DECOMPOSING THE LONG WAVE THROUGH A NON-PARAMETRIC STOCHASTIC MEASUREMENT MODEL

By
Ronald Schoenberg
National Institute of Mental Health
and
Albert J. Bergersen
University of Arizona
March, 1986

ABSTRACT

Earlier work demonstrated the presence of two long waves of colonial administration of different lengths (1490 - 1825 and 1826 - 1969). Whether these were separate episodes or examples of deeper underlying cyclical dynamics has implications for the existence of a common system dynamic over the long wave. To further inquire into the existence of a common cyclical rhythm these larger waves were decomposed through the use of a non-parametric stochastic measurement model.

To do this a 490 year time series of colonial administrations is divided into ten episodes. A conditional Normal-Poisson model is proposed based on the assumptions of a stochastic process. The mean number of colonial administrations established and terminated over each episode are estimated, controlling for a quadratic time trend which would be induced if the system was not constant throughout an episode as assumed. Two sub-cycles are observed withing each of the two long waves of colonial administrations previously reported (Bergesen and Schoenberg, 1980).

The presence of these matching sub-cycles provides strong evidence supporting a common system dynamic in not only economic but political aspects of international life.

In 1415 Portugal established the first European colony on the coast of North Africa. This launched the first of two long waves of European colonization that eventually swept the globe. The first wave began very slowly, a second colony was established 65 years after the first in 1480, another in 1482 and 1485, and two more in 1496. From there the European empire steadily mounted, increasing to the apex of the first wave at the end of the 18th century. The colonial system then collapsed with the dissollution of the French and Spanish empires, and shortly afterwards a new, second wave of colonization began.

An interesting feature of the European colonization is that it is borne out of the competition of independent states. Eligible states, members of the "core," vie with one another in the subjegation of alien territories. And this process is not cooperative. The colonized territories become elements in a centuries long dispute among the core countries that was (and perhaps will continue to be) more devastating to human life than anything that the planet has ever seen.

Figure 1 contains a diagram of the 490 history of European colonization from 1480 to 1969 (the 65 years from 1415 to 1480 are omitted since there were no colonies established in those years and for analytical purposes we shall begin the colonial system with 1480). The curve in Figure 1 traces the net number of colonial governors established by European core states in existence in each year from 1480 to 1969 as reported in Henige (1970). This curve is the joint outcome of a process of the establishment of colonies and the termination of colonies. This in any given year the net number of colonies is the number of colonies that have been established minus the number of colonies terminated or transfered to other states.

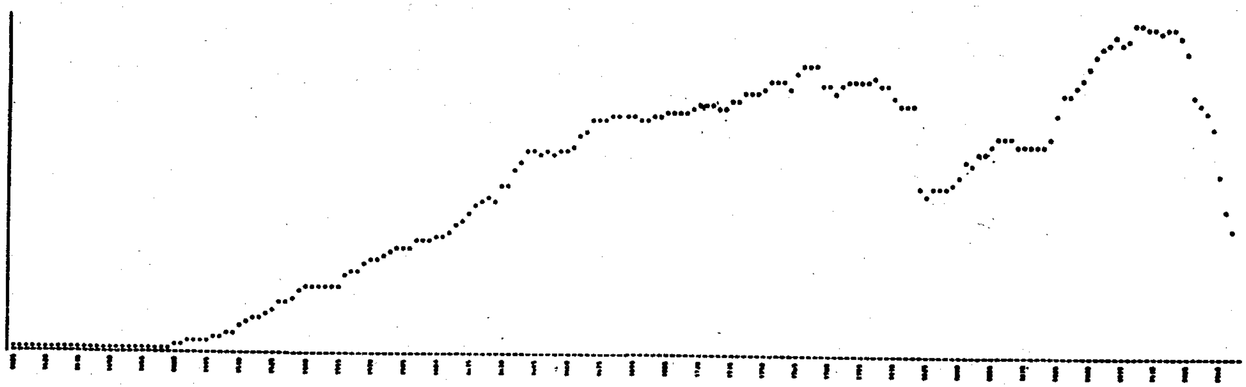

FIGURE 1: Net number of colonies of all European states from 1480 to 1969.

We have used the establishment of a colonial governor as a measure of colonies, first, because this is the only data available for the entire history of colonialism, and second, because a political, as opposed to some geographical, definition is more consistent with our intended purposes.

