Eindhoven Centre for Innovation Studies, The Netherlands

Working Paper 00.15

Intellectual property rights, strategic technology agreements and market structure. The case of the GSM

R. Bekkers, G. Duysters, B. Verspagen

Faculty of Technology Management

Eindhoven University of Technology, The Netherlands

December 2000

BREAKING THE WAVES: A POISSON REGRESSION APPROACH TO SCHUMPETERIAN CLUSTERING OF BASIC INNOVATIONS

by

Gerald Silverberg & Bart Verspagen*

ECIS, Eindhoven University of Technology & MERIT, Maastricht University

Third draft.

We are indebted to Bart Los and Witold Kwasnicki for helpful discussions. Only the authors are to be blamed for any controversial opinions or remaining errors, however.

Abstract:

The Schumpeterian theory of long waves has given rise to an intense debate on the existence of clusters of basic innovations. Silverberg and Lehnert have criticized the empirical part of this literature on several methodological accounts. In this paper, we propose the methodology of Poisson regression as a logical way to incorporate this criticism. We construct a new time series for basic innovations (based on previously used time series), and use this to test the hypothesis that basic innovations cluster in time. We define the concept of clustering in various precise ways before undertaking the statistical tests. The evidence we find only supports the 'weakest' of our clustering hypotheses, i.e., that the data display overdispersion. We thus conclude that the authors who have argued that a long wave in economic life is driven by clusters of basic innovations have stretched the statistical evidence to far.

JEL Codes: O3, O4, N10, C2

* corresponding author. Please use the first of the two addresses below for correspondence:
ECIS - Eindhoven Center for Innovation Studies, PO Box 513, 5600 MB Eindhoven, the Netherlands, fax +31 40 2474646
MERIT - Maastricht Economic Research Institute on Innovation and Technology, PO Box 616, 6200 MD Maastricht, the Netherlands, fax +31 43 3884905
email gerald.silverberg@merit.unimaas.nl, b.verspagen@tm.tue.nl
Web: meritbbs.unimaas.nl/staff/silverb.html, www.tm.tue.nl/ecis/bart/

1. Introduction

Schumpeter was the first economist to place the mechanism of basic innovation in the center of the discussion about the long-run development pattern of global capitalism. His emphasis on the *clustering* of basic innovations[1], while challenged at the time by, for example, Kuznets (1942), has remained a key concept in the debate about long waves and their explication. While Silverberg and Lehnert (1993), in a formal model, demonstrated that such clustering was not logically necessary to produce long-period cyclical behavior of a Schumpeterian economy, their empirical work still left some questions open about how clusters do in fact occur over time, and cast considerable doubt on the conclusions of previous researchers.

The relationship between innovation and long waves is not just a matter of idle historical curiosity. The publicity surrounding the so-called "new" economy (information and communication technology, the Internet) as an apparent golden era which seemingly repeals the constraints of the "old" economy, is a case in point. Does ICT usher in an upswing in economic activity and productivity growth that suspends the orthodox wisdom about the relationship between, for example, employment and inflation, or the fair valuation of equities, or is it only a rerun of previous historical experience with revolutionary and pervasive innovations? Ten Raa and Wolff (2000), for example, find strong productivity effects related to ICT, while Gordon (2000) is much more skeptical after a comparison of the productivity gains of ICT with those of previous technological revolutions.

The perspective of long waves also raises the question as to whether the (supposed) productivity increases related to ICT represent a permanent increment in the long-term rate of technical change, or are simply a passing cyclical phenomenon. The long wave view would hold that the current period of strong growth is the upswing of the fifth long wave, and, hence, that a downswing would inevitably follow upon it.[2] The two views are not mutually exclusive. One may imagine the long-run rate of basic innovations as a process with both a positive trend and a cyclical component. In this case, the relevant question becomes whether a model can be formulated that identifies both elements simultaneously.

After a long period of quiescence, a new debate on Schumpeter's hypothesis arose in the 1970s and 1980s. In this debate, which is summarized in some detail below, much of the empirical argument centered around time series of basic innovations. These were collected by various authors (who all obtained different results) on the basis of an historical evaluation of the importance of innovations. The main research question for these authors was whether the number

[1] "[Innovations] are not evenly distributed in time, but that on the contrary they tend to cluster, to come about in bunches, simply because first some, and then most firms follow in the wake of successful innovation" (Schumpeter, 1939, p. 75).

[2] In his historical analysis, Schumpeter observed three long waves since the beginning of modern capitalism. Freeman and Soete (1997), in a similar vein, observed five long waves, i.e., they argue that two new ones occurred since the end of Schumpeter's analysis.

of basic innovations was higher in depression periods than in prosperity periods (a more specific hypothesis than Schumpeter put forward, since he did not relate clustering to specific phases of the economic cycle). The possibility of a (very) long-run trend in the rate of basic innovation did not play any explicit role in this debate.

Silverberg and Lehnert (1993) argue that the statistical methods used by the researchers in this debate were not valid due to the count nature of the underlying data. They point out that innovations are discrete events and the distribution used to describe the occurrence of innovations must take this into account. The *t*- or *z*-tests applied by Kleinknecht (1990) and Solomou (1986) cannot be applied when the data are not normally distributed. Silverberg and Lehnert instead applied nonparametric tests appropriate to the null hypothesis of a Poisson distribution.

As early as 1974, Sahal (1974) had already suggested using a statistical method based on the Poisson and negative binomial distributions to describe time series of incremental innovation. In the mean time, Poisson regression has become a standard part of the econometric toolbox (see, e.g., Green, 1995). Applying this method, the question of whether Schumpeter was right (i.e., basic innovations tend to cluster) can be analyzed simultaneously with the question of whether the long-run rate of occurrence of basic innovations is rising. This is what this paper intends to do, working with the same basic innovation databases used by authors such as Kleinknecht and Mensch in the 1970s and 1980s.

The rest of this paper is organized as follows. In Section 2, the empirical literature on basic innovation time series and long waves will be briefly reviewed. This section will also introduce the critique of this debate by Silverberg and Lehnert and their results in more detail. On the basis of these results a number of explicit statistical hypotheses are formulated to capture the ideas that have only been implicit until now in the debate on Schumpeterian basic innovations. Section 3 presents the methodology of the Poisson regression model. Section 4 implements this model in the context of several time series for basic innovations, estimating a model admitting both overdispersion (clustering) and trend components. Section 5 summarizes the main findings, and draws some conclusions.

2. Innovation time series. Review of the literature and shortcomings

The 1970s and 1980s witnessed a rekindling of the debate on the Schumpeterian hypothesis of clustering of innovations rooted in statistical analysis of real data. The first contribution to this debate was Mensch (1979), who argued strongly in favour of the Schumpeterian hypothesis that says that basic innovations tend to be clustered. The specific hypothesis he advanced was that basic innovations tend to cluster in the depression periods of the long wave. Mensch's theoretical explanation for such clustering was that only in the despair of the depression phase do firms resort to the highly risky strategy of introducing basic innovations (drawing on inventions that may have been made previously). During the upswing and upper turning phase, in contrast, firms focus on squeezing profits out of the dominant technologies, and search activities aimed at basic innovations are at a low level (whether this is perfectly rational or myopic behavior with respect to the relative costs of the various technological opportunities over the cycle remains open). His approach was to construct a database of basic innovations, and to use a runs test on the resulting time series of the number of innovations occurring each year (using year of innovation rather than year of invention) to test whether the series clustered. A runs test is a test of independence on the assumption of identical distribution, and thus does not control for a trend. On the other hand, it does not require the researcher to impose a periodization on the data, a defect which, as we shall see, marrs the work of later researchers.

Haustein and Neuwirth (1982) compiled a different time series on basic innovations, and used spectral analysis to test for periodicity in this and other time series related to the long wave. They concluded that there are “some doubts when looking at the regular patterns of inventions and innovations by Mensch and Marchetti” (p. 67). In other words, their results suggest a less strictly periodic pattern than was suggested by Mensch.

In the mean time, the interpretation of the innovation time series data by Mensch gained support from van Duijn (1983), Kleinknecht (1981) and Kleinknecht (1987), who provided more extensive data sets. However, the original results obtained by Mensch were severely criticized by Freeman, Clark and Soete (1982), who argued that the data used by Mensch were not representative and were wrongly dated. A further contribution that was critical of the clustering hypothesis was Solomou (1986). He applied a *z*-test to the null hypothesis that the mean number of basic innovations in adjacent periods were drawn from a normal distribution with the same mean, which could not be rejected.

Kleinknecht (1990) responded to this criticism by constructing a new time series of basic innovations that was a compilation of three different time series used earlier (Mensch, Haustein and Neuwirth, Van Duijn). He used a similar null-hypothesis as Solomou, but relied on a *t*-test rather than a *z*-test. His conclusions were strongly in favour of clustering of basic innovations in the depression periods of the long wave. An important element of the differences in the results between Kleinknecht and Solomou lies in the different periodizations they use (as admitted by Kleinknecht, 1990, p. 86). Both Solomou and Kleinknecht use a division of a full cycle into two periods (up and down), and they rely on periodizations proposed in the earlier literature. These periodizations were based on examining (partly) the same time series for basic innovations as were employed by these authors in testing the clustering hypothesis. Solomou (1986) uses the periodization proposed by Mensch (1979), Kleinknecht (1990) relies on two different periodizations (for different lead times of innovation to economic time series), which were both based on his own work (Kleinknecht, 1987).

