CHAPTER 3

DEMOGRAPHY

WALTER SCHEIDEL

Demographic conditions are a major determinant of economic performance. In the following, I focus on the relationship between demographic structures and macro-economic features. The economics of household and gender are discussed in the next chapter. After a brief outline of the fundamental demographic characteristics of the Greco-Roman world (I–II), I present a theoretical model of the interdependence of economic and demographic development, and explore its principal variables in the context of ancient Mediterranean economies (III–VIII). This introductory survey is meant to provide a conceptual framework for the more specific discussions in Chapters 7 to 28, and more generally seeks to contextualize the study of Greek and Roman economic and demographic history within the wider ambit of historical demography and population theory. While many of the issues raised in the following sections cannot be satisfactorily addressed on the basis of ancient evidence, they are nevertheless essential to our understanding of ancient economies.

I LIFE EXPECTANCY AND ITS CORRELATES

(a) *Mortality and morbidity*

In recent years, researchers have established a broad consensus regarding the basic structure of ancient populations.[1] Continuing controversies are now largely confined to particular interpretations of the evidence, which are of limited relevance here.[2] The populations of the ancient world were characterized by a regime of high fertility and mortality. While mean life expectancy at birth is conventionally put in a range from about twenty to thirty years, the actual age structures of ancient populations (and thus age-specific survival rates) are generally unknown. Age records from some

[1] Much of recent work deals with the Roman world: Parkin 1992; Bagnall and Frier 1994; Frier 2000. For Greece, see Hansen 1985; Sallares 1991: 42–293. Corvisier and Suder 2000 give a general overview. For critical evaluations of recent research, see Golden 2000; Scheidel 2001b; and cf. Suder 1988, 1991, and Corvisier and Suder 1996 for earlier bibliography. Scheidel forthcoming, b will provide a comprehensive overview.

[2] More substantial disagreements over population numbers are addressed below, II.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

300 census returns filed in Roman Egypt during the first three centuries AD have been used to reconstruct female and male age distributions that are broadly consistent with model life tables suggesting a mean life expectancy at birth of twenty-two to twenty-five years.[3] An alternative reading of these records points to significant differences between rural and urban populations, and to exceptionally high attrition rates in large cities.[4] Mean life expectancies at birth in the low twenties have also been extrapolated from select cemetery populations, Roman tombstones in north Africa, and a Roman schedule used to calculate annuities known as "Ulpian's life table."[5] However, demographic readings of these sources remain controversial: age records in epitaphs are often vitiated by age and gender preferences; skeletal samples cannot always be reliably aged and need not accurately mirror the age structure of actual past populations; and the evidentiary basis of the annuities schedule is unknown.[6] At the same time, in supporting the notion of a low mean life expectancy at birth in the twenties, these estimates converge with literary evidence that suggests comparably low survival rates among the Roman elite and, more importantly, with comparative data from more recent high-mortality regimes ranging from eighteenth-century France to early twentieth-century China, India, and Egypt.[7] In consequence, we may reasonably assume that throughout the Greco-Roman world, average life expectancy at birth fell in a bracket from about twenty to thirty years, although these limits may occasionally have been exceeded in particularly hazardous (e.g., malarious) or healthy (e.g., high-altitude and low-population density) environments.

Age-specific death rates and therefore overall age structures were likely to have varied within a broad band bounded by constraints set by the probable levels of mean life expectancy at birth. While modern model life tables are predicated on the simplifying notion of predictable ratios between life expectancies at different ages which allow a given life expectancy at birth to be linked to a particular set of normative age distributions,[8] the specifics of the actual age composition of pre-modern populations were contingent on ecological conditions. In the long term, the nature and prevalence of endemic infectious disease acted as the principal environmental determinant of local age structures. The considerable degree of epidemiological variation within the ancient Mediterranean has been reviewed in the previous chapter: the correlate of climate, population density, and the evolution of infective agents, ancient disease environments were subject to profound change over space and time. Local demographic conditions would have

[3] Bagnall and Frier 1994: 75–110; Bagnall, Frier and Rutherford 1997: 99–104.
[4] Scheidel 2001a: 118–80. [5] Frier 1982; 1983b; 2000: 790–2.
[6] For criticism, see Hopkins 1966; Parkin 1992: 5–58, 75–8, 82–3; Scheidel 2001b: 17–19; 2001c.
[7] Roman elite: Scheidel 1999. Comparanda: Bagnall and Frier 1994: 88.
[8] Coale and Demeny 1983. For criticism, see Scheidel 2001c: 3–11.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

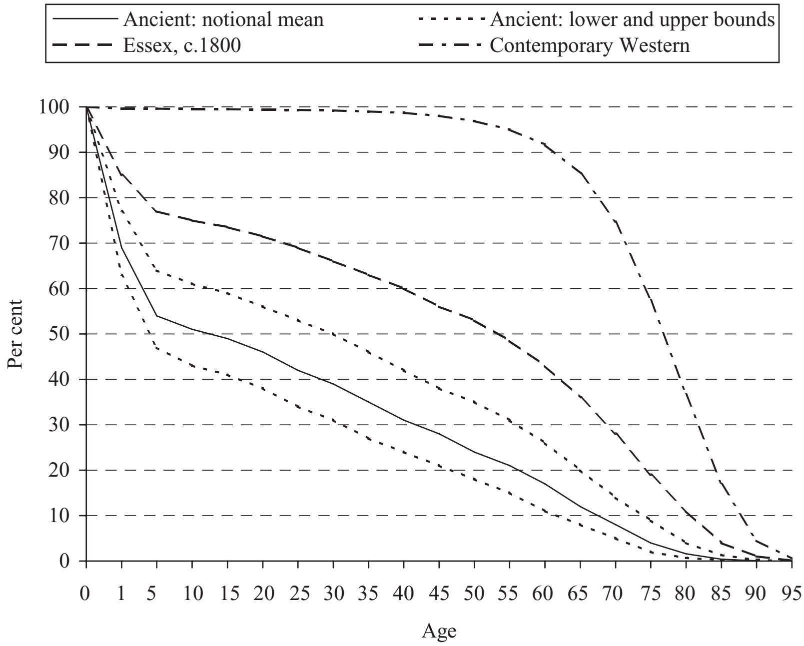


Figure 3.1 Percentage of survivors to age *x*
Source: Coale and Demeny 1983: Model West Females, Levels 1 and 6 (Ancient: lower and upper bounds), Level 3 (notional mean), Level 11 (Essex, c.1800), Level 25 (Contemporary Western)

varied accordingly.[9] In addition, cultural conventions (from collective violence to gender biases) affected differential mortality.[10] For these reasons, the survival rates for Greco-Roman populations set out in Figure 3.1 are best understood as rough approximations that map out the limits of the plausible. The "notional mean" in particular is suggested for computational purposes only and should not be taken to apply to any given regional or local environment.

Despite these uncertainties, the basic outlines are clear: although mortality rates in infancy and early childhood are particularly volatile and may arguably be exaggerated by standard model life tables, there can be no doubt that attrition was strongly concentrated in the first years of life.[11] In addition, mortality remained substantial throughout the life-cycle, thereby depressing the number of those surviving to a more advanced age: up to three or four times as many people may have died before age ten as after age sixty. Death was as much a phenomenon of childhood and maturity

[9] E.g., Scheidel 1996c: 139–63; 2001a: 1–180; Sallares 2002; Woods forthcoming.
[10] Scheidel 2001b: 30. Cf. below, Chapter 4. [11] Woods 1993; Scheidel 2001b: 21–3.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

as of old age. As is shown in the next chapter, mortality on this scale hampers economic development by discouraging investment in human capital, creating large numbers of orphans and widows, and disrupting long-term economic strategizing. It is true that what is by modern standards considered "high" mortality is not an insuperable impediment to economic growth: after all, mean life expectancy at birth in England did not consistently exceed forty years until the mid-nineteenth century, when industrialization was well underway. However, an average rate in, say, the low twenties differs profoundly from one in, say, the mid-thirties, translating to adolescent mortality rates that are two-thirds higher in the former than in the latter scenario. The incidence of familial and economic dislocation would vary accordingly.[12] Moreover, the scale of mortality is positively correlated with that of morbidity. In the most developed nations today, "health-adjusted life expectancy" (HALE) – the average number of years lived in good health – falls short of overall life expectancy by no more than 8 percent, whereas in the most disease-ridden countries, up to one-sixth of all years lived are spent in a state of ill health or disability.[13] A plausible HALE of twenty years or less would have significantly depressed economic productivity in ancient societies.

(b) *Fertility*

Most importantly, high mortality usually went hand in hand with high fertility.[14] At ancient levels of life expectancy, the average woman surviving to menopause must have given birth to between 4.5 and 6.5 children in order to maintain the existing population. In view of likely levels of divorce, widowhood, and sterility, mean marital fertility must have been higher still, around 6 to 9.[15] Local long-term fertility rates would have differed depending on the specifics of local mortality regimes. Within any given region, however, average fertility must have been closely tied to the death rate: a population in which mean birth rates had exceeded (or fallen short of) mean death rates by one-fifth (equivalent to annual growth rates of plus or minus 0.7 percent) would have doubled (or been halved) in size every century. Shifts of that nature could only occur locally and for a limited amount of time. Annual growth rates of the order of either −1 or +1 percent or more are feasible mostly in the short run or in specific locales: while death rates could exceed birth rates either temporarily in the event of epidemic outbreaks or more consistently in large cities where negative

[12] England: Wrigley and Schofield 1981: 235. Estimates based on Coale and Demeny 1983: 56, 61.

[13] WHO, *World Health Report 2003*.

[14] Bagnall and Frier 1994: 135–59; Frier 1994, 2000: 797–808; Scheidel 2001b: 32–46, forthcoming, b.

[15] Frier 1994: 325.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

growth rates had to be counterbalanced by immigration, growth spurts – especially in the wake of demographic contractions or in colonial settings – would raise fertility over baseline mortality.[16] Nonetheless, in the long term (such as the period covered by this volume), local, regional, and temporal variations largely canceled each other out, converging into a central trend of very slow net growth over time.[17]

II POPULATION SIZE

(a) *Long-term trends and contexts*

Recorded population numbers are few in number and often of doubtful quality: rhetorical stylization and symbolic figures permeate ancient sources, and although official counts in Greek *poleis* or the Roman empire may well have yielded reasonable approximations of gross population number, such findings rarely entered the surviving literary tradition.[18] The serial statistics of Roman citizen numbers from the early Republic through the first century AD are the main exception but subject to problems of interpretation (see below). The most reliable evidence is derived from papyrological documents from Ptolemaic and especially Roman Egypt which record the number of residents of various cities and villages but no regional totals.[19] In contrast to the contemporaneous Han empire in China, no global census totals for the Roman empire are known. As a consequence, the figures given in this and the following sub-section range from rough approximations to probabilistic guesses and should not normally be taken to indicate more than a certain order of magnitude.[20]

During the period under review, from the middle of the second millennium BC to the early first millennium AD, all parts of the Mediterranean and its hinterlands experienced significant demographic growth. With all due caution, we may assume that between the twelfth century BC and the second century AD, the population of the part of Europe that was eventually taken over by the Roman empire approximately quadrupled in size, at a long-term average annual growth rate of around 0.1 percent. Growth was probably slower in the already more developed eastern half of the Mediterranean but far from negligible, perhaps of the order of 0.07 percent per

[16] For epidemics and urban mortality, see below, II, VIII. For growth following contraction, see below, Chapter 8, and Scheidel 2003b.

[17] This principle is easily confirmed by the observation that at an annual net growth rate of 1 percent, a Mediterranean population of about 50,000 in 1000 BC would have expanded to 6.5 billion (the current world population) by AD 200. Ancient baseline growth must have been much lower, around 0.05–0.1 percent p.a.: see below, II.

[18] Scheidel 2001b: 49. [19] Rathbone 1990; Alston 2002: 332.

[20] For the (hazardous) derivation of population estimates from probable carrying capacity, see esp. Beloch 1886; Corvisier 1991: 229–92; de Angelis 2000. Cf. Sanders 1984 on potential margins of error.

year during the same period.[21] By comparison, the mean annual growth rate for the European part of the Roman empire amounted to between 0.06 and 0.07 percent for the period from the late second-century AD peak to 1800, while it was substantially higher, probably approaching 0.1 percent, for Europe as a whole.[22] During the same period, virtually no net growth occurred in Rome's former Asian and African provinces. In China proper, annual growth averaged 0.1 percent between 1 and 1800.[23] These comparisons reveal both fundamental continuities and divergences. Long-term demographic growth in the temperate parts of Europe was fairly stable from the late Bronze Age to the onset of modernity. After the depression of population numbers following the disintegration of the western and much of the eastern Roman empire in the fifth and sixth centuries, the formerly Roman part of Europe (with the exception of Greece) generally re-attained peak Roman population levels by the twelfth or thirteenth centuries, and after another slump caused by the Black Death consistently exceeded them from the mid-fifteenth century onwards. The underlying secular trend rate was closely tracked by the population of China at the opposite end of the Eurasian land mass. By contrast, demographic – and therefore economic – development in north Africa and the Levant reached its pre-modern saturation point in the Roman imperial period, and did not permanently cross this threshold until the nineteenth century.[24]

To the extent that the production of people is a function of the production of goods, these underlying growth rates also provide a rough index of economic development. In the long term, and especially in subsistence economies which cannot accommodate substantial declines in conventional living standards, population cannot grow faster than total output. The ability of a system of production to support a given increase in population over a given period of time at a constant real wage is known as the absorption rate. Judging by the above estimates, the minimal long-term absorption rate of ancient – or indeed any pre-modern agrarian – economies appears to have been close to 0.1 percent per annum. This rough estimate is in line with the mean absorption rate calculated for late mediaeval and early modern England.[25] In addition, due to ongoing technological innovation (see below, III), per capita consumption must have risen very gradually in the long term. If we were to posit, entirely speculatively but perhaps not unreasonably, a 50 percent increase in per capita output and consumption

[21] From c. 8–10 million to 35–40 million in Europe and from c. 8–9 million to 20–23 million in the Near East: McEvedy and Jones 1978, and below, Table 3.1.

[22] From 35–40 million for the Roman provinces (and 45–50 million for Europe as a whole) to 100 million (and 180 million) in 1800: Table 3.1; de Vries 1984: 36.

[23] McEvedy and Jones 1978: 171. Cf. Wilson and Airey 1999: 119 for medium-term variation.

[24] McEvedy and Jones 1978; Frier 2000: 814.

[25] Lee 1980: 525 (0.089 percent p.a. 1250–1700, 0.46 percent p.a. 1705–1789, 0.88 percent p.a. 1810–1839).