In our earlier report (Bergesen and Schoenberg, 1980), we did not propose any kind of systematic model to account for variation in the number of colonies. a sophisticated theory sufficient for the specification of a math model did not then exist, nor does one seem to exist at this time. An attempt to apply series analysis of the Box-Jenkins type (Box and Jenkins, 1970) was made by McGowan (1985). In our opinion, however, the application of such a methodology was inappropriate (Bergesen, 1985, and Schoenberg, 1985b). First, the curve of net colonies is clearly nonlinear in both frequency and amplitude. Box-Jenkins time series analysis requires an assumption of stationarity which is clearly violated. These defects may be corrected by the appropriate selection of a "filter" or transformation of the data, but this would remove from the data precisely what we want to explain.

In fact it is not the net curve of colonies that we want to explain anyway. First, the underlying process is revealed in both the establishment and the termination of colonies. We would want, then, to study these outcomes of the process separately. To analyze them combined in a single number would be throwing away information. Second, the true dependent variable is the *rate* of colonization. As in the study of population dynamics in which the birth *rate* (that is, the number of children born per unit time) is the essential structural measure of productivity rather than the sheer number of people in existence, so it will be with colonization: a constant level of the underlying generating process over, say, 10 years, will produce a constant probability or "propensity" to generate a colony in each of those years, and thus though the net number of colonies changes each year (from one in the first year to ten in the tenth year) the rate of change of the number of colonies, one per year, is constant as is the propensity.

The fundamental dependent variables, then, will be the propensities to establish and to terminate colonies. And as we shall see, these propensities may be interpreted as instantaneous rates of change of the number of colonies, That is, the number of colonies established and terminated per year.

THE TIME SERIES MODEL

We shall assume that the time series of colonial governors is an outcome of a stochastic process. Denote the propensity for the establishment of a colonial governor by $\lambda(t)$. The propensity for an event to occur is the instantaneous rate of change of probability of that event occuring:

$$(1) \qquad \lambda(t) = \lim_{dt \downarrow 0} \Pr\{N(t+dt)-N(t)=1\}/dt$$

where $N(t)$ is the number of colonial governors established by time t. From the definition in equation (1) and the assumption that no more than one event can occur in dt, we may conclude that the probability of k colonial governors being established during time interval s is

$$\Pr(N(t+s)-N(t)=k) = Q(t,s)^k \exp\{-Q(t,s)\}/k! \qquad k = 0,1,2,\ldots$$

where

$$Q(t,s) = \int_0^{t+s} a(\tau)d\tau - \int_0^t a(\tau)d\tau$$

Over the 490 year time series of the establishment of colonial governors there appears to be wide variation in the propensity, and $\lambda(t)$ is therefore a complicated function of time. For example, the first 75 years of the 17th century seems to be a period of very high propensity to establish colonies and is followed by another 75 year period of a relative lower propensity. It seems reasonable to suppose, however, that the change in propensity over short periods of time is quite small relative to larger periods of time, decades, say, and that we may therefore assume that the propensity is approximately constant over the range of a single year. It follows that in any given year $\lambda(t) = \lambda$, and

$$Q(t,1) = \int_0^{t+1} \lambda d\tau - \int_0^t \lambda d\tau = \lambda$$

The distribution of the observed number of established colonial governors within any one year, then, is a Poisson distribution with expected value equal to its propensity. Let $X_t = N(t+1)-N(t)$ be the number of colonial governors established in year t, then

$$Pr(X_t=k) = \lambda_t^k exp(-\lambda_t)/k! \qquad k = 0,1,2,...$$

where λ_t is the propensity of year t.

At this point a model that specified the behavior of λ_t for t = 1480, 1481,...1969 would be advantageous. However, a theory that specifies the behavior of the propensity across years does not exist at this time. This behavior is clearly nonstationary and nonlinear and thus the application of any of several well-known linear models of time series would be inappropriate. We shall here attempt, then, to specify a model that will entail very limited claims while providing enough description of the process to permit some conclusions about what is happening.

