LIVES WORKING PAPER 2012 / 16

TITLE

Time Dependency in Diffusion Models: Gamma-Diffusion Models as an Alternative to the Hernes Model

Author

Jean-Marie Le Goff

Research Paper

FNSNF
Swiss National Science Foundation

The National Centres of Competence in Research (NCCR) are a research instrument of the Swiss National Science Foundation.

LIVES
Swiss National Centre of Competence in Research

Author

Le Goff, J.-M.

Abstract

The marriage process in a cohort is sometimes analyzed with the Hernes model. This model presents a convenient property that clearly allows one to distinguish the quantum effect in the cohort from the tempo effect. Hernes (1972) formulated originally his model in terms of diffusion or contagion of marriage from people who were already married to those not yet married. The spread of the proportion of married people in the cohort has an increasing effect on the risk of marriage for singles. However, this increasing effect is slowed down due to facts that unmarried people progressively cease to be attractive on the marriage market when aging and that because the number of potential partners decrease. Hernes mentions that this decreasing force could be related to another mechanism in which each individual in the cohort could be heterogeneous in his or her susceptibility to marriage. This assertion, however, does not anymore correspond to his model, as it is formalized. This paper first develops models corresponding to this mechanism, the gamma-logistic and gamma-mixed influence diffusion models in which heterogeneity in susceptibility is hypothesized to be distributed as a gamma function. The second part is devoted to developing an application of the gamma-mixed-influence model on data from the Wisconsin Longitudinal Study and to comparing results with those obtained by the estimation of the Hernes model.

Keywords

Diffusion models | Hernes model | Gamma-diffusion models | Event history analysis.

Author's affiliation

NCCR LIVES, LINES, Université de Lausanne.

Correspondence to

Jean-Marie.Legoff@unil.ch

A first version of this paper has been presented at the European Population Conference in Vienna in September 2010. Some earlier results were presented at the IMA-FORS-MISC seminar on methods at the University of Lausanne in April 2009. I wish to thank Joshua R. Goldstein for his advises and comments.

1. Introduction

The Hernes model is classically used to analyze marriage across cohorts (Diekmann, 1989; Goldstein & Kenney, 2001; Hernes, 1972). This model presents a property to clearly distinguish the quantum effect in the cohort (i.e., the proportion of married people) from the tempo effect (i.e. the timing of marriage) (Billari & Toulemon, 2006). Unlike other models used to analyze the marriage processes—for example, the Coale-McNeil model—the fraction of persons in the cohort that remains unmarried is not considered a subpopulation predetermined from the starting point process not to wed (Coale & McNeil, 1972). The Hernes model was originally estimated on aggregated data. It has been incorporated in the corpus of event history analysis models and can be estimated on individual life course data (Wu, 1990; Rohwer & Pötter, 2002).

In his seminal paper, Hernes (1972) formulates his model in terms of diffusion or contagion of the idea of marriage from people already married to those not yet married. Contagion is determined by a mechanism of imitation by non-married persons or by a mechanism of persuasion of married persons on non-married individuals. Whatever the mechanism, the spread in the proportion of married people in the cohort has an increasing effect on the risk of marriage. Hernes postulates that an opposing force however slows down this diffusion process because unmarried people progressively cease to be attractive on the marriage market as they age or because the number of non-married persons decreases. As a consequence, there is a decreasing effect on the hazard of marriage. The overall hazard of marriage then results in the following two components: an increasing component in relation to the diffusion of marriage in the cohort and a decreasing element as a consequence of the depreciation of marriageability. Later, Diekmann (1989) proposed the log-logistic model as an alternative to the Hernes model, with a similar opposing mechanism of diffusion and depreciation of marriageability.

In his 1972 paper, Hernes mentions that the decreasing force can result from another mechanism; aging people who remained unmarried are those, for example, who have never held a prestigious job and, for this reason, are unattractive on the marriage market. The point here is about individual heterogeneity. In another and more precise paper, Hernes (1976) uses the term *structural heterogeneity*, which is defined as "when a capacity is differentially distributed in the population" (p. 428). The mechanism underlying the distribution of the risk of marriage for the cohort is different from the precedent described one but is similar to those described with notions of unobserved heterogeneity or of frailty in unemployment studies and mortality studies, respectively (Heckman & Singer, 1982; Vaupel, Manton, & Stallard, 1979). In the present case, those in the cohort who have a better ability to marry will wed earlier, and the weight of people with unfavorable capacities in the unmarried sub-population will become progressively higher and higher as time goes on. As a consequence, the differential ability results in a negative effect on the risk to marry at the level of the population, which slows the increasing effect due to the mechanism of diffusion. A similar effect could be described if each individual in the cohort differs by his or her own “susceptibility” to adopt the behavior when in contact with someone already married who transmits the idea of marriage (Strang & Tuma, 1993).

As in unemployment or mortality studies, differential ability or susceptibility may be due to unobserved characteristics of an individual. This unobserved heterogeneity then has to be incorporated in the models. In this paper, we propose two diffusion models that introduce an unobserved ability or susceptibility of persons to adopt behavior or an innovation that can be estimated on individual retrospective data. In the first model, the *gamma-logistic model*, the diffusion mechanism is described by the classical logistic curve, while the unobserved ability or susceptibility of individuals is patterned by a gamma distribution. In the second model, the *gamma-mixed influence diffusion model,* the unobserved heterogeneity is previously patterned by a gamma distribution, while there are two kinds of influence (Coleman, 1964; Bass, 1969): the first is due to internal influence—that is, influence of persons who already adopted the behavior on those who have not; the second is due to external influences such as media, advertising, and institutions that diffuse norms about marriage.

In the first section of the paper, after a recall of the Hernes model specification, we present the gamma-logistic and the gamma-mixed influence models. In the second section, we apply the second of these models to the case of marriage of men and women interviewed in the Wisconsin Longitudinal Study (Hauser, 2009). We are especially interested in comparing the fit of the gamma-mixed influence model with the fit of the Hernes model.

2. Time Dependency in Diffusion Models

2.1 The general Hernes model

This section discusses a general family of diffusion models proposed by Hernes in his second influential paper on diffusion models (Hernes, 1976). This general formulation is interesting because all models evoked in the introduction of the present paper—the Hernes, log-logistic model, logistic, and mixed influence models—are particular forms of this general formulation. This model could be formalized for each case where there is a diffusion process, such as the propagation of an innovation, a behavior, a rumor, and, eventually, a contagious illness. However, in particular, we have in mind the diffusion of marriage in a cohort or a group of persons. The model can be read as a general mixed-influence diffusion model (Mahajan & Peterson, 1985) in which transmission coefficients associated with external and internal influences vary with time. As originally formulated by Hernes (1976, p 434), this model does not include unobserved susceptibility or the ability to adopt the behavior, but we will include it later. The model assumes that there are no social barriers between groups of unmarried and married people. Let $F(t)$, the cumulative proportion of persons already married between t_0 and t; $f(t)$, the probability density to get married, that is the derivative of *F(t)*; *S(t)* the complementary of *F(t)*, i.e., the proportion of people who are not married at time *t*:

$$f(t) = p(t)S(t) + q(t)F(t)S(t) \qquad (1)$$

The product *F(t)S(t)* represents the probability for two persons, one unmarried the other married, to interact. *q(t)* is the rate of diffusion or contagion at time *t*, given that an unmarried person is in contact with a married person during a unit of time. In this general formulation, this rate is a function of time. In the case of marriage, the coefficient of internal diffusion *q(t)* can be understood as the rate for a single person to get married, given he or she receives information about marriage from an already married person in the cohort. The level of this rate at time *t* can depend on several elements related to the predisposition of the person to marriage and to his or her position on the marriage market. As formulated in this model, each person already married, even for a long time, is considered to have a potential influence on an unmarried person, and this influence is equal among all already married people. *p(t),* also a function of time, is the rate of adoption of the behavior due to external influences such as media or norms on marriage. Strang & Tuma (1993) suggest another interpretation in which *p(t)* is no longer related to external influence but to the effect of individual endogenous characteristics on the behavior adoption rate.

This model can be rewritten as a hazard rate function instead of a probability density function. If *h(t)* symbolizes this hazard rate, as *h(t)=f(t)/S(t)*:

$$h(t) = p(t) + q(t)F(t) \qquad (2)$$

In the case of *q(t)=0*, adoption of diffusion depends only on the intrinsic characteristics of each individual or on external influence. In this case, *p(t)* can be shaped by one of the usual functions applied to a parametric event history model. For example, if *p(t)* is considered constant, i.e., *p(t)=p* , then an exponential model is estimated. But if *p(t)* is considered always increasing or decreasing, it can be estimated by a Weibull or a Gompertz function.