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

between 1000 BC and AD 200, extensive economic growth would have averaged 0.16 percent per annum, translating to a baseline rate of intensive growth of 0.06 percent. The latter ties in with recent guesstimates of corresponding rates of 0.05 percent for Greece from 1300 to 300 BC (including a 0.15 percent spurt from 800 to 300) and of up to 0.1 percent for the last two centuries BC in the Roman empire.[26]

(b) *Regional developments*

While these figures are unlikely to be wide of the mark overall, they conceal significant variation in the medium term and between regions. At the end of the Bronze Age, the future predominance of Greece and Italy was already foreshadowed by the fact that their average population densities exceeded those of other parts of Europe.[27] In mainland Greece, a massive demographic contraction in the wake of the collapse of the Mycenaean palace system was followed by a prolonged recovery that took off in the ninth century, gathered momentum in the eighth and generated substantial net growth relative to late Bronze Age levels before petering out in the fourth or third centuries BC (see Chapter 8). By the classical period, this process of natural growth, together with the influx of large numbers of chattel slaves, had resulted in particularly high population densities in the major *poleis* of central Greece, above all Athens, Corinth, and Aegina, which to varying degrees came to depend on the large-scale import of grain to support their local populations (see Chapters 12–14). By contrast, populations in highland and other peripheral areas expanded later and did not peak until the Hellenistic and Roman periods.[28] Modern estimates of the peak population of the core of mainland Greece and the Aegean islands in the classical period converge on about 2 million, or 3 to 3.5 million including Thessaly, Epirus, and Macedonia.[29] Overseas settlements, first in Ionia on the west coast of Asia Minor and from the eighth century onwards in Sicily, southern Italy, and the Black Sea region, may have accounted for another 2 million Greeks or thereabouts.[30] Modern debates have mostly focused on the size of the Athenian population in the fifth and fourth centuries BC. Drawing on scarce source material and divergent assumptions about local carrying capacity and the scale of food imports, modern estimates range

[26] Morris 2004: 736; Saller 2002: 259–60. Cf. Jones 1988: 30–4 on traditional long-term growth.

[27] McEvedy and Jones 1978: 19–20, with 20 fig. 1.3.

[28] Bintliff 1997b. Cases include Thessaly, Macedonia, Epirus, Aetolia, Achaea and Crete.

[29] Corvisier and Suder 2000: 34. Cf. Map 3.1.

[30] Beloch 1886: 261–305; Scheidel 2003b. Hansen 2006 estimates a somewhat higher grand total of at least 7 million, based on a series of extrapolations from the known size and number of walled Greek cities and assumptions about urban settlement density and the ratio of urban to rural residents (for the latter, see below, n. 188).

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

from 250,000 to over 300,000 residents (including 60,000 adult male citizens and perhaps 100,000 slaves) at the late Periclean peak in the 430s BC, and from 150,000 (including 20,000 adult male citizens and 50,000 slaves) to 250,000 residents in the 320s BC.[31] Our ignorance about the likely number of resident aliens (metics) and slaves is responsible for much of this uncertainty. In the early fifth century BC, the Spartan citizen population peaked at around 25,000, sustained by up to 100,000 helots.[32]

In its heyday in the early fifth century BC, the Persian empire of the Achaemenid dynasty stretched from Egypt and the Aegean to the Indus valley and may have comprised some 20 to 25 million subjects. Its principal successor, the Seleucid empire, hardly exceeded 15 million in the third century BC and subsequently kept losing territory, while the population of Egypt fluctuated between perhaps 4 and 7 million during the Greco-Roman period.[33] Field surveys frequently indicate increasing urbanization as well as a spread of settlement in the Hellenistic East (except in Greece itself) that is indicative of ongoing population growth and agrarian intensification.[34]

Estimates of the number of Roman citizens (and of the population of Italy as a whole) hinge on our interpretation of reported census totals for adult males. Figures are given from 508 BC onwards but may not be trusted until the mid-third century; fourteen totals (from 258,318 to 394,736) are available for the second century but only four for the next, featuring a jump from 910,000 in 70/69 BC to 4,063,000 in 28 BC.[35] As the free population of peninsular Italy probably stood at around 3 million in 225 BC and only a few hundred thousand adult male citizens are thought to have resided outside Italy in 28 BC, the total for 28 BC would imply a citizen population of over 10 million for Italy, a higher gross total including aliens and slaves, and an even more massive aggregate tally of perhaps closer to 20 million for the mid-Principate.[36] The implied population density for the Italian peninsula in particular of up to 100 people/km^2 is very high and was not reached until the late nineteenth century.[37] This has led the majority of scholars to follow Beloch's conjecture that the Augustan census recorded all citizens instead of adult males only.[38] This otherwise unsupported assumption reduces the total Italian population to a more plausible 4 million citizens plus aliens and slaves in 28 BC. It deserves attention that even in this conservative scenario,

[31] See my addenda in Garnsey 1998: 195–200 for a full survey of the literature.

[32] Figueira 2003; Scheidel 2003c.

[33] My estimates for the Achaemenid and Seleucid empires fall between those by McEvedy and Jones 1978: 125 (lower) and Aperghis 2004: 56–8 (higher; cf. also Aperghis 2001: 73–7). For Egypt, see Scheidel 2001a: 181–250.

[34] Alcock 1994: esp. 187: see below, Chapter 15, Table 15.1.

[35] Brunt 1987: 3–120. Figures for the late first century are necessarily higher because of the enfranchisement of the Italian allies and the Cisalpina.

[36] Brunt 1987: 60, 263; Morley 2001: 50–2; Scheidel 2004c: 6–7.

[37] Scheidel 1996c: 168. [38] Beloch 1886: 370–8; Brunt 1987: 113–15.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

Map 3.1 Population densities in Greece in the fifth and fourth centuries BC
Source: adapted from Corvisier and Suder 2000: 36 fig. 1

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

Italy's average population density was still much higher than in any other region except Egypt and Syria (see below, Table 3.1). Even so, the higher estimate is not strictly speaking impossible and has recently been defended by Lo Cascio.[39] Evidence of endemic conflict over land in the late Republic, the large number of cities in Italy and other indications of population pressure (as well as the strikingly high levels of military mobilization implied by the lower estimate) might be taken to support this revisionist argument.[40] At the same time, the "high count" is seriously undermined by its mismatch with comparative data from other parts of the Roman empire and from later periods, its implications regarding the size and distribution of the imperial population as a whole, and its logical incompatibility with well-documented developments such as the emergence of a central Italian slave economy and the eventual geographical peripherization of military service. Thus, while the possibility of a substantially larger Italian population (which might require upward adjustments for other parts of the Roman empire as well) cannot be ruled out entirely, the balance of probability favors the lower estimate.[41]

If we accept the lower estimate of the size of the Roman citizenry, the total population of the Roman empire at its peak on the eve of the "Antonine plague" of AD 165 probably numbered between 60 and 70 million. While approximately 55–60 percent of them resided in the European provinces and around 20 percent each in Asia and Africa, the demographic split between the "Latin" western and the "Greek" eastern halves of the empire was about 60–65 to 35–40 percent (Table 3.1).[42] By coincidence, the Roman empire broadly resembled the Han empire in terms of population number and density, with each of them controlling about a quarter of the world's population.

In Late Iron Age western Europe, considerable economic innovation and therefore demographic growth occurred for several centuries prior to the Roman conquests, and continued into the early imperial period.[43] Under

[39] Lo Cascio 1994a, 1994b, 1999b, forthcoming.

[40] See below, VI and VIII, and Lo Cascio 2001a (mobilization; but see now Rosenstein 2004); Morley 2001: 59–61 (land hunger).

[41] Lower estimate: Morley 1996: 46–50; Scheidel 1996c: 167–8 and esp. 2004c: 2–7. Cf. also De Ligt 2004: 728–38. The likely impact of urban excess mortality, above all in Rome, also implies a low total: Morley 1996: 33–54; Jongman 2003a; Scheidel 2003a. For a brief summary of this conundrum, see Scheidel 2001b: 52–7. The suggestion that the Augustan census figures may refer to all household heads, orphans, and widows (Hin [in progress]) supports a plausible intermediate scenario that is consistent with an early imperial Italian population of perhaps 11–12 million.

[42] Figures in Table 3.1 adapted from Frier 2000: 814 table 6 (for a grand total of 61.4 million), albeit with a preference for somewhat higher tallies where appropriate: cf., e.g., Woolf 1998: 138; Millett 1990: 181–6; Scheidel 2001a: 242–8. The numbers in parentheses are derived from the alternative interpretation of imperial census figures in n. 41.

[43] See below, Chapters 9 and 24. This process can be linked to climatic improvements from about 400 BC onwards (cf. Galloway 1986: 17).

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

Table 3.1 *The estimated population of the Roman empire in AD 165*

Region	Population (in millions)	Area (in km²)	Population density (per km²)
Italy & islands*	8–9 (?12–13)	310,000	26–29 (?39–42)
Iberia	7–9	597,000	12–15
Gaul & Germany	9–12	680,000	13–18
Britain	1.5–2	160,000	9–13
Danubian region	5–6	660,000	8–9
Greek peninsula	2.5–3	161,000	16–19
European provinces	33–41 (?37–45)	2,568,000	13–16 (?14–18)
Anatolia	9–10	650,000	14–15
Greater Syria**	5–6	140,000	36–43
Asian provinces	14–16	790,000	18–20
Egypt	5–6	30,000	167–200
North Africa	7–8	415,000	17–19
African provinces	13–14	445,000	–
Total	59–72 (?63–76)	3,800,000	16–19 (?17–20)

* Sicily, Sardinia and Corsica
** Including Cyprus

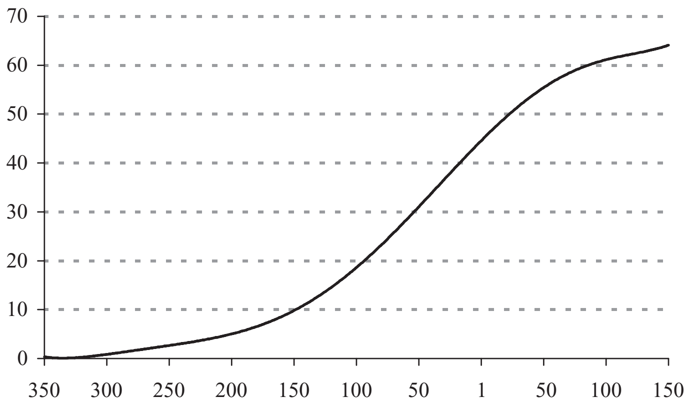


Figure 3.2 Approximate size of the population under Roman control, 350 BC–AD 150 (trendline; in millions)
Source: Scheidel forthcoming, b

Roman rule, the Maghreb and Tripolitania experienced major expansions in output and population (cf. Chapter 27). In general, it is reasonable to assume that the peaceful conditions during the first two centuries of the Principate resulted in ongoing growth throughout the empire and particularly in its western half.[44] Widespread epidemics from the 160s to the

[44] Frier 2000: 813–15; Scheidel 2001b: 63–4. The "frontier hypothesis" predicts higher fertility in sparsely settled frontier provinces: Easterlin 1976; Vanlandingham and Hirschman 2001.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

190s AD, probably smallpox, caused a demographic contraction which may have been substantial, although its actual scale remains uncertain.[45] Further pandemics in the 250s and 260s AD probably contributed to this process. Despite ancient and modern concerns about general "manpower shortage" in late antiquity, the reality or extent of population decline is difficult to ascertain and was probably most pronounced in unstable frontier areas in the western provinces.[46] For this late period, from the fourth to the sixth centuries AD, field surveys and other archaeological data suggest expansions in parts of the eastern half of the empire from mainland Greece to Syria and Palestine.[47] Large-scale drops in population accompanied the disintegration of the western Roman empire in the fifth century and the arrival in the Mediterranean in the 540s AD of a plague pandemic akin to the Black Death.[48]

(c) *Migration*

Migratory flows are usually impervious to any but the crudest probabilistic attempts at quantification.[49] According to Cavalli-Sforza's reconstruction of the genetic history of the region, no massive population shifts occurred after the beginning of the Iron Age, which means that the bulk of all growth may be ascribed to the gradual expansion of local populations with improving productive technology.[50] At the same time, the geography of the Mediterranean facilitated inter-regional mobility, exemplified by the creation of hundreds of settlements all over the Mediterranean and the Black Sea by Phoenicians and Greeks during the first half of the first millennium BC, and the transfer of Roman and Italian merchants to the Greek East and the reverse flow of migrants from the Levant to Italy.[51] Local migration between villages may well have been substantial, and the establishment of hundreds of cities in previously un-urbanized regions must have necessitated the (mostly short-range) relocation of perhaps 20 to 40 million people during the period under review.[52] In quantitative terms, the slave trade was the most important conduit of inter-regional migration, involving as it did the transfer of millions of slaves to the core areas of

45 Littman and Littman 1973; Duncan-Jones 1996 (with Greenberg 2003); Scheidel 2002; Zelener 2003; Rijkels 2005: 22–76.

46 See esp. Lewit 1991; also Whittaker 1976; Fischer 1993. Salmon 1974: 114–79 fails to produce compelling evidence of general population decline.

47 Bintliff and Snodgrass 1988; Tate 1992; Safrai 1994; Foss 1995; and see below, VI.

48 Christie 1996; Russell 1968; Biraben 1975: 25–48; Conrad 1981.

49 Scheidel 2001b: 46–8. For a comprehensive parametric reconstruction of population transfers in Roman Italy, see Scheidel 2004c, 2005a.

50 Cavalli-Sforza, Menozzi and Piazza 1994: 277–80.

51 Horden and Purcell 2000: 377–400 stress the lateral mobility of Mediterranean populations; cf. Purcell 1990 on Greek mobility in particular. Settlements: Aubet 1993 (Phoenicians); Chapter 10 (Greeks).

52 Villages: Osborne 1991b; cities: see below, VIII.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

ancient slave society, Greece and especially Roman Italy.[53] By contrast, not more than a few hundred thousand individuals are likely to have emigrated from mainland Greece from the archaic to the Hellenistic periods, and by AD 14, the number of Roman citizens of Italian extraction who resided in the provinces need not have exceeded 750,000.[54] In contrast to the Later Han empire, where population shifted southwards, and various centrifugal colonization movements under subsequent dynasties, the Roman empire does not appear to have experienced mass migration between its constituent parts.[55]

III POPULATION REGULATION

(a) *A general model*

In order to appreciate the significance of demographic conditions for economic development, we need to address several key questions: in traditional societies, what was the reciprocal relationship between population growth and economic change; was the former in any sense regulated, and by what mechanisms; and how did population size or growth relate to general well-being?

