Working Paper Series, N. 1, March 2016

Modelling Dynamic Market Potential: Identifying Hidden Automata Networks in the Diffusion of Pharmaceutical Drugs

Renato Guseo
Department of Statistical Sciences
University of Padua
Italy
renato.guseo@unipd.it

Reinhard Schuster
Medizinischer Dienst der Krankenversicherung Nord
and Institut für Mathematik, University of Lübeck,
Ratzeburger Allee 160, 23562 Lübeck,
Germany
schuster@math.uni-luebeck.de

Abstract: This paper proposes an extension of the Guseo–Guidolin model (GGM; Guseo and Guidolin, 2009), with reference to the communication component, which is based on an assumption of complete connectivity of the hidden network supporting the growth of awareness focussed on specific pharmaceutical drugs or products with wide communication investments. The basic ideas are grounded on the extension of the Fibich–Gibori distribution (Fibich and Gibori, 2010) obtained for a minimally connected one dimensional, 1D, network topology and a subsequent inclusion, in a convex combination, with the Bemmaor–Lee distribution (Bemmaor and Lee, 2002), which takes into account unobserved heterogeneity aspects of agents in a market under a complete connectivity. Based on a continuum between opposite extremes, the extended final model, the network automata NA-GGM, allows the modulation of the involved network in a communication process that determines dynamic market potential within topologies of different complexities over time. Various applications of GGM are then given for some antidiabetic drugs in Italy. A specific application of the new model is discussed in detail with reference to a statin, Rextat, in the central part of Italy. The proposed extension, NA-GGM, is statistically globally significant with respect to the GGM and is more coherent in commercial behaviour forecasting.

Keywords: Diffusion of innovations models, dynamic market potential, network modulation, pharmaceutical drugs diffusion

Final version (2016-03-03) Submitted to tfsc.

Contents

Department of Statistical Sciences
Via Cesare Battisti, 241
35121 Padova
Italy

tel: +39 049 8274168
fax: +39 049 8274170
http://www.stat.unipd.it

Corresponding author:
Renato Guseo
tel: +39 049 827 4146
renato.guseo@unipd.it
http://www.stat.unipd.it/guseo

Modelling Dynamic Market Potential: Identifying Hidden Automata Networks in the Diffusion of Pharmaceutical Drugs

Renato Guseo
Department of Statistical Sciences
University of Padua
Italy
renato.guseo@unipd.it

Reinhard Schuster
Medizinischer Dienst der Krankenversicherung Nord
and Institut für Mathematik, University of Lübeck,
Ratzeburger Allee 160, 23562 Lübeck,
Germany
schuster@math.uni-luebeck.de

Abstract: This paper proposes an extension of the Guseo–Guidolin model (GGM; Guseo and Guidolin, 2009), with reference to the communication component, which is based on an assumption of complete connectivity of the hidden network supporting the growth of awareness focussed on specific pharmaceutical drugs or products with wide communication investments. The basic ideas are grounded on the extension of the Fibich–Gibori distribution (Fibich and Gibori, 2010) obtained for a minimally connected one dimensional, 1D, network topology and a subsequent inclusion, in a convex combination, with the Bemmaor–Lee distribution (Bemmaor and Lee, 2002), which takes into account unobserved heterogeneity aspects of agents in a market under a complete connectivity. Based on a continuum between opposite extremes, the extended final model, the network automata NA-GGM, allows the modulation of the involved network in a communication process that determines dynamic market potential within topologies of different complexities over time. Various applications of GGM are then given for some antidiabetic drugs in Italy. A specific application of the new model is discussed in detail with reference to a statin, Rextat, in the central part of Italy. The proposed extension, NA-GGM, is statistically globally significant with respect to the GGM and is more coherent in commercial behaviour forecasting.

Keywords: Diffusion of innovations models, dynamic market potential, network modulation, pharmaceutical drugs diffusion

1 Introduction

The prediction of the market potential of a new product or service is not a simple problem. Potential is not directly measurable, and some estimates are only possible

under theoretical frameworks that allow an indirect identification to some extent.

Diffusion of innovation theories (Rogers, 2003) and related methodologies, developed within technological forecasting, quantitative marketing, operations research, physics and statistics, may give a reasonable framework, particularly in the characterisation of finite life cycles that represent the temporal behaviour of the majority of the existing products or services. Relevant review papers, among other topics, have focussed on the pioneering Bass model (BM, Bass, 1969) and related extensions (Mahajan and Peterson, 1978; Mahajan and Wind, 1986; Meade and Islam, 2006; Peres et al., 2010). The equations characterising these diffusion models may have a parallel description in the context of Complex Systems, based on the local interaction of the involved agents, under a mean field approximation (Boccara, 2004; Guseo and Guidolin, 2009; Liu et al., 2013). The basic directions of research in extending the BM cover different elements, as follows: the introduction of external controls and interventions, the generalised Bass model (GBM; Bass et al., 1994), the introduction of relevant multi-cycles (Mahajan and Muller, 1998; Muller and Yogev, 1998; Robertson et al., 2007; Vakratsas and Kolsarici, 2008; Guseo and Guidolin, 2015), and the presence of discrete or continuous latent heterogeneity among agents in the system (Bemmaor, 1994; Bemmaor and Lee, 2002; Goswami and Karmeshu, 2004; Guseo and Guidolin, 2015).

To take into account some aspects of interaction among diffusion processes that share a common environment, competition is a central topic. However, its mathematical viability is essentially limited to simple duopolistic markets with diverse approaches in terms of asymmetrical communication effects and the different levels of availability of residual markets (Krishnan, Bass and Kumar, 2000, Savin and Terwiesch, 2005, Guseo and Mortarino, 2012, 2014, 2015; Guidolin and Guseo, 2015).

A relevant aspect pertaining to the study of diffusion of innovations in relatively complex markets with a huge number of competitors is the topic of prediction of the market potential that may vary during the corresponding life cycle. A basic introductory reference is Mahajan and Peterson (1978). A particular characterisation is based on the dual–market hypothesis that, with different terminologies, identifies two convenient sub-populations of adopters over time (Mahajan and Muller, 1998; Muller and Yogev, 1998; Vakratsas and Kolsarici, 2008).

