

Mathematical Requirements for Homeostasis in Prehistory

KENNETH W. WACHTER

University of California, Berkeley

24 September 1985

1 Homeostasis

If demographers could ask one question of the wisest of our prehistoric ancestors, they would surely ask, "How was the growth of your population kept so slow?" The average rate of growth in the numbers of our ancestors over a million years cannot by any stretch of the imagination have exceeded one thousandth of our own world rate of growth in the last two decades, and this low long-term average growth rate is the one undisputed empirical fact about the hunter-gatherer epoch in human demography.

Two kinds of explanations compete with each other. There are those who believe that average levels of death rates compared to average levels of birth rates themselves suffice to account for the slowness of growth. High mortality in the "state of nature" never gave humans' biological potential for rapid growth a chance. Opposing these views are those who believe in homeostasis. Humans' biological potential for rapid growth had many chances, but ever again came up against the limits of the carrying capacity of the environment exploitable by existing technologies. Natural regulatory mechanisms made growth rates respond to the density or size of population, so that average levels of birth and death rates are not the appropriate concepts for analysis, but should give place to elasticities of response defined with regard to changes in density or size. These two competing points of view, needless to say, have ideological concomitants for contemporary population controversies concerning the nature of limits to growth. Historically, the ghost of Thomas Hobbes confronts the ghost of Thomas Malthus. Today, exemplars of the two points of view are found recently respectively in Ester Boserup and Ansley Coale (Note 1).

Empirical evidence is in all respects indecisive. Must paleolithic mortality have been much worse than mortality under, say, Louis XIV? Skeletal

evidence argues yes, but it is fraught with selection biases (Note 2). Data on contemporary hunter-gatherers suggests no; it is carefully collected but of contestable relevance (Note 3). That humans' physiological potential for large families or long lifespans dates back to paleolithic times at all is only a working hypothesis (Note 4). On the other hand, the notion of carrying capacity as it might apply to humans is problematic (Note 5), and even in historical times the operation of Malthusian feedback is not easy to measure directly (Note 6).

These empirical uncertainties justify attention to any *a priori* arguments that might limit the range of possibilities. An important *a priori* argument for the necessity of homeostasis in prehistory runs as follows: The time spans of prehistory are very long. They are measured in tens of thousands of generations. Suppose, contrary to the assumption of homeostasis, that the process governing births and deaths at the family level had been independent of the number of families over such long time spans. Then the inescapable randomness in births and deaths and in the climatic and other determinants of birth and death rates would have cumulated, eventually imparting such random variability to population numbers as to have driven the human race to extinction or to incredible size. Thus, by *reductio ad absurdum*, one should conclude that population growth rates were not wholly independent of population numbers, and that homeostatic mechanisms in the broadest sense applied.

This article examines that argument and similar *a priori* arguments in the light of advances in the mathematical theories of population growth. Branching processes in varying environments and their associated martingales and diffusion approximations have seen rapid development in the last dozen years (Note 7). Asymptotics, questions of growth or extinction over very long periods of time tending toward infinity, are precisely the part of the theory most actively pursued, so the vast long course of prehistory offers itself as the application of this subject *par excellence*. However, it turns out that even 40,000 generations in the context of slow growth are not enough to make some of the asymptotic theorems apply, and formulas for actual calculation of probabilities and bounds have hardly been a chief product of the discipline so far. Thus cautious and intricate juggling and enhancement of existing theory are required for the results presented in this article. Future developments, it is to be hoped, will make more refined conclusions soon practicable.

The appeal of homeostasis as a way out is no absolute.

Perhaps some number like 0.000015 is no less natural than zero or than any other number. Put another way, the time it took the human race to reach into the millions, having survived, has to be something, and why not nearer a million than ten thousand or ten million years? Nonetheless, astronomers in the same predicament with the age of the universe have persisted in seeking in "inflationary" theories an explanation for the age of the universe fixed by the dimensionless curvature (Note 9), and demographers might similarly be disposed to seek in homeostasis an explanation for the age of the race. Unlike the argument from cumulated randomness, the argument about mean levels encroaches quickly on metaphysics rather than on statistics, and here it will not be pursued.

It is a matter of great complexity, definitional as well as empirical, to say when the human population began, and what numbers should be posited for it at what times (Note 10). The argument from cumulated randomness might be urged in any of several contexts. It might be applied to all hominids, separating from apes over five million years ago, encompassing the australopithicenes together with their successors. Or it might be urged for the genus *Homo*, from the appearance of *Homo habilis* around 2.3 million years ago onward. At the other extreme, it might be restricted to our own species *Homo sapiens*, dating from about 300,000 years ago. The context chosen here lies midway between extremes, encompassing the growth in numbers of the large-brained primate species from the advent of *Homo erectus*, as much perhaps as 1.2 million years ago, through the exclusively hunter-gatherer epoch of *Homo sapiens* down to the origins of agriculture among *Homo sapiens sapiens* in the Neolithic Transition about 8000 B.C.

The shorter the period, the more stringent the test of the argument, since a shorter period allows randomness less time to cumulate. A million years seems a safe underestimate for the temporal range of large-brained primates. By poetic license, all females of the genus will be called "women", and all models will be one-sex models restricted to females. A safe lower bound for the starting female population is certainly 1. The effect on the calculations of the choice of starting size larger than 1 is one theme of the following sections.

Size estimates in the midst of prehistory mainly derive from density multipliers based ultimately on modern hunters and gatherers, multipliers which are applied to areas believed inhabited from archaeological evidence. These

estimates thus build in certain assumptions of homeostasis, and the present study will not rely on them. They do, however, provide upper bounds for final population size on the eve of the Neolithic. Estimates of final population size range up to 15 million individuals (Note 11) with most writers favoring figures closer to 6 million individuals or 3 million women.

