

# UC Irvine

## Structure and Dynamics

### Title

Dynamical Feedbacks between Population Growth and Sociopolitical Instability in Agrarian States

### Permalink

<https://escholarship.org/uc/item/0d17g8q9>

### Journal

Structure and Dynamics, 1(1)

### Author

Turchin, Peter

### Publication Date

2005-09-18

### DOI

10.5070/SD911003259

Peer reviewed

## INTRODUCTION

The decline and fall of the Roman Empire, recurrent collapses of Mesopotamian civilizations, and rise and demise of great powers—such historical events excite lay public's imagination and provide fodder for controversies about possible causes among historians (Jones 1964, Kennedy 1987, Yoffe and Cowgill 1988). Over the centuries, historians and political thinkers advanced many explanations of such recurrent waves of state breakdown. Some were quite particularistic, specific for each instance of collapse, such as the role of Christianity in the fall of Rome (Gibbon 1932), others more general, for example, Joseph Tainter's (1988) theory of diminishing returns on sociopolitical complexity. The explanation of breakdown in the agrarian states and empires that is arguably the best grounded in sociological mechanisms is the demographic-structural theory (Goldstone 1991, Turchin 2003c).

According to this theory, population growth in excess of the productivity gains of the land has several effects on social institutions. First, it leads to persistent price inflation, falling real wages, rural misery, urban migration, and increased frequency of food riots and wage protests. Second, rapid expansion of population results in an increased number of aspirants for elite positions. Increased intraelite competition leads to the formation of rival patronage networks vying for state rewards. As a result, elites become riven by increasing rivalry and factionalism. Third, population growth leads to expansion of the army and the bureaucracy and rising real costs. States have no choice but to seek to expand taxation, despite resistance from the elites and the general populace. Yet, attempts to increase revenues cannot offset the spiraling state expenses. Thus, even if the state succeeds in raising taxes, it is still headed for fiscal crisis. As all these trends intensify, the end result is state bankruptcy and consequent loss of the military control; elite movements of regional and national rebellion; and a combination of elite-mobilized and popular uprisings that manifest the breakdown of central authority (Goldstone 1991).

Sociopolitical instability resulting from state collapse feeds back on population growth. Most obviously, when the state is weak or absent, the populace will suffer from elevated mortality due to increased crime, banditry, and internal and external warfare. Additionally, the times of troubles cause increased migration rate, as refugees flee war-affected areas. Migration may lead to emigration (and we can simply add that to mortality) and to spread of epidemics. Increased vagrancy spreads the disease by connecting areas that would stay isolated during better times. As vagabonds and beggars aggregate in towns and cities, increasing their population size, they may tip the density over the epidemiological threshold (a critical density above which a disease spreads). Finally, political instability causes lower reproduction rates, because during uncertain times people choose to marry later and to have fewer children. People's choices about their family sizes may be reflected not only in birth rates, but also in increased rates of infanticide.

Instability can also affect the productive capacity of the society. First, the state offers protection. In a stateless society people can live only in natural strongholds, or places that can be made defensible. Examples include hillfort chiefdoms in preconquest Peru (Earle 1997), and the movement of settlements to hilltops in Italy after the collapse of the Roman Empire (Wickham 1981). Fearful of attack, peasants can cultivate only a small proportion of productive area that is near fortified settlements. The strong state protects the productive population from external and internal (banditry, civil war) threats, and thus allows the whole cultivable area to be put into production. Second, states often invest in increasing the agricultural productivity by constructing

irrigation canals, roads, and flood control structures. A protracted period of civil war results in a deterioration and outright destruction of this productivity-enhancing infrastructure.

I investigated the theoretical relationships between population numbers and social structures, described above, with a suite of dynamical models, ranging from very simple to moderately complex (Turchin 2003c, Turchin and Korotayev 2004). The mathematical theory suggests two general insights. First, population numbers should oscillate with a period of roughly two-three centuries. Second, sociopolitical instability also oscillates with the same period, but shifted in phase with respect to population numbers (Figure 1).

The theory, thus, yields quantitative predictions about the dynamical relationship between population and instability that can be empirically tested using historical data. My goal in this paper is to test the theory on data from several empirical case-studies where data are available on the dynamics of both population and instability: late medieval–early modern England (1450–1800), ancient and medieval China (the Han and Tang periods), and the Roman Empire.

## **METHODS**

### *England 1450-1800: Population Data*

Population numbers for the period 1540–1800 were taken from Table A.9.1 in Wrigley et al. (1997). The quinquennial data of Wrigley et al. were resampled at decadal intervals. For the period 1450–1525 population data were taken from Hatcher (1977), also sampled at 10-y intervals (all data analyzed here were sampled at 10-y intervals). The value for 1530 was interpolated. The population data show an increasing long-term trend. Such nonstationarity violates one of the most important assumptions of nonlinear time-series analysis; thus, data need to be detrended (Turchin 2003a:175).

### *Detrending the English Population Data*

Agrarian revolution in England started during the seventeenth century (Grigg 1989, Allen 1992, Overton 1996). We can trace this revolution using data on long-term changes in grain yields (Grigg 1989, Overton 1996). Average wheat yields in the thirteenth century were around 10 bushels of grain per acre. Yields declined slightly during the fourteenth and fifteenth centuries to 8 bushels per acre (perhaps as a result of the worsening global climate). Even as late as the 1580s, the yields were still at their late medieval level. During the seventeenth century, however, yields began improving, increasing to ca. 15 in 1700 and 20-21 in the early nineteenth century (Grigg 1989:69). Net yields (subtracting seed corn) were lower. For example, the typical late medieval seeding rates were 2 bushels per acre; thus, the net yield was only 6 bushels per acre.

Net yields from Grigg and Overton are plotted in Figure 2a. To capture the rising trend, I fitted the data after 1580 with a straight line (see Figure 2a, note the log-scale). The linear relationship appears to be an adequate description of the trend (thus, adding a quadratic term failed to better the regression in a statistically significant fashion).