A more recent proposal was inspired by some basic ideas concerning absorptive capacity (Cohen and Levinthal, 1990) and a dual effect between two main drivers in diffusion, separating the communication phase from the adoption stage. The Guseo–Guidolin model (GGM; Guseo and Guidolin, 2009) is based on the characterisation of the dynamic market potential through the evolutionary latent behaviour of a thematic network in a system with complete connectivity amongst its agents. The topic of network structures and their developments has been examined in physics and other disciplines, with many contributions in the literature (e.g., Barabasi, 2012; Handcock and Gile, 2010; Newman et al., 2011; Fibich and Gibori, 2010; Schuster and Schuster, 2015). Applications and extensions of the GGM model include energy issues, pharmaceutical drug performance, and factors related to seasonality (Guidolin and Guseo, 2012, 2014; Guseo and Guidolin, 2010, 2011).

In this paper, we examine the possibility of modelling the latent network included

in the GGM through a more realistic description of the structure of the diffusion of knowledge process in a constrained network, avoiding the implicit assumption of the GGM, based on a complete connectivity of the involved agents in a system. The paper draws on the construction of a special shifted distribution by Fibich and Gibori (2010), introduces Bemmaors's heterogeneity effects and a convex combination with the Bemmaor–Lee model. The obtained diffusion of a modulable network between a minimal connected and a maximally connected framework is introduced in a GGM by substituting its communication driver. The extended model, the network automata GGM (NA-GGM), emphasises the role of network automata.

The paper is organised as follows. Section 2 introduces the general GGM model with free dynamic market potential. Section 3 summarises the main aspects of the GGM and describes some applications in the study of the commercial trajectories of various pharmaceutical drugs in Italy. Section 4 proposes some extensions of the Fibich–Gibori approach by combining it in a convex continuum. Section 5 proposes the new extension of the GGM, the NA-GGM, and provides an illustrative application to the diffusion of a statin, Rextat, in the central part of Italy. Section 6 provides final comments, the discussion and conclusions.

2 Interventions in a model with dynamic potential: The dynamic GBM (D-GBM)

A dynamic potential, $m(t)$, can be easily implemented in the GBM and allows the double control of time allocation of adoption events and a variable scale of the potential market. Both controlling tools are relevant and act on 'orthogonal' dimensions of diffusion of innovations or growth processes.

The inclusion of interventions requires the identification of the central element of a diffusion process; in this respect, the concept of a residual market or available resource, $R(t) = m(t) - z(t)$, is a rational reference, while it would not make much sense to associate the intervention with $z(t)$, the cumulative sales at time t, which essentially describes the past. The available resource $R(t)$ may be expanded or reduced in the adopters' perception through function $x(t)$. In Equation (1), function $x(t)$ will therefore operate on the first addend, and not on the second one:

$$z'(t) = m(t)\left(p + q\frac{z(t)}{m(t)}\right)\left(1 - \frac{z(t)}{m(t)}\right)x(t) + \frac{z(t)}{m(t)}m'(t). \tag{1}$$

By setting $F(t) = z(t)/m(t)$ and $f(t) = F'(t)$, directly yields:

$$f(t) = (p + qF(t))(1 - F(t))x(t), \tag{2}$$

that is, the GBM equation (Bass et al., 1994). By imposing $F(0) = 0$, we obtain the explicit solution:

$$z(t) = m(t)\frac{1 - e^{-(p+q)\left(\int_0^t x(\tau)d\tau\right)}}{1 + \frac{q}{p}e^{-(p+q)\left(\int_0^t x(\tau)d\tau\right)}}, \qquad t \geq 0, \quad p, q > 0, \tag{3}$$

and introduce, as usual, the further constraint $z(t) = 0$ for $-\infty < t \leq 0$, which confirms the result obtained for the constant case, $m(t) = m$.

If

$$E = e^{-(p+q)\int_0^t x(\tau)d\tau},$$

the instantaneous diffusion is

$$z'(t) = m(t)\frac{p(p+q)^2x(t)E}{(p+qE)^2} + m'(t)\frac{1-E}{1+\frac{q}{p}E} \quad t \geq 0;\ p,q > 0, \tag{4}$$

and $z'(0) = 0$ for $t < 0$.

Some remarks are needed for a proper assessment of the new dynamic GBM (D-GBM) model. In this case, we face the possibility of controlling the diffusion dynamics in two different directions. First, the timeline is governed by the usual intervention function $x(t)$. Second, the 'orthogonal' direction associated with a dynamic carrying capacity depends upon convenient choices of the latent function $m(t)$, which is a totally free function like function $x(t)$. The specification of $m(t)$ is an interesting theoretical problem. Some results may be found in Guseo and Guidolin (2009, 2010, 2011).

3 The GGM: A dynamic network of knowledge and market potential

Innovation diffusion processes are usually conceived as emergent aggregate dynamics. The standard BM and the GBM are well-known paradigms in this area. They are defined through specific differential equations that establish the level of a rate as a function of initialising mechanisms combined with parallel interaction effects between two classes of agents with different attitudes and preferences. Interaction is defined between adopters, who are directly informed of the performance of a product, and non-adopters who are susceptible to change their state and are interested in acquiring relevant information to reach a decision.

The BM and GBM assume the homogeneous behaviour of agents within the two classes to simplify the description using tractable mathematics. Conversely, Complex Systems analysis emphasises the description of agents and their interpersonal relationships to depict an emergent macro-behaviour grounded on *local transition rules* governing the change of state of each agent over time. Models involving Cellular Automata or Network Automata are known examples in this field; these are usually based on computer simulation.

In social and economic systems, the change of the state of agents depends upon sufficient levels of knowledge, personal preferences, interaction and cultural effects. The diffusion of knowledge is a prerequisite or a necessary condition that determines an adoption or choice. The *absorptive capacity* of members of a Complex System interact with the collective knowledge and contribute to the accumulation of prior knowledge that expands the ability to assimilate an innovation (Cohen and Levinthal, 1990). For this reason, it is crucial the study of evolving networks related to the knowledge surrounding an innovation. Usually, these networks are only a

thematic non-observable (latent) transitory framework. Below, we summarise the properties of the GGM, and in particular, the evolution of knowledge in a particular network (Guseo and Guidolin, 2009) through elementary tools of graph theory; moreover we transform this process to attain a dynamic market potential for a particular good.

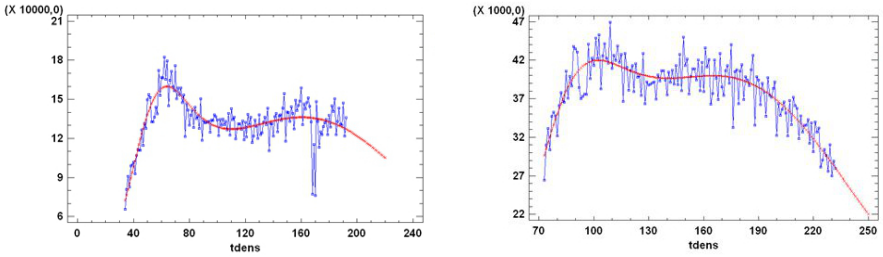

Figure 1: Glucophage 30cpr riv 500mg (left) and Metbay 30cpr 500mg (right). Packages sold in the Italian market (Source: IMS-Health Italy). The estimated GGM model is denoted by a red smooth curve.