Population size equilibrates with resources at a level mediated by technology and conventional living standards.[56] Following Malthus, the principal equilibrating mechanisms can be divided into preventive checks (acting on fertility) and positive checks (acting on mortality). Homeostatic regulation is achieved by the interplay of these variables regardless of whether the members of a population are aware of or attempt to influence this process. The theoretical concept of demographic equilibrium does not envisage an inflexibly static system: rather, it seeks to identify sustainable population size in the absence of exogenous forces, such as famines, epidemics and wars and, in the long run, climatic change, technological innovation, and the evolution of endemic disease patterns, which in reality account for much of observed variation in population size and growth. Thus, by emphasizing underlying constraints, the concept of the equilibrium state helps us assess the relative input of exogenous factors.

Figure 3.3 illustrates the governing principles of homeostatic control. Following demographic contractions that pushed population size below the equilibrium level (as in the case of the Early Iron Age depression in

[53] Finley 1981: 167–75; Harris 1980b, 1999; Bradley 1987; Scheidel 1997, 2005a, forthcoming, c.

[54] Scheidel 2003b; Brunt 1987: 264–5.

[55] China: Bielenstein 1947; Lee and Wang 1999: 115–19; cf. Scheidel 2004c: 25. Frier 2000: 813 speculates that between AD 14 and 164, 20,000 persons may have annually migrated from the eastern to the western half of the empire (i.e., at a rate equivalent to 0.1 percent of the eastern population).

[56] For the theoretical underpinnings of this section, see esp. Lee 1986a; Wood 1998.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

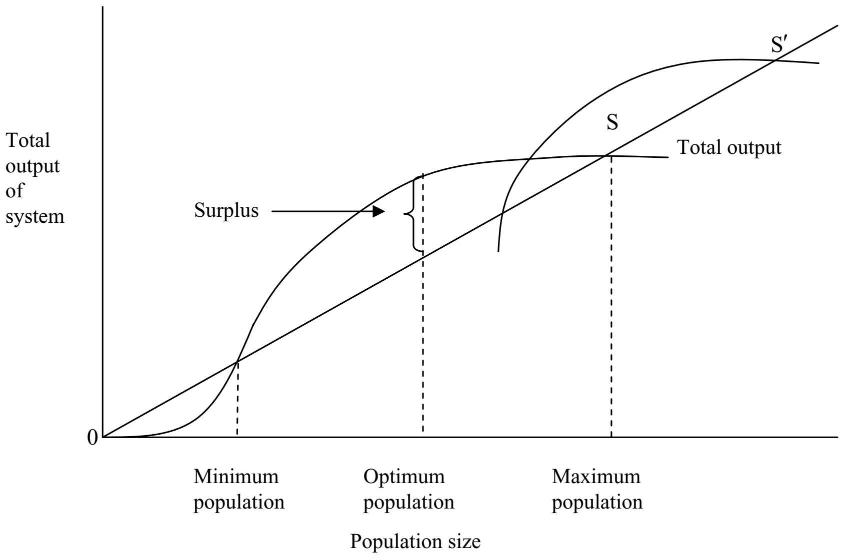


Figure 3.3 Population size and surplus for given resources and technology
Source: Lee 1986a: 101 fig. 1 and Wood 1998: 113 fig. 9

twelfth- and eleventh-century BC Greece discussed in Chapter 8 or, perhaps, in Roman Egypt in the wake of the "Antonine plague" of the late second century AD) or improvements in productive capacity caused by the opening up of new resources (such as during Greek "colonization" in the archaic period or Roman expansion across the western Mediterranean) or by technological progress (less common in the case of ancient economies), gross and per capita output may initially rise concurrently with population number. Eventually, growth in output and therefore population will decline as population size increases, especially once it approaches the sustainable equilibrium level. This constraint on population growth is a function of declining marginal productivity (also known as the "law of diminishing returns"): without ongoing technological change, output (i.e., production) will gradually rise more slowly than input (i.e., labor). When labor is the main input, diminishing returns will reduce average consumption (indicated by the shrinking distance between output and subsistence requirements in Figure 3.3). In theory, a population may reach a demographic saturation point (here denoted as S) beyond which further inputs yield negative returns and additional growth is not feasible.[57] At this stage, owing to the lack of capacity reserves, the population is increasingly rendered vulnerable to environmental shocks such as disease or food crises. Crucially, however,

[57] Wood 1998: 105–7.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

any notional equilibrium state or saturation point (equivalent to the common concept of "carrying capacity") is entirely contingent on the current system of production, and will change with any modification of the technological and institutional framework. Thus, equilibrium levels or carrying capacity are relational variables that are in practice subject to ongoing variation. Assuming with Boserup that technological change is itself spurred by increases in population, the model in Figure 3.3 suggests that further improvements in the system of production and factor endowments (such as new land, climatic change, better tools, more efficient ways of organizing labor and capital, etc) may shift the theoretical saturation point (from S to S') well before average surpluses have in fact disappeared.[58] The equilibrium level changes concurrently with the system of production, albeit not always at the same pace. In the very long run, these parallel movements facilitate net population growth.

(b) *Productivity change*

That this long-term trend obtained for Greco-Roman antiquity as a whole is not in doubt (see below, IV–VIII).[59] At the same time, it is the specific correlation between output, population size, and surpluses in particular areas and periods that is pivotal to ancient economic history. In this, the pace of innovation in the system of production and the variability of conventional living standards are the critical variables. While minor adaptations in productive technology and organization would occur on a regular basis, major improvements that permitted rapid demographic expansions were probably rare or even non-existent. Why do systems of production change at all? Technical inertia – the rational resistance to change that entails costs and especially novel risks – tends to keep systems of production from changing unless and until they are compelled to do so: decline in per capita output associated with rising population size raises the utility gain from innovation and therefore may well be the principal inducement to productivity-enhancing adaptations.[60] Under these circumstances, the pressure for innovation is a function of the average level of well-being relative to the subsistence level which is in turn determined by the ratio of population number to current carrying capacity. Moreover, the rate of innovation is positively correlated with population size *per se*. It has been formally demonstrated that in the long term, population size and the stock of knowledge (which governs productivity) are not only mutually interdependent but serve as the two fundamental causal determinants of economic progress, while any other

58 Boserup 1965, 1981. In practice, S is a curve rather than a point: Wood 1998: 108.

59 *Pace* Horden and Purcell 2000: 267, emigration – i.e., the lateral expansion of a given system of production – cannot mitigate Malthusian pressures in the long term.

60 Boserup 1965: 65–9; Grigg 1980: 144; Wood 1998: 108, 111.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

factors are ultimately reducible to secondary endogenous variables.[61] The main reason is that larger populations create larger markets for information about technology and institutional arrangements.[62] Although the effects of this linkage are at best dimly perceptible for ancient economies, it seems that the rate of population growth in Greece and Roman Italy may have been positively correlated with the rate of inventions.[63] The principle that population size also contributes to economic development independently of population pressure is borne out by the observation that population growth tends to be proportional to population size in the sense that in the long term, rates of increase depend on previous population levels, a correlation that is empirically supported from at least 4000 BC onwards.[64] As a consequence, low-density areas would have been less conducive to innovation, whereas growing population density and concurrent urbanization (see below, VIII) can reasonably be expected to have accelerated technological and institutional progress. However, the *extent* of any such developments in any period or place within the Greco-Roman world must not be overestimated.[65] On the whole, processes that facilitated gross quantitative growth, above all the extension of cultivated land enabled by colonization and imperial pacification, may well have been of greater import than qualitative improvements of the existing resource base. In so far as systems of production expanded laterally rather than in terms of intensity, inducements for innovation must have been comparatively weak. Indications of technological responses to population pressure are only rarely discernible in the record: mainland Greece in the fifth and fourth centuries BC is the main example (see below, VI). The issue of demographically induced institutional change still awaits investigation.[66]

A variety of causal relationships may be consolidated into a comprehensive model of the interdependence of demographic and economic development (Figure 3.4). It is clear from this chart that without the flow from population size to technology, the whole system would tend towards a stable equilibrium, perturbed by exogenous forces but always returning to a steady state. Thus, as noted above, it is the capacity for building up and applying knowledge that is the pivotal determinant of long-term change, and, in so far as innovation is in turn a corollary of population size, the latter is an equally critical variable. Figure 3.4 takes a long-term view: the inversion of the conventional positive correlation between increasing knowledge and real wages on the one hand and rising fertility on the other is a late development, as is the amelioration of health hazards through knowledge.

[61] Simon 2000. Cf. Johnson 2000. [62] Jones 1988: 89.
[63] Simon 1986: 66–7. [64] Kremer 1993: esp. 706–7.
[65] Cf. Elvin 1973: 298–315; Jones 1988: 141–6; Mokyr 1990: 218–38; Deng 2000: 13–14, on the limitations of this process in pre-modern China. See below, Chapters 5–6.
[66] Cf. North 1981: 133, on mediaeval Europe.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

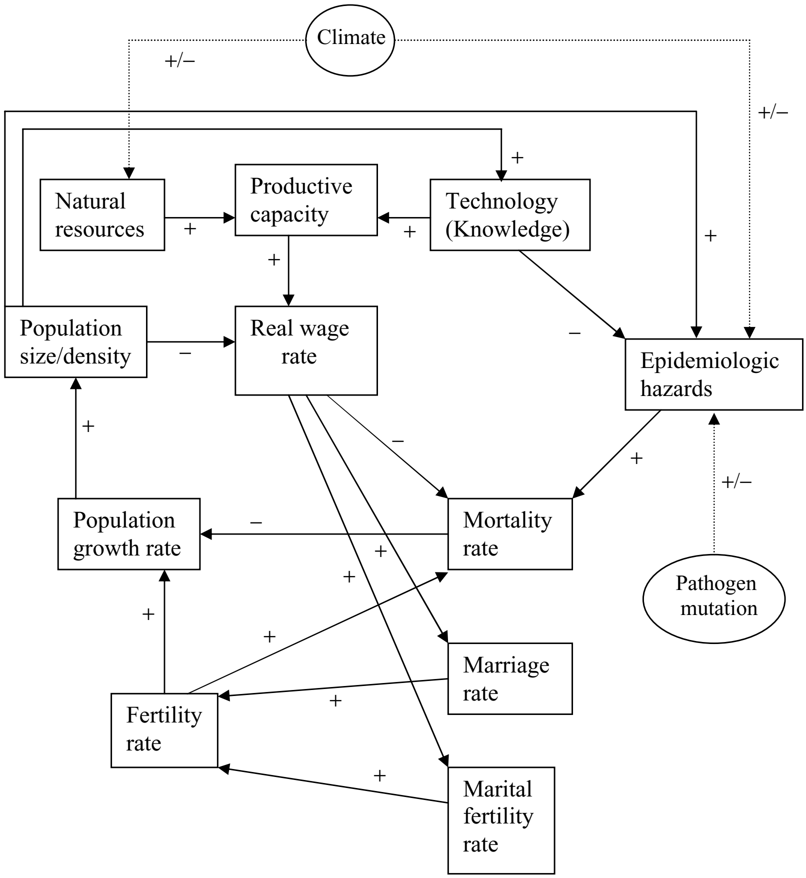


Figure 3.4 Causal relationships in a schematic model of homeostatic population regulation in pre-transitional populations
Source: Considerably expanded and modified version of Schultz 1981: 31 fig. 2.9

Key: Arrows indicate the effect of an increase in the base variable on the target variable (+for increase/improvement, – for decrease/deterioration); a decrease in the base variable has the opposite effect. Variables represented by squares are fully or partially endogenous, while variables in ellipses are entirely exogenous. Institutions are integrated into the model in that they are instrumental in mediating the relationship between various variables (such as between natural resources/technology and real wages, or between real wages and the fertility rate), and are best visualized as "acting on the arrows." The envisaged population as a whole is closed, and migratory flows within the system will on average cancel each other out. Migration and urbanization processes are subsumed within the relationship between population density, productive capacity and real wages. Relationships that are in the first instance confined to industrialized settings (such as the impact of technology on climate or pathogen mutation) have been disregarded.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

Ancient economies, just as any other pre-modern systems, were ultimately governed by the causal relationships mapped out in the model.

(c) *Living standards*

The relationship between normative living standards and physiological subsistence is the other main variable that mediates pressure for innovation. To the extent that minimally acceptable living standards exceed bare subsistence (defined as the level of consumption enabling survival and demographic reproduction at replacement level), demand for improvements in productivity will emerge before population growth reaches a final saturation point. The greater the average difference between these two consumption levels, the more capital reserves will be available, and the easier it is for innovation to occur. In principle, and in the very long run, normative living standards rise with technological advances in as much as the latter (which require human capital) increase the value of time and thus the price of children, thereby constraining fertility and permitting intensive growth (see below, VII). However, family farms would have been left largely unaffected by such developments.[67] In traditional agrarian societies, therefore, this process must have been unfolded very slowly, especially since in the long term, upward adjustments of average living standards during phases of productivity growth could be offset by rising inequality in resource distribution or reversed in periods of declining marginal productivity (see below, V). Given low levels of technological change, living standards may more commonly have been manipulated through institutional arrangements that governed the rate of surplus extraction via rents, taxes, tribute, serfdom and slavery, without significantly altering average consumption levels in the population as a whole.

(d) *Conclusion*

For all these reasons, ancient economies were probably caught in what has been described as a "low equilibrium trap," in which in the long term, limited increases in output will raise surpluses less than population size and the latter will eventually offset intermittent productivity gains (as envisaged in Figure 3.3). A lasting escape from this trap is not feasible merely with cumulative increases in modest investment over an extended period; rather, a "critical minimum effort" is required to generate self-sustaining intensive growth.[68] This is corroborated by the fact that technological progress depends partly on the quantity of the available surplus (also known as the savings rate): as innovation emerges primarily from the non-agrarian sector,

[67] Lee 1986b: 99. [68] Leibenstein 1954; Nelson 1956.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

the likelihood of successful innovation is a function of the size of the agrarian surplus that supports the non-agrarian sector (i.e., urban and rural industry), which equals the extent to which the average product of labor exceeds normative living standards. This principle holds true regardless of whether surplus is coerced by rent-seeking elites (arguably the most common pattern in antiquity) or created through the operation of competitive markets.[69] However, as indicated in Figure 3.3, while the preconditions for innovations are best at optimal population size – when the size of the non-agrarian sector could, at least in principle, peak in a given system of production – the demand for innovation grows only as productivity declines. Moreover, in the absence of systematic family limitation (see below, VII), optimal population size is inherently unstable.[70] (Barring effective birth control, an optimal population size could only be stabilized through increased surplus extraction without food re-circulation.) This particular configuration of factors in turn undermines any "critical minimum effort" to escape the low-equilibrium trap. In general, therefore, innovation was a self-limiting escape mechanism: in the words of James Wood, "over long stretches of time, population and food supply may leapfrog over each other, generating ever larger population sizes and more intensive systems of production but never gaining any permanent improvements in well-being."[71] The following three sections sketch out how and why this general model applies to the Greco-Roman world.