The methodology of these studies was strongly criticized by Silverberg and Lehnert (1993). They argue that

> “*z* and *t* tests ... are only applicable to a normally distributed random variable. On a priori grounds we have argued that the null hypothesis on innovations must be that they are

homogeneous Poisson distributed, however, and a histogram of, for example, the Haustein and Neuwirth data (as well as any of the other series we have examined) confirms that they are anything but normally distributed ... both authors apply their tests to sub samples they claim have been selected on a priori criteria ... In general, the periodisations employed derive from previous authors such as Mensch, whose use of a runs test did not depend on it, or on the examination of growth rates and the addition of a time lag ... growth rates and a moving average of the innovation data may be highly (cross) correlated, so that the selection of a proper lag against variations in the growth rate series may simply be a method to select sub periods of above- and below-average innovation activity even from a completely random series. This fact would further invalidate any means test (even one appropriate to a Poisson process, such as a binomial statistic)". (p. 31)

Thus, Silverberg and Lehnert argue that the statistical methodology used by Solomou and Kleinknecht is flawed, both with regard to the assumed distribution underlying the test statistics, and the particular choice for a periodization scheme, and go on to apply a number of nonparametric tests of a homogeneous Poisson process without imposing a periodization on the time series. This is convincingly rejected for all series. They then test for the existence of an exponential trend, again using a nonparametric method appropriate for count (Poisson) data, and show such a trend to be highly significant for all series. The growth rate of the trend is estimated (on the assumption of a nonhomogeneous Poisson process with exponentially growing arrival rate) and found to be in the range ½ to 1% per annum, depending on the series examined. After eliminating the trend from the data by a process of exponential slowing down of the original time series, the resulting trendless series are still shown to deviate significantly (although less so than non-detrended data) from a time-homogeneous (i.e., nonclustering) Poisson distribution. They concluded that, while to a first approximation these count data appeared to be generated by a Poisson process with exponentially growing trend, it could not be determined with these methods whether the statistically significant remaining deviations where indeed periodic or were random clustering resulting from some point process other than Poisson (e.g., negative binomial). They then suggested further research using a more elaborate modeling strategy based on the Poisson distribution (pp. 33) to carry this line of inquiry further.

This advance in the methodological discussion forces us to confront the question of what explicit statistical hypothesis actually corresponds to the Schumpeterian theory of basic innovation, since different variants are conceivable that necessitate different statistical tests. One particularly strong hypothesis, which we will term the Schumpeter Mark I.0, seems to be implicit in most of the empirical work, although it is not necessarily attributable to Schumpeter himself. This hypothesis might be formulated as follows. Basic innovations are generated by a stochastic process, but one whose arrival rate fluctuates deterministically and strictly periodically with a period of 50-60 years.

A more involved version of this hypothesis, which we will refer to as Schumpeter Mark I.1, also claims that the arrival rate varies deterministically, but as a function of macroeconomic variables such as the profit rate. The arrival rate will thus have time series properties similar to those of such variables and correlate with them, but not necessarily be strictly periodically. Which way causality runs will be difficult to determine (with Mensch arguing from macroeconomic variables to innovation back to macroeconomic variables, while Schumpeter, as near as we can make out, would only argue from innovation clusters to macroeconomic growth). If the macroeconomy is in fact more or less periodic, then the Schumpeter Mark I.0 and I.1 hypotheses will be almost impossible to distinguish.

Schumpeter (1939) himself appears to argue only that innovations cluster, without any reference to a specific deterministic structure determining when exactly clusters of innovations will occur in time (i.e., in which phase of the economic wave, or even in regularly spaced spells of high and low innovation activity). Such a 'weak' variant of clustering may be consistent with a random clustering patterning, i.e., periods of high and low innovation activity do occur in the data, but there is no regular and predictable mechanism governing their occurrence. Hypotheses of this kind we will classify as Schumpeter Mark II: the stochastic process generating innovations is more complicated than Poisson (in particular: it is overdispersed) and clusters arrive stochastically. Two variants can be differentiated. Mark II.0 invokes purely random clustering from some additional distribution (such as a negative binomial process where the arrival rate fluctuates according to a Gamma distribution; see below). Mark II.1 calls for the clustering to be initiated by random events, but then obey some sort of causal mechanism that makes this random event 'persist' for some time. An example of such a process would be an autoregressive process, in which (with positive autocorrelation) a single random event tends to be followed by higher activity in the periods immediately following. To see how such a hypothesis might consistent with at least part of Schumpeter's arguments, recall the quotation from Schumpeter (1939) in footnote 1.

In the remainder of this paper we shall attempt to adapt our statistical methodology to allow us to address these fours types of hypotheses explicitly by drawing on recent advances in the analysis of count data.

3. Poisson regression

As we have argued above, innovations are intrinsically count data, and thus will be generated by a point process. While the statistical properties of point processes have been researched since the beginnings of mathematical statistics, many economists are still relatively unacquainted with this branch of the subject. Moreover, many relevant modeling tools have only been developed recently.

The starting point for all subsequent analysis is the (time-homogeneous) Poisson process, which makes the simplifying assumption that the probability of occurrence of an innovation within a given interval of time is independent of previous innovations and independent of time.

The probability of y innovations during an interval of time T is given by

$$Prob(y) = \frac{e^{-ëT}(ëT)^{y}}{y!},$$

where y å ℕ, and ë is a parameter, often referred to as the arrival rate of the Poisson process. Note that ë is not necessarily an integer number. It is easily shown that the expected number of events per unit time is ë, which also happens to be the variance of the distribution. Note that time series generated from a time-homogeneous Poisson process will not display a completely uniform pattern of occurrences of the random event. In other words, to the naive eye some clustering will characterize even this simplest point process. The model presented by Silverberg and Lehnert demonstrates that such an exogenous and unstructured innovation process is sufficient to generate long waves of economic growth.

The parameter ë may also be specified endogenously, for example as lnë = **â**'**x**, where **x** is a vector of independent variables, and **â** is a parameter vector. Other specifications for ë are possible, but the example above is often used because it is convenient for estimation purposes (using lnë as the independent variable ensures non-negative arrival rates). In this case, the parameter vector **â** can be interpreted as a vector of elasticities of the arrival rate with respect to the independent variables. Such an approach allows to test the assumption that the Poisson process is time-homogeneous by setting up the null hypothesis that all elements of **â**, including one corresponding to time itself, are equal to zero.

Sahal (1974) proposed using the Poisson model to examine the characteristics of various time series of innovations in different industries. His conclusion was that "invention is properly characterized as a Poisson random process, but [...] its rate is a function of economic forces" (p. 403). Although Sahal did provide some estimates of **â**-type parameters, these were based on ordinary least square methods. However, when the data contain many zero and small integer values, a maximum likelihood approach based explicitly on the Poisson distribution is more appropriate. Such a procedure was introduced into the literature on innovation by Hausman, Hall and Griliches (1984). They estimated a model in which the number of patents of a firm is related to the firm's R&D expenditures. Elaborations on this approach were presented by Crepon and Duguet (1997), Crepon and Duguet (1997) and Cincera (1997).

One problem with the Poisson model is its characteristic that the mean and variance of the distribution are equal. The empirical data often show a larger variance than mean for the dependent variable, a phenomenon termed 'overdispersion'. Hausman, Hall and Griliches (1984) observed overdispersion in their firm level patent database. A model that can account for overdispersion may be obtained by adding an unobserved random effect to the mean of the Poisson distribution (Hausman, Hall and Griliches, 1984). This leads to a modified probability distribution of the type:

$$Prob(Y = y|u) = \frac{e^{-ëu}(ëu)^y}{y!},$$

where u is a random variable for which some distribution must be assumed (see Greene, 1995, p. 939). The variable u may, for example, reflect random noise, or cross-sectional heterogeneity (when the model is estimated in the cross-sectional dimension). Assuming that u is gamma distributed, one obtains the following unconditional distribution (Cameron and Trivedi, 1998, p. 71):

$$Prob(Y = y|\mathbf{x}) = \frac{Ã(á^{-1} + y)}{Ã(y+1)\,Ã(á^{-1})} r^{y}(1-r)^{á^{-1}}, \quad \text{where } r = \frac{ë}{ë + á^{-1}}.$$

This distribution is known as the negative binomial distribution, and has mean ë and variance ë(1+áë), for á>0. When á approaches 0, the model reduces to a standard Poisson model, and the variance becomes equal to ë again. The negative binomial model can also be estimated using a maximum likelihood method. A test of the Poisson against the negative binomial distribution can be implemented by means of a Likelihood Ratio test, Wald test or *t*-test (Green, 1995) of the null hypothesis á=0.

In the context of time series and the hypothesis of clustering of events (innovations) in specific time periods, the correlation structure of the residuals becomes of interest. Cameron and Trivedi (1998) suggest investigating the standardized residual $z_t = (y_t - m_t)/\sqrt{ó_t}$, where y_t is the observed (integer) value, m_t is the sample mean value and $ó_t$ is the sample variance. When a Poisson model is fitted, $m_t = ó_t$ is equal to the estimated arrival rate. Cameron and Trivedi (1998) then suggest applying either the Box-Pierce portmanteau, the Box-Ljung statistic (which has better small sample properties) or a slightly modified statistic that guards against incorrect standardization to test for the null-hypothesis that all autocorrelations of the residuals up to lag *k* are zero.