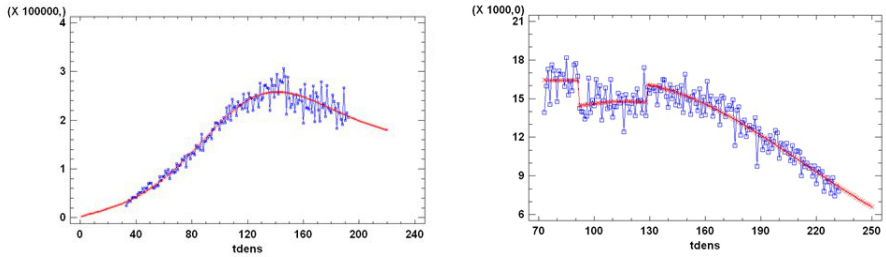

Figure 2: Metforal 30cpr riv 850mg (left) and Gliben 30cpr 5mg (right). Packages sold in the Italian market (Source: IMS-Health Italy). The estimated GGM model is denoted by a red smooth curve. A local shock is included through the control function of GGM.

A finite directed graph $G = (V, E)$ is characterised by a set $V = \{1, 2, \ldots, N\}$ of N *vertices* and a corresponding set $E \subset V \times V$ of ordered pairs (i, j) or *edges*, with $i, j \in V$. The cardinality of E is $U = C(E) \leq N^2$. We assume that each edge (i, j) may be active (existent) or non-active (non-existent) at time t. The indicator function $c(i, j; t)$ is time dependent; it equals 1 if and only if the edge $(i, j) \in E$ is active and zero otherwise.

The induction of activation of edge (i, j) at time t may be determined by a local pressure of neighbouring edges defined through the function:

$$\sigma_c(i, j; t) = \sum_{(r,s)\neq(i,j)} c(r, s; t)\, p_{r,s}, \qquad r, s \in V, \tag{5}$$

where $p_{r,s}$ is a probability measure with $\sum_{r,s\in V} p_{r,s} = 1$. Equation (5) may be generalised by considering a non-uniform distance $d(r, s)$ between vertices r and s, for instance $1 \leq d(r, s) \leq N$ to penalise the connectivity measure $p_{r,s}$. We may define a local pressure acting on (i, j) accordingly, for instance through $\sigma_c(i, j; t) =$

$\sum_{(r,s)\neq(i,j)} c(r,s;t)\, p_{r,s}/d(r,s) = \sum_{(r,s)\neq(i,j)} c(r,s;t)\, \tilde{p}_{r,s}$, which is essentially analogous to the Equation (5) that we follow here for simplicity.

Under a mean field approximation, the local pressure on edge (i,j) is a mean value, namely a probability:

$$\sigma_c(i,j;t) \simeq v(t) = \sum_{r,s} \frac{c(r,s;t)}{U}. \tag{6}$$

The *transition rule* that determines the state of an edge (i,j) at time $t+1$ is

$$c(i,j;t+1) = c(i,j;t) + Bi(1,p_c)I_{c(i,j;t)=0} + Bi(1, q_c\ \sigma_c(i,j;t))I_{c(i,j;t)=0}, \tag{7}$$

where $Bi(1,p_c)$ denotes a binomial experiment conditional upon $c(i,j;t)=0$, such that with probability p_c it may be determined by an interaction with an external information. Similarly, $Bi(1, q_c\ \sigma_c(i,j;t))$ represents a binomial experiment, conditional upon $c(i,j;t)=0$, which depicts a possible change of state in edge (i,j) through the interaction between the local pressure $\sigma_c(i,j;t)$ and the probability q_c to exploit this internal information channel.

The average number of active edges within E at time t is as follows:

$$Uv(t+1) = U\{v(t) + [p_c + q_c v(t)][1 - v(t)]\}. \tag{8}$$

The corresponding continuous approximation is an autonomous Riccati equation, specifically,

$$v'(t) = -q_c v^2(t) + (q_c - p_c)v(t) + p_c, \tag{9}$$

whose solution is the well-known BM under the initial condition $v(0)=0$.

Function $Uv(t)$ depicts the cumulative growth of active edges (i,j) within E over time. We may rescale this function to describe the transmitted knowledge through edges with reference to the activated vertices, as follows:

$$k(t) = \sqrt{U}\sqrt{v(t)}. \tag{10}$$

As sustained in Guseo and Guidolin (2009) a simple generalisation of the $Uv(t)$ transformation may be based on $v(t)^\alpha$ to take into account a possible dimensional collapse of $E \subset V \times V$ and recognised empirically through the estimation of the unknown parameter α. Nevertheless, as we propose in Sections 4 and 5, we may modulate the growing network topology of relationships through a parametric mixture of the corresponding extremal structures that depict opposite connectivities amongst agents (vertices).

More specifically, a dynamic market potential, driven by previous evolutionary expansion of knowledge in an unconstrained network, may be represented as follows:

$$m(t) = K\sqrt{\frac{1 - e^{-(p_c+q_c)t}}{1 + \frac{q_c}{p_c}e^{-(p_c+q_c)t}}}, \tag{11}$$

i.e., a monotone increasing transformation of a BM for cumulative active relationships where K denotes the asymptotic market potential.

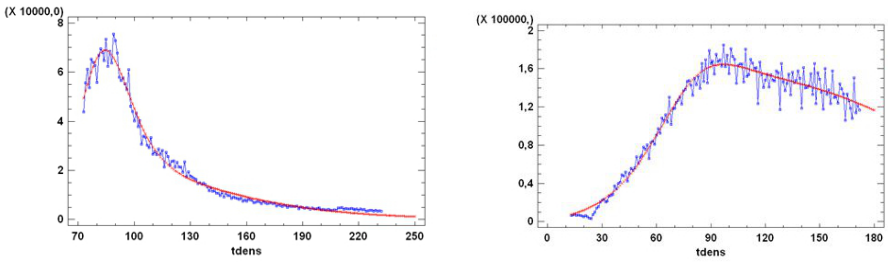

Figure 3: Diabrezide 40cpr 80mg (left) and Solosa 30cpr 2mg (right). Packages sold in the Italian market (Source: IMS-Health Italy). The estimated GGM model is denoted by a red smooth curve.