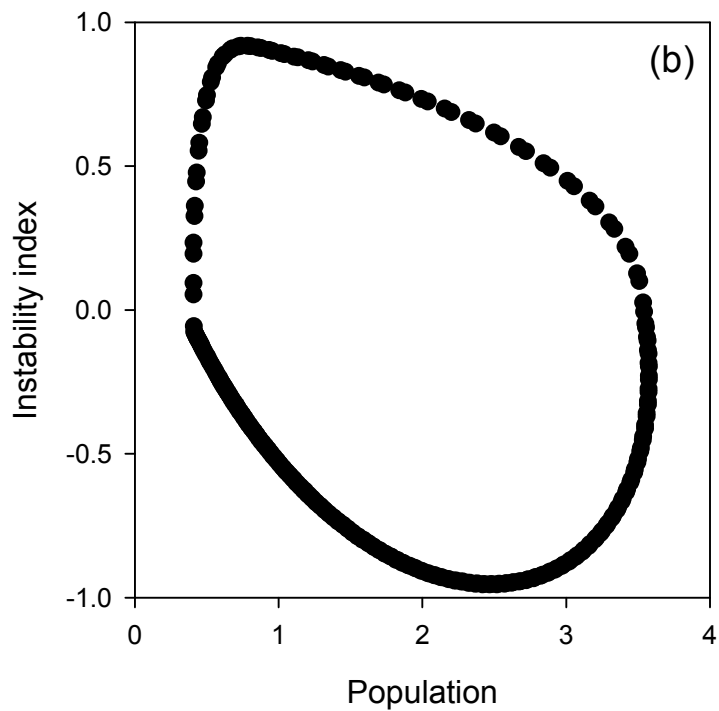
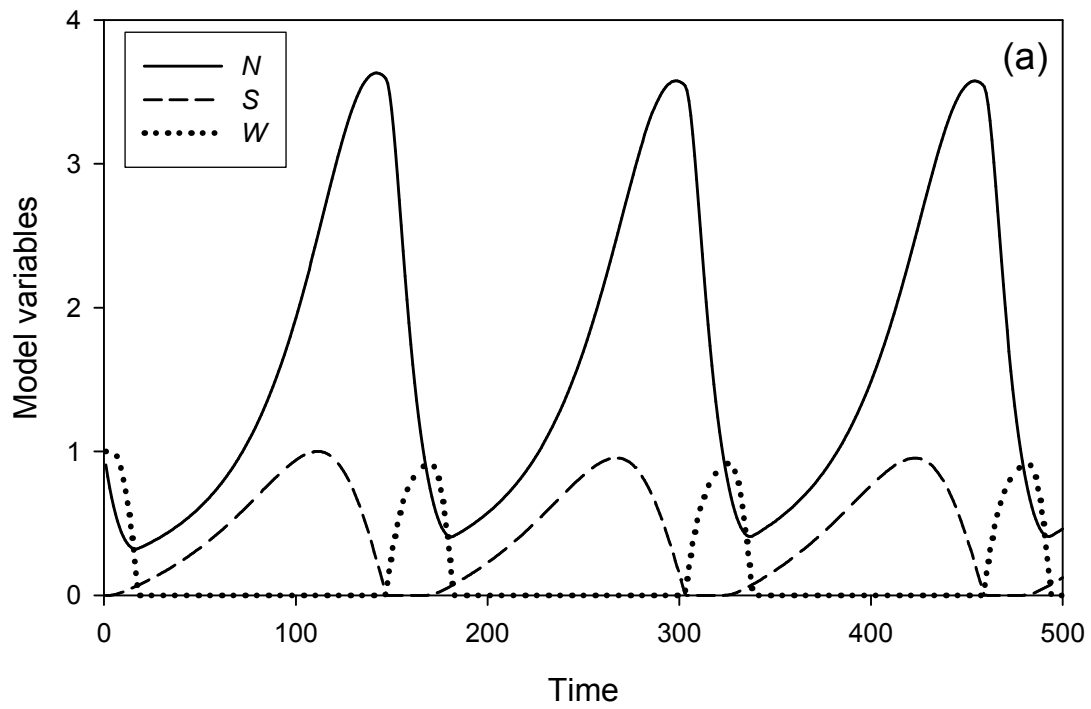


Figure 1. (a) Typical dynamics of population ( $N$ , solid curve), state strength ( $S$ , broken curve), and warfare intensity ( $W$ , dotted curve), illustrated with the output of the model investigated in Turchin and Korotayev (2004). (b) The trajectory in the phase space of population and sociopolitical instability (defined as the difference between  $W$  and  $S$ ).

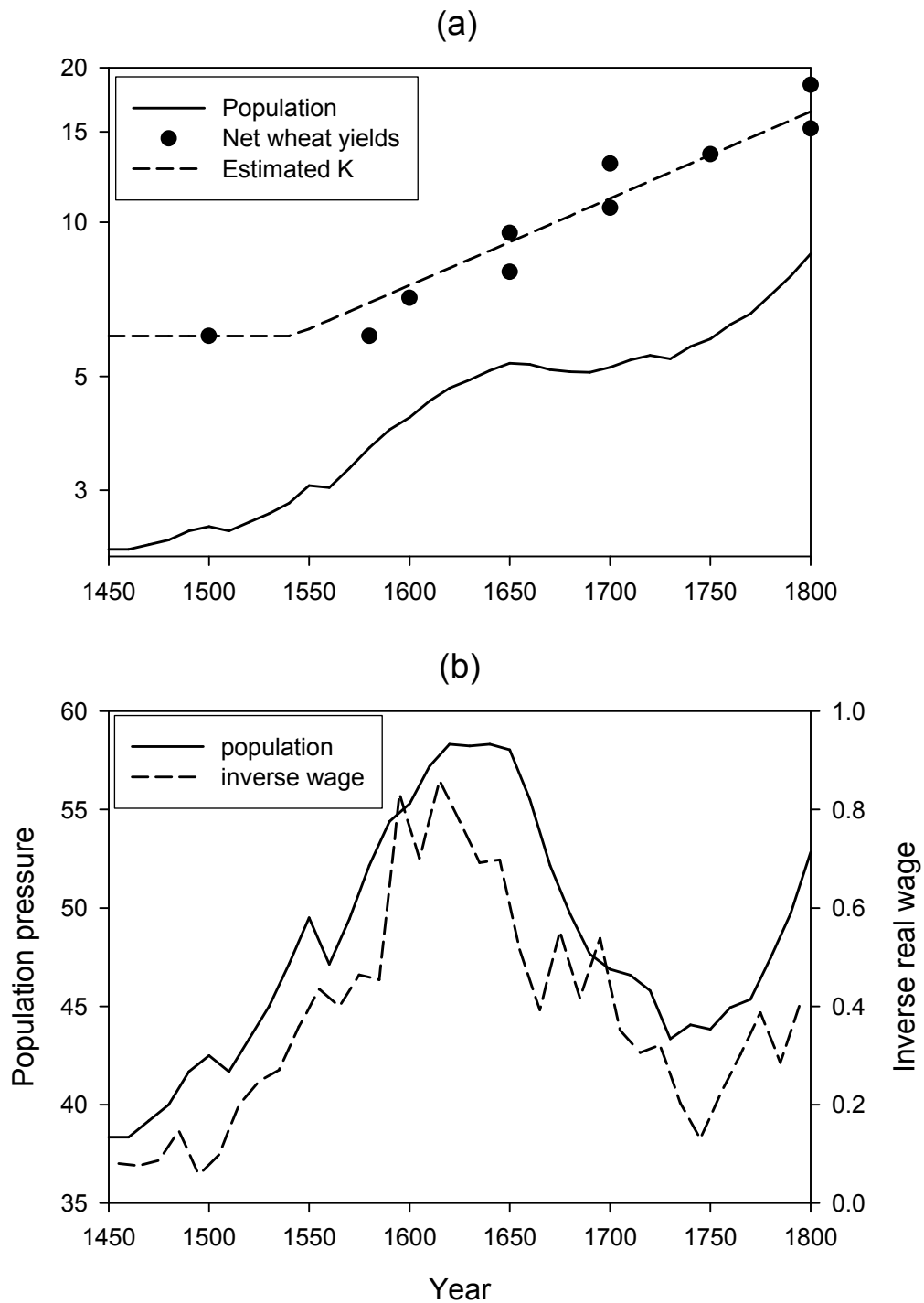


Figure 2. Detrending population trajectory for England. (a) Population numbers (in million), net yields (in bushels per acre), and the estimated carrying capacity (in million of people) in England from 1450 to 1800 (all variables plotted on a log-scale). (b) Detrended population (“population pressure”) trajectory (solid curve) and inverse real wages (broken curve).

We can obtain an approximate estimate of the carrying capacity by assuming that it was proportional to the net yield. Assuming the total potentially arable area of 12 mln acres (Grigg 1989) and that one individual (averaging over adults and children) needs a minimum of one quarter (8 bushels or 2.9 hectoliters) of grain per year, I calculated the carrying capacity of England shown by the broken line in Figure 2a (by coincidence 1 bushel of net yield per acre translates exactly into 1 million of carrying capacity).

We can now detrend the observed population numbers by dividing them by the estimated carrying capacity. The detrended population, which can also be thought of as “population pressure on resources,” is defined as  $N'(t) = N(t)/K(t)$ . Note that the critical assumption here is that  $K$  is proportional to the net yield,  $Y$ ; since  $Y$  is the only quantity varying with time in the formula, other components (total arable area, consumption minimum) being constant multipliers,  $K$  will wax and wane in step with  $Y$ . In other words, the exact values of constant multiples do not matter, since we are interested in relative changes of population pressure. Note that the estimate of  $K$  is based not on the area that was actually cultivated (this fluctuated up and down with population numbers), but on the potentially arable area. The latter quantity fluctuated little across the centuries (for example, as a result of some inundation of coastal areas during the Middle Ages or more recent reclamation using modern methods) and can be approximated with a constant without a serious loss of precision.

