An expansion of the demographic transition model: the dynamic link between agricultural productivity and population

Russell Hopfenberg

Psychiatry and Behavioral Sciences Department, Duke University

(Received 22 March 2014; final version received 4 August 2014)

The classic demographic transition model illustrates the pattern of birth and death rates over time, shifting from high and equivalent to low and equivalent, with population increasing sharply during this transition as a society industrialises. However, the model has a limited temporal frame and cultural scope. It also overlooks that human population trends follow agricultural productivity. Because food is an essential carrying capacity variable and a fundamental economic driver, as food availability is increased the population increases leading to severe biodiversity loss. The current analysis expands the classic model, taking into account all of human history, and highlighting the basic carrying capacity foundations of fertility changes. This comprehensive model shows birth and death rates in Stage A as low and equivalent before the advent of the agricultural revolution. Stage A is followed by Stages B and C, in which the increasing birth rate precedes the increasing death rate, causing a rise in population. The stages then progress as in the classic demographic transition model.

Keywords: agriculture; fertility; food availability; food production; mortality; population

Introduction

Arguably one of the greatest threats to biodiversity is the accelerating environmental degradation due in large part to the near exponential growth of the human population. As population size and consumption levels rise, basic natural resources are depleted. Habitat destruction through agricultural and urban expansion continues to cause ecosystem simplification through species extinction. It is well understood that escalating human population is fueling the acceleration of all environmental problems (Pimentel et al., 1998; Hopfenberg and Pimentel, 2001). A widely held point of view is that transitioning the world through the demographic transition will lead to population stabilisation (Wilson, 2002). The demographic transition model (DTM) shows the pattern of changes in rates of fertility and mortality, illustrating the proximate causes of population growth in an industrialised society. Classic demographic transition (DT) theory predicts human population stabilisation, after modernisation induces the reduction in both mortality and fertility (Notestein, 1945). The DTM is a generally accepted illustration of the birth and death rate changes that occur as a society shifts its economy from agrarian to industrial. This model continues to be conceptualised as the necessary process for population stabilisation (Wilson, 2002). Thus the DTM places population dynamics in the context of a country’s or a society’s economic development.

Steiner (2006) pointed out that economic development goes hand in hand with, and can increase only under, conditions of intensive agricultural expansion. He lamented that, for example, ‘In much of sub-Saharan Africa and large parts of Asia – according to estimates compiled by the Millennium Ecosystem Assessment (MA) – almost no highly productive land is left’. He further stated that ‘Assisting poorer countries to intensify their agriculture may seem the most obvious and sensible solution...’. Agricultural development is clearly essential to economic growth, leading to the perception of expanding opportunity (Båge, 2006).

The DTM, as currently conceived, addresses fertility and mortality as a society industrialises. However, it does not take into account the effects of agricultural intensification, which is the foundation of expanding development and industrialisation. The present analysis describes the agricultural context in which the DT occurs and expands the DTM to include a greater historical time frame. Using this comprehensive model, an alternative understanding of sustainable human population dynamic processes is then indicated.

The current perspective

The DTM describes population change over time. It is based on an analysis begun in 1929 by the American demographer Warren Thompson. He observed changes, or transitions, in birth and death rates in industrialised societies over the previous two hundred years. The DTM illustrates the stages of
transformation of countries from having high birth and death rates to having low ones as they become industrialised (see Figure 1).

Countries characterised as pre-industrial are considered to be in Stage 1 of the DT. Both birth and death rates are high and generally in balance. The high birth rate is offset by a high death rate and this means that overall population growth from the current level is stable, and low or near zero. The rates ebb and flow based upon circumstances such as drought or disease, and these fluctuations are represented by the waves in Stage 1 of the model.