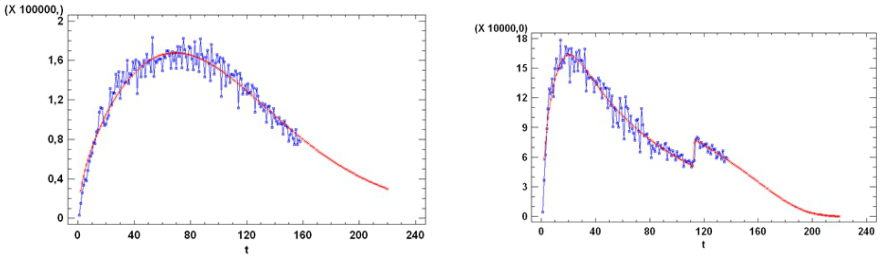

Figure 4: Gliconorm 36cpr riv 5mg 500mg (left) and Gliclazide Molteni 40cpr 80mg (right). Packages sold in the Italian market (Source: IMS-Health Italy). The estimated GGM model is denoted by a red smooth curve. A local shock is included through the control function of GGM.

Under a GBM with dynamic potential, D-GBM, as described in Section 2, we obtain the GGM, as follows:

$$z(t) = K\sqrt{\frac{1-e^{-(p_c+q_c)t}}{1+\frac{q_c}{p_c}e^{-(p_c+q_c)t}}}\cdot\frac{1-e^{-(p_s+q_s)t}}{1+\frac{q_s}{p_s}e^{-(p_s+q_s)t}}, \tag{12}$$

where the second factor depends upon parameters p_s and q_s and describes the specific adoption process. In other words, the normalised GGM is the product of two distribution functions,

$$z(t) = K\sqrt{F(t)}\,G(t) = KH(t), \tag{13}$$

and separates two driving sub–processes, as follows: the awareness component, which is mainly related to the communication aspects, and the adoption phase, which describes the decision timing. Note that Equation (12) may be generalised by including control or intervention functions, as expressed in Equation (3).

The application of the model to oral antidiabetic drug sales in Italy (see Fig. 1, 2, 3 and 4) emphasises the flexibility of the GGM model and the careful fitting depicting, in particular, different behaviour in the launch phase that may be interpreted through various perspectives, as explained below. Note that the right side of Fig. 2 and 4 show the presence of specific shocks due to changes in regulations; these are recognized through the intervention function.

The density $h(t)$ of $H(t)$ is a dynamic mixture of the basic densities of the two

sub–processes, that is:

$$h(t) = H'(t) = \frac{1}{2}F^{-1/2}(t)G(t)f(t) + F^{1/2}(t)g(t) = k_1(t) + k_2(t), \qquad (14)$$

where $f(t) = F'(t)$, $g(t) = G'(t)$, and $k_1(t) > 0$, $k_2(t) > 0$.

Decomposition (14) is discussed in Guseo and Guidolin (2011). There are two main results. First, the density $h(t)$ may be unimodal or bimodal and this aspect emphasises the composition of the two main forces in the diffusion of innovations, namely communications and adoption. The functions $k_1(t)$ and $k_2(t)$ may be associated accordingly and may produce an interesting *slowdown*. This effect establishes a kind of pause in sales after the first effort to launch the product.

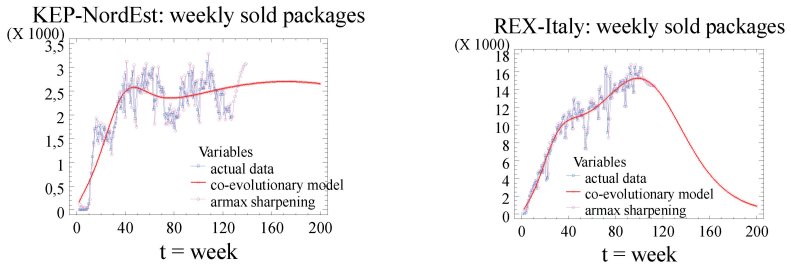

Figure 5: Keplat North–East of Italy (left) and Rextat in Italy (right). Packages sold in the Italian market (Source: IMS-Health Italy). The estimated GGM model is denoted by a red smooth curve.

The second result is that the two positive contributions in $h(t)$, $k_1(t)$ and $k_2(t)$, may be ordered over time through a *likelihood ratio order* that is stronger than the usual *stochastic order*. A *weaker order* that is implied by the likelihood ratio order may be obtained through location indices, as follows:

$$\begin{aligned} t^+_{com} &= \frac{\ln d}{c}; \qquad {}_Ft_{0.5} = \frac{1}{c}\ln(2+d); \qquad E(T_F) = \bar{t}_F = \frac{1}{q_c}\ln(1+d), \\ t^+_{ado} &= \frac{\ln b}{a}; \qquad {}_Gt_{0.5} = \frac{1}{a}\ln(2+b); \qquad E(T_G) = \bar{t}_G = \frac{1}{q_s}\ln(1+b), \end{aligned} \qquad (15)$$

where $a = p_s + q_s$, $b = q_s/p_s$, $c = p_c + q_c$, $d = q_c/p_c$. Usually, the communication component dominates the adoption component in the sense that it comes first. The institutional effort during the launch is essential for products that are not radical innovations and require communication investments. In special situations, a new product may be strongly expected in the launch phase by potential customers that are well-informed about its performance, and this context may produce a timing inversion with the adoption component that drives the commercial behaviour and a communication effort that acts as a maintenance driver. A typical example is a pharmaceutical drug for severe pathologies that elicits an accumulation of demand at the beginning of the commercial process. See in particular Fig. 5, which depicts the performance of Keplat is an over-the-counter (OTC)) drug for minor pain and inflammations monitored in north–east Italy, and Rextat is a new statin launched in

Italy. The former is a traditional product with the normal behaviour of 'first communication, then adoption'. The latter gives evidence for an accumulation of demand confirmed by an inversion in driving factors – 'first adoption, then communication'.

4 Connectivity, seeding and heterogeneity effects

In this section, we propose a definition of a parameterised continuum that allows the description, in the next section, of a latent network of relationships belonging to a wide class that includes two extremal connectivity hypotheses.

The logistic component in the standard BM (Bass, 1969), $\frac{q}{m}z(t)(m - z(t))$, is characterised by a perfect connectivity assumption amongst informed agents, $z(t)$, and susceptible agents, $(m-z(t))$. In real–world contexts, connectivity is modulated by physical, social, cultural, technological and economic constraints. See, among others, Pastor-Satorras and Vespignani (2007), Milgram (1967), McCubbins et al. (2009), Dorogovtsev and Mendes (2003), Schuster and Schuster (2015), and Fibich and Gibori (2010).

