Notes on *The Economic History of Europe*

Levi van Boekel

july 27, 2024

Contents

This page is intentionally left blank.

1 The Neolithic revolution: The rise of civilization

1.1 Hunter-gatherer societies

When talking about the *Economic History of Europe*, we shall keep in mind that we are in fact only describing a brief moment of history for mankind as a whole. In fact the before before the first Neolithic revolution covers around 99 pct. of human history, but is only described scarcely, as reliable data of sources of high quality simply do not exist. Studies have, however, suggested population densities of around 1 person per 0,6 per km^2 with a total world population estimated to be around 6-10 million. During the period technological progress remained extremely slow, as there, historically, seems to be a rough positive relationship between population size and density and technological progress. The latter is explained by the fact that innovators are sparse, and are thus simply more likely to emerge when more people are pooled together. The scarce technological progress that occurred did, however, not increase the yield of the game they were hunting, but rather increased leisure time, as humankind was simply bound by the size of the stock available for hunting. Thus population growth seems to have been unusual, as the population was simply limited by the available food resources. It also seems reasonable to assume a high degree for egalitarianism, as hunting was often done in teams.

1.2 The Neolithic Revolution

Even though data on Hunter-gatherers is scarce, we know that the Neolithic Revolution had its origin in southwest Asia and began around 12,000 years BCE, reaching Europe approximately 4000 - 3500 BCE. There is furthermore sound archaeological evidence showing how agricultural practices developed independently, highlighting a possible link between agriculture itself and some general human characteristic.

While it for modern readers may seen like a trivial fact, that early mankind went from hunting and gathering food in nature to harvesting crops and managing farms, the consequences of the change are immense. First and foremost, technological progress emerged as a something that could in fact increase population, whereas hunter-gatherers were bound by the growth of the stock of game they were hunting. Thus even though technological progress most certainly made life easier for hunter-gatherers by e.g. reducing the time needed to hunt, mankind was still bounded by the size of the prey population. However, with the Neolithic revolution unfolding, farmers could suddenly choose the best seeds to increase yields, while simultaneously domesticating animals and thus managing the quality, size and reproduction of their stock through deliberate selection and breeding. Through learning-by-doing technological progress now suddenly could increase the production of goods with given resources, what we know as TFP growth. Being somewhat independent of *natural constraints*, furthermore, introduced a trade-off between labour and leisure, highlighting how farmers suddenly could indeed experience population growth by simply having a relative preference towards labour and income (in a broad sense). Perhaps somewhat counter intuitively, this does not necessarily imply a welfare increase (e.g. higher output per capita), as we will discuss later.

Having a society centered around agriculture also marked a remarkable shift from egalitarianism towards private property rights, which suddenly gave room inequality (e.g uneven distribution of land and/or food). Moreover, scholars (e.g. Diamond, 1997 and Weisdorf, 2000) argue, that the food surplus that farmers were able to generate through cultivating nature, allowed for the early establishment of a non-food-producing sector, including craftsmen, chiefs, bureaucrats, early *scientists* and priests, which in turn demanded innovations such as writing, city planning, markets and ultimately

institutions (Weisdorf, 2005). The theoretical explanation is straightforward and concerns Engels Law, stating that as household income increases, the proportion of income spent on food decreases, even if the absolute amount of money spent on food increases. Thus there was suddenly *room* for non-essential sectors to form, fostering early labour division. Increased population, moreover, created so called *agglomeration effects*, as knowledge (a non rival good) could spread, stimulating some degree of convergence while fostering services like mathematics, writing, and governance.

Weisdorf (2005) thus regards the establishment of the non-food sector as the main reason reason that the Neolithic Revolution was decisive to economic growth and ultimately made the Industrial Revolution possible, thousands of years later.

1.3 Why farm?

As discussed above, the Neolithic revolution introduced a labour-leisure trade-off where humankind for some reason had a relative preference towards labour. In fact we see from archaeological evidence that many of the first farmers were worse of than their hunter-gatherer ancestors, as they were shorter, were more prone to disease and lacked important vitamins. In the textbook associated with this course, Person et al. (2023) argue that risk aversion is the main determinant for this seemingly paradoxical behaviour, arguing that humankind was willing to accept a decline in living standard and an increased labour effect to reduce their risk in acquiring food. Other scholars, however, yield different views on the sudden increase in labour effort. Smith (1975) argues that the extinction of herding animals by hunters in the old stone age led to the rise of agriculture, while North and Thomas (1977) argue that population pressure in infertile zones played an important role. Furthermore, Olsson (2001) finds empirical support for a hypothesis claiming that environmental factors and genetic changes in the species suitable for agriculture, made farming possible and attractive.

