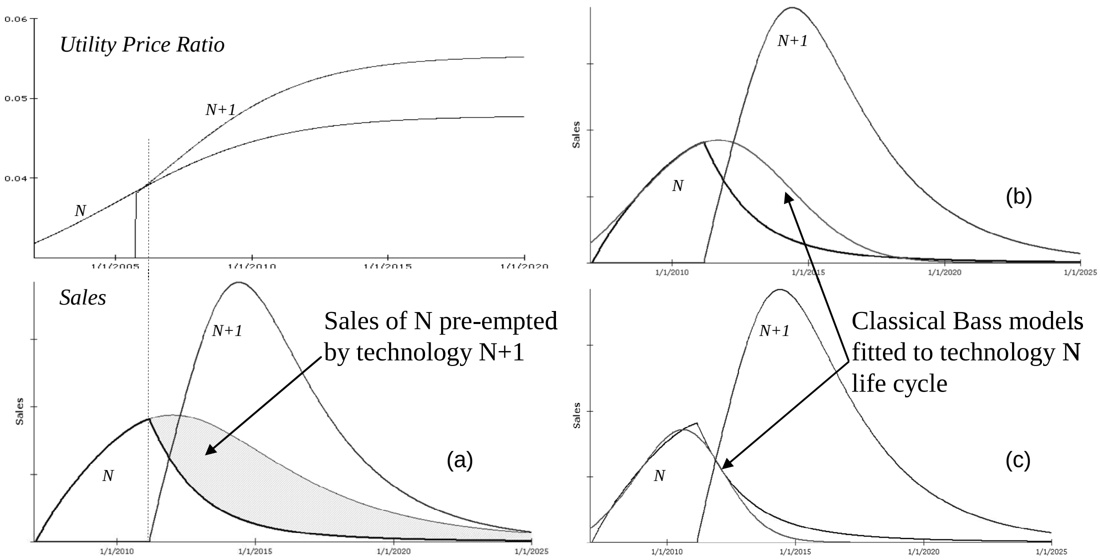


Figure 12: Simulation of the substitutive drop vs. the Bass model

Complete dynamics hypothesis

The experience curve assumes that every doubling of the cumulative production volume is associated with a cost reduction by a constant percentage. In our model we follow the simple pricing strategy of “full cost coverage”, i.e. “standard cost per unit plus a constant profit margin to assure prices above cost level even during the early stages of the life cycle” (Milling, 1996).

Our framework considers the evolution of both the performance and the utility per price (UPR) of each technology in order to investigate the diffusion, substitution, renewal and upgrading dynamics. The reconstitution of the evolution of a variable must identify the different time intervals which compose the process and its evolution over time. Figure 13 presents the evolution over time of the performance and the UPR of two competing technologies under a general perspective. Let us define the following times *T* that correspond to particular events:

- T0 as the introduction time of the current technology *CT*,
- T1 as the time at which UPR_{CT} has attracted the entire category of innovators, i.e. $F_k(UPR_{CT}) = 2.5\%$,
- T2 as the introduction time of the new technology *NT*,
- T3 as the time at which the performance – but not the UPR – of technology *NT* exceeds the performance of technology *CT*,
- T4 as the time at which UPR_{NT} exceeds UPR_{CT}.

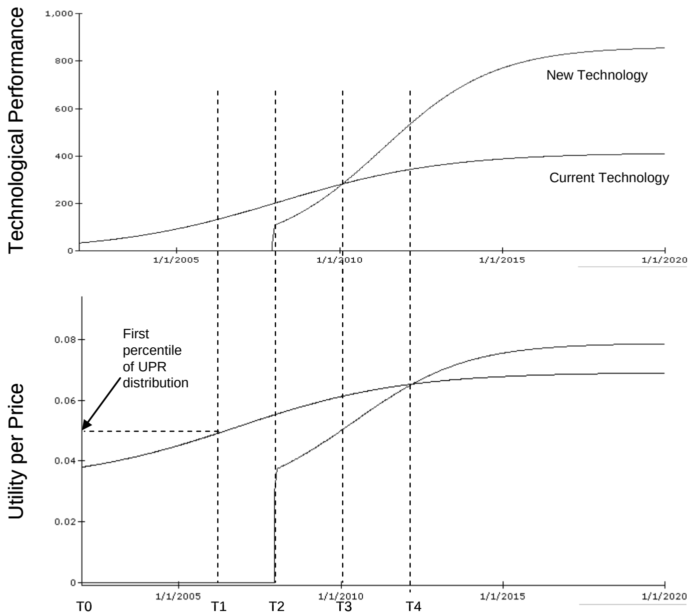


Figure 13: Performance and Utility Price trajectories of two competing technologies

This longitudinal perspective highlights five time periods with particular dynamics:

$T0 \leq t \leq T1$

Although the current technology CT offers a relatively low performance, the innovators who have the lowest UPR requirement in the market are the only ones interested in buying it.

$T1 \leq t \leq T2$

All the innovators have adopted, now the UPR of the current technology is high enough to be of interest to the followers. The innovators who have already adopted generate an information stream about this technology.

$T2 \leq t \leq T3$

The new technology NT is introduced at *T2*. It is still a crude version with performance and UPR significantly lower than for CT. Therefore, in the mass market, the situation is the same than during time period $T1 \leq t \leq T2$. However, if the new technology NT offers a new functionality which is valued in a niche market, then we could observe the first sales of this new technology in that niche market. This is in compliance with the view that new technologies very often emerge from outside the mass market.

$T3 \leq t \leq T4$

From *T3*, the new technology offers a better performance but still has a lower UPR than the current technology due to a higher price. Nevertheless, the innovators are performance hungry, so they generally tend to upgrade to the leading edge performance. This upgrading process of the innovators is controlled by the relative performance of their current technology in comparison to the leading edge NT and their uniform distribution of requirement for upgrading. These lead users, alongside with the niche markets' adopters, will generate the first stream of market information concerning this new technology. However, during this period the mass market followers are still joining the potential group based on the development of the UPR of the current technology and are not yet considering buying the second technology.