An inspection of the total number of colonial governors (established minus terminated colonial governors) suggests that the two global cycles may be crudely characterized by the following sequence: (a) an initial low level of colonization (the 16th century and middle 19th century), (b) rapid colonization (17th century and later 19th century), (c) a tapering off of colonization (18th Century and and 1920 to 1940), (d) fluctuating colonization (late 18th century and 1940's, actually this is a period of primarily the exchange of colonies as a product of armed conflict), and (e) the collapse of colonial empires (early 19th century and middle 20th century).

Let us postpone for a moment the discussion of the determination of the precise boundaries of these ten time segments. Instead we shall at this time develop the time series model. First, let us distinquish between the *establishment* and the *termination* of colonial governors. The total number of colonial governors existing at time t is the net outcome of the number of established colonial governors minus the number of terminated governors. Since it is likely that the propensity to establish colonial governors is subject to a different set of determinants than the termination of colonial governors, we shall model them separately. Therefore, let X_{1tj} denote the number of colonial governors *established* in the t-th year of the j-th time segment, and X_{2tj} the number of colonial governors *terminated* in the t-th year of the j-th segment.

$$Pr(X_{ajt}=k|\lambda_{aj}) = (1/k!)\ \lambda_{ajt}{}^k exp(-\lambda_{ajt})$$

$k = 0,1,2,\ldots$
where $a = 1,2,\ t = 1,2,\ldots T_j,\ j = 1,2,\ldots 10$, and T_j is the number of years in the j-th time segment.

The model at this point assumes constant propensities within time segments. This is a rather strong assumption which can be relaxed. To account for variation in the propensity within a segment we may assume that the propensity has some distribution within a segment. Propensity is a positive real variable and therefore let us propose that it has a log-normal distribution:

$$\Pr(\lambda_{ajt} = \Lambda) = (2\pi\Lambda^2\sigma^2_{aj})^{-1/2}\exp[-(\log\Lambda-\mu_{aj})^2/2\sigma^2_{aj}]$$

where μ_{aj} is the mean of the log of the propensity of the a,j-th segment, and σ^2_{aj} is the variance of the log of the propensity of the a,j-th segment.

The assumption that the *means* of the propensities are constant throughout the time segments is also unrealistic since the segments are cut out of a time series which clearly changes smoothly with time. Thus we might expect that a segment cut from the rising part of the curve would have a higher mean at the end of the segment than at the beginning. Segments cut from the top parts of the curve might have higher means at the middle than at either end. For these reasons we shall also specify a quadratic time trend for the propensities within segments:

$$(2) \qquad \mu_{ajt} = \gamma_{aj0} + \gamma_{aj1}\tau + \gamma_{aj2}\tau^2 + \varepsilon_{ajt}$$

where ε_{aj} is $LN(1,\sigma^2_{aj})$, and $\tau = t - t_{aj}$, where t_{aj} is the mean of t within the a,j-th segment.

This model requires 60 parameters to describe 980 observations (= 2 X 490 years). The results reported here were produced from the maximization of the marginal likelihood through the EM method. The marginal likelihood function is

$$(3) \quad \log L = \sum_{a=1}^{2} \sum_{j=1}^{10} \log \int_{-\infty}^{+\infty} \Pr(X_{ajt}|\lambda_{ajt})\Pr(\lambda_{ajt})d\lambda_{ajt}$$

In the EM method (a more complete description of which, as applied to the type of problem presented here, may be found in Schoenberg, 1985a), equation (3) is divided into two parts, the first part containing sufficient statistics, treating the log-propensities as latent or unmeasured variables, as parameters in a log-likelihood conditional on the observations and the unknown parameters, that is, the 60 regression coefficients, 20 residual variances, and 20 auto-correlation coefficients; and the second part containing the unknown parameters conditional on the sufficient statistics. The first step in the EM method, then, computes the sufficient statistics of the latent log-propensity given the observations and either initial guesses of the parameters or estimates from previous iterations, and the second step computes maximum likelihood estimates of the parameters given the sufficient statistics.

Finally, we must consider the determination of the boundaries of the time segments. Such a determination amounts to the estimation of the elements of T_j. The simplest procedure, and the one applied here, is to vary the size of the time segments at the boundaries beginning with a best guess for the T_j. The likelihood function in (3) is then maximized for each variation of the T_j. The values for T_j are those chosen that are associated with the largest maximum.