A test of the appropriateness of this detrending was obtained by regressing the estimated population pressure on real wages reported by Allen (2001). There was a very close inverse relationship between these two variables, and not a very good one if we were to use the non-detrended population numbers. As Figure 2b shows, population pressure and inverse real wage fluctuated virtually in perfect synchrony.

#### *England 1450-1800: Sociopolitical Instability Data*

For the period 1492-1800 I used the list of civil wars and rebellions compiled by Tilly (1993:Table 4.2). The list reports on revolutionary situations in all British polities. Since my focus is on England, I excluded all rebellions in Ireland, as well as in Scotland prior to the unification under the Stuarts. For the period prior to 1492, I used the compendium of Sorokin (1937: Appendix to Part III), which essentially added the data on the Wars of the Roses. The complete list is given in Table 1.

#### *Smoothing Socio-Political Instability*

I constructed an index of sociopolitical instability by assigning “1” to years with rebellion or civil war and “0” to years without (Boswell and Chase-Dunn 2000). To translate this discontinuous index into a smoothly varying one, I used the technique known as the kernel regression. The kernel regression is a nonparametric function estimator. The degree with which the estimated curve interpolates vs. smooths over a scatter in the data is determined by a single parameter  $h$  called the bandwidth (Härdle 1990). I used an exponentially weighted kernel (that is, the contribution of a data point to the smoothed point declines exponentially with distance between the two points). The choice  $h = 50$  was determined by a prior empirical observation (see Turchin 2003c: Section 9.1.2) that the times of trouble in European polities tended to be “lumpy” on a human generation scale. That is, revolutions and bouts of civil war tend to skip generations: if fathers participate in bitter internal fightings, their sons tend to value stability at almost any cost, while the grandsons exhibit a renewed willingness to revolt. Interesting as this pattern may be, it is a different phenomenon from the one we are investigating (the average periodicity of this

lumpiness is two human generations, or 50 years, compared to 200-300 year secular cycles). Using the bandwidth of 50 y smooths out any bigenerational cycles that may be present in the data, and allows us to focus on the secular oscillations (Figure 3).

**Table 1. Civil wars and rebellions in England 1450-1800.**

<i>Years</i>	<i>Description</i>
1455-6	The Wars of Roses: 1st phase
1460-5	The Wars of Roses: 2nd phase
1467-71	The Wars of Roses: 3rd phase
1483-5	The Wars of Roses: 4th phase
1495	Rebellion of Perkin Warbeck
1497	Insurrection in Cornwall
1536-7	Pilgrimage of Grace
1549	Kett's rebellion
1554	Wyatt's rebellion
1569	Rebellion of catholic lords of the North
1639-40	The Bishops' Wars
1642-7	Civil War
1648-51	Second Civil War
1655	Penruddock rising in Salisbury
1660	Monk's coup; restoration of James II
1666	Revolt of Scottish Covenanters
1679	Revolt of Scottish Covenanters
1685	Monmouth and Argyll rebellions
1687-92	Glorious Revolution, with intervention by France
1715-6	Jacobite rebellion in Scotland
1745-6	Scottish rising (Jacobite pretender)

#### *China: Population Data*

The situation with population data for China is complex. On one hand, the central authority in China (when it existed), conducted detailed censuses for tax purposes. On the other hand, corrupt or lazy officials often falsified or fabricated population data (Ho 1959). Conversion coefficients between the number of taxable households and the actual population are often unknown, and what is worse, these coefficients probably changed from dynasty to dynasty. The area controlled by the state also continually changed. Finally, it is often difficult to determine whether the number of taxable households declined during the times of trouble as a result of demographic change (death and emigration), or as a result of the state's failure to control and enumerate the subject population. Thus, there is a certain degree of controversy among the experts as to the precise levels that population numbers achieved at the highs and lows (Ho 1959, Durand 1960, Song et al. 1985). However, the controversy primarily concerns the absolute population levels, and there is a substantial degree of agreement on the *relative* changes in population density (which are, of course, of primary interests to a dynamical analysis).

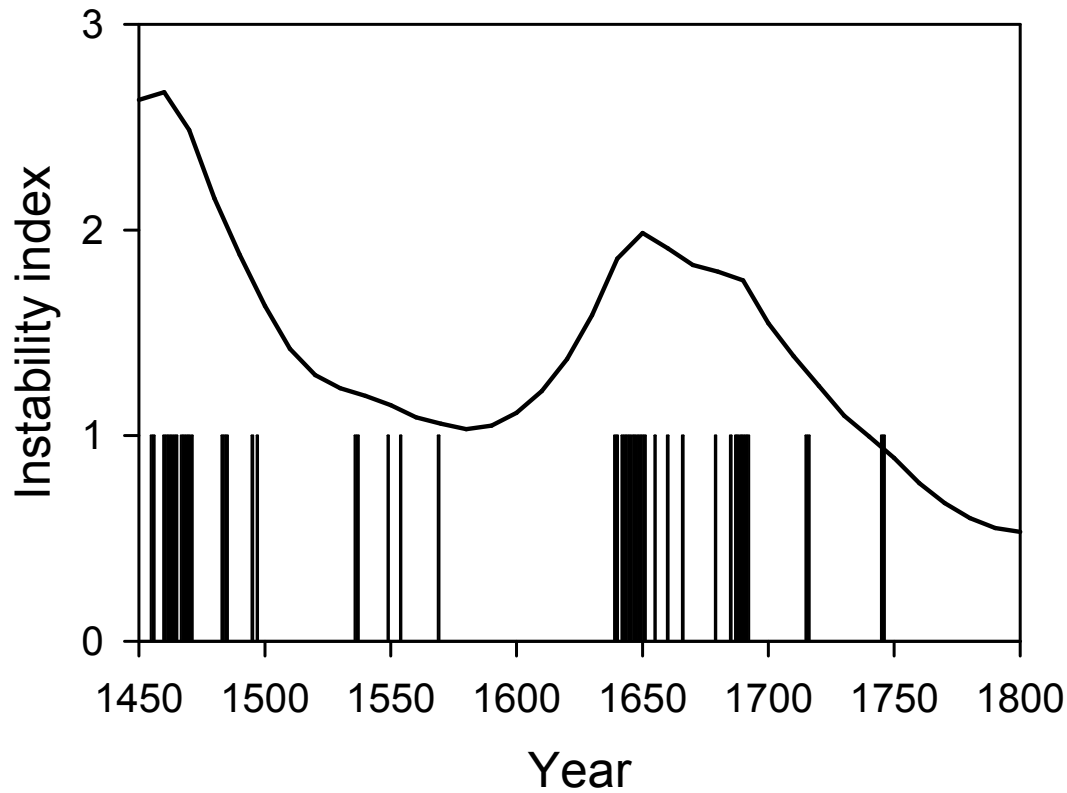


Figure 3. Sociopolitical instability in England, 1450-1800. Spikes: years in civil war or rebellion. Curve: yearly data smoothed by an exponential kernel with bandwidth of 50 y.

The most detailed trajectory of population dynamics in China known to me was published by Zhao and Xie (1988). These authors give estimates of Chinese population numbers at irregular time intervals. In order to make the data suitable for time-series analysis, I interpolated Zhao and Xie data using an exponential kernel with bandwidth of 10 years, and then subsampled the resulting smoothed trajectory at 10 year intervals. Setting  $h = 10$  y, same as the sampling interval, results in a minimal smoothing of the data.