$T4 \leq t$
From time=*T4*, the new technology NT offers both a better performance and a better UPR than CT. Therefore, followers in the untapped market will become interested in buying the second technology and will join the potential adopters group of NT. However, it is assumed that followers who have already joined the potential adopters group of CT will keep on considering buying the first technology as they require more information to adopt and there will be more information about the first than the second technology. Followers that already possess CT will be interested in upgrading to the NT based on relative UPR.

Reference mode

We now present a more generic case where successive generations of technology emerge and compete against each other in the market place. The complete simulation model considers the evolution of the performance and the UPR of four generations of technology. Figure 15 to figure 18 give for each types of adopters the profiles of first purchases, physical renewal, and upgrading. All these are aggregated to obtain the complete life cycle of each technology.

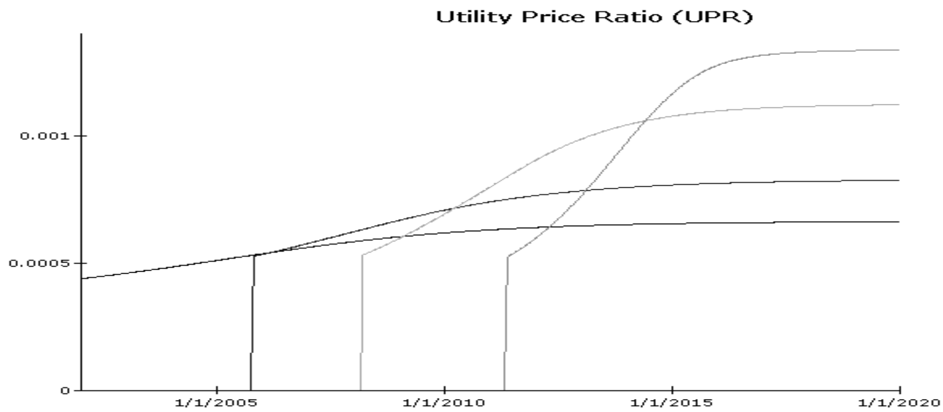


Figure 14: Evolution of Utility Price Ratio of four competing technologies

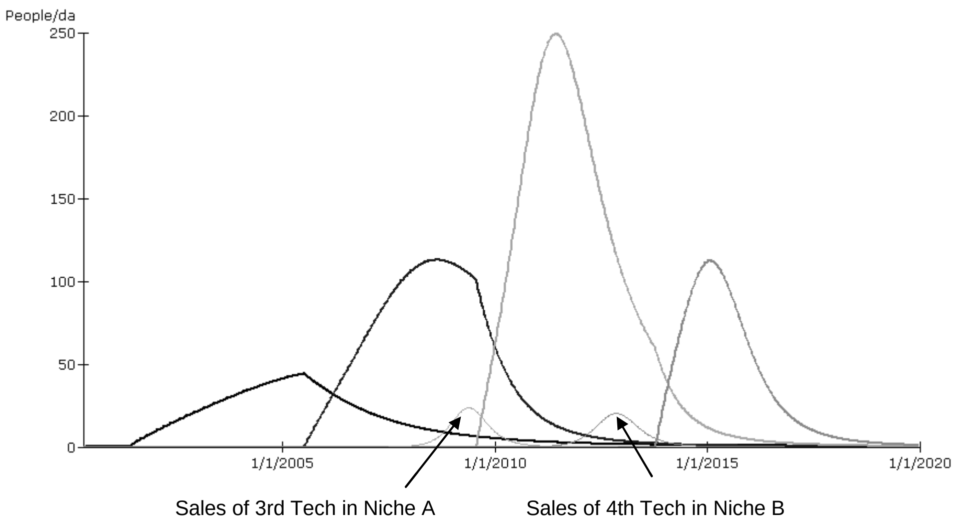


Figure 15: First time sales of four competing technologies

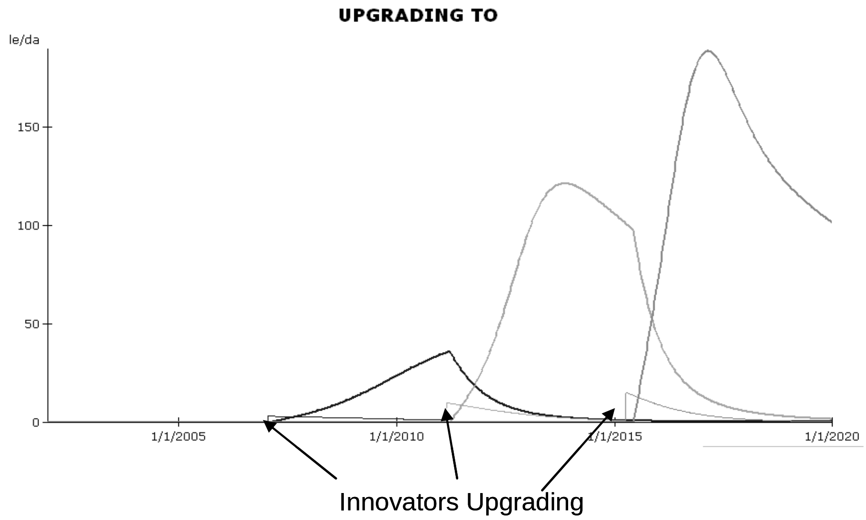


Figure 16: Upgrading to each of the four competing technologies

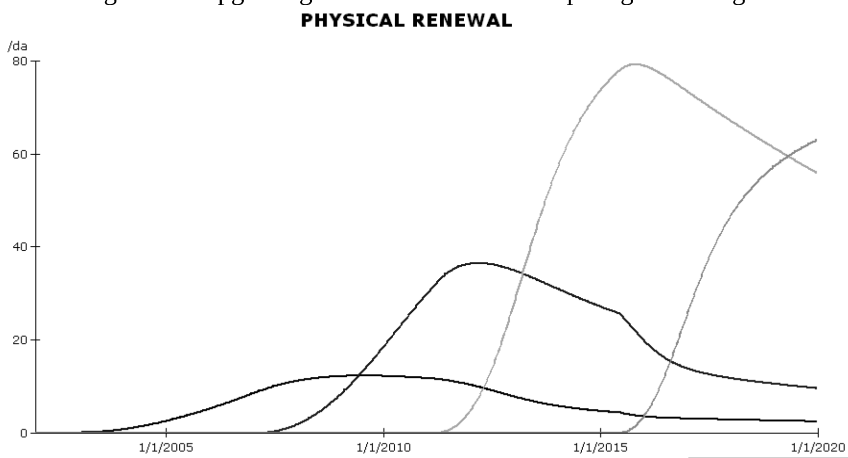


Figure 17: Physical renewal of four competing technologies among innovators and followers