In the case of *p(t)=0*, the process of adoption of the behavior depends only on internal influences. In this case, if *q(t)* is constant, i.e., *q(t)=q*, the model corresponds to the well-known logistic growth with *q* as the coefficient of diffusion (Coleman, 1964; Griliches, 1957; Mahajan & Peterson, 1985). In his first paper on the diffusion of marriage in American cohorts, Hernes (1972) supposes that *q(t)* is a decreasing function of time because unmarried people progressively lose their ability to attract potential partners as they grow older. This depreciation of the "marriageability" of a person on the marriage market draws a force opposed to the force that results from the process of diffusion. An alternative meaning of this decreasing force on the hazard rate of marriage is proposed by Diekmann (1989), who argues that it corresponds to a process of isolation of single persons as they get older due to the rarefaction of potential partners on the marriage market. A third mechanism can be related to the progressive habituation of a person to remain single and to have one's one way of doing things incompatible with a life with a partner[i]. The formulation proposed by Hernes for *q(t)* is:

$$q(t) = Ab^t \qquad (3)$$

A is the initial average of "marriageability" (*A>0*), while *b* is the constant of deterioration of this ability (*0<b<1*). The Hernes model owns the property to be defective—the cumulated proportion *F(t)* of married persons does not necessary reach 1 at the end of the marriage process. A fraction of persons in a cohort is excluded from marriage due to the negative force on the marriage hazard rate becoming higher in absolute value than the positive force due to the increase of already married persons. It is to worth noting that a difference between the number of men and women would not correspond to Hernes hypothesis of a progressive negative force: In this case, if the interest is on the marriage of men and if there are more men than women in the population, it would result in a classic cure model in which a fraction of men would be determined to remain unmarried from the staring point of the marriage process (Box-Stephensmeier & Jones, 2004; Schmidt & Witte, 1988).

Another decreasing function with time has been proposed by Diekmann (1989), for whom the most used log-logistic model can be interpreted as a diffusion model with a decreasing coefficient of diffusion with time. In the case of the log-logistic model:

$$q(t) = \frac{b}{t} \qquad (4)$$

Where *b>0*. The log-logistic model is more parsimonious than the Hernes model as it only has two parameters to estimate: one related to the initial conditions and the second to the decrease in time. However, unlike the Hernes model, the log-logistic model is not defective, which means that everyone is considered to have adopted the innovation by the end of the process. Immunity can however be considered with the hypothesis that a fraction of persons in a cohort will never adopt the innovation (Brüederl & Diekmann, 995). This hypothesis, which corresponds also to estimate a cure model, means that according to unobserved characteristics, some people are determined from the beginning of the process to remain unmarried. Such a model could work, for example, if there is a difference between the number of men and women in the cohort. Generalizations of the Hernes and/or the log-logistic models have been proposed by Banks (1994), Braun & Engelhard (2004), Diekmann (1992) and Yamaguchi (1994). Whatever these generalizations, the principle of a diffusion process countered by a loss of abilities remains.

The classic mixed influence diffusion model corresponds to the hypothesis in which in equations (1) and (2), $p(t)$ and $q(t)$, are constant ($p(t)=p$; $q(t)=q$) (Mahajan & Peterson, 1985). This model was introduced in mathematical sociology by Coleman (1964), and it has been rather diffused in marketing research, as a consequence of the work of Bass (1969). In this discipline, the peculiarity of this model is that it is generally estimated on aggregated data on the diffusion of innovation products (Mahajan & Peterson, 1985). This model has been less estimated on individual data of adoption. Its adaptation to the corpus of parametric methods of event history analysis with a procedure of estimation based on the maximization of a likelihood equation does not present many difficulties (Bass, Jain, & Krishnan, 2000; Roberts & Lattin, 2000). This model is written:

$$f(t) = \left[p + qF(t)\right]S(t) \tag{5}$$

Figure 1 presents the density distribution $f(t)$ of the mixed influence model when p=0.01 and q=0.03, with the contribution to this density of each influence. The contribution of the external influence on density is decreasing while the contribution of internal influence first progressively increases and then decreases. The mixed-influence model has sometimes been interpreted to allow the distinction between two groups of persons in a cohort: a first fraction of persons, called innovators, adopt the behavior under external influences, while others, called imitators, adopt it through channels of internal influences (Bass, 1969). This point of view, however, has been criticized for the fact that the mixed-influence diffusion model, as specified, does not allow a clear distinction between two groups with different behaviors of adoption. As it is formalized, this model means that one person can adopt the innovation either under external influences or internal influences, the weight of each of these influences varying over time (Tanny & Derzko, 1988).

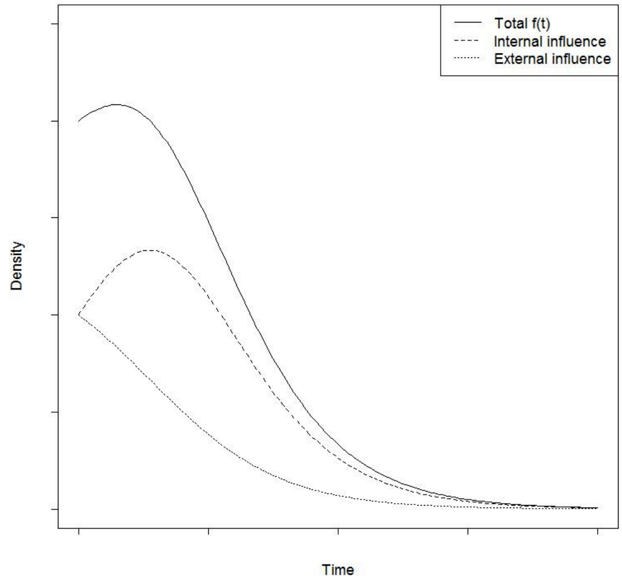

Figure 1. Density function *f(t)* of a mixed-influence diffusion model and contribution of external and internal influences on density

The mixed-influence model has been much used in marketing research in order to analyze diffusion of innovation in consumption products. We think that this model is also interesting to be applied on the analysis of demographic process such as marriage or union formation. In such cases, external influences can be interpreted as social pressure coming from persons other than peers who already adopted the behavior, for example, parents or relatives. These social pressures can also come from institutions or network channels that diffuse social norms on marriage. Moreover, the distinction between internal and external influences can be related to the distinction between the two elements of social interactions as theorized in family planning studies (Bongaarts & Watkins, 1996; Montgomery & Casterline, 1996; Kohler, 2001). The first of these elements is social learning, which corresponds to the acquisition of information about the innovation or the behavior through others. In the case of marriage, information is, for example, about others already married. Social learning can then be assimilated as internal influence. The second element of social interactions is social influence, which refers to social conformity pressures. It can be considered as external influence. The mixed influence model, as initially formulated by Coleman (1964), is not defective; it does not allow that a fraction of individuals in a cohort remain unmarried. Such a fraction can be introduced by estimating a split-population model in which, for unknown reasons or characteristics, a portion of individuals is determined to remain unmarried from the beginning of the process or if there is a difference between the number of men and women in the population. However, the introduction of general unobserved heterogeneity can also be envisioned.

2.2 *Unobserved Individual Susceptibility*

The Hernes and log-logistic models are based on the observation that for most diffusion processes, the shape of the growth curve is an asymmetric S shape, “with the upper shank of the S being more extended” (Lekvall & Wahlbin, 1973, p. 364). Such an asymmetric shape is usually observed in the case of demographic behaviors and events of the transition to the adulthood, such as living home, birth of the first child, and especially marriage. Hernes (1972) and Diekmann (1989) hypothesize that this asymmetry corresponds to a decrease in the transmission rate *q(t)* when time increases. However, an alternative hypothesis can be proposed. This hypothesis leans on notions of unobserved heterogeneity or frailty as it is developed in the analysis of unemployment and mortality (Aalen, Borgan, & Gjessing, 2008; Heckman & Singer, 1982; Vaupel et al., 1979). In the domain of mortality, frailty models assume that the general shape of the hazard rate is the same for each individual of a population but that each individual is characterized by his or her own frailty, which remains invariant as time passes (Vaupel et al., 1979). The hazard rate of death for an individual corresponds to the product of the general shape of mortality hazard and individual frailty. Such a model means that at the beginning of the process, most frail people die while less frail people take more and more weight in the population to survive. As a consequence, the hazard rate of decease at the level of the population is decreasing.