IV VARIATION IN POPULATION GROWTH AND SIZE

(a) *Short-term variation*

The baseline upward trend in ancient population number is modulated by two distinct layers of variation in the short and medium terms. Short-term variation is commonly caused by exogenous shocks such as climatic fluctuations that affect harvest quality. Comparative evidence from other pre-modern societies suggests that temporary output fluctuations are more likely to activate preventive checks (either via nuptiality, as in England, or via marital fertility, as in China) than to raise mortality. While fertility tends to be highly sensitive to grain prices, the responsiveness of mortality is inversely correlated with the level of development.[72] Dearth was often a necessary but not a sufficient cause for higher death rates: in fact, the effect of harvest variation on mortality appears to have been mainly indirect, via migration and exposure to disease, and was mediated by social mechanisms

[69] Lee 1986a: 101–2. [70] Wood 1998: 113.
[71] Wood 1998: 113. Jones 1988: 85–146 surveys depressants of intensive growth.
[72] Galloway 1988 is the most comprehensive survey; see also Galloway 1994. Different regional responses: Lee in Wrigley and Schofield 1981: 356–401; Campbell and Lee 2000.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

that govern the distribution of the impact of scarcity.[73] Sweeping famines associated with less differentiated starvation were generally rare.[74] Although there are no data from the Greco-Roman world that would allow us to trace the relationship between economic and demographic variation in the short term, it is reasonable to assume that the picture was very similar. Inter-annual fluctuations in precipitation and harvests are well documented for the modern Mediterranean and must have been equally common in antiquity.[75] Sporadic evidence from Roman Egypt reveals local inter-annual price variations.[76] Buffering mechanisms such as long-distance transfers of foodstuffs must have been largely confined to the urban sphere.[77] Temporary food crises greatly outnumbered famines that led to mass mortality.[78] While there is no way of knowing whether harvest failures primarily affected marriage rates or marital fertility, we may speculate that given early and universal marriage for women and social acceptance of post-partum interventions (see below, VII), responses to scarcity may very well have been concentrated within marriage. At low levels of development, the negative effects of variability of food consumption on survival are conducive to risk-averse behavior among poor agriculturalists: indeed, the scale of inter-annual food storage has been found to be positively correlated with average life spans.[79] Because of this, it is preferable to view risk-averse behavior as a function of pre-existing economic constraints rather than an exogenous, cultural brake on development. High levels of vulnerability to scarcity and the benefits from conservative allocation strategies would have reinforced the "low-equilibrium trap" by discouraging productive innovation.

(b) *Medium-term variation*

For the historian, variations in the medium term – on the scale of generations or centuries – are more important because they help account for differentiation within and between conventionally defined historical periods. For heuristic purposes, we may distinguish between four ideal types of population growth: first, negative growth – i.e., contraction – brought about by a variety of exogenous (climate, certain kinds of disease) and endogenous factors (war, systems collapse); second, growth during the recovery from demographic contractions or depressions; third, growth in response to endogenous productivity gains that raise local carrying capacity; and fourth, "lateral" growth caused by the transfer of more efficient systems of production (technology, organization of labor) to other regions.

[73] Walter and Schofield 1989b: 48–57. Gender preferences also play a role: e.g., Tsuya and Kurosu 2000, and esp. Bengtsson, Campbell, Lee et al. 2004.
[74] Watkins and Menken 1985. [75] Garnsey 1988: 10–14; Ruschenbusch 1998.
[76] Rathbone 1997a: esp. 213. [77] Cf. Hopkins 1983a.
[78] Garnsey 1988: 3–39; cf. 1998: 272–92. [79] Ortega Osona 2000.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

Negative growth varied widely in terms of scale and scope. Greco-Roman history as a whole is bounded by two major contractions: at the end of the Mycenaean period in the Aegean (c. 1200–1000 BC; see Chapter 8), and more universally from the fifth to the seventh centuries AD in the wake of the disintegration of the western Roman empire and the subsequent pandemic of bubonic plague (see Chapter 2, and above, II). In between, widespread epidemics in the late second and the mid-third centuries AD may have depressed population levels in many parts of the Roman empire. The loss of perhaps more than half of Athens' adult male citizenry during the Peloponnesian War in the late fifth century BC is by far the most conspicuous case of severe but strictly localized attrition.[80] Although massive and prolonged contractions may cause productive techniques to regress and carrying capacity to drop, demographic recovery not only created pressure to restore earlier productivity levels but commonly propelled output and population size beyond any previously attained levels.[81] This scenario implies a combination of the second and third categories of population growth. The ongoing growth of output and population in Greece from the Early Iron Age through the classical period – for a total of 500–600 years – is the principal example from antiquity. Notwithstanding its exceptional duration, this process structurally resembled the European economic and demographic expansion in the High Middle Ages, from about AD 1000 to 1300, or the shorter recovery after the Black Death and subsequent net expansion in the late fifteenth and the sixteenth centuries. China's growth phases during the Song and Ming dynasties conform to the same pattern.[82] Overall, however, these instances of "efflorescence" – simultaneous increases in population and productivity – were comparatively rare in pre-modern history and ordinarily failed lastingly to overcome the constraints of the "low-equilibrium trap" (see above, III).[83] More frequently, demographic recovery events would merely compensate for prior population losses or even fall short of this measure: examples include fourth-century BC Athens and parts of the Roman empire after the epidemics of the second and third centuries AD. Unlike the previous two types, "lateral" growth is exogenous in the sense that it entails changes in the system of production that do not occur in response to internal inducements, and thus falls outside the scope of homeostatic models. The establishment of Greek overseas settlements in the Mediterranean and Black Sea regions from the eighth to the fifth centuries BC, the migration of Greeks and Macedonians to the Near East during the Hellenistic period (e.g., the Fayyum in Egypt), and Roman colonization within and later outside Italy all belong in this category.[84]

[80] Hansen 1988: 14–28. [81] Boserup 1965: 62–3.
[82] Grigg 1980: 7; 53; Goldstone 1991: 25. [83] Goldstone 2002.
[84] See below, Chapters 10, 16, 18.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

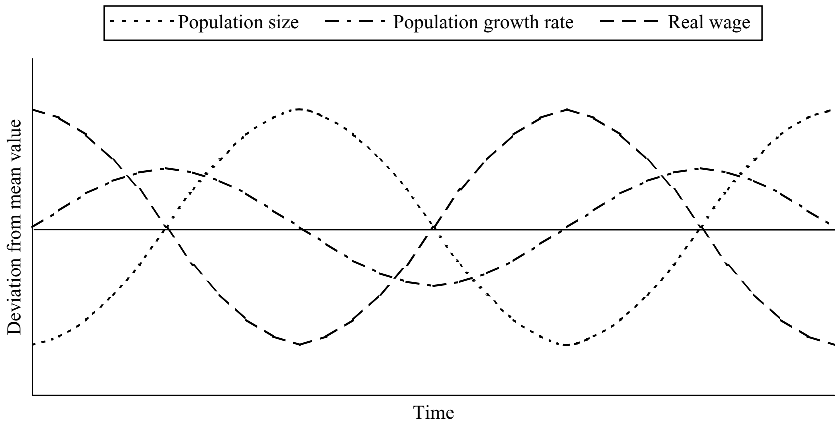


Figure 3.5 Model of population regulation driven by exogenous change in the population growth rate
Source: Lee 1986b: 87 fig. 3

In keeping with the general model of homeostatic population regulation outlined in the previous section, the most reliable comparative data suggest an inverse correlation between population growth (and life expectancy) on the one hand and real wages (i.e., average surpluses) on the other. This relationship is best documented for England from the thirteenth up to the early nineteenth centuries.[85] The specifics of this correlation (presented in idealized form in Figure 3.5) indicate that while wages respond sensitively to deviation of population size from the equilibrium level and fertility responds weakly but positively to ensuing changes in real wages, thereby acting as a proximate determinant of population growth, the observed swings are ultimately driven by exogenous variation in the growth rate caused by the incidence of mortality rather than by endogenous forces.[86]

This causal interconnection underscores the homeostatic tendencies of pre-transitional economic-demographic systems and the importance of mortality shocks and secular shifts in the severity of the disease environment, and shows that population size and wage levels cannot deviate too widely from the equilibrium level without triggering countervailing mechanisms.[87] For all we can tell, exogenous shocks – in the form of climatic and epidemiological changes – were similarly important in the Greco-Roman period. The impact of epidemics has been briefly considered above (II). More generally, long-term fluctuations in population size are strongly

[85] Lee 1973, 1980, 1986b, 1987: 447–8. Real wages are important indicators of well-being: Schofield 1989: 289–90 shows an inverse correlation between real wages and mean life expectancy at birth, and a positive match between the Crude Marriage Rate and real wages.

[86] Lee 1980: 536–40, 1986b: 100. See also Tsoulouhas 1992; Reher and Osona 2000: 185–6.

[87] Wood 1998: 110; Reher and Osona 2000: 205.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

influenced by long-term climatic variation.[88] In middle latitudes, warmer winters were historically associated with accelerations in population growth: in early modern Europe, for example, a 1 degree C increase in mean winter temperature would raise the annual natural growth rate by 0.1 percent.[89] Warming periods induced similar expansions in Greece after the post-Mycenaean depression and in the western Mediterranean from about 200 BC to AD 400.[90] Unfortunately, evidence of the movement of real wages or overall well-being relative to population size or growth is almost entirely missing from the record. In what may arguably be the only exception, price and wage data from Roman Egypt in the second and third centuries AD seem to indicate a rise in real incomes in the wake of the demographic contraction caused by the "Antonine Plague."[91] If correct, this reading is consistent with the model derived from later European data in Figure 3.5. Overall, in the absence of comparable statistics, we must rely on the interpretation of features that can be taken as proxy variables suggestive of levels of average well-being.

V VARIATION IN WELL-BEING

Secular changes in ancient living standards are only beginning to be systematically explored. Assessing a bundle of features including mortality, morbidity, nutrition, housing and clothing, Morris argues for significant improvements in well-being in the Greek Aegean between the tenth and the fourth centuries BC.[92] No comparable survey currently exists for the Roman world, and although we may suspect similar developments to have occurred in Roman Italy during the Republican period, different institutional characteristics (associated with higher levels of socio-economic differentiation and political inequality) may conceivably have limited the degree to which the majority of the Italian population benefited from economic growth. These issues will be explored in several of the following chapters. In general, the increasing availability of certain goods, such as metal tools and implements, and the distribution of manufactured wares serve as a rough index of long-term growth in per capita output.[93]

(a) *Nutritional status*

Yet more than anything else, in an environment in which food intake would account for perhaps two-thirds of the consumption budget of many commoners, it is changes in food availability and diversity that are crucial to any appraisal of living standards. Despite the introduction and dissemination

[88] Galloway 1986. [89] Lee 1987: 456–7. [90] See Chapters 2 and 8.
[91] Scheidel 2002. [92] Below, Chapter 8, and Morris 2004, 2005.
[93] Hopkins 1978b: 55–9; Gunderson 1982: 247–50.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

of new crops (see Chapter 2), it remains doubtful whether the majority of the population gradually came to enjoy improved diets.[94] Moreover, nutritional status arises from the balance of all energy inputs and outputs, and since food not only maintains the body but also sustains work and combats parasites and infections, food intake *per se* does not determine nutritional status or correspond to physiological well-being. Fluctuations occur in food supply as well as energy demand. The impact of disease is a key factor for the latter: the presence of disease determines the nutritional status which in turns affects susceptibility to disease. At the same time, some diseases are so virulent that they act independently of nutritional status. Moreover, in demographic terms, moderate levels of malnutrition may even be beneficial by affording protection against certain infections and by reducing fecundity. Because of the complexity of these interactions, evidence of poor nutritional status can be no more than a vague indicator of the quality of food intake.[95] Ancient data for nutritional status are available in the form of trace elements in skeletons, osteological markers of deficiency conditions, and measurements of adult body height.[96] Widespread evidence of deficiency disease as indicated by stress markers on bones and teeth clearly point to a high incidence of low nutritional status. However, the fact that average height for both sexes from Greek and Roman sites in the Mediterranean matches the means for Naples in the 1960s where life expectancy was much higher inspires little confidence in extrapolations from stature to demographic conditions.[97] Perhaps most importantly, information on the Body Mass Index (BMI) – the principal indicator of well-being – does not exist for antiquity. As a result, we can plausibly assume but not independently demonstrate that Robert Fogel's claim that poor nutritional status rendered a significant proportion of early modern European populations unfit for sustained work, or observations on economically disruptive energy conservation efforts of low-BMI populations in developing countries today, also hold true for the impoverished strata of ancient populations.[98] It is uncertain to what extent improved nutrition could substantially reduce mortality in populations that were exposed to numerous sources of infection.[99] In more recent pre-modern societies, higher incomes ensured higher food intake but did not necessarily alleviate the impact of infectious disease. This is borne out by data concerning European elites prior to the mid-eighteenth century when survival

[94] Garnsey 1999: 118–20. See Chapters 14 and 22.

[95] See Walter and Schofield 1989b: 17–21 for a summary; Carmichael 1983 (protection); Rotberg and Rabb 1985: 308 (virulence); Waterlow 1996: 106–7 (fecundity).

[96] Morris 1992: 97–100; Garnsey 1999: 43–61.

[97] Ibid. 58. There are additional reasons to doubt a universal correlation between height and life expectancy: Riley 1994.

[98] Fogel 1993: 10–13; Waterlow 1996. [99] Johansson 1994.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

rates still depended more on location (i.e., climate and population density) than on socioeconomic status.[100] Evidence of low life expectancy among the urban top echelons of Roman imperial society implies a – in any case *a priori* plausible – similar scenario for the ancient world.[101] For the same reason, temporary economic-demographic "efflorescences" need not have resulted in substantial improvements in nutritional status or overall well-being, especially in as much as they led to higher population densities and levels of urbanization and thus to new health hazards (see below, VIII): after all, even modern economic growth, although it unfolded on a much more dramatic pace and scale, initially failed to improve general well-being.[102]

(b) *Inequality*

In any pre-modern population, where the majority of the population produce their own food, the direct link between average per capita output and average well-being is fairly well defined. Nevertheless, social and political differentiation complicates this relationship except in the most egalitarian systems.[103] The impact of differential resource allocation on carrying capacity (i.e., the equilibrium size of a population) is positively correlated with the extent to which normative living standards exceed minimum physiological subsistence. In the kind of largely agrarian subsistence economy that was typical of most of the Mediterranean for most of the period under review, transferable surpluses were generally small and their manipulation by rent-seeking elites could not greatly affect the demographic makeup of the rural majority. Moreover, changing levels of surplus extraction would not affect average per capita consumption as long as food rents were recirculated into the general population.[104] Although the Roman imperial world-system transformed parts of the Mediterranean into net exporters of staple foodstuffs, the relative share of such transfers in total local production must normally have been small (unlike for cash crops).[105] It is only when institutional arrangements interfere with normative living standards that they begin to affect gross population size. Chattel slavery is the most important example: in so far as a slave could be coerced to work harder and/or consume less than the average non-slave worker, a slave economy would support a larger population than a traditional family-based system of production. This mechanism facilitated the influx of large numbers of

[100] Livi-Bacci 1991: 63–78; Johansson 1994: 113–14; see also Lee, Wang and Campbell 1994: 401 (China). For contrasting findings, cf. Flinn 1981: 18 (Geneva).