A member of susceptible agents $(m-z(t))$ may interact at time t with a variable and limited number of neighbours, $0 \leq n(t) \leq z(t)$. To some extent, parameter q may absorb this limitation in connectivity, but it works by assuming a complete network of relationships.

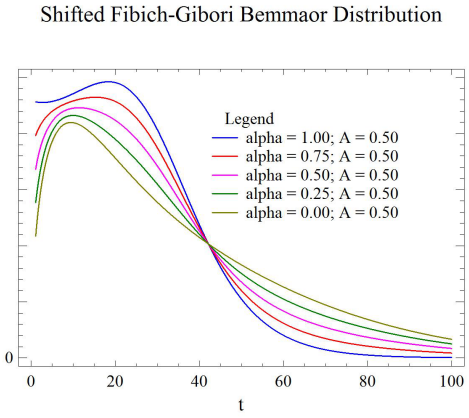

Figure 6: Shifted Fibich–Gibori Bemmaor distribution with heterogeneity parameter $A = 0.50$ and different levels α of mixing between extremes.

Fibich and Gibori (2010) examined two minimally connected networks in a circle, namely: a one–sided one-dimensional (1D) model where each individual can be influenced by his/her left-neighbor or a two-sided 1D model where each individual can be influenced by two neighbours. Under previous connectivity assumptions, there exists a common limiting distribution in the diffusion process, namely:

$$F_{FG}(t) = 1 - e^{-(p+q)t+\frac{q}{p}(1-e^{-pt})}, \tag{16}$$

where parameter p expresses the external effect acting on an agent, and q denotes the internal effect due to the neighbour(s). A second result mentioned in Fibich and

Gibori (2010) established that the normalised distribution $F(t)$ in an agent-based model with an intermediate spatial structure of relevant relationships is dominated by the standard Bass distribution $F_B(t)$ and dominates $F_{FG}(t)$, as follows:

$$F_{FG}(t) \leq F(t) \leq F_B(t). \tag{17}$$

It is well known that the standard Bass distribution $F_B(t)$ is the product of a monomolecular by a logistic distribution, that is:

$$F_B(t) = (1 - e^{-(p+q)t}) \frac{1}{(1 + \frac{q}{p} e^{-(p+q)t})} = M(t)L(t). \tag{18}$$

The initialisation mechanism of $F_{FG}(t)$ may be distributed over time through the multiplication by a monomolecular model $M(t)$ determining a shifted Fibich–Gibori distribution:

$$F_{SFG}(t) = M(t)F_{FG}(t). \tag{19}$$

Obviously, $F_{SFG}(t)$ is dominated by $F_{FG}(t)$, and therefore:

$$F_{SFG}(t) \leq F_B(t), \tag{20}$$

or, by eliminating $M(t)$:

$$F_{FG}(t) \leq L(t). \tag{21}$$

In other words, the logistic distribution dominates the corresponding Fibich–Gibori's.

The presence of heterogeneous behaviour of adopters in Bemmaor and Lee's (2002) approach, in contrasts, determines an exponentiation of the logistic component, $L^A(t)$, giving rise to the specific Bemmaor–Lee distribution, $F_{BL}(t)$:

$$F_{BL}(t) = M(t)L^A(t). \tag{22}$$

We may define a more general shifted Fibich–Gibori distribution with heterogeneity effects through an exponentiation of $F_{FG}(t)$ by A and a shifting based on $M(t)$, as follows:

$$F_{SFGB}(t) = M(t)F^A_{FG}(t). \tag{23}$$

Clearly, the Bemmaor–Lee distribution dominates the shifted Fibich–Gibori with heterogeneity effects:

$$F_{SFGB}(t) \leq F_{BL}(t). \tag{24}$$

Minimally connected networks, 1D and the complete connectivity assumed in the Bm and Bemmaor and Lee model determine extremal configurations. Intermediate spatial networks, which are not observable, may be represented though a parametric mixture of the extremes, as follows:

$$G(t) = \alpha F_{BL}(t) + (1 - \alpha)F_{SFGB}(t), \qquad \alpha \in [0, 1]. \tag{25}$$

Model (25) is a normalised distribution over time depending upon several parameters, as follows: p is the external effect, q is the internal effect, A represents heterogeneity and α represents an indirect measure of connectivity. Moreover, m is an

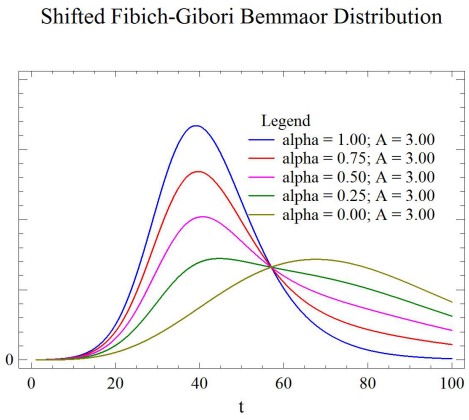

Figure 7: Shifted Fibich–Gibori Bemmaor distribution with heterogeneity parameter $A = 3$ and different levels α of mixing between extremes.

external scale parameter representing the absolute market potential value so that the cumulative sales may be described explicitly, by setting $b = p + q$ and $\beta = q/p$ through:

$$\begin{aligned} z(t) &= mG(t) \\ &= m\left\{\alpha\frac{(1-e^{-bt})}{(1+\beta e^{-bt})^A} + (1-\alpha)(1-e^{-bt})\left(1-e^{-bt+\beta(1-e^{-pt})}\right)^A\right\}. \end{aligned} \tag{26}$$

Figures 6 and 7 illustrate the high flexibility of the shifted Fibich–Gibori Bemmaor distribution with heterogeneity parameter $A = 0.50$ or $A = 3$ and different levels α of the amount of mixing between the opposite extremes.

A more general mixture of the two driving distributions, $F_{BL}(t)$ and $F_{SFGB}(t)$, may be obtained with local free parameters by setting $b_1 = p_1 + q_1$, $\beta_1 = q_1/p_1$, $b_2 = p_2 + q_2$ and $\beta_2 = q_2/p_2$, that is:

$$\begin{aligned} z(t) &= m\tilde{G}(t) \\ &= m\left\{\alpha\frac{(1-e^{-b_1 t})}{(1+\beta_1 e^{-b_1 t})^{A_1}} + (1-\alpha)(1-e^{-b_2 t})\left(1-e^{-b_2 t+\beta_2(1-e^{-p_2 t})}\right)^{A_2}\right\}. \end{aligned} \tag{27}$$

Nevertheless, generality has to be carefully balanced to avoid contradictory results in applications. The augmented flexibility due to hierarchical extension provides a better global fitting (monotonicity theorem) within the observed region, but this does not imply a gain in out-of-sample forecasting.