By analogy with frailty models, the diffusion processes of an innovation or behavior can be decomposed into the following two elements: first, the general shape of the diffusion, which could be called—following Strang & Tuma (1993)—the *infectiousness* from those who have adopted the behavior or, preferably, the transmissibility of the innovation to persons who have not yet adopted the behavior; second, the susceptibility or the ability of an individual to adopt the innovation or the behavior. The susceptibility is the equivalent of the frailty or the unobserved heterogeneity with the common property to be unobserved. This unobserved susceptibility can be related to the ability of a person to be in contact with persons who have already adopted the innovation or to a person’s ability to accept an innovation or a behavior when there is such a contact. By hypothesis, it remains invariant as time goes on. As in the case of frailty models, we suppose a proportional effect of individual susceptibility to the risk of adoption of the behavior. A first model incorporating individual susceptibility that can be envisaged is the simplest model of logistic growth:

$$h_i(t \mid u_i) = u_i\left[qF(t)\right] \qquad (6)$$

Where u_i represents the individual unobserved susceptibility to adopt the innovation and $h_i(t \mid u_i)$ represents the hazard rate for an individual given that he has susceptibility u_i to adopt. q expresses the coefficient of transmissibility from a person who already adopted it. If we suppose that u_i is distributed such that its mean is equal to 1, then q represents the coefficient of transmission of the behavior to a person with an average susceptibility of adoption.

The meaning of individual susceptibility has to be better specified. It is important to underline that this susceptibility is related to a person and not to possible transmitters and their "infectiousness" (Strang & Tuma, 1993). As in the Hernes or log-logistic models, the model expressed in Eq. (6) supposes that everyone who adopted the behavior has the same infectiousness, whatever the moment of the adoption or the social proximity to potential adopters. Susceptibility can be related to two series of factors: The first factor is related to the level of contact with others and, more generally, to the openness to receive information or, referring to social interactions approach (Bongaarts & Watkins, 1996), to learn from others. The second factor is related to the probability that someone adopts the innovation after he/she has acquired information about it. For example, in the marketing research tradition, susceptibility is related to the ability of a person to purchase a given product (Jeuland, 1981, qtd. in Mahajan & Peterson, 1985; Roberts & Lattin, 2000). This suggests that susceptibility depends on the properties of the innovation, more generally of the object of diffusion. Individual susceptibility is specific to the innovation and can be different according to what is diffused. It then can be related to the openness toward the innovation but also to the context in which the person is living. For example, if the behavior is marriage, susceptibility can depend on the degree of aversion to the marriage of the person and his/her attractiveness on the marriage market. The more a person is isolated from others or the more aversion he/she has toward the behavior—or the less the context is favorable for him/her—the lower his/her susceptibility toward adopting the behavior.

In Eq. (6), the transmissibility process follows a logistic growth. Two opposite "forces" play a role on the process of adoption of the behavior. As before, the first force is related to the increase of people who have already adopted the innovation with the effect of increasing the hazard rate. The second force is related to the differentiation of susceptibility among individuals. By analogy with frailty models, most susceptible persons will first experiment with the event. Consequently, less susceptible individuals will progressively take more and more weight in populations that did not yet experiment with the event. For some of them, susceptibility u_i can be so low that the hazard rate to adopt the behavior will approach zero over the time, which means that they will not experiment with the event. Such a model with a constant rate of transmission from people who already adopted the behavior to those who did not, and with individual susceptibility to adopt the behavior, can explain a S growth curve with a more extended upper shank as well as the Hernes or the log-logistic models.

An individual susceptibility element can also be added to the classic mixed influence model (Jeuland, 1981, qtd. in Mahajan & Peterson, 1985). The hypothesis here is that an individual's susceptibility to adopt is similar whether under external or internal influence.

$$h_i(t \mid u_i) = u_i\left[p + qF(t)\right] \tag{7}$$

In this case, *p* and *q* represent the average susceptibility of adoption under each of the influences. As before, the model displays two opposing forces—one related to the increase of the force of adoption due to increase of persons that already have adopted the behavior, while the second is related to the increase of the weight of those less susceptible to adopt in the population. Finally, through similarity with the general model expressed in Eq. (1) and (2), we can write a general model in which external and internal diffusion coefficients are expressions of time and into which is introduced unobserved susceptibility:

$$h_i(t \mid u_i) = u_i\left[p(t) + q(t)F(t)\right] \tag{8}$$

2.3 *From Individual Hazard to Population Hazard*

Equations (6) to (8) are expression of an individual hazard rate that depends on the adoption susceptibility of *i*. Following our analogy between the susceptibility to adopt an innovation or behavior and the frailty in mortality research or unobserved heterogeneity in unemployment studies, we now assume that u_i is gamma distributed with a mean and a variance equal to 1 and κ, respectively. With the assumption of a gamma distributed unobserved heterogeneity model, it has been shown that whatever the shape of the underlying or basic hazard rate (Aalen et al., 2008):

$$h(t) = \frac{\alpha(t)}{1 + \kappa C(t)} \tag{9}$$

and:

$$S(t) = \left(1 + \kappa C(t)\right)^{-\frac{1}{\kappa}} \tag{10}$$

h(t) and *S(t)* represent, respectively, the hazard rate at the level of the population and the probability of not having experienced the event or the behavior at time *t,* while *α(t)* represents the basic hazard rate and *C(t)* represents the cumulated basic hazard rate from 0 to *t*:

$$C(t) = \int_0^t \alpha(u)\,du \tag{11}$$

Expressions (9) and (10) mean that:

$$h(t) = \alpha(t)\left(S(t)\right)^{\kappa} \tag{12}$$

As *S(t) = 1-F(t),* and if we consider that *α(t)* is shaped by the Hernes mixed-influence diffusion model as expressed in Eq. (2)*,* i.e., *α(t) = [p(t)+q(t)]F(t)*, then:

$$h(t) = \left[p(t) + q(t)F(t)\right]\left(1 - F(t)\right)^{\kappa} \tag{13}$$

The density of adoption then becomes:

$$f(t) = \left[p(t) + q(t)F(t)\right]\left(1 - F(t)\right)^{1+\kappa} \tag{14}$$

As the Hernes (1976) mixed-influence diffusion model is the more general expression we have, this property of diffusion models with a gamma-distributed susceptibility remains when *q* is a constant and when *p* is equal to 0 (logistic diffusion model) or is constant (classic mixed-influence diffusion model).

3. Application to Wisconsin Longitudinal Study Data

In this section, we wish to compare the gamma-mixed-influence model with parameters *p* and *q* constant with the Hernes model. We estimate these models on marriage behaviors of people interviewed in the Wisconsin Longitudinal Study Sample (WLS) (2006). The first question we wish to answer is which model fits better with data. The second question is about eventual links between different parameters of the gamma mixed-influence model.

3.1 Models integration and estimation

After integration (Hernes, 1972) and reparametrization (Wu, 1990), the Hernes model is specified with three parameters to be estimated:

$$F(t) = \frac{\sigma^{-1}\exp(-\beta\lambda^{t})}{1+\sigma^{-1}\exp(-\beta\lambda^{t})} \tag{15}$$

and:

$$h(t) = \frac{-\sigma\beta\lambda^{t}\log\lambda\exp(\beta\lambda)}{1+\sigma^{-1}\exp(-\beta\lambda^{t})} \tag{16}$$

Where *λ=b, β=-A/log(b)*.

σ is related to the quantum effect of marriage, *λ* to the speed of the decline of marriageability, while *β* is inversely related to the initial proportion of persons adopting marriage (Billari & Toulemon, 2006). As *λ* and *β* are positive definite and the proportion *σ* is bounded between 0 and 1 included, a better way to fulfill these conditions is to estimate the logarithm of the two first parameters and the logit of the third:

$$\begin{aligned} \lambda &= \exp(l) \\ \beta &= \exp(m) \\ \sigma &= \exp(\mathrm{s})/\left[1+\exp(\mathrm{s})\right] \end{aligned} \tag{17}$$

Parameters can be estimated by maximizing the logarithm of a likelihood equation; this equation corresponds to traditional likelihood equations for survival data (Wu, 1990).