The identifying feature of Stage 2 is a decrease in the death rate, though the birth rate remains high. For example, by the mid-1700s, the death rate in Western European countries began to decrease due, in large part, to improvements in sanitation and medicine, though the birth rate remained high. This dropping death rate, coupled with the high, stable birth rate of Stage 2, contributed to skyrocketing population growth, noted in Figure 1 as ‘natural increase’. In Stage 3, the death rate continues to decrease and the birth rate also decreases. One reason for the trend of a declining birth rate is the economic circumstances of family life. Over time, as a society becomes a more industrial one, children become an added expense and are less able to contribute to the wealth of a family. In other words, in Stage 1, perceptions of increasing resource availability sustains high fertility rates. Through Stage 2, industrialisation leads to improved sanitation and health care. As the society continues to develop economically, it transitions to Stage 3, and there is a trend of societal expectations for higher per capita resource consumption. These expectations impinge on fertility as there is a direct trade-off between number of children and resource allocation per person (Abernethy 1997; Moses and Brown 2003). Because of the trade-off between fertility and per capta resource consumption, the birth rate of Stage 3 was reduced through the twentieth century in developed countries, though unexpected periodic rapid expansions of wealth still led to ‘baby booms’ as experienced in post-World War II United States and Europe. Thus, populations still grew significantly but the growth rate began to slow, leading to Stage 4. In Stage 4, birth and death rates are low and, therefore, total population size is stable. However, because of the population growth through the previous stages, this now stable population is highly elevated. It is the current belief that the world population will pass through a demographic transition and then stabilise. In 1800, the world population reached one billion. In 1950, the number was 2.5 billion. The world population more than doubled to 6.5 billion in 2005. Many demographers expect this trend to continue then peak at about 10 billion around 2070 (Bongaarts, 2009).

In summary, the DTM is descriptive of the transformation within countries and regions from having high birth and death rates to having low birth and death rates as the society industrialises. Although the model is a descriptive one, it is many times seen as explaining population growth as a particular country or region moves from a pre-industrial economy to an industrial one. As stated by the Population Reference Bureau (PRB, 2010), the DTM shows ‘typical changes in the birth rate and death rate that happen as a country industrialises’. This shift has occurred throughout Europe, North America, and a number of other areas in the nineteenth and early twentieth centuries, and started in many developing countries in the middle of the twentieth century.

Left unaccounted in the DTM are the carrying capacity underpinnings of human population growth. In the DTM, population growth is divided into birth and death rates. As birth and death rates are the defining elements of population size at any given point in time, they are proximate causes and do not ultimately or independently account for population size. In other words, birth and death rates are synonymous with population growth rate.

To reach a scientific understanding of population changes through the stages of the DT, the independent variables which precipitate changes in the birth and death rates must be understood. The proximate causes speak to the ‘how questions’, i.e. how does the population grow, stabilise or decline? This question can be answered by looking at changes in birth and death rates. The ultimate causes of population changes speak to the ‘why questions’, i.e. why do birth and death rates change over the DT stages?
Human population dynamics: the impact of carrying capacity

As population dynamics are essentially a biological phenomenon, the ultimate causes of population changes lie in the ecosystem’s life-sustaining variables. These include, for example, food, air, water, space, interspecies competition, predation, disease, and are together referred to as the carrying capacity. Human carrying capacity has been typically understood as the ‘number of us, living in a given manner, which a given environment can sustain indefinitely’ (Catton 1982). Catton (1982) refers to the number of people living in a given manner as the ‘environmental load’. As the environmental load changes, fertility adjusts. Catton (1982) further stated that when the load is less than the carrying capacity, ‘there is room for an expansion of numbers, for an enhancement of living standards, or both’.

There is also the possibility of exceeding the local carrying capacity. When this occurs, the current solution is to ‘prop up’ the population with resources from elsewhere. Big cities are a clear example as food has to be imported. However, this ‘propping up’ of a population is not sustainable in the long run. Deforestation in the service of establishing cropland has enormous detrimental impact on biodiversity which undermines the ecological foundation of human life. The aforementioned cities, as well as surrounding housing, roads, etc., eradicate billion of tons of flora in addition to countless eliminated and displaced fauna (Pimm, 2001). These factors cause many to see the human population as being in overshoot, a circumstance in which the carrying capacity is at risk of precipitously diminishing.