In case such a null-hypothesis is rejected, various ways are suggested (Cameron and Trivedi, 1998, Section 7.5) of specifying a model to deal with the autocorrelated residuals. The method closest to our original Poisson regression approach is to estimate an autoregressive model, i.e., to include lagged values of the dependent variable in the regression as independent variables. The simplest model is called exponential feedback, and assumes $\ln ë_t = â'\mathbf{x}_t + ñy_{t-1}$. This, however, implies explosive behaviour for ñ > 0. In order to rule out this undesirable property, Cameron and Trivedi suggest using $ë_t = \exp(â'\mathbf{x}_t)\sum_k (y^*_{t-k})^{ñ}{}_k$, where y^* is a transformation of y to ensure positive values, i.e., $y^* = \max(½, y)$, or $y^* = y + ½$.

4. Empirical results

Like Kleinknecht (1990), we will use a time series for basic innovations that is created from several other time series introduced by other researchers. This new innovation time series was created by merging the two longest time series available in the literature, i.e., Haustein and Neuwirth's data with Van Duijn's. Kleinknecht's's time series is too short compared to the two other basic innovation time series. The Baker data were not included in this series because they are based on patent data rather than innovation data, and hence are less compatible with the other two time series. Instead, the Baker time series will be used separately in the statistical analysis.

The construction of the merged innovation time series used here differs substantially from the Kleinknecht approach. The main difference refers to the overlap, i.e., those innovations which are covered in both sources. Kleinknecht constructs a time series in which the innovations that occur more than once in the three time series he considers are double counted. In other words, he simply adds up the numbers of innovations per year in the three innovation time series. Kleinknecht justifies this procedure by arguing that it provides some implicit weighting scheme, in which the important innovations (i.e., those on which all sources agree) are weighted more heavily.

It is clear that such an implicit weighting procedure is not adequate in the context of a Poisson regression approach. This is why a different approach was chosen here. This approach consists of identifying the innovations that are covered by both samples, and counting these only once. A complication in this procedure is that, as noted by Kleinknecht, the innovation dates of the same innovation often differ between the two sources. The majority of overlap cases are dated in a range of 10 years, but differences of up to 50 years exist. In all cases, the earliest date was used to assign the innovation to the merged sample. The merged sample contains 88 innovations that occur in both samples, 90 innovations that only occur in the Haustein and Neuwirth sample, and 70 cases that are only listed in the Van Duijn sample. The merged series thus has 248 innovations, dating from 1764 to 1976. The complete listing of all innovations in the merged sample, as well as the original Haustein and Neuwirth and Van Duijn sources, together with their assignment to the merged sample, are given in the appendix. Following Kleinknecht's terminology, we will refer to this merged time series as the ‘supersample time series'.

Figure 1 shows histograms for the Baker patent time series, the two time series we used to create the ‘supersample time series' (Haustein & Neuwirth and Van Duijn), and the ‘supersample time series'. The Baker time series runs from 1769 until 1970, and thus comprises 202 years. The Haustein and Neuwirth (1982) data run from 1764 until 1975 (212 years), while Van Duijn (1983) covers the period 1811 - 1971 (161 years). The ‘supersample' comprises the period 1764 - 1976 (213 years).

All three histograms show that the highest frequency is found for zero innovations. Also, all histograms show declining frequencies for larger numbers of innovations per year. No time series shows more than seven innovations per year (this occurs twice in the Haustein and Neuwirth series). Silverberg and Lehnert (1993) constructed similar histograms and concluded that these show that the assumption of normally distributed data must be rejected at face value. As noted by Silverberg and Lehnert, this invalidates the statistical methods used by Kleinknecht (1990), and suggests that fitting a Poisson or negative binomial model to the time series might be a more useful approach.

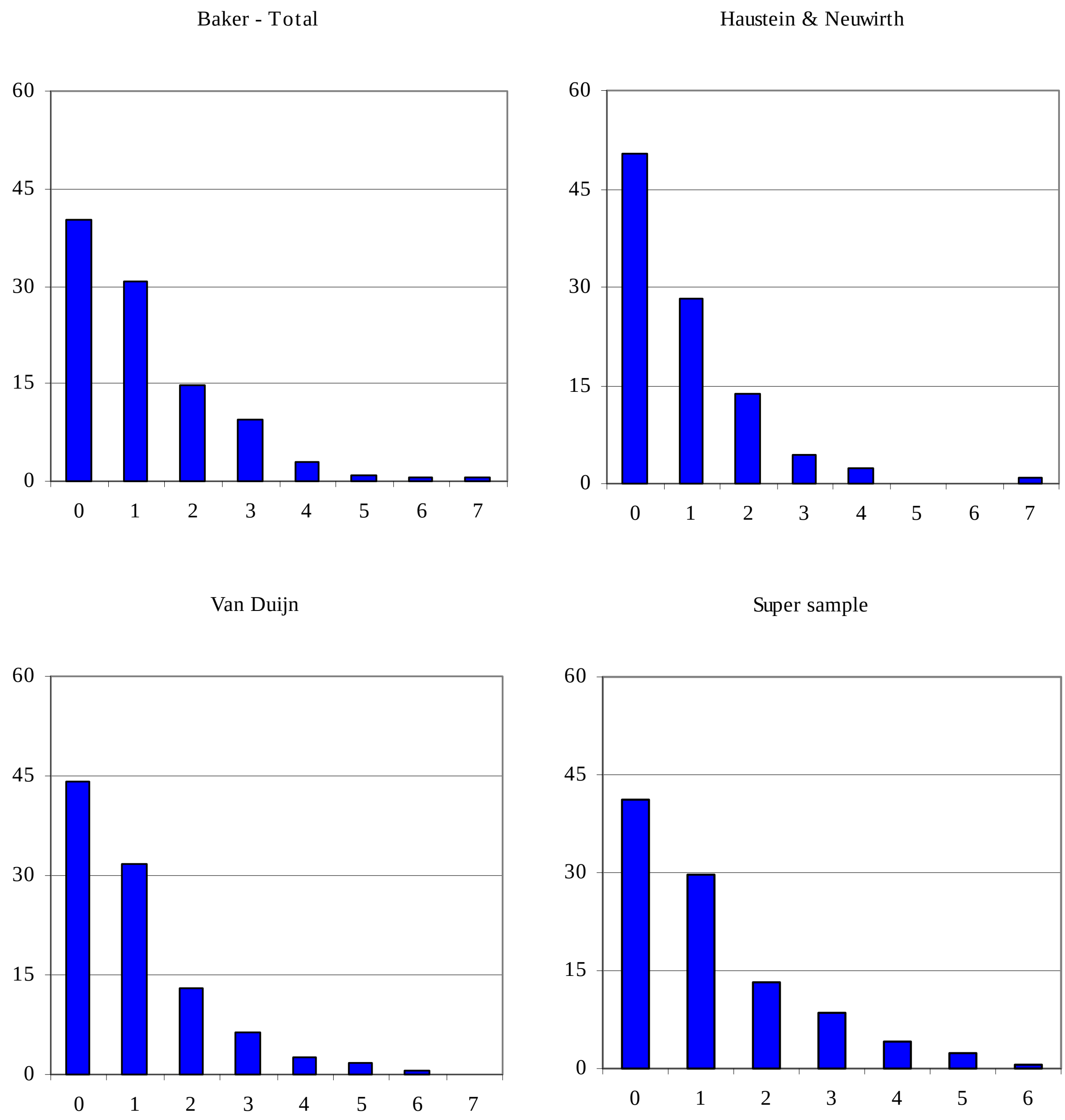


Figure 1. Histograms of basic innovation time series data, on horizontal axis: number of innovations per year; on vertical axis: percentage share of total sample

We start by fitting a simple Poisson model with a constant arrival rate (i.e., a time-homogenous process), and then proceed to implement two departures from this simple model. The first is to allow for overdispersion by fitting a negative binomial model. The second is to estimate the arrival rates (in both the Poisson and the negative binomial model) as following a time trend. The latter is specified by equating the log of the arrival rate to a polynomial function of time, where we experiment with polynomial degrees up to 3 (a log formulation is used to ensure that the arrival rate is always nonnegative). Thus we estimate lnë = c + â$_1$ t + â$_2$ t^2 + â$_3$ t^3 (where t is time and we set some of the higher order âs to zero in some estimations). The results of these estimations are displayed in Table 1.

The hypothesis of a constant arrival rate (time-homogenous Poisson process, or all âs equal to zero) is obviously rejected (the numbers in brackets are p-values associated with the t-tests on the estimated coefficients). Both for the supersample innovations time series and the Baker patents time series, the exponential quadratic trend emerges as the ‘best’ model: entering the third-order term only leaves the linear trend significant, while a second-order polynomial emerges with both âs significant. We thus proceed on the assumption that a quadratic trend captures the long-run growth rate of the number of basic innovations in a reasonable way.

Figure 2 displays the raw data and the fitted linear and quadratic exponential trends for the two models (Poisson and negative binomial). The monotonically increasing lines are the fitted linear trends, the lines leveling off towards the end of the period are the estimated quadratic trends. The estimated trends do not differ greatly between the pure Poisson and the negative binomial models. There is, however, a major difference between the linear and quadratic trends. The latter shows a higher level during the period (roughly) 1850 until 1900, and levels off around 1930 (supersample) or 1920 (Baker patents). This would seem to indicate that the rate of basic innovation is slowing down in the 20th century after a period of relatively rapid increase in the last part of the 19th century. This phenomenon may, however, be caused by an end of sample bias in the time series caused by the fact that at the time when the time series were constructed (during the 1970s), it was not yet clear which recent innovations would prove to be basic.

