If Not Malthusian, Then Why? A Darwinian Explanation of the Malthusian Trap

Lemin Wu*

June 2015†

Abstract: This paper shows that the Malthusian mechanism alone cannot explain the pre-industrial stagnation of living standards. Improvement in luxury technology, if faster than improvement in subsistence technology, would have kept living standards growing. The Malthusian trap is essentially a puzzle of balanced growth between the luxury sector and the subsistence sector. The author argues that balanced growth is caused by group selection in the form of biased migration. It is proven that a tiny bit of bias in migration can suppress a strong growth tendency. The theory re-explains the Malthusian trap and the prosperity of ancient market economies such as Rome and Sung. It also suggests a new set of factors triggering modern economic growth.

Keywords: Malthusian Trap, Industrial Revolution, group selection, very long run growth, source-sink migration

JEL Classification: N3, O41, B12

*School of Economics, Peking University. I would like to thank David K. Levine, Nicholas Papageorge, and Rohan Dutta, who co-authored a related paper with me, and George Akerlof, Gerard Roland, Joel Mokyr, Ronald Lee, Bradford DeLong, Pranab Bardhan, Oded Galor, Martha Olney, Gregory Clark, Peter Lindert, Jan De Vries, Barry Eichengreen, Yew-Kwang Ng, Yingyi Qian, Yang Xie, Yuriy Gorodnichenko, Jian-Cheng Lai, Tuan-Hwee Sng and the attendees of seminars at UC Berkeley, UC Davis, Peking University, Tsinghua University, Fudan University, Shanghai University of Finance and Economics, National University of Singapore, Nanyang Technological University and Singapore Management University for thoughtful suggestions and essential encouragement. I especially thank Gregory Clark for his insightful and provocative introduction to Malthusian theory in *A Farewell to Alms*, which inspired this work. The usual disclaimers apply. To contact the author, please email: leminwu@outlook.com

†First draft: November 2013. An earlier version of this paper, "Does Malthus Really Explain the Malthusian Trap?" is the first chapter of the author's doctoral dissertation at UC Berkeley.

1 Introduction

> All other arguments are of slight and subordinate consideration in comparison of this. I see no way by which man can escape from the weight of this law, which pervades all animated nature.
>
> Thomas Malthus (1809, chap.1)

> But in economics, the admission that mankind need not live at the margin of subsistence [...] meant that, the very long run limit of wages was not physiological subsistence, it was *psychological subsistence*—a much more complicated and difficult matter to formulate exactly.
>
> Lionel Robbins (1998, pg.174)

Life was miserable and stagnant for most who lived before the Industrial Revolution. "The average person in the world of 1800 was no better off than the average person of 100,000 BC" (Clark, 2008). According to Malthus, poverty lingered because economic progress only led to faster population growth. A larger population depresses average income and brings society back to persistent poverty. For two hundred years, Malthus's wisdom has been held as one of the most infallible doctrines in economics.

However, this paper shows that Malthus's mechanism alone cannot explain the Malthusian trap. Technological improvement in the production of luxury goods, if faster than technological improvement in the production of subsistence goods, would have kept living standards growing. To suppress the upward trend of living standards growth, it takes the Darwinian force of selection in a social context: cultures and technologies spread faster if they favor group prosperity at the expense of individual welfare. The trap is no less Darwinian than Malthusian.

Explaining the Malthusian trap is at the core of Malthusian theory. Despite empirical weakness of the relationship between average income and population growth, most economists hold a Malthusian view of history because Malthus explains the persistent poverty of the pre-industrial world, and hardly any competing theory has been available. However, Malthus's explanation for the Malthusian trap bears further investigation. The prediction is right—the trap exists—but the mechanism is open to challenge.

Malthusian theory is often inconsistent with data, and, in addition, relies on a crucial assumption that is disputed in this paper. Implicitly, Malthusian theory assumes away the conflict between individual welfare and group fitness. The traditional theory ignores the fact that, in the real world, what promotes the individual's welfare does not always expand the

group's population. Diamonds, spas, circuses, monuments—be they for sexual attraction, comfort, entertainment or vanity—are immune from the Malthusian force. Increases in these types of consumption can raise living standards without limits.

Take flowers and bread as a metaphor. Flowers attract the mates; bread feeds the mouths. Population increases with the bread but not with the flowers. Choice-theoretically, when there are more flowers, people are simply better off. With the average bread consumption anchored by the Malthusian force, the equilibrium living standards are fully determined by how many flowers the average person has, which in turn depends on the ratio of flowers to bread in an economy.

I call bread "subsistence" and flowers "luxury". Both contain hedonistic value, but one dollar worth of subsistence has a larger demographic effect than one dollar worth of luxury (the division is relative: beef is a luxury relative to potatoes, yet is a subsistence good compared to diamonds).

The two-sector model allows culture and technology to change equilibrium living standards while the Malthusian constraint is still binding. What matters is not the size of the economy but the structure of production (say, flowers versus bread); not the aggregate demand but the relative preference, i.e., how much value people place on one thing—flowers—as compared with another—bread. By adjusting the ratio of luxury to subsistence, variation in production structure and social preference can now explain a large portion of fluctuations in living standards, which are misattributed to changes in fertility culture and disease environment when the two-sector model is unavailable.

It follows that the Romans were rich not because technological progress temporarily exceeded population growth—as the traditional Malthusian theory would claim—but because Rome had a business-friendly legal system and an active market economy. Well-functioning courts and marketplaces boost industry more than they boost agriculture; this biases the production structure toward luxury, and thereby raises the average living standards of the whole society.

Conversely, after the Agricultural Revolution, the hunters-and-gatherers-turned-peasants failed to achieve the level of leisure and nutrition their ancestors once enjoyed because agriculture biased the production structure toward subsistence. The tragedy recurred when potatoes dominated the Irish diet in the late 18^{th} century.

Yet the two-sector model raises a serious puzzle. If luxury productivity had been growing slightly faster than subsistence productivity, living standards would have been rising steadily, but this never happened until the modern era. The absence of a trend in the average living standards implies that the luxury sector and the subsistence sector somehow grew at the same speed. The balance of growth lies at the heart of the question of why the Malthusian

trap ever existed, but Malthus did not have the opportunity to take this puzzle into account.

I show that balanced growth is caused by Darwinian selection of cultures and technologies at the group level. Here is how it works: because living standards increase with the ratio of luxury to subsistence, migration, not necessarily peaceful, is usually from places relatively rich in subsistence to places relatively rich in luxury.[1] Subsistence cultures and technologies are thus more likely than luxury cultures and technologies to be carried around by migrants and conquerors—ideas are spread in a selective way. Even if luxury technologies intrinsically grow faster, the advantage of subsistence technologies in spread might offset the advantage of luxury technologies in growth, thus keeping the two sectors in balance. Simulations confirm that living standards would grow steadily if there were only the Malthusian mechanism but not technological selection, whereas a tiny bit of selection is enough to suppress a strong tendency of growth. Selection, no less than the Malthusian mechanism, is crucial to the existence of the Malthusian trap.

The Darwinian theory of the Malthusian trap provides a new perspective on the onset of modern economic growth. Instead of treating growth as the result of "an unprecedented beam" that "suddenly lit up the British scene" (in Eric Jone's words), this paper sees growth as a omnipresent and recurring phenomenon that was usually — though not always — suppressed. In the same spirit as Jones (1988), this paper asks "what had been suppressing the urge towards betterment everywhere and, second, what eventually came along to lift the heavy weight." However, the answer I give is different from those discussed by Eric Jones, who emphasized marriage patterns (Jones, 2003) and the rulers' rent-seeking (Jones, 1988). Rather, my view of historical growth is more in line with Mokyr (2005)'s observation that

> "Before 1750, growth had been limited to relatively small areas or limited sectors, often a successful city state, a capital of a powerful monarchy, or a limited agricultural region. These niches had to spend much of their riches to protect their possessions against greedy neighbors, real-life manifestations of Mancur Olson's 'roving bandits' who often killed entire flocks of golden-egg-laying geese. After the Industrial Revolution, it became a more aggregative phenomenon, with a substantial number of economies becoming members of the much-coveted 'convergence club'. "

Building on the new theory of the Malthusian trap, this paper also sketches possible mechanisms how group selection switched from a force that suppressed growth into a force that favors growth. It is hypothesized that the reversal of selection is crucial to understanding the Industrial Revolution. Further elaboration is left to my other work.

[1]Subsistence, as is defined in the paper, includes military skills, equipments and organizations.

The paper is composed of two parts. The first part is the two-sector Malthusian model, which raises the balanced growth puzzle (section 3). The second part addresses the puzzle with the group selection theory (section 4). To revise Malthusian theory, I need to demonstrate three things. First, I show that the luxury-subsistence division can explain a number of irregularities in historical data (section 3.4). Second, I uncover the source-sink pattern in historical migrations (section 4.2). Third, I simulate and mathematically prove that living standards would grow in a Malthusian economy unless selection is introduced, and that a tiny bit of selection can suppress a strong tendency of growth (section 4.4 and 4.5).

In section 5, I consider the implications of the theory for the Agricultural Revolution (section 5.1), the rise and fall of ancient market economies (section 5.2), the long-term impacts of wars and migrations (section 5.3), and the Industrial Revolution (section 5.4). Section 6 concludes and discusses what the research means to Economic History as a discipline.

2 Literature review

Does the hypothesized existence of the Malthusian trap fit historical facts? Figure 1 shows Maddison (2003)'s estimates of the world's GDP per capita for the last two millennia. Using Maddison's data, Ashraf and Galor (2011) confirm that, by year 1500, the level of a country's technology explains the country's population density, but not the income per capita. These estimates, however, might be biased by a Malthusian presumption. Further work in the past few years has largely wiped away the humdrum picture of ancient economic life in the original Maddison series. It has been shown that, during the pre-modern centuries, while Italy and Spain experienced stagnation and even declines (Malanima, 2011; Álvarez-Nogal and De La Escosura, 2013), the per capita GDP of England doubled between 1270 and 1700, and that of the Netherlands almost tripled between 1000 and 1500.

Research has questioned when sustained growth actually started (Hersh and Voth, 2009; Persson, 2010; Fouquet and Broadberry, 2015), but, except Dutta et al. (2014), has taken the Malthusian trap as a given. The take-home message of most of the research is that living standards fluctuation was much larger than previously thought—the findings disturb the Malthusian interpretation of short-run fluctuations but not the Malthusian fact of long-term stagnation. Even Dutta et al. (2014) agree that the European per capita GDP failed to recover from the collapse of the Roman Empire until perhaps the dawn of the Industrial Revolution. Hence, it is fair to say that the existence of the Malthusian trap is still widely held as a fact. This paper does not take a stand on whether the Malthusian trap actually exists or not. I assume the trap and explain why it would ever arise if it did arise.

There are questions about the extent to which Malthusian theory is applicable to ex-

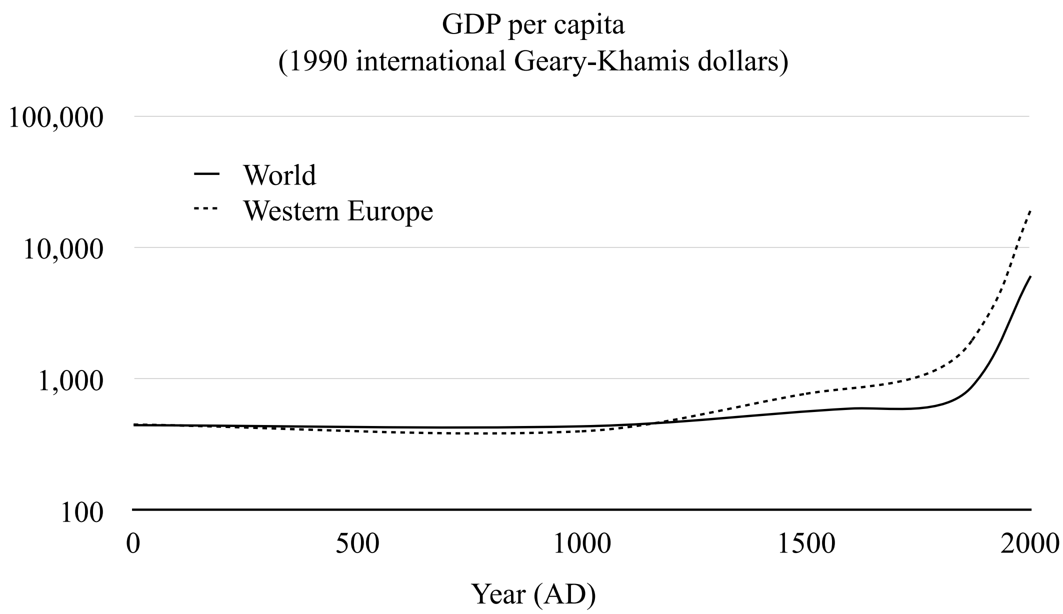

FIGURE 1: World GDP per capita 1-2001 AD. Data source: Maddison (2003)

plaining short-run changes. As a matter of fact, average income and population growth were poorly correlated in the English data. Yet, the Malthusian force is believed to dominate in the long run by its persistence (Lee, 1987, p.452). Allen (2008) worried about the conjecture but mentioned no alternative solution.

Another post-Malthusian development is the study of modern growth. Presuming that stagnation was caused by the Malthusian force, theorists have endogenized the acceleration of growth to reconcile the Malthusian stagnation with modern economic growth (Simon and Steinmann, 1991; Jones, 2001; Hansen and Prescott, 2002; Galor and Weil, 2000; Galor and Moav, 2002). Such research describes how a Malthusian shackle would have been broken, but maintains the assumption that the shackle was truly Malthusian.

This paper has two key concepts: sectoral duality and group selection. Two-sector Malthusian models have appeared in Restuccia, Yang, and Zhu (2008); Voigtländer and Voth (2013), and Yang and Zhu (2013). In these studies, agriculture is different from industry in two aspects. First, agricultural production relies on land more than industrial production. Second, the income elasticity of demand for agricultural products is lower than the income elasticity of demand for industrial products[2]. The population growth rate depends on per capita income instead of food intake. Subsistence, when mentioned in these

[2]To capture the difference in income elasticity of demand between agriculture and manufacturing, Restuccia, Yang, and Zhu (2008) and Voigtländer and Voth (2013) use a Stone-Geary utility function; Yang and Zhu (2013) assumes a fixed level of food per capita.

papers, is a threshold of income that people must meet before consuming goods that are less land-intensive to produce. In contrast, this paper abstracts from these differences to focus solely on the fact that agricultural products and manufactured products have different demographic effects. The meaning of subsistence is thus different. Unlimited growth is possible in my model but not in theirs. As will be discussed later, incorporating the differences in the intensity of land use and the income elasticity of demand would complicate the analysis but actually favor my result.

That said, this paper is not the only study that distinguishes sectors in terms of demographic effects. Rudimentary models with similar assumptions have been built in Davies (1994) and Taylor and Brander (1998). More sophisticated versions appeared in Lipsey, Carlaw, and Bekar (2005, chap.9), Weisdorf (2008), Strulik and Weisdorf (2008) and Vollrath (2011). These theories are useful at comparing living standards across ancient societies. However, except Strulik and Weisdorf (2008), they do not explore the possibility that a sustained directional change in production structure would produce an upward trend in living standards.

Strulik and Weisdorf (2008) is aware of the possibility and provides a endogenous growth theory to explain why the Malthusian trap still arose. The idea is that, in the early phase of economic development, the labor force was concentrated in the agricultural sector. As technologies grew mostly through learning by doing, the relatively larger size of agricultural labor meant a faster growth in agricultural technology than in manufacturing technology. But agricultural expansion is constrained by fixed land endowment, hence the development of agriculture drove more and more labor into the manufacturing sector, which finally overtook agriculture and unleashed a steady growth of per capita income. Strulik and Weisdorf (2008)'s theory nicely captures some crucial aspects of the transition from the Malthusian trap to the modern economic growth — for example, the steady decline in the relative price of agricultural products in Europe on the eve of the Industrial Revolution. However, as section 3.6 will discuss in detail, the endogenous growth theory is not without limitations. For one thing, it cannot explain why the Hittite kingdom, the Roman Empire, and the Sung China, all of which were highly developed in industry and commerce, failed to bring about "modern" economic growth at an earlier time. For another, the division of agriculture and manufacturing is merely a metaphor. Each category is composed of hundreds of, if not thousands of, sectors. The mere fact that agriculture is a larger category than manufacturing does not mean that each sector in agriculture must hire more labor than sectors in manufacturing. I show that, to complete the re-explanation of the Malthusian trap, it takes group selection to complement Strulik and Weisdorf (2008)'s endogenous growth mechanism.

The concept of group selection was once a taboo but has revived in the past few decades.

Academia once resisted the idea because it reminded one of Nazi eugenics. But modern biological research has demonstrated the absurdity of racism based on genetic reasons. Research on group selection is and should be apolitical. Nevertheless, it is worth noting that, unlike Galor and Moav (2002) and Clark (2008), my notion of selection is not genetic but cultural. Another reason why group selection was dismissed as pseudo-science is that researchers once thought group selection to be invariably weak compared to individual-level selection. In social contexts, emphasizing group selection seemed to contradict individual rationality. However, since the 1970s, biologists and economists have found numerous ways for group selection to exert a significant influence despite individual selection. Bowles and Gintis (2002) and Bowles (2006) use it to explain how cooperation gets rooted in human nature. The sociobiologist Wilson (2015) provides a book-length introduction to the modern literature on multilevel selection and how it explains altruism. In this paper, selection operates among Nash equilibria that emerge out of individuals' rational choices. It does not even require people to be altruistic to do things that matter for the group's survival and expansion. Therefore, to the extent that critiques on group selection are valid, this paper addresses those critiques. This paper also contributes to the sociobiological literature of group selection. Multilevel selection theorists have been overwhelmingly concerned with the origin of altruism; hundreds of books and papers have been published on the topic. But this paper is the first to apply the idea to explaining the Malthusian trap, an issue of no less importance.

A parallel study that I participate in, Dutta, Levine, Papageorge, and Wu (2014) (henceforth DLPW), is the closest research to this one. DLPW has a similar two-sector Malthusian model, but there are at least three differences between DLPW and this paper. First, DLPW covers only one dimension of comparative statics; this paper covers three. Second, DLPW takes the inter-sector difference in demographic effects as a presumed fact, while this paper explains how the difference arises from an evolutionary biological perspective. Third, although both DLPW and this paper come to the same conclusions — that either the Malthusian trap is an inexact description of historical facts, or the traditional Malthusian explanation of the Malthusian trap is incomplete — DLPW explores the first possibility, while this paper focuses on the second: re-explaining the Malthusian trap.

Similarly to DLPW, another paper, Levine and Modica (2013), treat the Malthusian trap not as a fact to explain but as an assumption that deserves scrutiny. Instead of dividing goods into sectors, they focus on the allocation of resources between the people and the authority—a special case of luxury and subsistence within my two-sector Malthusian theory. Their equilibrium is the maximization of authority-controlled resources meant for wars between states. It is also a special case of my group selection theory.

3 A two-sector Malthusian theory

In this section, I begin with a two-sector Malthusian model that adds two new comparative statics to the classical theory. The classical Malthusian theory is shown to be a special case that assumes away the conflict between individual welfare and group fitness. Next, I discuss the historical relevance of multi-sectorality, and its implications for long-term growth.