After 1000 C.E. the trajectory becomes clearly nonstationary, and requires detrending. For this reason, I focus on the pre-1000 data (analysis of the second millennium data will be reported in Turchin and Nefedov 2006). After the fall of the East Han dynasty and before the Sui re-unification China was divided among a number of warring states. I excluded this period because the demographic-structural theory is state-centered (rather than focusing on state systems, as China was during the Han-Sui interregnum). This gave me two periods (using centuries as convenient break-points): 200 BCE – 300 CE and 600 – 1000 CE.

#### *China: Instability Data*

The index of sociopolitical instability in China comes from the remarkable publication by J. S. Lee (1931), who during the 1920s set out to calculate the frequency of internecine wars in Chinese history (ranging from fairly localized uprisings to wide-spread rebellions and civil wars).



For the period of interest to us (up to year 1000) Lee largely extracted his data from the Tih Wang Nien Piao by Chih Shao-nan. Checks with independent sources demonstrated the high accuracy of this source (Lee 1931:114). Lee presented the data as counts of internecine wars per 5-year interval. I smoothed his data using an exponential kernel with bandwidth  $h = 30$  y, and resampled the data at 10-y intervals. I reduced the bandwidth (compared to  $h = 50$  y used in the analysis of the English data) because the Chinese data did not appear to exhibit bi-generational cycles. In general, Chinese dynamics operated on a faster time scale, so using a bandwidth of 50 y would result in oversmoothing (however, when I redid the analyses with  $h = 50$  y the results were essentially the same).

#### *Rome: Population and Instability Data*

Population history of the Roman Republic and Empire remains a highly contentious topic (Scheidel 2001). Archaeological data, however, begins to throw light on this obscure aspect of Roman history. Recently Lewit (1991) integrated the results from numerous archaeological sites within the Western Empire and presented summaries indicating the proportion of archeological sites occupied in a 50-year period for Britain, Belgica, Northern and Southern Gaul, Northern and Southern Spain, and Italy. The data suggest that there were two periods of settlement expansion and two periods of settlement abandonment. I constructed a crude index of population dynamics by averaging provincial occupation curves.

Data on internal warfare in the Roman Empire was published by Sorokin (1937:Table 26). The data points are given for each 25-y interval. Smoothing the data using kernel with  $h = 50$  y reveals two periods of intense sociopolitical instability. One is the first century B.C.E., which was a period of transition between the Republic and Empire. During the first half of the Principate (the Early Roman Empire), after internal warfare subsided, population exhibited a long period of sustained growth. The population peak was achieved just before 200. The second period of instability occurred during the third century, when the Empire was convulsed by a series of internal wars, which were accompanied by population decline. Another period of stability and population growth occurred during the first half of the Dominate (the fourth century). After the decline and fall of the Roman Empire in the West, population decreased. Note that Sorokin's index of internal warfare underestimates the extent of actual sociopolitical instability during the fifth century, because he treated barbarian invasions as external warfare.

The population data are too crude to analyze using standard time-series methods (the main problem is the length of the sampling period, 50 years). Thus, I did not fit models with the population index as the dependent variable, but only used it as an independent variable in the analysis of sociopolitical instability. The population data were smoothed using an exponential kernel with bandwidth  $h = 30$  y, and resampled at 10-y intervals.

#### *Statistical Analysis: Regressions*

The conceptual framework of the statistical analysis is explained in the accompanying *Primer on Statistical Analysis of Dynamical Systems*.

Prior to analysis I log-transformed all data:  $X(t) = \log N(t)$  and  $Y(t) = \log W(t)$  where  $N(t)$  and  $W(t)$  are population and internal war (instability) data. As explained above, the English population data were detrended by calculating population pressure. English internal war data were also non-stationary (see Figure 3). I detrended instability data by calculating

$Y'(t) = Y(t) - (a_0 + a_1t)$ , where  $a_0$  and  $a_1$  are parameters of linear regression of  $Y(t)$  on  $t$ . (This is equivalent to dividing the untransformed data by the temporal trend, which is the same as the procedure used in detrending population.)

I fitted a simple time-series model to the data, the linear autoregressive process

$$X(t) = a_0 + a_1X(t - \tau) + a_2Y(t - \tau) + \varepsilon_t \quad \text{Model (1)}$$

(and an analogous Model (1) equation for  $Y(t)$ , reversing the definitions of  $X$  and  $Y$ ). Here  $a_i$  are parameters to be estimated, and  $\varepsilon_t$  is an error term, assumed to be normally distributed (Box and Jenkins 1976). The time delay was chosen as  $\tau = 30$  y, which approximates a human generation length. This particular time delay is also a reasonable choice that optimizes the tension between *redundancy* and *irrelevance* (see Turchin 2003c: Section 7.2.2). To check how my conclusions were affected by the specific value of the time lag, I fitted all models using an alternative choice of  $\tau = 20$  y, and obtained essentially same results.

As another check I fitted a model that used a quadratic polynomial:

$$X(t) = a_0 + a_1X(t - \tau) + a_2Y(t - \tau) + a_{11}[X(t - \tau)]^2 + a_{22}[Y(t - \tau)]^2 + a_{12}X(t - \tau)Y(t - \tau) + \varepsilon_t$$

The purpose of this model was to determine whether the process had a strong nonlinear component. There was statistical evidence for nonlinearity in some series, but using the quadratic model instead of the linear one (where it fit better) did not change any conclusions discussed below, so I do not report these results here.

To quantify any reciprocal effects of population and instability on each other I employed the stepwise regression. Thus, in order to estimate the effect of instability on population change, I first regressed  $X(t)$  on  $X(t - \tau)$ , and then tested whether adding the term  $Y(t - \tau)$  significantly reduced unexplained variance. The effect of population density on instability was investigated by regressing  $Y(t)$  on  $Y(t - \tau)$ , and then adding the term  $X(t - \tau)$ .

#### *Analysis: Cross-Validation*

The ultimate test of any model is its ability to predict independent data (data that were not used to develop the model and estimate its parameters). To assess the ability of the demographic-structural model to predict out-of-sample data I split each data set into two equal-sized parts. Then I fitted model (1) to the first half (the “fitting set”) and used the estimated coefficients to predict each data point in the second half (the “testing set”). Thus, the predicted population values,  $X^*$ , (the asterisk denotes prediction) were calculated as follows:

$$X(t)^* = a_0 + a_1X(t - \tau) + a_2Y(t - \tau)$$

where  $X(t - \tau)$  and  $Y(t - \tau)$  are the observed values of the independent variables in the *second half*, while the parameters  $a_0$ ,  $a_1$ , and  $a_2$  were estimated using data in the *first half*. The correspondence between the observed  $X(t)$  and predicted  $X(t)^*$  was assessed by linear correlation.

After using the first half of the data set to predict the second, I reversed the procedure and used the second half to predict the first. This procedure allowed me to use the complete data set

for testing the model performance. Finally, I repeated the complete procedure with the instability data ( $Y$ ).

One possible objection to the procedure outlined above is that there is some positive autocorrelation between  $X(t)$  and  $X(t-\tau)$  due to the time-series nature of the data, and it is conceivable that the excellent correlations between the observed  $X(t)$  and predicted  $X(t)^*$  are entirely due to this “inertial” effect. To eliminate this possibility, I redid the analyses with a different dependent variable,  $\Delta X(t) = X(t) - X(t-\tau)$ .  $\Delta X(t)$  is a measure of the rate of change, and by using it we break the autocorrelation arising from the time-series nature of the data. In fact,  $\Delta X(t)$  is none other than the realized per capita rate of population change, which is the standard dependent variable in the analyses of population data (Turchin 2003a). There can still be some predictive relationship between  $\Delta X(t)$  and  $X(t)$ , so we need to compare two alternative models:

$$\Delta X(t) = a_0 + a_1 X(t-\tau) + \varepsilon_t \quad \text{Model (2)}$$

which I call the *inertial* model (with an analogous Model (2) for  $Y(t)$ ), and

$$\Delta X(t) = a_0 + a_1 X(t-\tau) + a_2 Y(t-\tau) + \varepsilon_t \quad \text{Model (3)}$$

which I call the *interactive* model (with an analogous Model (3) for  $Y(t)$ ). The interactive model has an extra parameter, but in a cross-validation setting this does not matter (if the extra independent variable does not have a systematic influence on the dependent variable, then adding it to the model actually *decreases* to the ability of the model to predict out-of-sample data).