[101] Scheidel 1999. [102] Kuznets 1966; Fogel 2000: 139–63. [103] Wood 1998: 107–8.

[104] If higher extraction detracts from normative living standards, the population will work harder to create additional income, and initially population size may rise as a result; in the long term, fertility will decline in response to this squeeze, and population growth will be reversed: Lee 1986a: 109.

[105] Cf. Hopkins 1983a.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

slaves into areas that were already experiencing population growth, specifically central Greece in the archaic and classical periods and parts of Roman Italy and Sicily during the Republican period.[106]

Whereas average well-being rises during an expansion before dropping with declining marginal productivity (as indicated in Fig. 3.3), variation in well-being may also rise but then become "crystallized" instead of abating under population pressure. In this case, increasing material inequality could be regarded as an endogenous outcome of economic and demographic growth.[107] To what extent this process accounts for economic differentiation in ancient societies remains to be investigated: Greece in the late classical and early Hellenistic periods and Roman Italy in the late Republic and the early Empire are likely candidates. This trend towards rising inequality would be exacerbated by generalized "demographic differentiation" as defined by Chayanov, when even with identical access to resources, stochastic variation in household histories gradually causes living standards to diverge.[108] Since inter-family differences in birth rates are positively correlated with the overall level of fertility, periods of growth are necessarily conducive to rising variance in well-being.[109] At the same time, however, actual – as opposed to predicted – variance is strongly influenced by institutional arrangements. Thus, the Athenian democratic system arguably constrained tendencies towards increasing differentiation much more effectively than Roman institutions, especially under the Principate. This is illustrated by the fact that uncommonly egalitarian landholding patterns in classical Attica point to high normative living standards, whereas extremely high indices of inequality in landholding in parts of the Roman empire are likely to reflect rural poverty.[110] Even so, it deserves attention that even the supposedly significant gap between physiological subsistence and normative living standards (which in turn determined the demographic equilibrium level) in classical Greece did not in the long run engender sufficiently rapid innovation to outpace population growth and sustain lasting intensive growth.

VI POPULATION PRESSURE

Evidence of population pressure is of particular relevance to our understanding of the interaction of economic and demographic developments and the

[106] See below, Chapters 10, 12, 19.

[107] Wood 1998: 115, 128. Cf. ibid. 115–16 for a model of the development of variation in well-being.

[108] Chayanov 1986: 254. For (computationally deficient) models of ancient household life cycles, see Gallant 1991.

[109] Wood 1994: 33–6, 1998: 116. See also Goldstone 1991: 32–3.

[110] For Gini coefficients, see Morris 2000: 140–1 (Athens); Duncan-Jones 1990: 138–9 (Roman empire). See also Hopkins 1978a: 1–96 on the process of differentiation in Roman Republican society.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

constraints of a given system of production on sustainable growth. In the almost complete absence of statistical information on real wages, we must fall back on indirect diagnostic signs of declining marginal productivity and relative overpopulation, such as the fragmentation and subdivision of farms, rising land prices and rents, landlessness, and especially agricultural intensification and increased cultivation of marginal land.[111] For much of the period and region under review, existing productive technologies appear to have been sufficiently flexible and expandable to accommodate lengthy phases of moderate demographic growth. It is worth remembering that "overpopulation" is independent of population size *per se* but reflects average well-being relative to the subsistence level. In fact, given the positive correlation between population density and innovation rates, low-density populations may be more vulnerable to population pressure than others.[112] With regard to Greek and Roman societies, the conspicuous growth of chattel slavery is perhaps the most telling sign that productive arrangements could readily be adapted to sustain a growing population. Although precise numbers are unavailable, there can be no doubt that the number of foreign slaves entering central Greece in the archaic and early classical periods greatly exceeded the number of settlers leaving the region for other parts of the Mediterranean or the Black Sea.[113] The same is even more true for Roman society and Italy in general during the last two centuries BC: on a cautious estimate, several million slaves must have been transferred to Italy while the total number of Roman citizens of Italian origin residing overseas probably did not exceed 750,000 by the beginning of the Augustan period.[114] Boserup has argued that in the absence of sufficient population pressure, intensification requires coerced labor.[115] The Roman villa system may arguably be understood in these terms. The fact that in both central Greece and the core areas of Roman Italy, slaves dominated the domestic service sector and were disproportionally represented in the cities also speaks against widespread population pressure and falling living standards in the countryside. Even Egypt, traditionally the most densely populated region of the ancient Mediterranean, could support an expansion of domestic slavery in the Ptolemaic and Roman periods, alongside immigration from the Aegean and the creation of a new capital city.[116]

111 See Grigg 1980: 20–8 for a summary of comparative evidence.

112 Wood 1998: 111, 114.

113 Scheidel 2003b. Greek overseas settlement may have been motivated by relative advantages in factor prices but is not necessarily indicative of saturated production systems in the source regions. Moreover, emigration seems to have absorbed only a moderate share of total natural growth in the Aegean.

114 Scheidel 2005a, forthcoming, c (slave numbers); Brunt 1987: 263 (citizens).

115 Boserup 1965: 73–4.

116 Roman Italy: e.g., Hopkins 1978a: 68; Jongman 2003a. In Egypt, slavery is associated with Hellenization: e.g., Clarysse and Thompson 2006.

Nevertheless, natural growth together with the influx of coerced labor would eventually outstrip local capacity for ongoing economic growth. In mainland Greece, the dispersal of settlement into the countryside from the sixth to the fourth centuries BC is indicative of intensification. Signs of terrace-agriculture in late classical Attica can be taken to represent the logical progression of this process.[117] The most extreme case among Greek city states is the island of Aegina which could feed only a minority of its population and, while also depending on imports, may have practiced marling to increase agricultural productivity.[118] Greek preference for barley (with higher outputs than wheat) has also been linked to the need to support the largest possible number of people.[119] Overpopulation would help account for endemic warfare from the 430s BC onwards.[120] In Roman Italy, the social and political instability of the last century of the Republican period has recently been associated with rising population pressure and competition for cultivable land.[121] This trend may have continued into the early imperial period when even marginal land appears to have been in high demand.[122] While Roman agronomists described systems of crop rotation that reduce fallowing and raise total output, it remains unclear to what extent they were in fact adopted, especially outside large profit-oriented estates.[123] Given that the shift to the more productive three-field system in late mediaeval Europe provides powerful evidence of intensification and population pressure, it is all the more unfortunate that the existence or scale of comparable processes in the Roman empire cannot be determined.[124] Later ancient evidence suggestive of overpopulation is limited to the eastern Mediterranean in the fourth to sixth centuries AD, including Greece, Syria, Palestine, and Egypt: the expansion of villages in northern Syria is merely the best-known example of a more widespread trend.[125]

In the broadest terms, cycles of demographic growth in the Greco-Roman world appear to have conformed to the model set out in sections III–V, with parallels in high mediaeval Europe in general and early modern England,

[117] Morris 1994b: 363–4. For the problems of extrapolating relative population density from field survey data, see Sbonias 1999a, 1999b.

[118] Figueira 1981: 23–43; Strabo 8.6.16.

[119] Sallares 1991: 313–16.

[120] Goldstone 1991 argues that population pressure leads to societal breakdowns. Cf. also Chu and Lee 1994.

[121] Morley 2001: 59–61. "High" estimates of the size of the Roman citizenry (above, II) logically imply population pressure in Italy: Lo Cascio 1996b: 294–6. Cf. De Ligt 2004 for population pressure in the context of the "low" count.

[122] Evans 1980: 34–5; Dyson 1992: 115. For the possibility of population pressure in the early empire in general, see Frier 2001.

[123] White 1970a; Pleket 1993: 73–4, 75–9; Kron 2000. For large-scale grain farming on Roman slave estates, see Scheidel 1994b.

[124] But cf. Pleket 1993: 322–32.

[125] See esp. Tate 1992: 273–342 (Syria); Safrai 1994: 446–57 (Palestine); cf. also Villeneuve 1985 (Hauran); Bintliff and Snodgrass 1988 (Greece); Alston 2002: 357–8 (Egypt).

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

France and Ireland in particular.[126] The base-line trend of very slow population growth was partly overlaid by distinct growth cycles, above all in archaic and classical Greece, in Roman Italy in the Republican and early imperial periods, and in the eastern Mediterranean in late antiquity, that were eventually checked by a variety of factors: endemic warfare and subsequent emigration in mainland Greece, epidemics in the second and third and again in the sixth centuries AD in the Roman empire. The paucity of pertinent evidence makes it hard to tell whether other periods of growth likewise resulted in population pressure and exposure to positive checks, including the demographic expansions in western Europe in the Late Iron Age, in the Eastern Mediterranean and the Middle East in the Hellenistic period, in the "Barbaricum" of central Europe beyond the Roman frontiers, or in Mesopotamia in the Sasanian period. These uncertainties notwithstanding, what matters most for our understanding of ancient economic history is that there is no single case on record in which mounting population pressure precipitated a breakthrough to a markedly superior system of production that would have re-kindled intensive economic growth.

VII FERTILITY CONTROL

(a) *Natural fertility*

The preceding discussion of the interdependence of demographic and economic developments is predicated on the assumption that increases in production or real income tend to result in population growth that curbs intensive economic growth. Because of this linkage, fertility control is a critical factor in economic development. Prior to the modern demographic transition, virtually all known historical populations shared a characteristic age-specific distribution of fertility, known as a "natural fertility" regime.[127] In this scenario, the average incidence of childbirth rises steeply after menarche, peaks in the early twenties and subsequently declines at first gradually and then more rapidly, petering out in the forties. This pattern is a direct function of life-cycle changes in the physiological fecundity of the female body. By contrast, family limitation, conventionally defined as the deliberate cessation of procreation following the births of a particular number of children, often confines childbirths to the most fecund years and generates a different distribution pattern (Figure 3.6). The age distributions of fertility with and without family limitation differ regardless of the average level of fertility (i.e., the Total Fertility Rate, defined as the average

[126] Grigg 1980: 49–144. The same is true of China: Elvin 1973. No documented system could break this mold until the eighteenth century, except for small Holland from the sixteenth century onward (Grigg 1980: 145–234).
[127] Henry 1961; Coale and Watkins 1986; Wood 1994: 23–112.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

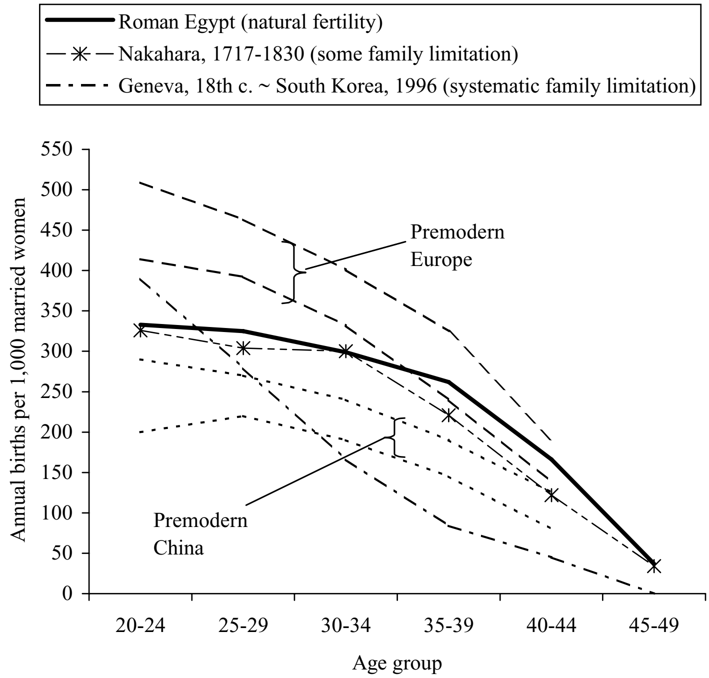


Figure 3.6 Levels of marital fertility in different populations
Source: Frier 1994: 325 table 1 (Egypt); Flinn 1981: 31 table 3.3 (Europe); Lee and Wang 1999: 87 fig. (China); Eng and Smith 1976: 175 (Nakahara); Henry 1961: 89–90, Korea Institute of Health and Social Affairs (Korea)

number of children born to each woman surviving to menopause). The only available sample of pertinent quantifiable evidence from the Greco-Roman period comes from the census returns of Roman Egypt. The documented incidence of live births is strongly suggestive of a natural fertility regime (Figure 3.6).[128] There is no indication of a "stopping strategy" governed by parity or maternal age.

(b) *Means of fertility control*

According to Ansley Coale's classic definition, birth control will not be applied unless it is "within the calculus of rational choice" of couples who want smaller families and have the means to act on their preference.[129] However, apparent absence of family limitation is not to be mistaken for a

[128] Frier 1994. On natural fertility in antiquity, see also Sallares 1991: 129–60.
[129] Coale 1973: 65.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

lack of fertility control in general. As Figure 3.6 demonstrates, age-specific birth rates may vary considerably under conditions of natural fertility. This is because the concept of natural fertility refers to reproduction that is socially controlled rather than by individuals on a parity-specific basis, and excludes instances of fertility control unrelated to parity, such as birth spacing or marital abstinence. Historically, birth spacing has been much more common than stopping. Thus, large variations in average reproductive performance have been achieved in the absence of anything like "conscious control" over procreation. Moreover, while the *concept* of rational choice in reproduction was widespread in pre-modern societies, *motivation* appears to have been the pivotal factor in the actual application of fertility control.[130] Considerable evidence from pre-transitional societies suggests that couples have always strategized about the configuration of their offspring, at least in terms of preference, and that they may often have tried to achieve their preferences even in the absence of deliberate stopping behavior, primarily by means of child abandonment and infanticide, sale, wet-nursing, adoption, loan, service arrangements, etc.[131] While none of these interventive measures reduces total fertility (as measured in live births), several of them would have a significant impact on early survival rates. Cumulatively, they can be instrumental in shaping the composition of individual families at various stages of the household life cycle, thereby adjusting the economic consequences of reproductive outcomes in ways not necessarily very different from preventive family limitation.

