Discussion Paper No. 09-080

Knowledge Diffusion and Knowledge Transfer: Two Sides of the Medal

Torben Klarl

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 09-080

Knowledge Diffusion and Knowledge Transfer: Two Sides of the Medal

Torben Klarl

Non-technical summary

The new growth theory considers knowledge to be a decisive engine of economic growth. More precisely, knowledge is not used solely to the benefit of its originator, but generates positive side effects also for others, provided they have the capability to understand the transferred knowledge potential. Knowledge generation has also welfare implications for a country or a region. On a macroeconomic level, the implication of knowledge diffusion for growth seems straightforward, whereas on the microeconomic level the effect of knowledge diffusion seems more complex.

Before dealing with the question of how knowledge diffusion can be adequately modeled in a concise microeconomic approach, two important aspects have to be distinguished from each other: knowledge diffusion and knowledge transfer. Although these terminologies are well established in the relevant literature, it is not or rather inaccurately acknowledged within the knowledge diffusion modeling context. Knowledge transfer is associated with the exchange of knowledge within networks, which consists of innovators and imitators of knowledge. On the contrary, knowledge diffusion describes the diffusion of knowledge within the group of innovators and imitators. Apparently, knowledge transfer can accelerate but is not a necessary condition for knowledge diffusion. From this point of view, welfare implications are mainly expected from knowledge diffusion, which can be indirectly enforced by knowledge transfer. Therefore, the intensity of knowledge networks should also affect the diffusion pattern of knowledge.

This paper proposes a model that can explain endogenously the knowledge diffusion patterns induced by network effects. In this way, to the best of my knowledge this is the first attempt to discuss both, knowledge diffusion and knowledge transfer in a comprehensive framework. A key result shows that unimodal diffusion patterns are generated by strong network effects, whereas bimodal diffusion patterns occur due to weaker network effects. Thus, the stronger network effects, the faster is knowledge diffusion. Furthermore, this model assumes that the knowledge diffusion process is embedded in a stochastic environment. Particularly, at the beginning and in the middle the uncertainty of adopting new knowledge is larger than at the end

of the diffusion process. From an econometric point of view, this can be modelled via heteroscedastic errors in the error term. A further pleasant feature of this model is that it can be directly estimated with a seemingly unrelated regression (SUR) approach.

Nicht-technische Zusammenfassung

Wissen als Produktionsfaktor wird in der neuen Wachstumstheorie als eine elementare Erklärungsgröße für wirtschaftliches Wachstum gesehen. Dabei wird der Diffusion von Wissen ein hohes Maß an Aufmerksamkeit gewidmet, da Wissen nicht nur vom Produzenten genutzt werden kann, sondern auch über Transferkanäle anderen Nutzen stiften kann. Je stärker die Wissensdiffusion dabei ausgeprägt ist, desto mehr Wirtschaftssubjekte können dieses Wissenspotential nutzen, vorausgesetzt die Wirtschaftssubjekte können das gesendete Wissen verstehen und verarbeiten. Dies hat natürlich auch Implikationen für die Gesamtwohlfahrt eines Landes. So einfach diese Wirkungskette auf der makroökonomischen Ebene klingen mag, so komplex ist die Beantwortung der Frage der Wissensdiffusion auf der mikroökonomischen Ebene.

Bevor man sich der Frage widmet, wie Wissensdiffusion auf der mikroökonomischen Ebene modelliert werden kann, sind zunächst zwei Aspekte stringent voneinander zu trennen: Wissensdiffusion auf der einen, und Wissenstransfer auf der anderen Seite. Obwohl diese Termini in der Wissensdiffusionsliteratur weitestgehend bekannt sind, fehlt in der Wissensdiffusionsmodellierung bis dato eine wirkliche Trennung beider Aspekte. Wissenstransfer meint im Wesentlichen den Austausch von Wissen in Netzwerken zwischen Innovatoren und Imitatoren von Wissen, während Wissensdiffusion die Verbreitung von Wissen innerhalb der Gruppe von Innovatoren und Imitatoren bezeichnet. Es ist offenkundig, dass Netzwerkeffekte die Wissensdiffusion beschleunigen können, andererseits ist auch ohne Wissenstransfer über Netzwerkeffekte eine Verbreitung von Wissen möglich. Wohlfahrtstheoretische Implikationen sind daher nur von der Wissensdiffusion zu erwarten, die indirekt über Wissenstransfers über Netzwerke verstärkt werden können. Die Intensität von Netzwerkeffekten sollte sich demnach auch im Diffusionsmuster von Wissen widerspiegeln.

Mit dem in diesem Aufsatz vorgestellten Modell ist es mithin möglich, Diffusionsmuster endogen durch Netzwerkeffekte zu erklären. Es zeigt sich, dass stärkere Netzwerkeffekte unimodale Diffusionsmuster von Wissen erzeugen, während schwache Netzwerkeffekte auf bimodale Diffusionsmuster schließen lassen. Starke Netzwerkeffekte führen damit zu einer schnelleren Verbreitung von Wissen. Zudem wird im Rahmen dieses Aufsatzes der Tatsache Rechnung getragen, dass der Wissensdiffusionsprozess stochastisch ist und insbesondere zu Beginn und in der Mitte die Adoptionsunsicherheit größer ist als am Ende des Diffusionsprozesses. Im ökonometrischen Kontext kann diese Unsicherheit mittels heteroskedastischer Störterme modelliert werden. Ein weiterer Vorteil des theoretischen Modells liegt darin, dass es sich mittels eines *Seemingly Unrelated Regression* (SUR)-Ansatzes direkt schätzen lässt.

Knowledge diffusion and knowledge transfer: two sides of the medal

Torben Klarl[*][†]

December 8, 2009

Abstract

Understanding the way in which knowledge is technically produced and transferred, and how its diffusion path can be characterized is of fundamental importance for the performance of an economy. Although this fact seems to be plausible ex ante, the relevant literature so far has paid less attention investigating the microeconomic link between knowledge transfer and knowledge diffusion in a comprehensive approach. The aim of this paper is to highlight the link between knowledge transfer, knowledge diffusion and network effects in a stochastic environment, because the adoption decision of new knowledge should be treated as a stochastic event. For this reason, a new knowledge diffusion model in the line of Bass (1969) has been put forward, which integrates knowledge diffusion and knowledge transfer. The advantage of the proposed model is twofold. From a theoretical point of view, not only the so-called unimodal diffusion phenomena can be modelled, but also bimodal diffusion phenomena can be obtained. From an empirical point of view, the model which incorporates heteroscedastic errors and mean reverting behaviour can be theoretically estimated directly within a standard SUR context.

Keywords: Knowledge diffusion, knowledge transfer, SUR

JEL Classification Number: C50, C51, C61, D83

[*]Address: Centre for European Economic Research (ZEW), P.O. Box 103443, D-68034 Mannheim, Germany. Email: klarl@zew.de

[†]Address: Department of Economics, University of Augsburg, Universitätsstraße 16, D-86159 Augsburg, Germany. Email: torben.alexander.klarl@wiwi.uni-augsburg.de

1 Motivation

To know how knowledge is technically produced and to understand its diffusion path is of fundamental importance in the innovative process. In an economic context, knowledge itself can be embodied in new products, or it can be approximated by citation of scientific publications[1], but loosely spoken there is no clear-cut definition of what knowledge constitutes. What we know, first, is that technological knowledge is often not transferred as itself, but instead within new technologies - via licensing or through FDI - for instance. Thus if we talk about knowledge diffusion, it is either a direct transfer in the sense of human capital transfer or more indirectly linked with the diffusion of new technologies and of intermediate and capital goods, as Rivera-Batiz and Romer (1991) have argued. This model follows the first direction and assumes direct knowledge diffusion.

Second, we know that knowledge diffusion and adoption is not a homogeneous process over the entire distribution of potential adopters. In a simplistic *homo economicus* world, where everybody knows everything from the beginning, or in a world with a less strict assumption that everybody can learn anything with probability one, diffusion of knowledge can be associated with a picture of dropping colour in a glass of water and waiting until the colour is more or less uniformly distributed. In such a world, the question of what kind of knowledge diffuses easily and what kind diffuses with difficulty is obsolete.

Assuming that the world is not perfect with respect to learning abilities and information potentials for instance, however, makes the question of what determines and fosters knowledge diffusion relevant and important. Polany (1967) takes this question seriously and separates implicit knowledge from explicit knowledge. Explicit knowledge can be transferred without any limits, whereas implicit knowledge -labelled tacit- can not. If knowledge is partly or completely tacit, its diffusion depends on the specific characteristics of the individuals. Thus, some people are more in touch with new developments than others. This is especially the case for two important groups of adopters that play an important role in the diffusion process:

[1] Refer to Fok and Franses (2007), for instance.

the innovators and imitators of new knowledge.

Bass (1969) in his seminal work and others such as Easingwood et al. (1983), Mahajan and Peterson (1985) and Mahajan and Wind (1986), p. xiii, have mentioned that innovators and imitators behave differently in the diffusion process. This assumption is reasonable because innovators and imitators have different intentions to adopt. Following Kalish (1985), one can differentiate between the so called *search attributes* and *experience attributes.* The innovators need only search information to adopt the new knowledge, while imitators require experience type information before they adopt. As noted by Gatignon and Robertson (1985) and Rogers (1983) the speed of diffusion of knowledge depends on several characteristics, including complexity, relative advantage, status value or observability etc.. These characteristics influence innovators and imitators in different ways. However, most of existing studies fail to highlight the different behaviour of these two groups.

Although Schmalen (1982) has mentioned that innovators' and imitators' behaviour regarding their adoption decision differs, he does not capture this fact in a notational form. The famous so-called *two compartment model*, proposed by Tanny and Derzko (1988) goes in line with the model of Schmalen (1982), but their definition of *innovators* and *imitators* seems not to be clear-cut: *innovators* adopt because of learning effects driven by external information, whereas imitators adopt because of external knowledge by prior adopters. This model hypothesises that innovators adopt because of search information, while imitators adopt due to experience but also due to knowledge transfer which can be justified with the existence of networks.

It is therefore assumed that the adoption decision is also influenced by networks, which are a necessary condition for knowledge transfer between both groups. If a dense network structure is available, knowledge transfer is easier and thus the imitator should adopt faster. On the contrary, if networks do not exist, knowledge transfer is excluded and thus adoption takes place later. The latter scenario often leads to the so called *chasm* pattern between early and late adoptions, which is extensively discussed in Moore (2002) and sometimes mentioned in diffusion related literature[2].

Therefore, network effects should also have an influence on the shape of the adoption curve, which is in the latter case not necessarily unimodal but bimodal for the entire market. The novelty is that the introduced model treats the chasm pattern as an endogenous number. The literature is still silent about this topic, and only a few papers take the network effects into account, including Van den Bulte and Lilien (2001), Van den Bulte and Joshi (2007), Hill et al. (2006) and Golendberg et al. (2006).

The aim of this paper can be laid out as follows: on the basis of Golendberg et al. (2006), Van den Bulte and Joshi (2007), Boswijk and Franses (2005) and Boswijk et al. (2006) a knowledge diffusion model is set up, which includes the behaviour of both innovators and imitators. The model is able to replicate unimodal and bimodal adoption patterns. Which pattern occurs depends on whether or not network effects play a crucial role within the diffusion process. Additionally, the model will be extended into a stochastic knowledge diffusion model to capture the idea that uncertainty of adoption is a function of time, which means at the end of the diffusion process uncertainty regarding the adoption should tend to zero, while at the beginning and in the middle of the diffusion process uncertainty of adoption is high. Another feature of the proposed model is that it can be tested empirically within a SUR context.