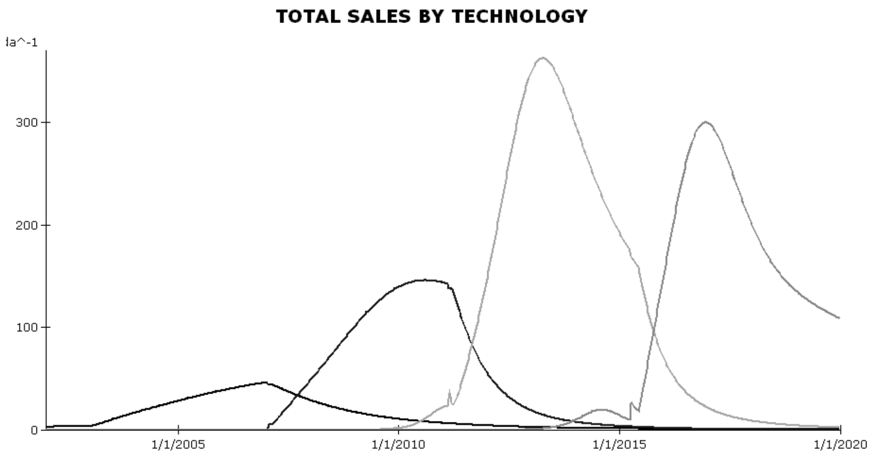


Figure 18: Total Sales of the four competing technologies

Our framework creates a link between the performance trajectories of successive generations of technology and market penetration. By aggregating all these user bases across all technologies, it also captures the growth of the market size as illustrated by figure 19, which shows a S-shaped diffusion pattern at the category level. Moreover, by normalizing by the total market size, we get the market share view of classical studies of technological models (e.g. Fisher and Pry models) as illustrated by figure 20. These models also have an inherent view of technological cycles which follow the sequence "emergence-growth-dominance". However, figure 19 also shows that when there are some overlaps between the successive generations of technology, each technology may not have time to reach complete dominance. This results in multi-level substitutions as discussed by (Kabir, Sharif et al., 1981).

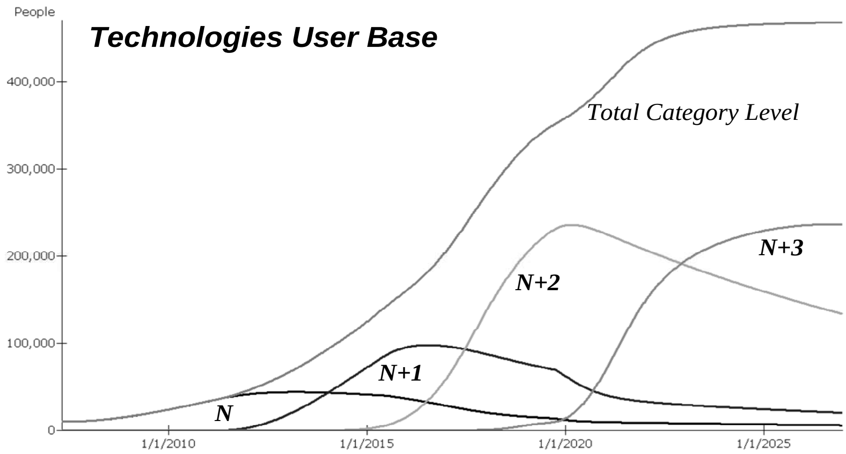


Figure 19: User bases for successive generations of technologies and diffusion at the category level

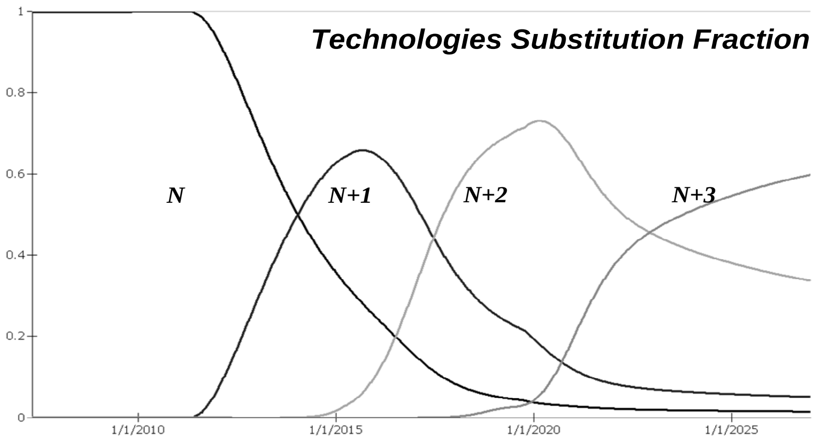


Figure 20: Fractional base of successive generations of technologies

Figure 19 and figure 20 show that our diffusion framework operationalized through a system dynamics simulation model can generate the diffusion pattern at the product category level, the diffusion of individual technologies as well as the substitution effects between successive generations.