From a global comparative perspective, we may distinguish between three principal systems of household formation that correspond to particular modes of fertility regulation. In the most general terms, the historical "northwestern European" pattern was characterized by late marriage for both sexes (typically in the mid-twenties) coupled with high rates of both celibacy and marital fertility. In the "Mediterranean" pattern, women married earlier while men married late, whereas in the "Eastern" (i.e., East Asian) scenario, universal and early marriage was common for both sexes.[132] In the first case, given the lack of fertility control within marriage, marital birth rates were high (cf. Figure 3.6), and overall fertility had to be regulated proximately through the incidence and length of marital unions, as determined by average age at first marriage, the proportion of people ever marrying, and rates of remarriage. Limited differences in levels of marital fertility were largely a function of differences in breastfeeding practices which determined post-partum fecundability.[133] In China, the main example of the "Eastern" pattern, universal and early marriage for women meant that nuptiality could

[130] Spacing: Saito 1996: 545–6. Variation: Kirk 1996: 366. Concept vs. motivation: Alter 1992: 22–3.
[131] Mason 1997: 447–8. [132] Hajnal 1982. Cf. Viazzo 2003 for qualifications.
[133] Saito 1996: 544–6.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

not reliably serve as a preventive check.[134] Instead, comparatively low rates of marital fertility (cf. Figure 3.6) appear to have been ensured by a combination of prolonged breastfeeding, female infanticide, and male celibacy. Adoption, mostly of relatives, helped balance household needs.[135] In Japan, femicide may possibly have depressed post-partum survival rates.[136] Broadly speaking, the "Mediterranean" regime occupied an intermediate position between these two ends of the spectrum, with (moderately) early female marriage raising the importance of marital fertility control.

With regard to nuptial practices and household structure, Greek and Roman families appear to have conformed mostly to the "Mediterranean" pattern (see Chapter 4), though arguably in conjunction with some "(Far) Eastern" features. In the case of Roman Egypt, cultural preferences for extended breastfeeding may have contributed to extended birth spacing.[137] At the same time, the fact that the documented fertility distribution for Roman Egypt resembles that in the early modern Japanese village of Nakahara where about half of the residents practiced family limitation in the form of stopping behavior while the (wealthier) other half did not (Figure 3.6), suggests that even a putative "natural fertility" regime may in fact accommodate a certain degree of parity- and/or sex-specific birth control.[138] More importantly, postnatal interventions may arguably have been more prevalent outside Egypt.[139] Child exposure and infanticide are recurrent motifs in the Greek and Roman literary and legal traditions.[140] While these practices are unlikely to have been rare, even rudimentary quantification remains beyond our means, and reported sex ratios do not normally permit us to estimate the likely scale of femicide.[141] The latter may be a function of the size of dowries. There is little useful information on this issue from outside elite circles: dowries appear to have been generally important in Greece, where they represented the only (pre-mortem) inheritance for daughters and other relatives were obligated to provide them if necessary, whereas daughters' rights of inheritance in Roman society created more flexibility, and dowries among commoners may have been modest.[142] Compared to East Asian societies, post-partum measures were uncommon in later periods of European history owing to a set of circumstances that curbed fatal neglect of children, including religious doctrine, fear of punishment,

[134] Lee and Wang 1999: 63–82.

[135] Low fertility: Lee and Wang 1999: 47–51; adoption: 107–9. Their claim that these strategies were supplemented by deliberate birth control (including stopping behavior) is rejected by Wolf 2001.

[136] This is likewise controversial: for doubts, see Saito 1992.

[137] Masciadri and Montevecchi 1984: 32–5; cf. Tyldesley 1995: 69.

[138] Eng and Smith 1976.

[139] Scheidel 2001b: 45, on Diod. 1.80; Strabo 17.2.5.

[140] Eyben 1980/1; Boswell 1988; Harris 1994.

[141] *Contra* Pomeroy 1983; Brulé 1992. But cf. Clarysse and Thompson 2006 for potential indirect evidence for sex ratio manipulation in Ptolemaic census lists for Greek and Macedonian settlers in Egypt.

[142] Fox 1998: 166–20; Saller 1994: 204–24.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

collective institutions for foundling care, child labor in richer households, and large-scale wet-nursing.[143] It may well be relevant here that these constraints were weak or absent in Greek and Roman culture.

A broad array of putative contraceptives and abortifacients were discussed in ancient sources, and at least some of the recommended chemicals may have been effective.[144] However, the actual frequency of deliberate preventive or invasive means of birth control within marriage remains obscure: documented interest in the subject or knowledge of herbal agents fail to tell us about motivation or the context of actual application.[145] Ancient and comparative evidence suggests that these types of birth control may have been concentrated within competitive elites concerned about intergenerational status preservation.[146]

Divorce acts as a preventive check by reducing lifetime fertility. Quantifiable data are limited to the census documents of Roman Egypt, which report fairly frequent and often early marital dissolution.[147] In Athens and Rome, divorce was easy to obtain by both sexes. While its overall incidence remains unknown, we may suspect that it was significantly more common than in later Christian Europe.[148] Low rates of female remarriage must also have depressed total fertility.

Although unrelated to fertility control *per se*, adoption was another way to adjust family composition. Again, the evidence focuses on aristocratic practice and quantification remains impossible. Conceivably more common in Rome than in the Greek world, adoption was easier to undertake than in later periods of European history but there is no compelling reason to accord it a particularly important role in ancient household strategies.[149]

(c) *Households and fertility*

Recent research on the composition of ancient households has revealed a range of types from putatively nuclear (urban) households in the western provinces of the Roman empire to more complex families in the eastern Mediterranean, especially in Roman Egypt and Syria.[150] Family and household structures are critical in determining the nature of fertility regimes and their interaction with economic development. Broadly speaking, small nuclear families practicing neolocal marriage enjoy greater economic independence; for them, age of first marriage and even the number of children

[143] Lynch 2000. Even if some forms of childcare were thinly veiled instruments of infanticide – cf. Boswell 1988 – there is no doubt that overt infanticide had by then become culturally marginalized.

[144] Riddle 1992, 1997.

[145] Thus Frier 1994. Cf. Hopkins 1965.

[146] Johansson 1987. Cf. Salmon 1999; Caldwell 2004.

[147] Bagnall and Frier 1994: 123–4.

[148] Cox 1998: 71; Treggiari 1991b; Saller 1994: 219–20.

[149] Pomeroy 1997: 122–3; Saller 1994: 43. More common among Romans?: cf. Wentzel 1930.

[150] Saller and Shaw 1984 (cf. below, Chapter 4); D. B. Martin 1996; Bagnall and Frier 1994: 57–74; Sadurska and Bounni 1994.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

matter more than for extended families that buffer risk and help channel resource flows within rather than between households. In extended families, where spousal units are embedded in larger kinship groups, the cost of raising children is more dispersed, favoring early non-neolocal marriage for both sexes. In this latter category, the timing of marriage is less sensitive to economic conditions.[151] These differences may have consequences for fertility control, in as much as the welfare functions of the extended family create an incentive structure that inhibits responsiveness of marital fertility to economic opportunities or secular mortality decline. While the latter factor would not normally have been present to any significant extent in ancient populations (see below), reduced sensitivity to population pressure would have made intensive growth even more difficult to achieve. However, because in nuclear families, children act as risk insurance – in lieu of collective buffers provided by other co-resident kin – marital fertility need not decline even when per capita output improves.[152] Conversely, the linkage between economic and demographic developments among the nuclear families of early modern England depended on distinct features such as strict neolocality, life cycle servitude and systems of communal risk devolution that do not appear to have existed in a comparable way in ancient societies.[153] Rather, conditions in the Greco-Roman world resembled those in pre-transitional southern Europe (or modern Bangladesh), where disproportionate emphasis on kin relations through the male line fostered virilocality that limited the economic freedom of new couples; low levels of labor circulation between rural households ensured a strong commitment to family labor; patron–client links constituted the main welfare agency; and socially approved minimum living standards for marriage were low. This combination of features tends to facilitate high fertility.[154]

It therefore seems unlikely that ancient nuclear – let alone extended – families benefited from social or economic institutions that would have restrained marital fertility in the event of rising economic output. While the causes of secular fertility decline in modern history continue to be hotly debated,[155] none of them seems to have acted on ancient populations. The Greco-Roman world conforms best to the ideal type of a relatively undifferentiated economy of family farms and rural crafts in which economic activity is largely a family affair, labor is applied to capital in the family's control, and inheritance is the principal means of access to the means of production, so that the level of nuptiality is a function of mortality in the

[151] Cain and McNicoll 1988. [152] Cain 1981. [153] Smith 1981.

[154] Smith 1981: 617–18; Cain and McNicoll 1988. Cf. Viazzo 2003.

[155] Economic theories have centered on the changing cost and value of children, whereas more recent critiques emphasize cultural and ideational factors. For comprehensive overviews, see Kirk 1996: 367–81; van de Kaa 1996: 402–28; more briefly Alter 1992; Hirschmann 1994.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

previous generation, and the young tend to fill existing niches. It is only when a significant proportion of the population sells their labor that intergenerational links weaken and both inheritance and children lose value, and growing markets in goods and labor begin to serve as a preventive check on fertility.[156] No comparable changes in intergenerational wealth flows are visible in antiquity. Although urbanization may have been conducive to such changes, lower urban fertility would merely have reinforced the cities' demographic function as "population sinks" that helped regulate rural population growth (see below, VIII). Without a concurrent mortality decline, migration to the cities and the resultant increase in the specialization of labor might even have raised rather than depressed the fertility of rural households.[157]

The economic "low-equilibrium trap" that inhibits sustained intensive growth (see above, III) has an important demographic dimension: a low-productivity, high mortality and high fertility regime typically rests in an equilibrium state that can only be perturbed by a boost in capital and the stock of knowledge that favors increased investment in offspring.[158] For most of human history, the default position was a low-level equilibrium with little human capital and low rates of return on investment in human capital (i.e., education), associated with large families and low investment. Conversely, it takes high rates of return on human capital relative to return on children *per se* to increase investment to an extent that ultimately depresses fertility. Although the existence of large families does not necessarily imply demand for them, and the diffusion of cultural preferences may arguably matter more than micro-economic forces,[159] there can be little doubt that these crucial economic preconditions must obtain for changing preferences to translate to secular shifts in fertility levels. In the absence of dramatic changes in productive technology or (non-urban) labor markets, no transformation of this kind can have occurred in antiquity.[160] Fertility declines in developing countries have been correlated with improvements in the status of women: women need to benefit from a fertility decline in so far as their wage rate is inversely correlated with their fertility.[161] Again, no comparable process can be posited for Greek or Roman societies (see Chapter 4).

[156] Caldwell 1982; Schofield 1989. [157] E.g., Stark 1981.

[158] Becker 1988; Becker, Murphy and Tamura 1990.

[159] E.g., Bulatao and Lee 1983; Easterlin and Crimmins 1985.

[160] Although choice theories are predicated on the possibility of conscious fertility control by couples – Cleland and Wilson 1987 – this option may well have been available in principle, as suggested above. But see Scheidel 2001b: 37–44 for a critique of the ancient and modern construct of alleged Greco-Roman preference for small families or childlessness.

[161] Women's status: e.g., Mason 1985; Handwerker 1991; Federici, Mason and Segner 1993. Benefits: Schultz 1981: 6, 150–90.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

(d) *Mortality and fertility*

It is true that according to some comparative data, a fertility transition is not a necessary precondition for economic development; nor is the latter indispensable to the former.[162] Nevertheless, if mortality rates drop, intensive growth is unlikely to be sustained in the long term unless fertility follows suit. Although fertility transitions are now known to have occurred under diverse conditions, mortality decline is undeniably a necessary – if not a sufficient – condition for them to unfold.[163] Mortality decline is of crucial importance also because of its correlation with HALE (see above, II): lower death rates will necessarily be accompanied by improved health among survivors that raises returns on investment in human capital, thereby curbing fertility and facilitating productivity growth. No substantial mortality or morbidity decline is credibly attested for the Greco-Roman world or indeed any pre-modern population prior to European and Chinese elites in the eighteenth century.[164] Moreover, since a mortality decline may initially trigger higher fertility, the long-term depression of fertility by lowered mortality is largely a function of the labor system and the sources of income.[165] In the absence of fundamental economic transformations along the lines described above, even an increase in mean life expectancy – had it indeed occurred in certain phases of ancient history – would not have triggered a fertility transition or supported sustained intensive growth. In consequence, we must conclude that the principle of a relatively inelastic ratio of productivity to demographic growth that lies at the core of the model outlined above (III) is fully applicable to Greco-Roman economies.

(e) *Change and continuity*

In the most general terms, we can observe a mixture of continuity and change between antiquity and more recent periods of Mediterranean history. The intermediate level of marital fertility in Roman Egypt (Figure 3.6) is consistent with a mixture of "European" and "Asian" features of fertility control as outlined above. While it is unclear whether the "northwestern European" pattern of late marriage pre-dates the time of the Black Death, it does not appear to have prevailed in the western provinces of the Roman empire. It is possible that as Christianization increasingly marginalized traditional fertility depressants such as post-partum family limitation, adoption, and divorce, the age and frequency of first marriage gained importance

[162] Coale and Watkins 1986.
[163] Stressed by Kirk 1996: 368–9; van de Kaa 1996: 405–9. Chesnais 1992 refutes apparent exceptions to this rule.
[164] Livi-Bacci 1991: 63–7; Lee, Wang and Campbell 1994: 401. [165] Schultz 1981: 5.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

as the main preventive check.[166] More importantly, we have to allow for substantial diversity within the ancient Mediterranean and its hinterlands, and even between local communities, even if it defies empirical investigation.[167] Comparative evidence from Europe and East Asia shows that because demographic variables such as age of first marriage and celibacy rates are correlated with economic opportunities, their relative importance is prone to fluctuate over time.[168] The lack of serial data from any part of the classical world prevents us from tracking any such developments, in so far as they did in fact occur. However, notwithstanding considerable underlying demographic inertia, it would be misleading to reckon with rigorously stable demographic regimes that constrained economic development in an unchanging fashion. Rather, demographic and economic structures would co-vary over time. These shifts most likely took place over the medium term, as measured in generations or centuries, and would therefore be of particular relevance to our understanding of differences between conventionally defined periods of Greco-Roman history. Our inability to calibrate our analysis at the required level of resolution is perhaps the biggest impediment to the synthetic study of ancient economic and population history.