3.1 The division of sectors

Suppose there are H identical people living on an isolated island that has M types of commodities, $j = 1, 2, ..., M$. The representative agent consumes $E \in \mathbb{R}_{+}^{M}$, a bundle she chooses to maximize a utility function that is differentiable and strictly increasing:

$$\max_{E \in C(H)} U(E). \tag{1}$$

The island is a Malthusian economy. Given resources, her choice set, C, shrinks with population: $\forall H_1 < H_2$, $C(H_1) \supset C(H_2)$. The population growth rate depends on the average consumption E:

$$\frac{\dot{H}}{H} = n(E). \tag{2}$$

Assume that $n(E)$ is continuous, differentiable and strictly increasing, and that there exists a set $\mathbb{S}$ on which population stays constant. Call $\mathbb{S}$ the constant population set. Any isolated economy that finds itself on the constant population set is at a Malthusian equilibrium.

If $U(\cdot)$ is not a transformation of $n(\cdot)$, there must exist some bundle of consumption E, at which one commodity is more luxurious than another, i.e., $\exists j_1, j_2 \in \{1, 2, ..., M\}$ such that

$$\frac{\frac{\partial U(E)}{\partial E_{j_1}}}{\frac{\partial n(E)}{\partial E_{j_1}}} > \frac{\frac{\partial U(E)}{\partial E_{j_2}}}{\frac{\partial n(E)}{\partial E_{j_2}}} \tag{3}$$

Compared with j_2, commodity j_1 marginally contributes more to individual utility than to population growth. This makes j_1 a luxury relative to j_2. We can define the "luxuriousness" of each commodity in the following way: $\forall E \in \mathbb{R}_{+}^{M}$, $\forall j \in \{1, 2, ..., M\}$, commodity j's luxuriousness at E is

$$\frac{\frac{\partial U(E)}{\partial E_{j}}}{\frac{\partial n(E)}{\partial E_{j}}} \tag{4}$$

Order all commodities by their luxuriousness, and we have a spectrum from the most luxurious commodity to the most "subsistential" commodity. In fact, we can always distinguish luxury goods from subsistence goods as long as $U(\cdot)$ is not a transformation of $n(\cdot)$.

3.2 Production structure and social preference

Here I provide a graphical representation of the two-sector Malthusian model. To start with, draw the representative agent's consumption space (Figure 2A). Besides the familiar indifference curve and production possibility frontier, the diagram has a "constant population curve". If the agent's consumption bundle is right on the curve, population stays constant. If the bundle lies to the left—consumption is less than reproduction requires—population decreases. If the bundle lies to the right, population increases.

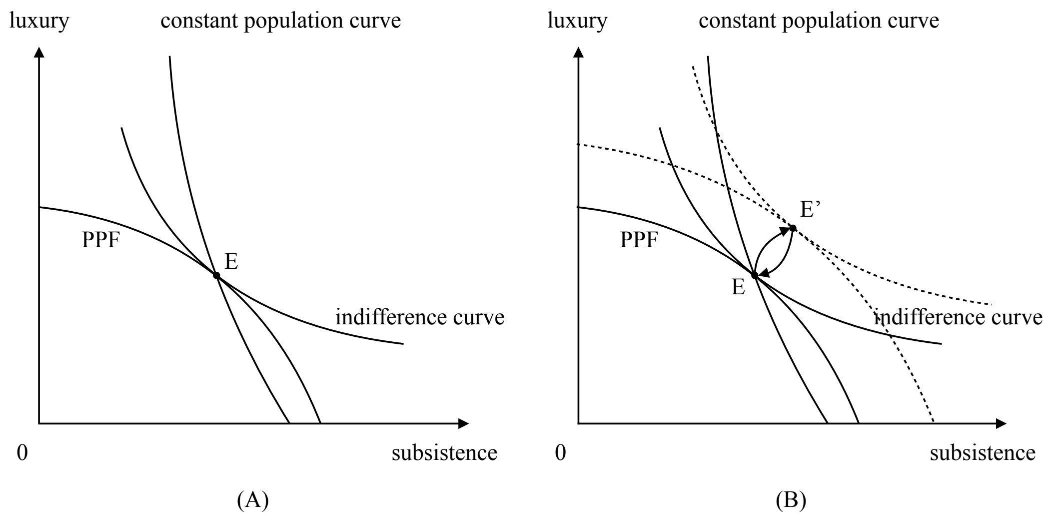

FIGURE 2: The two-sector Malthusian equilibrium

Population changes shift the production possibility frontier. When population declines, the frontier expands: each person is endowed with a larger choice set. When population rises, the frontier contracts—the economy has diminishing returns to labor. The returns to labor diminish because land is crucial to production and its supply is inelastic.

Assume that the expansion and contraction of the production possibility frontier are shape preserving, that is, the shape of the production possibility frontier is independent of the size of the population. It would be interesting to allow luxury to be more labor-intensive than subsistence, but such a relaxation of assumption would only complicate the model in ways that are favourable to my hypothesis—the equilibrium living standards would then become even more responsive to technological changes. Hence I stick to the shape-preserving assumption[3]. The idea is that I do not equate subsistence with food. Luxury and subsistence are more fundamental a pair of concepts than are food and non-food, or agriculture and industry.

[3]One may consult Dutta et al. (2014) for the case where the assumption is dropped.

The constant population curve crosses the indifference curve from above because, by definition, subsistence goods are more important than luxury goods to population growth.[4] The curve does not have to be vertical though. Beef is more luxurious than potatoes, but beef contains calories too.

The Malthusian equilibrium must lie on the constant population curve. As Figure 2B shows, if the economy expands, the temporary affluence will raise the density of population. The production possibility frontier will then contract until the economy returns to the constant population curve.

Unlike the classical Malthusian theory, the two-sector framework allows technological shocks to change equilibrium living standards. A positive shock in luxury technology expands the production possibility frontier vertically (Figure 3A). After population adjusts, the economy returns to the constant population curve (Figure 3B). The new equilibrium (E'') is above the old one (E) because the production possibility frontier has become steeper.

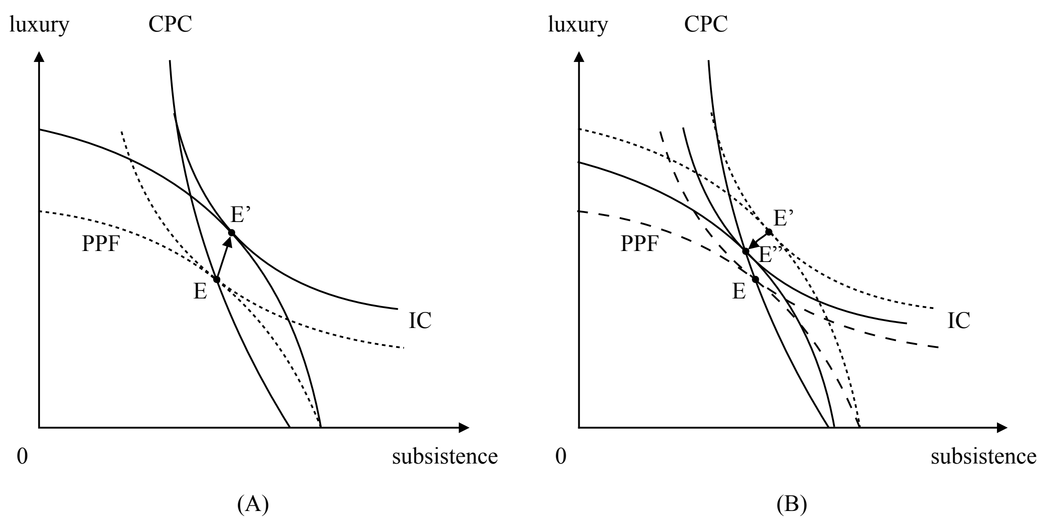

FIGURE 3: Progress of luxury technology improves equilibrium living standards.

However, long-term living standards will decrease if progress occurs in the subsistence sector instead. As Figure 4A shows, the production possibility frontier expands horizontally as the subsistence sector expands. The abundance of subsistence goods increases population.

[4]The definition of luxury determines the direction of crossing. If we denote the consumption of subsistence as E_A and the consumption of luxury as E_B, the definition of luxury implies

$$\frac{\frac{\partial U(E)}{\partial E_A}}{\frac{\partial n(E)}{\partial E_A}} < \frac{\frac{\partial U(E)}{\partial E_B}}{\frac{\partial n(E)}{\partial E_B}} \Longrightarrow \frac{\frac{\partial U(E)}{\partial E_A}}{\frac{\partial U(E)}{\partial E_B}} < \frac{\frac{\partial n(E)}{\partial E_A}}{\frac{\partial n(E)}{\partial E_B}} \quad \text{i.e., } MRS_U < MRS_n, \tag{5}$$

so the constant population curve must be steeper than the indifference curve at E.

After the economy returns to the constant population curve, the new equilibrium stays below the old one, because the production possibility frontier has become flatter (Figure 4B). In the long run, what matters for living standards is not the size but the shape of the production possibility frontier.

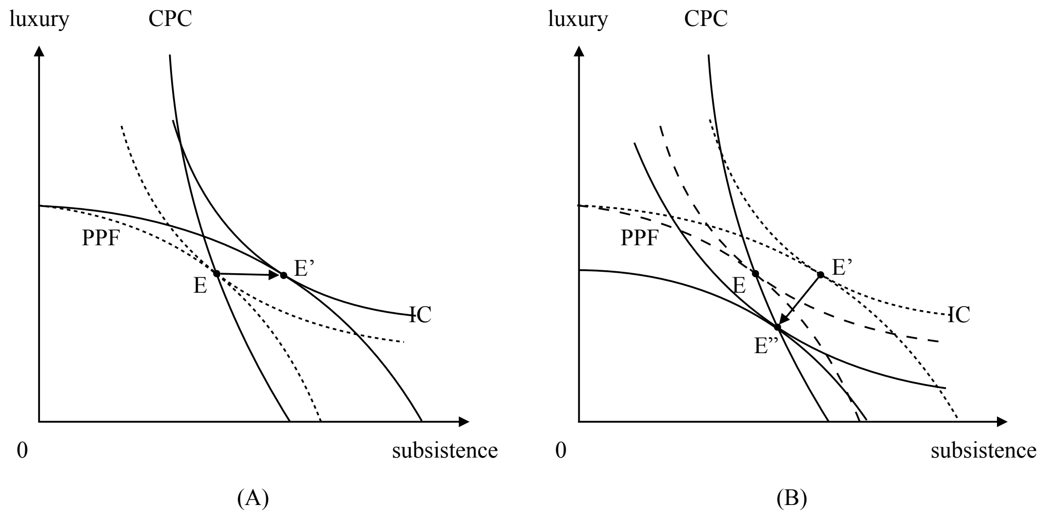

FIGURE 4: Subsistence technological progress decreases equilibrium living standards.

Luxury expansion characterizes the impact that market-oriented economic policies would have on an ancient economy, such as the Roman Empire and the Sung dynasty of China. Markets boost both agriculture and manufacturing, but manufacturing usually reaps more benefits from markets than agriculture does, so the equilibrium living standards will rise as a result of the tilted production structure. The Malthusian force would have no way to check the improvement.

Paradoxically, subsistence expansions, such as the Agricultural Revolution, only drag an economy into deeper poverty. Archaeological evidence shows that the ancient peasants lived a worse life than their hunter-gatherer ancestors: leisure time was shortened; diet became less diversified; and harvest failures caused more frequent starvation (Diamond, 1987).

Besides technology, the two-sector model also gives culture an important role to play. Suppose there is a cultural shock that makes people desire more luxury—the indifference curve becomes flatter (Figure 5A). Imagine the luxury culture as one that promotes a new conspicuous consumption used to signal unobserved income (Moav and Neeman, 2008). Those who dared reject the cultural norms might have more food to eat but would be less attractive on the marriage market. Individual reproductive rationality requires people to spend resources on vanity. As people trade subsistence for luxury, population undergoes a gradual decline.

When the adjustment is over, those who remain turn out to enjoy higher equilibrium living standards (Figure 5B).

In a word, luxury is *socially free*: so long as everyone desires more, more is granted to each who survives. Who pays for the extra luxury? It's those who would have been born and those who would not have died that pay with lives.

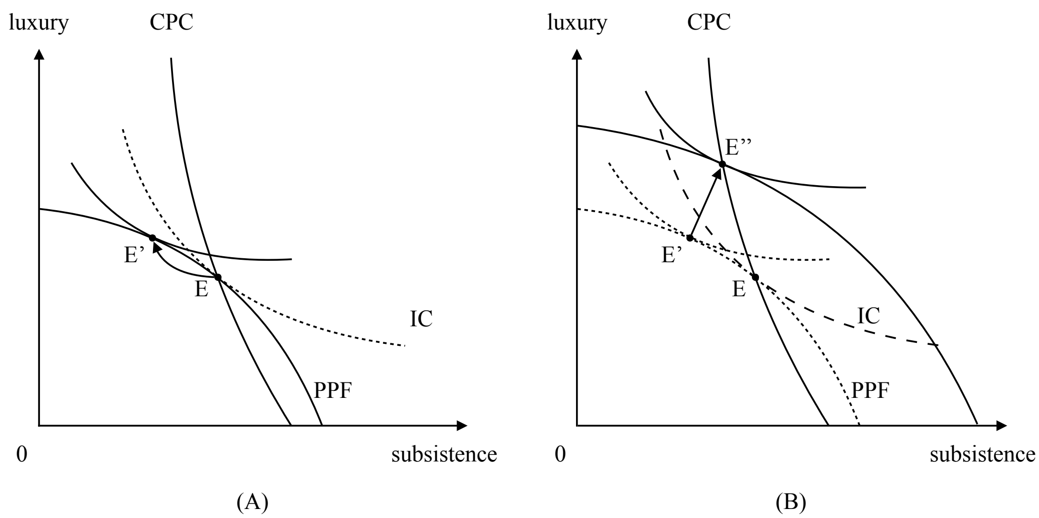

FIGURE 5: Luxury culture shock increases equilibrium living standards.

The above results seem to suggest that technologies and cultures that are more biased toward luxury always yield higher living standards. The correlation holds most of the time but exceptions exist. The monotonicity is broken if there exist multiple equilibria.

If subsistence goods are Giffen, it is possible that the economy can be locked at an inferior equilibrium after a subsistence-biased technological *decline*. Despite a more luxurious production structure, the temporary poverty, through the income effect, drives people to a lifestyle so subsistence-oriented that living standards become permanently lower. There are two ways to rule out this scenario. The first way is to narrow the focus to technological progress only. Hence the following theorem:

Theorem 1 (The First Production Structure Theorem)**.** *For an economy already on a stable equilibrium, a positive luxury technological shock always improves equilibrium living standards.*

The second way is to avoid Giffen goods.

Theorem 2 (The Second Production Structure Theorem)**.** *Other things being equal, if subsistence goods are not Giffen, a more luxury-biased production structure always means higher equilibrium living standards.*

The first theorem is useful in evaluating the long-term impacts of technological changes on a society's living standards. The second theorem helps compare living standards and production structures across societies.

Similarly, to establish the monotonicity between social preference and equilibrium living standards, I only need to rule out the case where technologies decline sharply when the whole family of indifference curves becomes flatter in the meantime. When subsistence is Giffen, the income effect is likely to dominate the luxurious shift in culture. By ruling out either the declines or the possibility of subsistence being Giffen, I establish the following theorems about social preference.

Theorem 3 (The First Free Luxury Theorem)**.** *For an economy already on a stable equilibrium, a luxury cultural shock always improves equilibrium living standards.*

Theorem 4 (The Second Free Luxury Theorem)**.** *Other things being equal, if subsistence goods are not Giffen, a more luxury-biased culture always means higher equilibrium living standards.*

The proofs are available in the online appendix[5]. The theorems add two new dimensions of comparative statics to the classical Malthusian theory, namely the production structure and the social preference. The classical theory, in contrast, has only one exercise in comparative statics, the one that is associated with the constant population curve. This single "classical" comparative statics, it turns out, works well in the two-sector model too. When the disease environment worsens, warfare becomes more frequent, or people decide to postpone marriage and have fewer children, the constant population curve simply shifts rightward: population grows more slowly at each level of consumption (Figure 6A). The ensuing decline of population expands the production possibility frontier. People who survive the changes are better off in the new equilibrium than in the old equilibrium (Figure 6B). The two-sector model thus preserves the merit of the classical theory.

3.3 Classical Malthusianism as a special case

The classical theory is actually a special case of the two-sector model. When $U(\cdot)$ is a transformation of $n(\cdot)$, all commodities have the same level of luxuriousness, and the constant population curve will coincide with the indifference curve. As the two curves coincide, a luxury technological shock can change the consumption bundle but not the equilibrium living standards. The return of the economy to the constant population curve will bring about the same level of utility as in the original equilibrium (Figure 7B).

[5]http://wulemin.weebly.com/uploads/1/4/6/2/14620598/malthusonlineappendix.pdf

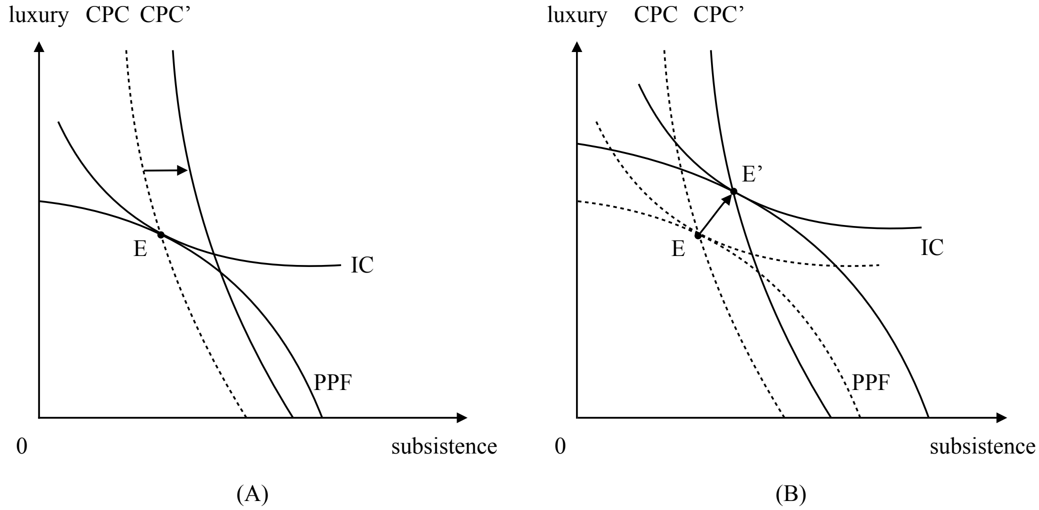

FIGURE 6: Diseases, wars and delayed marriage increase equilibrium living standards.