## RESULTS

### *England: 1450-1800*

Between the late fifteenth century and 1800 England went through several phases in which population growth and sociopolitical instability were in inverse relationship to each other (Figure 4a). There were two periods of endemic civil war (the Wars of the Roses of the late fifteenth century, and the revolutionary period of the seventeenth century) during which population stagnated or even declined. There were also two periods of internal stability (roughly, the sixteenth and the eighteenth centuries) during which population grew at a rapid pace. When plotted in the phase space, the trajectory moved in a cyclic manner (Figure 4b). Time-series analysis of these data provides strong evidence for reciprocal influences of population and instability on each other (Table 2). In fact, a simple linear time-series model (Model 1) explains a remarkable 85-93% of variance. Furthermore, the signs of the estimated coefficients (all highly statistically significant, see Table 2) correspond to those predicted by the theory: instability has a negative effect on population, while population has a positive effect on instability.

The strength of the effect of sociopolitical instability on population growth rate can be illustrated as follows. I took the data on the compound annual growth rate (CGR) of the English population from Table A.9.1 of Wrigley et al. (1997) and smoothed them using the 25-year running average suggested by Wrigley et al. Note that CGR is also known as the realized per capita rate of population growth; it is the most common measure of population growth used by population ecologists (Turchin 2003a). Plotting this measure of population growth rate against the instability index we observe a very tight relationship (Figure 5).

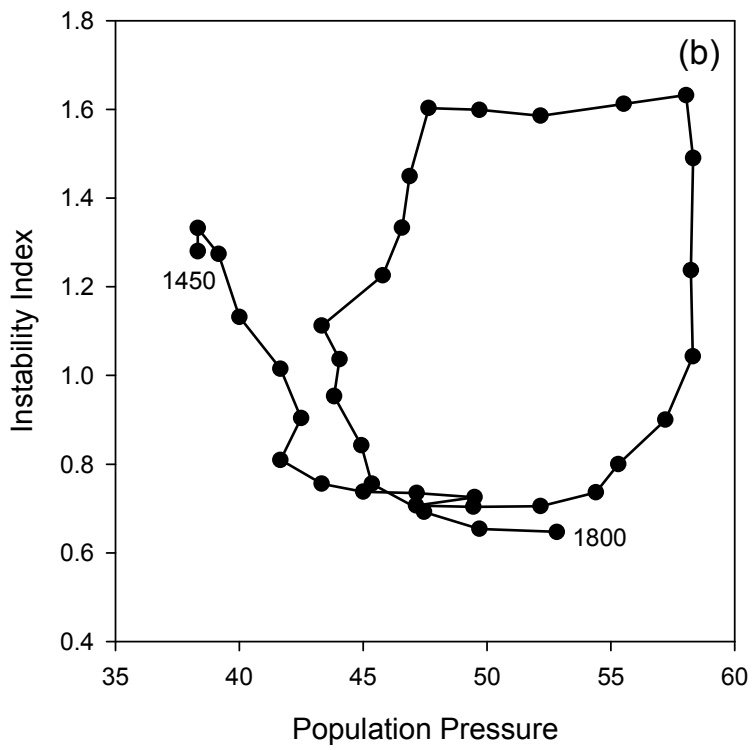
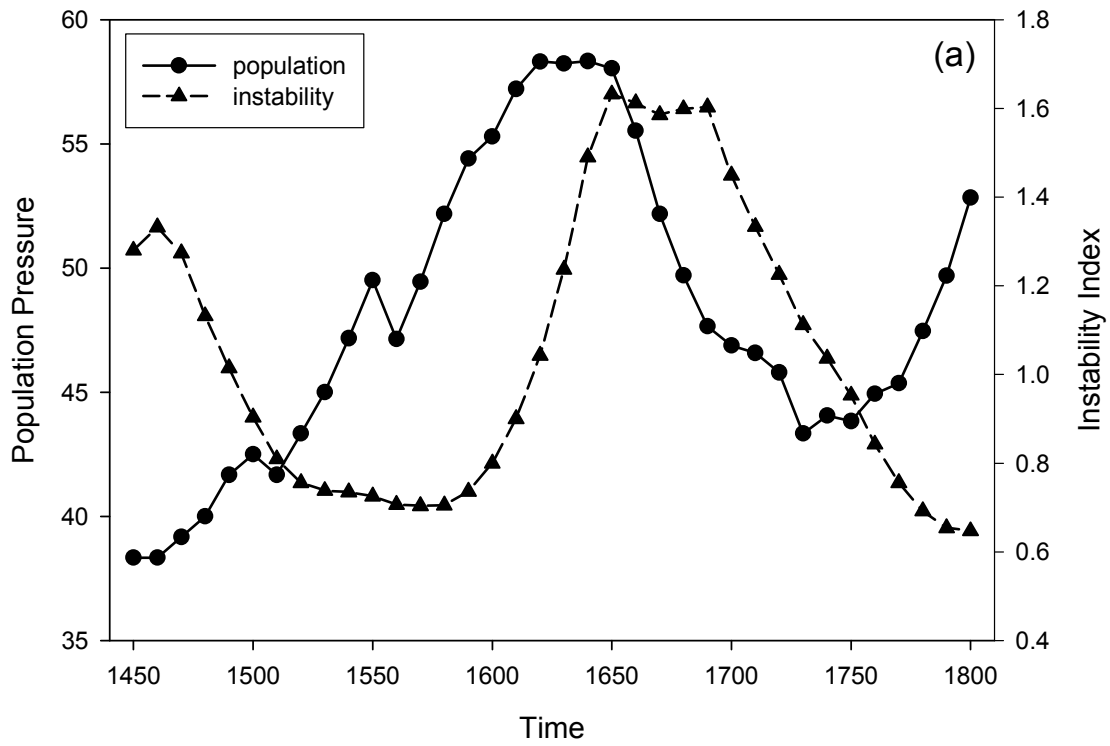


Figure 4. Dynamical interrelation between population pressure and instability index in England 1450-1800. (a) Time plot of the two variables. (b) The empirical trajectory in the phase space.