5 Substitutive Drop

Our model's structure generates a recurring pattern whereby the sales of a technology *n* drop when technology *n+1* offers a better utility per price. In our model, technology *n+1* starts diffusing among niche markets despite lower performance or among innovators despite lower UPR. Once *n+1* offers a better UPR there is no more inflow into the group of potential adopters of technology *n* which thus starts depleting through actual adoption of technology *n*. It is the discontinuity induced by the sudden stoppage of inflow which generates the substitutive drop in the life cycle.

Norton and Bass multigenerational model

Norton and Bass have developed a formulation for the classical model which accounts for some substitution dynamics (Norton, 1986; Norton and Bass, 1987). The analytical formulation of the Norton and Bass model yields the generic behavior illustrated by figure 21 (Norton and Bass, 1987). We can see that their model does not replicate the dynamics described in our framework. Therefore, while it has a relatively good fit to historical data, the structure of their model seems to lack explanatory power.

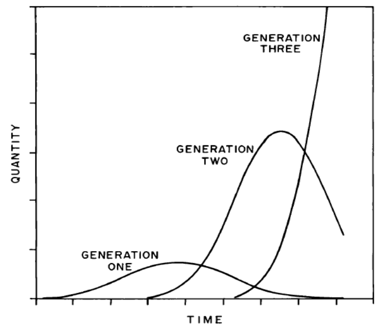


Figure 21: Generic behavior of the Norton and Bass multigenerational model

Norton and Bass applied their model to the successive generations of DRAM. The historical data of DRAM sales feature sudden drops for each generation that their model's structure does not replicate. Because they use an adaptation of the traditional Bass model, their analytical formulation cannot account for such a sudden drop in sales. We could say that classical models reproduce "dome" shape life cycles while strong substitution effects generate life cycles which look more like the "Sydney opera". Therefore, as shown in figure 22, the classical pattern of their curve can only be fitted to the data by anticipating the actual peak of sales by many quarters and by ignoring the pre-empted volume in their estimate of market potential. This is exactly the bias illustrated by figure 12.

Their model considers a smooth decrease in the rate of sales as for any usual Bass model diffusion curve. However, figure 23 shows that at quarter 23, the first derivative of sales (thousands/quarter²) of 4k DRAM started to drop during three quarters more drastically than anticipated by their model. This dramatic behavior is also evident for 16k DRAM at quarter 37.

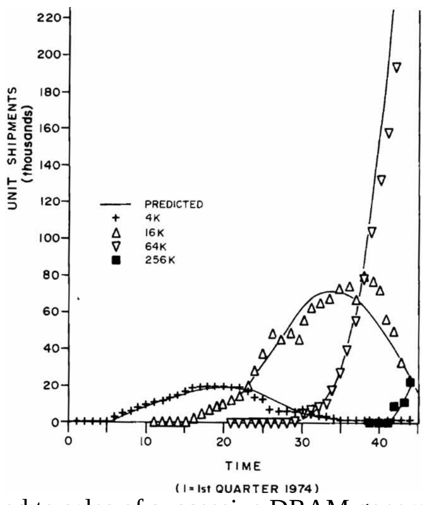


Figure 22: Classical model fitted to sales of successive DRAM generations (Norton and Bass, 1987)

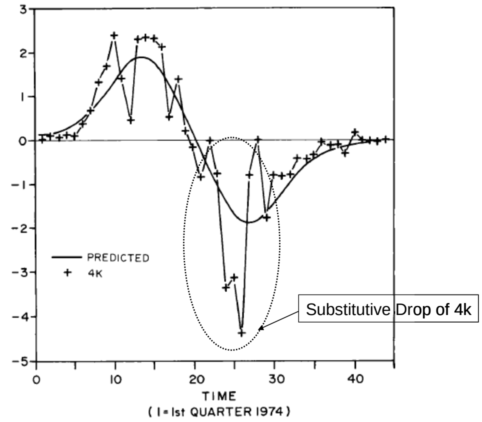


Figure 23: 4k DRAM – Missing the drop in the first derivative of actual sales (Norton and Bass, 1987)

Our model's explanation is that a substitutive drop occurs when technology *n+1* overtakes technology *n*. As illustrated by figure 24, our model replicates the same phenomenon for the first derivative of sales as observed in figure 23.

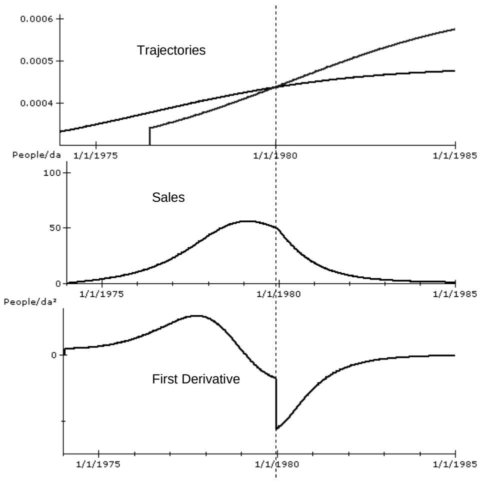


Figure 24: Our model: Structural evidence of substitutive drop (sales and first derivative of sales)

Historical data

To investigate this hypothesis, we turn, like Norton and Bass, to the historical data of performance, price and sales for the successive generations of DRAM since 1974. We obtained secondary data from Semico Research on the performance, price, and sales evolution for DRAM devices for the period 1974-2004.