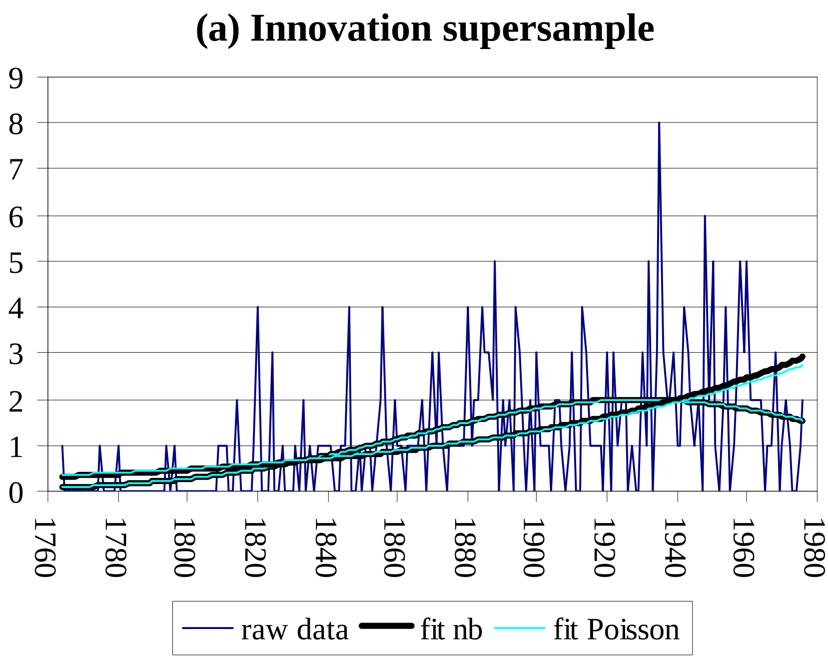
With regard to overdispersion, the results differ somewhat between the two time series. In the case of the supersample of innovations, the negative binomial model is always preferred to the simple Poisson model, as is clear from the fact that the á parameter is always significant (at the 10% level). In the case of the Baker patent time series, the quadratic and cubic models both yield ás that are significant at a level just above 10%, while the linear exponential trend (10% level) model and the time-homogenous process (1% level) yield significant ás.

These results have implications for the hypotheses that were introduced above. Specifically, the fact that the negative binomial model is generally preferred over the simple Poisson model points to the fact that the data are in fact overdispersed. In other words, compared to a (time-homogenous) Poisson process, the data display clustering in the sense of random spells of high and low innovation activity. We interpret this as evidence in favour of the Schumpeter Mark II.0 hypothesis.

Table 1. Regression results, Poisson and Negative Binomial models, arrival rate function of time

Innovation Data Source	N	Start year	*c*	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	á	logl	SQ	Q
Innovations	213	1764	0.1521					-331.654	***	***
			(0.000)							
Innovations	213	1764	0.1521				0.6087	-316.186	***	***
			(0.069)				(0.000)			
Innovations	213	1764	-1.0568	0.9692[a]				-291.823	**	**
			(0.000)	(0.000)						
Innovations	213	1764	-1.1420	0.0104			0.2662	-286.994	**	**
			(0.000)	(0.000)			(0.023)			
Innovations	213	1764	-2.5467	0.0389	-0.1168[b]			-279.317		
			(0.000)	(0.000)	(0.000)					
Innovations	213	1764	-2.5702	0.0393	-0.1186[b]		0.1744	-276.876		
			(0.000)	(0.000)	(0.000)		(0.077)			
Innovations	213	1764	-2.6012	0.0408	-0.1347[b]	0.4946[e]		-279.311		
			(0.000)	(0.028)	(0.410)	(0.912)				
Innovations	213	1764	-2.6159	0.0410	-0.1351[b]	0.4669[e]	0.1742	-276.872		
			(0.000)	(0.044)	(0.465)	(0.928)	(0.077)			
Baker patents	202	1769	0.1123					-299.213	***	***
			(0.041)							
Baker patents	202	1769	0.1123				0.4254	-291.537	***	***
			(0.166)				(0.005)			
Baker patents	202	1769	-0.7487	0.7256[a]				-279.997	***	***
			(0.000)	(0.000)						
Baker patents	202	1769	-0.7867	0.7577[a]			0.2422	-276.740	***	**
			(0.000)	(0.000)			(0.059)			
Baker patents	202	1769	-1.6670	0.0269	-0.8347[c]			-274.230	**	*
			(0.000)	(0.000)	(0.000)					
Baker patents	202	1769	-1.6616	0.0268	-0.8268[c]		0.1876	-272.086	**	*
			(0.000)	(0.000)	(0.001)		(0.110)			
Baker patents	202	1769	-1.7662	0.0308	-0.1225[b]	0.1133[d]		-274.204	**	*
			(0.000)	(0.029)	(0.366)	(0.769)				
Baker patents	202	1769	-1.7267	0.0294	-0.1093[b]	0.7774[d]	0.1870	-272.076	**	*
			(0.000)	(0.068)	(0.491)	(0.866)	(0.112)			

Notes: estimated coefficients are equal to documented coefficients divided by the following factors (absence of a note indicates estimated coefficient is equal to documented coefficient): [a] 100, [b] 1000, [c] 10000, [d] 1000000, [e] 10000000, N is number of observations, numbers between brackets are p-values associated with the *t*-statistics. One, two and three starts point to rejection at the 10, 5 or 1% level, respectively, for a ÷^{2}-test of the null-hypothesis "all autocorrelations of the residuals up to this order are zero" ,Column SQ gives the statistic suggested by Cameron & Travedi, p. 229, Column Q gives the Box-Ljung statistic, all residuals used in these tests are standardized.



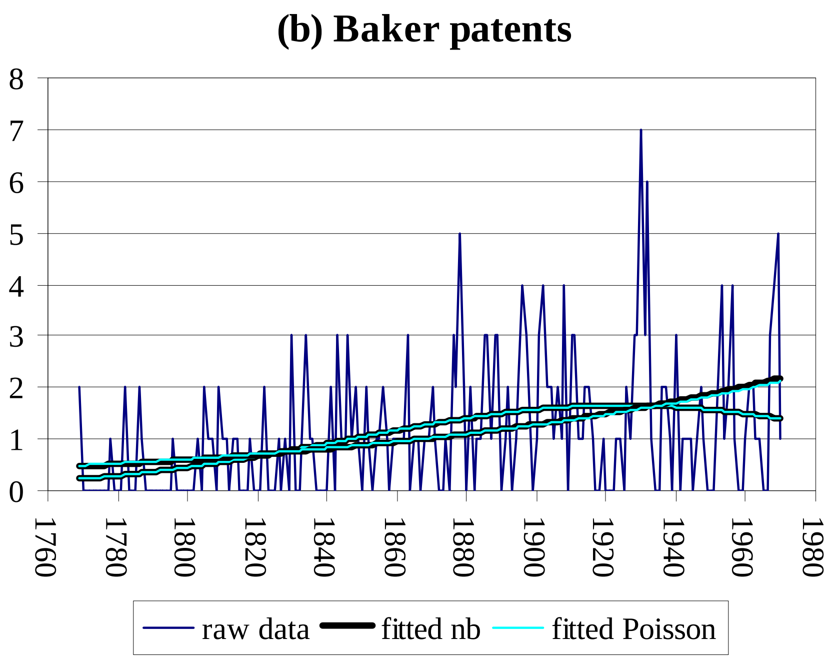


Figure 2. Raw data and fitted linear and quadratic trends (details of estimation are in Table 1)

In order to get a first impression of the deviations from the long-run trend growth rate which bears on the three other hypotheses, an analysis of the autocorrelation of the residuals of the regressions was performed. To this end, the residuals were standardized by dividing them by the estimated standard deviation (square root of the variance). As explained above, in the case of a Poisson process, this implies dividing by the square root of the arrival rate itself (which varies over time when a trend is included). In the case of a negative binomial model, the variance is calculated as ë(1+áë), where ë again varies over time in the event of a time trend. We test for the null-hypothesis that all autocorrelations up to (and including) order *k* (= 1..20) are zero using two tests. The first is suggested by Cameron and Trivedi (p. 229), and “guards against incorrect standardization” of the residuals. The second is the standard Box-Ljung statistic. Both tests apply a statistic that is ÷^{2} distributed with k degrees of freedom (the maximum order of the autocorrelation process).

The last column in Table 1 gives the significance level at which (if at all) we reject the null-hypothesis of no autocorrelation in the residuals. We used orders up to 20 to test for this. In the case of a time-homogenous process, there is clearly very significant autocorrelation, for both time series (patents and innovations) and for the pure Poison model as well as for the negative binomial model. When time trends are included, this diminishes, although the extent to which this happens differs greatly between the two time series. With the innovation data, a linear time trend still leaves significant autocorrelation at the 5% level, for the Baker patent time series at the 1% level for 3 of 4 cases. For higher order polynomial trends, however, autocorrelation vanishes completely in the case of the innovation data. For the patent data, the second and third order polynomial trends leave significant autocorrelation in the residuals at the 10% or 5% level. These results suggest that for the Baker patent time series, the models in Table 1 leave at least some aspects of the long-run dynamics of the time series unexplained, and hence that there is scope for investigating the three other Schumpeterian hypotheses. We now proceed in two ways.