5 The NA-GGM: Modelling of hidden Network Automata

In this section, we propose a more comprehensive description of the 'communication' component in the GGM model that takes into account the main result described in Section 4. Connectivity, seeding and heterogeneity are relevant components in determining a flexible framework for a hidden growing thematic network. The two opposite and extreme topologies refer to the 1D approach by Fibich and Gibori

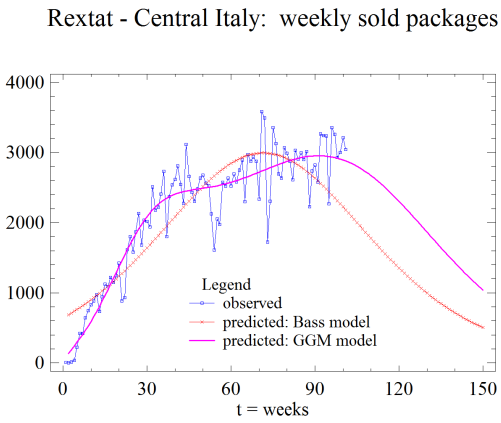

Figure 8: The standard GGM model based on a complete connectivity assumption in the awareness process determining a dynamic market potential. The model is applied to weekly sold packages of Rextat in Central Italy. The figure compares the GGM with a standard BM model based on a fixed market potential assumption.

(2010) and the fully connected network represented via Bemmaor and Lee's (2002) asymmetric approach. The mixed distribution of Equations (26) and (27) may be introduced in the 'communication' component of the GGM as follows, by setting $b_c = p_c + q_c$, $\beta_c = q_c/p_c$ and considering two separate sources of heterogeneity, A_1 and A_2:

$$z(t) = K\sqrt{\alpha\frac{(1-e^{-b_c t})}{(1+\beta_c e^{-b_c t})^{A_1}} + (1-\alpha)(1-e^{-b_c t})\left(1-e^{-b_c t+\beta_c(1-e^{-p_c t})}\right)^{A_2}}.\frac{(1-e^{-(p_s+q_s)t})}{(1+\frac{q_s}{p_s}e^{-(p_s+q_s)t})}. \quad (28)$$

Parameter α controls the mixture between extremal topologies and related dynamic growth. In particular, for $\alpha = 1$ and $A_1 = 1$, we obtain the GGM model (Guseo and Guidolin, 2009). We call this new framework the network automata GGM (NA-GGM) to emphasise the more general nature of the expanding hidden network in real contexts where cultural, historical and traditional barriers may limit a full expansion of awareness towards a specific item, drug, product or service.

To evaluate the performances of the extended NA-GGM model with reference to the special cases of GGM and the classical standard BM, we have focussed on various drugs. Here, we present the diffusion of Rextat, a statin, in the central part of Italy. The source of weekly data is IMS-Health and the series depicts the number of packages sold during the period of April 2005 - August 2007. A sequence of models of BM, GGM and NA-GGM is suggested to use convenient estimations for the initial values of more complex models. The obtained results are based on the nonlinear least squares (NLS) methodology through the Levenberg–Marquardt algorithm (Seber and Wild, 1989) and are summarised in Tables 1, 2, 3 and 4.

The BM model presents a sufficient global fitting with $R^2_{cum} = 0.998089$, but it fails in different directions. In Fig. 8 we observe the initial overestimation followed by the failure in recognising the evident typical 'slowdown' of many pharmaceutical drugs, and finally, the evident underestimation due to an incoherent quick closure of the cycle.

Table 1: Parameter estimates of a standard BM model for the Rextat series of packages sold in Central Italy; ( ) marginal linearised asymptotic 95% confidence limits. Estimates performed on cumulative data.

m	p	q	
287465	0.00227369	0.0369784	$R^2_{cum} = 0.998089$
(273978)	(0.00218444)	(0.034726)	$RSS = 918,362,000$
(300953)	(0.00236294)	(0.039231)	$DW = 0.0229153$

Table 2: Parameter estimates of a GGM model for the Rextat series of packages sold in Central Italy; ( ) marginal linearised asymptotic 95% confidence limits. Estimates performed on cumulative data.

K	p_c	q_c	p_s	q_s	
345752	0.000386819	0.0426465	0.00980359	0.0786105	$R^2_{cum} = 0.999921$
(326342)	(0.000372499)	(0.0405032)	(0.00917997)	(0.0725633)	$RSS = 37,936,500$
(365163)	(0.000401139)	(0.0447897)	(0.01042720)	(0.0846578)	$DW = 0.262115$

The GGM performs well with a high global determination index, $R^2_{cum} = 0.999921$, and a partial correlation coefficient defined with respect to the BM, $\tilde{R}^2 = 0.95866$, which confirms a noticeable gain. With a standard terminology, the F-ratio, whose critical value at %95 level is about 4, reaches an excellent value, $F = 1113.11$, as summarised in Table 4. These assessments are confirmed by a graphical analysis as displayed in Fig. 8, where the principal drawbacks of the BM model are widely exceeded by the the properties of the GGM.

In this special situation, the residual space to improve the GGM is quite challenging. In Tables 3 and 4, we show that $R^2_{cum} = 0.999949$ defines a natural improvement of NA-GGM due to the monotonicity theorem. The main issue is expressing a convincing test of the statistical significance of the NA-GGM extension.

In this case, with a minor emphasis on the allowable limited space, the partial correlation coefficient $\tilde{R}^2 = 0.354430379$ and the corresponding F-ratio, $F = 17.57$, confirm the presence of a dynamic network sustaining the diffusion of Rextat that is not fully connected. Parameter $\alpha = 0.4752$ denotes the presence of a network far from the usual Bass or Bemmaor and Lee models, with an intermediate connectivity with respect to the Fibich–Gibori extreme.

Parameter $A_1 = 93.779$ denotes the heavy heterogeneity of agents in the Bass-

Table 3: Parameter estimates of a NA-GGM model for the Rextat series of packages sold in Central Italy; ( ) marginal linearised asymptotic 95% confidence limits. Estimates performed on cumulative data.