At the same time, it deserves notice that in the long term, different configurations of fertility control tend to generate similar homeostatic equilibria: in the final analysis, historical Total Fertility Rates do not vary much regardless of the typical age of first marriage or the average length of birth intervals.[169] In the same way, irrespective of the culturally mediated incidentals of proximate agency, ancient populations would necessarily regulate their fertility in a homeostatic manner. There is no sign that at the documented level of economic development and organization of labor, ancient Mediterranean populations were in a position to reduce average fertility in order to preserve intermittent productivity gains in a way that would have sustained long-term intensive growth. Among agrarian economies, "only populations blessed with the most advantageous institutions governing reproduction, surplus extraction, and use of surplus, would be able to (. . .) progress into the next-higher technological regime."[170] Ancient societies were not among them.

VIII URBANIZATION

(a) *The scale of urbanization*

Yet even in the absence of major breakthroughs, some intensive growth could and did of course occur, but is virtually impossible to measure and

[166] See Shaw and Saller 1984 for a partial rebuttal of Goody's 1983 thesis of sweeping demographic changes in the wake of Christianization.
[167] Cf. Wilson and Airey 1999: 122 for striking differences between contemporaneous villages in Tokugawa Japan.
[168] Schultz 1981: 29; Saito 1996: 542. [169] Wilson and Airey 1999. [170] Lee 1986a: 123.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

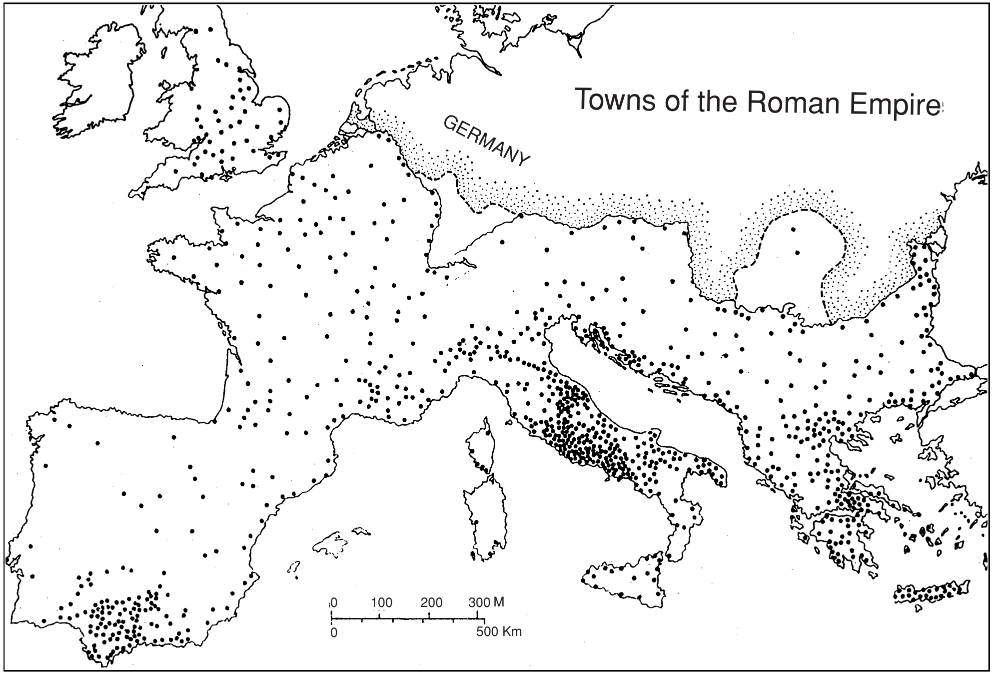
can only be estimated in vague terms (cf. above, II). This raises the question to what extent shifts in the distribution of population between countryside and cities, or more precisely between the agrarian and the non-agrarian sectors of the economy, were indicative of and correlated with real economic growth. Urbanization, whilst mediated by political and cultural factors, is ultimately a function of population density. Even if labor productivity and individual surpluses are small, rising density will be conducive to the expansion of non-agrarian production and urbanization. Although levels of urbanization also depend on prevailing transport technologies (to supply cities with vital goods) and modes of sociopolitical organization, in so far as these variables are themselves shaped by population size and density, they may to a certain degree be considered endogenous epiphenomena of underlying demographic conditions. And while transport by river and sea lowers the density threshold for urbanization by increasing the catchment area for urban sustenance, low population densities forestall long-distance transfers in as much as the supplying periphery needs to be endowed with sufficiently high concentrations of population to enable mass transport of goods to ports.[171] This precondition helps explain the role of high-density Egypt as the leading exporter of foodstuffs in the ancient Mediterranean, and may account for the fact that the Maghreb only gradually assumed a similar position, presumably as local population densities rose under Roman rule.

The average share of urban residents in the total population cannot be reliably determined for any particular region or period of the Greco-Roman world. It is however clear that for much of the period under review, the Mediterranean witnessed a gradual westward spread of cities. After the collapse of Mycenaean civilization, no genuinely urban communities appear to have existed anywhere west of Anatolia, Crete, and Egypt.[172] With the formation of urban nuclei in the Aegean and the foundation of Phoenician and Greek settlements in the west, the coastal areas of the Mediterranean came to be sprinkled with small cities, some of which flourished while others eventually contracted or failed. Major inroads were made under Roman rule, first with the proliferation of cities in Italy and subsequently with the expansion of urbanism into Spain, Gaul, Britain, the Danubian region and the interior of the Maghreb.[173] However, even at the climax of Roman urbanism, cities were largely clustered on the coast or along major rivers (Map 3.2). In the meantime, urbanism also thrived in the eastern half of the Mediterranean (Map 3.3).

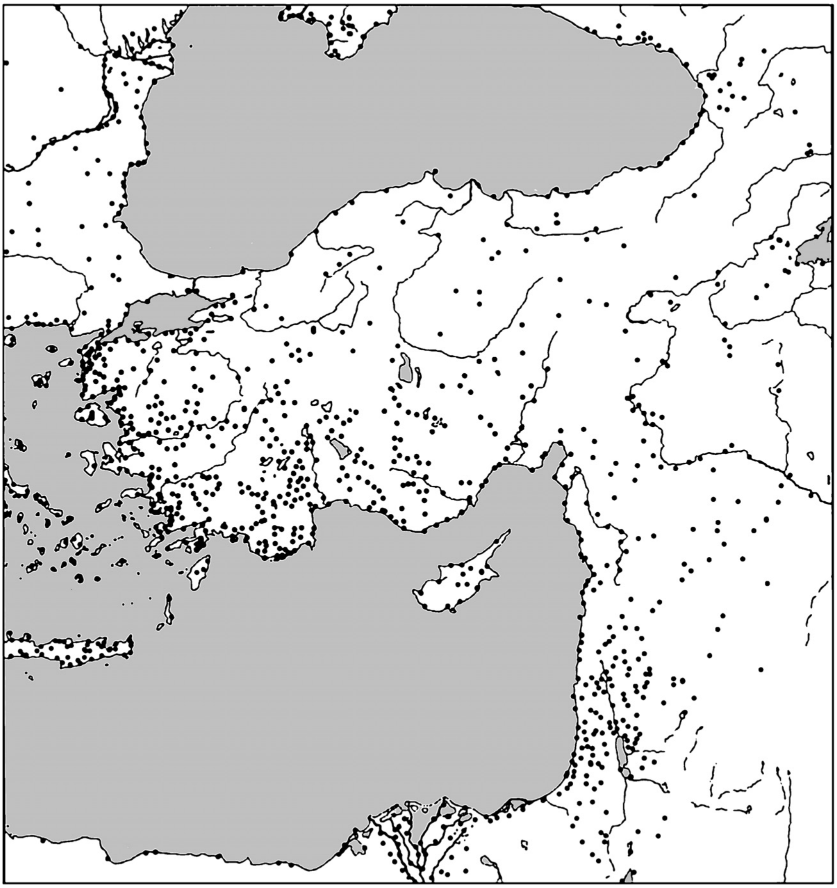
It would make little sense to estimate urbanization ratios for Greek *poleis*. By the classical period, a few million Greeks had come to be spread out across some 1,000 *poleis* in the Aegean and other parts of the Mediterranean and the

[171] Boserup 1981: 68, 74. [172] Cf. Morris 2006 for conditions in Greece.
[173] E.g., Lassère 1977; Bekker-Nielsen 1989.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press



Map 3.2 The distribution of cities in the western provinces of the Roman empire
Source: Pounds 1969: 155 fig. 10



Map 3.3 The distribution of cities in the eastern provinces of the Roman empire
Source: adapted from Jones 1937: maps II–IV; graphics by A. T. Wilburn

Black Sea. As a consequence, most communities were small and the urban–rural divide was exceedingly permeable. According to one guesstimate, fewer than half of all *poleis* comprised more than 2,000 members, and only 15 percent more than 5,000.[174] Under these circumstances, only the largest *poleis* (which did however account for a substantial share of the total Greek population) would be endowed with urban centers that predominantly consisted of non-farmers.

Large cities were newcomers to the scene: up to the fourth century BC, no city of 100,000+ residents is safely attested, and the permanent urban

[174] Number: Hansen and Nielsen 2004: 53–4. Size: Ruschenbusch 1985: 262; but cf. Nixon and Price 1990.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

population of Athens definitely, and that of Syracuse almost certainly, fell short of this mark.[175] While we cannot tell for sure when (or even if) Carthage reached this size, it was not until the third century BC that the city of Rome crossed this threshold, accompanied by Alexandria in Egypt and Antioch in Syria.[176] In the following centuries, these metropoles grew faster than the general population: by the second century AD, between 1.5 and 2 million people lived in the five largest cities of the Roman empire – Rome, with up to one million; Alexandria, with up to 500,000; Antioch, with at least 150,000; and Carthage and Ephesus. At its peak, Rome was as populous as the twelve largest cities of Christian Europe in 1500 taken together, on a par with the capitals of Song China and Tokugawa Japan, and unsurpassed in Europe until London reached the same size around 1800.[177]

Cities of that size had to be supported by catchment areas that only a very large empire could provide. Owing to divergent paths of historical development, the urban system of the early imperial period was more intensely primate in the western half of the empire than in the east. The main urban centers of the Levant – Alexandria, Antioch, Pergamon – had originally attained their size as the capitals of separate empires during the Hellenistic period and subsequently retained or even added to their substantial populations under Roman rule. In the west, by contrast, the Romans had expanded into comparatively weakly urbanized regions (and destroyed their main urban rival Carthage). As a result, with the exception of the re-established regional African metropolis of Carthage, the Latin half of the empire long lacked substantial secondary population centers: we cannot reliably identify a single western city other than Rome and Carthage whose urban population reached 50,000, and it was not until the emergence of Mediolanum and Ravenna as new centers of the late imperial period that significant changes seem to have occurred. In AD 330, the designation of Constantinople as a new capital city added a third mega-city to the eastern half of the empire.[178]

The total number of cities in the Roman empire is not known with precision but probably approached 2,000.[179] Urban status was a legal construct, and some small settlements may have been urban more in name than in

[175] See Morris 2006, who thinks that classical Athens probably peaked at c. 40,000.

[176] Alexandria: Scheidel 2004a; Carthage: Lancel 2000. Seleucia on the Tigris belonged in the same category, in a sense succeeding Babylon, which ought to have reached 100,000 (or possibly a much higher total) by the middle of the first millennium BC: cf. van de Mieroop 1997: 95–7. The size of Memphis is another unknown variable (cf. below, Chapter 16).

[177] Rome: Hopkins 1978a: 96–8; Morley 1996: 33–9. Lower estimates (e.g., Storey 1997) are unlikely: Lo Cascio 1997a; Scheidel 2001b: 51–2. Alexandria: Delia 1988; Antioch: Downey 1958; Liebeschuetz 1972: 92–5; Carthage: Gros 2000a: 542. Cf. in general Nicolet, Ilbert and Depaule 2000. Europe in 1500: de Vries 1984: 270–8.

[178] Dagron 1974. [179] Scheidel forthcoming, b.

nature.[180] Official bias in the recognition of urban status accounts for substantial regional imbalances: despite their similarly sized populations, Italy boasted over 400 cities while Egypt had only around 50; because of this, large Egyptian "villages" could be more populous than small Italian "cities."[181] Since the Roman empire grew out of a conglomerate of highly diverse sociopolitical formations ranging from earlier empires to tribal peripheries, it did not develop a single unified urban system that could profitably be subjected to rank-size analysis. As is typical of pre-modern economies, disproportionately large primate cities coexisted with numerous small urban communities while intermediate population centers were comparatively rare.[182] Italy is necessarily the most extreme example, where the capital may have accounted for up to two-thirds of the total population of cities of more than 10,000 residents. In Roman Egypt, Alexandria was perhaps six to eight times as populous as the next-largest city.[183] This pattern reflects the gravitational pull of rent-absorbing political centers and the corresponding weakness of the economic integration of urban economies within a given region.[184]

Relying on necessarily speculative and over-schematic calculations, I am inclined to posit an aggregate urban population of the Roman empire of the order of 7 to 9 million, approximately one-eighth or one-ninth of the total imperial population.[185] Regardless of its reliability, any notional average is bound to conceal substantial regional differences.[186] Over time, nucleation would often alternate with dispersal.[187] In earlier periods, very large proportions of the inhabitants of many Greek communities seem to have resided in the urban centers instead of their rural territories: Bintliff has argued that almost three-quarters of the population of classical Boeotia lived in cities, and drawing on an exhaustive survey of the size of walled cities throughout the Greek world, Hansen projects an overall urban/rural ratio of 1 to 1.[188] Most of these "urban" residents would have farmed land in the surrounding area. It is unclear to what extent Roman settlers in Republican *coloniae* at least initially resided behind city walls and whether comparable "agro-towns" also existed in Roman Italy.[189] More generally, we need to allow for the possibility that the full sociopolitical fusion of town and country that was a defining characteristic of Greek and Roman

[180] Cf. Paus. 10.4.1. [181] Beloch 1886: 438; Hopkins 1978a: 68; Rathbone 1990: 124–37.

[182] Smith 1982 (a "concave" pattern, as opposed to the lognormal distribution of city size found in developed economies). Cf. de Vries 1984: 85–120.

[183] Morley 1996: 182; Rathbone 1990: 119–21.

[184] Cf. Ades and Glaeser 1995 for the correlation between political inequality and urban primacy.

[185] Scheidel forthcoming, b.

[186] Urbanization rates were highest in the most densely populated parts of the empire: see Hopkins 1978a: 68 for Italy, and Scheidel 2001a: 247–8 for Egypt.

[187] E.g., Alcock 1993. [188] Bintliff 1997a: 235; Hansen 2006: 24.