According to Weisdorf (2005), the only widespread theory thus remains, that there simply is no single explanation that seems to introduce a compelling explanation for the radical shifts from millions of years of successful hunting to, at least initially, debilitating farming.

2 The emergence of early Europe

2.1 From Greece to Rome

Greek civilization is by many scholars regarded as having laid the foundation for later western civilization. Thus despite limited land, Greece achieved population densities (and thus agglomeration effects and innovation etc.) comparable to sixteenth-century Holland (what we describes as the *birthplace* of modern economy). This was only possible due to an advance trading network between the different cities (polis), exploiting the benefits of international trade and division of labour. From historical records, we e.g. know that Greece imported grain, while exporting manufactured goods and quality foodstuffs. However, as the trade related institutions were still in their infantry, the model was vulnerable to distributions in transport (e.g. piracy) and political turmoil. Nevertheless, the trading network linked the entire Mediterranean, and the profits associated with it, fostered *inventions* in material production and intellectual development within philosophy (e.g. Socrates), science (e.g. Pythagoras) and arts (e.g. Homer).

The decline of Greece was ultimately more political than economic, why the Hellenized regions large parts of their Greek heritage, while shifting their focus from the Mediterranean to Italy and thus shaping early European civilization. The geographical boundaries of Europe have in fact shown perhaps surprising continuity since the Roman and Carolingian empires.

2.2 Trade as a unifying force

An important similarity between Greece, Rome and modern day Europe, is trade. Throughout history we have seen countless of examples of how trade can be seen as a unifying force, as it tends to erode border effects [1] and thus leads to a degree of institutional convergence. This can e.g. be seen in command weights and measures, harmonization of commercial laws, a lingua franca, trust, etc. Under the right initial institutional framework trade furthermore, promotes sigma convergence (i.e. that the variance in economic prosperity between the trading partners falls). The latter is the case, as trade leads to knowledge transfer, which itself is a non rival good and thus can be *used* by the *receiver* without limiting the *sender* of the knowledge. Trade thus serves as a positive feedback loop in the sense that higher knowledge leads to even more trade, which in turns leads to higher knowledge transfer. Finally, it is worth noting that trade unions (like the EU) today include an element of geo-politics, essentially arguing that mutual interdependence reduces the chance of war between trading partners, as the cost (lost trading revenue) is simply too great. However, recent developments in Russia have challenged this belief, showing that trading partners may indeed accept a recession in order to expand their borders.

To understand the dynamics of trade, we use the following simple *gravitation model of trade*, implying a positive relationship between economic size and trade, and a negative relationship between distance and trade:

$$T_{ij} = \frac{G \cdot M_i \cdot M_j}{D_{ij}^{\beta}} \cdot C_{ij} \tag{1}$$

where T_{ij} is the trade flow from country i to country j, G is a constant, M_i and M_j refers the economic activity (often GDP) of country i and country j respectively, D_{ij} is the distance between country i and country j, β is the elasticity of trade flows with respect to distance and C_{ij} refers to the border effects, which can be both positive and negative. The intuitive reasons as to why trade is negatively correlated with distance is been coined the *fundamental problem of exchange*, and refers to the fact that imperfect and asymmetric information, lead to greater risk and uncertainty when conducting long distance trade, compared with spot exchange.

2.3 The Dark Ages

During the Dark Ages (the century after the decline of the Roman Empire), cities lost population and skills, trade declined, while infrastructure such as roads wore down due to a lack of proper maintenance. Moreover income fell for both ordinary and the rich, while social order was itself difficult to maintain. However, it is important to keep in mind, that even though trade declined it did not cease. The same goes for technological progress; it was slow, but it did certainly not stagnate.

2.4 Smiths' Theory on Pre-industrial Growth

When studying pre-industrial economy, it is important to keep in mind that these economies were based on a scarce amount of capital and knowledge (i.e. A and K in a classic Cobb-Douglas production function). The latter is explained by the fact that scientific progress was not made, and knowledge thus was developed through learning-by-doing and simple trial and error. The main determinants of pre-industrial growth according to Smith do thus fall back to division of labour (specialization), learning-by-doing, and trade based on differences in resource endowments. The argument concerning division

[1] Border effects refer to cultural, religious and jurisdictional differences that serve as barriers to trade.

of labour is similar to Ricardo's proposition of comparative advantages, with the only difference being the fact that we are not dealing with countries but individuals. Division of labour thus increases the production possibility frontier by allowing each individual to specialize into the field where she has a comparative advantage. A by-product of the specialization is the fact that it fosters efficiency gains in two ways within the specialized area. Firstly, (1) as the individual specialising simply spends quite some time doing her task, she will become gradually better (a term coined *economies of practice*). However, the gains from repeating a single task multiple times are tied to the individual producer and cannot be transferred between individuals. Secondly, (2) learning-by-doing meant that producers learn from the similarities they observe and thus change the production ever so slightly, when they learn something new. Even though this is not formally written, it still outlives the individual worker, as it is incorporated into the production process. Imagine e.g. a cloth production that gradually changes the share of linen in each cloth to increase durability, as this correlation is observed by the workers. When the workers who initially observed the former die, their idea will outlive them, as the next generation can do the same, thus effectively increased technology, A. Furthermore, certain investments in e.g. the watermill were associated with high fixed costs and were thus only profitable, if the demand was sufficiently high. This leads to the insight, that demand itself can be a positive feedback loop, as it allows for more sophisticated technologies, which in turn increase demand and income even more...