K	α	p_c	q_c	$A1$	p_s	q_s	$A2$	
431021	0,47527	0,053613	0,10668	93,779	0,002313	0,022438	2,598	$R^2 = 0.999949$
(403595)	(0,31221)	(-0,02274)	(0,00837)	(-51,9511)	(0,002248)	(0,021040)	(-0,078)	$RSS = 24,313,200$
(458446)	(0,63833)	(0,12997)	(0,20499)	(239,509)	(0,002378)	(0,023836)	(5,274)	$DW = 0,388617$

Table 4: Performance comparisons amongst standard BM, GGM and NA-GGM. $\tilde{R}^2 = (R^2_{m2} - R^2_{m1})/(1 - R^2_{m1})$ is the squared multiple partial correlation coefficient between two nested models $m1$ and $m2$, where $m1$ is included in $m2$ and R^2 is the standard determination index. The F-ratio $F = [\tilde{R}^2(n-v)]/[(1-\tilde{R}^2)u]$ has a critical threshold about 4 for $u = 1$ or a lower value for $u > 1$. In addition, n is the number of observations, v is the number of parameters of the extended model $m2$, and u is the incremental number of parameters from $m1$ to $m2$.

	BM	GGM	NA-GGM
BM	$R^2_{cum} = 0.998089$	$\tilde{R}^2 = 0.958660$ $F = 1113.11$	$\tilde{R}^2 = 0.973312401$ $F = 678.353$
GGM		$R^2_{cum} = 0.999921$	$\tilde{R}^2 = 0.354430379$ $F = 17.57$
NA-GGM			$R^2_{cum} = 0.999949$

like communication component, while on the opposite side, the more constrained network structure is weakly heterogeneous or practically homogeneous; here, $A_2 = 2.59839 \simeq 1$, representing the equilibrium point.

The graphical analysis showed in Fig. 9 explains the nature of the improvement both in the initial part of the process, where some deviations are systematically perceived by NA-GGM, and in the final out-of-sample projection, where NA-GGM fixes one of the most frequent problems of BM, namely the incoherent quick closure of the cycle.

6 Discussion and concluding remarks

In this paper, we examined the communication component of the general D-GBM in Equations (1) and (3) to properly include some networking constraints that limit the diffusion of knowledge and related awareness about the properties of a product, and particularly, of a pharmaceutical drug. The centrality of communication efforts of pharmaceutical companies is well known and is provided through parallel channels, as follows: institutional and advertising investments in meetings, conferences, publications in specialised professional magazines and direct investment in the activities of sales representatives (detailmen). These channels are completed through internal communication and word-of-mouth amongst physicians, as well as indirectly through patients' perceptions.

To model different diffusion processes of specific knowledge on a continuum, we combined the diffusions that pertain to opposite possible topologies of the dynamic network. The Fibich–Gibori model, with monomolecular shifting and an extension expressing heterogeneous agents' behaviour, corresponds to the most constrained network (1D). The Bemmaor and Lee model referred to a completely accessible

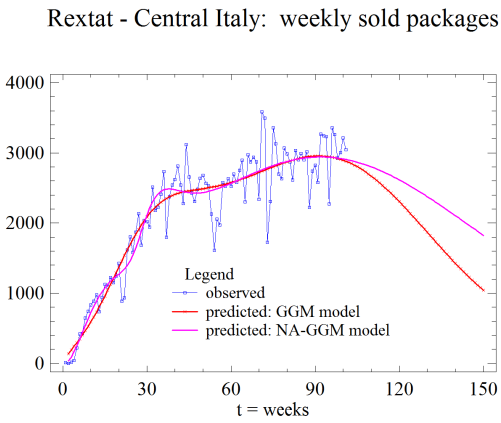

Figure 9: NA-GGM: Shifted Fibich–Gibori Bemmaor mixed control of network automata in the definition of the dynamic market potential in a GGM model with modulated connectivity. The figure applies to weekly sold packages of Rextat in Central Italy. The NA-GGM is compared with the standard GGM based on the complete connectivity assumption.

system. The convex combination of these extremes with a continuous parameter α, a probability in mixing, allowed flexible modulation in the empirical identification and estimation.

Due to the monotonicity theorem in nested models, the obtained theoretical result in general is relatively good in terms of fitting properties . Nevertheless, a specific simple test may establish whether the extension is significant. The case of Rextat for central Italy soundly confirms this and similar examples that are not reported here for brevity (e.g. Lyrica), exhibit a similar significant gain.

The proposed model, NA-GGM, is surely of interest for fine tuning when it comes to monitoring specific drugs at a company level for the specific strategies and the related managerial implications. A different aspect is related to the systematic application at the category level, where large numbers of different specialties are evident (different AIC codes in Italy). For instance, there are about 200 non-insulin antidiabetic drug specialties, ATC code A10B, in Italy, about 500 in Sweden and about 2000 in Germany. In this more complex framework, where a third party payer – the Ministry of Health or insurance companies – has to tackle simultaneous supply counterparts, the simpler GGM be a sufficient approximation for practical large–scale evaluations.

References

A.-L. Barabasi, Network science: Luck or reason, Nature 489(7417) (2012) 507–508.

F.M. Bass, A new product growth model for consumer durables, Management Science 15 (1969) 215–227.

F.M. Bass, T. Krishnan, D. Jain, Why the Bass model fits without decision variables, Marketing Science 13 (1994) 203–223.

A.C. Bemmaor, Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity. In: G. Laurent, G.L. Lilien, B. Pras (Eds.) Research Traditions in Marketing, 201–229, Kluwer Academic, Boston, MA, 1994.

A.C. Bemmaor, J. Lee, The impact of heterogeneity and ill-conditioning on diffusion model parameter estimates, Marketing Science 21 (2002) 209–220.

N. Boccara, Modeling complex systems, Springer-Verlag, New York, 2004.

W.M. Cohen, D.A. Levinthal, Absorptive capacity: a new perspective on learning and innovation, Administrative Science Quarterly 35 (1990) 128–152.

L. Donetti, P.I. Hurtado, M.A. Munoz, Entangled networks, synchronization, and optimal network topology, Physical Review Letters 95(18) (2005) 188701.

S.N. Dorogovtsev, J.F.F. Mendes, Evolution of Networks: From biological nets to the Internet and WWW, Oxford University Press, Oxford, 2003.

G. Fibich, R. Gibori, Aggregate Diffusion Dynamics in Agent-Based Models with a Spatial Structure, Operations Research 58(5) (2010) 1450–1468.

D. Goswami, Karmeshu, Modelling data uncertainty in innovation diffusion model: simulation based on simulated annealing, Technological Forecasting and Social Change 71 (2004) 705–722.