The gamma-mixed influence model with parameters *p* and *q* constant can be integrated as a standard mixture model (Rohwer & Pötter, 2002). In this case:

$$F(t) = 1-\left[1-\kappa\log\left[\frac{\left(1+\frac{q}{p}\right)\exp\left[-(p+q)t\right]}{1+\frac{q}{p}\exp\left[-(p+q)t\right]}\right]\right]^{-\frac{1}{\kappa}} \tag{18}$$

$$h(t) = \frac{\dfrac{p+q}{1+\frac{q}{p}\exp\left[-(p+q)t\right]}}{1-\kappa\log\left[\dfrac{\left(1+\frac{q}{p}\right)\exp\left[-(p+q)t\right]}{1+\frac{q}{p}\exp\left[-(p+q)t\right]}\right]} \tag{19}$$

Where κ represents the variance of the gamma distribution of individual susceptibility. As these three parameters are positive definite, the logarithm of these parameters have to be estimated.

$$\begin{aligned} p &= \exp(a) \\ q &= \exp(b) \\ \kappa &= \exp(\mathrm{c}) \end{aligned} \qquad (20)$$

These three parameters can be estimated by the maximum likelihood method. Note that the likelihood equation (not presented here) is two times derivable which means that variances of parameters can also be estimated.

Hernes models are often estimated on birth cohort samples at a country level (Billari & Toulemon, 2006; Dieckmann, 1989; Goldstein & Keney, 2001; Hernes, 1972). It is not a problem to estimate them on such samples with the aim to isolate quantum and tempo effects. However, it seems to us to be more awkward to estimate these types of models on a whole cohort if the theoretical frame is related to the analysis of a diffusion process in a behavior. The implicit hypothesis here is that all persons of the cohort that do not adopt the behavior at time t can be in contact or have information about those who already adopted the behavior, despite spatial and social distances between persons. This hypothesis does not seem to be realistic, and it should be better to estimate models in which weights are given according to distances between persons (Hedström, 1994; Montgomery & Casterline, 1996; Palloni, 2001; Strang & Tuma, 1993) or, like sometimes proposed in epidemiology, models in which is integrated a "mixing" parameter which measures the average probability for one person to have contact with others (Garnett, 2000). Models like the Hernes model and those we propose with the theoretical frame of social diffusion can however be estimated in the case of groups of people who are living in the same micro-local place or who were socialized during a period in the same institution, for example, in a school. If the analyzed behavior, for example, marriage, generally occurs after that all members of the group left school, we can suppose that, first, all people know one each other, at least superficially, and second, that all persons can have information about others, even if there is no direct contact between them: the network of peers, i.e., people socialized in the same school, forms a channel of diffusion (Strang, 1991) in a sense that it is not necessary for someone to be directly in contact with another to learn that this other adopted the behavior. Such a hypothesis of a diffusion channel is compatible with models like the Hernes model in which the adoption of the behavior depends on the proportion of those who already adopted the behavior. It is also compatible in models like the logistic-gamma and the mixed-influence gamma models in which the adoption of a behavior does not depend on the proximity to or the infectiousness of a person that already adopted the behavior but from the ability of a person to be informed about behaviors of others or their ability to imitate them.

3.2 Wisconsin Longitudinal Study Data

Data of the Wisconsin Longitudinal Study Sample allows distinguishing people according to the high school they attended (Hauser, 2009; WLS, 2006). The Wisconsin Longitudinal Study is a panel composed of one-third of men and women who graduated from a Wisconsin high school in 1957 (N = 10,317). Several interviews were conducted between 1957 and 2005 on this sample, sometimes with their sibling or marital partner. An interview in 1975 on a subsample of 4,330 men and 4,808 women (N = 9,138) reconstitutes the marital history of each individual, especially the date of the first marriage. Despite the ancientness of data that may not reflect contemporary behaviors of union formation, the Wisconsin Longitudinal Study Sample corresponds to our desire to use data for which the theoretical frame of a diffusion process can be developed.

As there is the possibility to distinguish persons according to the school they followed[ii], we can suppose that peers that leave a school in the same year form a channel of diffusion. We estimate the Hernes model and gamma-mixed-influence model in each group in which at least thirty men and/or women were sampled in the survey with the aim to see which of these two models better fits marriage processes[iii]. It should be noted that the estimation of the diffusion model on each subsample necessitates the hypothesis that the marriage diffusion process measured with selected persons in the sample reflects ones in the entire school cohort. A first examination of data shows that marriages are rare before June 1957 but start to increase at this date, especially in the case of women. We suppose that this is because most youth left school after their graduate degree. We decided to consider May 1957 as the starting time (t_0) of the marriage process and discounted all persons who married before this date. Size and numbers of censored persons for each school are indicated in result tables in the annex.

4. Results

Models were first estimated with the TDA version 6.4 software, especially with the use of the *frml* command, which allows programming likelihood functions for event history models[iv] (Rohwer & Pötter, 2002). However, we used the function *mle* (maximum likelihood estimation) of the library *stats4* in the R package for definitive estimations (Venable & Ripley, 2002). In the case of the Hernes and the Gamma-mixed influence model, parameters were estimated with a quasi Newton-Raphson method of optimization (method BFGCS). Estimates of both models according to gender and schools are reported in tables in the appendix (tables a1 to a4). Note that in some cases, estimations gave inaccurate results, in the sense that a parameter was estimated to be very low with a large variance. This is the case of 3 schools in 41 with women and 2 in 30 with men. We removed these schools in our further comments.

There are no formal possibilities to compare the Hernes and the Gamma- mixed influence models with usual criterions like, for example, the likelihood ratio test, since models are very different while they share the same number of parameters. However, if we nevertheless compare the maximum logarithm of likelihood obtained for each model, it is highest in the case of the gamma mixed-influence for 35 schools in 38 for women (table 1). Results are more mitigated in the case of men as estimations show that the maximum of the logarithm of the likelihood for the gamma-mixed influence model is higher than the one of the Hernes model in 19 out of 28 schools only. The estimated cumulative function of marriage of each model fit with a non-parametric Kaplan Meier estimation in the case of the school 1 (55 men and 62 women), as shown in figure 2. Similar patterns are found in other schools for both men and women. The hypothesis of the diffusion process of marriage with two mixed influences and individual susceptibility then appears to be very plausible. The diffusion process thus seems to be an interesting competitor to the Hernes process in which people lose progressively their attraction on the marriage market with time. But the results could have another interpretation. Indeed, the model formulated by Hernes in 1972 was created in a context in which the idea of an individual heterogeneity was not yet introduced in survival models[v]. In this perspective, the Hernes model can be then considered as a good approximation of the hypothesis that there is an individual susceptibility to adopt the behavior.

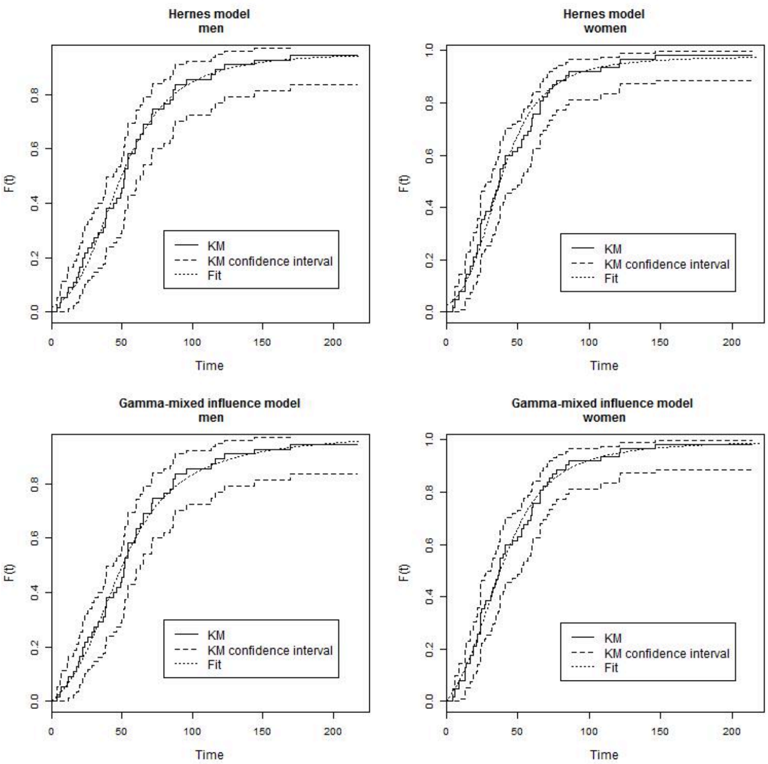

Figure 2**.** Fit of the Hernes model and the gamma mixed influence diffusion model in the case of the school 1

A synthesis of estimations in each school of gamma mixed influence models for men and women is proposed in Table 1. It is interesting to note that the *q* parameter of internal influence does not seem to differ between men and women. This result means that there is no gender difference in the internal process of diffusion. However, the *p* parameter of external influence is, on average, larger for women than for men, which means that there is greater normative pressure on women to marry quickly after leaving school than for men. If the mean of the *κ* parameter for the variance of the gamma distribution is similar between men and women, this parameter seems to be more dispersed for women than for men. The distribution of individual susceptibility appears then to be more strongly dependent on the school in the case of women.