The paper is structured as follows: In the second section, I start off with an introduction and discussion of the Bass (1969) model. In the third section, a deterministic knowledge diffusion model is set up. After discussing the solution of this model the solution's stability has been investigated. The fourth section embeds the deterministic knowledge diffusion model into a stochastic frame. Additionally, a sensitivity analysis of the stochastic model is performed to derive some model implications. The fifth section conducts a simulation study of both the deterministic and stochastic models. Before giving final remarks and highlighting avenues for further research in the seventh section, I provide some econometric remarks regarding the estimation

[2] Refer to Van den Bulte and Joshi (2007) for instance.

of the stochastic knowledge diffusion model in section six.

2 The Bass diffusion model

The Bass (1969) model, loosely spoken, describes how a new product or technique is adopted over time via interaction between potential and de facto adopters or users. Adoption stops if the market saturation level m has been reached, which means that every potential adopter has become a de facto adopter. For each potential adopter the time of adoption is random, i.e., ex ante the potential adopter does not know when he will adopt the product. In statistical terms, time of adoption is a random variable with a distribution function $F(t)$ and the corresponding density $f(t)$. The Bass (1969) model assumes that the portion of potential adopters who adopt at time t, given that they have not adopted yet, can be written as a linear function of adopters as:

$$\frac{f(t)}{[1-F(t)]} = p + qF(t). \tag{1}$$

The left hand side of equation (1) can also be interpreted as the hazard rate of potential adopters. The parameter p denotes the probability that a potential adopter adopts at t influenced by external factors, such as word of mouth influence through the de facto adopters. On the contrary, q can be interpreted as the internal probability that a potential adopter adopts at t. The latter adoption decision depends solely on the internal influence caused by the group of de facto adopters.
The time-dependent diffusion process[3] of the Bass (1969) model can be written as a differential equation:[4]

$$\frac{dF(t)}{dt} = f(t) = [p + qF(t)][1 - F(t)], \tag{2}$$

[3] A mathematical diffusion function can be expressed as the solution $y = y(t)$ of a deterministic differential equation $\frac{dy}{dt} = f(y,t)$. $f(\cdot)$ describes the pattern of the diffusion path and y gives information about the evolution of the diffusion process over time. Thus $f(\cdot)$ is a dependent function of y and diffusion time t. This is the basic idea of modelling diffusion paths.

[4] Refer to Kalish and Sen (1986) and Mahajan et al. (1984), for instance.

which can be interpreted as follows. On the left hand side of equation (2) we can find the rate of change with respect to time t of the cumulative number of adopters. This is equal to the hazard rate $[p+qF(t)]$ times the adopters which have not adopted yet in t. Consequently, $[1-F(t)]$ denotes the fraction of potential adopters. If $p \to 0$ we obtain a diffusion process which is entirely driven by internal influence of adopters in t, whereas for $q \to 0$ the diffusion process depends solely on external influence. In general, a mixture influence model is assumed, i.e., $\{p,q\} \in (0,1)$.
Labelling the cumulative number of adopters at t as $N(t) = mF(t)$, the alteration rate of adopters is given by

$$n(t) \equiv \frac{dN(t)}{dt} = m\frac{dF(t)}{dt} = mf(t), \tag{3}$$

or inserting (2) in (3) and noting that $N(t) = mF(t)$ yields

$$n(t) \equiv \frac{dN(t)}{dt} = m\left\{\left[p + q\frac{N(t)}{m}\right]\left[1 - \frac{N(t)}{m}\right]\right\}, \tag{4}$$

or

$$n(t) \equiv \frac{dN(t)}{dt} = \left[p + q\frac{N(t)}{m}\right][m - N(t)] = \chi(t)[m - N(t)], \tag{5}$$

with $N(t) = \int_{t_0}^{t} n(t)dt$. The last derived equation is the so called *Ricatti*-differential equation with constant coefficients. Equation (5) can be interpreted as the rate of change with respect to time t of the cumulative number of adopters, which is equal to a time dependent variable $\chi(t) = \left[p + q\frac{N(t)}{m}\right]$, which covers the mixture influence of adoption, governed by $\{p,q\} \in (0,1)$, times the cumulative number of potential adopters in t given by $[m-N(t)]$. From equation (5) we observe that the change rate of cumulative adopters is zero, given the number of potential adopters equals the number of cumulative adopters which is equal to the postulation that $[m-N(t)] = 0$. The solution of (5) for the cumulative number of adopters is given by:

$$N(t) = mF(t) = m\left[\frac{1 - exp\{-(p+q)t\}}{1 + \frac{q}{p}exp\{-(p+q)t\}}\right], \tag{6}$$

and for the adoption in t:

$$n(t) = mf(t) = m\left[\frac{p(p+q)^2 exp\{-(p+q)t\}}{(p+qexp\{-(p+q)t\})^2}\right]. \tag{7}$$

The question that remains is how to translate this theoretical model into practical application. The number of adopters is usually measured in discrete values, whereas the above derived diffusion equation (5) is written in continuous time. Therefore, Bass (1969) applies a simple *Euler*-discretisation scheme to obtain the following discrete time difference equation of the continuous time differential equation (5):

$$N(t) = N(t-1) + \left[p + q\frac{N(t-1)}{m}\right][m - N(t-1)]. \tag{8}$$

Due to its parsimonious specification, the Bass (1969) diffusion model and its extensions are popular in diffusion research[5]. It should be mentioned that from equation (8) it is quite clear that the Bass (1969) model is primarily very attractive for empirical application, especially for out-of-sample forecasts[6], because equation (8) can be estimated and tested without any relevant[7] modifications[8]. Although the Bass (1969) model seems to be very intuitive and well established both in theoretical and empirical application, there are several drawbacks, particularly when it comes to incorporating the diffusion of knowledge in new products.

Bass (1969) mentions, that innovators' and imitators' behaviour is driven by different aspects in the diffusion process. This assumption is reasonable ex ante, because it is justified that each subgroup of adopters, the innovators and imitators, have different intentions to adopt. Following Kalish (1985), one can differentiate between the so called search attributes and experience attributes. As a consequence of that, the innovators need only search information to adopt knowledge, while the imitators

[5]Refer to Parker (1994), Mahajan et al. (1990), Mahajan et al. (1993), Sultan et al. (1990) and Mahajan et al. (2000) for an overview.

[6]For instance, refer to Bass (1993) and Bass (1995) for this topic.

[7]Relevant in this context means that an error term must be added to this model.

[8]It should be mentioned that there is a large bulk of paper which discuss estimation strategies for the Bass model. Refer to Boswijk and Franses (2005) for a discussion of that topic.

require experience type information before they adopt. As noted by Rogers (1983) among others, the speed of diffusion of knowledge depends on several characteristics, such as complexity, relative advantage, status value and observability. These characteristics influence innovators and imitators in different ways.

Additionally, communication between these two types and thus network effects create a second channel which influence imitators adoption decision.[9] This fundamental assumption is not reflected in equation (8), although it is of central importance for a micro founded theory of knowledge diffusion.[10]

A second limitation of the Bass (1969) model stems from the fact that it inherently assumes a bell-shaped, single-peak adoption curve. Certainly, this could be the case for some specific kinds of products, but this assumption is not universally valid. As Kluyver (1977) has pointed out, that one major drawback of such (diffusion type) models is that only unimodal phenomena can be fitted. If one refers to the empirical literature, there is strong evidence that life cycle of innovations fits to a more bimodal pattern[11]. This is due to the fact that in the early stages of an innovation life cycle innovators' demand leads to a sharp rise, followed by a plateau or a fall in adoption until imitators cause a second but delayed rise of adoptions.

What causes the second, delayed rise of adoption? One possible answer is an aspect, which is not yet implemented in the existing diffusion frameworks: due to the phenomenon of knowledge transfer, which is often closely related to network effects. As highlighted by Audretsch and Feldman (1996), strong network effects should considerably enhance knowledge transfer. From this point of view, it seems reasonable to incorporate the aspect of knowledge transfer into a microeconomic model of knowledge diffusion and to investigate further the effects of the degree of tightness of networks on knowledge diffusion. In other words, one has to distinguish between pure knowledge transfer, which is for instance practiced via face-to-face communication and knowledge diffusion. Knowledge transfer must not necessarily influence the adoption decision but it can. In this context it is meant that knowledge

[9] Refer to Gladwell (2000), Moore (1995), Rosen (2000) and Slywotzky and Shaprio (1993).

[10] Already Jeuland (1981) has pointed out this fact.

[11] See Rink and Swan (1979) and Tellis and Crawford (1981).

transfer is only possible if the knowledge is transferable, for instance, via face-to-face communication. This second limitation defines the central focus of this paper: does and if yes in which direction does knowledge transfer via networks influence the adoption decision of imitators and innovators in microeconomic model of knowledge diffusion?

To address this question, in the next section a more general Bass (1969) type model is set up, which first includes a heterogeneous potential adopter group that is split in innovators and imitators. Furthermore, the new model formulation also includes network effects, which are not symmetric: it is assumed that imitators can benefit from information about the adoption of knowledge from the innovators. Thus, the often mentioned effects of knowledge transfer via network effects and its effect on knowledge diffusion, embodied by the adoption of a new knowledge are incorporated in the model setup. In this manner, it is possible both to replicate unimodal as well as bimodal shapes of the adoption curves. The shape of the curve only depends on the easiness of knowledge transfer. The easier knowledge transfer, the faster should knowledge diffusion be and the lower the probability that bimodal adoption pattern or so-called chasm pattern between early and the later parts of the adoption curve occurs[12].

3 Deterministic knowledge diffusion model

In this section a deterministic diffusion model with its relevant elements is introduced.

3.1 Setup

The group of adopters $N(t)$ is separated into subgroups, $N(t)_k$ of innovators and imitators with $k = \{1, 2\}$.[13] $k = 1$ represents the subgroup of innovators, whereas $k = 2$ symbolises the group of imitators. The key idea is to incorporate a communication channel between these two subgroups, which should cover the tightness of

[12] See to Van den Bulte and Joshi (2007).

[13] In the following time index t is only used if clarity demands it.

the knowledge exchange network. In this way, an asymmetric communication flow is created, because per definition the subgroup of innovators could not learn anything about the subgroup of imitators regarding their adoption decision. Innovators by definition are the first entering the market. To be more precise, the model contains a tightness parameter $q_{12} \in (0,1)$ which stands for the communication probability between the group of innovators and the group of imitators in the knowledge exchange network. The diffusion process for innovators $N(t)_1$ is similar to the Bass diffusion equation (5) and can be written as:

$$\frac{dN_1}{dt} = \left[p_1 + \left(q_1 \frac{N_1}{m_1}\right)\right][m_1 - N_1]. \tag{9}$$

The diffusion process for the imitators $N(t)_2$ instead should be written as:

$$\frac{dN_2}{dt} = \left[p_2 + \left(q_2 \frac{N_2}{(m_1 + m_2)}\right) + \left(q_{12} \frac{N_1}{(m_1 + m_2)}\right)\right][m_2 - N_2]. \tag{10}$$

Therefore, the change rate of cumulative group of imitators $\frac{dN_2}{dt}$ is also affected by network effects. If q_{12}=0, then innovators' and imitators' adoption are independent from each other, but still not symmetric, because even if $q_{12} \to 0$, the entire market saturation level $m_1 + m_2$ is of importance for the imitators.