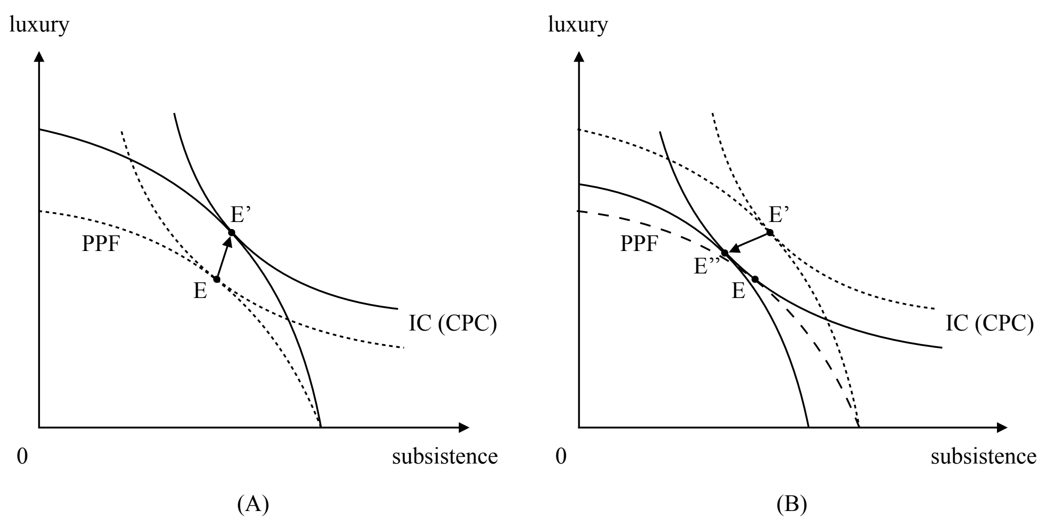

FIGURE 7: Malthusian theory is a special case where population growth and individual utility are fully aligned.

This is why the classical theory predicts stagnation. By the coincidence of curves, two sectors are reduced to one; the difference of demographic effects between bread and flowers is ignored; and the conflict of reproductive interest between individual and group is assumed away.

The last point bears emphasizing. The conflict between individual and group is the ultimate reason why the curves cross. The constant population curve is an iso-group-fitness curve, along which population grows at the same rate; the indifference curve is an iso-utility curve, or approximately an iso-individual-fitness curve. Millions of years' natural selection shapes human beings' preference system into a maximizer of one's own reproductive success. The conflict of reproductive interests between individual and group prevails in both culture and nature.

Luxury in human society is analogous to peacock's tails and guppies' spots. As signals of health, these traits promote the owners' chance of mating. But mating is a zero-sum tournament. One's gain is another's loss. Diverting nutrition to signalling limits the resources available for feeding and sheltering. The whole species' density is lower if sexual competition escalates.

For the individual, there is no distinction between luxury and subsistence. One way or another, both enhance one's own chance of survival and reproduction. But for the group, the division matters. A simple way to endogenize the difference is to specify an "individual population growth rate function" (akin to biological fitness) that has the following form:

$$n_i = \delta \left[\ln \left(\frac{x_i}{\bar{x}} \right) + \beta \left(y_i - \frac{\Sigma y_i}{H} \right) \right] \tag{6}$$

where x_i and y_i are the individual's consumption of subsistence and luxury, $\bar{x}$ is the threshold level of subsistence that keeps population balanced, and $\frac{\Sigma y_i}{H}$ is the average luxury of the whole group. Natural selection entails that the individual's preference should maximize her own fitness. So long as the group is large enough for the individual to treat the average luxury as given, the individual's preference can be represented as

$$u_i = \ln \left(\frac{x_i}{\bar{x}} \right) + \beta y_i. \tag{7}$$

Equation 7 rules out envy as a source of negative utility, which misses an important aspect of human psychology, but is closer to how economists measure income and welfare in modern society. It is crucial that we should apply a consistent criterion to evaluating living standards both in the ancient times and in the modern times.

The group's population growth rate is the sum of the individuals' population growth rates.

As the luxury terms are canceled out, the group's population growth rate is independent of luxury consumption:

$$n = \frac{\Sigma_{i=1}^{H} n_i}{H} = \delta \left(\frac{\Sigma \ln x_i}{H} - \ln \bar{x} \right). \tag{8}$$

The above exercise illustrates how a commodity can serve the individual's fitness maximization while having little impact on the group's fitness. For a full model that motivates luxury as a sexual signal, please consult De Fraja (2009).

The individual's and the group's reproductive interests are aligned over subsistence consumption but in conflict over luxury consumption. As surely as the conflict persists, the division of sectors is a perpetual human condition. Assuming the two interests to be perfectly aligned is simply unrealistic.

There are still many who believe there is nothing wrong with accepting an unrealistic assumption so long as the theory makes the right prediction, and that among all theories that predict correctly, the simplest is the best (Friedman, 1953). In their view, Malthusian theory is great. It might fail to predict short-term changes, but *ad hoc* remedies are available; most importantly, the theory predicts the Malthusian trap. So why bother to find an alternative?

The above view makes a serious methodological mistake. Unrealistic assumptions are fine only if they are not crucial. Otherwise—if making the assumptions more realistic would lead the prediction astray from reality—the theory must be wrong (Solow, 1956). Malthusian theory assumes the harmony of reproductive interests between individual and group. It is a very strong assumption that is obviously unrealistic. If we relax the assumption, the model predicts a trend of growth in living standards (this part will be shown in section 3.6). Hence, Malthusian theory alone cannot explain the Malthusian trap. The millennia of economic stagnation must be caused by something else.

3.4 Evidence of multi-sectorality

How relevant is sectoral division in actual history? The short answer is: very much. Whether the division is recognized decides views as important as when the Great Divergence occurred and how real income inequality evolved.

Based on the similarity of wages in terms of calories, Pomeranz (2009) argued that the Yangzi Delta of China was on the same level of development as Northwest Europe as late as the end of the 18$^{\text{th}}$ century. His work implicitly assumed that the per capita calories are fixed by the Malthusian force and did not consider that most of the difference in living standards comes from the difference in non-food consumption. As Broadberry and Gupta (2006) showed, if we measure the purchasing power of wage by grams of silver, the "silver wage" of the Yangzi Delta was only comparable with the level of the central and eastern parts

of Europe (Table 1). Regions varied widely in the ratio of silver wage to grain wage. That the ratio was higher in England and the Netherlands than in other places reflected these two countries' relative advantage in tradable goods production, and it was by this advantage that the Northwest European economies stood out in the pre-modern centuries.

Table 1: The grain wage and silver wage of different regions, 1750-99

Regions	Grain wage (kg/d.)	Silver wage (g/d.)	Ratio (Sw/Gw)
Southern England	7	8.3	1.2
Antwerp	9.6	6.9	0.7
Vienna	7	3	0.4
India	2.3	1.2	0.5
Yangzi delta, China	3	1.7	0.6

Data source: Broadberry and Gupta (2006).

Hoffman et al. (2002) studied the implication of sectoral division for real income inequality. Because the rich spent a larger portion of income on luxury than did the poor, the decline of the prices of luxuries relative to the prices of subsistence goods enlarged real inequality in Europe between 1500 and 1800.

More direct evidence for multi-sectorality is provided in Wu (2012). In that paper, I did a simple exercise: instead of regressing birth and death rates on real wages, as most people do when testing the Malthusian relationship, I regress these demographic variables on "sectoral wages", that is, the purchasing power of wages measured in terms of goods of certain sectors. During the three centuries before 1800, the period conventionally believed to have provided the strongest evidence in support of Malthus, multi-sectorality turned out to be a salient feature of the English economy. If pasture goods became more affordable, population growth barely changed; if arable goods were more affordable, population growth rate increased a lot. Within the category of arable goods, the affordability of wheat had little impact on population growth, but the affordability of barley and oats—the poor people's staple foods—almost solely explained the impact of real wages on birth and death rates. However, barley and oats were merely 10% of the English economy, much smaller than the share of wheat. The remaining 90%, including wheat, beef, cotton and candles, hardly mattered demographically. Productivity improvement in the 90% sector surely would have increased the long-term living standards, with more families switching their diet from porridge to bread, and starting to call tea, sugar and coffee "necessities" of life. Changes like this might not be reflected in the real wage series historians have built, but, with the proper data and method, it is still possible to assess this part of the welfare increase.

Hersh and Voth (2009) estimated that, by the end of the 18$^{\text{th}}$ century, tea, coffee and

sugar had added at least the equivalent of 16% (and possibly as much as 20%) of income to English welfare. Contrast it with two other New World crops, maize and potatoes. Chen and Kung (2013) estimated that maize accounted for 18% of the population increase in China during 1776-1910, but had no significant effect on economic growth. Nunn and Qian (2011) estimated that potatoes accounted for about a quarter of the growth in Old World population between 1700 and 1900. Potatoes triggered a Malthusian crisis in Ireland in the late 18th century; the explosion of population drove the Irish to extreme poverty that culminated in the Great Famine of 1845 (Mokyr, 1983). The difference between tea, coffee and sugar on the one hand, and maize and potatoes on the other hand, is none other than the difference between luxury and subsistence. While the abundance in the former improved the quality of life, the abundance in the latter increased the population.

3.5 The algebraic two-sector model

This section builds a simple algebraic model that captures three comparative statics. Assume the representative agent maximizes a Cobb-Douglas utility function over her subsistence consumption, x and luxury consumption, y.

$$\max U(x,y) = x^{1-\beta} y^{\beta}. \tag{9}$$

The constant returns to scale makes the magnitude of utility meaningful: utility doubles when consumption doubles.

Specify the subsistence production function as $X = AL_A^{1-\gamma_A} H_A^{\gamma_A}$ and the luxury production function as $Y = BL_B^{1-\gamma_B} H_B^{\gamma_B}$. L_A and L_B are land used in the production of subsistence and luxury respectively. Their sum is the total endowment of land, $L_A + L_B = L$. H_A and H_B are labor employed in the respective sectors, and $H_A + H_B = H$.

Assumption 1. $\gamma_A = \gamma_B \equiv \gamma < 1$.

I assume $\gamma_A = \gamma_B$ so that population growth affects the two sectors in proportion, equivalent to the shape-preserving assumption in the graphical model. γ_A and γ_B are smaller than one because of diminishing returns to labor.

By maximizing the agent's utility under land and labor constraints, we can derive her consumption bundle:

$$x = A(1-\beta)\left(\frac{H}{L}\right)^{\gamma-1} \tag{10}$$

$$y = B\beta\left(\frac{H}{L}\right)^{\gamma-1} \tag{11}$$

Substitute Equation 10 and Equation 11 into $U = x^{1-\beta}y^{\beta}$. The level of utility is

$$U = A\left(\frac{H}{L}\right)^{\gamma-1}\left(\frac{B}{A}\right)^{\beta}(1-\beta)^{1-\beta}\beta^{\beta}. \tag{12}$$

Because $A(H/L)^{\gamma-1}(1-\beta) = x$ (Equation 10), U can be expressed alternatively as

$$U = x\left(\frac{B}{A}\right)^{\beta}\left(\frac{\beta}{1-\beta}\right)^{\beta}. \tag{13}$$

The economy converges to equilibrium by population adjustment. In an isolated economy, the *net* growth rate of population, g_H, is equal to the *natural* growth rate of population, n. Assume that n depends on the average consumption of subsistence only.

Assumption 2. *$g_H \equiv \dot{H}/H = n = \delta(\ln x - \ln \bar{x})$, and $\delta > 0$.*

$\bar{x}$ is the level of average subsistence at which population remains constant. The assumption means a vertical constant population curve—population growth is independent of average luxury consumption. In the equilibrium, $x = \bar{x}$. Therefore,

Proposition 1.

$$U^E = \bar{x}\left(\frac{B}{A}\right)^{\beta}\left(\frac{\beta}{1-\beta}\right)^{\beta} \tag{14}$$

The equilibrium utility increases with

(a) the relative luxury productivity, $\frac{B}{A}$,

(b) the relative preference for luxury, β,

(c) and the required consumption for population balance, $\bar{x}$.

It is worth noting that the above model uses an exogenously given function of population growth. However, the new standard of the literature is to endogenize the birth rates from households' fertility choices. Yet I stick to the exogenous functions because, as Weisdorf (2008) demonstrates, adding such a microfoundation will not change any of the above results. When subsistence is relatively cheaper than luxury, raising children becomes less costly to parents. The result is the same: the equilibrium income per capita depends on production structure. I abstract from fertility choices to save unnecessary complications and to highlight the key mechanisms at work—it's not fertility choice but the physiological nature of commodities that matters.

3.6 The balanced growth puzzle

Proposition 1 implies that living standards will rise steadily if luxury productivity grows faster than subsistence productivity. Denote g_A as the growth rate of subsistence productivity and g_B as the growth rate of luxury productivity. Appendix A.1 proves that g_U converges to $\beta(g_B - g_A)$ in the long run.

In the classical Malthusian theory, there is only one sector ($\beta = 0$); therefore $g_U = 0$. In the two-sector model, β is positive. The equilibrium living standards will have an upward or downward trend unless $g_B = g_A$.

Given Malthusian stagnation as a fact, the implied balanced growth is an extraordinary phenomenon. The world population had grown from several million at the dawn of the agricultural revolution, to three hundred million at the birth of Christ, and to almost one billion on the eve of the industrial revolution—it went up by a factor of at least 1000. To keep up, subsistence production must have grown by a thousand-fold; the subsistence technology A would have grown by about thirty-fold if $\gamma = 0.5$.

Throughout the thousand-fold growth of subsistence, luxury must have grown in exact proportion. What could keep things balanced over ten thousand years during a thousand-fold growth? In comparison, world population has grown only six-fold since the industrial revolution. In the meantime, the world GDP has grown by a factor of 150—the progress in luxury productivity has been much faster than the progress in subsistence productivity. If the balance of growth is so difficult to maintain within such a short period, isn't it extraordinary that growth was once balanced for not decades or centuries, but for millennia and even longer?

Moreover, it is natural to expect luxury production to intrinsically grow faster than subsistence production. There are at least four reasons why we should expect so. First, manufacturing and commerce are usually more labor-intensive than agriculture[6]. Population growth, by increasing labor supply, naturally expands luxury production more than it expands subsistence production.

Second, industrial innovations are much less constrained by the possibilities of nature

[6] I assume land plays equal parts in luxury production and subsistence production, i.e., $\gamma_A = \gamma_B$ (a simplifying assumption that is unfavourable to my hypothesis). If instead I allow subsistence to rely on land more than luxury does—which would be a very reasonable assumption, for subsistence is mostly about food and basic shelter—then balanced growth alone is enough to cause steady progress of living standards. Nevertheless, I keep the stricter assumption of $\gamma_A = \gamma_B$ for model tractability. That said, it is admittedly imprecise to call balanced growth the stagnation condition, for the condition should take into account the gap between γ_A and γ_B. Dutta et al. (2014) derive the γ-adjusted balanced growth condition as

$$g_B - \frac{1-\gamma_B}{1-\gamma_A} g_A = 0. \tag{15}$$

than are agricultural innovations.

Third, the incentives for industrial innovations are better protected than the incentives of agricultural innovations. An ancient farmer who succeeded with a new crop could hardly reap any of the social benefit that spilled out of her own land. By contrast, in manufacturing and commerce, keeping trade secrets for monopoly rent was feasible most of the time.

Last but not the least, manufacturing allows a larger extent of division of labor. As Adam Smith (1887, Chapter I, Book I) put it,

> "The nature of agriculture, indeed, does not admit of so many subdivisions of labour, nor of so complete a separation of business from another, as manufacture. [...] This impossibility of making so complete and entire a separation of all the different branches of labour employed in agriculture, is perhaps the reason why the improvement of the productive powers of labour, in this art, does not always keep pace with their improvement in manufactures. The most opulent nations, indeed, generally excel all their neighbors in agriculture as well as in manufactures; but they are commonly more distinguished by their superiority in the latter than in the former."

When the above reasons are combined, the Malthusian trap appears to be a most unlikely coincidence. Therefore, the Malthusian fact is still a fundamental puzzle of history. Malthus has not solved the most mysterious part of the puzzle. Below, I discuss five possible solutions. Some of them have serious limitations.

Evolutionary adaption

Long exposure to a luxury good might cause genetic adaption that allows people to use it as a subsistence good. For example, lactose intolerance is relatively rare among Northwest Europeans, whose ancestors, as a conjecture goes, had a higher reliance on milk as a source of nutrition than did people in Asia, who did not develop the gene. The problem is that, even if the conjecture is correct, genetic adaption is usually slow, and the mechanism that works for food does not work for manufactured products, which are a more important category of luxury goods.

Positional goods

Diamonds are precious because they are rare. Positional goods become worthless when there are too many: people value how much they own compared with others instead of what they own *per se*. The abundance of a particular luxury drives people away from the luxury. The

problem with this hypothesis is that the Malthusian fact is not about the lack of desire, but the shortage of goods. The mechanism might explain why being rich does not make one much happier, but it does not explain why physical deprivation lasts.

Constant returns to scale

Solow and Samuelson (1953) showed that, in a dynamic system described as

$$\begin{aligned} A_{t+1} &= F_A(A_t, B_t) \\ B_{t+1} &= F_B(A_t, B_t), \end{aligned}$$

if $F_A(\cdot)$ and $F_B(\cdot)$ have constant returns to scale, then A and B will grow in balance on a stable path. But the theorem only pushes back the question one step further. It is doubtful whether the theorem is applicable to luxury and subsistence growth. Even if it is applicable, we still have to answer why the functions have constant returns to scale. So far, I have seen no reason why that should be so.

Group selection

This hypothesis holds that, for most of ancient history, there was a systematic bias in the direction of migration and conquest: people and civilizations from relatively subsistence-rich groups tended to conquer and replace those from relatively luxury-rich groups. High-living-standard societies were rare in the ancient times because the only way to achieve high living standards under the Malthusian constraint was to have a highly developed luxury sector, but if a society spends too many resources on luxury, it invites invasion and immigration from relatively subsistence-rich groups that finally erode the luxury culture. Because of the bias, groups that value austerity are more likely to linger than groups with licentious lifestyle. The bias also enables subsistence technology to spread faster than luxury technology. Even if luxury technology intrinsically grows faster than subsistence technology, the spread advantage of subsistence might offset the growth advantage of luxury. This is the hypothesis that I propose as the solution to the balanced growth puzzle. But before I move on to the details of the group selection theory, I shall address one last hypothesis, Strulik and Weisdorf (2008)'s theory, the only explanation that has been proposed in the literature.

Learning by doing

Strulik and Weisdorf (2008) hypothesized that agricultural technological progress was once dominant over manufacturing technological progress because agriculture employed a larger

share of labor. The larger size of labor employed in agriculture gave subsistence an early advantage in growth because, supposedly, most innovations in the ancient times came from learning by doing. But agriculture is subject to decreasing return to labor more than manufacturing is; over time, the development of agriculture diverted population, hence the chance of learning by doing, to the manufacturing sector, which finally triggered the Solow-style economic growth. The hypothesis captures some aspects of historical reality but it has four limitations, which makes it unlikely to be the complete explanation of the Malthusian puzzle.

First, agriculture and manufacturing are only a special case of subsistence and luxury. The luxuriousness of commodities is correlated with the degree of reliance on land, but they do not always coincide. Demographically speaking, as Wu (2012) shows, 90% of the 17^{th}-century English economy was effectively luxury, yet if we narrowly define luxury as non-agricultural goods, the proportion was less than half. Strulik and Weisdorf (2008)'s theory cannot explain why the luxury goods in the agricultural sector were as suppressed as the luxury goods in the manufacturing sector, and why subsistence goods in the manufacturing sector were as prevalent as the subsistence goods in the agricultural sector.

Second, agriculture and manufacturing are two large categories that each contains hundreds of, if not thousands of, smaller fields. The fact that agriculture initially employed more labor does not mean that each field in agriculture was larger than the fields in manufacturing. If we ignore the inter-field spillover of technologies, there is no reason why a larger category should systematically grow faster than a smaller category at all. Even if we allow spillover, it bears explanation why the spillover is more effective within categories than across categories, which, at least in the modern times, does not seem to be the case.