The fact that the carrying capacity defines a population size’s upper limit seems to have eclipsed the other important reality of population dynamics: that the carrying capacity of any species, including the human, is, in effect, an ecological magnet that draws population numbers to it (Hopfenberg, 2003). In fact, all logistic population growth models clearly indicate that population growth proceeds as a function of carrying capacity. The one carrying capacity variable that has been drastically manipulated for thousands of years is food production. The prodigious increase in food production has its roots in the beginning of the agricultural revolution 10,000 years ago. This has led to near exponential human population growth, in keeping with logistic equation models. As Cohen (1995a) stated, ‘The ability to produce food allowed human numbers to increase greatly and made it possible, eventually, for civilisations to arise’.

Widening the scope of the demographic transition model

As a society industrialises, agricultural production significantly intensifies. This intensification produces massive economic benefits ‘as documented in hundreds of studies’ (Reifschneider, 2006). Though the resulting food supply boon fuels population growth, the powerfully adhered to cultural perspective is that food production must be increased to feed the growing population (Hopfenberg, 2009; Lee and Zhang, 2010). The result has been a vicious cycle of escalating food production and population growth. Also, because agriculture is a cornerstone of an industrialised society’s economy, food production increases are
intertwined with industrialisation itself. As Wilson (2002) stated: ‘What happened was, with the green revolution, we started a large part of the world through the demographic transition’.

In order to embed the population trend outlined by the DTM in the greater context of the ecological laws governing population dynamics, a clear view of the time period before Stage 1 must be elucidated. Stage 1 encompasses a society’s pre-industrial period. In Stage 1, the perception of the society is that many children are needed for farming and are seen as economic assets. That Stage 1 occurs in an agrarian society indicates that farming practices have already been well established before Stage 1, and that the food supply is already relatively high. If this were not the case, ‘many children’ could not be produced as people cannot be created without the biological building blocks, i.e. food, required to make them.

It is clear that a reorganisation of human demographic systems similar to the classic DTM occurred prehistorically following the conversion of societies to a primary reliance on agriculture (Gage and DeWitte, 2009). As humans existed worldwide for millennia prior to societies’ almost total reliance on agriculture, an expansion of the DTM would necessarily encompass the changes in fertility and mortality rates that accompany the transition of a population from pre-Stage 1 to Stage 1 of the DTM.

The time period before Stage 1 encompasses a population’s transition to a primary reliance on agriculture vs., for example, hunting and gathering. In general, evidence indicates that among hunter-gatherers, mortality and fertility were both relatively low, leading to stable population size with no or slight increases over long periods of time (Gage and DeWitte, 2009). This is due to the fact that their population had reached the limits of the long-term carrying capacity (Lee, 1969; Lee and DeVore, 1976; Pimentel and Pimentel, 1996; Quinn, 1996; Hopfenberg and Pimentel, 2001). In a comprehensive DTM (C-DTM), a pre-agricultural stage must be included and is referred to as Stage A in Figure 2.

Empirical archeological evidence and secondarily inferred genetic evidence point directly to population expansions dated to the transition to a primary reliance on agriculture. Similarly, recent analyses have shown human population growth to be a direct result of agricultural increases (Hopfenberg and Pimentel, 2001; Hopfenberg, 2003). Diamond (1997) noted that ‘the first connection is the most direct one: availability of more consumable calories means more people’. Farb (1978) stated that ‘intensification of production to feed an increased population leads to a still greater increase in population’.

![Figure 2. The comprehensive demographic transition model (C-DTM).](image)

Again, Stage A shows birth and death rates as being both low and stable. Also, a society in Stage A has low population numbers, i.e., numbers that are sustainable indefinitely at the local carrying capacity. In general, when people changed to intensive agricultural production, either through adopting this practice or being overrun by peoples who already practised it, fertility increased rapidly as indicated in Stage B of Figure 2.
Early on in the process of converting to a primarily agrarian lifestyle, declines in health and increases in mortality follow the initial increases in fertility (Gage and DeWitte, 2009). With acute birthrate increases and the resultant large population, the pre-industrial society moves from Stage B to Stage C (see Figure 2). During this stage, the death rate increases as environmental, medical and sanitation problems are exacerbated. In fact, there is sufficient evidence to conclude that the adoption of agriculture was generally detrimental to human health and quality of life (Cohen and Armelagos, 1984; Gage and DeWitte, 2009).