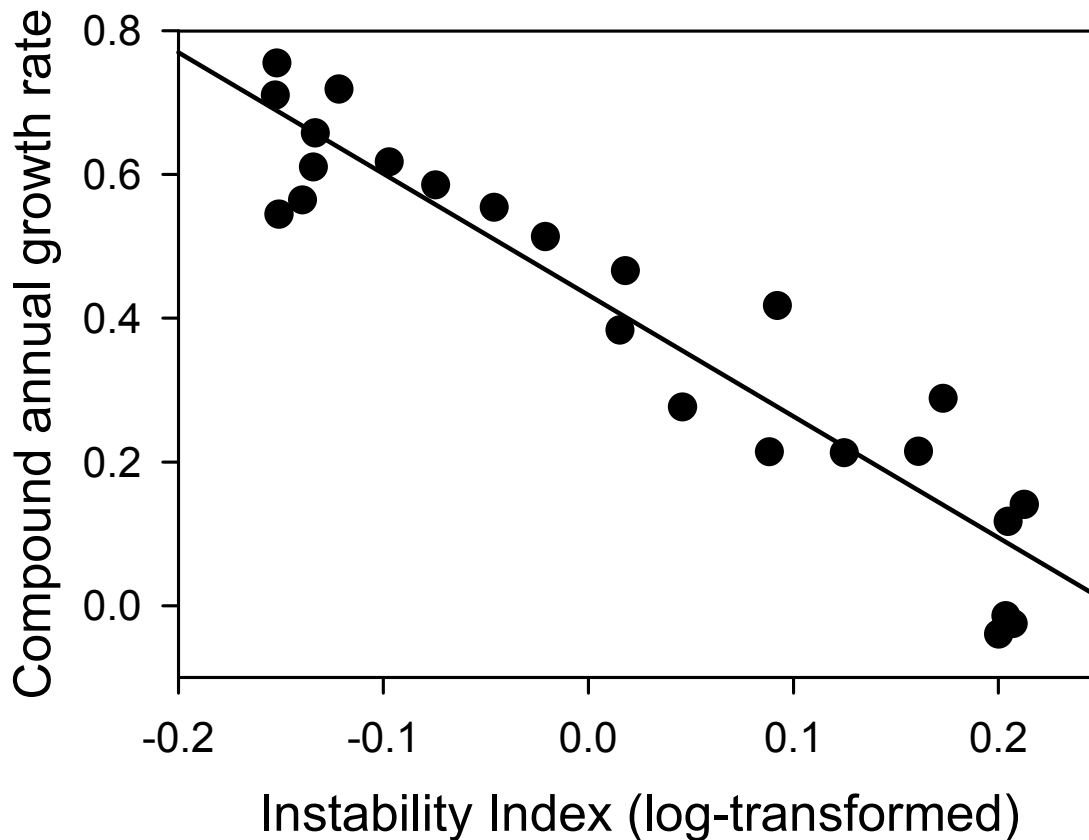


Figure 5. Relationship between the per capita rate of population growth and sociopolitical instability in England, 1540-1800.

#### *China*

The Chinese data (see Figure 6) are measured with much less accuracy, compared to the English ones, and time-series models with the same  $\tau = 30$  y resolve a smaller proportion of variance. Nevertheless, the coefficients of determination fall in the 0.4–0.8 range (Table 2), a very respectable result for what are quite imperfectly measured data. Coefficients associated with reciprocal feedbacks between population and instability are of correct sign and are all highly statistically significant (Table 2).

**Table 2. Results of regression analyses using Model (1). The regression correlations are negative for effects of instability on population and positive for effects of population on instability. The time lag is  $\tau = 30$  y in each of these regressions.**

Source of data	Dependent variable	$F$ -statistic for the reciprocal effect <sup>1</sup>	Regression $R^2$	Correlation between data and predictions	
				Fitted on 1st half, tested on 2nd half	Fitted on 2nd half, tested on 1st half
England	population	137.91***	0.93	0.90***	0.96***
England	instability	76.46***	0.85	0.92***	0.41*
Han China	population	15.55***	0.42	0.86***	0.52**
Han China	instability	36.47***	0.78	0.86***	0.79***
Tang China	population	50.48***	0.64	0.76***	0.84***
Tang China	instability	13.96***	0.63	0.80***	0.94***
Roman Empire	instability	8.04**	0.63	0.66**	<i>NS</i>

<sup>1</sup> $F$ -statistic of adding the  $Y(t-\tau)$  term to the  $X(t)$  regression in a stepwise fashion, and analogously for the  $Y(t)$  regression.

\* $P < 0.05$

\*\*  $P < 0.01$

\*\*\*  $P < 0.001$

#### *Rome*

Despite the inadequacies of the data (see *METHODS*), the qualitative dynamics of the variables are clearly consistent with the pattern predicted by the theory. Oscillations have a period of 2-3 centuries, instability lags in phase behind population (Figure 7). The effect of population index on instability with the same  $\tau = 30$  y is statistically significant (Table 2).

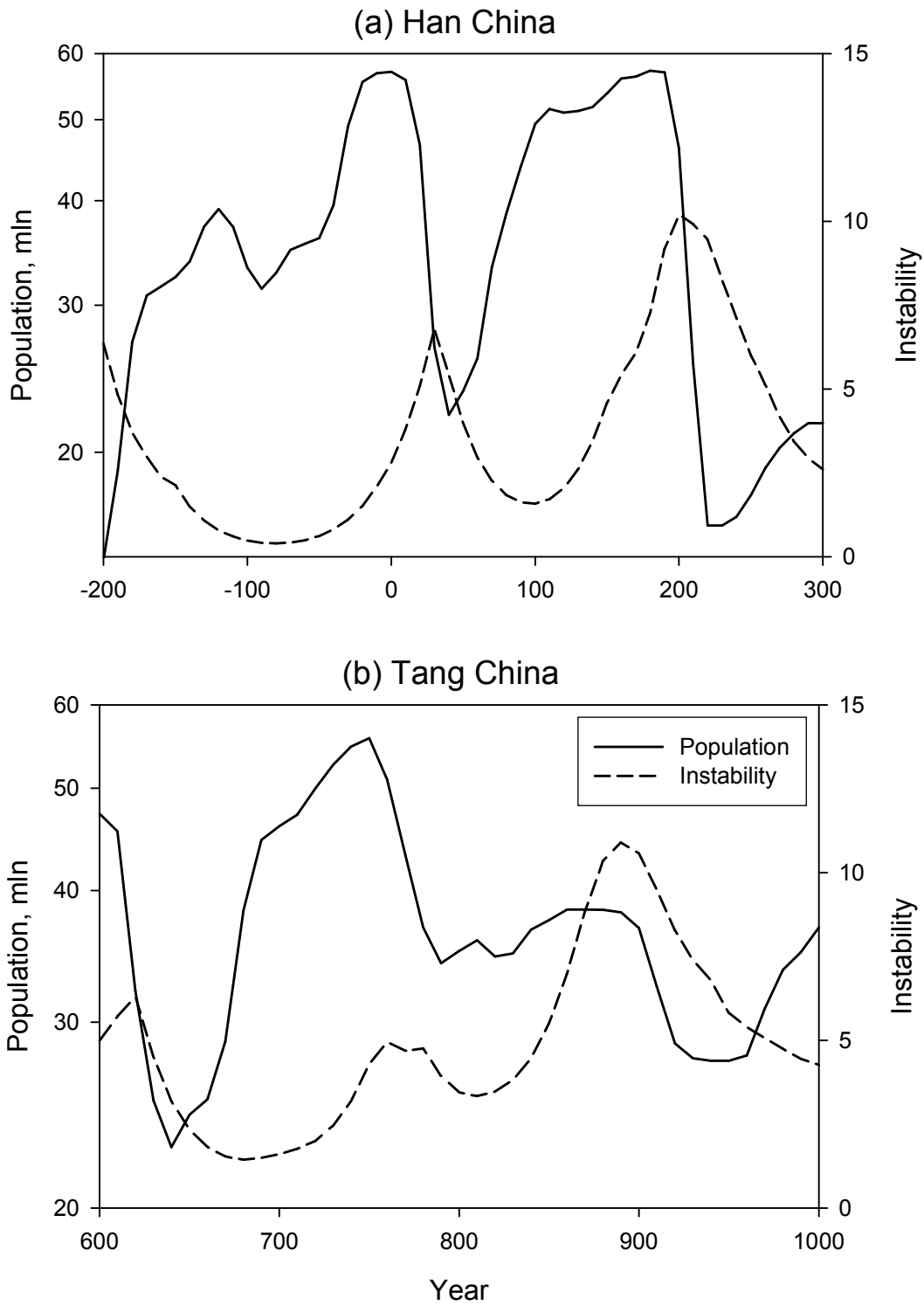


Figure 6. Dynamics of population (solid curve) and sociopolitical instability (broken curve) in China: (a) the Han period (200 BCE to 300 CE); (b) the Tang Period (600 – 1000 CE).