These two model segments (9) and (10) can be stacked into a system of equations as follows:

$$\begin{bmatrix} \frac{dN_1}{dt} \\ \frac{dN_2}{dt} \end{bmatrix} = \begin{bmatrix} \left[p_1 + \left(q_1 \frac{N_1}{m_1}\right)\right] & 0 \\ 0 & \left[p_2 + \left(q_2 \frac{N_2}{(m_1+m_2)}\right)\right] \end{bmatrix} \begin{bmatrix} [m_1 - N_1] \\ [m_2 - N_2] \end{bmatrix} +$$

$$+ \begin{bmatrix} 0 & 0 \\ 0 & \left(q_{12} \frac{N_1}{(m_1+m_2)}\right) \end{bmatrix} \times \begin{bmatrix} [m_1 - N_1] \\ [m_2 - N_2] \end{bmatrix}, \tag{11}$$

or in a compact manner:

$$\dot{\mathbf{N}} = \mathbf{\Xi} a + \mathbf{\Pi} b. \tag{12}$$

From (11) we should notice that information flow is asymmetric, because the first diagonal element of $\mathbf{\Pi}$ in (12) is zero and thus $\mathbf{\Pi}$ is an upper triangular matrix. The next section deals with the solution of system (12).

3.2 Solution

Given $N(0)_1 = 0$, the solution of our differential equation system for $N(t)_1$ can be written as:

$$N(t)_1 = mF(t) = m\left[\frac{1 - exp\{-(p_1 + q_1)t\}}{1 + \frac{q_1}{p_1}exp\{-(p_1 + q_1)t\}}\right]. \tag{13}$$

In contrast to the solution of $N(t)_1$, the derivation of solution for $N(t)_2$ is cumbersome and can be found in the appendix represented at the end of the paper.

3.3 Stability

Before we proceed the equilibrium points of model (11) or (12) must be identified, and, additionally, their stability should be examined.

Proposition: Given the assumption that the partial derivatives, $\frac{dN_1}{dt}$, and $\frac{dN_2}{dt}$ exist and that $\frac{dN_1}{dt}$ and $\frac{dN_2}{dt}$, hold simultaneously $\forall t$, system (11) has a unique steady state vector $\mathbf{S}$, which contains N_1^* and N_2^* in the long run.□

Proof: An optimal steady state vector $\mathbf{S}$ exists if and only if $\frac{dN_1}{dt} = \frac{dN_2}{dt} = 0$ holds. This is realized, if

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \left[p_1 + \left(q_1 \frac{N_1}{m_1}\right)\right] & 0 \\ 0 & \left[p_2 + \left(q_2 \frac{N_2}{(m_1+m_2)}\right) + \left(q_{12} \frac{N_1}{(m_1+m_2)}\right)\right] \end{bmatrix} \begin{bmatrix} [m_1 - N_1] \\ [m_2 - N_2] \end{bmatrix}. \tag{14}$$

To find the elements for the steady state vector, the first equation from the derived system (12) has been examined first. Given $\frac{dN_1}{dt} = 0$, this equation can be written as follows:

$$\left[p_1 + \left(q_1 \frac{N_1}{m_1}\right)\right][m_1 - N_1] = 0. \tag{15}$$

An equilibrium is found if $\frac{dN_1}{dt} = 0$ holds. Thus, if $m_1 = N_1^*$, then $\frac{dN_1}{dt} = 0$. If $m_1 = N_1^*$ then the number of innovators of new knowledge have realized their market saturation level m_1, which implies that every potential innovator has adopted the new source of knowledge.

Second, if

$$N_1^* = \frac{-m_1 p_1}{q_1} < 0, \tag{16}$$

equation (15) is zero again and thus $\frac{dN_1}{dt} = 0$ also holds. Note that this equilibrium can be ruled out because $N_1 > 0$ by definition.

Let us now turn to the second equation of system (12), which can be written as

$$\left[p_2 + \left(q_2 \frac{N_2}{(m_1 + m_2)}\right) + \left(q_{12} \frac{N_1}{(m_1 + m_2)}\right)\right][m_2 - N_2] = 0, \tag{17}$$

given $\frac{dN_2}{dt} = 0$.

Again, if $m_2 = N_2^*$ then $\frac{dN_2}{dt} = 0$, which implies again, that the number of imitators have reached their market saturation level m_2. Additionally, if

$$N_2^* = -\frac{1}{q_2}\left[p_2 m + q_{12} N_1^*\right] < 0, \tag{18}$$

then $\frac{dN_2}{dt} = 0$ holds again. This equilibrium can be ruled out ex ante because $N_2 > 0$ by definition.

From this discussion it is possible to derive four long run equilibria: the first equilibrium is given by

$$m_1 = N_1^* \text{ and } m_2 = N_2^*. \tag{19}$$

This is the case, when both, the innovators and imitators have reached their specific market saturation levels.

The second equilibrium is obtained if

$$N_1^* = \frac{-m_1 p_1}{q_1} \text{ and } \; m_2 = N_2^*. \tag{20}$$

The third equilibrium is characterised by

$$N_1^* = \frac{-m_1 p_1}{q_1} \text{ and } \; N_2^* = -\frac{1}{q_2}\left[p_2 m + q_{12} N_1^*\right]. \tag{21}$$

Noting the fact, that $N_1^* = \frac{-m_1 p_1}{q_1}$ and inserting this expression in $N_2^* = -\frac{1}{q_2}\left[p_2 m + q_{12} N_1^*\right]$ yields $N_2^* = \frac{m_1(q_{12}p_1 - p_2 q_1) - p_2 m_2 q_1}{q_2 q_1}$.

Obviously, the sign of N_2^* for the latter case is not clearly determined. For a given value of N_1^*, N_2^* can be positive or negative.
The last equilibrium is defined by

$$N_1^* = m_1 \text{ and } N_2^* = -\frac{1}{q_2}\left[m_1(p_2 + q_{12}) + m_2 p_2\right]. \blacksquare \tag{22}$$

Next, system (12) is linearised around the steady state values to establish the stability of the obtained equilibria. After linearising the entire system, the Jacobian matrix for each equilibrium of system (12) has been evaluated. Table 1 provides a summary of the obtained equilibria and further reports the equilibrium specific Eigenvalues with their corresponding signs.
It is obvious that the first equilibrium is a stable nod. The stability of the remaining equilibria is not of importance, because from an economic point of view only the first equilibrium ensures a plausible result, which means that both N_1^* and N_2^* are positive, given the parameter definition above. These result can be fleshed out also graphically in figure 1 for positive values of $N_1^* > 0$ and $N_2^* > 0$.
From figure 1 we can once again conclude that only for the first equilibrium an economic interpretation is possible. In the long run, the market saturation level will be reached for both groups of adopters. Moreover, this equilibrium is stable. The third equilibrium is ex ante not clearly determined, because for a given parameter constellation, positive as well as negative values for N_2 are possible. Theoretically, the $\dot{N}_2 = 0$ straight line, the slope and location of which is determined by E_3 and E_4, can result in another feasible solution. But if we take a closer look at our model, we can rule out this possibility. If we refer again to figure 1, we observe that both, $-\frac{p_2 m}{q_2}$ and $-\frac{p_2 m_2}{q_{12}}$ determine the location of $\dot{N}_2 = 0$ straight line, $N_2 = -\frac{q_{12}}{q_2}N_1 - \frac{p_2 m}{q_2}$. The maximum limit expression of $\dot{N}_2 = 0$ straight line is given by: $N_2 = \frac{q_{12}}{q_2}N_1$, which is graphically replicated by a dashed straight line through the origin, as can be seen in figure 1. This is the maximum limit because, $-\frac{p_2 m}{q_2}$ cannot be positive by definition, as all parameters in expression $\frac{p_2 m}{q_2}$ are positive. This is also true for $-\frac{p_2 m_2}{q_{12}}$. Note also that the $\dot{N}_2 = 0$ straight line will not be translated parallel, because the upper limit for N_2 is given by $\tilde{N}_2 = \frac{q_{12} m_1 p_1}{q_2 q_1}$ and hence the difference between the upper limit of $\tilde{N}_2$ and N_2 is given by $\Delta N_2 = \frac{q_1 p_2 m_1}{q_2 q_1} + \frac{p_2 m_2}{q_2} = \frac{q_1 p_2 m_1}{q_2 q_1} +$

$\Delta^+ N_2$ which is by expression on modulus greater as $\Delta^+ N_2 = \frac{p_2 m_2}{q_2} > 0$ if we refer to equilibrium four.

From this discussion it can be concluded that E_3 cannot be a possible candidate for a relevant economic equilibrium. Again, from an economic point of view we only focus on the first equilibrium which is given by: $N_1^* = m_1$ and $N_2^* = m_2$. Thus from any given and feasible starting point within the rectangular area bounded by the parallel $\dot{N}_2 = 0$ line to the hypotenuse and the parallel $\dot{N}_1 = 0$ line to the ordinate we can always realize the equilibrium point E_1 for given starting values, $N(0)_1 \geq 0$ and $N(0)_2 \geq 0$. Referring again to figure 1, a steady state path for given but arbitrary starting values, $N(0)_1 > 0$ and $N(0)_2 > 0$, has been drawn.

Model (11) derived so far has to be criticised as it assumes a short and long run deterministic behaviour of the adoption process, which means that once being on the S-shaped diffusion path, no deviations from this path are possible, even in the short run. The implication is that uncertainty regarding the adoption process should not be treated as constant over time or even totally neglected, as the Bass (1969) model does. Particularly, at the beginning or in the middle of the diffusion process, say around the inflection point, uncertainty should be much more higher than at the end, which implies that fluctuations of the adoption curve should be largest around the inflection point. From this point of view, a stochastic expansion of system (11) is required which will be derived in the next section.

4 Stochastic knowledge diffusion model

4.1 Setup

In this section a stochastic expansion of system (11) will be derived. I follow Boswijk and Franses (2005) and Boswijk et al. (2006) who derived a stochastic counterpart of the Bass (1969) model by assuming short-run deviations from the deterministic diffusion process in a one-dimensional model. To arrive at our stochastic counterpart of system (11) it has to pointed out first that the cumulative numbers of innovators and imitators are both random variables with

$$\bar{N}(t)_k = E[N(t)_k] = mF(t),\ k = \{1,2\}, \tag{23}$$

where $k = 1$ stands for the innovators, and $k = 2$ for imitators, and t is measured still in continuous time.

Defining $\frac{d\tilde{N}(t)_k}{dt} = \bar{n}(t)$ for $k = \{1,2\}$ we can theoretically derive two different systems: the first system assumes that mean reverting takes place from the mean number $\bar{n}(t)$ or from the actual number of adoptions $\tilde{n}(t)$. The difference is that the mean number of adoptions $\bar{n}(t)$, is treated as an exogenous variable, whereas $\tilde{n}(t)$ is endogenous. For this reason, we should prefer to work with $\tilde{n}(t)$.