Thus one arbitrary but reasonable standard scenario, for which the results of the various models of random population growth will be compared, posits growth from at least 1 female to upwards of 3 million females, over a time frame of a million years consisting of 40,000 generations averaging 25 years each. The period ends at the Neolithic Transition when evidence of higher growth rates suggests a sweeping change in demographic patterns. This scenario implies a maximum growth ratio per generation of $m = 3000000^{1/40000}$, or a maximum yearly crude growth rate per year of

$$(1/1000000) \log(300000) = 0.000015.$$

These choices give one standard case. Readers will surely have other preferences, and the calculations to follow deserve to be repeated as these preferences are expressed.

2 Sources of Randomness

Randomness in population processes is of two kinds. There is specifically demographic randomness – when people are subject to fixed risks of conception or of dying, the occurrence of these events is still a matter of chance, the luck of the draw, so to speak, and family sizes vary. On top of this variability is the randomness that economists are prone to study – risks themselves vary haphazardly in response to circumstances and conditions in the community, area, and world, so that the rates that govern family sizes vary.

Random draws (given fixed rates) and random rates – the distinction is not a sharp one, but it is serviceable. Random draws dominate in small populations such as those studied by anthropologists and historical demographers. Standard deviations track the square root of population size, the same rule as governs random sampling, though the interrelationships between mothers and daughters and among others in a tribe or village violate the independence assumption behind random sampling and full statistical analysis

is complex (Note 12). Random rates overshadow random draws for groups of thousands or more. The random rate set is like a deck of cards chosen at random from a shelf of decks. The group consists of those who draw cards at random from the same deck. The standard deviation tracks the group size, not its square root, so that the proportional effect does not disappear even for large groups. Spencer as well as Lee (Note 13) has operationalized the distinction with preindustrial villages with a variety of sizes by plotting the standard deviation in birth counts against village size and separating linear from square root trends. Without such data, the problem is to specify what is the group that "draws from the same deck" in some respect, when the rates vary with local flood and drought, with huntable herds, with tribal wisdom, or with global climate. The problem will be an important one in Section 5.

3 Random Draws and the Extinction Threat

The chief long-run contributions from random draws compose the randomness modeled by Galton-Watson-Bienyame Branching Processes. Correlations between the risks affecting different families are neglected, and each new-born woman has chances of surviving and bearing no daughters, one daughter, two daughters, etc., independent of every other woman. Her daughters are, in the simple version, independently subject to the same probabilities, and each generation produces the next one independently in accordance with the same probability law. In more complicated versions the probability law varies by age or varies from generation to generation. In all versions, childbearing rates may vary from woman to woman, but propensities toward large or small families may not be inherited by descendants, as they may in the more complicated models mentioned in Section 6. If rates vary from group to group in ways that persist over generations, the model should be taken to apply to the group with the highest average rate of growth, which comes over a long time period to dominate the population. The key assumption is always the independence of different draws for completed family size.

The classic result of Branching Process Theory is the one that lends initial credence to the *reductio ad absurdum* argument for homeostasis, the result mentioned by Jagers in the quoted passage: With probability one, as n increases, the size of the n -th generation either converges to zero or increases

to infinity. Even if the expected mean growth rate is zero, there is zero probability of the population continuing to fluctuate within reasonable bounds. This result holds true when age structure is built into the model and, in all interesting cases, even when the probability law goes to growth corresponding to m near unity, starting with one original ancestress, the probability of non-extinction is close to $2(m-1)/V$. For m near 1, V should fall between 1 and 2, since variances in human completed family sizes are typically less than twice the means, and since higher V would demand less than one chance in three of any surviving daughters at all. With $V = 2$, the standard case for a million years of prehistory with $m = 1.000373$ demands a probability of ultimate extinction $q = .99926$. To keep the extinction chance as "small" as $99m = 1.005$, giving $4(10^8)$ women after 40,000 generations, or reaching 3 million women in a mere 75,000 years.

These overwhelming extinction probabilities, however, tell less than they seem to tell. Although the mean time for extinction is very late, and approaches infinity as m approaches unity, most extinctions happen early when the population is still at low size. There is a famous theorem (Note 14) that the distribution of time to extinction has probabilities for large times that come to match a geometric series. Unfortunately, in this case, that approximation is still poor after 5000 generations. The exact formulas for a very special shape of family size distribution, the reduced geometric distribution, turn out to give a much better approximation in general for m near 1. The probability that extinction has not yet happened by the n -th generation is close to

$$2(m-1)/(V(1-m^{-n}+2(m-1)m^{-n})).$$

Taking $V = 2$, it follows that extinction chances reach 99.9% after only 850 generations or 21,000 years. Conditional on extinction not yet having occurred, the expected population size then would be about 1372 women. Once a population reaches such a size, the probability of extinction, though never negligible, drops to tolerable levels.

The high extinction risk is wholly a consequence of starting with a single ancestress. Keeping $m = 1.000373$ and $V = 2$, the probability of non-extinction could be raised to 25% by starting with 771 women, or to 90% by starting with 6173, taking respectively a little more or a little less than half-a-million years to reach 3,000,000 from these starting sizes. This is an important lesson. The extinction threat built into the process of random

draws does not directly force a recourse to homeostasis. Instead, it forces the adoption of a different model for humanity's beginning period.

The meaning of an initial population size, of course, raises a host of conceptual questions about the definition of the human line. Underlying the demographic phenomena under study there is an evolving population of genes which would need to be modelled for a full account. But any genetic model for the processes of transformation and replacement that led to *Homo erectus* from *Homo habilis* and from *Homo erectus* to *Homo sapiens* and to us would involve so many complexities as to divert attention away from demographic randomness altogether.