Table 1. Synthesis of estimated parameters of gamma-mixed influence models

	Men				Women			
Parameter	Min	Max	Mean	Var(log(parameter))	Min	Max	Mean	Var(log(parameter))
κ	0.2163	2.0375	0.6960	0.3195	0.3835	1.8810	0.8527	0.1320
p	0.0002	0.0049	0.0018	0.5716	0.0010	0.0251	0.0069	0.5856
q	0.0385	0.1249	0.0611	0.0803	0.0377	0.2013	0.0668	0.1226

Q-plots of estimated parameters show a negative correlation between the *p* and the *q* parameter, which is especially strong in the case of men (Figures 3a and 3b). The higher is internal influence and the lower is external influence. Such a result could indicate that there is a kind of competition between the two influences or that the lack of one influence, for example, external influence, is compensated by a surplus of the other influence, for example, internal influence. There is no strong correlation between parameters of external influence and dispersion of individual susceptibility. However there is a positive correlation between the external influence and dispersion parameters. Higher is the external influence in a school and higher is the dispersion of the individual susceptibility.

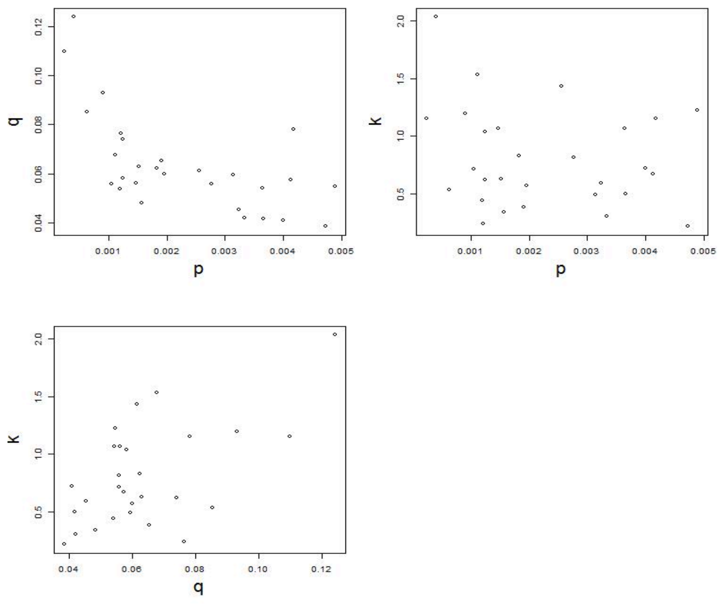

Figure 3a. q-plots of parameters of gamma mixed-influence models-men

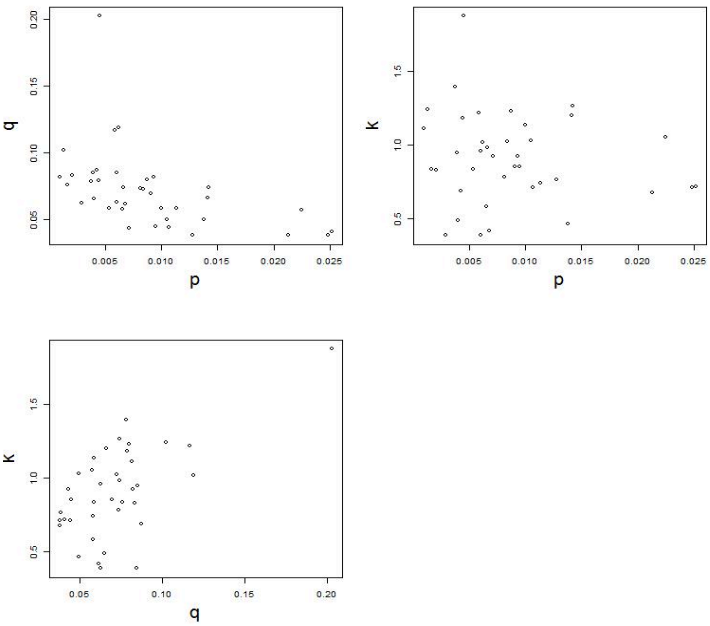

Figure 3b. q-plots of parameters of gamma mixed-influence models-women

5. Conclusion

The Hernes model postulates that a diffusion effect is slowed down by the decrease of the marriageability of persons as time goes on. The alternative models we propose are based on another suggestion of Hernes (1972, 1976) in which the diffusion effect is progressively counterbalanced by heterogeneity in the susceptibility of persons to adopt the behavior. Persons with higher susceptibility have a higher risk to adopt, and those who have a lower risk take more and more weight in the population of unmarried persons as time passes. The models are similar to those with a gamma frailty in mortality studies. Substantially, the gamma-mixed-influence model also presents the great potential to allow understanding processes of marriage and other demographic behaviors. As it makes the distinction between an external and an internal influence, it does not bind the diffusion process only to the influence effect of the cohort or the group. Moreover, it introduced the idea that susceptibility to be influenced is not the same for all persons.

The estimation of the gamma-mixed influence model on the Wisconsin longitudinal data on marriage shows that this model seems to fit a little better than the classical Hernes model. But, in our opinion, both models approximate the data well. The gamma mixed influence can be further developed by taking into account individual characteristics of persons that can have an influence on each of the two influence parameters as well as on the individual susceptibility variability. There are potentially no difficulties in adding fixed covariates. Moreover, we presented a generalization of the model that can be a basis for investigations in order to introduce time dependant covariates and to allow the different parameters of internal and external influences to be time dependent.

Notes

[i] Joshua R. Goldstein, personal communication.
[ii] Each individuals surveyed was attributed an ID number in which the first three digits was a code for the school. Note that we do not have information about the school in standards data files (WLS, 2006).
[iii] We chose a limit of 30 because models cannot be estimated if a sample is under this limit.
[iv] Note that an example for the estimation of the Hernes model is proposed in the user's manual (Rohwer & Pötter, 2002).
[v] It will be later by Manton, Vaupel & Stallard (1979).

References

Aalen O., Borgan, O., & Gjessing, H. K. (2008). *Survival analysis and event history analysis. A pro cess point of view*. New York: Springer.

Banks, R. B. (1994). *Growth and diffusion phenomena. Mathematical frameworks and applications*. Berlin: Springer Verlag.

Bass, F. M. (1969). A new-product growth model for consumer durables. *Management Sciences*, *15*, 215−227.

Bass, F. M., Jain, D., & Krishnan, T. (2000). Modeling the marketing-mix influence in new-product diffusion. In Mahajan Vihay, Muller Eitan, & Wind Yoram (Eds), *New-product diffusion model* (99−122). New York: Springer.

Billari, F., & Toulemon, L. (2006). *Cohort childlessness forecasts and analysis using the Hernes model*. Paper presented at the European Population Conference 2006. Liverpool, UK.

Bongaarts, J., & Watkins, S. C. (1996). Social Interactions and Contemporary Fertility. Transitions. *Population and Development Review, 22*(4), 639-682.

Box-Stephensmeier, J. M., & Jones, B. S. (2004). *Event history modeling. A guide for social scientists*. Cambridge: Cambridge University Press.

Braun, N., & Engelhard, H. (2004). Diffusion processes and event history analysis. *Vienna Yearbook of Population Research*, 2, 111−131.

Brüederl, J., & Diekmann, A. (1995). The log-logistic model. Two generalizations with an application to demographic data. *Sociological Methods and Research, 24*(2), 158−186.