Keeping this in mind, the following stochastic expansion of system (11) is defined:

$$\begin{bmatrix} dn(t)_1 \\ dn(t)_2 \end{bmatrix} = \begin{bmatrix} \zeta[\tilde{n}(t)_1 - n(t)_1]dt + \sigma n(t)_1^\gamma & 0 \\ 0 & \zeta[\tilde{n}(t)_2 - n(t)_2]dt + \sigma n(t)_2^\gamma \end{bmatrix} \begin{bmatrix} dW(t)_1 \\ dW(t)_2 \end{bmatrix}, \tag{24}$$

where $W(t)_k$ is the standard Wiener process, and $\zeta > 0$ is the adjustment speed. Please note that $W(t)_1$ and $W(t)_2$ are eventually correlated. Furthermore, it is assumed that $\sigma > 0$ and $\gamma \geq 0.5$. Therefore, the speed of mean reversion depends on the value of ζ. System (24) is a generalised stochastic version of system (11), because it contains an error term in continuous time with a standard deviation which equals to $\sigma n(t)_k^\gamma$. As $n(t)_k \to 0$, the error term $\sigma n(t)_k^\gamma \to 0$, and thus it is guaranteed that $n(t)$ takes non negative values. It should be clear that system (11) is obtained if $\zeta \to \infty$ and $\sigma \to 0$. For $\gamma = 1$ it can be shown that $n(t)_k$ is strictly positive. As a degree of freedom $\gamma = 1$ has been set. In this work, the examination of system (24) dynamic behaviour is done on the basis of simulation experiments. Alternatively, one can show formally the existence and solution of system (24). One aspect which can be easily seen from system (24) is that, given the starting value $N(0)_k = 0$, $N(t)_k$ increases monotonically to $N(t)_k = m$ for $t \to$ large T. Please additionally note that the speed of adjustment, ζ, is assumed to be the same for both the innovators and imitators. That is also the case for σ.

Embedding system (11) in system(24) yields the following system of stochastic differential equations (SDE):

$$\begin{bmatrix} dn(t)_1 \\ dn(t)_2 \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} dW(t)_1 \\ dW(t)_2 \end{bmatrix} , \tag{25}$$

with

$$A \equiv \zeta \left\{ \Theta_1 - n(t)_1 \right\} dt + \sigma n(t)_1, \tag{26}$$

and

$$B \equiv \zeta \left\{ \Theta_2 - n(t)_2 \right\} dt + \sigma n(t)_2, \tag{27}$$

and

$$\Theta_1 \equiv \left[p_1 + \left(q_1 \frac{N(t)_1}{m_1} \right) \right], \tag{28}$$

and

$$\Theta_2 \equiv p_2 + \left[\left(q_2 \frac{N(t)_2}{(m_1 + m_2)} \right) + \left(q_{12} \frac{N(t)_1}{(m_1 + m_2)} \right) \right]. \tag{29}$$

For simulation purpose, system (24), which is a continuous time model has to be translated into a discrete time model with discrete observations $N_{i,k} = N(t_i)_k$ for $i = 1, 2, ..., T$ and $k = \{1, 2\}$. Thus, adoption of new knowledge over the interval $(t_{i-1,k}, t_{i,k}]$ is given by $\Psi_{(i,k)} \equiv N_{i,k} - N_{i-1,k}$. Therefore, $\Psi_{(i,k)}$ can be interpreted as the discrete approximation of $n(t)_k$.

4.2 Euler-Maruyama approximation

The discretisation of model (25) is based on the so called *Euler-Maruyama* approximation[14]. On a given interval $[t_0, T]$ and for a given discretisation $t_0 < t_1 < ... < t_i <$

[14]Refer to for Kloeden and Platen (1992), p. 305.

... $< t_N = T$ of $[t_0, T]$, an *Euler-Maruyama* approximation of an one-dimensional *Ito* SDE $dX_t = f(X_t, \theta) + g(X_t, \theta)dW_t$ is a so called time stochastic process which satisfies the proposed iterative scheme

$$y_{i+1} = y_i + h_i f(y_i) + g(y_n)\Delta W_i, \tag{30}$$

with $y_0 = x_0$ for $i = 0, 1, ..., N-1$, where $y_i = y(t_i)$, $\xi_i = [t_i - t_{i-1}]$ is the step size and $\Delta W_i = W(t_i) - W(t_{i-1}) \sim \mathcal{N}(0, \xi_i)$ with $W(t_0) = 0$.

The last follows because due to the definition of a *Wiener* process we conclude that these increments are independent *Gaussian* random variables with mean 0 and variance h_i. The increments ΔW_n can be computed as $\Delta W = \int_{t_i}^{t_{i+1}} dW_t = W(t_{i+1}) - W(t_i)$. It is straightforward that the proposed *Euler-Maruyama* approximation still holds for systems like (25).

It is known that the *Euler-Maruyama* method converges with strong order $(\gamma = 1)$ for additive noise, and for a constant diffusion term g the *Euler-Maruyama* method should provide a reasonable approximation.[15] For other cases, however, the method provides eventually a poor estimate of the solution, especially if the coefficients of interest have to be treated as non-linear, a fact, which is known from the deterministic *Euler*-approximation. In this case, higher order schemes, like the *Milstein* scheme should be consulted to obtain a satisfying approximation in terms of higher accuracy. It has to be pointed out that the order of the *Euler-Maruyama* scheme is only satisfactory regarding approximation results if a fine time span $\xi_i = \frac{H}{T}$ is used.[16]

Applying the *Euler-Maruyama* approximation for system (25) and using $n(t_i)_k - n(t_{i-1})_k$, the following expression is obtained:

$$\begin{bmatrix} n(t_i)_1 - n(t_{i-1})_1 \\ n(t_i)_2 - n(t_{i-1})_2 \end{bmatrix} \approx \begin{bmatrix} \zeta \left\{ \left[p_1 + \left(q_1 \frac{N(t_{i-1})_1}{m_1} \right) \right] - n(t_{i-1})_1 \right\} \xi + \vartheta_1 \\ \zeta \left\{ p_2 + \left[\left(q_2 \frac{N(t_{i-1})_2}{(m_1+m_2)} \right) + \left(q_{12} \frac{N(t_{i-1})_1}{(m_1+m_2)} \right) \right] - n(t_{i-1})_2 \right\} \xi + \vartheta_2 \end{bmatrix}, \tag{31}$$

[15] On general, the *Euler-Maruyama* method has strong order of convergence $\gamma = 0.5$ and for weak order of convergence $\gamma = 1$.

[16] Refer to Kloeden and Platen (1992), p. 345.

with $\xi = [t_i - t_{i-1}]$ and $\vartheta_k = \sigma[W(t_i)_k - W(t_{i-1})_k] \sim$ i.i.d. $\mathcal{N}(0, \sigma^2\xi)$.

The approximation of adopting new knowledge $\Psi_{i,k}$ over the time interval $(t_{i-1}, t_i]$ can be written as

$$\Psi_{i,k} = N(t_i)_k - N(t_{i-1})_k = \int_{t_{i-1}}^{t_i} n(t)_k dt \approx n(t_i)_k(t_i - t_{i-1}) = n(t_i)_k\xi. \tag{32}$$

Thus the alteration rate of adopting new knowledge is given by $\Delta\Psi_{i,k} \equiv \Psi_{i,k} - \Psi_{i-1,k}$, or

$$\Delta\Psi_{i,k} \equiv \Psi_{i,k} - \Psi_{i-1,k} \approx \xi[n(t_i)_k - n(t_{i-1})_k]. \tag{33}$$

Using model (33) together with model (24) and model (12) or model (25) we derive at

$$\begin{bmatrix} \Delta\Psi_{i,1} \\ \Delta\Psi_{i,2} \end{bmatrix} \approx \begin{bmatrix} \xi\zeta\left\{\left[p_1 + \left(q_1 \frac{N(t_{i-1})_1}{m_1}\right)\right] - \frac{\Psi_{i-1,1}}{\xi}\right\}\xi + \xi\frac{\Psi_{i-1,1}}{\xi}\vartheta_{i,1} \\ \xi\zeta\left\{\left[p_2 + \left(q_2 \frac{N(t_{i-1})_2}{(m_1+m_2)}\right) + \left(q_{12}\frac{N(t_{i-1})_1}{(m_1+m_2)}\right)\right] - \frac{\Psi_{i-1,2}}{\xi}\right\}\xi + \xi\frac{\Psi_{i-1,2}}{\xi}\vartheta_{i,2} \end{bmatrix}, \tag{34}$$

or

$$\begin{bmatrix} \Delta\Psi_{i,1} \\ \Delta\Psi_{i,2} \end{bmatrix} \approx$$

$$\approx \begin{bmatrix} \zeta\xi^2 p_1(m_1 - N_{(i-1),1}) + \xi^2\frac{q_1}{m_1}N_{(i-1),1}(m_1 - N_{(i-1),1}) - \zeta\xi\Psi_{i-1,1} + \Psi_{i-1,1}\vartheta_{i,1} \\ \zeta\xi^2 p_2(m_2 - N_{(i-1),2}) + \xi^2\frac{q_2}{(m_1+m_2)}N_{(i-1),2}(m_2 - N_{(i-1),2}) + \varsigma - \zeta\xi\Psi_{i-1,2} + \Psi_{i-1,2}\vartheta_{i,2} \end{bmatrix}, \tag{35}$$

with

$$\varsigma \equiv \xi^2\zeta\frac{q_{12}}{(m_1+m_2)}N_{(i-1),2}(m_2 - N_{(i-1),2}), \tag{36}$$

or

$$\begin{bmatrix} \Delta\Psi_{i,1} \\ \Delta\Psi_{i,2} \end{bmatrix} \approx$$

$$\approx \begin{bmatrix} \phi_{0,1} + \phi_{1,1} N_{(i-1),1} + \phi_{2,1} N^2_{(i-1),1} + \phi_{3,1} \Psi_{i-1,1} + \Psi_{i,1} \vartheta_{i,1} \\ \phi_{0,2} + \phi_{1,2} N_{(i-1),2} + \phi_{2,2} N^2_{(i-1),2} + \phi_{3,2} N_{(i-1),1} + \phi_{4,2} N_{(i-1),1} N_{(i-1),2} + \phi_{5,2} \Psi_{i-1,2} + \Psi_{i,2} \vartheta_{i,2} \end{bmatrix}, \tag{37}$$

with $\vartheta \sim$ i.i.d. $\mathcal{N}(0, \sigma^2 \xi)$ and

$$\phi_{0,1} = p_1 m_1 \zeta \xi^2, \tag{38}$$

$$\phi_{1,1} = \zeta \xi^2 (q_1 - p_1), \tag{39}$$

$$\phi_{2,1} = \frac{-q_1 \zeta \xi^2}{m_1}, \tag{40}$$

$$\phi_{3,1} = -\zeta \xi, \tag{41}$$

$$\phi_{0,2} = p_2 m_2 \zeta \xi^2, \tag{42}$$

$$\phi_{1,2} = \zeta \xi^2 \left[\left(\frac{m_2}{m_1 + m_2} q_1 \right) - p_2 \right], \tag{43}$$

$$\phi_{2,2} = \frac{-q_2 \zeta \xi^2}{m_1 + m_2}, \tag{44}$$

$$\phi_{3,2} = \frac{m_2}{m_1 + m_2} \zeta \xi^2 q_{12}, \tag{45}$$

$$\phi_{4,2} = \frac{-q_{12} \zeta \xi^2}{m_1 + m_2}, \tag{46}$$

$$\phi_{5,1} = -\zeta \xi. \tag{47}$$

In this notational form we can interpret $\phi_{3,2}$ as the knowledge transfer parameter function, which depends among other values on q_{12}. If $q_{12} \to 0$, then $\phi_{3,2} \to 0$, and thus, no knowledge transfer from innovators to imitators takes place.