Some of the relevant concepts are actually easier to formulate at the demographic level than at the deeper genetic level. Every daughter has one and only one mother. In a one-sex demographic model the set of ancestresses a million years back from whom today's women descend in the female line is sharply defined. The wider set of females a million years back sufficiently genetically similar to the actual ancestresses that, but for the vagaries of demographic chance, they could have had unexceptionably human descendants by today's standards, is a hazier conceptual construct. But the size of this set is less elusive than its nature. The calculations of the preceding paragraphs can be rephrased to show how big a stock of potential ancestresses a million years back with descendants subject to the same growth rates as the actual ancestresses would be required to keep the odds of extinction under control. The growth-rate endowment is partly a matter of selective advantage *vis a vis* a wider population of female hominids, partly a matter of the fortunes of mate choice, but whatever the sources, the genetic and environmental processes did produce some trajectory of growth rates, those growth rates do imply some level of cumulating randomness, and those levels of randomness are subject to evaluation.

Given a final population of three millions, a time span of a million years, and variance V of completed female family size close to 2, consider requiring that the probability of extinction from random draws fall below a value q . This requirement essentially fixes the (method of moments) estimates of the minimum expected number of actual ancestresses in the wholly female line, the minimum starting population size of potential ancestresses in the wholly female line, and the maximum long-term average growth rate. For 50 these estimates are less than one expected actual ancestress, 4221 potential an-

cestresses, and $m = 1.000164$. For 10 actual ancestresses, 18,001 potential ancestresses, and $m = 1.000155$. The lesson is that the extinction threat would recede if one were willing to assume that tens of thousands of the later representatives of *Homo habilis* were qualified in the appropriate sense, whatever it might be, to generate descendant *Hominines erecti*. Whether such a radically "polyphyletic" picture is tenable is a question that must be left to anthropological debate.

The other way out of the extinction threat is to allow a small number of progenitors but to posit higher growth rates during an initial period before safe sizes were attained. One can imagine a story in which the earliest human forbear or forbears escaped immediate extinction by finding themselves in an unusually favorable environment, a small safe and bountiful corner of the world, or a brief interlude of beneficently balanced sun and rain. Growth rates in the earliest generations could have been high enough to carry the species out of danger of extinction from the cumulative effects of random draws. As humans spread out to less favorable homes or the climate returned to its usual mixture of good or bad, average growth rates could have dropped down to their long-term modest levels, but sizes could have been by then sufficient to keep the chance of extinction low. In such an account, the early years of humanity would bear some resemblance to a story of Eden, revised to put the serpent's appearance some generations later in the drama.

Strictly speaking, this alternative involves a kind of homeostasis. Population growth rates decrease as population size is increasing. But this kind of homeostasis is not the usual kind, in which growth rates respond to size or density over a wide range of sizes, not just near size zero, the lower boundary of possible population size. It is conceivable that there are some advantages, for instance, the avoidance of war, which are only likely to be grasped when the whole interacting population is small. In this case, it makes sense to speak of "lower-boundary homeostasis" and distinguish it from ordinary homeostasis. In the other case, where a favorable niche might have happened to exist at the right time and place to give the race its chance, perhaps after many brief unsuccessful tries, the connection between low size and favorable rates is more fortuitous and less repeatable, and it does not seem right to speak of homeostasis at all.

The period during which the expected population size remains small is still the chief determinant of extinction due to random draws even when

family size patterns are allowed to vary deterministically from generation to generation. Athreya and Karlin (Note 15) have shown, to good approximation, that 100 generations over which rates dictate an expected size of 500 people contribute the same amount to the cumulative probability of extinction as 10,000 generations over which rates dictate expected sizes of 50,000. The calculations deal with expected sizes. Of course extinction is occurring when the actual size is dropping below the expected size due to unfavorable random draws. The order of high and low expected sizes is irrelevant in the calculation. If an ice age drove the expected total population down into the thousands late in prehistory, the extinction threat from random draws would be the same as in the early years. But it is not plausible to imagine fluctuating conditions so extreme as ever to drive expected sizes low enough to cause trouble after the early fragile period had been well passed. Thus the conclusion stands: the threat of extinction from random draws necessitates a special story for the earliest period, but does not require homeostatic mechanisms over most of prehistory.

4 Random Draws and the Explosion Threat

The other threat is the threat of explosion. Does avoiding extinction commit one to too high a probability that random draws would generate an implausibly large population? The calculations are more difficult, but the answer is no. As far as random draws with fixed rates are concerned, taking care of extinctions also takes care of explosions.

The question of explosion has two aspects, the chance of sizes much larger than the final size on the way to the final size, and the chance of a final size itself much larger than the average value predicted by the model. By "explosion" here is meant any increase of population beyond an upper limit of plausible population sizes for prehistory, say, 9 million women or more. It does not mean growth faster than exponential growth like the population explosion of more recent history. The amount by which the maximum of a branching process up to generation n is likely to exceed the size at time n is not a quantity for which evaluable expressions have yet been found. Dividing the size by its predicted average value produces the kind of process called a martingale, and an inequality of Pitman (Note 16) does apply, but because of the division by average size the bounds are tight in the early years when

average size is small and loose in later years when average size is big and tight bounds are most needed. Heyde and Brown's (Note 17) special Central Limit Theorem for branching processes has no guarantee of multivariate convergence fast enough to cover the maximum. The best approach seems to be a long-winded argument with the ordinary Central Limit Theorem and Berry-Esseen Theorem (Note 18). It gives a figure less than .001 for the probability that the maximum up to the 40,000-th generation exceeds the size at the 40,000th generation by more than 20,000, given $m = .000373$ and given that the size at the 40,000th generation is two million or more. Thus the randomness is nearly all absorbed into the randomness in final size. The random variation around the final size in the path approaching it for such high final sizes is nearly negligible.