Our theoretical framework assumes that the decision process of innovation adoption is based on expected utility. Given the role of memory devices in the computing power of computers, we assume that the value function for these devices is uniattribute in memory size. Moreover, we assume that the expected utility function for these devices is a linear function of performance (i.e. memory size). Thus, we operationalize the UPR evolution by computing the performance/price ratio of the successive DRAM generations. Annex 2 shows the evolution of the Mbits/$ for each generation of DRAM for the period covered. Annex 2 also identifies the time at which the trajectory of technology *n+1* overtakes the trajectory of technology *n*. For example, we can see that after taking the lead of Mbits/$ in 1979, the 16k devices were overtaken by the 64k generation in 1983.

Annex 3 shows that the life cycles of all the generations exhibit at sudden decrease that could not be matched by a smooth bell shape life cycle from a classical model. We compute the first differences in sales and, as in figure 23 and figure 24, identifies the time at which the drop occurred. Table 2 shows that over the ten generational changes, a substitutive drop occurred nine times when generation *n* overtook generation *n+1*. The only exception is the substitutive drop of 4M for 1M that occurred in 1991-1992, while the overtaking had happened in 1990-1991. We have no contextual data to explain this delay.

	Overtaking Mbits/$	Substitutive Drop
1978 -1979	1	1
1982 - 1983	1	1
1984 - 1985	1	1
1988 - 1989	1	1
1990 - 1991	1	0
1995 - 1996	1	1
1997 - 1998	1	1
2000 - 2001	1	1
2001 - 2002	1	1
2004 - 2005	1	1

Table 2: Relationship between performance / price overtaking and substitutive drops

6 Exploring technological and social dynamics

By accounting for the interdependencies between technological developments and market heterogeneity, our model's structure replicates dynamics not accessible to classical models of diffusion. Therefore, we further explore the capacity of the model to handle other dynamics induced by technological development or the social topology.

Sailing ship effect

As Rosenberg noted, the emergence of a competing technology is often a more effective agent in generating performance improvements in an existing technology than the more diffuse pressures of intra-industry competition (Rosenberg, 1976). This defensive surge allows to maintain a performance advantage over the new technology. However, the usual effect of such advances is only to postpone the dominant technology's displacement (Smith, 1992). This phenomenon is usually described by presenting the surge in the performance trajectory; but data are rarely provided to substantiate the delayed substitution.

One of the most famous example of this phenomenon is the introduction of extremely fast Clipper ships in 1845 when sailing ships were being threatened by the substitution from steam boats. We collected historical data[2], presented in figure 25, which show that these Clipper ships induced a 30-years delay in the substitution process. In 1875, improvements in steam and steel making technologies (compound engine, open-hearth furnaces, etc..) resumed the substitution dynamics.

[2] *U.S. Bureau of the Census*. (Carter, Gartner, Haines, Olmstead, Sutch and Wright, 2004).

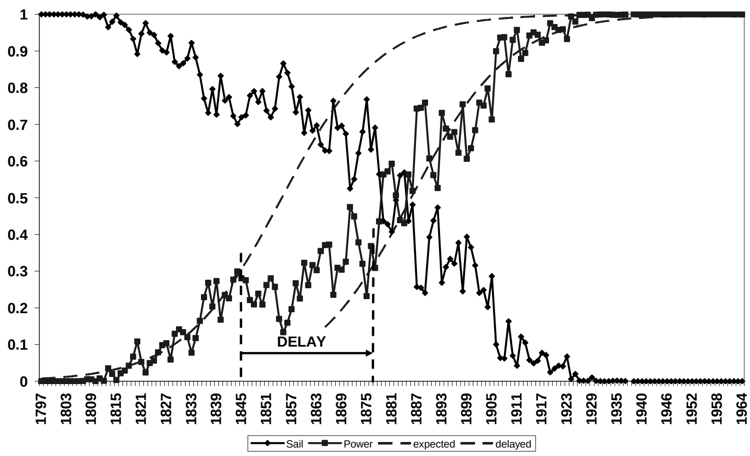


Figure 25: Example of sailing ship substitution – Sailing ships vs. Power (1797 – 1964)

We use our system dynamics model to explore the substitution dynamics induced by such a defensive surge in technological performance. We set up the model so that that, ceteris paribus, the emergence of technology *N+1* triggers a surge of performance from technology *N*, as illustrated by figure 26.

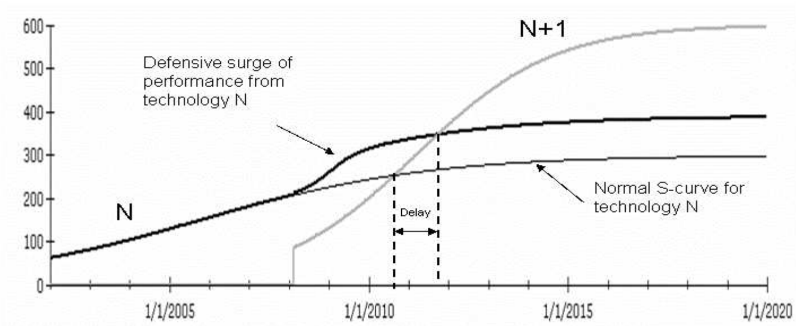


Figure 26: Model configuration for a technological defensive surge

Figure 27 presents the fractional base for each technology (normalized by total market size) generated by our model under these conditions. We can see that because our model operationalizes the link between technology trajectories and the innovation adoption decision process, it can generate a delayed substitution similar to the pattern observed in the historical data of marine cargo.