5 Simulation

5.1 Simulation of deterministic knowledge diffusion model

In this section the adoption curves of model (11) are simulated. For simulation purposes we first have to assign a set of parameters. The values of the external knowledge transfer coefficients p_1 and p_2 are set to $p_1 = 0.13$ and $p_2 = 0.01$, which means that the innovators are more influenced by external knowledge transfer than the imitators. The values for the internal knowledge transfer coefficients, q_1 and q_2, are determined at $q_1 = 0.75$ and $q_2 = 0.50$, which means that internal knowledge transfer matters more for the group of innovators.[17] For the meanwhile, the knowledge transfer coefficient q_{12} is set to $q_{12} = 0.07$. Later on in this paper a sensitivity analysis regarding parameter variations of q_{12} for the stochastic knowledge diffusion model is performed to determine the effect on the overall adoption curve for different parameter constellations of q_{12}. Table 3 summarises the calibrated values for the simulation study.

The simulation of model (11) has been conducted with `Matlab 6.5.0` with cross checks performed with `Mathematica 5.2.` Simulation results have been graphically represented in figure 3. In the left upper figure, the adoption curves for $N(t)_1$, $N(t)_2$ and the overall market $N(t)_{all}$ have been drawn with solid, dash-dotted and dashed lines respectively. As a result we can observe that the knowledge diffusion process of the innovators comes to an end after around 6 periods, because the entire population of innovators has adopted new knowledge, which means that $m_1 = N(6)_1 = 1$, and thus, $\dot{N}(t)_1 = 0$. On the contrary, the knowledge diffusion process of the imitator group stops after around 20 periods of time with $m_2 = N(20)_2 = 1$, and thus, $\dot{N}(t)_2 = 0$. Using the results from our stability analysis, we have realized a stable equilibrium at $m_1^* = m_2^* = 1$. Figure 2 gives a graphical representation of the equilibrium path for the simulated model based on parameter values in table 3 and with arbitrary starting values $N(0)_1 = N(0)_2 = 0$. Furthermore, figure 3 shows that the unique equilibrium $m_1^* = m_2^* = 1$ is stable.

The left lower figure contains the same information as the left upper figure, but in relative numbers related to the market potential m_1 and m_2, respectively. The inflection points of the innovators and imitators are realized at around 2 periods for the innovators and at around 9 periods for the imitators. The upper right figure depicts the diffusion process of the new knowledge, whereas the lower right panel shows the relative diffusion process. It is easy to see that for $m_1 = 1$ and $m_2 = 1$ the upper right and the lower right figures must coincide.

[17] As unreported simulation experiments have shown, the exact choice of p_k and q_k for $k = \{1, 2\}$ do not influence the general result of unimodal and bimodal diffusion patterns.

What impression can we get from figure 3 regarding the overall diffusion process $n(t)_{all}$? First, the knowledge diffusion process does not exhibit a bell shaped pattern, as in the original Bass (1969) model, but is unimodal with dent towards right. This is because the innovators still have adopted the entire knowledge and have realized the inflection point, whereas the imitators just start to adopt. Please note that we do not observe the typically bimodal chasm pattern because cross sectional external knowledge transfer ($q_{12} > 0$) takes place. As shown later, the typical chasm pattern of the knowledge diffusion process is only realized if $q_{12} \rightarrow 0$[18].

As mentioned before, one of the drawbacks of this model is that the adoption curves $N(t)_1$ and $N(t)_2$ still both exhibit the typical deterministic S-shaped pattern, as one can see from the upper left and lower left pictures of figure 3. This assumption seems to be to strict. Thus, this strict pattern structure has been relaxed by assuming that the diffusion process is a mean reverting event and hence, short term deviation from a deterministic sigmoid adoption path should be allowed. The simulation of this stochastic model is performed in the next subsection.

5.2 Simulation of stochastic knowledge diffusion model

In this section a simulation study of model (25) has been conducted. Inserting the calibrated values from table 3 in equations (38) to (47) leads to

[18]It is referred to the sensitivity analysis in section 5.2.

$$\phi_{0,1} = p_1 m_1 \zeta \xi^2 = 1.625 \times 10^{-3}, \tag{48}$$
$$\phi_{1,1} = \zeta \xi^2 (q_1 - p_1) = 7.75 \times 10^{-3}, \tag{49}$$
$$\phi_{2,1} = \frac{-q_1 \zeta \xi^2}{m_1} = -9.375 \times 10^{-3}, \tag{50}$$
$$\phi_{3,1} = -\zeta \xi = -0.25, \tag{51}$$
$$\phi_{0,2} = p_2 m_2 \zeta \xi^2 = 1.625 \times 10^{-3}, \tag{52}$$
$$\phi_{1,2} = \zeta \xi^2 \left[\left(\frac{m_2}{m_1 + m_2} q_1 \right) - p_2 \right] = 3.000 \times 10^{-3}, \tag{53}$$
$$\phi_{2,2} = \frac{-q_2 \zeta \xi^2}{m_1 + m_2} = -3.125 \times 10^{-3}, \tag{54}$$
$$\phi_{3,2} = \frac{m_2}{m_1 + m_2} \zeta \xi^2 q_{12} = 4.375 \times 10^{-4}, \tag{55}$$
$$\phi_{4,2} = \frac{-q_{12} \zeta \xi^2}{m_1 + m_2} = -4.375 \times 10^{-4}, \tag{56}$$
$$\phi_{5,1} = -\zeta \xi = -0.25. \tag{57}$$

As mentioned above, the simulation of the system of stochastic differential equations (25) has been performed with `Matlab 6.5.0`. The corresponding simulation results have been depicted in figures 4, 5 and 6. If we refer to figure 4, on the upper picture we can find the distribution functions $F(t)_1$ and $F(t)_2$ for the adoption process for innovators and imitators. Additionally, the discrete approximation of the distribution functions expressed by $\frac{N(t)_1}{m_1}$ for the innovators and $\frac{N(t)_2}{m_2}$ for the imitators have been plotted. The overall distribution function $F(t)_{all}$ exhibits a dint pattern which commemorates slightly on a S-Shaped pattern. This is also the case for the discrete approximation $\frac{N(t)_{all}}{m_{all}}$ in the same subpicture 5. Furthermore, the density functions $f(t)_1$ and $f(t)_2$ for the adoption process for innovators and imitators and the corresponding approximations $\frac{n(t)_1}{m_1}$ and $\frac{n(t)_2}{m_2}$ have been plotted in the middle placed picture of figure 4. In the lower figure we find the approximation of the density function $f(t)_k$ denoted as $\frac{\Psi_{(i,k)}}{m_k}$ for $k = \{1, 2\}$. Also, for the entire population $N(t)_{all}$ we observe a S-shaped mean reverting behaviour with the largest deviation from the mean around the inflection point, as we should expect. Additionally, overall adoptions $\Psi_{(i,all)}$ exhibit a mean reverting behaviour with the largest fluctuations around the peak for both innovators and imitators.

The interesting point is how changes of the knowledge transfer parameter q_{12} affects system (25) and how variations of knowledge transfer affect the cumulative and adoption curves of the model (25) for both groups. For this purpose, a sensitivity analysis for three scenarios has been performed: in the first scenario it is assumed that knowledge transfer from the group of innovators to the group of imitators is nearly prohibited, which coincides with $q_{12} \rightarrow 0$[19]. Please note again, that the knowledge transfer process is asymmetric, which means that knowledge transfer goes from the group of innovators to the group of imitators and not vice versa. The second scenario is characterised by a limited knowledge transfer, with $q_{12} = 0.07$, which corresponds to the already performed simulation. The last scenario assumes nearly complete knowledge transfer, which implicitly means that strong network effects are in place. For this simulation scenario, $q_{12} \rightarrow 1$[20].

The simulation results for the first and last simulation scenarios are depicted in figures 5 and 6.

If we refer to figures 4, 5 and 6, we come to the following result: the less important the network effects are - which coincides with parameter value of $q_{12} \rightarrow 0$ - the more realistic is the so called chasm pattern. In other words, the greater the discrepancy between the realization of the inflection point of innovators and the beginning of imitators' adoption, the more realistic is a bimodal shape of the adoption curve. On the other side, the stronger the network effects, the greater is the parameter value of q_{12} and the less realistic is the so called chasm pattern, because right before innovators have realized the inflection point imitators have nearly reached their inflection point. In this way we can conclude that a bimodal pattern of overall knowledge diffusion is more likely, if it is hard to establish knowledge networks, whereas unimodal but not necessarily a bell-shaped pattern in the sense of Bass (1969) of diffusion is more likely, if strong network effects are in place. This conclusion can be verified by a sensitivity analysis. This has been conducted for the stochastic model for the entire parameter range of $q_{12} \in (0, 1)$. The parameter range $q_{12} \in (0, 1)$, as well as the diffusion time t, defines a grid in $\mathbb{R}^2$. For a given point in t and a given param-

[19]For the simulation study q_{12} is set to $q_{12} = 0.01$.

[20]For the simulation study q_{12} is set to $q_{12} = 0.99$.

eter point $q_{12} \in (0,1)$, a point on the approximated density function of adopters is located on the grid. Again, for large q_{12} unimodal diffusion patterns are more likely than a bimodal diffusion pattern. The results of this study can be found in figure 7.

6 Econometric Annotations

The question which is unanswered yet is how knowledge diffusion can be measured empirically, especially the parameter q_{12}. First, one has to find suitable proxies for knowledge diffusion. One possibility is to assume that new knowledge is stored in scientific journals, and citations of specific articles could be a proxy for diffusion of this new knowledge. Citations typically often have similarities with the diffusion of new products. At the beginning, citations are low, then they start growing and reach a peak before the citations tend to zero.

As one can see, system (25) can be estimated and tested directly without manipulating the system itself. Obviously, a seemingly unrelated regression (SUR) seems to be appropriate for estimating system (25), because system (25) is block recursive. Note that this model assumes heteroscedastic errors because of the term $\sigma n(t)_k^\gamma \neq 0$. This again reflects the idea that diffusion is more certain at the end of the diffusion process.

Before performing the SUR regression, the question which should be answered is, whether the estimated coefficients are consistent or not. Boswijk and Franses (2005) have shown that the estimators do not exhibit the desired asymptotic normality behaviour by estimating their one dimensional stochastic version of the Bass (1969) model. More precisely, the authors have shown that even by increasing the sample period the estimators $\phi \in \Phi$ cannot be estimated consistently. This result seems to be reasonable, because after realizing the saturation levels m_1 or m_2 respectively, within sample information no longer increases which is necessary to obtain consistent estimators of the parameter vector. On the basis of Monte Carlo simulations for different time spans $H = \xi T$, Boswijk and Franses (2005) have further concluded that standard normal distribution can be consulted to approximate t-statistics of the estimated parameter vector, provided the inflection point lies within the sample period. Although it is helpful to know large sample properties of the estimators $\phi \in \Phi$, small sample properties are of importance because of the fact that sample information is limited towards t. This topic is not sufficiently addressed in the relevant literature and defines an avenue for further research.