Another known reason for the high death rate of stage C in Figure 2 is famine. Widespread famine, almost by definition, occurs primarily under conditions of intensive agricultural production. For example, all famines over the past 1000 years noted in the Encyclopaedia Britannica (2010) cite populations that rely exclusively on intensive agricultural production. Thus, increases in agricultural production precipitate a high birth rate in Stages B, C, 1, 2 and 3 of the C-DTM. The high birth rate is accounted for by the increase in agricultural production. The high death rate is accounted for by agriculture-related famine and health decline as well as high population-related disease.

Again, once health care and sanitation improve, the population enters Stage 2. In Stage 2, the birth rate remains elevated and the death rate decreases so the population naturally increases. As average resource consumption per individual increases, the population enters Stage 3. In Stage 3 of the C-DTM, the birth rate begins to decline. As this trend continues the population, theoretically, moves to Stage 4. In Stage 4, birth and death rates are low but, at this point, population size and resource consumption are highly elevated.

**Verifying the model**

According to the DTM, in Stages 2 and 3, birth rates are elevated relative to death rates and, as the model indicates, the population will grow dramatically. In Stages 1 and 4, both birth and death rates are nearly equivalent and, thus population size is stable. In other words, there is one phase where the population grows rapidly and two phases where population growth is at or near zero. The theoretical growth rate curve, based on the DTM, is sigmoidal in shape (see Figure 3) and follows most mathematical models of simple population growth. Studies show that increases in carrying capacity, primarily food supply, cause a near exponential growth in population which tapers off as the population reaches the limit of the carrying capacity (Hopfenberg and Pimentel, 2001; Hopfenberg, 2003). Thus, the sigmoidal shape of the theoretical population growth curve meshes with the DTM, as illustrated in Figure 3.

![Figure 3. The DTM with superimposed sigmoidal population](image-url)
However, the C-DTM indicates that there are, in fact, two phases where the population grows rapidly and three phases where population numbers are stable (see Figure 2). Again, a society moves through these stages as it transitions from non-agrarian to agrarian to industrial. In Stage A, the population size is low and sustainable, and birth and death rates are low and roughly equivalent. As the society becomes agrarian and moves through Stages B and C, the population numbers naturally increase as there is an increase in the birth rate followed by an increase in the death rate. The population then stabilizes as birth and death rates become equivalent. The overall population growth curve is sigmoidal, similar to the growth curve derived from the DTM. However, after this phase of stable population, the society then moves through Stages 1 through 4 of the C-DTM, and the population growth curve continues on in a second sigmoidal pattern. Thus the population curve derived from the C-DTM has two pulses, essentially combining two sigmoidal population functions (see Figure 4).

![Figure 4. The C-DTM with two superimposed sigmoidal population growth curves.](image)

It is important to remember that population is a function of carrying capacity (Cohen, 1995b). Hopfenberg (2003) demonstrated that actual food production numbers, used in the logistic equation as a dynamic carrying capacity, accurately predicts actual population numbers. That analysis provides additional evidence that human food availability is the critical salient variable in assessing human carrying capacity, and that human population numbers vary as a function of human carrying capacity.

In pre-industrial societies, empirical studies show that in some periods the population expanded rapidly while in others population growth became stagnated. At times a relative population collapse followed (Turchin, 2009). Yet all of these oscillations are directly connected to food, i.e., carrying capacity (Turchin, 2003). When these fluctuations are viewed over an extended time period, the overall effect is a general logistic population function, as shown by the curve fitted to the data points in Figure 5.

In 1999, Meyer and Ausubel demonstrated that the effect of a dynamic carrying capacity over the time period spanning a society’s change from agrarian to industrial, leads to a bi-sigmoidal population growth curve, i.e. a sigmoidal population growth curve with two pulses (see Figure 5). They mathematically derived the carrying capacity values using a logistically increasing function. The predicted numbers were plotted and concurred with the graph of actual population numbers.