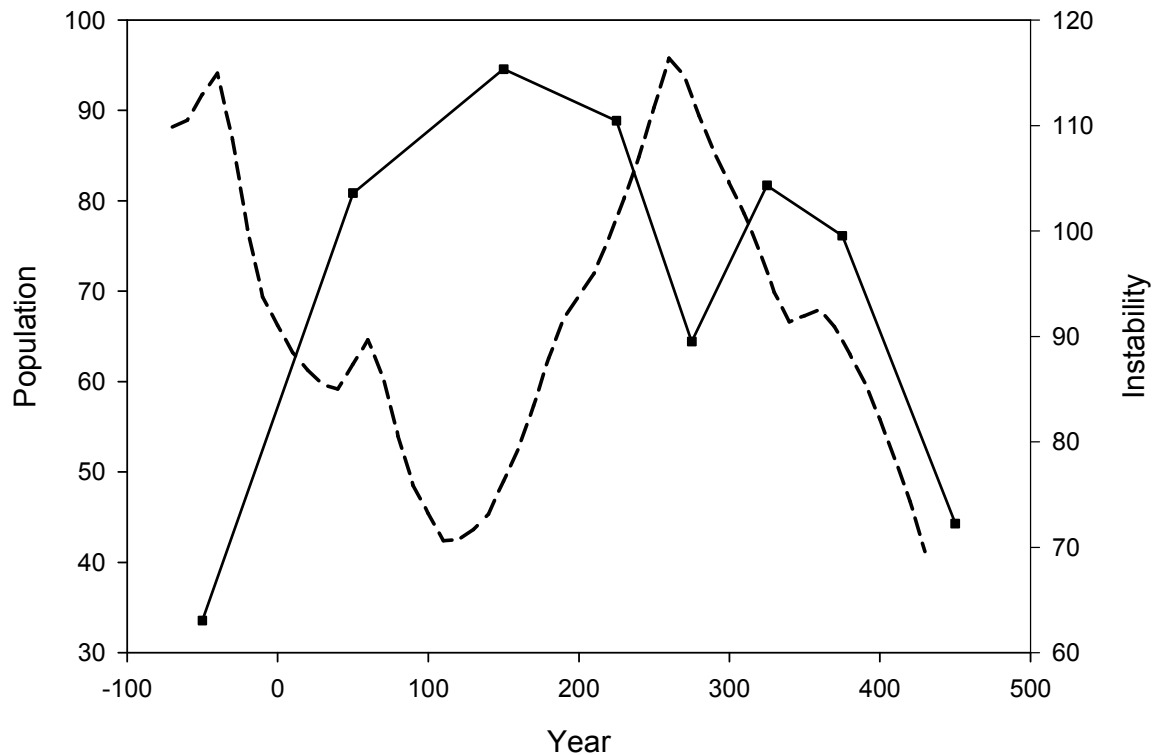


Figure 7. Dynamics of population (solid curve) and sociopolitical instability (broken curve) in the Roman Empire.

#### *Cross-validation results*

The population-instability model (1) was capable of making very accurate out-of-sample predictions (Table 2, last two columns). The correlation between model predictions and data was not statistically significant in only one case. In the majority of cases the correlations exceeded 0.8, and in some cases they were even greater than 0.9.

Ability to make accurate forecasts is not due to inertial dynamics of the observed variables (Table 3). The inertial model (2) does better than the interactive model (3) in only one case (indicated by italics in Table 3). In all other cases the prediction accuracy is substantially increased by using the interactive model. In fact, in half of the cases the correlation coefficient between the observations and predictions made by the inertial model is not significantly positive. In the English case the inertial model does so poorly that the correlations between the predictions and observations are actually negative (remember, we are *predicting*, not *fitting* data, and therefore negative correlations are possible).

In summary, knowledge of population dynamics significantly increases the ability to predict instability, and vice versa. This appears to be a very robust result, especially with respect to the best dataset, England. I checked its validity by analyzing with a variety of approaches: (1) detrended and not detrended; (2) smoothed strongly ( $h=50$  y), moderately ( $h=30$  y), and minimally ( $h=10$  y) or not at all; (3) using variables themselves ( $X$ ,  $Y$ ) or their rates of change ( $\Delta X$ ,  $\Delta Y$ ) as dependent variables; (4) fitting regression models or cross-validation; (5) using time



step  $\tau = 30, 20,$  or  $10$  y; and (6) using linear and quadratic forms of independent variables (to check for nonlinearities). In all cases, analyses suggested that instability is the dominant influence on population dynamics, and population is likewise on instability dynamics, although coefficients of determination (or prediction) and  $P$ -values varied depending on the approach employed.

**Table 3. Comparing out-of-sample predictive abilities of the inertial Model (2) and interactive Model (3). The time lag is  $\tau = 30$  y in each of these predictions.**

Source of data	Dependent variable	Correlation between predicted and observed			
		Fitted on 1st half, tested on 2nd half		Fitted on 2nd half, tested on 1st half	
		inertial	interactive	inertial	interactive
England	population	-0.57	0.94	-0.07	0.44
England	instability	-0.13	0.80	-0.53	0.89
Han China	population	0.45	0.57	0.73	0.48
Han China	instability	0.39	0.87	0.37	0.68
Tang China	population	0.56	0.80	0.61	0.90
Tang China	instability	0.57	0.78	0.66	0.92

## DISCUSSION

The best data set analyzed in this paper was for the early modern England. We are lucky to have quantitative data for many dynamical aspects of this social system (Turchin and Nefedov 2006), and some variables, such as population numbers, are quite accurately measured. The demographic and economic variables for this period have been ably analyzed by Wrigley and Schofield (1981), and by Lee (1973, 1985, Lee and Anderson 2002). These analytical results suggest that population dynamics plays the dominant role in setting the level of real wages. My analysis, focusing on “population pressure,” population numbers divided by the estimated carrying capacity, confirms this result (Figure 2b).

The feedback effect of real wage on vital rates is much more difficult to discern. Wrigley and Schofield thought that fertility responded to variation in real wages, but with a substantial lag time (variously estimated as 50 and 30 years). Lee (1985), however, made a very important point: if population numbers are driven in a cyclical fashion by some variable *other than real wage*, then we should expect a correlation between the real wage and the rate of population change (or its components, such as fertility) *shifted by one-quarter of the cycle period*, which in this case is around 50 years. For this reason, Lee concluded that fertility responds to some other factor than real wage. Both authorities agree that variation in mortality exhibits no discernible relation to the real wage.

Thus, Lee came to the conclusion “that extraneous variation in vital rates drove the system over this long swing.” Goldstone (1991) concurred, and suggested that the exogenous variables responsible for the population cycle could have been a combination of climate change and receding disease.

What I find particularly compelling in this discussion is the theoretical point made by Lee. Restating it in terms of the theory of nonlinear dynamics, let us suppose that we have two dynamic variables,  $X$  (in this case, population numbers) and  $Y$  (the inverse real wage). If the cycle is driven by the interaction between  $X$  and  $Y$ , then oscillations in  $Y$  must be shifted by about a quarter cycle with respect to  $X$ . If, on the other hand,  $Y$  responds to  $X$  without a time lag, that is, both variables oscillate synchronously (which is the case, see Figure 2b), then it is not the endogenous interaction between  $X$  and  $Y$  that drives the oscillations in the system; there must be some other variable  $Z$ . This basic fact is explained more fully in Turchin (2003b) in the context of predator-prey cycles. In that publication I also point out that a lag shift of one-half of the cycle period (which is what we get when using real wage instead of inverse real wage) is also inconsistent with the hypothesis of endogenously driven cycles.