Coale, A. J., & McNeil, D. R. (1972). The distribution by age of the frequency of first marriage in female cohort. *Journal of American Statistical Association, 67*, 743−749.

Coleman, J. S. (1964). *Introduction to mathematical sociology*. London: MacMillan, The Free Press of Glencoe.

Diekmann, A. (1989). Diffusion and survival models for the process of entry into marriage. *Journal of Mathematical Sociology*, *14*, 31−44.

Diekmann, A. (1992). The log-logistic distribution as a Model for Social Diffusion processes. *Journal of Scientific & Industrial Research*, 51, 285-290.

Garnett, GP. (2002). An introduction to mathematical models in sexually transmitted disease epidemiology. *Sexually Transmitted Infections*. 78: 7-12.

Goldstein, J. R., & Kenney, C. T. (2001). Marriage delayed or marriage forgone? New cohort forecasts of first marriage for US women. *American Sociological Review, 66*, 506–519.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica, 25*(4), 501–522.

Hauser, R.M. (2009). The Wisconsin Longitudinal Study: Designing a Study of the Life Course. in Elder Glen H. & Giele Janet Z. (Eds). *The Craft of Life Course Research*. New York: The Guilford Press.

Heckman, J., & Singer, B. (1982). *Population heterogeneity in demographic models*. Orlando, FL: Academic Press.

Hedström, P. (1994). Contagious collectivities: On the spatial diffusion of Swedish trade unions 1890-1940. *American Journal of Sociology*. *99*(5), 1157-1179.

Hernes, G. (1972). The process of entry into first marriage. *American Sociological Review*, *37*(2), 173−182.

Hernes, G. (1976). Diffusion and growth. The non-homogeneous case. *Scandinavian Journal of Economics*, *78*, 427−436.

Jeuland, A. (1981). *Parsimonious models of diffusion of innovation: Derivation and comparisons*. Working Paper, Marketing Department. Chicago: Graduate School of Business, University of Chicago.

Kohler, H.-P. (2001). *Fertility and social interaction: An economic perspective*. Oxford: Oxford University Press.

Lekvall, P., & Wahlbin, C. (1973). A study of some assumptions underlying innovation diffusion functions. *Swedish Journal of Economics*, *75*, 326−377.

Mahajan, V., & Peterson, R. A. (1985). *Models for innovation diffusion*. Newbury Park, CA: Sage.

Montgomery, M.R., & Casterline, J.B. (1996). Social learning, social influence, and new models of fertility. *Population and Development Review*, 22, 151-175.

Palloni, A. (2001). Diffusion in Sociological Analysis. In Casterline J.B. (Eds), *Diffusion Processes and Fertility Transition. Selected Perspectives* (66-114). Washington D.C.: National Academic Press.

Roberts, J., & Lattin, J. (2000). Disaggregate-level diffusion models. In Mahajan Vihay, Muller Eitan, & Wind Yoram (Eds), *New-product diffusion model* (207−236). New York: Springer.

Rohwer, G., & Pötter, U. (2002). *TDA user's manual.* Bochum: Ruhr Universität Bochum. Retrieved from http://www.stat.ruhr-uni-bochum.de/tda.html.

Schmidt, P., & Witte, A. D. (1988). Predicting criminal recidivism using split population survival time models. *Journal of Econometrics*, *40*, 141−159.

Strang, D. (1991). Adding social structure to diffusion models: An event history framework. *Sociological Methods and Research*. *19*(3), 324-353.

Strang, D., & Tuma, N. B. (1993). Spatial and temporal heterogeneity in diffusion. *American Journal of Sociology*, *99*(3), 614−639.

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

Vaupel, J. W., Manton, K., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, *16*, 439−454.

Venable, W. N, & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer.

WLS. (2006). *Wisconsin Longitudinal Study handbook. Tracking the life course*. Madison: University of Wisconsin. Retrieved from http://www.ssc.wisc.edu/wlsresearch/documentation/handbook/WLS_Handbook.pdf

Wu, L. L. (1990). Simple graphical goodness-of-fit tests for hazard rate. In Mayer Karl Ulrich & Tuma Nancy (Eds), *Event history analysis in life course research* (184−199). Madison: The University of Wisconsin Press.

Yamaguchi, K. (1994). Some accelerated failure-time regression models derived from diffusion process models. In Mardsen Peter (Ed), *Sociological Methodology,* 267−300. Washington D.C.: Blackwell.

Appendix

Table a1. Results of Hernes models-Men

		logit(σ)			log(λ)			log(β)			
School	N	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Log-lik
1	55	-2.9623	0.4864	18.0422	4.0622	0.4864	328.3602	1.9351	0.0135	276.4662	-265.5252
31	42	-2.6981	0.5146	14.1471	4.1284	0.5146	269.4392	1.9460	0.0175	216.2683	-206.0673
39	31	-1.9502	0.5904	6.4420	4.4311	0.5904	182.0973	1.9050	0.0306	118.4285	-146.6958
52	34	-3.7768	1.6717	8.5328	4.2359	1.6717	151.4941	2.1850	0.0217	219.5848	-165.5805
62	39	-2.1909	0.4062	11.8164	4.1169	0.4062	290.0043	1.9580	0.0224	171.3211	-185.3116
78	44	-3.2293	0.9256	11.2670	4.3371	0.9256	202.1574	1.9768	0.0172	227.6873	-221.7318
133	34	-2.7385	0.5097	14.7120	4.0150	0.5097	242.7974	2.0769	0.0248	173.9497	-167.8940
137	59	-2.8605	0.5106	16.0248	4.2081	0.5106	294.3074	2.1747	0.0137	344.2064	-289.8286
139	54	-4.1064	1.7270	9.7640	4.4819	1.7270	176.8371	2.2458	0.0135	373.4949	-273.1263
151	62	-2.5966	0.3513	19.1931	4.2307	0.3513	369.2866	1.8651	0.0119	293.1979	-308.7135
159	52	-4.2730	1.9184	9.5174	4.2728	1.9184	169.8576	2.2224	0.0169	291.7582	-249.2428
160	33	-3.5377	1.4961	8.3652	4.1547	1.4961	140.2962	2.0721	0.0227	189.1365	-159.7110
190	33	-4.0610	2.9396	5.6102	4.4790	2.9396	108.9641	2.0922	0.0341	128.5254	-164.5614
192	53	-3.1647	0.7118	14.0709	4.1536	0.7118	267.3501	2.1278	0.0139	326.6993	-254.8241
200	34	-3.7006	1.5316	8.9417	4.3238	1.5316	164.6825	2.1126	0.0224	199.3905	-169.4546
230	41	-4.7273	4.0626	5.5008	4.5977	4.0626	115.9538	2.1105	0.0491	90.7331	-203.6846
232	63	-1.8521	0.1552	22.1030	3.9594	0.1552	502.5194	1.8283	0.0145	231.1015	-298.4852
236	49	-2.6010	0.3530	19.1651	3.7790	0.3530	287.8027	2.3990	0.0299	192.5695	-230.2719
243	34	-4.9107	5.3260	4.5278	4.7076	5.3260	99.3948	2.2483	0.0427	118.5163	-170.9186
245	36	-2.0996	0.3445	12.7975	3.9497	0.3445	255.9237	1.8431	0.0222	152.8260	-171.6982
281	31	-7.7316	30.6573	1.9499	4.7176	30.6573	41.3276	2.5409	0.1415	45.6357	-146.0805
323	52	-2.2206	0.2373	20.7807	3.7266	0.2373	397.2853	1.8427	0.0152	223.8423	-239.6428
334	50	-2.3519	0.3469	15.9460	4.1375	0.3469	305.9168	2.0516	0.0182	231.3262	-244.4407
338	34	-6.0748	15.4807	2.3839	5.0155	15.4807	55.0178	2.2132	0.1565	31.3076	-177.8495
340	48	-1.6212	0.1524	17.2413	3.4889	0.1524	418.4141	2.2005	0.0331	146.1286	-208.4452
372	52	-1.7966	0.1906	16.9317	3.9761	0.1906	398.4982	1.9732	0.0217	179.5456	-244.4176
374	42	-4.5064	2.4657	8.2358	4.2610	2.4657	131.1492	2.3711	0.0163	344.1506	-203.9134
389	39	-2.6299	0.5426	12.7460	4.0492	0.5426	245.5586	2.0342	0.0198	209.0440	-187.5621
422	59	-2.8645	0.5148	15.9383	4.2170	0.5148	309.4356	1.9534	0.0126	303.7945	-289.5065
432	44	-2.1076	0.2582	17.2034	4.0062	0.2582	346.4517	1.7420	0.0177	171.8191	-213.5950