The other aspect of explosion is the randomness in final size itself. The average final size has been prearranged to be 3,000,000. But it might turn out that rates which made the average equal 3,000,000 also gave disquieting high probabilities to much higher sizes. Indeed, the branching process starting with a single ancestress has this bad property, either dramatically or moderately depending on whether the average is unconditional or conditional on non-extinction. If the unconditional average equals 3,000,000, the average size of nonextinct populations is some eight billion (Note 19) If rates are readjusted to make the average size of non-extinct populations equal 3,000,000, the random distribution of final population size due to different random draws is very close to an exponential probability distribution, with a 5% chance of more than 9,000,000 women at the end (Note 20). Starting with a large enough initial population, 6173, to bring the extinction probability down below 10between calculations with unconditional and conditional averages. The final size distribution is close to a sum of exponential random variables each of mean 1,303,184. The number of terms in the sum corresponds to the number of starters whose families do not go extinct in the pure female line, and is a random number with a nearly Poisson distribution with mean 2.3021. Then the probability of random draws producing a final size of more than 9,000,000, with rates that produce an average size of 3,000,000, with 6173 starters, is only 4

In sum, the famous fault of branching process models, the choice between extinction and explosion and nothing in between, is, for practical purposes, a misconception. The fault is not the branching model, which allows for random draws without homeostatic checks and balances, but rather the arti-

ficial assumption of a tiny starting population size. The randomness is largely built in when size is still tiny, and if a special story is allowed to provide for growth to some thousands, the remaining randomness from random draws is not enough to cumulate to alarming compass even over tens of thousands of generations. Framed in terms of random draws, the kind of randomness branching processes incorporate, the *a priori* argument for the necessity of homeostasis fails.

5 Random Rates

The examination of the *a priori* argument for homeostasis in prehistory now turns to the randomness of random rates. For two reasons, calculations must be more tentative. First, unlike random draws, where the model itself determines the scale of the random fluctuations, random rates may fluctuate on any scale, and the appropriate scale has to be guessed from empirical considerations. Second, solutions rely on approximations using processes called diffusions, and no theorems yet exist specifying how close in practical cases these approximations come. Thus the bounds obtained are indicative, not rigorous, and rigorous bounds await further progress in the field.

There are three great issues in modelling random rates – first, the extent to which luck varies, second, the size of the group which shares the same good luck or bad luck, and, third, how long the same stretch of luck persists. The last issue is the deepest one. The models of this section assume the random influences on rates to be ephemeral and volatile. They are ephemeral inasmuch as fortunate groups are not supposed to hold onto their superior good fortune from generation to generation. They are volatile, inasmuch as the mix of fortunes is supposed to change quickly, from era to era, over eras that last a generation or a century, rather than a millenium or more. More complicated models, in the next section, take up the possibility of durable fortunes passed on to descendants and only eventually abating, as well as the possibility of slow ground swells of change.

The second issue, how large a group shares the same luck, is another crucial issue, but it is often neglected. Models often assume immediately that luck is ubiquitous. They assume, for instance, that all members of the population, whether a hundred or a million strong, share the same lucky

or unlucky global climate. The simplest model with randomness introduced entirely on a global scale is the random walk. At the other extreme, a model in which randomness enters entirely at the local level is equivalent to the branching process of Sections 2 and 3 with a higher value for the variance of family size. Both global and local random rates can be combined into one model, using diffusion approximations, and these alternatives will be treated in turn.

A random walk model assumes that the crude growth rate is chosen each era (generation or century) by a random coin toss with a biased coin, a coin slightly weighted, in this case toward heads. The crude growth rate takes on only two values, one positive (for heads), the other negative by the same amount (for tails). Because this fixed variance is imparted to the rate of growth rather than to the numbers gained or lost, it is the logarithm of population size rather than the size itself that describes a random walk. The logarithm takes an upward or a downward step each era by the amount of the crude growth rate for the era. The variance equals the square of the step size times $2p(1-p)$, where p is the probability of heads.

A starting population of 6,000 women has its logarithm poised about half-way between the logarithms for populations of one woman (the brink of extinction) and 3,000,000 women. With no bias towards heads, that is, with an average growth rate of zero, the odds of extinction before reaching three million would be close to 50-50. With non-zero bias, much the same remains true as long as the variance is well above the bias. To keep the chances of non-extinction up to 90% final target size on a logarithmic scale, that is, a starting size of over 600,000 members (Note 21). The strategy of raising starting size to negate the extinction threat worked well with random draws, but it turns out to be untenable with significant global randomness in rates. The starting sizes required are too high. The alternative is to raise the long-term average growth rate, reflected in the bias of the coin. Higher average growth, however, implies shorter expected times to reach target size, and therefore less tenable models for prehistory.

To keep the probability of extinction down to 10% requires enough bias in the coin to make the average growth rate per era times the logarithm of starting size about equal to the variance, that is to say, about equal to the square of the size of each step. The expected time it takes to reach one boundary or the other, either extinction or the target total, is roughly

inversely proportional to the variance and directly proportional to the length of the era, the stretch of time over which the luck determined by one coin toss persists. Thus the threat of reaching high sizes in an implausibly short time is exacerbated about as much by halving the length of the era as by doubling the variance or scaling up the standard deviation of the rates or the step size of the random walk by the square root of two.

These tradeoffs imply that ubiquitous volatile random rates give a population considerable potential for explosion, though it is difficult to know what sorts of figures for the variance and the era length are appropriate. To keep the expected time to reach either 9,000,000 or extinction up to the figure of 20,000 generations would require a coin biased toward positive outcomes with odds of no more than .50289 to .49711 and a crude growth rate fluctuating no more widely than between $-.062$ and $+.062$ per generation, values which correspond to rates of $-.0025$ and $+.0025$ a year. With a standard deviation of .08 and eras of a generation each, keeping down the extinction chance to 10time to reach 3,000,000 from 6,000 would be only 6192 generations or 154,799 years. These times are short enough to make the model start to seem unrealistic, if the level of global variability incorporated is at all appropriate.