First, we estimate a set of autoregressive models specified in the way explained above. One problem in estimating these models is to determine the exact order of the autoregressive process. This is in fact common to all autoregressive models, and the approach most commonly found in the literature is an empirical procedure employing information criteria. These are simple statistics based on the log likelihood value of the models for different autoregressive orders. The simplest is the Aikake Information Criterion (AIC), which we will use here. The procedure is to find the minimum value of the AIC for a range of autoregressive orders, which in the present case we select as 0 - 20.

Table 2. Regression results, Poisson and Negative Binomial models, arrival rate function of time and lagged dependent variable

	Innovations		Baker patents	
coefficient	(1)	(2)	(3)	(4)
c	-2.7656 (0.000)	-2.6037 (0.000)	-1.6902 (0.000)	-3.3500 (0.001)
$\hat{a}_1$	0.0424 (0.000)	0.0395 (0.000)	0.0271 (0.000)	0.0526 (0.001)
$\hat{a}_2$	-0.1290[b] (0.000)	-0.1207[b] (0.000)	-0.8646[c] (0.001)	-0.1711[b] (0.003)
á	0.1762 (0.078)			
$\tilde{n}_1$	-0.0184 (0.871)	0.0431 (0.675)	0.2651 (0.010)	0.2342 (0.056)
$\tilde{n}_2$		0.6907[a] (0.947)	0.0933 (0.302)	0.0748 (0.527)
$\tilde{n}_3$		-0.0809 (0.422)	-0.1776 (0.056)	-0.1914 (0.122)
$\tilde{n}_4$		0.1680 (0.065)		0.0622 (0.617)
$\tilde{n}_5$		0.0516 (0.578)		-0.1807 (0.145)
$\tilde{n}_6$		-0.1155 (0.234)		-0.0612 (0.647)
$\tilde{n}_7$		-0.0417 (0.692)		-0.0209 (0.866)
$\tilde{n}_8$		0.0447 (0.646)		0.0767 (0.497)
$\tilde{n}_9$		-0.2083 (0.040)		-0.1685 (0.143)
$\tilde{n}_{10}$		0.2489 (0.010)		-0.1302 (0.323)
$\tilde{n}_{11}$		-0.2508[a] (0.983)		-0.1335 (0.337)
$\tilde{n}_{12}$		-0.0263 (0.801)		-0.1832 (0.123)
$\tilde{n}_{13}$				0.1076 (0.350)
$\tilde{n}_{14}$				-0.1912 (0.125)
$\tilde{n}_{15}$				-0.2145[a] (0.987)
$\tilde{n}_{16}$				0.0867 (0.480)
$\tilde{n}_{17}$				-0.1436 (0.233)
$\tilde{n}_{18}$				-0.0891 (0.479)
$\tilde{n}_{19}$				-0.0843 (0.474)
logl	-274.108	-264.790	-263.916	-240.489

Notes: estimated coefficients are equal to documented coefficients divided by the following factors (absence of a note indicates estimated coefficient is equal to documented coefficient): [a] 100, [b] 1000, [c] 10000, [d] 1000000, [e] 10000000, numbers between brackets are p-values associated with the *t*-statistics.

We estimated (and calculated the AIC) for a model with a quadratic exponential time trend plus autoregressive terms. Both a Poisson model and a negative binomial model were estimated for both time series and all autoregressive orders in the range 0 - 20. For the innovation supersample, the Poisson model has minimum AIC at order 1, while the negative binomial model has minimum AIC at order 12. When the AIC is plotted against the autoregressive order, the Poisson model also has a local minimum that does not differ much in value from the global minimum at order 1, at order 12, while the negative binomial model has a local minimum (again not so much different from the global minimum) at order 1. We therefore document the full set of results for the model at both orders 1 and 12. When the negative binomial term á is significant, we only document this model, while if it is not significant, we only document the Poisson model.

For the Baker patent time series, the minimum value of the AIC was found at orders 19 (Poisson) and 12 (negative binomial). The Poisson model has a local minimum of the AIC at order 3, which is also documented. In this case, the negative binomial term á is not significant for either the autoregressive order 19 or 3, which is why we document only pure Poisson models. The results for the autoregressive model are documented in Table 2.

We note from the table that the significance of the time-parameters â is largely unaffected, i.e., that the time trends estimated in Table 1 are robust against the inclusion of autoregressive terms. This does not hold for the negative binomial term á. When autoregressive terms are included in the model, this term becomes insignificant in three of the four cases in Table 2 (and, in fact, in most of the undocumented autoregressive models that were run; these results are available on request). The general loss of significance of the negative binomial model points to the fact that the clustering effects that we have attributed above (in our Schumpeter Mark II.0 hypothesis and the discussion of Table 1) to overdispersion may in fact also (and statistically more satisfactorily) be captured by an autoregressive model.

This raises the question as to what is the specific pattern of clustering implied by the estimated autoregressive structure, and whether this pattern is consistent with, for example, our Schumpeter Mark II.1 hypothesis). The answer to this question obviously depends on the value of the estimated ñ-parameters. As in a conventional autoregressive model, the values of these parameters have implications for the speed at which 'errors' leave their trace in the fitted time series. This speed can be visualized by plotting the impulse response function, which, because of the model we have specified in Section 3, must be a multiplicative function rather than the commonly used additive function (see, e.g., Hamilton, 1994, ch. 1).

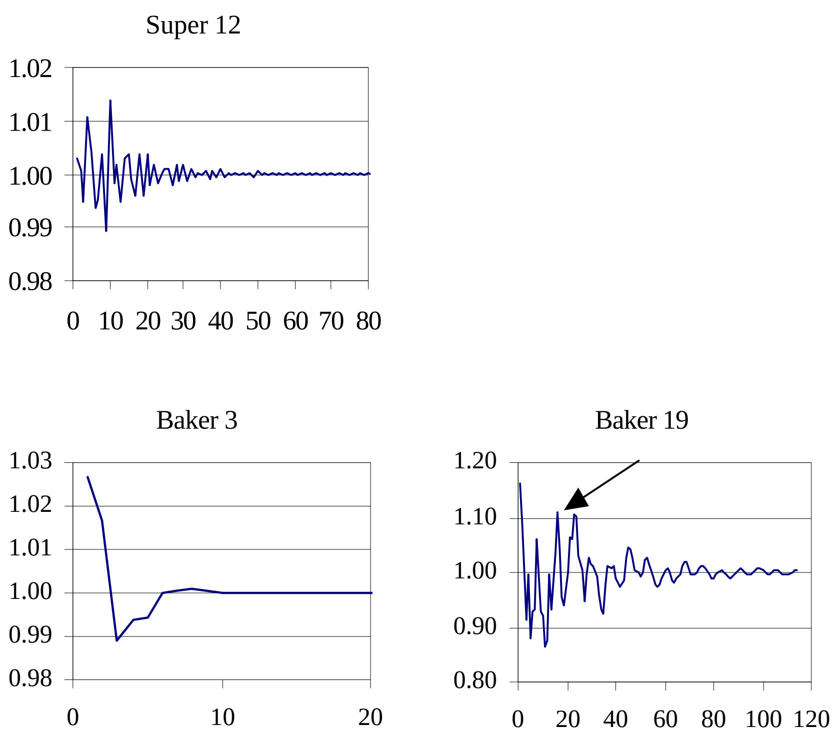


Figure 3. Impulse-response functions for the autoregressive processes in Table 2

Figure 3 shows the impulse response function for columns (2) - (4) in Table 2. The function for column (1) is almost completely flat and is therefore not documented. These graphs display the net multiplicative effect of a random innovation in period 0 on the arrival rate in the subsequent periods. For example, the figure for the Baker patent time series at autoregressive order 19 shows that the effect of one 'random extra' innovation in period 0 is to raise the arrival rate in period 19 by slightly more than 10% compared to a situation without such a random extra innovation (value of 1.1 at the peak is indicated by the arrow). Note that the fact that all three graphs converge to one implies that the effect of a random extra innovation ultimately dies out, i.e., that the process is (trend) stationary.

All three graphs show oscillatory behaviour following a 'random extra' innovation. This is caused by the negative autoregressive parameters (ñs), and leads us to reject the Schumpeter Mark II.1 hypothesis introduced above. This hypothesis argues that a random extra innovation leaves a slowly dying trace of (relatively) high innovation activity. Such a pattern would result if (most of) the estimated ñs had high positive values. With the large number of negative values for the autoregressive parameters we obtain in Table 2, the result is the highly irregular and random-looking pattern observed in Figure 3. This is consistent to some extent with the Schumpeter Mark II.0 hypothesis of random clustering, but not with the Schumpeter Mark II.1 hypothesis.

This still leaves both Schumpeter Mark I hypotheses to be investigated. As a preliminary remark, note that if a strictly periodic pattern (possibly surrounding a trend of some form) existed in the data, one would expect the autoregressive specification to pick this up in some form. In other words, the irregular patterns observed in Figure 3 also have implications for the Schumpeter Mark I hypotheses, in particular, they cast doubt on these hypotheses. To clarify this further, we employ an additional test to reach a final verdict on Schumpeter Mark I.