Table a2. Results of gamma-mixed influence models-Men

		log(κ)			log(p)			log(q)			
School	N	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Log-lik
1	55	-0.3953	0.1678	0.9312	-5.4900	0.2165	139.1940	-2.8578	0.0832	98.1637	-264.3192
31	42	-0.2049	0.1914	0.2194	-5.8896	0.3342	103.8064	-2.8863	0.1009	82.5489	-205.3663
39	*31*	*-20.2257*	*10.6170*	*38.5303*	*-5.6938*	*0.1773*	*182.8204*	*-4.4789*	*0.2217*	*90.4871*	*-147.8478*
52	34	-0.4734	0.2954	0.7587	-6.6842	0.6872	65.0146	-2.6047	0.0986	68.8223	-165.2476
62	39	0.0608	0.1565	0.0236	-6.5268	0.4634	91.9324	-2.8814	0.0770	107.8159	-184.8665
78	44	-0.6976	0.5166	0.9421	-5.6096	0.2835	110.9964	-3.1781	0.1888	53.5104	-220.7413
133	34	0.1808	0.1439	0.2272	-7.0066	0.8267	59.3870	-2.3756	0.0956	59.0026	-166.5815
137	59	-0.3380	0.1769	0.6456	-6.8566	0.3126	150.4021	-2.8834	0.0539	154.1176	-290.7695
139	54	-0.8102	0.4203	1.5616	-6.7288	0.4005	113.0542	-2.9192	0.0832	102.4810	-273.4750
151	62	-0.3225	0.2220	0.4685	-5.5235	0.2020	151.0163	-3.1933	0.1421	71.7782	-306.9494
159	52	-0.9628	0.3401	2.7253	-6.2603	0.2862	136.9489	-2.7295	0.0545	136.7555	-249.2819
160	33	-0.7141	0.4678	1.0899	-5.7633	0.3856	86.1338	-2.8233	0.1347	59.1839	-159.5108
190	33	-1.1968	1.0041	1.4263	-5.7051	0.3258	99.9136	-3.1675	0.1509	66.5020	-163.8517
192	53	-0.4618	0.1827	1.1674	-6.4926	0.3145	134.0530	-2.7653	0.0528	144.9481	-255.0187
200	34	-0.5569	0.3167	0.9792	-6.2374	0.4692	82.9168	-2.8147	0.1011	78.3379	-168.8710
230	41	-1.5310	1.1132	2.1055	-5.3562	0.2317	123.8228	-3.2567	0.1353	78.3926	-201.7476
232	63	0.3583	0.0905	1.4178	-5.9709	0.2742	130.0250	-2.7906	0.0880	88.5282	-297.4653
236	49	0.1440	0.1026	0.2022	-8.3386	1.0126	68.6652	-2.2096	0.0542	90.0048	-230.9734
243	34	-1.0678	0.7133	1.5984	-6.4568	0.4620	90.2417	-3.0322	0.0992	92.7117	-170.3473
245	36	0.0652	0.2163	0.0196	-5.6170	0.3625	87.0411	-2.9153	0.1786	47.5906	-171.4998
281	31	-1.4160	1.4357	1.3966	-6.7188	0.7907	57.0919	-2.5738	0.1075	61.6247	-145.7426
323	52	0.1419	0.1080	0.1864	-5.4808	0.2947	101.9146	-2.5483	0.0990	65.5944	-239.4599
334	50	0.0370	0.1635	0.0084	-6.6837	0.4325	103.2977	-2.8439	0.0949	85.1932	-244.7426
338	*34*	*-7.5728*	*726.6166*	*0.0789*	*-5.1221*	*0.1630*	*160.9520*	*-3.8005*	*0.1292*	*111.7644*	*-174.7609*
340	48	0.7117	0.0708	7.1563	-7.8292	0.8429	72.7235	-2.0871	0.0527	82.7031	-209.0629
372	52	0.4290	0.0979	1.8797	-6.8047	0.4828	95.8998	-2.6939	0.0770	94.2339	-244.3363
374	42	-0.6194	0.3101	1.2371	-7.3694	0.7974	68.1088	-2.4621	0.0763	79.4199	-204.0764
389	39	-0.1840	0.1889	0.1793	-6.3034	0.4029	98.6256	-2.7765	0.0791	97.4621	-187.6983
422	59	-0.5256	0.2096	1.3177	-5.7331	0.1915	171.6271	-3.0935	0.0755	126.7576	-288.7924
432	44	0.2029	0.1618	0.2544	-5.3215	0.3231	87.6380	-2.9063	0.2095	40.3129	-211.3404

Note: In italics problems in estimation

Table a3. Results of Hernes models-Women

		logit(σ)			log(λ)			log(β)			
School	N	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Log-lik
1	62	-3.7708	0.8738	16.2733	4.0057	0.8738	241.8952	2.0097	0.0142	285.1032	-292.1862
31	44	-2.8410	0.4062	19.8711	3.5283	0.4062	261.9430	1.9032	0.0159	228.1106	-197.5023
47	34	-4.1348	2.2456	7.6133	4.5423	2.2456	131.6321	1.9504	0.0381	99.7981	-178.2180
52	46	-1.9974	0.2151	18.5438	3.4278	0.2151	317.2360	1.6519	0.0163	167.6227	-201.2647
62	51	-2.9315	0.4270	20.1265	3.9302	0.4270	291.1080	1.6612	0.0169	163.4715	-247.3624
78	54	-2.8034	0.3557	22.0968	3.8534	0.3557	368.0943	1.9079	0.0136	266.8936	-255.6021
81	30	-4.6472	3.3454	6.4556	4.0018	3.3454	94.7500	2.1315	0.0393	115.7189	-137.2257
108	34	-3.6071	1.4204	9.1607	4.0608	1.4204	155.8016	2.0483	0.0239	175.8502	-160.8620
133	53	-2.2146	0.2129	23.0383	3.7575	0.2129	397.0877	1.6912	0.0137	208.2761	-249.3336
134	30	-3.0474	0.6992	13.2828	3.9219	0.6992	194.9441	1.8185	0.0242	136.5619	-146.3294
137	53	-2.2956	0.2451	21.5055	3.6665	0.2451	407.0047	2.0694	0.0174	246.2372	-240.1882
151	67	-2.3721	0.1878	29.9669	3.4286	0.1878	439.0833	1.8492	0.0112	304.1220	-296.1258
153	39	-2.4629	0.3238	18.7309	3.5392	0.3238	285.2757	1.6425	0.0188	143.5372	-176.2331
159	64	-3.1498	0.4840	20.4993	3.9742	0.4840	359.7727	1.9050	0.0128	284.2046	-301.3927
160	42	-3.2386	0.5640	18.5979	3.9801	0.5640	271.9583	1.8199	0.0178	185.8351	-206.9863
163	31	-2.2503	0.3889	13.0204	3.8146	0.3889	254.2428	1.6799	0.0244	115.5470	-145.8473
192	46	-2.9309	0.4483	19.1618	3.7678	0.4483	293.8123	1.8578	0.0157	219.9285	-214.8226
200	38	-3.1188	0.7664	12.6917	3.7549	0.7664	221.7116	1.9888	0.0207	190.8099	-169.8220
201	37	-2.7329	0.4707	15.8689	3.6266	0.4707	234.7586	1.6839	0.0213	133.0239	-168.4411
230	41	-2.5920	0.4393	15.2946	3.9306	0.4393	283.4075	1.8455	0.0181	188.0616	-195.0364
232	66	-2.4672	0.2221	27.4045	3.7131	0.2221	479.0285	1.8132	0.0113	291.7855	-304.3171
236	77	-2.6265	0.2334	29.5614	3.8889	0.2334	506.6553	1.7540	0.0098	315.0224	-366.1275
243	41	-7.5265	32.2251	1.7579	4.6419	32.2251	35.8345	2.3301	0.2800	19.3892	-189.5062
245	49	-2.1364	0.2113	21.6043	3.4773	0.2113	361.8309	1.6565	0.0153	179.8072	-217.6534
250	42	-2.2247	0.3082	16.0579	3.9662	0.3082	295.9784	1.7361	0.0174	173.3025	-203.3372
255	35	-2.2594	0.3224	15.8360	3.6456	0.3224	253.3644	1.7316	0.0208	144.0652	-162.2680
268	33	-2.5528	0.4972	13.1068	3.5027	0.4972	183.6890	1.6655	0.0236	117.3558	-146.2536
276	49	-2.1231	0.2307	19.5427	3.6375	0.2307	363.5849	1.8541	0.0162	212.2346	-222.3382
281	31	-2.5921	0.5134	13.0872	3.5729	0.5134	201.6333	1.8067	0.0235	138.8726	-139.1441
294	35	-2.3420	0.3976	13.7939	3.7847	0.3976	249.4800	1.8121	0.0210	156.4820	-163.1392
323	51	-2.4433	0.3019	19.7727	3.8573	0.3019	357.5547	1.6932	0.0150	191.3463	-240.8719
338	36	-2.2147	0.2808	17.4669	3.3873	0.2808	228.6359	1.7620	0.0207	149.7724	-160.9695
340	69	-1.9196	0.1396	26.3927	3.6716	0.1396	596.0396	1.8065	0.0124	262.2413	-311.5339
365	33	-2.3413	0.4392	12.4797	3.8446	0.4392	234.0941	2.1453	0.0305	150.6543	-154.7239
372	63	-2.8469	0.3549	22.8402	3.8553	0.3549	424.0568	2.0497	0.0121	347.9914	-291.6180
374	39	-4.4592	2.6750	7.4336	4.3343	2.6750	148.1226	2.1503	0.0318	145.2549	-186.5377
389	45	-2.3054	0.2915	18.2317	3.5997	0.2915	327.9196	1.7665	0.0167	186.2961	-201.8213
409	32	-3.4681	1.1734	10.2501	4.2744	1.1734	177.6906	2.0493	0.0230	182.8208	-161.0443
419	35	-2.3046	0.3911	13.5792	3.8475	0.3911	243.7828	1.7212	0.0210	141.1098	-165.7378
422	59	-2.7287	0.2950	25.2381	3.6853	0.2950	364.7411	1.6975	0.0129	223.9121	-273.2600
432	52	-3.2764	0.6880	15.6026	3.8300	0.6880	251.0206	1.8412	0.0189	179.1713	-236.6495