[189] Garnsey 1998: 107–31, with my addenda 131–3.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

civilization raised the number of urban residents well above the size of the non-agricultural population.[190] For this reason alone, ancient urbanization tallies may not readily lend themselves to comparative assessments. Moreover, while computations of urbanization ratios for later historical periods do not normally recognize settlements of fewer than 5,000 (or even 10,000) residents as "urban," the majority of all "cities" of the Roman period must have fallen short of this threshold.[191] The magnitude of the resultant incompatibility of ancient and later urbanization estimates is highlighted by the observation that in Europe west of the Balkans in 1500, some 3,000–4,000 settlements enjoyed some form of urban rights whereas only about 500 of them had at least 5,000 inhabitants and are included in modern calculations of urbanization levels.[192] Nevertheless, in the context of my global estimate, it seems reasonable to assume that probably as early as 1300, and certainly by 1500, the formerly Roman territories west of the Balkans had re-attained and surpassed Roman urbanization levels above the conventional threshold of 5,000, especially if we control for the anomalous size of the city Rome in antiquity.[193] In terms of the total number of "urban" settlements, the Roman empire once again appears to have broadly matched the contemporaneous Han empire.[194]

(b) *Urbanization and economic development*

These tentative quantifications are of little relevance to our understanding of economic development unless we can correlate urbanization levels with the share in economic output of the non-agrarian sector. In the absence of usable statistics, we can only guess at the ratio of non-farmers to city-dwellers. The observation that in pre-modern societies, the former tend to outnumber the latter by up to 5 percent need not apply to economies in which the minimum size threshold for "cities" is very low and the population of many small towns of one or two thousand residents must have included a strong complement of farmers.[195] The presence of larger "agro-towns" would have skewed the picture even further.[196] Nonetheless, there is no good reason to believe that more than one person in eight would have been permanently or predominantly engaged in non-agrarian labor. In any case, whilst allowing for exceptions in particularly developed or

[190] Cf. below, n. 195. [191] Cf. de Vries 1990: 44–8 on definitional problems.

[192] de Vries 1984: 28, 67; cf. Pleket 1990: 145. Roman city sizes often cannot be determined with the requisite precision: cf. Duncan-Jones 1982: 259–77; Scheidel 2001b: 60–1.

[193] de Vries 1984: 42–3, 70, 72 (c. 9? percent in 1300); 36, 67, 270–8 (c. 11 percent in 1500).

[194] Nishijima 1986: 574. However, my estimate of the urban proportion of the Roman population is twice as high as (low) estimates for T'ang and Song China: Rozman 1973: 279–80.

[195] Bairoch 1989: 266. Cf. de Vries 1984: 22. About 6 percent of men in several (large) villages in Roman Egypt were tradesmen: Alston 2002: 335–6.

[196] See above, n. 188–9, and below, n. 202.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

privileged regions such as classical Attica or Roman Italy, we must assume that the proportion of non-farmers in the total population fell short of the 20 percent estimated for sixteenth-century England.[197]

More generally, urbanization may be envisaged as the outcome of any one of four processes: the concentration of a previously dispersed non-agrarian population of rentiers, craftsmen, traders, and even farmers in cities, without concurrent changes in population density or productivity; increasing population density at constant per capita output, creating a larger cumulative surplus that sustains larger, urban, settlements for the non-agrarian population; increasing per capita output at constant population levels, expanding the relative share of the non-agrarian sector and encouraging urban residence; and, finally, concurrent increases in population and productivity. In practice, none of these ideal types occurs in pure form, and the same is true for antiquity: the likely convergence of these inducements alone makes it impossible to gauge the contribution of each of them. Even so, the scale of the urban expansion in some parts of the Greco-Roman world, most notably in the north-western provinces of the Roman empire, suggests that gross population growth *per se* was a necessary but not a sufficient condition. Rather, institutional arrangements and even moderate levels of intensive economic growth appear to have been the main driving force behind the success of urbanism. For instance, the hundreds of new urban centers that sprang up in archaic Greece may have owed their existence to the strengthening of local government and the emergence of a nucleated rentier class.[198] In later periods, while the collection and re-allocation of resources by urban elites remained fundamental to the existence of cities, transcendent imperial power gradually became an even more critical determinant of urban development. By the early Roman imperial period, many cities had come to function as nodal points of a larger system of exploitation and transfers, converting local taxes and rents into exportable items of trade and cash. Without the exaction of resources that was caused or facilitated by imperial authority, elongated lines of trade and the resultant network of cities that was ultimately centered on the capital would not have emerged in the same way.[199] Because of these obvious differences in context, it would not make much sense to treat the Greco-Roman or "ancient" city as a stable and uniform institution or attribute to it a single function or location within a particular system of production as a whole.

Nevertheless, modern observers have frequently been preoccupied with attempts to define ancient cities in terms of ideal types and to identify

[197] Wrigley 1987: 170. [198] Morris 2006.

[199] Hopkins 1995/6. This network was best developed in Italy (Morley 1996), although overall levels of integration remained poor (cf. above).

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

their shared characteristics. In the wake of Finley's work, much debate has revolved around the distinction between Sombart's and Weber's concepts of the "consumer city" (in which a rentier class of landowners or officials draws wealth generated in the countryside in the form of produce from their own holdings, rents, and taxes, and by spending this income on retainers and artisans in the city creates an urban market for food and labor) and the "producer city" (supported by the production of goods that are exchanged for food and raw materials), and the application of these ideal types to the study of ancient economies. For Finley, Greek and Roman cities typically – though not exclusively – belonged in the former category because they relied much more on non-reciprocal rents than on trading or manufacturing for external markets, whereas critics are at pains to demonstrate the supposed significance – yet *de facto* often just the mere existence – of urban commercial activities.[200] This is not the place to revisit this increasingly stale controversy, and readers are referred to the more specific discussions of the urban economy of different periods and regions in later chapters. Suffice it to say that I find myself in agreement with Erdkamp's recent observations that since the model of the "consumer city" is primarily concerned with the economic foundation of the urban economy in its relation to the outside world (i.e., the mechanism employed in the extraction and transfer of agricultural surplus), it does not predict a particular level of economic development and is not logically associated with the notion of a "primitive" economy; that the concept is readily consistent with the presence of numerous urban artisans and merchants, and with "complex" urban economies in general; that evidence of inter-regional trade or manufacturing for export would not impinge on the model; and that even though reciprocal exchange through the market between food production and other sectors (urban or rural) did of course occur, non-reciprocal relationships predicated on social and political entitlement were sufficiently dominant to determine the nature of ancient economies within and beyond the urban sphere.[201] Once again, however, we must reckon with significant change over time. As Hansen has argued, many of the *poleis* of archaic and classical Greece conformed more closely to Weber's concept of the "farmer-citizen-city," in which the majority of community members resided in an urban core but farmed their own land and were sustained by its products.[202] In parts of the Roman empire, on the other hand, urban rentiers appear to have played a more dominant role.[203] Thus, political regime, overall levels

[200] Finley 1981: 3–23, 1999: 123–49. Jongman 1988a is the fullest case study. For criticism, see, e.g., Engels 1990; Parkins 1997; Horden and Purcell 2000: 105–8; Mattingly et al. 2001. On the debate, Whittaker 1995.

[201] Erdkamp 2001.

[202] Hansen 2004, with reference to Weber 1999 (1921): 67–8 ("Ackerbürgerstädte").

[203] See above, n. 110.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

of inequality, and the structure of landownership were critical determinants of the economic character of ancient cities.

It will always be easier to count cities and track shifts in their numbers than to reach agreement on the nature of the urban economies of the ancient world. From a pragmatic standpoint, therefore, what matters most in the present context is whether urbanization and economic development tend to be correlated in a predictable fashion, and more specifically whether cities, by their very existence, are conducive to economic growth. In principle, urbanization creates new problems – at the most mundane level in terms of the technology and organization of food transport, and beyond that regarding institutional arrangements – and therefore new opportunities and an intrinsic impetus for innovation.[204] On average, the mere presence of cities can be expected to have raised agricultural efficiency by creating novel inducements.[205] Thus, to the extent that urban elites increased rent extraction to support urban spending and the perceived amenities of cities encouraged migration, cities might be said to have acted in the same way as population density: while growing density raises demand per unit of land (an absolute increase), urbanization raises it by increasing the non-agrarian share of the population (a relative increase). Needless to say, these underlying mechanisms cannot be formally demonstrated for ancient economies, and the scale of such effects is a crucial issue that is hardly susceptible to empirical investigation.

Thanks to the predominance of non-reciprocal resource flows between the agrarian and non-agrarian sectors, ancient "consumer cities" might be regarded as "parasitical" and therefore considered an impediment to (rural) economic growth.[206] However, even parasitic cities may have positive effects on the countryside and on economic development in general: as Wrigley has pointed out, they are parasitical only with regard to the division of current flows of goods, not with regard to the creation of circumstances in which the flow of goods can be increased over time. For instance, barring unilateral exploitation, rural populations might increase in size without being better off per capita. In fact, the absorption of population by cities is an important safety-valve of high-fertility regimes, and might even prepare the ground for economic development: "The growth of towns tended to increase the likelihood of achieving a low-pressure equilibrium because the high death rates in towns meant that they were normally consumers of men."[207] Although urban excess mortality would need to be accompanied by relatively moderate mortality levels in lower-density rural areas to facilitate an eventual fertility transition – which was the case in early

204 Lee 1986a: 100; cf. North 1981: 132–5. 205 Wrigley 1990: 102.

206 On "generative" vs. "parasitic" cities, see Hoselitz 1954/5; Wrigley 1978. Cf. Ringrose 1990.

207 Wrigley 1978: 306.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

modern England but hardly in the ancient Mediterranean – Greek and Roman cities could at the very least act as a brake on population growth, thereby delaying the negative impact of declining marginal productivity. From this perspective, it is tempting to attribute the considerable duration of ancient growth phases in part to the concurrent growth of cities and the strong urban focus of Greco-Roman civilization in general. Although the precise determinants of urban excess mortality continue to be debated, comparative evidence from later periods suggests that urban death rates usually exceeded birth rates and that cities were sustained by immigration from the countryside.[208] The pull of the city of Rome is merely the most conspicuous example: just as early modern London absorbed most natural growth in England and in the process reconfigured the economic system of its hinterland, late Republican and early imperial Rome has been said to have had a comparable impact on the social and economic structure of Italy.[209] In conjunction with the more than 400 other cities of Italy, the growth of the capital may arguably account for the apparent demographic stagnation of the Roman citizenry in the late Republic, even if slaves from overseas made up a substantial proportion of all urban residents.[210] In Egypt, another heavily urbanized region, urban excess mortality was equally likely to have constrained overall population growth: census data from some of its largest cities point to much higher attrition rates than in the surrounding villages.[211] In less urbanized regions, the pull of urban "population sinks" would have been correspondingly weaker but hardly non-existent. In the broadest terms, we may conjecture that in the Roman empire, urban excess mortality could easily have absorbed rural natural growth of the order of 0.1 percent per year, and possibly more.[212] By implication, over the course of about a millennium, the creation and maintenance of some 2,000 cities would consequently have entailed the movement of perhaps up to 40 million people from the countryside to urban environments.[213] Thus, from a demographic perspective, ancient cities, regardless of the nature of their economic foundations, were bound to make a contribution to economic

[208] de Vries 1984: 175–98; Scheidel 2001b: 28 n. 106; Woods 2003.

[209] Wrigley 1987: 133–56 (and cf. Rozman 1974 on Edo); Morley 1996. For health hazards in the capital, see Scobie 1986. Although Rome's infrastructural provisions for water supply and waste disposal may have alleviated morbidity, the exceptional severity of the local endemic disease environment (dominated by malaria) was bound to offset any such benefits: see Scheidel 2003a, *contra* Lo Cascio 2001b, 2001c: 187–92, 2006a. Contrast Hanley 1987, on Edo.

[210] Jongman 1990, 2003a; Scheidel 2004c: 14–19.

[211] Scheidel 2001a: 142–62.

[212] Annual rates of natural decrease of 1 percent for one-fifth of the gross urban population (residing in the five largest cities of the empire) and 0.5 percent for the remainder could be compensated for by the annual transfer of 0.1 percent of the rural population to the cities. Growing cities would require higher inputs.

[213] If the aggregate urban population increased by 9 million between 800 BC and AD 200 and annual natural decrease averaged 0.6 percent, 36 million would have been required to establish and maintain these cities. At half that attrition rate, the total is 22.5 million.

development by alleviating population pressure and slowing down slides in marginal productivity.

In consequence, pre-industrial cities are best viewed both as parasitic and as a stimulus; these functions cannot be separated. The extent to which urbanization stimulated growth is a question of degree, and critically depends on how much it led to increases in functional specialization.[214] By the time of the Roman empire, even small towns displayed a striking degree of differentiation in crafts and trades.[215] At the same time, the nature of the division of non-agrarian labor between city and countryside remains relatively obscure. We may assume that much as in the late mediaeval and early modern periods, most rural demand was met locally but supplemented by city–country trade, often mediated by periodic fairs. It has also been argued that ancient economies may have experienced less division of labor between cities and regions than mediaeval western Europe, and less inter-regional specialization in general.[216] These issues will be addressed in the discussion of particular regions and periods.

IX CONCLUSIONS

Throughout the Greco-Roman period, structural continuity on some levels coincided with significant change on others. Despite a potentially considerable degree of ecologically determined but ultimately multi-directional variation over space and time, the basic patterns of mortality and fertility are unlikely to have changed much – or at all – in the long term. At the same time, the distribution of population within the Mediterranean and its hinterlands underwent a lasting transformation. With the extension of short-fallow farming and urbanism to the western reaches of Europe and to north Africa, the demographic center of gravity gradually shifted from the eastern to the western half of the Mediterranean region. Hellenistic and Roman imperialism gave rise to the first mega-cities of the region and created a far-flung network of cities in the south-western half of Europe that, a millennium later, provided the template for urban revival and further expansion in the High Middle Ages. In the absence of major transitions in productive technology, much of the discernible increase in population must have been the result of extensive growth, facilitated by climatic change and the lateral dissemination and adaptation of crops, techniques and institutions. The true extent of intensive growth is largely obscured by these

[214] E.g., Wrigley 1978: 298–304; Pleket 1990: 79–86.

[215] Patlagean 1977: 156–81 notes that 110 different trades were recorded in the small town of Korykos in Cilicia. See also below, Chapter 25. Treggiari 1980: 56 counts 225 different jobs for urban laborers in the western Roman empire.

[216] De Ligt 1990, 1991; and see 1993a for fairs. Cf. Jongman 2000b on textiles. On Roman Egypt, see Alston 2002: 334–42.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press

developments: fluctuating over time, it may have accounted for the more conspicuous phases of socioeconomic development of the period – above all in archaic and classical Greece and in late Republican Italy – but was bound to progress very slowly in the long term. From a global perspective, the demographic and economic developments of the Greco-Roman period were part of a wider upward trajectory of undulating growth and contraction that extended into the mediaeval and early modern periods of European history.

https://doi.org/10.1017/CHOL9780521780537.004 Published online by Cambridge University Press