My contribution to the debate is the suggestion that the “factor  $Z$ ” is sociopolitical instability. It is an exogenous variable in the theoretical framework of Wrigley, Schofield, and Lee, who have focused exclusively on demographic and economic variables. Ironically enough, it is also exogenous in the theory of Goldstone because, although he postulated and empirically supported the causal connection from population growth to sociopolitical instability, he did not close the causal loop. In the version of demographic-structural theory tested in this paper, instability is endogenized by postulating the feedback connection from instability to the rate of population growth.

From the point of view of nonlinear dynamics, real wage and sociopolitical instability respond to population numbers in fundamentally different ways (both in theory and in data). Real wage adjusts to population numbers essentially instantaneously (we can ignore a time lag of few years when cycles take two or three centuries to unfold). By contrast, population affects not instability itself, but its rate of change. Because instability is a slow dynamical variable, it takes a long time for it to reach its peak, leading to a phase shift between it and population numbers. Similarly, instability affects population via its growth rate (see Figure 5). Endogenous oscillations in population and instability are, thus, generated by a classical dynamical mechanism, which also operates in such disparate phenomena as planetary orbits and predator-prey cycles.

### *General Conclusions*

Quantitative time-series analysis of several empirical case-studies, ranging spatially across the breadth of Eurasia and temporally over two millennia, suggests that the demographic-structural theory does an excellent job of capturing dynamic relationships between population dynamics and sociopolitical instability. Significantly, more precise data resulted in better-resolved relationships (as measured by the coefficients of determination). It should be stressed, however, that there is one special attribute of all case studies that I analyzed: they all are characterized by a high degree of “endogeneity”, and thus it should be easier to detect feedbacks between different variables interacting within these “low-dimensional” dynamical systems. For example, early modern England was largely insulated from other European states—by virtue of its insular position. By contrast, preliminary analysis of the medieval cycle in England reveals a much greater impact of exogenous forces: the effects of the Black Death on population dynamics, and of the cross-channel involvement in French affairs on the rise of instability. As was noted by Guy Bois (1985) export of the “surplus elites” to France during the Hundred Years War reduced social pressures for internal war in England. It is, thus, not surprising that as soon as the English were finally expelled from France (by 1450), England went into the convulsions of the Wars of the Roses. This analysis will be further pursued in Turchin and Nefedov (2006).

The Roman and Chinese empires were largely “closed” military-political systems by virtue of their size and lack of significant rivals. Even their “barbarians” can be thought of as an integral part of the system. An excellent case for this interpretation of the Chinese-nomad relationship is made by Barfield (1989).

In cases involving non-insular medium-sized states and empires, therefore, we would expect to find the relationship between population dynamics and sociopolitical instability to be partially obscured. Despite these caveats, one remarkable finding here was that strong dynamical feedbacks can be detected at all in the historical record. This result has broad implications for the study of history, suggesting that historical societies can be profitably analyzed as dynamical systems.

### ***References cited***

- Allen, R. C. 1992. *Enclosure and the yeoman*. Clarendon Press, Oxford, UK.
- Allen, R. C. 2001. The great divergence in European wages and prices from the Middle Ages to the First World War. *Explorations in Economic History* **38**:411-447.
- Barfield, T. J. 1989. *The perilous frontier: nomadic empires and China*. Blackwell, Oxford, UK.
- Bois, G. 1985. Against the neo-Malthusian orthodoxy. Pages 107-118 in T. H. Aston and C. H. E. Philpin, editors. *The Brenner debate: agrarian class structure and economic development in pre-industrial Europe*. Cambridge University Press, Cambridge.
- Boswell, T., and C. Chase-Dunn. 2000. *The spiral of capitalism and socialism: toward global democracy*. Lynne Rienner, Boulder, CO.
- Box, G. E. P., and G. M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control*. Holden Day, Oakland, CA.
- Durand, J. D. 1960. The population statistics of China, AD 2-1953. *Population Studies* **13**:209-256.
- Earle, T. 1997. *How chiefs came to power: the political economy of prehistory*. Stanford University Press, Stanford, CA.
- Gibbon, E. 1932. *The decline and fall of the Roman Empire*. Modern Library, New York.
- Goldstone, J. A. 1991. The causes of long-waves in early modern economic history. In, J. Mokyr, editor. *Research in Economic History. Supplement 6. The Vital One: Essays in Honor of Jonathan R. Hughes*. JAI Press, Greenwich, CT.
- Goldstone, J. A. 1991. *Revolution and rebellion in the Early Modern world*. University of California Press, Berkeley, CA.
- Grigg, D. 1989. *English agriculture: An historical perspective*. Blackwell, Oxford, UK.
- Härdle, W. 1990. *Applied nonparametric regression*. Cambridge University Press, Cambridge, UK.
- Hatcher, J. 1977. *Plague, population and the English economy: 1348-1530*. Macmillan, London, UK.
- Ho, P. T. 1959. *Studies on the Population of China: 1368-1953*. Harvard University Press, Cambridge, MA.
- Jones, A. H. M. 1964. *The later Roman Empire, 284-602: A social, economic, and administrative survey*. University of Oklahoma Press, Norman, OK.
- Kennedy, P. 1987. *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*. Random House, New York.
- Lee, J. S. 1931. The periodic recurrence of internecine wars in China. *The China Journal* (**1931: March-April**):111-163.

- Lee, R., and M. Anderson. 2002. Malthus in state space: Macro economic-demographic relations in English history, 1540 to 1870. *Journal of Population Economics* **15**:195-202.
- Lee, R. D. 1973. Population in preindustrial England: an econometric analysis. *Quarterly Journal of Economics* **87**:581-607.
- Lee, R. D. 1985. Population homeostasis and English demographic history. *Journal of Interdisciplinary History* **15**:635-660.
- Lewit, T. 1991. *Agricultural productivity in the Roman economy A.D. 200-400*. Tempus Reparatum, Oxford, UK.
- Overton, M. 1996. *Agricultural revolution in England*. Cambridge University Press, Cambridge.
- Scheidel, W., editor. 2001. *Debating Roman demography*. Brill, Leiden, Netherlands.
- Song, J., C. H. Tuan, and J. Y. Yu. 1985. *Population Control in China: Theory and Applications*. Praeger Scientific, New York.
- Sorokin, P. A. 1937. *Social and cultural dynamics. Vol. III. Fluctuations of social relationships, war, and revolution*. American Book Company, New York.
- Tainter, J. A. 1988. *The Collapse of Complex Societies*. Cambridge University Press, Cambridge, UK.
- Tilly, C. 1993. *European revolutions: 1492-1992*. Blackwell, Oxford, UK.
- Turchin, P. 2003a. *Complex Population Dynamics: A Theoretical/Empirical Synthesis*. Princeton University Press, Princeton, NJ.
- Turchin, P. 2003b. Evolution in population dynamics. *Nature* **424**:257-258.
- Turchin, P. 2003c. *Historical dynamics: Why states rise and fall*. Princeton University Press, Princeton, NJ.
- Turchin, P., and A. Korotayev. 2004. Relationship between population density and internal warfare in prestate societies.
- Turchin, P., and S. Nefedov. 2006. *Secular cycles*. Book manuscript.
- Wickham, C. 1981. *Early medieval Italy: Central power and local society, 400-1000*. Macmillan, London.
- Wrigley, E. A., R. S. Davis, J. E. Oeppen, and R. S. Schofield. 1997. *English population history from family reconstruction: 1580-1837*. Cambridge University Press, Cambridge, UK.
- Wrigley, E. A., and R. S. Schofield. 1981. *The Population History of England 1541-1871: A Reconstruction*. Harvard University Press, Cambridge, MA.
- Yoffee, N., and G. L. Cowgill, editors. 1988. *The Collapse of ancient states and civilizations*. University of Arizona Press, Tucson, AZ.
- Zhao, W., and S. Z. Xie. 1988. *Zhongguo ren kou shi: China Population History* (in Chinese). People's Publisher, Peking.