Meyer and Ausubel (1999) tested their model in two cases, extending the theoretical dynamic carrying capacity to include several centuries of pre-industrial society in both England and Japan. The resulting two-pulsed population curve, derived from the dynamic carrying capacity, fit with the available census data in both cases, thus providing strong support for their model. The two-pulsed population curve, which fit with the census data in Meyer and Ausubel’s (1999) two test cases, directly corresponds with the course of population growth predicted from the C-DTM pictured in Figure 4.
Flexibility of the model

In the C-DTM, as with the DTM, the time periods for each stage in the figures are uniform for the sake of illustrating the DT phenomenon. However, the actual duration of the stages in both models are highly varied. For example, in the C-DTM, the time period for Stage A might represent hundreds of centuries whereas the time period for Stage 2 might be measured in decades (PRB, 2010).

In plotting actual population data during a society’s transition from agrarian to industrial, Meyer and Ausubel (1999) showed that the second pulse of the actual population curve is typically highly elevated relative to the first, indicating a short transition time between Stages 1 and 4 relative to the time period before Stage 1. The flexibility and usefulness of the C-DTM is shown by altering the model to depict a more realistic representation of relative timeframes. The condensed timeframe of Stages 1–4 relative to Stages A–C is illustrated in Figure 6. By highlighting a more realistic representation of relative timeframes, the second pulse of the bi-sigmoidal growth curve takes on the elevated characteristic seen in actual population data of Figure 5.

Figure 5. Population of England 1541–1975 (left) and Japan 1100–1992 (right) indicated by the P(t) lower lines, showing bi-sigmoidal

Figure 6. Condensation of the DTM timeframe within the C-DTM.
When thinking about the slope of the population curve it is also important to take into account the relativity of rate calculations. For example, if at a time in Stage 3 a population consists of 2 million people and the growth rate is 3% per year, the population the following year will be 2.06 million, an additional 60 thousand people. At a later point in Stage 3, through the ‘natural increase’ because of elevated birth rate compared to death rate, let’s say that the population has increased to 6 million people. According to the model, the growth rate will have dropped. A growth rate decrease of one-third would bring the rate to 2% per year. With a 2% growth rate applied to 6 million people however, the population the following year would be 6.12 million – an additional 120 thousand people. That’s twice as many added people than at the 3% rate, even though the growth rate is lower by one-third. These rate calculations also apply to the death rate, though in an inverted manner.

In addition to attending to the relativity of timeframe representations and rate calculations, the most salient independent causative variable, agricultural production, is highly elevated in industrialised society relative to production levels of agrarian society. These agricultural production levels are, of course, even more extremely raised compared to pre- or non-agrarian societies. The effect of the elevated rate of agricultural production is the maintenance of a high birth rate, which can be accounted for in the model by a temporal extension of Stage 3.

Raising the availability of human food through conversion to more intensive agricultural practices ushered societies through Stages B, C and 1–3 of the C-DTM (Hern, 1990; Wilson, 2002; Gage and DeWitte, 2009). Agricultural production was, and continues to be, the driver of societies through the DT and a basis for industrialisation. The level of agricultural production has been shown to act as an ecological magnet, drawing population numbers up to it through the maintenance of a high birth rate relative to the death rate (Hopfenberg, 2003). Population is maintained in dynamic equilibrium with the level of food availability or carrying capacity.

Conclusion
The comprehensive DTM (C-DTM) is an extension of the classic DTM and uncovers the carrying capacity variable of food production that is hidden by the DTM’s limited historical scope and focus on industrialisation per se. The C-DTM is a model that encompasses all of human history and reveals the impact of societies’ conversion to a primary reliance on agriculture and subsequent industrialisation and agricultural escalation. The conversion of societies to intensive agricultural production has had a profound impact on global human population numbers as well as on resource consumption, species extinction and climate change (Pimm et al., 1995; Pimentel et al., 1998; Hopfenberg and Pimentel, 2001; Hopfenberg, 2003, 2009). With an understanding of the impact of the conversion to agriculture on human population growth, the C-DTM makes clear that agricultural intensification has been part and parcel of societies’ industrialisation. The Green Revolution is the quintessential exemplification of this process (Wilson, 2002).