Third, the generation of ideas depends not only on the aggregate number of practitioners but also on the density of artisans. It takes density for innovators to inspire ideas in each other. Even if manufacturing and commerce employed fewer people than agriculture[7], did the non-agricultural sector necessarily grow more slowly? Consider how many ideas the ancient Greeks have left to the world. When it comes to industry and commerce, the size of labor force alone does not seem to matter that much.

Last but not the least, Strulik and Weisdorf (2008)'s theory predicts that once the manufacturing sector has reached a decent size, the Industrial Revolution and the Demographic Transition are inevitably to come the next. The theory cannot explain the regression of technology following the collapses of the Hittite Kingdom, the Roman Empire, the Sung China, and many other civilizations that achieved high level of development in manufacturing and commerce. As will be discussed in section 5, the Sung Chinese government collected two-

[7] I doubt that the ratio of non-agricultural labor to agricultural labor would be lower than a third in Ancient Greece, the Roman Empire and the Sung China.

thirds of tax revenue from industry and commerce; both Sung and Roman Italy achieved a GDP per capita of nearly $1500, while the contemporary agrarian economies lived on about $450 only. The manufacturing sector reached a huge size quite long ago. In fact, as Jones (1988) observed, "Intensive growth arose in times and places that to all intents and purposes were unrelated to one another [...] We may be almost more surprised by the power of negative forces to suppress the shoots of growth than by the fact that they were occasionally overcome."

The group selection mechanism, proposed in this paper, is complementary to Strulik and Weisdorf (2008)'s mechanism. While the latter accounts for the accelerating switch of labor out of agriculture and the resulting acceleration of growth in industrial production, the group selection theory addresses why the process was usually interrupted so that the Malthusian trap could have lasted for thousands of years.

4 A group selection theory of the Malthusian trap

I have shown that it takes balanced growth between the luxury and subsistence sectors for living standards to remain constant, and that the balanced growth puzzle cannot be explained by the evolutionary adaptation hypothesis, the positional goods hypothesis or the constant returns to scale hypothesis. This section elaborates on a fourth alternative, the group selection mechanism.

Selection works through biased migration. I first explain what biased migration is, and discuss historical evidence in support of the migration pattern. Next, I build a group selection model that allows me to derive the threshold condition of stagnation. Finally, simulations are conducted to compare the paths of global average living standards with and without selection. Both model and simulation confirm that a Malthusian world without selection would have an upward trend of growth in average living standards, but when selection is introduced—a tiny bit of selection would suffice—the trend is gone.

4.1 Biased migration

Group selection is related to the phenomenon of "biased migration". This section uses a simple model to explain how certain characteristics in culture and technology bias the direction of migration.

Suppose there is a sea of identical villages, all at the equilibrium state. Following Tiebout (1956), I assume free migration across villages but forbid trade between them.[8] Bread and

[8] Trade substitutes for migration. If trade is free of cost, different regions will face the same relative price

flowers are the only commodities. Suddenly, one of the villages discovers a better way to grow flowers. Its production possibility frontier expands vertically. If migration were forbidden, the flower village would end up with higher living standards. But free migration equalizes utility across the villages (Figure 8B). With a steeper production possibility frontier tangent with the same indifference curve, the flower village stays to the left of the constant population curve in the migration equilibrium—its death rate is higher than its birth rate. The natural decrease of population does not expand the production possibility frontier because the under-reproduction is filled up with continuous immigration from the other villages. The flower village becomes a demographic *sink* and the surrounding villages a demographic *source*.

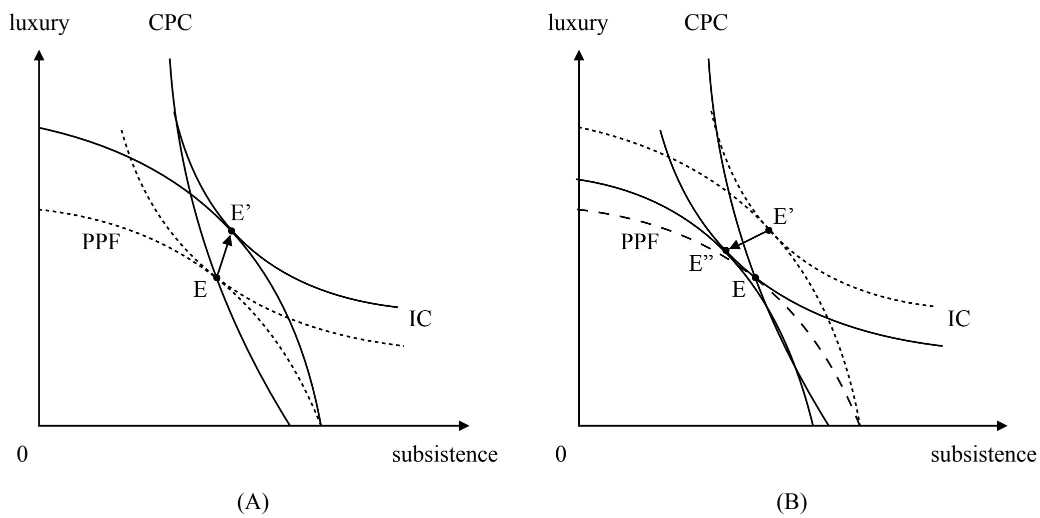

FIGURE 8: Source-sink migration emerges out of difference in production structure.

Differences among cultures cause source-sink migration too. Suppose in one of the villages, young women begin to ask for more flowers from their suitors: the indifference curve becomes flatter (Figure 9A). If migration were forbidden, the women who survive would get what they demand for free (remember the free luxury theorem). However, in the migration equilibrium, the equality of utility means demographic imbalance.

The imbalance does not arise immediately after the social preference change. People in the surrounding villages do not move in the beginning. They stay put until the population of the flower village decreases enough for the economy to move from E' to E'' (Figure 9B). After E'', the continuous immigration keeps the flower village to the left of the constant

of luxury to subsistence. The Malthusian force then equates consumption across regions, and there will be no need to migrate. But if trade has a cost, the relative price will differ and migration will emerge. I forbid trade only to simplify the analysis. The assumption is not crucial. The model applies so long as trade has a cost.

population curve. The source-sink pattern emerges again.[9]

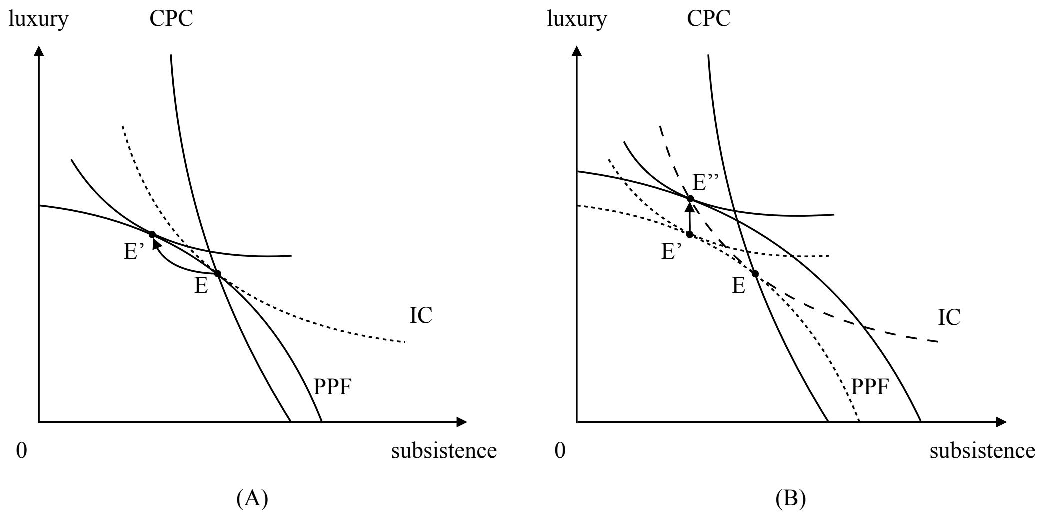

FIGURE 9: Source-sink migration emerges out of difference in social preference.

The craze for flowers will not last forever, because the immigrants who come from the places that do not value flowers as much will be diluting the flower culture[10]. The arms race of conspicuous consumption is constrained not by Malthusian forces, but by source-sink migration, and the selection that follows. Selection dissipates luxury cultures and diffuses subsistence cultures.

4.2 Evidence of biased migration

Source-sink migration is best documented in the context of rural-urban migration, where the phenomenon is sometimes called "urban natural decrease": in pre-modern Europe, the urban death rate was higher than the birth rate, and the natural decrease coincided with the natural increase in the surrounding rural area.

De Vries (2006) decomposes the net changes of pre-modern European urban population into net immigration and natural growth. As Figure 10 shows, during most of the time between 1500 and 1800, urban population had been growing in both Northern and Mediterranean Europe. Despite the *net* increase, urban population had been declining *naturally*, that is, the death rate was higher than the birth rate in the cities. During the half century between 1600 and 1650, Northern Europe had an annual growth of 0.32% in its urban

[9]Here the migrants are assumed to keep their old preference. If they convert to new cultures, the diagram is slightly different but the source-sink pattern still remains.

[10]The migrants might be assimilated into the host culture to some extent, which slows the dilution.

population; meanwhile, the urban death rate exceeded the birth rate by 0.33%. A flow of rural migrants that amounted to 0.65% of the size of urban population per year had been replenishing the cities.

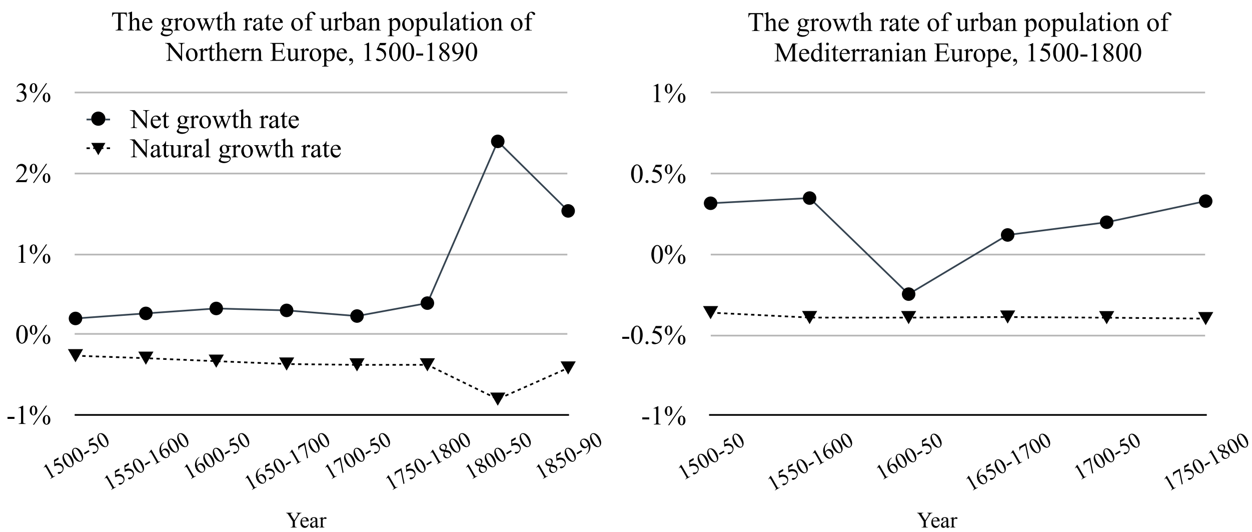

FIGURE 10: The source-sink pattern of migration in pre-modern Europe.
Data source: De Vries (2006, p.203-208).

Malthusian theory is inconsistent with the migration pattern. The period 1800 – 1850 witnessed a spike in the growth of urban population in Northern Europe. According to the classical theory, the spike suggests a rise in urban living standards, which would cause a faster natural growth in the urban population. In fact, the gap between the urban death rate and the urban birth rate only widened in this period.

The anomaly can be easily explained by the biased migration model. After 1800, the growth of manufacturing and commerce accelerated in the urban areas; the growth of agriculture accelerated in the rural areas. The polarization of production structures spurred more migration into the cities than in the previous centuries. The flood of immigrants lowered the average subsistence, including hygiene and workplace safety, by so much that the natural growth rate of the urban population dropped further. As Proposition 2 later shows, the depth of the demographic sink increases with the distance of production structures.

Another piece of evidence for biased migration comes from Ravenstein (1885), who calculates the gap between a county's population and the number of its natives, enumerated throughout England and Wales in 1881. When there were more residents than natives, he regards the county as one of absorption—more people moved in than moved out; otherwise, it was a county of dispersion. Ravenstein also marks whether a county was "agricultural" or "industrial". He calls a county "agricultural" if the county's proportion of agricultural population exceeded the sample average, and "industrial" otherwise.

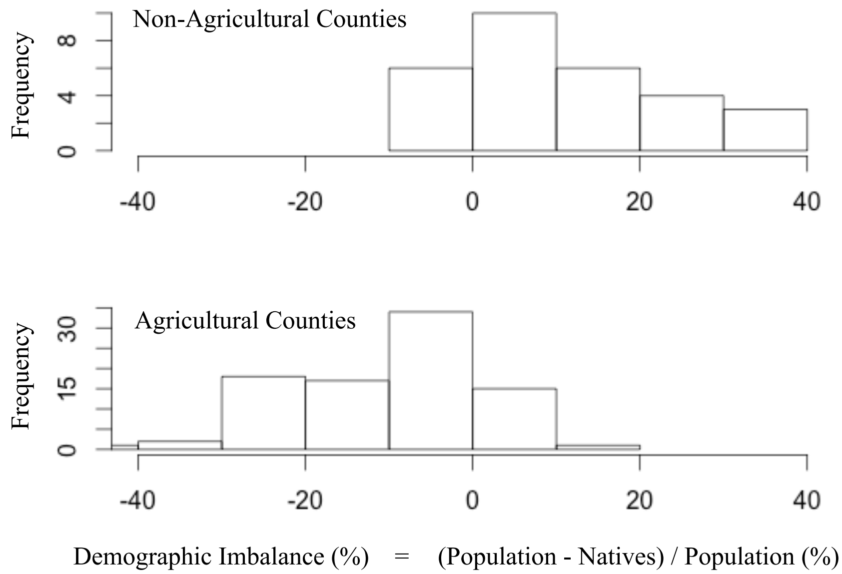

FIGURE 11: The demographic imbalances of non-agricultural and agricultural counties. Data source: Ravenstein (1885, p.185-186).

As Figure 11 shows, the overwhelming majority of the non-agricultural counties were counties of absorption, and the overwhelming majority of the agricultural counties were counties of dispersion. The pattern is consistent with the hypothesis that people usually move from subsistence-rich regions to luxury-rich regions.

Source-sink migration is not limited to the rural-urban context. Production structure can differ over a much larger scale, say, between the south and north of a continent. The difference will trigger biased migration too, but data are hardly available on such a large scale. That is why I focus on the rural-urban context to demonstrate the source-sink migration.

4.3 The spread of ideas through migration

So long as ideas move with people, the bias of migration naturally gives rise to bias in the spread of ideas. Because people usually move from subsistence-rich regions to luxury-rich regions, a subsistence culture (technology) is easier to spread than a luxury culture (technology), other things being equal.

Today, anyone with electricity and Internet can watch lectures taught by the best scholars in the world. Technological spread and migration are largely disentangled from each other. But in ancient times, when the supply of books was limited, migration was crucial to the spread of ideas. It took thousands of years for agriculture to spread from the Fertile Crescent to Northern Europe, and the process coincided with the spread of the original Neolithic

groups' genes in both timing and spatial extent (Cavalli-Sforza, Menozzi, and Piazza, 1993). If the hunter-gatherers beyond the frontiers had learned agriculture without the immigrants' help, the reproductive advantage of the first farmers would have been quickly lost, and the observed pattern of genetic spread would be impossible. Likewise, the Indo-European family of languages originated in the Caucasian steppe, where the early domestication of horses lent serious military advantage to the herders. If horse-herding had spread without migration, the Caucasians would have found it hard to conquer the neighbouring peoples who had learnt the new arts of war, and the proto-Indo-European language would have had no chance to diffuse at all.

Even in societies with decent literacy levels, migration still played a crucial role in the spread of ideas. In the 15^{th} century, the learned Byzantine exiles who fled out of the falling Constantinople revived Greek studies in Renaissance Italy. Considering the potential demand for Greek letters and thoughts, it is extraordinary that the knowledge had not spread earlier to Italy by means other than migration. Another example is the revolutionary impact of Jewish emigres from Nazi Germany on U.S. science. Moser, Voena, and Waldinger (2014) estimated that the arrival of German Jewish emigres brought a 31 percent increase in patenting by U.S. inventors in the emigres' research fields after 1933. Even today, complaints about "brain drain" are frequently heard. Again, people are concerned about the spread of technology through migration.

It is fair to say that, at least in ancient times, migration was an important channel of spreading ideas. The bias of migration means a bias in technological diffusion. The question is: can this bias explain the Malthusian trap?

4.4 The group selection models

I build two group selection models to answer the above question. The first model studies a single village that is surrounded by an infinite number of villages. I call the case "partial equilibrium" because the model treats the population and technology of the surrounding villages as fixed. The relatively simple result highlights the key mechanism at work.

The second model studies the general equilibrium of two villages, allowing them to fully interact with each other. A threshold condition is derived to tell when selection dominates growth and when growth overpowers selection. Both models assume away cultural selection by keeping β fixed. Focusing on technological selection alone, I can easily extend the results to cultural selection.

4.4.1 The partial equilibrium

Suppose there are an infinite number of villages, as described in section 3.5. All villages begin with the same technological levels and population sizes. As time elapses, bread and flower technologies stagnate at A' and B' in all villages except one. Use an asterisk to denote that special village. Its subsistence technology stagnates too, $A^* = A'$, but its luxury technology B^* tends to grow at the rate of g. We call it the flower village and the others the bread villages. When B^* exceeds B', there will be continuous migration from the bread villages to the flower village.

Assumption 3. *Trade is forbidden across the villages but migration is free of cost.*[11]

Free migration equalizes the level of utility, $U^* = U'$, which by Equation 13 means

$$x^* \left(\frac{B^*}{A^*}\right)^\beta \left(\frac{\beta}{1-\beta}\right)^\beta = x' \left(\frac{B'}{A'}\right)^\beta \left(\frac{\beta}{1-\beta}\right)^\beta$$

where x^* and x' are the average consumption of bread. Rearrange the equation and take the logarithm:

$$\ln x^* - \ln x' = -\beta \left[\ln\left(\frac{B^*}{A^*}\right) - \ln\left(\frac{B'}{A'}\right)\right]. \tag{16}$$

The net emigration rate from the flower village, m, is equal to the *natural* growth rate of population, n, which in turn depends on the average bread, x^*, that is,

$$m = n = \delta(\ln x^* - \ln \bar{x}). \tag{17}$$

$\bar{x}$ is the level of average bread that keeps the population in natural balance. Because $\delta > 0$ and $x^* < \bar{x}$, m is negative: migrants move from the bread villages into the flower village. The emigration has a negligible effect on each bread village, because their number is infinite and migration between them is frictionless. So the bread villages still have $x' = \bar{x}$.