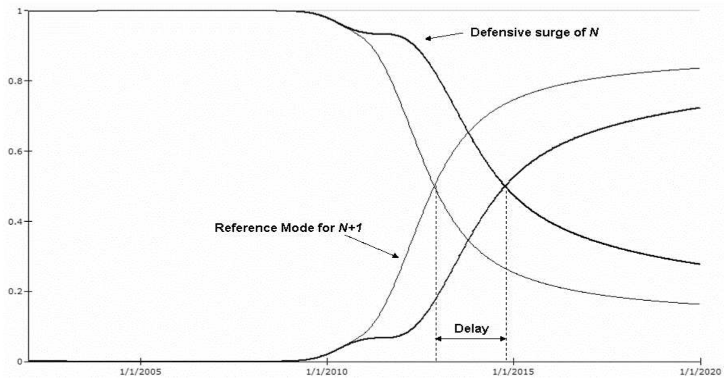


Figure 27: Simulation of a sailing ship substitution pattern

Accounting for social dynamics

In our model, the requirement distribution is considered normally distributed across the market population. We have assumed that the parameters of this distribution (μ_k, σ_k^2) are fixed. Our discussion has so far focused on the link between technological evolution and substitution dynamics through the multi-attribute expected utility framework. However, consumers' requirement thresholds are not static and evolve alongside their experience of and familiarity with a particular technological paradigm. The implications are that any estimation of parameters will be fit for a period of time only. In the model, the μ_k and σ_k^2 could be considered dynamics but a valid structural explanation could be difficult to construct. In his system dynamics model, Lyneis has modeled an exponential customer requirement trajectory and has derived conclusions with regards to either technology push or market pull dynamics (Lyneis, 1993).

The importance weights, w_k, of technological attributes are also considered to be fixed over the entire time period. However, it is clear that very strong social dynamics are at work here. Mary Tripsas has proposed a very interesting concept of preference trajectories and has shown that preference discontinuities could trigger technological change (Tripsas, 2005). These importance weights could also be linked to requirement thresholds. These thresholds could also be driven by technological improvements; this should be considered when studying substitutions over long time periods.

In our framework the utility derived from a technology $n+1$ is based on the perception of attribute levels and each technology is evaluated independently. However, Pistorius and Utterback have illustrated that there may exist multiple modes of interaction between technologies. To account for such effects our model would need to integrate the value of the installed base of technology $j<n+1$ into the utility function of technology $n+1$, such as in the classical case of network externalities. Schilling has shown how companies could try to influence the consumers' perception of the installed base to artificially increase the perceived utility (Schilling, 2003) and lead to a self fulfilling prophecy.

We have assumed that upgrading consumers will go through the same information gathering process before adopting a newer generation of technology. This in fact assumes that there is no familiarity effect, and that for each generation there is a knowledge or know-how switching cost that increases the perceived risk of adoption. We could easily assume that even though a technology is disruptive in terms of industrial dynamics, its diffusion among consumers could profit from familiarity acquired with previous technology. One such example is the diffusion of Internet in France which occurred later but quicker, maybe because the Minitel had been widely used and had developed familiarity with accessing distant services over a server. This is an occurrence of a prey-predator scenario based on familiarity in the Pistorius-Utterback framework. This familiarity could be integrated into our model by changing the critical amount of information i^* required for an upgrade. Hence depending on the interaction mode, one could change the a_{i^*} distribution parameter for upgrading as a function of the installed base of the previous technology.

Another assumption was that there was no forgetting during the process of gathering information for each technological innovation, i.e. $\frac{\partial i_n(t)}{\partial t} \geq 0$. Therefore, all the information gathered for a generation n will still be available even if the market is now dealing with generation *n+5*. A forgetting flow could be added to the stock of collected information.

Other important assumptions were made with regards to the communication patterns and the topology of the social system. Traditional models consider a fully interconnected system with a rather uniform view. In our model, this assumption was also made in that the information streams from different segments were flowing across the entire system and increased the amount of information available about a given technology. It is not evident that the information generated in a niche market may reach the followers into the main market. Rather, it has been often observed that a radical technology often emerges outside an industry and when it has sufficiently evolved it reaches the shores of an unaware industry. Our model can easily explore the implications of these communications interconnections by crossing the flow of information generated by respective groups into stocks of collected information. This is done by changing the λ_n parameter of technology *n* into a flexible $\lambda_{n,i,j}$ where *i* and *j* are subscripts of the social group of the emitter and receptor of word of mouth respectively. This allows to account for the effect of opinion leadership and the relevance of information (Dattee and Weil, 2005).

7 Conclusion

Our model's structure embeds a broader theoretical framework which accounts for the interactions between technological evolution, market adoption, and some social dynamics. The system dynamics methodology allows to build upon several advanced diffusion models. These models have powerful theoretical foundations but their analytical formulations impose solvability constraints and restrict their scope. Moreover, our model offers theoretical bases and clear formulation for some of the dynamics that were aggregated into generic coefficients by previous system dynamics models (Kabir, Sharif et al., 1981; Lyneis, 1993; Milling, 1996; Maier, 1998).

Our system dynamics model considers the heterogeneity of the market population (innovators, followers, niche), the innovation decision process based on expected utility (distribution of UPR requirement), the reduction of perceived risk of adoption through collecting information, plus renewal and upgrading. Our model's structure – multistage, multi innovation, restricted individual parameter, dynamic potential – is capable of generating both diffusion and substitution dynamics.

The link between technological evolution and market dynamics has allowed to identify what we call substitutive drops. The substitution dynamics between competing technologies generate life cycles with "Sydney opera" shape. When a new technology *n+1* emerges in the market, it does not offer the best performance, utility or price among all the alternatives. But when it does so after evolving along its trajectory, a sudden drop is visible in the sales of technology *n*. These dynamics were substantiated by historical data for the successive generations of DRAM devices from 1974 to 2004.

By broadening the scope of analysis and bringing together various theories of technological change, our model can explore more complex dynamics. Sterman insists that the goal of modeling is to expand the boundary of our models so that more and more of the unexplained variation in the behavior of a system is resolved into the theory (Sterman, 2000, p. 363). We showed that our model can replicate a delayed substitution induced by a defensive surge of the dominant technology; these are the classical dynamics of the "sailing ship effect".