7 Conclusion

In this paper the link between knowledge transfer, knowledge diffusion and implicit network effects has been investigated. For this reason, a new diffusion model was put forward which focuses on those before mentioned aspects. The relevant literature has paid less attention investigating the link between knowledge transfer and knowledge diffusion. Particularly, the question of how knowledge transfer has an influence on the behavoiur of innovators and imitators within the adoption process is of interest. The basis for this stochastic differential equation (SDE) model is the well known Bass (1969) model. Although Bass (1969) mentioned that communication between innovators and imitators is relevant for adoption decision, this fact is not reflected in his mathematical derivations. Following Kalish (1985) and assuming that innovators need only search information to adopt new knowledge, while the latter imitators require experience type information before adopting, a model which includes both the adoption decisions of innovators and imitators is set up. In this way, the group of adopters has to be treated as heterogeneous. Furthermore it was assumed that information flows only in one direction, from innovators to imitators. Thus, the information flow is asymmetric.

After an appropriate discretisation, in a simulation study it was shown that the shape of the adoption pattern depends on, whether knowledge diffusion occurs or not. If knowledge transfer occurs, the stronger the network effects, the more probable are the so-called unimodal patterns, because right before innovators have realized their inflection point, imitators have nearly reached themselves their inflection point. On the contrary, the greater the discrepancy between the realization of the inflection point of innovators and the beginning of imitators' adoption, the less important are network effects, and the more probable are the so called bimodal adoption phenomena. Thus chasm patterns of adoption curves occur if it is hard to establish knowledge networks.

The advantage of this new model is twofold: from a theoretical point of view, not only unimodal diffusion phenomena can be modelled, but also bimodal diffusion phenomena can occur. From an empirical point of view, the model which incorporates heteroscedastic errors and mean reverting can be estimated and tested directly with a SUR approach.

So far this study suggests some avenues for further research. First, the assumption that the market saturation level is exogenous and constant over time is very strict. Second, from a technical point of view, mean reverting is assumed to be the same over the entire population. Thus another source of heterogeneity can be introduced in the model by assuming intra-group individual values for the adjustment speed to the steady vector ζ. Third, after examining the large and small sample properties of the derived model the forecasting ability should be of interest.

8 Appendix

Given $N(0)_2 = 0$, the solution for $N(t)_2$ can be expressed as:

$$
\begin{aligned}
N(t)_2 = \frac{1}{\Theta}\Big\{&[m_1(p_2+q_{12})+m_2(p_2+q_2)]\times \\
&\Big[-(exp(-(p_1+q_1)t)^{\frac{m_1(p_2+q_{12})+m_2(p_2+q_2)}{(m_1+m_2)(p_1+q_1)}}(m_1+m_2)(p_1+q_1)+ \\
&+m_2(p_1+exp(-(p_1+q_1)t)^{\frac{m_1q_{12}}{(m_1+m_2)q_1}}q_2C\Big]+ \\
&+(exp(-(p_1+q_1)t)^{\frac{m_1(p_2+q_{12})+m_2(p_2+q_2)}{(m_1+m_2)(p_1+q_1)}}m_2(m_1+m_2)(p_1+q_1)\left(1+\frac{exp(-(p_1+q_1)t)q_1}{p_1}\right)^{\frac{m_1q_{12}}{(m_1+m_2)q_1}}\times \\
&q_2\tilde{H}\left[\frac{m_1(p_2+q_{12})+m_2(p_2+q_2)}{(m_1+m_2)(p_1+q_1)},\frac{m_1q_{12}}{(m_1+m_2)q_1},\Phi,-\frac{exp(-(p_1+q_1)t)q_1}{p_1}]\right]\Big\},
\end{aligned}
\tag{58}
$$

with Θ defined as:

$$
\Theta \equiv q_2\Big\{(p_1+exp(-(p_1+q_1)t)q_1)^{\frac{m_1q_{12}}{(m_1+m_2)q_1}}(m_1(p_2+q_{12})+m_2(p_2+q_2))C+
$$

$$+(exp(-(p_1+q_1)t)^{\frac{m_1(p_2+q_{12})+m_2(p_2+q_2)}{(m_1+m_2)(p_1+q_1)}}(m_1+m_2)(p_1+q_1)\left(1+\frac{exp(-(p_1+q_1)t)q_1}{p_1}\right)^{\frac{m_1 q_{12}}{(m_1+m_2)q_1}}\times$$

$$\times\tilde{H}\left[\frac{m_1(p_2+q_{12})+m_2(p_2+q_2)}{(m_1+m_2)(p_1+q_1)},\frac{m_1 q_{12}}{(m_1+m_2)q_1},\Phi,-\frac{exp(-(p_1+q_1)t)q_1}{p_1}\right]\Bigg\}, \tag{59}$$

with

$\Phi \equiv 1+\frac{m_1(p_2+q_{12})+m_2(p_2+q_2)}{(m_1+m_2)(p_1+q_1)}$

and

$$C \equiv \frac{1}{m_2 q_2[m_1(p_2+q_{12})+m_2(p_2+q_2)]}\Bigg[(m_1+m_2)(p_1+q_1)^{1-\frac{m_1 q_{12}}{(m_1+m_2)q_1}}\times$$

$$\times(m_1(p_2+q_{12})+m_2(p_2+q_2)-m_2\left(\frac{p_1+q_1}{p_1}\right)q_2\times$$

$$\times\tilde{H}\left[\frac{m_1(p_2+q_{12})+m_2(p_2+q_2)}{(m_1+m_2)(p_1+q_1)},\frac{m_1 q_{12}}{(m_1+m_2)q_1},\Phi,-\frac{exp(-(p_1+q_1)t)q_1}{p_1}\right]\Bigg]. \tag{60}$$

Note, that $\tilde{H}(\cdot)$ is a hypergeometric function, which series expansion is given by

$$\tilde{H} \equiv {}_2F_1(a,b,c,x) = \sum_{w=0}^{\infty} \frac{(a)_w (b)_w}{(c)_w} \frac{x^w}{w!} =$$

$$= 1 + \frac{abx}{c1!} + \frac{a(a+1)b(b+1)x^2}{c(c+1)2!} + \frac{a(a+1)(a+2)b(b+1)(b+2)x^3}{c(c+1)(c+1)3!} + ..., \tag{61}$$

where $(i)_w$ is the Pochhammer symbol defined as $(i)_0 = 1$, and $(i)_w = i(i+1)...(i+w-1) = \frac{\Gamma(i+w)}{\Gamma(i)}$ for $i = a, b, c$ where $\Gamma(\cdot)$ is called the *Euler-Gamma* function.[21] Furthermore, note that $\tilde{H}$ has a branch cut discontinuity in the complex z plane from 1 to ∞ and terminates if a and b are non positive integers. Of course, $n(t)_k = \frac{dN(t)_k}{dt}$ for $k = \{1, 2\}$.

[21] Abramowitz and Stegun (1972) p. 255.

References

Abramowitz, M. and Stegun, C. A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* New York.

Audretsch, D. and Feldman, M. P. (1996). *R&D Spillovers and the Geography of Innovation and Production.* In: The American Economic Review, 86, pp. 630-640.

Bass, F. M. (1969). *A new product growth for model consumer durables.* In: Management Science, 15, pp. 215-227.

Bass, F. M. (1993). *The Future of Research in Marketing: Marketing Science.* In: Journal of Marketing Research, 30, pp. 1-6.

Bass, F. M. (1995). *Empirical Generalizations and Marketing Science: A Personal View.* In: Marketing Science, 14, G6-G19.

Boswijk, H. P. and Franses, P. H. (2005). *On the Econometrics of the Bass Diffusion Model.* In: Journal of Business & Economic Statistics, 23, pp. 255-268.

Boswijk, P., Fok, D., and Franses, P. H. (2006). *A New Multivariate Product Growth Model.* In: Tinbergen Institute Discussion Paper No. 06-027/4.

Easingwood, C. J., Mahajan, V., and Muller, E. (1983). *A non-uniform influence innovation diffusion model of new product acceptance.* In: Marketing Science, 2, pp. 273-295.

Fok, D. and Franses, P. H. (2007). *Modeling the diffusion os scientific publications.* In: Journal of Econometrics, 139, pp. 376-390.

Gatignon, H. and Robertson, T. S. (1985). *A propostitional inventory for new diffusion research.* In: Journal of Consumer Research, 11, pp. 849-867.

Gladwell, M. (2000). *The Tipping Point. How Little things Can Make a Big Difference.* Little Brown and Company, London.

Golendberg, J., Libai, B., Muller, E., and Peres, R. (2006). *Blazing saddles: the early and mainstream markets in the high-tech product life cycle.* In: Israel Economic Review, 4, pp. 85-108.

Hill, S., Provost, F., and Volinsky, C. (2006). *Network-Based Marketing: Identifying Likely Adopters via Consumer Networks.* In: Statistical Science, 21, pp. 256-276.

Jeuland, A. (1981). *Parsimonious models of diffusion of innovation: Part a, derivations and comparison.* Working paper, Graduate School of Business, University of Chicago.

Kalish, S. (1985). *A new product adoption model with price, advertising and uncertainty.* In: Management Science, 31, pp. 1569-1585.

Kalish, S. and Sen, S. K. (1986). *Diffusion models and the marketing miex for single products.* In: Mahajan, V., Wind, Y. (eds.). Innovation Diffusion Models of New Product Acceptance, V, pp. 87-116. Cambridge.

Kloeden, P. E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations.* Applications of Mathematics 23, Berlin.

Kluyver, C. A. (1977). *Innovation and industrial product life cycles.* In: Californian Management Review, 20, pp. 21-33.

Mahajan, V., E., M., and Kerin, R. A. (1984). *Introduction Strategy for New Products with Positive and Negative Word-of-Mouth.* In: Management Science 30, pp. 1389-1404.

Mahajan, V., Muller, E., and Bass, F. M. (1990). *New Product Diffusion Models in Marketing: A Review and Directions for Research.* In: Journal of Marketing, 54, pp. 1-26.

Mahajan, V., Muller, E., and Bass, F. M. (1993). *New Product Diffusion Models.* In: Eliashberg, J., Lilien, G.L. (eds.). Handbooks in Operations Research & Management Science, V, pp. 349-408.

Mahajan, V., Muller, E., and Wind, Y. (2000). *New-Product Diffusion Models.* Norwell.

Mahajan, V. and Peterson, R. A. (1985). *Models for Innovation Diffusion.* Beverly Hills.

Mahajan, V. and Wind, Y. (1986). *Innovation Diffusion Models of New Product Acceptance.* Cambridge.

Moore, G. (1995). *Inside the Tornado. HarperBusiness.* New-York.

Moore, G. (2002). *Crossing the Chasm.* New-York.

Parker, P. M. (1994). *Aggregate Diffusion Models in Marketing: A Critical Review.* In: International Journal of Forecasting, 10, pp. 353-380.

Polany, M. (1967). *The Tacit Dimension.* New-York.

Rink, D. and Swan, J. (1979). *Product life cycle research: a literature review.* In: Journal of Business Research, 7, pp. 219-42.

Rivera-Batiz, L. A. and Romer, P. M. (1991). *International trade with endogenous technological change.* In: Economic Review, 35, pp. 971-1004.