Denote $s^* \equiv \ln(B^*/A^*)$ and $s' \equiv \ln(B'/A')$, the relative luxury productivities. Substituting $x' = \bar{x}$ and Equation 16 into Equation 17, we get

Proposition 2.

$$m = -\beta\delta(s^* - s') \tag{18}$$

The net emigration rate is proportional to the distance of production structures. Having a higher relative luxury productivity than the neighbouring villages causes net immigration.

[11] See footnote in section 4.1 for a justification of the assumption.

Migrants spread ideas. Assume that migration affects B^* by displacing hosts' technology with immigrants' technology in proportion to the number of immigrants:

Assumption 4. *From time t to $t+\Delta t$, B^* updates by taking the weighted geometric average of B^* and B' and growing at the rate of g.*

$$B^*(t+\Delta t) = B^*(t)^{1-m\Delta t}(B')^{m\Delta t}(1+g\Delta t) \tag{19}$$

Divide both sides of Equation 19 by A', take logarithms, and calculate the limit as $\Delta t \to 0$. We can rewrite the equation into the motion function of s^*:

$$\dot{s^*} = m(s^* - s') + g \tag{20}$$

Substitute Equation 18 into Equation 20:

$$\dot{s^*} = -\delta\beta(s^* - s')^2 + g \tag{21}$$

The differential equation has a stable equilibrium:

Proposition 3. *In the long run, even if B^* has an intrinsic tendency to grow at the constant rate g, the flower village's relative productivity, $s^* = \ln(B^*/A^*)$ will stabilize at*

$$s' + \sqrt{\frac{g}{\delta\beta}}$$

Note that g has a level effect but no growth effect on the equilibrium level of s^*.

4.4.2 The general equilibrium

Simplicity often entails some shortcomings. The partial equilibrium model assumes away the influence of the flower village on its neighbours. In this section, we study a two-village model to take into account this general equilibrium effect.

Suppose village 1 and village 2 start as identical. Their bread technologies, A_1 and A_2, grow at the same constant rate g_A, and their flower technologies, B_1 and B_2, drift with noise:

$$d\ln B_i = (g_A + g)dt + \sigma dz_i \tag{22}$$

Here $g > 0$ captures the growth advantage of flower productivity over bread productivity. The error terms z_i's ($i = 1, 2$) are Brownian motions. z_1 and z_2 are independent of each other, and $\text{Var}(\sigma dz) = \sigma^2 dt$. I introduce the stochastic growth of technology as the source

of inter-regional variation. Variation is the basis of technological selection as mutation is the basis of natural selection. I fix the growth rates of A_1 and A_2 to keep population equal between the regions. The equality of population makes the model tractable without loss of generality. With $s_i \equiv \ln(B_i/A_i)$, Equation 22 can be rewritten as

$$ds_i = gdt + \sigma dz_i. \tag{23}$$

The assumption that $g > 0$ allows both villages, if isolated, to grow steadily in living standards. However, selection cancels out growth by adding a "drag" term to the motion of s_i. The drag appears when $s_1 \neq s_2$. Following Assumption 4, the drag term is a quadratic of the difference between s_1 and s_2 as in Equation 21:

$$ds_i = [g - I_{\{s_i > s_j\}}\beta\delta(s_i - s_j)^2]dt + \sigma dz_i.$$

Here $I_{\{s_i > s_j\}}$ is an indicator function that equals 1 if $s_i > s_j$ and 0 otherwise. If $s_1 > s_2$, village 1 is relatively rich in flowers. It attracts immigration from village 2, which drags s_1 closer to s_2. If instead $s_1 < s_2$, village 1 is relatively rich in bread. The bread village receives no immigration and selection will not affect its relative luxury productivity.

Because utility depends on s, the most interesting variables are the global average of s_i's, $\mu = \frac{1}{2}(s_1 + s_2)$, and the inter-regional variation, $\nu = \frac{1}{2}(s_1 - s_2)^2$.[12].

Applying Itō's lemma, we get

$$d\mu = (g - \beta\delta\nu)dt + \frac{\sqrt{2}}{2}\sigma dz \tag{24}$$

$$d\nu = (\sigma^2 - 2\sqrt{2}\beta\delta\nu^{\frac{3}{2}})dt + 2\sqrt{\nu}\sigma dz \tag{25}$$

where z is a Brownian motion.

Taking the long-term expectation of both sides of Equation 24, we have

$$\mathop{\mathbb{E}}_{t\to+\infty}\left(\frac{d\mu}{dt}\right) = g - \beta\delta \mathop{\mathbb{E}}_{t\to+\infty}(\nu) \tag{26}$$

Denote $S \equiv \beta\delta\,\mathbb{E}_{t\to+\infty}(\nu)$. S is the force of selection. $(g - S)$ captures the race between growth and selection. If $g > S$, growth overcomes selection and living standards grow; otherwise, selection dominates growth.

[12] ν is the sample variance: $\nu \equiv \frac{1}{2}(s_1 - s_2)^2 = [s_1 - \frac{1}{2}(s_1 + s_2)]^2 + [s_2 - \frac{1}{2}(s_1 + s_2)]^2$

Appendix A.2 proves that the variation ν will always converge to a finite value, and

$$S = \frac{k}{2}(\beta\delta)^{\frac{1}{3}}(\sigma^2)^{\frac{2}{3}}, \tag{27}$$

where $k \equiv \left[3^{\frac{1}{3}} \text{Gamma}\left(\frac{4}{3}\right)\right]^{-1} \approx 0.78$ is a constant. Comparing g and S, we get the threshold condition:

Proposition 4. *Growth overcomes selection, if*

$$g > \frac{k}{2}(\beta\delta)^{\frac{1}{3}}(\sigma^2)^{\frac{2}{3}}. \tag{28}$$

Otherwise, selection dominates growth.

When $g < S$, $\mathbb{E}_{t\to+\infty}\left(\frac{d\mu}{dt}\right) < 0$. The possibility that μ could decline seems to contradict Malthusian stagnation. But in the real world, the decline of μ has a natural limit. Luxury consumption cannot decrease further when it reaches zero. In this sense, the case of $g < S$ already explains why the average luxury consumption was almost nil in ancient times. That said, if one feels uncomfortable with zero utility under a Cobb-Douglas utility function, a simple remedy is to assume, not unreasonably, that luxury productivity growth accelerates if average luxury is close to zero (demand is huge when luxury is rare). Then the equilibrium will have a positive amount of average luxury. The later simulations will use this method.

As Proposition 4 indicates, two sets of variables determine how strong selection is. The first is the variance of technological growth σ^2, which provides the necessary heterogeneity for selection to work on. The second is the product of two exogenous variables, $\beta\delta$. Denote $\lambda \equiv \beta\delta$, and call it the intensity of selection. In a richer setting, λ would further incorporate the migrants' willingness to move and the hosts' susceptibility to the migrants' influence. A little calibration can help gauge the relative strength of selection. If $\beta = 0.5$, $\delta = 0.1$ and $\sigma = 0.02$, the threshold $\hat{g}$ is 0.78‰. In actual history, the world population had been growing at about 1‰ per year between the Agricultural and Industrial Revolutions—$g_H \approx$ 1‰. Since appendix A.1 proves that g_A converges to $(1-\gamma)g_H$, if $\gamma = 0.5$, g_A is roughly 0.5‰. Therefore, if $g_B \leq g_A + \hat{g} =$ 1.28‰, or less than about 2.5 times the level of g_A, the global average living standards will not have an upward trend of growth.

It is worth noting that Assumption 4 allows technology to regress when "barbarians" invade. Though largely neglected in the growth literature, there were quite a few well-documented instances of technological regress in history. Some of them were apparently caused by invasions from less civilized groups. One example, which Aiyar, Dalgaard, and Moav (2008) provide in their study of technological regress, is a mortar called 'Pozzolana

cement' that the Romans used to construct large and durable structures such as baths, pantheons, and aqueducts. The technology was lost after the fall of the Roman Empire and was only relearned in the early 13$^{\text{th}}$ century. As the case of Pozzolana cement illustrates, luxury technologies are particularly vulnerable to social disorder following "barbarian invasions". Assumption 4 does capture part of reality.

However, there is no reason to expect peaceful immigration to bring about serious technological regress. While having lasting impacts on the host society, migrants tend to be gradually assimilated, and the mutual learning between natives and migrants often pushes forward the technological frontiers for both. Assumption 4 might characterize the consequences of invasions but it overstates the power of selection in the case of peaceful migrations.

The question is: is Assumption 4 crucial? In the next section, I will relax the assumption, allowing natives and migrants to learn from each other with no chance of technological regress. Even then, the group selection theory still holds. Selection is so strong that even if wars are ruled out, peaceful migration alone is enough to suppress the upward trend of living standards growth. However, I keep the assumption for the algebraic model; otherwise, the model would be intractable. As shown above, the assumption allows me to analytically derive the condition under which Malthusian stagnation arises. Qualitatively, the condition applies to peaceful migrations as well.

4.5 Simulations

The general equilibrium model has three limitations. First, the two-village setup may fail to capture the real-world intense competition among hundreds of countries. Second, the model assumes migration is free of cost and occurs instantly, so there is no difference of utility across the regions. Third, the model assumes that immigrants' technologies are able to displace natives' technologies, no matter whose technologies are better. As discussed above, the assumption is unreasonable for peaceful migrations.

To ensure robustness, I relax all of the three assumptions in this section. Because the algebraic models are no longer tractable, I turn to computer simulations. The simulations have a hundred regions instead of two. Migration is gradual and its speed increases with the utility gap across regions. Moreover, besides the baseline case where technologies are indiscriminately substituted, I also study the case where learning occurs only if the immigrants have a better technology (peaceful migrations with no wars).

4.5.1 The baseline simulation

Imagine a world composed of 100 sites, arrayed on a 10×10 grid. Each grid represents a region that has the same population dynamics and the same production and utility functions as the baseline model specifies. Time is discrete. At period t, the state of grid (i, j) is characterized by $\{A_{ijt}, B_{ijt}, H_{ijt}\}$, the subsistence technology, the luxury technology and the population size.

Assume A_{ij} and B_{ij} evolve the following way:

$$A_{ij}(t+1) = A_{ij}(t)(1 + g_{Aij} + \sigma_A \epsilon_{Aij}) + \text{selection effect} \quad (29)$$

$$B_{ij}(t+1) = B_{ij}(t)(1 + g_{Bij} + \sigma_B \epsilon_{Bij}) + \text{selection effect} \quad (30)$$

The error terms ϵ_A and ϵ_B have normal distributions, $\epsilon_A, \epsilon_B \sim N(0, 1)$, i.i.d. g_{Aij} is the same across all grids: $g_{Aij} = g_A$, but g_{Bij} increases with the relative rarity of luxury:

$$g_{Bij} = g_B \left[1 + \left(\frac{B_{ij}}{A_{ij}} \right)^{\alpha} \right]. \quad (31)$$

The additional term in the bracket is meant to prevent the downward trend. α is arbitrarily set to be a large negative number to minimize its impact when $B_{ij}/A_{ij} > 1$ ($\alpha = -10$). Though appearing *ad hoc*, adding the term increases the growth rate of luxury, which is unfavourable to my hypothesis and therefore makes the theory more robust.

At each period, residents of each grid decide whether they should move to a neighbouring grid for higher living standards. For two grids next to each other, if grid 1 has a higher utility than grid 2, some residents of grid 2 will move to grid 1, and the migration rate is proportional to the difference of utility:

$$\frac{\text{Migrants}}{\text{Population of the Origin}} = \theta(\ln U_1 - \ln U_2) \quad (32)$$

Unlike the previous model, here θ is finite.

I simulate two scenarios. In the first scenario, barbarian invasions and technological regress are allowed. The immigrants' technologies are assumed to displace the natives' technologies, no matter whose technologies are better. I call the scenario "indiscriminate substitution".

In the second scenario, people only learn from those who do better: if the immigrants are better at producing things, the natives will update their technology in the same way as "indiscriminate substitution"; but if the immigrants' technologies are inferior, the natives will keep their old ways to produce, and the immigrants will convert to the natives' technologies.

I call the scenario "selective learning". There is no chance of technological regress under "selective learning".

The force of selection is weaker under selective learning, yet it still favors the spread of subsistence technologies. To see this, suppose there are two identical regions. If one has a positive shock in subsistence productivity, people will emigrate to the other region to spread the improved subsistence technology. But if it is the luxury technology that has improved, no emigration will occur and the luxury technology has to remain local (contrast it with the indiscriminate substitution scenario where selection happens either way).

I simulate both scenarios. It turns out that, over a vast range of parameters, there is no upward trend of growth under either scenario. Selection is dominant over growth. To save computer time, I treat each simulated period as a decade. As Table 3 in appendix C summarizes, I parameterize $g_A = 0.5\%$ and $g_B = 1\%$ (per decade) for the baseline cases. The size of g_A guarantees a growth rate of population close to historical rates. g_B is twice as big as g_A, promising a strong tendency of luxury growth. The (subsistence) income elasticity of population, $\delta = 0.2$, matches the estimation from the English demographic and price data. The other crucial parameters include the standard deviation of the growth errors, $\sigma = 5\%$, and the migration propensity, $\theta = 0.1$. At a first approximation, $\theta = 0.1$ means that, if there is an opportunity to move to a twice richer place next to home, only 1% of people will take the opportunity in a typical year.

Figure 12(A) presents the key result of the simulations. It compares the global average utility, weighted by regional population[13], both with and without the indiscriminate-substitution type of selection. The Malthusian assumption alone fails to deliver the Malthusian result. Over ten thousand years, the global average utility increases about tenfold without migration (the result is the same if migration is allowed but migrants are assumed to carry no technologies). But when the knowledge-spreading migration is introduced, the trend is gone.

The absence of a growth trend under selection is not a result of technological stagnation. Rather, technologies grow faster when migrants are allowed to carry them around. This is how the "Malthusian + migration" case achieves a faster population growth than the purely Malthusian case in Figure 12(B).

What's more, the distribution of regional utility under biased migration is stationary (Figure 18 in appendix B). The richest region's utility never exceeds twice the poorest region's utility. Selection keeps all regions interlocked. In contrast, if there is only Malthusian

[13] The weighted global average utility is the "true" average that assigns equal weight to each person of the world. If I drop the weighting and use the average of regional utility instead, the path of global utility will only be more stable, as utility is negatively correlated with population. I will stick to the weighted average, which is unfavourable to my hypothesis, throughout all simulations.

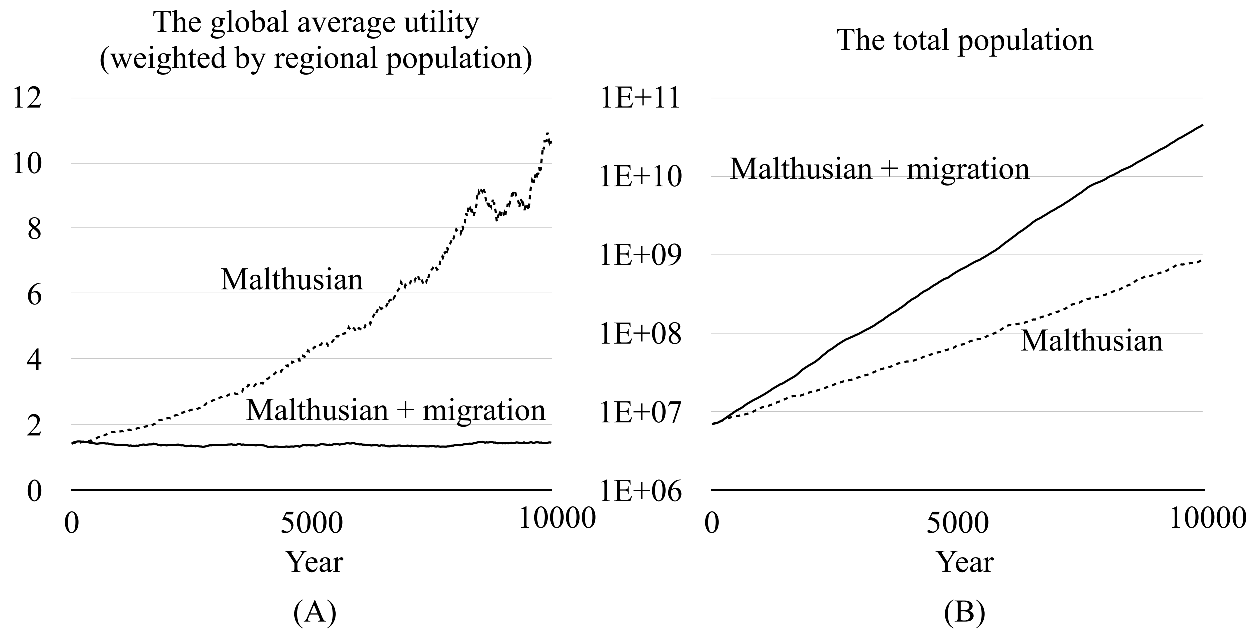

FIGURE 12: A purely Malthusian simulation does not produce the stagnation of living standards. To ensure stagnation, the Malthusian mechanism has to be combined with biased migration.

force but no selection, the variation is enormous and ever more divergent over time. Figure 17 in appendix B further traces the utility of three representative regions—one at the corner of the world, one on the side, and one in the middle. Despite cycles spaning thousands of years, there is no trend of growth in any single region.

Figure 13 shows the result under the selective learning scenario. Selection, weakened as it is, still dominates growth. The global average utility climbs up slowly before it stabilizes at a plateau. In the long run, there is no upward trend of growth either.

4.5.2 Robustness checks

In this section, I show that the dominance of selection over growth is robust to variation in three sets of parameters, including (a) the standard deviation of growth errors, σ, (b) the side length of the simulated world, w and l, and (c) the migration propensity, θ.

First, I vary σ from 0% to 15% with each step equal to 1%, and g_B from 0% to 2% with each step equal to 0.1%, keeping all the other parameters the same as in the baseline case. For each pair of σ and g_B, I run the simulations five times. I adopt a stringent criterion of stagnation. If the global average utility grows more than 25% from the 300$^{\text{th}}$ period to the 600$^{\text{th}}$ period—over a length of 3000 years—I treat it as a trend of growth, and if more than one simulations have a trend (excluding one), I mark the pair of parameters as "progressive";

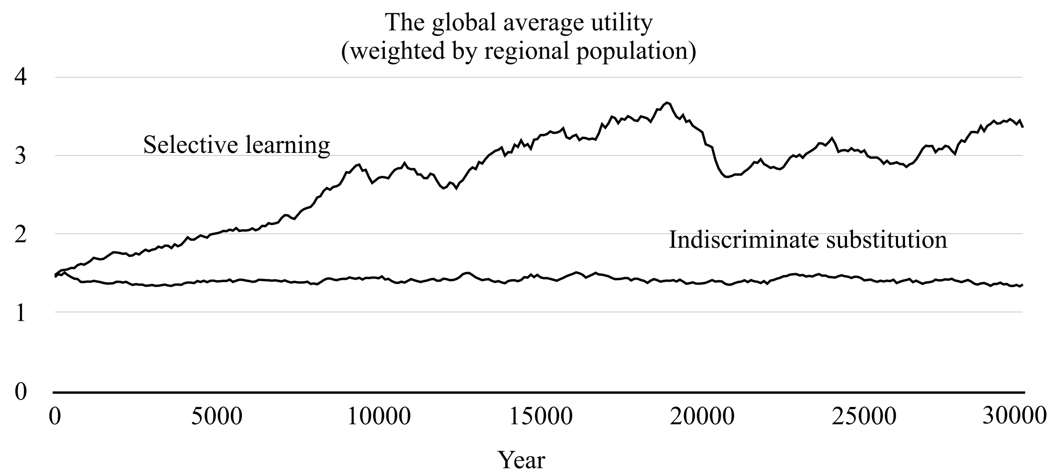

FIGURE 13: If learning is selective, the global average utility will stabilize at a higher level than if technology is indiscriminately substituted.

otherwise, "stagnant".