RESULTS

Table 1 contains the estimates of the mean propensities to establish and terminate colonies for the five time segments with each cycle of colonization. These estimates reveal a set of two sub-cycles within each global cycle that seem to be comparable across the global cycles. An important obstacle to such a comparison, though, is a nonlinearity in the global cycles, that is, an acceleration in the frequency of the global cycles and the sub-cycles. To improve comparability of the sub-cycles across the global cycles, an exponential transformation was performed on the time axis the parameters of which were chosen to equalize the length of the two main global cycles. Figure 2 contains a plot of the mean propensities to establish and terminate colonial governors, as well as the net propensity, on the transformed time axis.[1]

The sub-cycles appear to have a period of .44 time units (i.e., in units of global cycles: 1 = 1480, 2 = 1825, 3 = 1970, etc.) that is roughly constant. A constant period or frequency in the transformed time scale would seem to be justified if we are willing to concede that the first sub-cycle in the second global cycle finished a little late--the sub-cycle actually ended in 1921, but should have ended

		time segment	γ_{1j0}	γ_{2j0}
Cycle 1	1	1480-1606	.476	0.000
	2	1607-1674	1.186	.282
	3	1675-1742	.347	.148
	4	1743-1811	.696	.638
	5	1812-1825	.341	2.108
Cycle 2	1	1826-1874	.9168	.314
	2	1875-1921	2.014	.707
	3	1922-1932	.091	.371
	4	1933-1962	.512	1.931
	5	1963-1969	.142	2.833

TABLE 1: Mean propensities to establish (γ_{1j0}) and to terminate (γ_{2j0}) colonial governors for 10 time segments spanning 490 years from 1480 to 1969.