Are standard deviations on the order of .06 or .08 per generation conceivable for global variability in random rates? For England between 1541 and 1811, the standard deviation seems to be about 0.08 (Note 22). The English rates were probably themselves subject to homeostatic control, and might have varied more widely without it. Furthermore, the lives of hunters and gatherers must have been much less predictable than those of English villagers. At first glance, then, it does look as if the randomness of global random rates would be sufficient to drive prehistoric populations to extinction or explosion in the absence of homeostatic control.

The flaw in this argument so far is the assumption that the random rates are ubiquitous, that prehistoric populations spreading over Africa, Europe, and Asia shared the same stretches of good times and of bad times. English villagers were all affected to some degree by a national market and threatened by many of the same waves of plague and infection, as well as by the shared vicissitudes of English weather, whereas the food and water, predators, and prey of hunters and gatherers in distant areas must have been largely independent. Widespread epidemics were probably unknown. What the whole

population shared was mainly not even local weather, but global climate, with its slower- changing pace. Thus it makes more sense to attribute most of the variability in the growth rates of prehistoric population to local bands or groups, and to attribute only a small portion of the variance to global influences shared by all (Note 23).

Distinguishing between local and global components of variance in demographic rates means relaxing the assumption that rates are ubiquitous. Completed net female family size can be modeled by adding together an individual random component , with variance V_I plus a random tendency toward larger or smaller surviving families shared with the L members of a local group, a tendency whose level varies with a variance V_L , plus a tendency shared by the whole population of size x , consisting of x/L independent local groups, with a small global variance V_G . The overall variance in size next generation as a function of size x in a given generation is

$$xV_I + (x/L)L^2V_L + x^2V_G$$

or

$$(V_I + L V_L) x + (V_G) x^2 = B x + Gx^2$$

A hierarchy of components for local groupings in a hierarchy of sizes L can also be averaged, with the same result that the variance is the sum of a moderate multiple B of size itself, accounting for basic variability, plus a tiny multiple G of the square of size, accounting for global variability. More elaborate models are possible, involving correlations as functions of distance together with spatial distributions for expanding populations in expanding areas, such as Ammerman and Cavalli-Sforza (Note 24) have modeled for the later, Neolithic period. But the simpler quadratic model for variances seems sufficient for present needs.

The variance function is utilized with a diffusion approximation for branching processes in random environments, since the exact theory for times to extinction and explosion is not yet sufficiently detailed. The diffusion approximation, first proposed by Feller (Note 25) takes as units not people but clumps of people and as time intervals not generations but clumps of generations. It assumes both kinds of clumps to be close to infinite size, compromising the practical value of the theory. Even though the technical literature is silent on the problem, investigation suggests that with low average growth rates the diffusion approximation is already useful for bands of

dozens and more, except for the detailed behavior of populations hovering on the brink of extinction.

Diffusion theory derives its name from its applicability to the spatial diffusion of particles and creatures, but its use here has nothing to do with the spatial diffusion of populations, but rather to random spread in population numbers. Widely used in genetics, diffusion theory provides a formula for the odds of reaching a target size or upper boundary b before extinction, starting from size x . It also provides a differential equation which leads ultimately to probability distributions for extinction time or for the time of reaching target size, given variance coefficients B and G in the formula for variance as a function of size and squared size (Note 26). In applying the approximation, it is necessary to choose whether to use the observed long-term growth rate as an estimate of the expectation of the growth rate of population in the model or as an estimate of the growth rate of expected population in the model. The difference between the two has been a recent issue in the biological literature (Note 27) and corresponds here to the choice between Stratonovich solutions and Ito solutions to a stochastic differential equation (Note 28). The former alternative is chosen here, in accordance with a maximum likelihood criterion for estimation.

The larger the variance coefficients B and G , the larger is the starting population size required to keep the chance of extinction before reaching target size down to, say, 10%. The larger the starting size, the shorter is the stretch of prehistory to which the models apply, and, beyond a certain point, the more dubious their relevance. Exhibit Note 1 shows the bounds on B and G imposed by requiring no more than 10% no more than 20,000, and no more than 60,000, when the long-term expected growth rate is held down to the standard value of .000373 per generation. The ratio of the global component of variance to the expected growth rate is obviously very tightly constrained. For starting sizes below 20,000 and for $B = 4$, which corresponds to as much local variance as individual-level variance, the global variance could not exceed three-quarters of the growth rate. For starting sizes below 60,000, it could not exceed about thrice the growth rate. These are tight bounds.

Although the constraints on global variability are severe, they are not so severe as to rule out altogether the plausibility of models without homeostasis, given the present state of ignorance as to likely levels of global variability. Exhibit Note 2 shows the standard deviations in net reproduction rates when

populations attain various sizes, for pairs of B and G values permitted under the starting sizes considered in Exhibit 1 with nine to one odds of non-extinction. These are not really outside the range that would be plausible with moderate downward revision of the English figures to allow for hunters and gatherers experiencing less ubiquitous random rates.

Each combination of variance coefficients B and G with a growth rate and a starting size implies a different time to reach target size. What is easiest to calculate is the expected time to reach either target size or extinction, whichever is earlier. The expected time to reach target size conditional on non-extinction is likely to be longer, since extinction would tend to occur early. One might speculate that the random effects would drive the expected times down, and make the models less plausible, but this effect turns out not to be serious. With starting sizes of 20,000, the period remaining to reach 3,000,000 without randomness would be 12,095 generations, or a bit less than a third of the whole 40,000 generations. In this case, the random effects reduce the expected time to target size or extinction only to 11,837 when $B = 1$ and actually increase it to 13,262 when $B = 6$, given the largest G compatible with 10

In summary, it appears that it is not easy but not impossible to counter the threat of extinctions from random rates by restricting the model to a period after a certain minimum size has been attained. It is not nearly so easy with random rates as with random draws. But it is not yet out of the questions. Further research on the local and global components of randomness in historical populations is urgently needed, not only for its own sake, but also to give a better basis for extrapolation back to prehistoric settings. It may be that when this information is garnered, the permitted levels in Exhibit Note 1 will appear distinctly low, and the *a priori* argument for homeostasis will gain cogency. But at the present state of knowledge, as a conclusive argument, it fails.