I conduct the robustness check for both scenarios of knowledge spread. The result is Figure 19 in the appendix. Selection gets stronger with a larger σ: under indiscriminate substitution, when $\sigma = 3\%$, selection dominates if $g_B - g_A \leq 0.6\%$; when $\sigma = 5\%$, selection dominates if $g_B - g_A \leq 1\%$. As expected, selection is weaker under selective learning, but there is a caveat. A simulated history treated as progressive does not necessarily have an upward trend in the long run. The path of selective learning in Figure 13 keeps rising until stabilized at about the 20,000$^{\text{th}}$ year. Applying the above criterion, I would treat the history as progressive but it actually has no trend in the long run.

To verify that the force of selection is robust to various sizes of the square world, I experiment with every integer value of side length from 3 to 20, running five simulations for each. Figure 20 in appendix B shows the cumulative growth from the 300$^{\text{th}}$ period to the 600$^{\text{th}}$ period of these experiments. The variation is bigger when the world is smaller, for the results are then more likely to be driven by the idiosyncrasies of individual grids. Nevertheless, there is hardly any difference between a world of 100 grids and a world of 400 grids. The baseline simulation, which assumes a 10×10 world, is representative in this respect.

The results are also robust to variation in θ. To verify this, I run 10 experiments under each scenario for each value of θ from 0 to 0.2 with the step equal to 0.01. The power of small θ's is extraordinary. The baseline simulation assumes $\theta = 0.1$. It is already a conservative estimate of people's willingness to move. But as Figure 14 shows, even if θ is as small as 0.01—only 0.1% of people would move each year to a neighbouring region that is twice as rich—selection still dominates. This by no means suggests that migration is unimportant.

If $\theta = 0$ (the pure Malthusian case), the cumulative growth is much larger than if $\theta = 0.01$. Growth disappears as θ slightly deviates from zero. A tiny bit of migration is strong enough to dominate a strong tendency of growth. Why this is the case I will leave to section 5.4 where I discuss the implications of the group selection model for the Industrial Revolution.

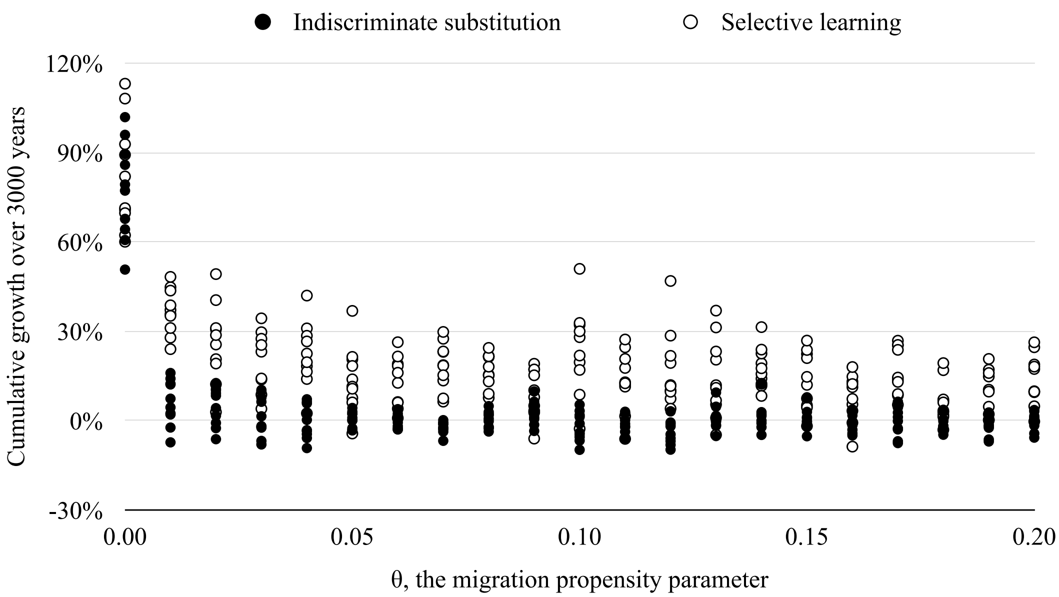

FIGURE 14: The cumulative growth of global average utility over 3000 years. Notes: each point represents an experiment at the corresponding level of migration propensity, θ.

4.5.3 Pointwise balance

Selection ensures the equality of long-run average growth rates between the sectors. But mere equality is not sufficient for the stagnation of living standards. World population growth had changed speed several times (Figure 15). Behind each change is the acceleration of subsistence technology growth. At these moments, for living standards to keep constant, the progress of luxury technology must accelerate to exactly the same speed—a pointwise balance.

To test the pointwise balance hypothesis, I fix g_B at 1%, and have g_A jump from 0.25% to 0.75% at the 1001$^{\text{st}}$ period of the simulation. If the pointwise balance exists, the global luxury technology growth will speed up to the same rate as the subsistence technology growth immediately after the juncture. This is exactly what happens in the simulation, as Figure 15 shows. I further conduct a Chow test:

$$\Delta \log(\text{luxury technology}) = \underset{(1e^{-3})}{5e^{-3}} + \underset{(0.6e^{-3})}{10e^{-3}} \times \text{break dummy}_{t=1001} + \epsilon \tag{33}$$

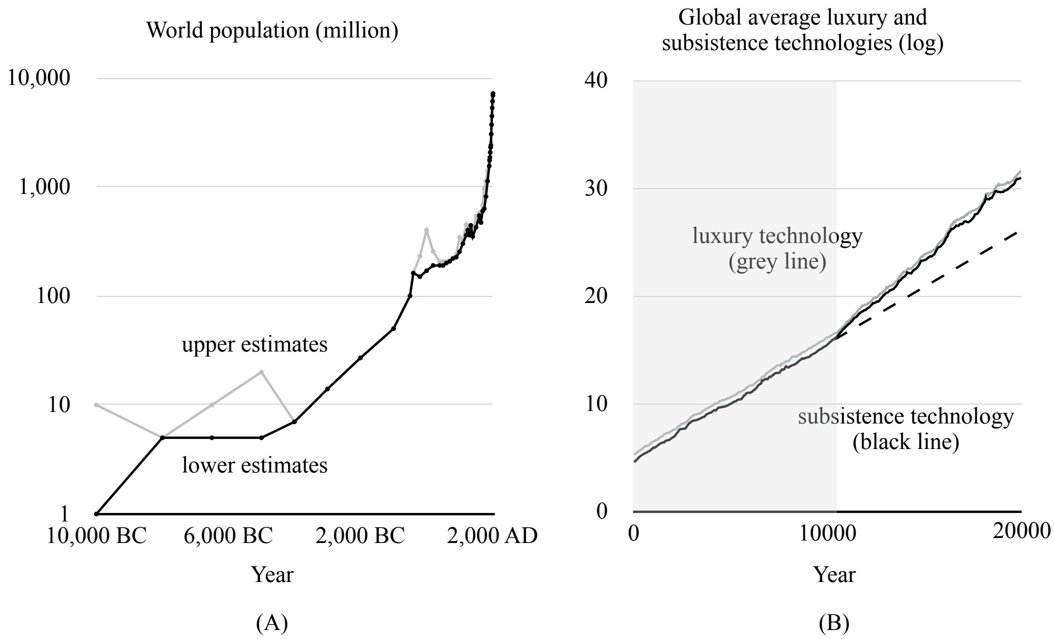

FIGURE 15: (A) The world historical population estimates. Data source: US census bureau. (B) When subsistence technology growth accelerates, luxury technology growth will accelerate to the same speed.

With a p-value as low as 10^{-6}, the test rejects the null hypothesis that there is no kink in luxury technology growth at the 1001st decade. The estimated coefficient of the break dummy, $10e^{-3}$, is exactly twice as large as the constant term, $5e^{-3}$. It means that when the growth rate of subsistence productivity triples from 0.25% to 0.75%, the growth rate of luxury technology triples from 0.25% to 0.75% too. Although g_B is fixed, luxury growth catches up fast and fully. Selection ensures balanced growth not only trend-wise but also point-wise.

5 Rethinking major events of economic history

The combination of the two-sector model and the group selection model paints a new picture of economic history. In what follows, I will discuss the implications of the theory for four issues, namely, the Agricultural Revolution, the ancient market economies, the welfare consequences of wars and migrations, and the Industrial Revolution.

5.1 Why farm?

The early farmers were worse off than their hunter-gatherer ancestors. They had less leisure, worse nutrition and greater inequality between sexes and across castes. The paradox of

immiserizing growth can be explained by the fact that agriculture is a subsistence technology. By tilting the production structure toward subsistence, it caused living standards to decline in the long run. Yet, if agriculture was so bad, "Why [did people] farm? Why work harder, for food less nutritious and a supply more capricious? Why invite famine, plague, pestilence and crowded living conditions (Harlan, 1975)?"

Farmers are faced with a prisoner's dilemma. People choosing what is best for themselves are hardly concerned with the prospect of the whole group's misery. That the farmers as a group would end up worse off would not bother one who saw agriculture as the dominant strategy to maximize her own chance of survival and reproduction. But even if there was a group of altruistic visionaries who coordinated to keep the hunting-gathering lifestyle, the group could not compete with one that had switched to agriculture. The latter was relatively richer in subsistence. The higher density of population, the consequent impoverished life, and the greed for new lands would drive the agricultural group to invade the hunting-gathering group instead of the other way around. Selection would wipe away whoever refused to farm (Cavalli-Sforza, Menozzi, and Piazza, 1993).

5.2 The rise and fall of the wealth of nations

First published in 1776, *The Wealth of Nations* declared the birth of modern economics. However, Gregory Clark (2008, chap.2, pg.35) commented, "[I]n 1776, when the Malthusian economy still governed human welfare in England, the calls of Adam Smith for restraint in government taxation and unproductive expenditure were largely pointless [... while] those scourges of failed modern states—war, violence, disorder, harvest failures, collapsed public infrastructures, bad sanitation—were the friends of mankind before 1800." Provocative as this sounds, it was the natural conclusion of the classical Malthusian theory. Without the two-sector model, there is no way to refute it.

In light of the two-sector Malthusian theory, Smith's prescriptions were sound, in ways that might not have been fully anticipated in the 18^{th} century. The policies suggested in *The Wealth of Nations* can improve living standards, not only in the short run but also in the long run, not only in Solow's time but also in Smith's time, and much earlier times as well. Laissez-faire, light taxation and the division of labor, if applied to economic policies, raise productivity in all sectors, but manufacturing and commerce benefit more than agriculture. The rise of the ratio of luxury to subsistence leads to higher equilibrium living standards.

This explains why the average Romans and Sung Chinese were richer than other peoples in history. According to Lo Cascio and Malanima (2009)'s estimation, the per capita GDP of Roman Italy reached $1400 in US 1990 dollars in 150 AD, and the per capita GDP of

the whole Roman Empire was as high as $1000. Among the many mentioned estimates, Temin (2013) regards this set of numbers as closest to reality. To put the estimates into perspective, consider that Maddison (2003) estimated the per capita GDP of most ancient societies at slightly above or around $450. $1400 per capita is what the Netherlands achieved as late as 1700. The reason why the Romans were rich is very similar to the reason why people living in modern developed countries are rich. As Temin (2013) shows, Rome had a functioning legal system, an active financial market, and a broad market network. The security of property rights stimulated investment; the scale of the market facilitated labor division; and standardised mass production improved the quality of consumer goods to a high level. None of these receive much attention in classical Malthusian theory, but in light of the new theory, they were as crucial to ancient living standards as they are to modern life.

On the contrary, the "friends of mankind"—wars, violence, disorder, collapsed infrastructures—often destroy more luxury than subsistence and decrease living standards in the long run.

5.3 A brief history of the long war against luxury

Fatal clashes on the group level have been a persistent human condition since primitive society. Of the fourteen groups studied in Mae Enga, a modern hunter-gatherer society in Papua New Guinea, five went extinct in tribal clashes over a 50-year period. In place of the extinct groups, new groups formed out of the old groups that survived and expanded (Soltis, Boyd, and Richerson, 1995). A group that spent too much on luxuries would hardly survive.

The domestication of animals and plants divided the world into nomadic zones and arable zones. Until the mass use of gunpowder, clashes between the two had disrupted growth over and over again. The three pre-modern peaks in Ian Morris (2011)'s social development index all ended in "barbarian" invasions. The sea peoples raided Anatolia, the Levant and Egypt; the Huns and Goths ruined the Western Roman Empire; the Jurchens and Mongols conquered the Sung Dynasty of China. A brief review of the three events can help us appreciate the crucial role migration plays in suppressing the trend of luxury growth.

Around 1000 BC, the sea peoples, arguably nomads from the hinterland of Europe, destroyed a number of highly developed kingdoms built by the Hittites, Minoans, and Mycenaeans. Urban centres, artistic representation, elaborate writing systems, and large-scale trading, shipping and construction vanished; civilizations were reduced to impoverished, illiterate, technically backward and violent small communities. The population of the largest cities in the West declined from 80,000 (Babylon and Thebes) in 1200 BC to 25,000 (Susa) in 1000 BC. "The invasions were not merely military operations, but involved the movements

of large populations, by land and sea, seeking new lands to settle" (Bryce, 1999).

The collapse of Rome was even more dramatic than the collapse of the Hittite kingdom. In the post-Roman Europe, production shrank to meet only local needs. Worldwide copper pollution, which approximates the production level, plummeted to a seventh of the Roman peak level (Hong et al., 1996). The post-Roman elites found it hard to afford the tiled roofs that once even the lowest class of Roman peasants had for their houses (Ward-Perkins, 2005). It is of course unfair to blame all of the loss and decline on the invaders—there was evidence of mild recession in the third and fourth centuries, arguably caused by civil wars and epidemics—but the invasions had certainly done most of the devastation.

This narrative diverges from the Malthusian version of Roman history. In that view, the average Roman lived beyond the verge of subsistence only because the Roman population had not caught up with technology for a short while; when it finally did, prosperity was gone (Temin, 2013). The problem is: if the Romans had lived a pinched life under population pressure, why would Emperor Valens have bothered to recruit armies from the "barbarian" immigrants, allowing the Gothic refugees from the Huns to reside within Roman territory in the first place? After the collapse of Rome and the decline of population that followed, why did average living standards not rise, as Malthusian theory predicts, but instead plunged? My answer is: the Romans were rich because their economic system encouraged luxury production and consumption; the post-Roman Europeans were poor because the new rulers—not so unlike bandits—adopted policies hurting commerce and industry. Europe later turned into a feudal society where obligation replaced profit as the guiding principle of economic life. In many parts of the continent, money transactions disappeared. Individuals' freedom gave way to group survival. It was then, by the contraction of commerce and industry, that Europe became a true "subsistence economy". The next time Europe recovered, it was a thousand years later when another round of commercial revolution began in the Italian cities.

The same catastrophe befell Sung China. Broadberry, Guan, and Li (2014) estimated that the per capita GDP of Sung was about $1500 in US 1990 dollars in the 11^{th} century. Manufacturing and commerce were so developed that they contributed two-thirds of the government's tax revenue (Liu, 2015). Sung's textile machinery was comparable with European designs in the eighteenth century. Its furnaces put out as much iron as the whole of Europe would produce in 1700. Sung's coal mines were large enough for hundreds of workers to work at the same time. However, while the combination of textile, iron and coal sent England onto the track of the Industrial Revolution, Sung failed prematurely.

Unlike England, Sung had few geographical barriers to protect itself from invasions. Its collapse is best viewed as one of several waves of group selection that surged in East Asia in

the 12th and 13th centuries. Before the Jurchens, Sung's rival was Liao, a country the nomadic Khitans built. After 25 years' war, Sung and Liao signed a peace treaty on the condition that Sung would pay an annual tribute to Liao. The peace lasted more than 120 years, bringing prosperity to both sides. Occupying part of China proper that included today's Beijing, Liao turned from a backward pasture economy into a civilized country that had a highly developed manufacturing sector. But the Khitans, now civilized, ended up an easy prey of the Jurchen barbarians. Two years after Liao fell, the Jurchens further conquered Kaifeng (Sung's capital) and annexed the northern half of China. A century later, the now-civilized Jurchens were in turn wiped out by the barbarian Mongols.

Though the Mongols inherited enormous wealth from the Jurchens and the Sung Chinese, they threw away many of the institutions and policies that had made the wealth possible. The Mongols divided subjects into four castes with institutionalized discrimination between them. For a cheap and stable supply of labor, the government forbade workers and their offspring from changing jobs. The system was later inherited by the Hong-Wu Emperor of the Ming dynasty, who concluded from the hyperinflation under the Mongols' rule that money was a dangerous thing, and that the best way to organise the economy was to fix people at preassigned places and jobs, discouraging movement of goods and people. The result was that Ming China became a predominantly agrarian economy. While Sung collected two-thirds of its tax from commerce and industry, agriculture provided 84% of Ming's tax revenue. Even so, Ming's total agricultural tax was still smaller than that of Sung China despite the former's larger territory and population. The living standards that the Sung Chinese once achieved were not reached again in China until perhaps Deng Xiaoping's reform.

Above, we have seen how invasions from less developed regions destroyed some of the greatest civilisations the world had ever seen. But that is only one aspect of wars' impact on luxury growth. In response to wars, groups often intentionally cut down on luxuries, and that might have been an even stronger force undermining luxury growth.

For example, during the Warring Period of China (476-221BC), restraint on luxury was the theme of a series of political and economic reforms[14]. In the face of constant nomadic harassment, King Wu-Ling of Zhao (340-295 BC) commanded his subordinates to take off their wide sleeves and long robes[15] and switch to nomadic uniform—pants, belts and boots—in order to fight as cavalry. Half a century before King Wu-Ling, *Shang Yang*'s reform swept another kingdom, Qin. The reformer punished commerce, rewarded cultivation, forbade migration and restricted entertainment. In a word, he cut down luxury and directed as

[14]To name a few, Li Hui's reform in Wei, Wu Qi's in Chu, Shen Buhai's in Han, Shang Yang's in Qin and King Wu-Ling's in Zhao.

[15]Veblen (1899) pointed out that the inconvenience in the clothing style reflects the need for conspicuous consumption.

many resources as possible to subsistence. The subjects were deprived, but the Qin kingdom defeated all of the six rival kingdoms and united China for the first time in history. A contemporary philosopher commented, "Qin is different from all the other kingdoms. The people are poor and the government is cruel. Whoever hopes for a better life can do nothing but fight hard. This makes the Qin army the strongest of all."[16]

Qin's idea of governance had a lasting impact on the later Chinese dynasties. Part of the influence is reflected in the mainstream of ancient Chinese economic thought, which emphasized restraints on luxury and commerce. The ancient thinkers thought differently than Adam Smith, not because they were blind to the benefits of commerce, but because they cared about the country's survival more than about the subjects' welfare. In a country where tax-fed mercenaries are not the backbone of military strength, the government had better sacrifice commercial gains to ensure ease of conscription.