Of course, the overwhelming perspective of the lay, political and even scientific community is that food production must be increased to feed the growing population. As stated in the United Nations Environment Programme (Nellemann et al., 2009): ‘The growth in food demand and need is the result of the combined effects of world population growth to over 9 billion by 2050...’. The report continues with: ‘Each day 200,000 more people are added to the world food demand’. Yet regarding the rest of the biological community, it is accepted without question that the population of every species increases to the level of its food supply (Pimentel, 1966; Hopfenberg and Pimentel, 2001).

Turchin (2009) stated that ‘as soon as population numbers reach the carrying capacity determined by the total number of territories, population growth rate is reduced to zero, without any time lag’. The phrase ‘without any time lag’ could lead to a perception that a species’ population reaching the carrying capacity limit is always fraught with highly aversive consequences, e.g. starvation (Turchin, 2009). However, the overwhelming evidence is that population growth and stabilisation proceeds in accord with the logistic mathematical function. This means that as a population approaches the carrying capacity limit, the growth rate diminishes asymptotically over time. In fact, the logistic function, with a variable carrying capacity can lead to bi-stability (see Figure 5), exponential growth that is seemingly unlimited (Hopfenberg, 2003), and periodic population oscillations (Turchin, 2009). Thus, population dynamics proceed logistically as a function of carrying capacity (Cohen, 1995b; Hopfenberg, 2003).

Along with human population varying as a function of carrying capacity, an additional reality is that a current ‘demand for more food’ cannot be met by increasing food production. An endeavour to increase
food production involves sowing and growing as well as distributing. Current demand cannot wait for these future processes. If people are starving at time A, a decision at time A to increase food production will not help the people at time A but would instead lead to providing an increased carrying capacity at future time B, fueling further population increase. Thus, as global food production has increased, the global human population has increased, and the number of starving and malnourished has similarly increased (Hopfenberg, 2009; Nellemann et al., 2009). Increasing food production clearly does not ameliorate starvation and malnutrition; it actually exacerbates this tremendous and unfortunate difficulty (Hopfenberg and Pimentel 2001).

Studies incorporated in the Millennium Ecosystem Assessment (2005) determined that by the year 2000, nearly 37% of the world’s terrestrial surface had been converted for agricultural use. This trend seems destined to continue as agri-businesses, academic institutions and world governments have a focus on increasing agricultural yields. However, as Farb (1978) noted ‘the population explosion, the shortage of resources, the pollution of the environment, exploitation of one human group by another, famine and war – all have their roots in that great adaptive change from foraging to production’. Farb’s statement makes clear that the ‘adaptive change from foraging to production’ is coming into focus as one that has provided some relatively short-term benefits and many long-term difficulties.

The DTM has been used to describe population trends and link these trends to industrialisation. It has also been a major driver regarding thoughts about societies’ courses of action as well as regional and international policy formation. Because the DTM is limited in scope, it masks the ecological factors that are the true drivers of population growth and size. With the limited view encapsulated in the DTM, the notion that population growth is mainly a function of industrialisation per se, furthers the perspective that we need to help foster a society through the DTM stages, i.e. to increase industrialisation, and especially agricultural expansion, in order to reach a stable population size. However, this line of thinking bypasses the strong evidence linking human population growth with agricultural production. Because intensification of agricultural production is ideologically fused with industrialisation, as industrialisation is promoted, intensification of agricultural production is promoted as well.

The C-DTM makes clear that agricultural increase is the independent variable which drives the population through its stages and maintains the near exponential population growth characteristic of Stage 3. This acute population growth is accompanied by other ecological and resource-consumption problems (Hopfenberg 2009). The model makes clear that, rather than leading to further problems, halting increases in agricultural production will lead societies through Stage 3 and on to Stage 4, with low birth and death rates and population stability. Thus, through a full understanding and appreciation of the effects of continually increasing agricultural production, elucidated through the C-DTM, more viable ways to attend to societal and global human well-being will be achieved.

Acknowledgements
I wish to thank Cole and Emily Hopfenberg for their patience and inspiration, Steven Salmony for his encouragement and Edie Hopfenberg for her tireless support and review of all drafts.

References