The model can thus be used to explore the substitution dynamics induced by relaxing some of the classical underlying assumptions (timing of emergence, performance limit, social topology, etc.). This allows exploring the capability of the generic model to generate various patterns of substitution. According to Sterman, “this kind of family member test is particularly helpful when the class of systems the model addresses includes a wide range of different patterns of behavior [...] the more diverse the instances of a system a model can represent the more general the theory it embodies” (Sterman, 2000 p. 881).

References

Adner, R. and D. A. Levinthal (2001). "Demand heterogeneity and technology evolution: implications for product and process innovation." Management Science **47**(5): 611-628.

Bass, F. M. (1969). "A New Product Growth Model for Consumer Durables." Management Science **15**: 215-227.

Carter, S. B., S. S. Gartner, et al. (2004). Historical Statistics of the United States: Colonial Times to 1970. U. S. B. o. t. Census.

Chatterjee, R. and J. Eliashberg (1990). "The Innovation Diffusion Process in a Heterogeneous Population: A Micromodeling Approach." Management Science **36**(9): 1057-1079.

Christensen, C. M., F. F. Suarez, et al. (1998). "Strategies for Survival in Fast-Changing Industries." Management Science **44**(12): 207-220.

Dattee, B. and H. B. Weil (2005). "Dynamics of social factors in technological substitutions." Massachusetts Institute of Technology, Sloan School of Management(Working Paper 4599-05).

DeGroot, M. H. (1970). Optimal Statistical Decision. New York, McGraw-Hill.

Dekimpe, M. G., P. M. Parker, et al. (1998). Global Diffusion of Technological Innovations: A Couple-Hazard Approach, INSEAD.

Easingwood, C. J., V. Mahajan, et al. (1983). "A Nonuniform Influence Innovation Diffusion Model of New Product Acceptance." Marketing Science **2**(3): 273.

Fisher, J. C. and R. H. Pry (1971). "A simple substitution model of technological change." Technological Forecasting and Social Change **3**: 75-88.

Homer, J. B. and S. N. Finkelstein (1981). Modeling the Development and Dissemination of an Emerging Medical Technology. Sloan School of Management Working Paper. Cambridge, MA, Massachusetts Institute of Technology

Kabir, C., M. Sharif, et al. (1981). "System Dynamics Modeling for Forecasting Technological Substitution " Computer and Industrial Engineering **5**: 7-21.

Lattin, J. M. and J. H. Roberts (2000). The Application of an Individual-Level Diffusion Model Prior to Launch. Research Paper Series, Stanford University.

Lyneis, J. M. (1993). A Dynamic Model of Technology Diffusion. International System Dynamics Conference.

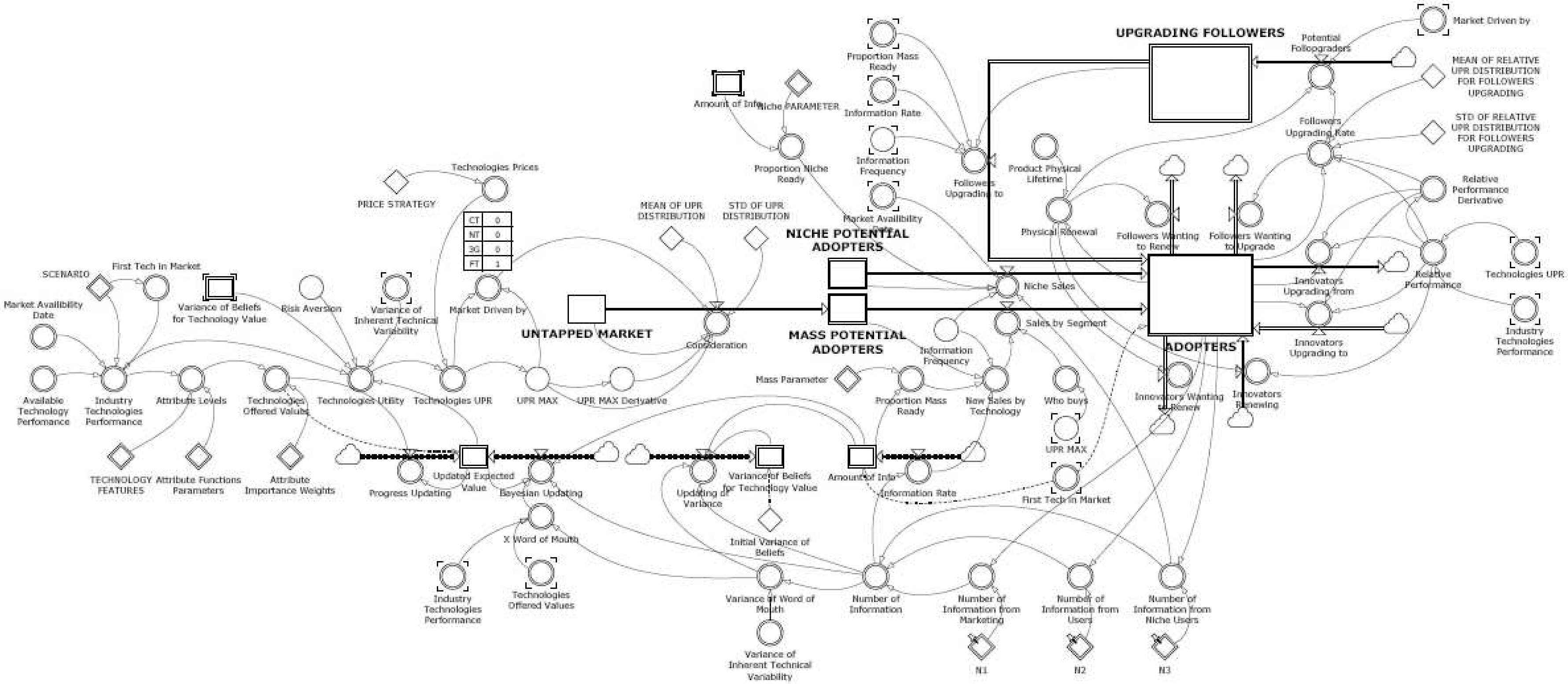
Maier, F. H. (1996). Substitution among Successive Product Generations - An almost Negelcted Problem in Innovation Diffusion Models, Universitat Mannheim.