6 Persistent Fortune

The calculations of Section 5, while relaxing the assumption of ubiquitous rates, maintain the assumption that fortune is ephemeral. How do the conclusions fare under models which allow groups with endowments that lead

to lower or higher than average growth rates to pass on some part of their differential fortunes to their descendants?

Fortunes which persist so long as to constitute a practically permanent selective advantage or disadvantage require no change in the analysis. The group with the largest growth rate comes to dominate the whole population, and, as mentioned in Section 3, calculations can proceed with the rates pertaining to this group. Temporarily persistent fortunes differ in their effects, depending on what determines the time during which advantages abate. Two extreme cases come immediately to mind. One extreme case postulates luck being dealt out to each matriarch and her descendants, to persist for a fixed number of generations, until a new deal reshuffles luck among all living individuals. Then the process repeats, with luck descending to the same number of generations, till a new reshuffling. This picture is an extreme version of what, more realistically, would be gradual processes, in which advantageous habits and technologies would be only imperfectly transmitted over each generational juncture or would only remain advantageous while a particular regime of environmental conditions prevailed. The key formal feature in this first picture is the independence of the process of persistence from the growth of groups or clans. The other extreme case postulates fortunes persisting until groups grow nearly extinct, and have to merge, or reach a target size and have to fission. In this picture, population mixing and dispersal would be the driving force behind the decay of runs of luck.

The picture with periodic new deals implies a slight tightening of the constraints derived in Section 5. This outcome is easiest to see if it is assumed that local randomness dominates global randomness at small sizes. Then the experience of a single clan during a single deal can be treated as a branching process with its own, randomly selected generation ratio m (Note 29). If the deal lasts long enough for the exponential growth of families dealt high values of m to dominate, then the expectation of the number of descendants, averaging over the randomness in m , could be much higher than it would be if all families shared the average m . The experience from deal to deal can be modeled as another branching process, with this higher expectation value as its value of m . It is this higher value that must match the observed long-term growth rate, so the contribution of m in the formula for the probability of non-extinction is unchanged, and the only change is an increase in the variance of completed female clan size, and a consequent mild increase in extinction probabilities. It is interesting to note, however, that a positive

long-term growth rate is compatible, under this model of persisting luck, with zero or negative short term average growth rates, a feature that might bear on the deterministic argument for homeostasis mentioned at the end of Section 1.

The other picture, with size-dependent reshuffling of luck, calls for detailed modeling of fission and fusion, which cannot be included here. In general, since the compounding effect of exponential growth for fast-growing clans is curtailed by fission, extinction rates and overall growth rates fall in between the case of ephemeral random rates treated in Section 5 and the case of periodic new deals. The most interesting aspect of this picture is the difference in extinction probabilities and long-term growth rates that can be produced by different splitting points and different splitting patterns, when both local and global randomness contribute. Populations which split into two groups of 500 when reaching 1000, and are endowed with new randomly chosen rates, differ from populations which split into two groups of 1000 when reaching 2000, and from populations which split into 500 and 1500 when reaching 2000. The interest of these patterns, which await treatment in a later study, lies in the mechanism for growth rates responding very slightly but still systematically to change in the typical size of settlements and their splitting propensities, changes that are easy to imagine accompanying the Neolithic Transition which ends the period under study here (Note 30).

7 The Early Years

This investigation of the argument from cumulating randomness to the necessity of homeostasis has found surprisingly little support for it as an argument about the long course of hominid prehistory. But it can also be read as offering strong support for it as an argument about the early stages of population growth before numbers reached substantial size. It is clear that if homeostasis is not to govern populations in the later periods, a very different story about rates and mechanisms of growth is required to avoid extinction in the early years.

The threat of extinction even from random draws, and all the more from the combined effects of random draws and random rates, is very severe in the early years. This severity is evident in the models used which have assumed a

constant long-term average growth rate for the whole of pre-Neolithic times. It is more usual to assume that rates of growth tended to accelerate during prehistory, in tandem with rising population and developing technology and social organization, either gradually or in sporadic bursts (Note 31). Gradual acceleration would imply that the early growth rates were even lower than those used here, and that the population remained smaller and even more direly threatened by extinction for an even longer period. Some emphatic mechanism is needed to explain why early extinction was eluded.

It seems as if (contrary in a broad sense to the picture of accelerating growth) average growth rates of families when the global or local population was tiny must have been higher than the long-term total average observed. This consideration harkens back to the reflections on lower-boundary homeostasis in Section 3. One alternative is to suppose that *Homo erectus* got its start, perhaps after many false starts, in a temporary timespan or small area particularly favorable to its survival. The other alternative is to suppose that very small populations enjoyed advantages mitigated fairly abruptly by modest population increase, a type of homeostasis, to be sure, but not the type most often held in mind.

The advantages of low size need to be mitigated abruptly, for the growth rates required to avoid extinctions from random draws while growing from, say, a hundred females to a thousand, are high enough to take the thousand to a million in as few as 8000 years. Such a deceleration in growth rates itself could not persist, however, since that would quickly take the average growth rates to levels below replacement. These arguments strongly suggest that the demographic mechanisms governing population growth at very low population sizes, locally or globally, must appear very different from the demography of populations at moderate size.

One answer is, of course, the classic Malthusian account: The growth potential of modern hominids far exceeded the rate at which new resources could be brought within its purview. Population quickly rose to levels where the limitations of resources kept it in check, and then rose slowly, tracking the expansion of its resource base. There was a golden age, an interim to be fruitful and multiply, but the golden age was brief. None of the calculations in this paper have anything to say against that account. They only argue against the necessity of adopting it.