M. Guidolin, R. Guseo, A nuclear power renaissance?, Technological Forecasting and Social Change 79(9) (2012) 1746–1760.

M. Guidolin, R. Guseo, Seasonality in Innovation Diffusion, Technological Forecasting and Social Change 86 (2014) 33–40.

M. Guidolin, R. Guseo, Technological change in the U.S. music industry: within-product, cross-product and churn effects between competing blockbusters, Technological Forecasting and Social Change 99 (2015) 35–46.

R. Guseo, M. Guidolin, Modelling a dynamic market potential: A class of automata networks for diffusion of innovations, Technological Forecasting and Social Change 76(6) (2009) 806–820.

R. Guseo, M. Guidolin, Cellular Automata with Network Incubation in Information Technology Diffusion, Physica A: Statistical Mechanics and its Applications 389(12) (2010) 2422–2433.

R. Guseo, M. Guidolin, Market potential dynamics in innovation diffusion: modelling the synergy between two driving forces, Technological Forecasting and Social Change 78(1) (2011) 13–24.

R. Guseo, M. Guidolin, Heterogeneity in diffusion of innovations modelling: A few fundamental types, Technological Forecasting and Social Change 90 (2015) 514–524.

R. Guseo, C. Mortarino, Sequential market entries and competition modelling in multi-innovation diffusions, Eur. J. Oper. Res. 216(3) (2012) 658–667.

R. Guseo, C. Mortarino, Within-brand and cross-brand word-of-mouth for sequential multi-innovation diffusions, IMA J. Man. Math. 25(3) (2014) 287–311.

R. Guseo, C. Mortarino, Modeling competition between two pharmaceutical drugs using innovation diffusion models, The Annals of Applied Statistics 9(4) (2015) 2073–2089.

M.S. Handcock, K.J. Gile, Modeling social networks from sampled data, The Annals of Applied Statistics 4(1) (2010) 5–25.

T.V. Krishnan, F.M. Bass, V. Kumar, Impact of a Late Entrant on the Diffusion of a new Product/Service, Journal of Marketing Research XXXVII (2000) 269–278.

Y.-Y. Liu, J.-J. Slotine, A.-L. Barabasi, Observability of complex systems, Proceedings of the National Academy of Sciences of the United States of America 110(7) (2013) 2460–2465.

V. Mahajan, E. Muller, When is it worthwhile targeting the majority instead of the innovators in a new product launch?, Journal of Marketing Research XXXV (1998) 488–495.

V. Mahajan, R. Peterson, Innovation diffusion in a dynamic potential adopter population, Management Science 24(15) (1978) 1589–1597.

V. Mahajan, R. Peterson, Models for Innovation Diffusion, SAGE Publications, Newbury Park, CA, 1985.

V. Mahajan, Y. Wind, Innovation Diffusion Models of New Product Acceptance, HarperCollins Publishers, New York, 1986.

M.D. McCubbins, R. Paturi, N. Weller, Connected coordination network structure and group coordination, American Politics Research 37(5) (2009) 899–920.

N. Meade, T. Islam, Modelling and forecasting the diffusion of innovation – a 25-year review, International Journal of Forecasting 22(3) (2006) 519–545.

S. Milgram, The small world problem, Psycology Today 1(1) May (1967) 61–67.

E. Muller, G. Yogev, When does the majority become a majority? Empirical analysis of the time at wich main market adopters purchase the bulk of our sales, Technological Forecasting and Social Change 73 (2006) 1107–1120.

M. Newman, A.-L. Barabasi, D.J. Watts, The structure and dynamics of networks, Princeton University Press, Princeton, 2011.

R. Pastor-Satorras, A. Vespignani, Evolution and Structure of the Internet: A statistical physics approach, University Press, Cambridge, 2007.

R. Peres, E. Muller, V. Mahajan, Innovation diffusion and new product growth models: a critical review and research directions, International Journal of Research in Marketing 27(2) (2010) 91–106.

A. Robertson, D. Soopramanien, R. Fildes, Segmental new-product diffusion of residential broadband services, Telecommunications Policy 31 (2007) 265–275.

E.M. Rogers, Diffusion of innovations, 5th ed., Free Press, New York, 2003.

S. Savin, C. Terwiesch, Optimal product launch times in a duopoly: balancing life-cycle revenues with product cost, Operations Research 53(1) (2005) 26–47.

R. Schuster, Stückweise epsilon-konvexe und epsilon-kokave Funktionen in der Arzneimittelökonomie unter Verwendung der internationalen ATC/DDD – Klassification., GDMS 2008. 53. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS) Düsseldorf: German Medical Science GMS Publishing House (2008) Available from: http://www.egms.de/en/meetings/gmds2008/08gmds100.shtml.

R. Schuster, E. von Arnstedt, Medizinisch–technischer Fortschritt und demografischer Wandel bei den GKV–Arzneimittelausgaben in Vertragsärztlichen Bereich, GDMS 2011. 56. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS), 6. Jahrestagung der Deutschen Gesellschaft für Epidemiologie (DGEpi) Düsseldorf: German Medical Science GMS Publishing House (2011) doi: 10.3205/11gmds007.

R. Schuster, M. Schuster, Graphentheoretische Analyse von Vernutzungsstrukturen im vertragsärztlichen Sektor einer Region der kassenärztlichen Vereinigung, GDMS 2015. 60. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS) Düsseldorf: German Medical Science GMS Publishing House (2015) doi: 10.3205/15gmds077.

G.A.F. Seber, C.J. Wild, Nonlinear regression, Wiley, New York, 1989.

D. Vakratsas, C. Kolsarici, A dual–market diffusion model for a new prescription pharmaceutical, International Journal of Research in Marketing 25 (2008) 282–293.

Acknowledgements

The authors want to express their gratitude to the colleagues invited to the "Workshop Predictive Modelling", AOK-WIdO, Berlin, 02/03 February 2016. Many stimulating discussions helped us to reach deep understanding of the open problems related to our topic, and we benefitted from fruitful suggestions for improving the basic GGM. In particular, we would like to thank Gisbert W. Selke, Cinzia Mortarino, Marion Bennie, Neil Hawkins, Ralf Neubauer, Björn Wettermark, Roberta Joppi, Irene Langner, Claudio Jommi and Thomas Car.

Working Paper Series
Department of Statistical Sciences, University of Padua

You may order paper copies of the working papers by emailing wp@stat.unipd.it
Most of the working papers can also be found at the following url: http://wp.stat.unipd.it

UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE STATISTICHE

Department of Statistical Sciences
University of Padua
Italy