Table a4. Results of gamma-mixed influence models-Women

		log(κ)			log(p)			log(q)			
School	N	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Estimate	Variance	Wald test	Log-lik
1	62	-0.8845	0.3177	2.4632	-4.9957	0.1577	158.2325	-2.7867	0.1072	72.4209	-290.2438
31	44	0.0197	0.1418	0.0027	-5.0880	0.3333	77.6692	-2.1310	0.1321	34.3691	-196.4451
47	*34*	*-7.0698*	*962.3678*	*0.0519*	*-4.5381*	*0.1285*	*160.2145*	*-4.4463*	*0.5694*	*34.7211*	*-174.2028*
52	46	0.1851	0.1246	0.2748	-4.2590	0.1942	93.3971	-2.7142	0.2104	35.0180	-200.5035
62	*51*	*-1.0810*	*0.6989*	*1.6719*	*-3.5539*	*0.0710*	*177.8649*	*-7.3904*	*513.2809*	*0.1064*	*-236.9257*
78	54	-0.0526	0.1176	0.0235	-5.5401	0.3191	96.1769	-2.4626	0.0948	63.9377	-253.8089
81	30	-0.9572	0.6249	1.4661	-5.1143	0.4204	62.2110	-2.4689	0.1722	35.4049	-136.0597
108	34	-0.7221	0.3361	1.5512	-5.5291	0.3243	94.2771	-2.7312	0.0962	77.5546	-160.4010
133	53	0.1266	0.1473	0.1088	-4.6022	0.1912	110.7860	-2.8410	0.2579	31.2935	-245.9989
134	30	-0.1612	0.3302	0.0787	-4.7051	0.4693	47.1680	-2.6672	0.4544	15.6553	-143.5792
137	53	0.2148	0.0874	0.5279	-6.6451	0.4362	101.2391	-2.2810	0.0533	97.6639	-240.9640
151	67	0.1954	0.0719	0.5310	-5.1394	0.2238	118.0187	-2.1484	0.0700	65.9007	-296.1056
153	39	0.0535	0.1772	0.0162	-3.7947	0.2082	69.1548	-2.8612	0.5090	16.0849	-171.7041
159	64	-0.5448	0.1599	1.8565	-5.0255	0.1554	162.5295	-2.8526	0.0833	97.6432	-299.2432
160	42	-0.2983	0.2854	0.3118	-4.4765	0.2618	76.5459	-2.8455	0.4022	20.1303	-202.2420
163	31	0.0303	0.2259	0.0041	-4.5545	0.3274	63.3574	-3.0065	0.3657	24.7152	-143.6757
192	46	-0.2443	0.1617	0.3692	-4.8128	0.2235	103.6463	-2.6120	0.1278	53.4033	-212.7647
200	38	-0.3782	0.1839	0.7777	-5.4654	0.3353	89.0793	-2.4395	0.0777	76.5640	-169.7918
201	37	-0.3389	0.2950	0.3894	-3.6966	0.2037	67.0791	-3.2702	0.9419	11.3544	-164.1486
230	41	-0.1828	0.1866	0.1791	-5.2302	0.2782	98.3188	-2.8389	0.1352	59.6084	-193.9221
232	66	-0.0170	0.0931	0.0031	-5.0186	0.1822	138.2316	-2.6018	0.0860	78.7238	-302.8154
236	77	-0.3420	0.1413	0.8282	-4.5409	0.1136	181.5113	-3.1231	0.1634	59.6882	-361.3828
243	*41*	*-8.6727*	*1272.4327*	*0.0591*	*-4.3373*	*0.1369*	*137.4270*	*-3.2875*	*0.1459*	*74.0801*	*-184.5004*
245	49	0.2361	0.1151	0.4843	-4.2525	0.2047	88.3421	-2.6034	0.2283	29.6818	-215.1431
250	42	-0.0782	0.2543	0.0241	-4.9419	0.2395	101.9805	-3.1414	0.3052	32.3309	-201.4717
255	35	0.2066	0.1943	0.2197	-4.7428	0.3604	62.4152	-2.5274	0.3188	20.0363	-160.6536
268	33	-0.3367	0.3065	0.3699	-3.6827	0.2216	61.2074	-3.1966	0.8392	12.1768	-143.4745
276	49	0.1662	0.1091	0.2532	-5.4203	0.2569	114.3543	-2.5416	0.0864	74.7629	-222.9686
281	31	-0.0787	0.2091	0.0296	-4.6765	0.3388	64.5482	-2.5028	0.2071	30.2525	-138.4490
294	35	-0.0405	0.1939	0.0084	-5.1115	0.2978	87.7243	-2.7703	0.1577	48.6817	-162.6326
323	51	-0.2686	0.1989	0.3627	-4.3632	0.1653	115.1661	-3.2666	0.3149	33.8870	-237.1138
338	36	0.6320	0.1263	3.1623	-5.4025	0.9165	31.8467	-1.5944	0.2568	9.8976	-159.2516
340	69	0.3314	0.0662	1.6579	-5.5763	0.2165	143.6014	-2.5469	0.0578	112.3010	-311.4319
365	33	0.1049	0.1624	0.0678	-6.9339	0.6733	71.4101	-2.5080	0.0795	79.1630	-155.2404
372	63	-0.1905	0.0947	0.3832	-6.1790	0.2529	150.9947	-2.4888	0.0378	163.9053	-291.6132
374	39	-0.9583	0.3591	2.5578	-5.8281	0.3371	100.7696	-2.7750	0.0738	104.2819	-185.5049
389	45	0.0237	0.1262	0.0045	-4.7785	0.2316	98.5903	-2.6278	0.1211	57.0045	-201.3019
409	32	-0.1792	0.2369	0.1355	-6.4314	0.8250	50.1376	-2.5764	0.1464	45.3337	-159.7454
419	35	-0.1619	0.2824	0.0928	-4.6571	0.2593	83.6548	-3.1127	0.3595	26.9502	-164.1451
422	59	-0.3939	0.2251	0.6893	-3.8497	0.1250	118.5931	-3.2760	0.5744	18.6832	-266.8423
432	52	-0.7708	0.2493	2.3830	-4.2814	0.1553	118.0041	-3.0064	0.1904	47.4699	-233.4154

Note: In italics problems in estimation