Another answer would seem to be one that emphasizes the higher growth

rates for very small sizes, either from intrinsic advantages or from some flexibility in intentional responses to the threat of extinction. The elasticity of response of growth to population size might be strongly positive at very low sizes, strongly negative at very high sizes, but growth and size or growth and density might be essentially decoupled over most of the range of hominid experience. The investigation in this paper suggests that random effects would not cumulate even over long periods to negate this alternative possibility.

A priori arguments of the kind analyzed here are only worth consideration when empirical knowledge is as sketchy as empirical knowledge of human prehistory continues to be. The general models used here cannot supply fine-grained information about the past, but they have served to bring to the fore issues which deserve more systematic consideration, both in the study of prehistory and in the study of modern population processes as they bear on possibilities for the distant past. Malthus himself seems to begin his essay with a belief that his principle of population had something like *a priori* status. That claim has not been generally accepted. Nor does the more quantifiable *a priori* argument for Malthusian homeostasis in the context of prehistory analyzed in this paper merit acceptance. At its strongest, that argument supports a much more general form of Malthusianism than Malthus went on to propound; it supports a theory about boundaries rather than about normal tendencies. But it may be that a homeostatic account of population at the boundaries is not a foolish nor a vacuous alternative, either as a theory of prehistory or as a theory of history itself.

Notes

The ideas presented here have been developed jointly in continual interchange with Ronald D. Lee, partly in agreement and partly in disagreement with his views, and I am greatly indebted to him. I also thank the National Institutes of Health who have supported this work under Grant R01-08-R1HD18107. I thank Carl Boe and Farzaneh Roudi for aid with calculations and Barbara Parrish for aid with word processing. I have used Fortran notation for powers and products: 2^{**7} means two raised to the seventh power, and $2*2$ is four.

1. Ester Boserup (1981) *Population and Technological Change*, University of Chicago Press, Chicago, especially pages 35-36.

2. Cf. J. Bocquet-Appel and C. Masset(1982) "Farewell to Paleodemography", *Journal of Human Evolution* 11, 321-333, and the reply by Jane Buikstra and L. Konigsberg (1985) "Paleodemography: Critiques and Controversies", *American Anthropologist* 87, 316-333.

3. Nancy Howell (1979) *Demography of the Dobe !Kung* Academic Press, New York, Chapters 4 and 5.

4. Nancy Howell (1976) "Toward a Uniformitarian Theory of Human Paleodemography" in R. Ward and K Weiss, eds., *The Demographic Evolution of Human Populations* Academic Press, New York.

5. Brian Haydon (1975) "The Carrying Capacity Dilemma: An Alternative Approach", *American Antiquity* 40,11-16.

6. Ronald D. Lee (1977) "Methods and Models for Analyzing Historical Series" in R. Lee, ed., *Population Patterns in the Past*, Academic Press, New York, pages 353-354.

7. References may be found in S. Asmussen and H. Hering (1983) *Branching Processes*, Birkhauser, Basel.

8. Peter Jagers (1975) *Branching Processes with Biological Applications*, Wiley, New York, page 70.

9 Keith Olive (1985) "Inflation and Cosmology", typescript.

10. I have followed Bernard Campbell (1985) *Human Evolution*, Aldine Press, New York, pages 89-117. Note the discussion in Fekri Hassan (1981) *Demographic Archaeology* Academic Press, New York, page 199, and in D. Pilbeam (1975) "Middle Pleistocene Hominids" in K. Butzer and G. Isaac, eds., *After the Australopithecines*, Mouton, The Hague, pages 812-814, 834-840.

11. Mark Cohen (1980) "Speculations on the Evolution of Density Measurement and Population Regulation in Homo-Sapiens" in M. Cohen, R. Malpass, and H. Klein, eds., *Biosocial Mechanisms of Population Regulation*, Yale University Press, New Haven, pages 275-303.

12. K. W. Wachter (1978) "Age Pyramid Covariances" in *Statistical Studies of Historical Social Structure*, Academic Press, New York.

13. Ronald D. Lee (1977) "Methods and Models for Analyzing Historical Series" in R. Lee, ed., *Population Patterns in the Past*, Academic Press, New

York, pages 353-354, and Barbara Spencer (1976) "Size of Population and Variability of Demographic Data", *Genus* 32, pages 11-40.

14. Krishna Athreya and Peter Ney (1972) *Branching Processes*, Springer, Berlin, page 40.

15. Krishna Athreya and S. Karlin (1971) "Branching Processes with Random Environments II", *Annals*, page 1857.

16. James Pitman (1980) "A Note on L-2 Maximal Inequalities", *Seminaire de Probabilities XV, Lecture Notes in Mathematics Number 850*, Springer, Berlin.

17. Chris Heyde and B. Brown (1971) "An Invariance Principle and some Convergence Rate Results for Branching Processes", *Zeitschrift fuer Wahrscheinlichkeitstheorie...* 20,271-278.

18. William Feller (1966) *An Introduction to Probability Theory and its Applications II*, Wiley, New York, page 515.

19. The expected value conditional on non-extinction times the probability of non-extinction has to equal the unconditional expected value.

20. The limiting variable W has a moment generating function which satisfies a functional equation of Hawkins and Ulam given on page 15 of T. E. Harris (1963) *Theory of Branching Processes*, Springer, Berlin. Expanding the right-hand-side of this equation in powers of $m-1$ proves the closeness to the exponential for m near 1.

21. Formulas are conveniently found in Samuel Karlin and H. Taylor (1975) *A First Course in Stochastic Processes*, Academic Press, New York, pages 263-364.