Smith's intellectual legacy lasts because he lived on the eve of the modern era, when individual welfare was finally reconciled with group survival and expansion—it was the richer Europeans who had moved to America, Africa and India, instead of the other way around. Smith prophesied the day at dawn.

5.4 Luxury explosions and the Industrial Revolution

Thanks to the Industrial Revolution, we have escaped the Malthusian trap. Understanding why the Malthusian trap existed is basic to explaining how we escaped it. The classical Malthusian theory predicts instability when birth rates decrease with income. So, most researchers have used demographic transition (multiple equilibria) to make sense of the Industrial Revolution. Demographic transition is *not* incompatible with the two-sector model, and the group selection theory adds to the conventional interpretation by pointing to a new set of factors triggering modern economic growth (Table 2).

The first factor is trade. Trade substitutes for migration. A decline in the cost of trade, combined with a rise in the cost of migration (political barriers to migration increased in the modern era), can slow down the selection that draws living standards downward.

The second factor is books. It is crucial to the group selection theory that migration should be a major channel to spread knowledge. The assumption no longer held after printing presses spread.

The next two factors appear in the threshold condition of growth, $g > \frac{k}{2}(\beta\delta)^{\frac{1}{3}}(\sigma^2)^{\frac{2}{3}}$. From 1500 to 1800, Northwest Europe experienced a steady decline in the price of luxuries relative to staple food and fuels (Hoffman et al., 2002). This implies that the gap in growth

[16] *Xun Zi*, chapter *Yi Bing* (On Wars).

TABLE 2: The "revolutionary" factors in the old and new theories

Model feature	Triggering event	Result
Classical Malthusian theory		
Birth rates decrease with income	Income rises	Switch to the higher equilibrium
Group selection theory		
Trade replaces migration	Trade cost drops	Selection slows down
Migrants spread knowledge	Printing Tech. develops	Tech. & migration disentangled
Threshold: $g > \frac{k}{2}(\beta\delta)^{\frac{1}{3}}(\sigma^2)^{\frac{2}{3}}$	g increases	Balance is tipped
Threshold: $g > \frac{k}{2}(\beta\delta)^{\frac{1}{3}}(\sigma^2)^{\frac{2}{3}}$	σ decreases	Growth dominates selection
Literacy was a luxury	It becomes subsistence	Literacy spurs growth

rates, g, had become larger. What is more, Fouquet and Broadberry (2015) shows that the variance of GDP growth rates of European economies decreased in the 19$^{\text{th}}$ century. The increase in g and the decrease in σ^2 might reverse the inequality relationship.

The last factor may be regarded as a theory by itself, the luxury explosion theory. The theory holds that a technology—akin to culture—that turns from a luxury technology into a subsistence technology will spread in an explosive way. To see this, go back to the group selection model. With a little abuse of notation, denote a negative selection, the case where a trait is selected against (luxury), as $\lambda < 0$, and a positive selection, the case where a trait is selected for (subsistence), as $\lambda > 0$. Following a similar derivation as in section 4.4.2, the relationship between the force of selection and the intensity of selection is still

$$S = \Phi\lambda^{\frac{1}{3}}, \tag{34}$$

except that the sign of λ now indicates whether the trait is selected for or selected against.

As Figure 16 shows, the "S" shape of $S(\lambda)$ means that even a tiny λ can produce a large force of selection. The "S" shape results from a subtle *variation effect*: when selection is less intense—because, say, people are more reluctant to move—regions tend to deviate further away from each other in terms of production structure and the level of utility. The increased gap of utility motivates people to move notwithstanding their inertia; and the enlarged difference of lifestyle means migrants have more surprises to offer to the host culture. Overall, the greater variation compensates for the loss of interest in migration. It makes weak selection still have a strong impact[17].

[17] A similar mechanism can explain why peoples who rarely intermarry can still be genetically close. Pinker (2003, p.143) notes that "Rare genes can offer immunity to endemic diseases, so they get sucked into one group from a neighbouring group like ink on a blotter, even if members of one group mate with members of the other infrequently. That is why Jews, for example, tend to be genetically similar to their non-Jewish neighbours all over the world, even though until recently they tended to marry other Jews. As little as one

The derivative of $S(\lambda)$ is infinite at 0. When environmental changes make a luxury trait less luxurious, little will change if the trait remains a luxury. But if the trait thereby turns into subsistence, even if the environmental change is extremely tiny, this process will trigger a big change in S—a luxury explosion.

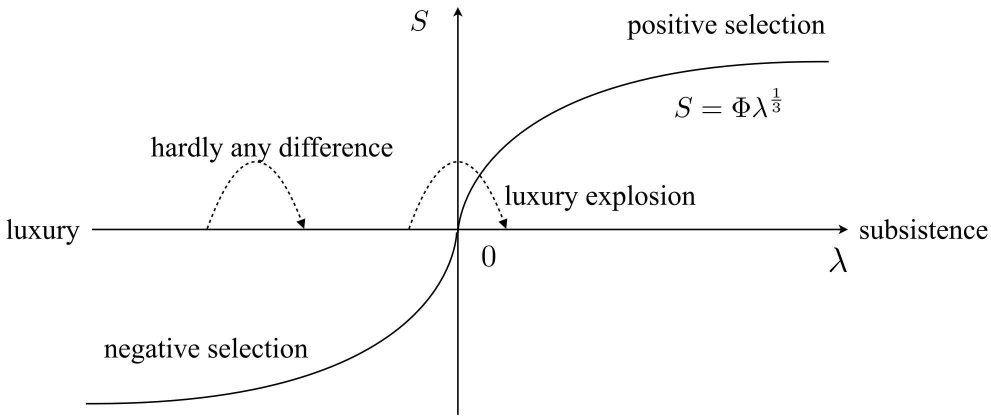

FIGURE 16: The relationship between the force of selection S and the intensity of selection λ. Here $\lambda < 0$ means the commodity is luxury, and it is subject to negative selection; $\lambda > 0$ means the commodity is subsistence, and it is subject to positive selection.

The luxury explosion has profound implications for triggering modern growth. Consider the relationship between literacy and fertility. The tradeoff between quantity and quality of children is a classic example of the choice between subsistence and luxury. Spending money on books and education for children—a choice of quality over quantity—might increase the number of grandchildren (Galor and Klemp, 2014), but, if all households do so, the density of population will decline in most cases. The quality of children is the pivot of transition in a number of unified growth theories (Galor and Moav, 2002; Galor and Weil, 2000; Clark, 2008; Galor, 2011). Transitions in these models are mostly driven by multiple equilibria of fertility choice. Here, selection provides a new mechanism of transition: literacy, which was meant for reading the Bible at the onset of the Religious Reformation, unintentionally equipped the masses with scientific knowledge, engineering knowhow and nationalist enthusiasm, by which Europe colonized the other parts of the world. What used to be a luxury turned into subsistence. The luxury explosion then made a revolution.

The Industrial Revolution is unique because human capital is a very special luxury. Most other luxuries, like diamonds and yachts, cannot switch into subsistence. Weapons switch but they do not improve living standards. Human capital not only enriches the individual but also strengthens a country. Moreover, human capital is immune to the "erosion" of

conversion, affair, or rape involving a gentile in every generation can be enough to blur genetic boundaries over time."

biased migration. If an immigrant wants to benefit from the "luxury", she has to learn it, whereas a learned emigrant can apply his knowledge away from home. Hardly any other luxury combines these wonderful features.

There have been plenty of explanations for the Industrial Revolution, some of them anecdotal. Indeed, it is unlikely that any single factor can account for the whole transition experience. The idea that has received the most rigorous modeling is demographic transition combined with multiple equilibria. The demographic explanation is rooted in the widely held presumption that the Malthusian trap is caused by the Malthusian mechanism. Questioning that presumption opens up new interpretations of the Industrial Revolution and allows rigorous modeling of a new mechanism that considers institutions, trade, social insurance, the Renaissance, the Reformation, the Scientific Revolution, and the Enlightenment (Mokyr, 2005).

6 Concluding remarks

Differing from the Malthusian view of history, this paper suggests the following basic story. Imagine a world where people live on two things: bread and flowers. Population increases with bread, hence the average consumption of bread is fixed in the long run by the Malthusian force. But population hardly responds to flowers. If the flower sector grows faster than the bread sector, people will live better and better by each having more and more flowers. Such had never happened until the Industrial Revolution. Throughout the thousands of years before that time, flower productivity had somehow grown at the same rate as bread productivity.

The cause of the balanced growth is group selection. People organize themselves into competing groups. When a group has comparative advantage at making bread, its average member will have fewer flowers than their neighbours do. Greed drives them to move abroad. As they move, they spread the technology of their hometown to other places. The consequence is that the bread technology tends to spread faster than the flower technology. Even if the flower sector intrinsically grows faster than the bread sector, a tiny bit of spread advantage of the bread sector can offset a large growth advantage of the flower sector. With the whole world interlocked in a network of migration and occasional conquests, living standards were stagnant almost everywhere. Thus the Malthusian trap was, at the same time, a Darwinian trap.

For all its novelty, the group selection theory of the Malthusian trap is a tautology, a tautology that makes the theory robust. Here is how: what the theory sets out to explain

is why the average pre-industrial person had so little luxury[18]—so few flowers. The theory ascribes this Malthusian fact to group competition. By definition, luxury contributes to individual utility at the expense of group fitness. Fitness matters only in the corresponding context of competition. Therefore, luxury must be constrained by group competition, the only context in which group fitness ever matters. So, the way I define luxury has already ensured that group competition is the main suppressor of luxury.

Overall, this paper has four contributions:

I It revises the traditional explanation of the Malthusian trap.

II It explains why the Malthusian relationship between average income and population growth is empirically weak: the classical theory misses two of the three determinants of long-run equilibrium, i.e., social preference and production structure.

III It explains the prosperity of ancient market economies such as Rome and Sung.

IV It suggests a new set of factors that might have triggered modern economic growth.

Malthusian theory is fundamental to our understanding of history. Revising it opens numerous possibilities for economic history research. Here, I discuss two of them.

First, the two-sector model allows living standards researchers to move beyond the Malthusian presumption. The presumption is evident in Maddison's series, where both Rome's and Sung's per capita GDP were estimated at $450, that is, only $50 above the lowest number in the data. The presumption is also evident in much empirical research that omits non-agricultural output when estimating income, based on the assumption, as Baumol (1990) put it, "[i]n a period in which agriculture probably occupied some 90 percent of the population, the expansion of industry [...] could not by itself have created a major upheaval in living standards." The presumption is even evident in many skeptics' work. Often, researchers provide strong evidence of high living standards in certain historical episodes yet conclude that the prosperity must be a temporary phenomenon that is doomed to disappear when population catches up with technology. Temin (2013, p.193), for example, said, "[i]t reveals even Malthusian economies can have economic growth, that is, can have rising standards of living. This can go on for a long time, even centuries, *even though without industrialization, it is doomed to end*." Now, with the two-sector model available, large swings of living standards becomes a serious theoretical possibility. Hopefully, the two-sector model will also direct more researchers' attention to the non-agricultural sectors. In the classical

[18] Daily calorie intake per person has hardly changed since the Industrial Revolution. The improvement of life is mostly reflected in the diversity and abundance of luxury.

model, commerce and industry are unimportant for living standards; but in the two-sector model, they become crucial.

Second, the group selection theory calls on the profession to embrace a "macro" view of the Industrial Revolution. What I mean is this: the classical Malthusian theory has led most studies to focus on demographic transition as the key mechanism of the Industrial Revolution. This led to a research focus on changes of fertility behaviour in pre-modern Europe—how fertility varied with income, status, education, etc. Fertility is a micro decision, made on the household level. Fertility is important. But it is only one of the three comparative statics in the two-sector model. The other two, social preference and production structure, are no less important than fertility. What determines these two? They are determined by politics, policies, institutions, wars, trade, migration, cultures, and geography. Incorporating social preference and production structure into the analysis means we need to rethink the roles these "macro" factors play in a society. When it comes to the Industrial Revolution, previous researchers asked why households changed their minds about children. Now the added question is: what had the prince done that made his country stand out?

Finally, I would like to further address the difference between Dutta, Levine, Papageorge, and Wu (2014) (DLPW) and this paper. Most related studies treat the Malthusian trap as a fact, but DLPW is an exception. DLPW argues that, because manufacturing and commerce usually grow faster than agriculture, income per capita had a slow yet still significant trend of growth before the Industrial Revolution.

As stated previously, the two-sector Malthusian theory leads to two possibilities. One is that the Malthusian fact is right, but requires a new explanation; the other is that the Malthusian trap did not exist at all. This paper explores the first possibility, while DLPW studies the second. The reality must lie in between. The two theories can actually be reconciled. I will leave the details of the reconciliation to another paper. Here, I only sketch the idea.

Many luxuries are culture-specific. They are desired within a culture but not without. Group selection has no way to eliminate the growth of such luxuries because migration never responds to the difference of consumption in these items. Therefore, a distinction should be made between "universal luxury" and "provincial luxury". Universal luxuries are desired by all human beings; provincial luxuries, only by a group of people. Group selection suppresses universal luxuries, but leaves provincial luxuries free to grow. This explains why culture was so diverse across pre-industrial societies despite the monotony of life (measured by universal luxury). This is also why most economic historians accept the Malthusian fact while DLPW comes to a different conclusion. The literature has focused on universal luxury only, but DLPW is concerned with all types of luxuries. The current paper explains the

Malthusian trap in its usual sense, that is, why the average consumption of universal luxury was constantly low throughout the pre-industrial era.

According to a famous anecdote in the history of science, Darwin and Wallace independently discovered natural selection, both by reading Malthus's essay on population[19]. While Malthus inspired Darwin, Darwin in turn contributed a piece to solving Malthus's puzzle.

[19] In his autobiography (1876), Charles Darwin wrote: "In October 1838, that is, fifteen months after I had begun my systematic inquiry, I happened to read for amusement Malthus on Population, and being well prepared to appreciate the struggle for existence which everywhere goes on from long-continued observation of the habits of animals and plants, it at once struck me that under these circumstances favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The results of this would be the formation of a new species. Here, then I had at last got a theory by which to work."

References

Aiyar, Shekhar, Carl-Johan Dalgaard, and Omer Moav. 2008. "Technological Progress and Regress in Pre-Industrial Times." *Journal Of Economic Growth* 13 (2):125–144.

Allen, Robert C. 2008. "A Review of Gregory Clark's a Farewell to Alms: A Brief Economic History of the World." *Journal of Economic Literature* 46 (4):946–973.

Álvarez-Nogal, Carlos and Leandro Prados De La Escosura. 2013. "The Rise and Fall of Spain (1270–1850)." *The Economic History Review* 66 (1):1–37.

Ashraf, Quamrul and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *The American Economic Review* :2003–2041.

Baumol, William J. 1990. "Entrepreneurship: Productive, Unproductive, and Destructive." *Journal of Political Economy* :893–921.

Bowles, Samuel. 2006. "Group Competition, Reproductive Leveling, and the Evolution of Human Altruism." *Science* 314 (5805):1569–1572.

Bowles, Samuel and Herbert Gintis. 2002. "Behavioural Science: Homo Reciprocans." *Nature* 415 (6868):125–128.

Broadberry, Stephen, Hanhui Guan, and David D. Li. 2014. "China, Europe and the Great Divergence: A Study in Historical National Accounting, 950-1850." Economic History Department Paper, London School of Economics.

Broadberry, Stephen and Bishnupriya Gupta. 2006. "The Early Modern Great Divergence: Wages, Prices and Economic Development in Europe and Asia, 1500–1801." *The Economic History Review* 59 (1):2–31.

Bryce, Trevor. 1999. *The Kingdom of the Hittites*. Oxford University Press.

Cavalli-Sforza, Luigi L, Paolo Menozzi, and Alberto Piazza. 1993. "Demic Expansions and Human Evolution." *Science* 259 (5095):639–646.

Chen, Shuo and James Kai-sing Kung. 2013. "Of Maize and Men: The Effect of a New World Crop on Population and Economic Growth in China." *Available at SSRN 2102295* .

Clark, Gregory. 2008. *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press.

Davies, John E. 1994. "Giffen Goods, the Survival Imperative, and the Irish Potato Culture." *Journal of Political Economy* :547–565.

De Fraja, Gianni. 2009. "The Origin of Utility: Sexual Selection and Conspicuous Consumption." *Journal of Economic Behavior and Organization* 72 (1):51–69.

De Vries, Jan. 2006. *European Urbanisation, 1500-1800*, vol. 4. Routledge.

Diamond, Jared. 1987. "The worst mistake in the history of the human race." *Discover* 8 (5):64–66.

Dutta, Rohan, David K Levine, Nicholas W Papageorge, and Lemin Wu. 2014. "Entertaining Malthus: Bread, Circuses and Economic Growth." Tech. rep., UCLA Department Of Economics.

Fouquet, Roger and Stephen Broadberry. 2015. "Seven centuries of European economic growth and decline." *Journal of Economic Perspectives* 29 (4).

Friedman, Milton. 1953. "The Methodology of Positive Economics." *Essays In Positive Economics* 3 (8).

Galor, Oded. 2011. *Unified Growth Theory.* Princeton University Press.

Galor, Oded and Marc Klemp. 2014. "The Biocultural Origins of Human Capital Formation." Tech. rep., University Library of Munich, Germany.

Galor, Oded and Omer Moav. 2002. "Natural Selection and the Origin of Economic Growth." *The Quarterly Journal of Economics* 117 (4):1133–1191.

Galor, Oded and David N Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review* 90 (4):806–828.

Hansen, Gary D and Edward C Prescott. 2002. "Malthus to Solow." *American Economic Review* :1205–1217.

Harlan, Jack Rodney. 1975. *Crops and Man.* American Society of Agronomy.

Hersh, Jonathan and Joachim Voth. 2009. "Sweet Diversity: Colonial Goods and the Rise of European Living Standards after 1492." .

Hoffman, Philip T, David S Jacks, Patricia A Levin, and Peter H Lindert. 2002. "Real Inequality in Europe since 1500." *The Journal of Economic History* 62 (02):322–355.

Hong, Sungmin, Jean-Pierre Candelone, Clair C Patterson, and Claude F Boutron. 1996. “History of Ancient Copper Smelting Pollution during Roman and Medieval Times Recorded in Greenland Ice.” *Science* 272 (5259):246–249.

Jones, Charles I. 2001. “Was an Industrial Revolution Inevitable? Economic Growth over the Very Long Run.” *Advances In Macroeconomics* 1 (2).

Jones, Eric. 2003. *The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia*. Cambridge University Press.

Jones, Eric Lionel. 1988. *Growth Recurring: Economic Change in World History*. University of Michigan Press.

Lee, Ronald. 1987. “Population Dynamics of Humans and Other Animals.” *Demography* 24 (4):443–465.

Levine, David and Salvatore Modica. 2013. “Anti-Malthus: Conflict and the Evolution of Societies.” *Research In Economics* .

Lipsey, Richard G, Kenneth I Carlaw, and Clifford T Bekar. 2005. *Economic Transformations: General Purpose Technologies and Long-Term Economic Growth*. Oxford University Press.

Liu, William Guanglin. 2015. “The Making of a Fiscal State in Song China, 960–1279.” *Economic History Review* 1 (68):48–78.

Lo Cascio, Elio and Paolo Malanima. 2009. “Ancient and Pre-Modern Economies: GDP in the Roman Empire and Early Modern Europe.” Presented at the conference on “Long-Term Quantification in Mediterranean Ancient History," Brussels.