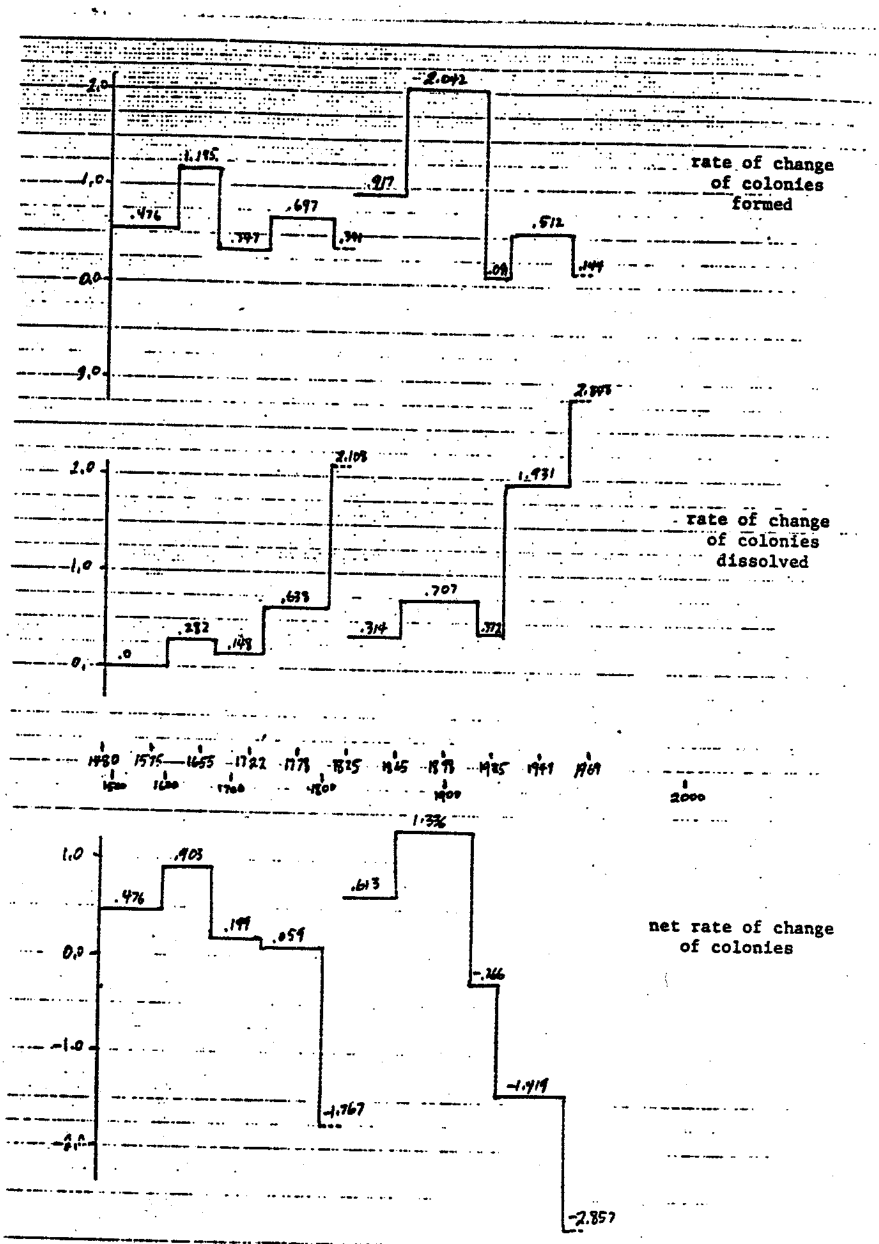

FIGURE 2: Estimated rates of change of colonial governors from 1480 to 1969 corrected for quadratic time trend.

about 1905. In the untransformed time scale, the one measured in calendar years, the period would be nonlinear, .44 units being equal to about 150 years in the 16th century and 70 years in the 19th century. The equivalent period for 1986 is 27 years; thus a complete subcycle today would complete its course in 27 years. If we are undergoing a third global cycle, and assuming that it started about 1975, then the first sub-cycle would be over by about the turn of the century.

An inspection of the net propensities, i.e., established minus terminated propensities, we see that the second wave of colonization peaked in the 19th century with a rather larger propensity (1.336) than the respective peak in the first wave of colonization in the 17th century (.903). However, the second wave clearly deteriorated more rapidly as well. Without more waves we cannot say whether there is a gradual change in the waves towards greater increased intensity of colonization accompanied by a greater deterioration of the system, or whether there is some compensating mechanism which penalizes waves with greater increasing intensity with later increasing deterioration.

Since we have proposed no model (other than the simple stochastic Poisson measurement model) our findings are primarily qualitative. And the most striking result is that the overall smooth curve of net colonization conceals oscillatory motion, that is, the sub-cycles, in its components. This suggests some kind of synchonization in the causal system. In future work on the global cycles we shall investigate models that will attempt to explain first the oscillatory motion of the colonization, and second the coordination of these components into the smooth curve of net colonies that we observe over the 490 years of European colonization.

FOOTNOTES

[1]The transformation of the time axis has only heuristic purposes in this paper. The formula is $t = 2072.173 - 1418.7183*\exp(-.8737*t_s)$, where t is calendar time in years and t_s is measured in global cycle units, i.e., 1 = 1480, the beginning of the first global cycle, 2 = 1825, the beginning of the second global cycle, and 3 = 1969, the beginning of the third global cycle. Taking this transformation seriously one could conclude that, (i) the system started in 653 ($t_x = 0$), (ii) the third global cycle will be over by 2029 ($t_s = 4$), and (iii) the system will have to re-calibrate itself by 2072 ($t_s = \infty$)

On the basis of this formula it is also tempting to conclude that a single year in the 1980's is equal to about 6.8 years in the 15th century, and about 2.8 years in the early 19th century.

REFERENCES

Bergesen, A.J.
1985 "How to Model the Cyclical Dynamics of the World-System: A Reply to Pat McGowan," θεχιεδ, 8(4), 501-512.

Bergesen, A.J., and Schoenberg, R.J.
1980 "Long Waves of Colonial Expansion and Contraction, 1415-1969," in Albert Bergesen (ed.) *Studies of the Modern World System*, Academic Press, 1980.

Box, G.E.P., and G.M. Jenkins
1970 *Time Series Analysis, Forecasting and Control.* San Francisco: Holden Day.

Henge, D.
1970 *Colonial Governors.* Madison: University of Wisconsin Press.

McGowan, P.
1985 "Pitfalls and Promise in the Quantitative Study of the World-System: A Reanalysis of Bergesen and Schoenberg's 'Long Waves' of Colonialism," *Review*, 8(4),477-500.

Schoenberg, R.J.
1985a "Latent Variables in the Analysis of Limited Dependent Variables," in N.B. Tuma (ed.) *Sociological Methodology 1985*, Jossey-Bass, Inc.

1985b "Statistical Models Must Be Appropriate: A Reply to Pat McGowan," *Review*, 8(4),513-516.