Rogers, E. M. (1983). *Diffusion of Innovations.* 3rd edt. New-York.

Rosen, E. (2000). *The Anatomy of Buzz.* New York.

Schmalen, H. (1982). *Optimal prica and advertising policy for new products.* In: Journal of Business Research, 10, pp. 17-30.

Slywotzky, A. J. and Shaprio, B. P. (1993). *Leveraging to beat the odds: The new marketing mind-set.* In: Harvard Business Review, 71, pp. 97-107.

Sultan, F., Farley, J. U., and Lehmann, D. R. (1990). *A meta-analysis of applications of diffusion models.* In: Journal of Marketing Research, 27, pp. 70-77.

Tanny, S. M. and Derzko, N. A. (1988). *Innovators and imitators in innovation diffusion modelling.* In: Journal of Forecasting, 7, pp. 225-234.

Tellis, G. J. and Crawford, C. M. (1981). *An evolutionary approach to product growth theory*. In: Journal of Marketing, 45, pp. 125-134.

Van den Bulte, C. and Joshi, Y. V. (2007). *New Product Diffusion with Influentials and Imitators*. In: Marketing Science, 26, pp. 400-421.

Van den Bulte, C. and Lilien, G. L. (2001). *Medical Innovation Revisited: Social Contagion Versus Marketing Effort*. In: American Journal of Sociology, 106, pp. 1409-1435.

Equilibrium	*Equilibrium conditions*	*Signs of Eigenvalues*
E_1	$N_1^* = m_1$ and $N_2^* = m_2$	$\lambda_1 < 0, \lambda_2 < 0$
E_2	$N_1^* = \frac{-m_1 p_1}{q_1}$ and $N_2^* = m_2$	$\lambda_1 > 0, \lambda_2 <> 0$
E_3	$N_1^* = \frac{-m_1 p_1}{q_1}$ and $N_2^* = \frac{m_1(q_{12}p_1 - p_2 q_1) - p_2 m_2 q_1}{q_2 q_1}$	$\lambda_1 > 0, \lambda_2 <> 0$
E_4	$\mathrm{N}_1^* = m_1$ and $N_2^* = -\frac{1}{q_2}\left[m_1(p_2 + q_{12}) + m_2 p_2\right]$	$\lambda_1 < 0, \lambda_2 > 0$

Table 1: Stability analysis of obtained equilibria from model (11) (I)

Equilibrium	*Imaginary part*	*Stability*
E_1	no	stable nod
E_2	no	saddle path or unstable nod
E_3	no	saddle path or unstable nod
E_4	no	saddle path

Table 2: Stability analysis of obtained equilibria from model (11) (II)

Parameter	Value	Parameter	Value
p_1	0.13	m_1	1.00
p_2	0.01	m_2	1.00
q_1	0.75	ξ	0.05
q_2	0.50	ζ	5.00
q_{12}	0.07	σ	0.50

Table 3: Parameter values for model (11) and model (37)

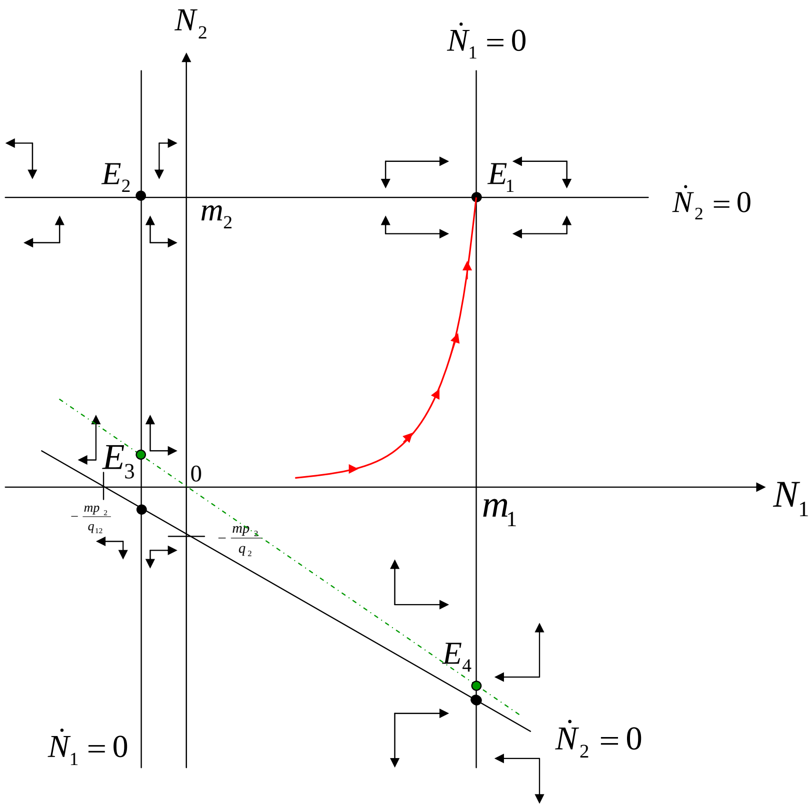

Figure 1: Phase plot of model (11) (I)

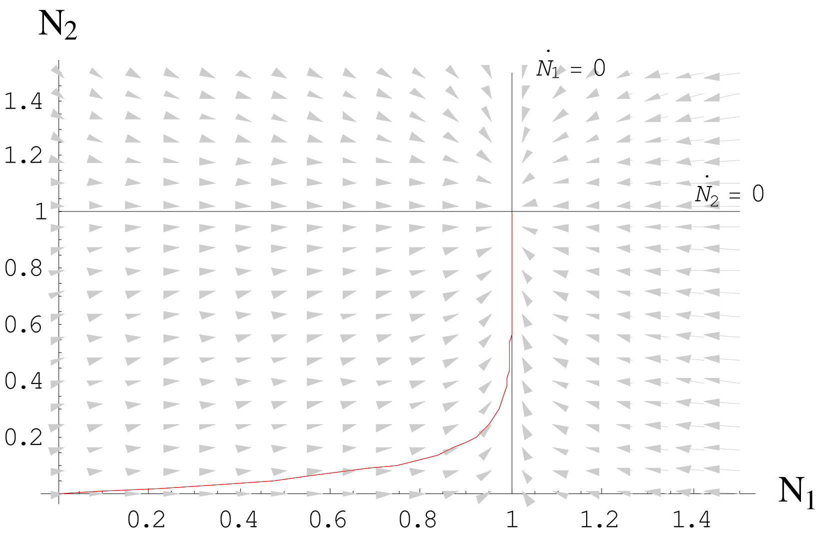

Figure 2: Phase plot of model (11) (II)