This test does not involve data on economic variables, as called for in the Mark I.1 hypothesis. Instead, we investigate the residuals from the regressions in Table 1 in a different way. Our main aim is to investigate whether any strictly periodic movements can be discerned in the residuals. To this end, we perform spectral analysis on the standardized residuals (defined above). If the Schumpeter Mark I hypotheses are valid, one would expect clear peaks in the spectral density plot for these residuals, as a result of strictly periodic movements around the estimated trend arrival rate.

Figure 4 displays the spectral density plots for residuals from the Poisson and negative binomial regressions in Table 1. Although there clearly are some peaks in the plots, these are consistent with white-noise data. The overall impression is one of a relatively flat spectrum, which indicates the absence of clear periodic components in the residuals. This holds especially for the supersample innovation time series, and to a somewhat lesser extent for the Baker patent time series.

We thus conclude that the evidence in favour of strictly periodic movements around the (exponential) quadratic trend in the arrival rate is weak. The finding that is closest to our Schumpeter Mark I hypotheses is the ‘weak’ peak at cycle length of 20 years in the Baker patent time series.

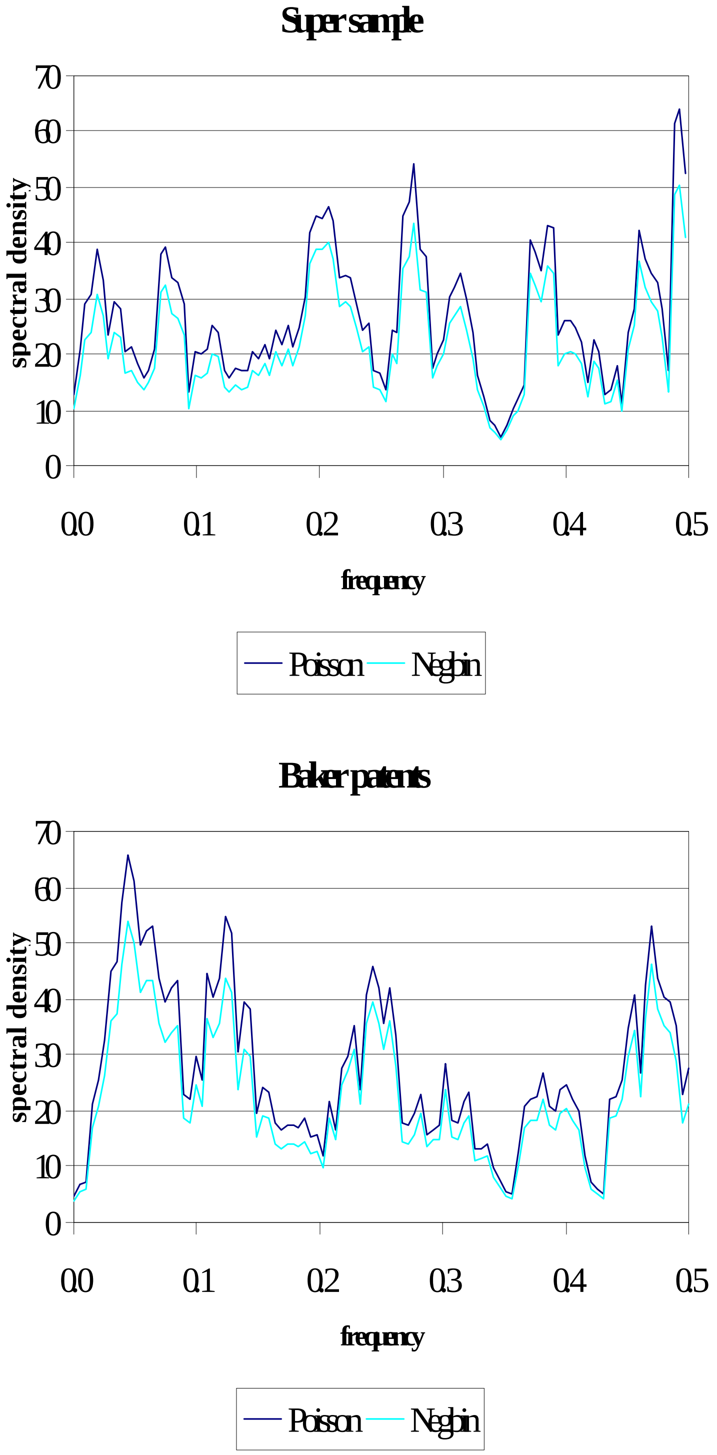


Figure 4. Spectral density plots for the residuals of the quadratic trend regressions in Table 1.

5. Summary and conclusions

This paper pursues a suggestion made by Silverberg and Lehnert (1993), in their discussion of the empirical evidence on the 'Schumpeterian hypothesis' that basic innovations tend to cluster. They proposed, in order to overcome the methodological shortcoming of previous work, modeling and estimating a stochastic Poisson process describing the occurrence of basic innovations. We formulate four alternative hypotheses bearing on the Schumpeterian clustering hypothesis that are amenable to statistical testing. We then estimated a model of the trend in the arrival rate of basic innovations, and found that a quadratic trend fits the data best (compared to a linear or third-degree trend). Investigating the residuals of this regression for periodic cycles around the trend, we found essentially no evidence in favour of such a strong clustering effect.

A weaker version of the clustering theory interprets clusters as themselves randomly distributed. Evidence for this interpretation is presented in the form of estimates of the negative binomial parameter, which is found to be significantly different from zero.

A different version of random clustering occurs when random events tend to 'persist' in time, for example due to an autoregressive process. We have estimated autoregressive Poisson and negative binomial models (again based on a quadratic trend), and found that these models may indeed improve the fit. However, by plotting the impulse response functions, we showed that innovation impulses do not lead to long periods of above-normal innovative activity. Rather, such impulses die out in an oscillatory way, and leave a rather irregular and random-looking trace. We interpret this as evidence against a knock-on effect being the source of clustering.

We are thus left with the conclusion that the only form of clustering of basic innovations that is consistent with the data is 'random clustering' superimposed on a second-degree trend. Basic innovations, while they arrive in clusters, do not engender higher rates of basic innovative activity (although they may initiate periods of higher incremental innovation not picked up in our data), nor are these clusters in any way periodic. To what extent they are influenced by the macroeconomic context is a subject we hope to study further.

References

Baker, R. (1976). *New and improved: inventors and inventions that have changed the modern world*. London, British Museum Publications.

Cameron, A.C., and P.K. Trivedi. (1998). *Regression Analysis of Count Data*. Vol. 30, *Econometric Society Monographs*. Cambridge: Cambridge University Press.

Cincera, M. (1997). "Patents, R&D and technological spillovers at the firm level: some evidence from econometric count models for panel data." *Journal of Applied Econometrics* **12**: 265-280.

Crepon, B. and E. Duguet (1997). "Estimating the innovation function from patent numbers: GMM on count panel data." *Journal of Applied Econometrics* **12**: 243-263.

Crepon, B. and E. Duguet (1997). “Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity.” *Journal of Econometrics* **79**: 355-378.

Freeman, C., J. Clark, and L. Soete (1982). *Unemployment and Technical Innovation*. London, Pinter.

Freeman, C. and L. Soete (1997). *The Economics of Industrial Innovation. 3rd Edition*. London and Washington, Pinter.

Gordon, R.J., (2000). “Does the "New Economy" Measure up to the Great Inventions of the Past?” NBER Working Paper No. W7833.

Greene, W. (1995). *Econometric Analysis, 3rd ed.* New Jersey, Prentice Hall.

Hamilton, J.D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

Hausman, J., B. H. Hall and Z. Griliches (1984). “Econometric Models for Count Data with an Application to the Patents-R&D Relationship.” *Econometrica* **52**: 909-938.

Haustein, H.-D. and E. Neuwirth (1982). “Long Waves in World Industrial Production, Energy Consumption, Innovations, Inventions, and Patents and Their Identification by Spectral Analysis.” *Technological Forecasting and Social Change* **22**: 53-89.

Kleinknecht, A. (1981). “Observations on the Schumpeterian Swarming of Innovations.” *Futures* **13**.

Kleinknecht, A. (1987). *Innovation Patterns in Crisis and Prosperity. Schumpeter's Long Cycle Reconsidered*. London, Macmillan.

Kleinknecht, A. (1990). “Are there Schumpeterian Waves of Innovations.” *Cambridge Journal of Economics* **14**: 81-92.

Kuznets, S. (1940). “Schumpeter's Business Cycles.” *American Economic Review* **30**.

Mensch, G. (1979). *Stalemate in Technology. Innovations Overcome Depression*. Cambridge, Ballinger.

Sahal, D. (1974). “Generalized Poisson and Related Models of Technological Innovation.” *Technological Forecasting and Social Change* **6**: 403-436.

Schumpeter, J. A. (1939). *Business Cycles: A theoretical, historical and statistical analysis of the capitalist process*. New York, McGraw-Hill (page numbers quoted in the text refer to the abridged version reprinted in 1989 by Porcupine Press, Philadelphia).

Silverberg, G. and D. Lehnert (1993). “Long Waves and ‘Evolutionary Chaos’ in a Simple Schumpeterian Model of Embodied technical Change”, *Structural Change and Economic Dynamics*, vol. 4, pp. 9-37.