22. This figure is based on calculations by Patrick Galloway of the Graduate Group in Demography at Berkeley with crude growth rates in non-overlapping intervals from Wrigley and Schofield (1981) *The Population History of England 1541-1871*, Harvard University Press, Cambridge, Mass., pages 208-209. It agrees with a figure obtained from 30-year averages of national average net reproduction rates from page 530 of the same source.

22a. It is worth noting, however, that Patrick Galloway estimates a global standard deviation as high as .08 from the covariance between European and Chinese series for 820 to 1780 in the extremely speculative series in McEvedy

and Jones (1978) *Atlas of World Population History*, Penguin, London. This matter deserves further study.

23. Albert Ammerman and L. Cavalli-Sforza (1984) *The Neolithic Transition and the Genetics of Populations in Europe*, Princeton University Press, Princeton, Chapter 5.

24. William Feller (1951) "Diffusion Processes in Genetics", *Proceedings of the Second Berkeley Symposium*, University of California Press, Berkeley, pages 227-246. Rigorous theorems are found in Anders Grimvall (1974) "On the Convergence of Sequences of Branching Processes", *Annals of Probability* 2, pages 1027-1045.

25. Samuel Karlin and H. Taylor (1981) *A Second Course in Stochastic Processes*, Academic Press, New York, pages 195 and 203.

26. Cf. Shripad Tuljapurkar and S. Orzack (1980) "Population Dynamics in Variable Environments I", *Theoretical Population Biology* 18, 314-342.

27. S. Karlin and H. Taylor (1981) pages 358-360.

28. Such a process is called a compound process.

29. The minuteness of the change in rates at the transition has been called the "Neolithic Paradox" by Ammerman and Cavalli-Sforza.

30. Cf. Fekri Hassan (1981) pages 193-199.

Bibliography

Cohen, Joel E. (1977) "Ergodicity of age structure in populations with Markovian vital rates, II and III," *Advances in Applied Probability*, 9, 18-37; 462-475.

Cohen, Joel E. (1981) "Stochastic Demography", in N. Johnson and S. Kotz, eds., *Encyclopedia of Statistical Sciences*, Wiley, New York.

Jagers, Peter (1975) *Branching Processes with Biological Applications*, Wiley, New York, p 70.

LeBras, Herve (1971) "Elements pour une theorie des populations instables", *Population (Paris)* 26, 525-572.

Lee, Ronald D.(1978) "Models of Preindustrial Population Dynamics, with Applications to England" in C. Tilley, ed., *Historical Studies of Changing Fertility*, Princeton University Press.

Lee, Ronald D. (1983) "Homeostasis in Human Populations", paper presented to the Stanford-Berkeley Colloquium in Demography, May 26, 1983.

Strong, Donald R.(1986) "Density VAgueness: Abiding the Variance in the Demography of Real Populations", in T. Case, J. Diamond, J. Roughgarden, and T. Schoener, eds., *Community Ecology*, Harper and Row, N.Y., to appear.

Wrigley, E.A. and R. Schofield (1981) *The Population History of England 1541-1871*, Harvard University Press, Cambridge, Mass.

Lee, Ronald D. (1977) "Methods and Models for Analyzing Historical Series" in R. Lee, ed., *Population Patterns in the Past*, Academic Press, New York, pp 353-354.

S. Asmussen and H. Hering (1983)*Branching Processes*, Birkhaueser, Basel.

Boserup, Ester (1981) *Population and Technological Change*, University of Chicago Press, Chicago, pp 35-36.

Bernard Campbell (1985) *Human Evolution*, Aldine Press, New York, pp 89-117.

Mark Cohen (1980) "Speculations on the Evolution of Density Measurement and Population Regulation in Homo-Sapiens" in M. Cohen, R. Malpass, and H. Klein, eds., *Biosocial Mechanisms of Population Regulation*, Yale Univerity Press, New Haven, pages 275-303.

Hassan, Fekri (1981) *Demographic Archaeology* Academic Press, New York, p 199.

Brian Haydon (1975) "The Carrying Capacity Dilemma: An Alternative Approach", *American Antiquity* 40,11-16.

Howell, Nancy (1979) *Demography of the Dobe !Kung* Academic Press, New York, Chapters 4 and 5.

J.R. Pilbeam (1975) "Middle Pleistocene Hominids" in K. Butzer and G. Isaac, eds., *After the Australopithecines*,Mouton, The Hague, pp 812-814, 834-840.

Wachter, Kenneth W. (1978) "Age Pyramid Covariances" in *Statistical Studies of Historical Social Structure*, Academic Press, New York.

Spencer, Barbara (1976) "Size of Population and Variability of Demographic Data", *Genus* 32, pp 11-40.

Athreya, Krishna, and Peter Ney (1972) *Branching Processes*, Springer, Berlin, p 40.

Athreya, Krishna, and S. Karlin (1971) "Branching Processes with Random Environments II", *Annals*, p 1857.

Feller, William (1966) *An Introduction to Probability Theory and its Applications II*, Wiley, New York, p 515.

T. E. Harris (1963) *Theory of Branching Processes*, Springer, Berlin. p 15.

Karlin, Samuel, and H. Taylor (1975) *A First Course in Stochastic Processes*, Academic Press, New York, pp 263-364.

Ammerman, Albert, and L. Cavalli-Sforza (1984) *The Neolithic Transition and the Genetics of Populations in Europe*, Princeton University Press, Princeton, Chapter 5.

William Feller (1951) "Diffusion Processes in Genetics", *Proceedings of the Second Berkeley Symposium*, University of California Press, Berkeley, pp 227-246.

Grimvall, Anders (1974) "On the Convergence of Sequences of Branching Processes", *Annals of Probability* 2, pp 1027-1045.

Karlin, Samuel, and H. Taylor (1981) *A Second Course in Stochastic Processes*, Academic Press, New York.

Shripad Tuljapurkar and S. Orzack (1980) "Population Dynamics in Variable Environments I", *Theoretical Population Biology* 18, 314-342.