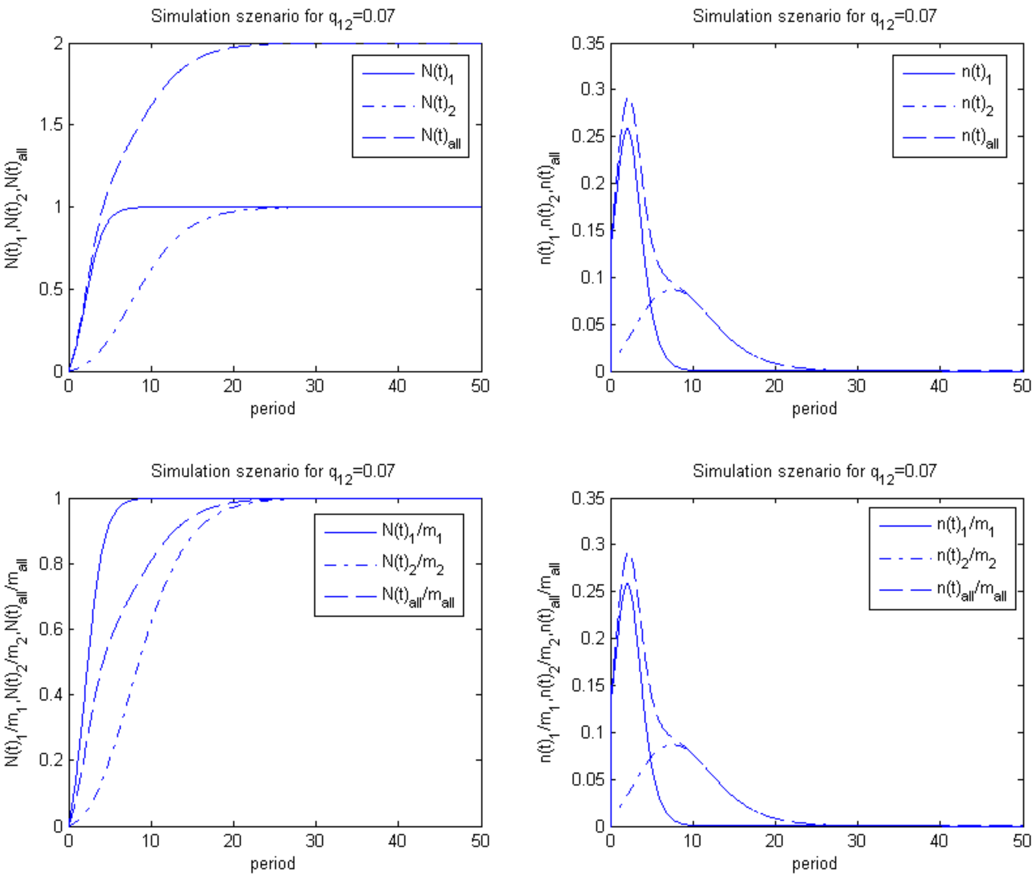

Figure 3: Graphical representation of simulated model (11) with $q_{12} = 0.07$

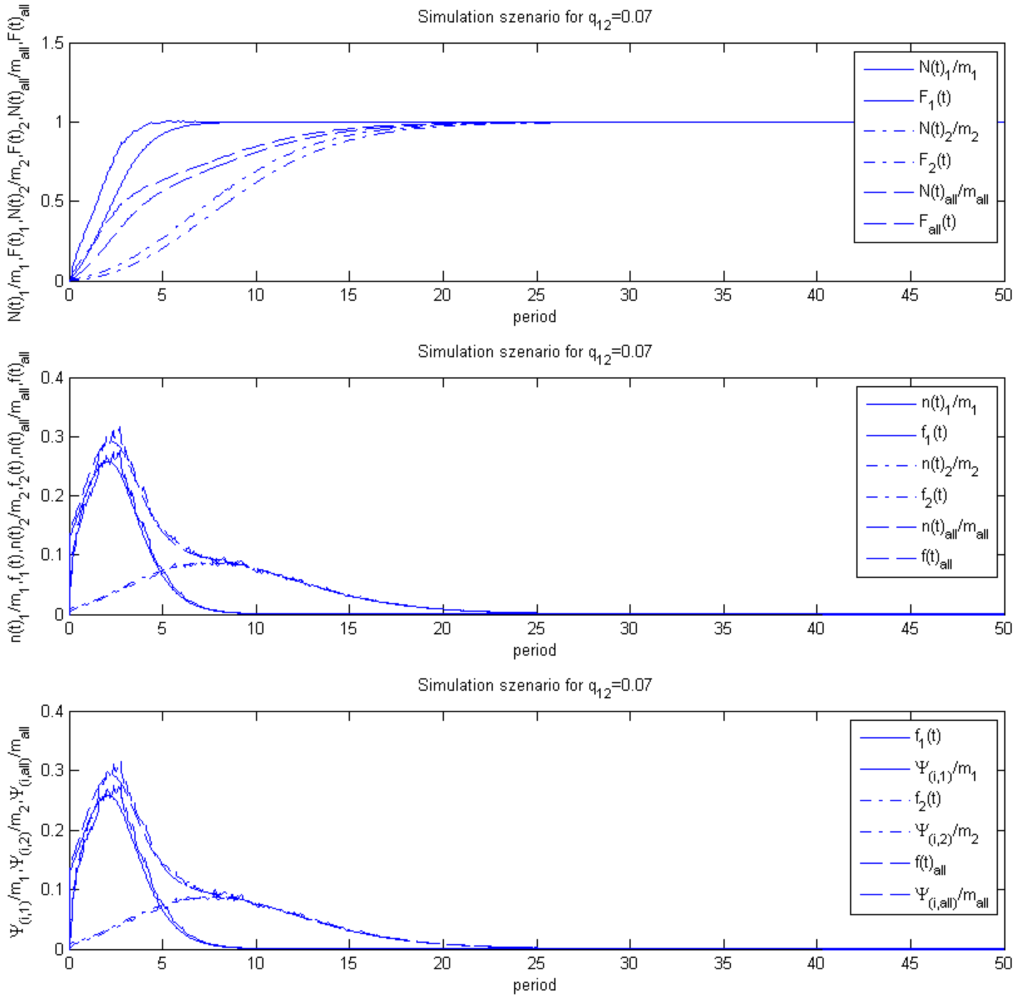

Figure 4: Graphical representation of simulated model (37) with $q_{12} = 0.07$

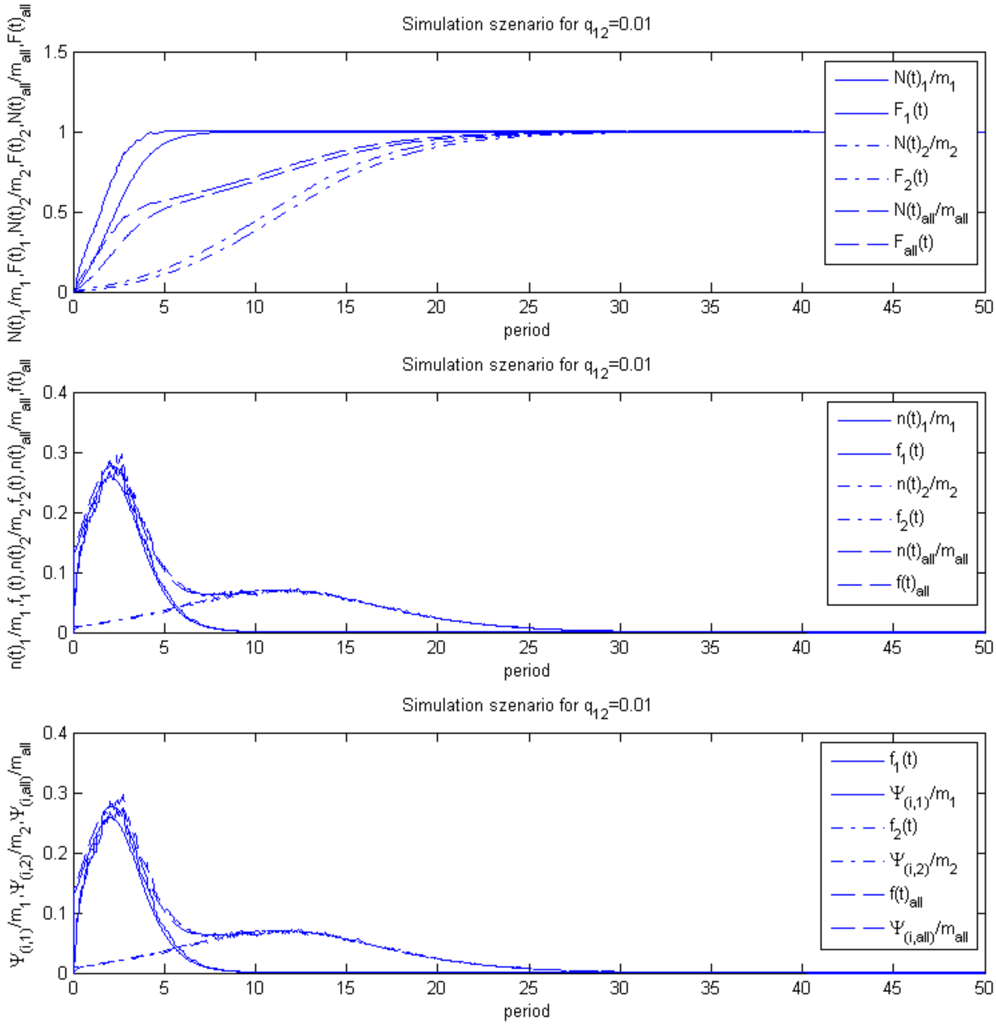

Figure 5: Graphical representation of simulated model (37) with $q_{12} = 0.01$

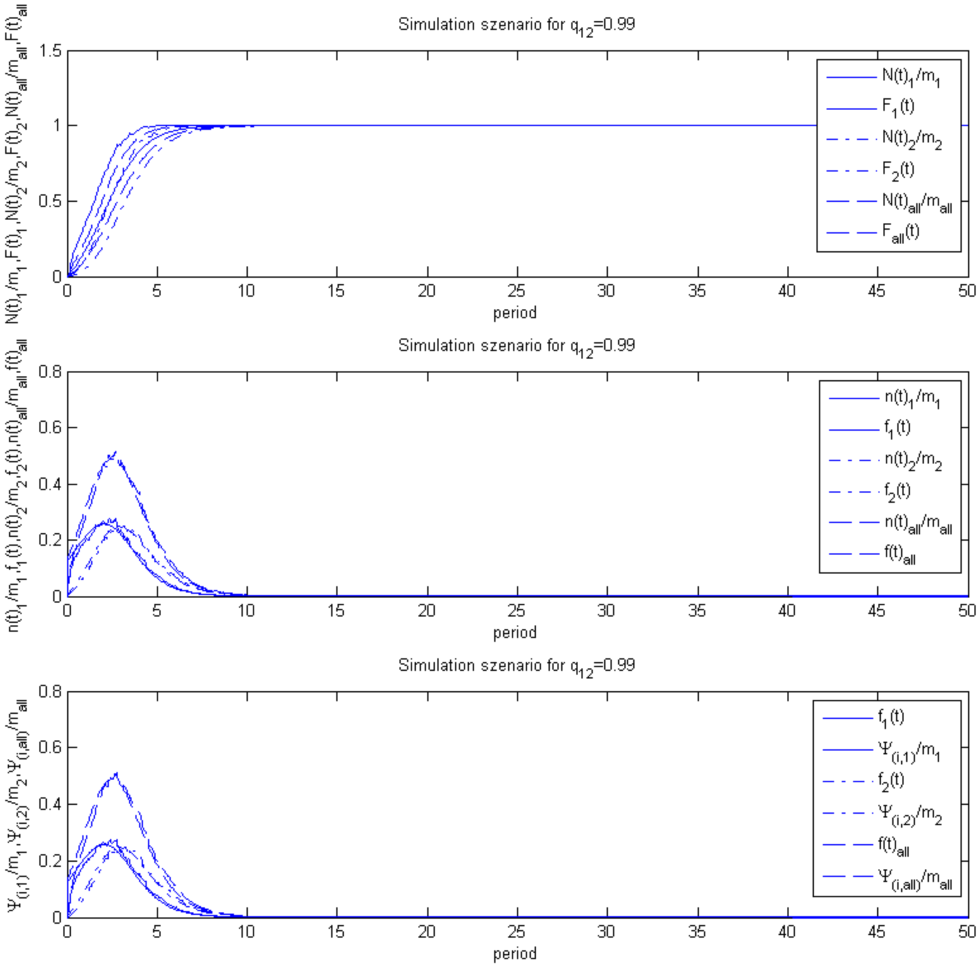

Figure 6: Graphical representation of simulated model (37) with $q_{12} = 0.99$

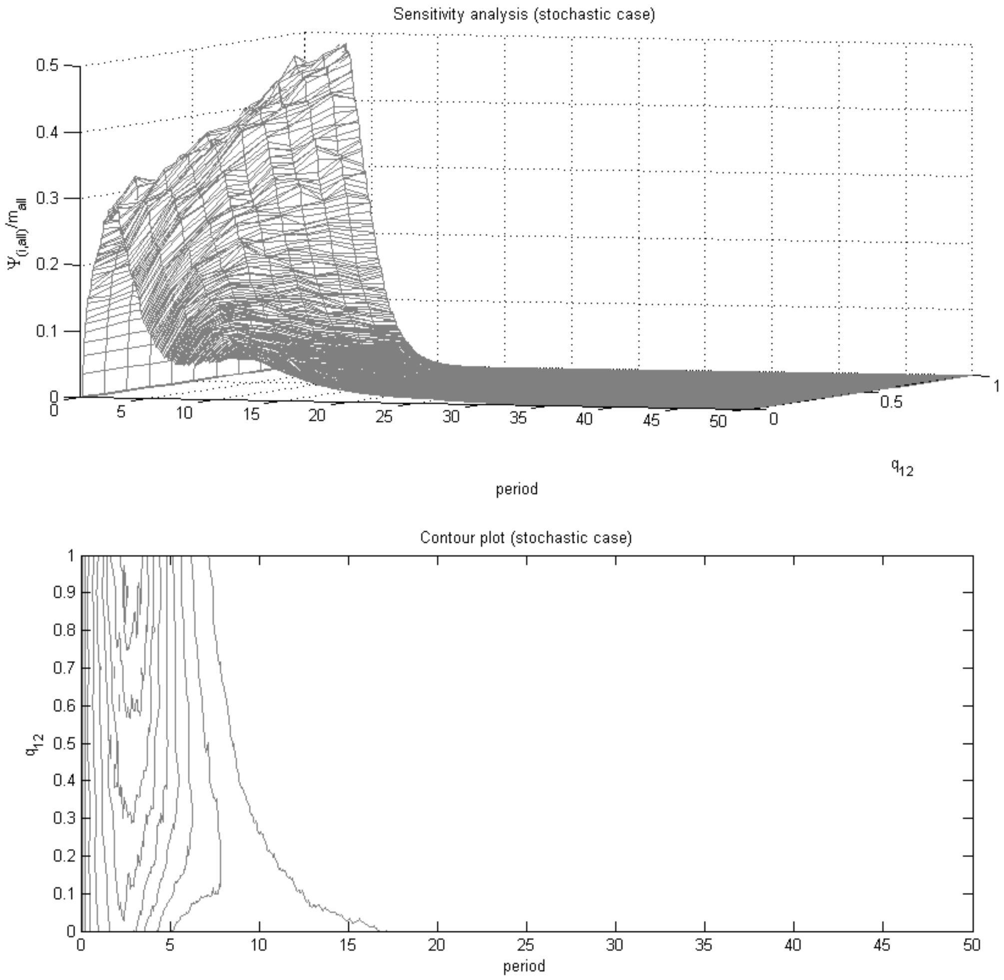

Figure 7: Approximated adoption´s density function for $q_{12} \in (0, 1)$ and corresponding contour plot based on model (37)