Solomou, S. (1986). “Innovation clusters and Kondratieff long waves in economic growth.” *Cambridge Journal of Economics* **10**, pp. 101-112.

Ten Raa, T. And E.N. Wolff (2000). “Engines of Growth in the U.S. Economy”, *Structural Change and Economic Dynamics*, vol. 11, forthcoming.

Van Duijn, J. J. (1983). *The Long Wave in Economic Life*. London, Allen & Unwin.

Annex. Data on basic innovations

Table A1. The supersample of basic innovations

Item number	Innovation	Year
89	Spinning machine	1764
90	Steam engine	1775
91	Automatic band loom	1780
92	Sliding carriage	1794
81	Blast furnace	1796
48	Steam ship	1809
95	Whitney's method	1810
80	Crucible steel	1811
183	Street lighting (gas)	1814
184	Mechanical printing press	1814
78	Lead chamber process	1819
77	Quinine	1820
98	Isolated conduction	1820
99	Rolled wire	1820
100	Cartwright's loom	1820
3	Steam locomotive	1824
61	Cement	1824
66	Puddling furnace	1824
101	Pharma fabrication	1827
102	Calciumchlorate	1831
79	Telegraphy	1833
103	Urban gas	1833
104	Rolled rails	1835
87	Electric motor	1837
67	Photography	1838
9	Bicycle	1839
88	Vulcanized rubber	1840
7	Arc lamp	1841
105	Jacquard loom	1844
106	Lathe	1845
107	Inductor	1846
108	Electrodynamic measuring	1846
185	Rotary press	1846
186	Anaestetics	1846
187	Steel (puddling process)	1849
188	Sewing machine	1851
109	Plaster of paris	1852
14	Aluminium	1854
40	Safety match	1855
189	Bunsen burner	1855
36	Refined steel/Bessemer steel	1856
84	Steel pen / Fountain pen	1856

110	Tare colours industry	1856
111	Baking powder	1856
190	Elevator	1857
76	Lead battery	1859
191	Drilling for oil	1859
54	Internal combustion engine	1860
39	Soda works	1861
19	Anilin dyes	1863
192	Siemens-Martin steel	1864
112	Paper from wood	1865
4	Deep sea cable	1866
50	Dynamite	1867
75	Dynamo	1867
113	Commutator	1869
60	Typewriter	1870
193	Celluloid	1870
194	Combine harvester	1870
85	Margarine	1871
6	Thomas steel	1872
46	Reinforced concrete	1872
114	Drum rotor	1872
115	Preservatives	1873
51	Sulphuric acid	1875
195	Four-stroke engine	1876
5	Telephone	1877
116	Nickel	1878
53	Electric Railway	1879
12	Incandescent lamp	1880
47	Water turbine	1880
117	Jodoforme	1880
196	Half-tone process	1880
197	Electric power station	1881
118	Veronal	1882
119	Cable	1882
120	Antipyrin	1883
121	Coals whisks	1883
10	Steam turbine	1884
122	Chloroforme	1884
198	Punched card	1884
199	Cash register	1884
38	Syntethic fertilizers	1885
52	Transformer	1885
123	Synthetic alcaloids	1885
124	Magnesium	1886
125	Electric welding	1886
200	Linotype	1886
72	Phonographe	1887

126	Electrolyse	1887
28	Motor car	1888
59	Pneumatic tyre	1888
127	Electric counter	1888
201	Portable camera	1888
202	Alternating-current generator	1888
86	Man-made fibres	1890
128	Chemical fibres	1890
129	Melting by induction	1891
83	Acetylene welding	1892
130	Accounting machine	1892
41	Cinematography	1894
131	Antitoxines	1894
203	Motor cycle	1894
204	Monotype	1894
8	Diesel engine	1895
132	Drilling machine for mining	1895
205	Electric automobile	1895
206	X-rays	1896
37	Aspirin	1898
133	Arc welding	1898
134	Air ship	1900
135	Synthesis of indigo	1900
207	Submarine	1900
136	Holing machine	1901
137	Electric steel making	1902
208	Safety razor	1903
209	Viscose rayon	1905
210	Vacuum cleaner	1905
138	Acetylen	1906
211	Chemical accelerator for rubber vulcanization	1906
212	Electric washing machine	1907
15	Gyro compass	1909
2	Airplane	1910
69	Bakelite (Phenol plastics)	1910
139	High voltage isolation	1910
65	Vacuum tube	1913
71	Assembly line	1913
213	Thermal cracking	1913
214	Domestic refrigerator	1913
140	Ammonia synthesis	1914
141	Tractor	1914
215	Stainless steel	1914
142	Tank	1915
32	Synthetic rubber	1916
42	Cellophane	1917

1	Zip fastener	1918
29	AM Radio	1920
216	Acetate rayon	1920
217	Continuous thermal cracking	1920
26	Synthetic detergents	1922
57	Insuline	1922
143	Synthesis of methanol	1922
35	Continuous rolling	1923
218	Dynamic loudspeaker	1924
219	Leica camera	1924
144	Deep frozen food	1925
220	Electric record player	1925
145	Coal hydrogenation	1927
17	Power steering	1930
221	Polystyrene	1930
222	Rapid freezing	1930
223	Freon refrigerants	1931
34	Crease-resistent fabrics	1932
224	Gas turbine	1932
225	Polyvinylchloride	1932
226	Antimalaria drugs	1932
227	Sulfa drugs	1932
56	Fluor lamp	1934
146	Diesel locomotive	1934
147	Fischer-Tropsch procedure	1934
11	Radar	1935
13	Ballpoint pen	1935
30	Rockets/guided missiles	1935
31	Plexiglas	1935
62	Magnetophone	1935
70	Catalytic cracking	1935
82	Colour photo	1935
148	Gasoline	1935
16	Television	1936
149	Photoelectric cell	1936
228	FM radio	1936
150	Vitamins	1937
229	Electron microscope	1937
20	Helicopter	1938
230	Nylon	1938
21	Polethylene	1939
55	Automatic gears	1939
151	Hydraulic gear	1939
27	Antibiotics (penicilline)	1940
152	Cotton picker	1941
43	Jet engine/plane	1942
45	DDT	1942

153	Heavy water	1942
231	Continuous catalytic cracking	1942
24	Silcones	1943
232	Aerosol spray	1943
233	High-energy accelerators	1943
44	Streptomycine	1944
154	Titanreduction	1944
22	Sulzer loom	1945
68	Oxygen steelmaking	1946
234	Phototype	1946
49	Numerically controlled machine tools	1948
58	Continuous steel making	1948
235	Orlon	1948
236	Cortisone	1948
237	Long-playing record	1948
238	Polaroid land camera	1948
155	Thonet furniture	1949
156	Polyester	1949
18	Computer	1950
23	Transistor	1950
25	Xerography	1950
239	Terylene	1950
240	Radial tyre	1950
157	Double-floor railway	1951
158	Cinerama	1953
241	Colour television	1953
33	Nuclear energy	1954
242	Gas chromatograph	1954
243	Remote control	1954
244	Silicon transistor	1954
159	Air compressed building	1956
160	Atomic ice breaker	1957
161	Space travel	1957
162	Stitching bond	1958
163	Holography	1958
164	Transistor radio	1958
165	Diffusion process	1958
245	Fuel cell	1958
166	Quartz clocks	1959
246	Polyacetates	1959
247	Float glas	1959
167	Maser	1960
168	Micro modules	1960
248	Polycarbonates	1960
249	Contraceptive pill	1960
250	Hovercraft	1960
64	Integrated circuit	1961

169	Planar process	1961
73	Laser	1962
251	Communication satellite	1962
170	Implementation of ions	1963
171	Epitaxy	1963
172	Synthetic leather	1964
173	Transistor laser	1964
174	Optoelectronic diodes	1966
252	Wankel-motor	1967
74	Video	1968
175	Light emitting fluor display	1968
176	Minicomputers	1968
177	Quartz watches	1970
63	Microprocessor	1971
178	Electronic calculator	1971
179	Light-tunnel technology	1972
180	16-bit microprocessor	1975
181	16384 bit RAM	1976
182	Microcomputer	1976

Note: items 1-88 occur in both databases, items 89-182 occur only in Haustein and Neuwirth data, items 183-252 occur only in Van Duijn data.