Maddison, Angus. 2003. *The World Economy: Historical Statistics*. OECD Publishing.

Malanima, Paolo. 2011. “The Long Decline of a Leading Economy: GDP in Central and Northern Italy, 1300-1913.” *European Review of Economic History* 15 (02):169–219.

Malthus, Thomas Robert. 1809. *An Essay on the Principle of Population, as It Affects the Future Improvement of Society*, vol. 2.

Moav, Omer and Zvika Neeman. 2008. “Conspicuous Consumption, Human Capital, and Poverty.” *Human Capital and Poverty* .

Mokyr, Joel. 1983. *Why Ireland Starved: An Analytical and Quantitative Study of Irish Poverty, 1800–1851*. London and Boston: George Allen and Unwin.

———. 2005. "The intellectual origins of modern economic growth." *The Journal of Economic History* 65 (02):285–351.

Morris, Ian. 2011. *Why the West Rules - For Now: The Patterns of History, and What They Reveal about the Future*. Random House LLC.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German-Jewish Emigrés and US Invention." Tech. rep., National Bureau of Economic Research.

Nunn, Nathan and Nancy Qian. 2011. "The Potato's Contribution to Population and Urbanization: Evidence From a Historical Experiment." *The Quarterly Journal of Economics* 126 (2):593–650.

Persson, Karl Gunnar. 2010. "The End of the Malthusian Stagnation Thesis." Tech. rep., Mimeo (University Of Copenhagen).

Pinker, Steven. 2003. *The Blank Slate: the Modern Denial of Human Nature*. Penguin.

Pomeranz, Kenneth. 2009. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton University Press.

Ravenstein, Ernest George. 1885. "The Laws of Migration." *Journal of the Statistical Society of London* :167–235.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis." *Journal of Monetary Economics* 55 (2):234–250.

Robbins, Lionel. 1998. *A History of Economic Thought: the LSE Lectures*. Wiley Online Library.

Simon, Julian Lincoln and Gunter Steinmann. 1991. "Population Growth, Farmland, and the Long-Run Standard of Living." *Journal of Population Economics* 4 (1):37–51.

Smith, Adam. 1887. *An Inquiry into the Nature and Causes of the Wealth of Nations*. T. Nelson and Sons.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *The Quarterly Journal of Economics* 70 (1):65–94.

Solow, Robert M and Paul A Samuelson. 1953. "Balanced Growth under Constant Returns to Scale." *Econometrica* :412–424.

Soltis, Joseph, Robert Boyd, and Peter J Richerson. 1995. "Can Group-Functional Behaviors Evolve by Cultural Group Selection? An Empirical Test." *Current Anthropology* :473–494.

Strulik, Holger and Jacob Weisdorf. 2008. "Population, Food, and Knowledge: A Simple Unified Growth Theory." *Journal of Economic Growth* 13 (3):195–216.

Taylor, M Scott and James A Brander. 1998. "The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource Use." *The American Economic Review* 88 (1):119–138.

Temin, Peter. 2013. *The Roman Market Economy*. Princeton University Press.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *The Journal of Political Economy* :416–424.

Veblen, Thorstein. 1899. *The Theory of the Leisure Class: an Economic Study of Institutions*. Macmillan.

Voigtländer, Nico and Hans-Joachim Voth. 2013. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe." *The Review of Economic Studies* 80 (2):774–811.

Vollrath, Dietrich. 2011. "The Agricultural Basis of Comparative Development." *Journal of Economic Growth* 16 (4):343–370.

Ward-Perkins, Bryan. 2005. *The Fall of Rome: and the End of Civilization*. Oxford University Press.

Weisdorf, Jacob L. 2008. "Malthus Revisited: Fertility Decision Making Based on Quasi-Linear Preferences." *Economics Letters* 99 (1):127–130.

Wilson, David Sloan. 2015. *Does altruism exist?: culture, genes, and the welfare of others*. Yale University Press.

Wu, Lemin. 2012. "Does Malthus Really Explain the Constancy of Living Standards?" *Available at SSRN 2187113* .

Yang, Dennis Tao and Xiaodong Zhu. 2013. "Modernization of Agriculture and Long-Term Growth." *Journal of Monetary Economics* 60 (3):367–382.

Appendices

A Proofs

A.1 Prove that g converges to $\beta(g_B - g_A)$ in the long run.

First, I prove the following lemma.

Lemma 5. *If an isolated economy has constant growth rates of technology g_A and g_B, then $g_A - (1-\gamma)g_H$ converges to 0.*

Proof:

Population evolves in the following way:

$$g_H = \delta(\ln x - \ln \bar{x})$$

Because $x = A(1-\beta)^\gamma H^{\gamma-1}$ (Equation 10),

$$g_H = \delta[\ln A + \gamma \ln(1-\beta) + (\gamma - 1)\ln H - \ln \bar{x}]$$

Denote $M \equiv \ln A + (\gamma - 1)\ln H$, then

$$g_H = \delta[M + \gamma \ln(1-\beta) - \ln \bar{x}]$$

The motion of M follows

$$\begin{aligned} dM &= g_A + (\gamma-1)g_H \\ &= g_A + (\gamma-1)\delta[M + \gamma\ln(1-\beta) - \ln\bar{x}] \end{aligned}$$

Because $(\gamma - 1)\delta < 0$, M will stabilize at

$$M^* = \frac{g_A}{(1-\gamma)\delta)} - \gamma \ln(1-\beta) + \ln \bar{x}$$

Hence $dM = g_A - (1-\gamma)g_H$ converges to 0. □

Proposition 6. *g_U converges to $\beta(g_B - g_A)$.*

Proof: Start by expressing U as a function of A and B. We cannot use the formula of equilibrium utility (Equation 14) because the continuous progress of technology will pull the

economy slightly away from the equilibrium state. So I turn to Equation 12, which applies to the dynamic scenario as well.

Suppose land is fixed. By log-linearizing Equation 12, we get

$$g_U = \beta(g_B - g_A) + g_A - (1-\gamma)g_H.$$

Lemma 5 holds that $g_A - (1-\gamma)g_H$ converges to 0. Therefore, g_U converges to $\beta(g_B - g_A)$. □

A.2 Prove that $S \equiv \beta\delta\,\mathbb{E}_{t\to+\infty}(\nu) = \frac{k}{2}(\beta\delta)^{\frac{1}{3}}(\sigma^2)^{\frac{2}{3}}$.

Proof: By Ito's lemma,

$$d\nu^x = \left[\sigma^2 x(2x-1)\nu^{x-1} - 2\sqrt{2}|\lambda|x\nu^{x+\frac{1}{2}}\right]dt + 2\sigma x\nu^{x-\frac{1}{2}}dz$$

Because $\mathbb{E}_{t\to+\infty}(d\nu^x) \to 0$, the long-run expectation of the drift term

$$\sigma^2 x(2x-1)\mathop{\mathbb{E}}_{t\to+\infty}(\nu^{x-1}) - 2\sqrt{2}|\lambda|x\mathop{\mathbb{E}}_{t\to+\infty}(\nu^{x+\frac{1}{2}}) = 0.$$

Let $f(x) \equiv \mathbb{E}_{t\to+\infty}(\nu^x)$ and denote $\frac{\sigma^2}{2\sqrt{2}|\lambda|}$ as a, then the above equation can be rewritten as a general term formula:

$$f\left(x+\frac{3}{2}\right) = a(2x+1)f(x)$$

with $f(0) = \mathbb{E}_{t\to+\infty}(\nu^0) = 1$.

The general solution is

$$f(x) = \frac{1}{3}(3a)^{\frac{2}{3}x}\,\text{Pochhammer}\left(\frac{4}{3}, \frac{2}{3}x - 1\right).$$

Let $x = 1$, then

$$f(1) = \frac{a^{\frac{2}{3}}}{3^{\frac{1}{3}}\,\text{Gamma}\left(\frac{4}{3}\right)}.$$

Denote $k \equiv \left[3^{\frac{1}{3}}\,\text{Gamma}\left(\frac{4}{3}\right)\right]^{-1} \approx 0.78$. Then $f(1)$ can be written as $ka^{\frac{2}{3}}$.

By definition,

$$\mathop{\mathbb{E}}_{t\to+\infty}(\nu) = f(1) = k\left[\frac{\sigma^2}{2\sqrt{2}|\lambda|}\right]^{\frac{2}{3}} = k\nu^*$$

Substituting $\mathbb{E}_{t\to+\infty}(\nu) = k\nu^*$ into $S \equiv \beta\delta\,\mathbb{E}_{t\to+\infty}(\nu)$, we get

$$S = \lambda k\nu^* = \frac{k}{2}(\beta\delta)^{\frac{1}{3}}(\sigma^2)^{\frac{2}{3}}.$$

□

B Figures

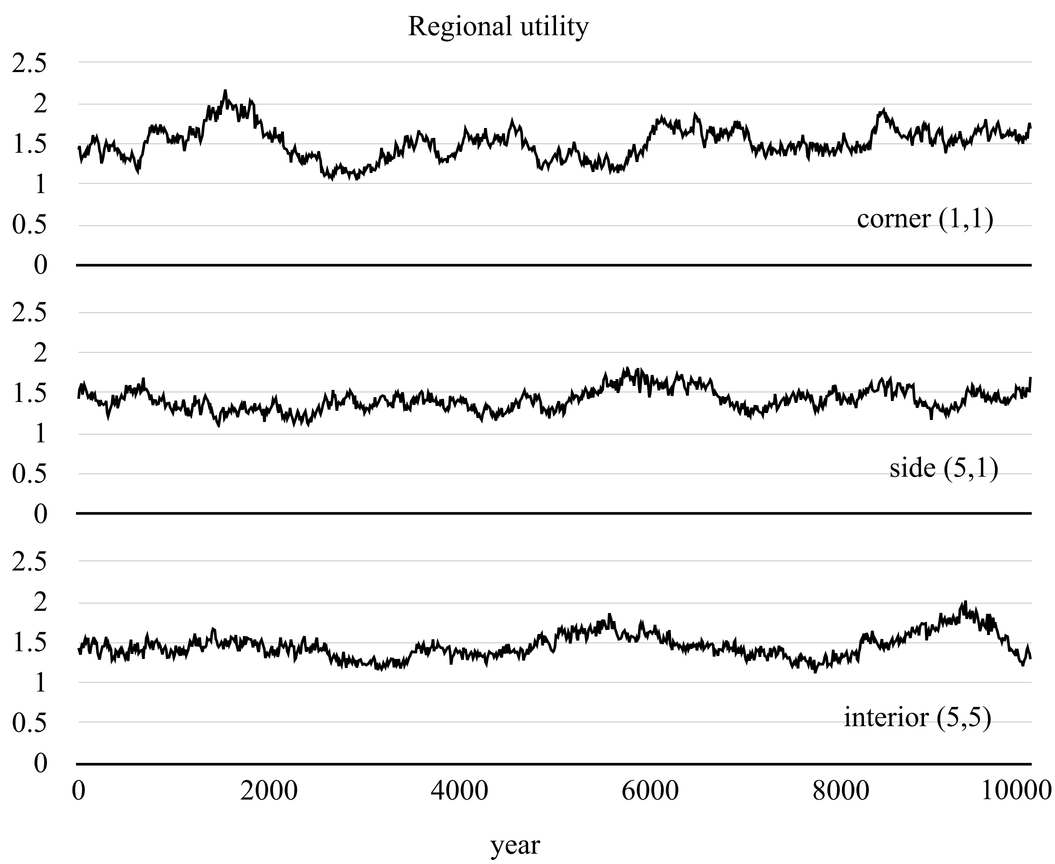

FIGURE 17: The regional utility fluctuates wildly but has no trend. Here is the history of regional utility of three representative regions: a corner region $(1,1)$, a side region $(5,1)$, and an interior region $(5,5)$.

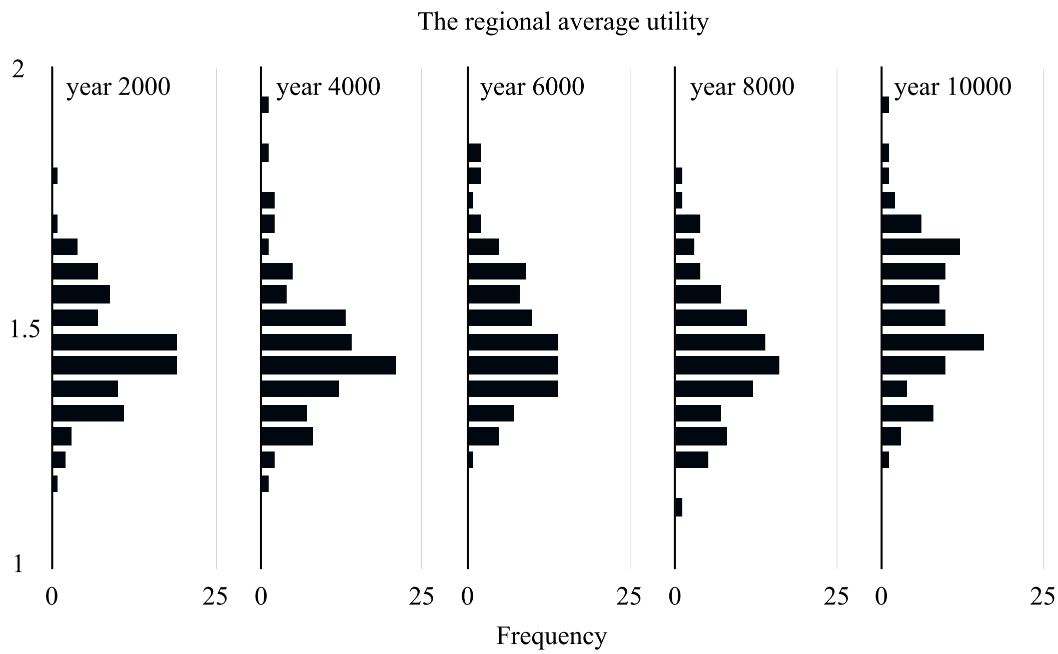

FIGURE 18: The distribution of regional average utility is stable over time.

C Tables

TABLE 3: Parameterization of the baseline simulation

Parameter	Value	Interpretation
g_A	0.5%	Subsistence growth rate
g_B	1%	Luxury growth rate
σ_A	5%	Std. of subsistence growth
σ_B	5%	Std. of luxury growth
δ	0.2	$n = \delta(\ln x - \ln \bar{x})$
γ	0.5	$X = AL_A^{1-\gamma}H_A^{\gamma}$, $Y = BL_B^{1-\gamma}H_B^{\gamma}$
$\bar{x}$	1	$n = \delta(\ln x - \ln \bar{x})$
θ	0.1	migrational rate
β	0.5	$U = x^{1-\beta}y^{\beta}$
α	-10	$g_{Bij} = g_B[1 + (B_{ij}/A_{ij})^{\alpha}]$

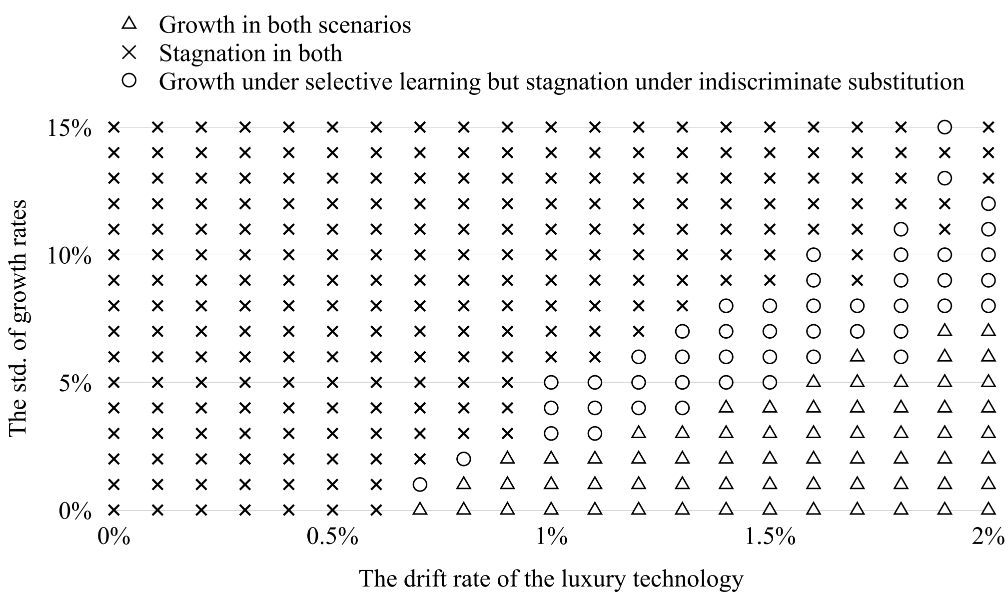

Figure 19: The progressive and stagnant areas on the parameter space. Notes: a point is counted as a "growth point" only if the global average utility grows more than 25% over 3000 years. There are three areas in the parameter space. The upper left area that is marked with crosses is where the global average utility stagnates in both the indiscriminate substitution case and the selective learning case. The middle area marked with circles is where the utility grows under selective learning but not under indiscriminate substitution. The lower right area marked with triangles is where growth occurs under both scenarios.

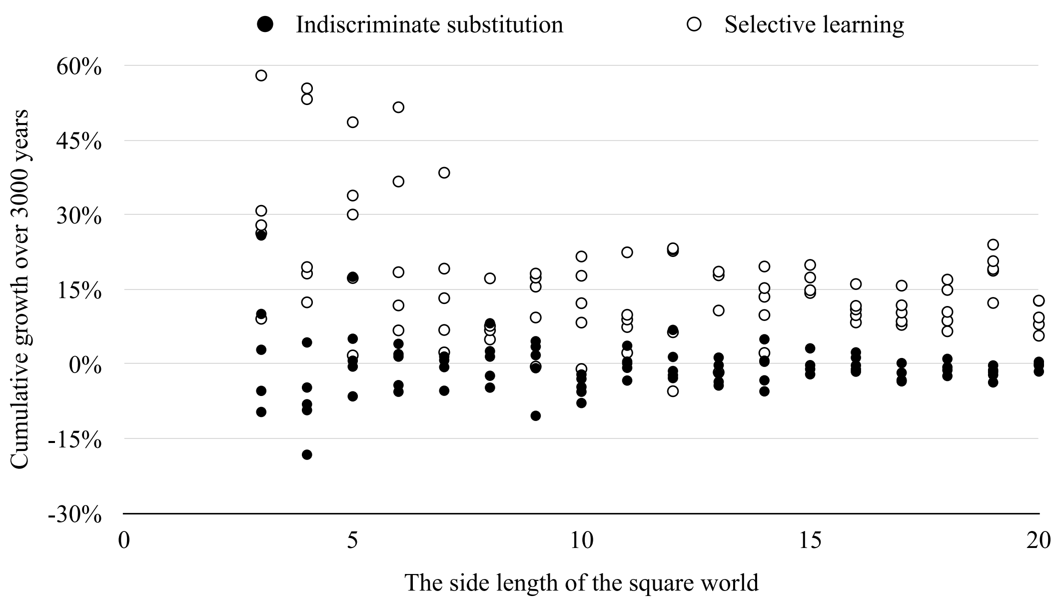

FIGURE 20: The cumulative growth of global average utility over 3000 years. Notes: each point represents an experiment at the corresponding side length of the world.