Maier, F. H. (1998). "New product diffusion models in innovation management - a system dynamics perspective." System Dynamics Review **14**(4): 285-308.

Milling, P. M. (1996). "Modeling Innovation Processes for Decision Support and Management Simulation." System Dynamics Review **12**(3): 221-234.

Milling, P. M. (2002). "Understanding and Managing the Innovation Process." System Dynamics Review **18**(1): 73-86.
Norton, J. A. (1986). Growth, Diffusion, and Technological Substitution in Industrial Markets: An examination of the semiconductor industry. Dallas, The University of Texas. **PhD**.
Norton, J. A. and F. M. Bass (1987). "A Diffusion Theory Model of Adoption and Substitution for Successive Generations of High-Technology Products." Management Science **33**(9): 1069-1086.
Norton, J. A. and F. M. Bass (1992). "Evolution of Technological Generations: The Law of Capture." Sloan Management Review **Winter**: 66-77.
Parker, P. (1991). Choosing among Diffusion Models: Some Empirical Guidelines, INSEAD.
Pistorius, C. W. I. and J. M. Utterback (1995). "The Death Knells of Mature Technologies." Technological Forecasting and Social Changes **50**: 133-151.
Pistorius, C. W. I. and J. M. Utterback (1997). "Multi-Mode Interaction Among Technologies." Research Policy **26**: 67-84.
Roberts, J. H. (1984). A Multiattribute Utility Diffusion Model: Theory and application to the pre-launch forcasting of automobiles. Sloan School of Management. Cambridge, MA, Massachusetts Institute of Technology. **PhD**.
Roberts, J. H. and G. L. Urban (1988). "Modeling Multiattribute Utility, Risk and Beliefs Dynamics for New Consumer Durable Brand Choice." Management Science **34**(2): 167-185.
Rogers, E. M. (2003). Diffusion of Innovations. New York, The Free Press.
Rosenberg, N. (1976). "On technological expectations." The Economic Journal **86**(34): 523-535.
Schilling, M. A. (2003). "Technological Leapfrogging :Lessons from the U.S. Video Game Console Industry." California Management Review **45**(3): 6-32.
Smith, C. (1992). "Understanding technological substitution: Generic types, substitution dynamics, and influence strategies." Journal of Business Research: 279-302.
Sterman, J. D. (2000). Business Dynamics: System Thinking and Modeling for a Complex World. New York, Irwin McGraw-Hill.
Tripsas, M. (2005). Customer Preference Discontinuities: A Trigger for Radical Technological Change, Harvard Business School.
Weil, H. B. (2007). Application of system dynamics to corporate strategy: An evolution of issues and frameworks. International System Dynamics Conference. Boston, USA.
Weil, H. B. and J. M. Utterback (2005). The dynamics of innovative industries. International System Dynamics Conference. Boston.

Annex 1: System Dynamics simulation model



Annex 2: Evolution of Mbits/$ for the DRAM generations (1974-2004)

Mbit per $	4k	16k	64k	256Kbit	1Mbit	4Mbit	16Mbit	64Mbit	128Mbit	256Mbit	512Mbit	1Gbit	2Gbit
1974	0,00016												
1975	0,00052												
1976	0,00087	0,00030											
1977	0,00142	0,00069											
1978	0,00218	0,00173	0,00036										
1979	0,00204	0,00272	0,00054										
1980	0,00209	0,00342	0,00101										
1981	0,00257	0,00769	0,00456										
1982	0,00194	0,01348	0,01142	0,00152									
1983	0,00138	0,01532	0,01686	0,00363									
1984	0,00128	0,01456	0,02049	0,01206									
1985	0,00096	0,01206	0,05795	0,06687	0,00591								
1986			0,06297	0,11709	0,04019								
1987			0,05884	0,10378	0,05645								
1988			0,04305	0,06612	0,06206	0,01473							
1989			0,03842	0,06768	0,07729	0,03241							
1990			0,04526	0,10847	0,16203	0,11680							
1991			0,04424	0,14900	0,22389	0,24212	0,05488						
1992			0,05022	0,16120	0,30148	0,34145	0,12082						
1993				0,15574	0,29700	0,31869	0,20439						
1994				0,14222	0,24776	0,32044	0,28497	0,13333					
1995				0,11692	0,24867	0,31095	0,30161	0,15371					
1996				0,20558	0,31847	0,70024	1,03534	0,63611					
1997					0,52845	1,46517	2,45000	1,88823		0,23742			
1998					0,65677	1,61493	5,86942	6,67883	2,82513	0,53149			
1999					0,68629	1,40899	4,86182	8,41668	6,74712	3,21147			
2000					0,66667	1,15499	4,08402	9,60535	9,69289	5,53647			
2001						2,34343	10,62183	31,25490	43,73359	31,28189			
2002						3,20430	15,64694	41,64881	35,61440	39,18557	1,43284		
2003						3,51745	14,09115	42,14909	49,75299	57,15269	4,75093		
2004							14,15581	30,23843	45,58645	57,86258	49,65588	6,34218	
2005							13,35441	43,98579	63,22857	79,22405	95,92399	49,83965	
2006							13,55932	64,64646	84,76821	89,82456	130,61224	88,49588	6,18716

Keys:

Leading Phase
Substitution Phase

Annex 3: Shipments of successive generations of DRAM (1974-2008 forecast)