Table A2. Innovations in the Haustein and Neuwirth time series that were matched to innovations in the Van Duijn time series

Innovation	Year	To which item in merged series?
Blast furnace	1796	to 81
Steamer	1809	to 48
Crucible cast steel	1811	to 80
Lead-chamber process	1819	to 78
Chinin fabrication	1820	to 77
Locomotive	1824	to 3
Puddling furnace	1824	to 66
Telegraphy	1833	to 79
Photography	1838	to 67
Bicycle (pedal)	1839	to 9
Cement	1844	to 61
Arc lamp	1844	to 7
Generator of current	1849	to 87
Hard rubber	1852	to 88
Aluminium	1854	to 14
Refined steel	1856	to 36
Steel pen	1856	to 84
Lead accumulator	1859	to 76
Soda works	1861	to 39
Production of analin	1863	to 19
Deep sea cable	1866	to 4
Safety matches	1866	to 40

Dynamite	1867	to 50
Dynamo	1867	to 75
Thomas steel	1872	to 6
Typewriter	1873	to 60
Sulphuric acid production	1875	to 51
Telephone	1878	to 5
Electric locomotive	1879	to 53
Incandescent lamp	1880	to 12
Cooking fat	1882	to 85
Electricity	1882	to 87
Electric heating	1882	to 87
Long distance conduction	1882	to 87
Synthetic fertilizers	1885	to 38
Transformers	1885	to 52
Combustion engine	1886	to 54
Phonograph	1887	to 72
Tyres with air compression	1888	to 59
Water turbine	1890	to 47
Welding by acetylene	1892	to 83
Steam turbine	1895	to 10
Automobile	1895	to 28
Cinematography	1895	to 41
Electric railway	1895	to 53
Diesel engine	1897	to 8
Aspirin	1898	to 37
Steel concrete	1902	to 46
Gyro compass	1909	to 15
Pheno plastics	1910	to 69
Airplane	1911	to 2
Conveyor belt production	1913	to 71
Synthetic rubber	1916	to 32
Electronic tubes	1920	to 65
Detergents/synthetic	1922	to 26
Radio	1922	to 29
Insuline	1922	to 57
Zip fastener	1923	to 1
Continuous rolling	1923	to 35
Cellophane	1926	to 42
Power steering	1930	to 17
Crease-resistant fabrics	1932	to 34
Fluorescent lamp	1934	to 56
Ball-point pen	1935	to 13
Rockets	1935	to 30
Plexiglass	1935	to 31
Magnetophone	1935	to 62
Catalytic cracking	1935	to 70
Colour film	1935	to 82

TV	1936	to 16
Radar	1939	to 11
Helicopter	1939	to 20
Automatic gears	1939	to 55
Synthetic fibres	1939	to 86
Antibiotics	1940	to 27
DDT	1942	to 45
Jet engine	1943	to 43
Streptomycine	1944	to 44
Sulzer loom	1945	to 22
Silicons	1946	to 24
Oxygen-process	1946	to 68
NC machines	1948	to 49
Continuous steelmaking	1948	to 58
Computer	1950	to 18
Transistor	1950	to 23
Xerographie	1950	to 50
Polyethylene	1953	to 21
Nuclear power station	1954	to 33
Integrated circuits	1961	to 64
Laser	1962	to 73
Video-tape recorder	1968	to 74
Microprocessor	1971	to 63

Table A3. Innovations in the Van Duijn time series that were matched to innovations in the Haustein and Neuwirth time series

Innovation	Year	To which item in merged series?
Crucible steel	1811	to 80
Sulphuric acid (lead chamber process)	1819	to 78
Quinine	1820	to 77
Portland cement	1824	to 61
Coke blast furnace	1829	to 81
Steam locomotive	1830	to 3
Puddling furnace	1832	to 66
Electric motor	1837	to 87
Steamship (Atlantic crossing)	1838	to 48
Photography	1839	to 67
Electric telegraph	1839	to 79
Vulcanized rubber	1840	to 88
Arc lamp	1841	to 7
Safety match	1855	to 40
Bessemer steel	1856	to 36
Lead battery	1859	to 76
Internal combustion engine	1860	to 54
Sodium carbonate	1861	to 39
Aniline dyes	1865	to 19
Atlantic telegraph cable	1866	to 4

Dynamite	1867	to 50
Dynamo	1867	to 75
Typewriter	1870	to 60
Margarine	1871	to 85
Reinforced concrete	1872	to 46
Sulphuric acid	1875	to 51
Telephone	1877	to 5
Electric railway	1879	to 53
Thomas oven	1879	to 6
Incandescent lamp	1880	to 12
Water turbine	1880	to 47
Steam turbine	1884	to 10
Fountain pen	1884	to 84
Transformer	1885	to 52
Bicycle	1885	to 9
Aluminium	1887	to 14
Motor car	1888	to 28
Cylindrical record player	1888	to 72
Pneumatic tyre	1889	to 59
Mechanical record player	1889	to 72
Rayon (nitro-cellulose pr.)	1892	to 86
Motion picture film	1894	to 41
Diesel engine	1895	to 8
Rayon (cuprammonium pr.)	1898	to 86
Aspririn	1899	to 37
Oxy-acetylene welding	1903	to 83
Airplane	1910	to 2
Bakelite	1910	to 69
Gyro compass	1911	to 15
Synthetic fertilizer (nitrogen)	1913	to 38
Vacuum tube	1913	to 65
Assembly line	1913	to 71
Cellophane	1917	to 42
Zip fastener	1918	to 1
AM radio	1920	to 29
Continuous hot strip rolling	1923	to 35
Insulin	1923	to 57
Synthetic detergents	1930	to 26
Synthetic rubber	1932	to 32
Crease-resisting fabrics	1932	to 34
Radar	1935	to 11
Plexiglas	1935	to 31
Magnetic tape recorder	1935	to 62
Colour photography	1935	to 82
Television	1936	to 16
Catalytic cracking	1937	to 70
Helicopter	1938	to 20

Fluorescent lamp	1938	to 56
Polyethylene	1939	to 21
Penicillin	1942	to 27
Guided missiles	1942	to 30
Jet airplane	1942	to 43
DDT	1942	to 45
Silicones	1943	to 24
Ball-point pen	1945	to 13
Streptomycin	1946	to 44
Automatic transmission (passenger cars)	1948	to 55
Sulzer loom	1950	to 22
Xerography	1950	to 25
Power steering (passenger cars)	1951	to 17
Electronic computer	1951	to 18
Transistor	1951	to 23
Continuous casting of steel	1952	to 58
Oxygen steel making	1953	to 68
Numerically controlled machine tools	1955	to 49
Nuclear energy	1956	to 33
Integrated circuit	1961	to 64
Laser	1967	to 73
Video cassette recorder	1970	to 74
Micro-processor	1971	to 63

Eindhoven Centre for Innovation Studies

W O R K I N G   P A P E R S

ECIS working papers (November 2000):

98.1: Per Botolf Maurseth & Bart Verspagen:
Knowledge spillovers in Europe and its consequences for systems of innovation

98.2: Jan Fagerberg & Bart Verspagen:
Productivity, R&D spillovers and trade

98.3: Leon A.G. Oerlemans, Marius T.H. Meeus & Frans W.M. Boekema:
Learning, innovation and proximity

99.1: Marius T.H. Meeus, Leon A.G. Oerlemans & Jules J.J. van Dijck:
Regional systems of innovation from within

99.2: Marcel P. Timmer: Climbing the technology ladder too fast?:
An international comparison of productivity in South and East- Asian manufacturing, 1963-1993

99.3: Leon A.G. Oerlemans, Marius T.H. Meeus, Frans W.M. Boekema:
Innovation and space: theoretical perspectives

99.4: A. Mukherjee & N. Balasubramanian:
Technology transfer in a horizontally differentiated product-market

99.5: Marius T.H. Meeus, Leon A.G. Oerlemans & J. Hage:
Sectoral patterns of interactive learning

99.6: Bart Verspagen:
The role of large multinationals in the dutch technology infrastructure

99.7: Leon A.G. Oerlemans & Marius T.H. Meeus:
R&D cooperation in a transaction cost perspective

99.8: Gerald Silverberg & Bart Verspagen:
Long memory in time series of economic growth and convergence

99.9: B. Bongenaar & A. Szirmai:
The role of a research and development institute in the development and diffusion of technology

99.10: M.C.J. Caniëls & B. Verspagen:
Spatial distance in a technology gap model

99.11: J.I.M. Halman, J.A. Keizer & X.M. Song:
Perceived risks in product innovation projects: development of a risk skeleton

99.12: A. Mukherjee:
Subsidy and entry: role of licensing

99.13: M.C.J. Caniëls & B. Verspagen:
The effects of economic integration on regional growth, an evolutionary model

00.01: O. Lint:
The primary assessment tool at Philips Electronics:
Capturing real options and organizational risk in technology portfolio management

00.02: O. Lint & E. Pennings:
The recently chosen digital video standard: playing the game within the game

00.03: A. Mukherjee & E. Pennings:
Imitation, patent protection and welfare

00.04: J. Hagedoorn & G. Duysters
The effect of mergers and acquisitions on the technological performance of companies in a high-tech environment

00.05: O. Lint & E. Pennings:
The V-shaped value evolution of R&D projects

00.06: H. Romijn:
Technology support for small industries in developing countries:
A review of concepts and project practices

00.07: B.Verspagen & W. Schoenmakers:
The spatial dimension of knowledge spillovers in
Europe: evidence from firm patenting data

00.08: B. Verspagen:
Growth and structural change: trends, patterns and policy options

00.09: O. Lint:
Retrospective insights from real options in R&D

00.10: O. Marsili:
Technological regimes and sources of entrepreneurship

00.11: A. Nuvolari:
The 'machine breakers' and the industrial revolution

00.12 B. Verspagen:
Economic growth and technological change: an evolutionary interpretation

00.13 M. Albaladejo & H. Romijn:
Determinants of innovation capability in small UK firms: an empirical analysis

00.14 O. Marsili:
Sources of Concentration and Turbulence in Evolutionary Environments:
Simulations of Learning and Selection

00.15 R. Bekkers, G. Duysters, B. Verspagen:
Intellectual property rights, strategic technology agreements and market structure. The case of the GSM

00.16 G. Silverberg & B. Verspagen:
Breaking the waves: a poisson regression approach to schumpeterian clustering of basic innovations

These papers can be downloaded from the ECIS website: http:/www.tm.tue.nl/ecis/.
On request hardcopies can be sent. Please contact ECIS.