Yet, Smith was very aware of the fact that division of labour demands certain economic forces which he coined *the extend of the market*. We might understand the latter as the aggregate demand, however, Smith was not only interested in the economic demand, but also the fact that a market with high demand increased available knowledge of technological possibilities and social factors. Increased population will thus increase the extend of the market, and thus the division of labour, as long as the aggregate income also increases. The latter is not always given, as we shall see later. It is, moreover, important to keep in mind that the extend of the market is defined over a geographical unit and can grow by extending the borders, which is an important explaining as to why division of labour decreased after the fall of the western Roman empire. Generally speaking, increased political has a positive effect on the extend of the market, as (1) safety within a given territorial unit increased and (2) the borders itself can increase.

After the fall of the western Roman empire, markets became increasingly imperfect, as information flowed slowly an was distorted on the way. One way to measure the former is to see, whether price differences for identical goods between regions where higher than their transportation costs, as this implies unclaimed arbitrage possibilities that would have been exploited in a market with better information flow. However, as trade increased with time, market performance and information flow improved, ultimately helping economies coming closer to their capacity constraint.

2.5 Malthus' Theory on Pre-industrial growth

Malthus developed a theory explaining what, to him, seemed like long-run population and income stability for mankind. His theory builds upon the following three assumptions (as described by Clark, 2007 and Persson, et al. 2023) and their feedback mechanisms in between;

1. The fertility-schedule in each society is coined the *preventive check*, and is (1) determined by social customs and (2) correlates positively with material living standards (e.g. GDP/capita).

2. The mortality-schedule in each society is coined the *positive check*, and correlates negatively

with material living standards (e.g. GDP/capita).

3. A binding land constraint implying diminishing returns of labour, leading to a negative relationship between population and income per person.

The following graph gives a visual understand of the assumptions and their implications:

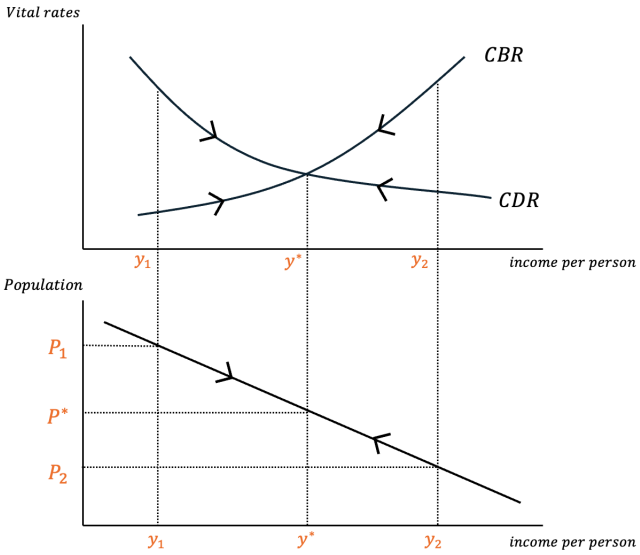


Figure 1: Equilibrating mechanism in the Malthusian model

Note: The figure is inspired by Clark, 2007 and Persson, et al. 2023. The convex and concave CBR/CDR schedules do not immediately follow from Malthus' primary texts.

The top graph shows relation 1 and 2, where CBR (crude birth rate) is increasing with income per person and CDR (crude death rate) is decreasing with income per person. We define the Malthusian equilibrium as the intersection between the two curves, as population is here at a steady-state ($CBR - CDR = 0$). The income associated with the equilibrium, is coined *subsistence income*, as this is the income necessary for the population to simply reproduce itself. However, as Clark (2007) notes, the term subsistence income leads to the incorrect nation that people (in equilibrium) barely survive. This is obviously not true, as differences in the location of the mortality and fertility schedules generate different subsistence incomes, and subsistence thus for one society could be starvation, while it e.g may be double the required calorie intake for another (Clark, 2007). According to Clark (2007) almost all Malthusian economies in fact had subsistence incomes that considerably exceed the income needed to avoid starvation. We shall thus understand the Malthusian equilibrium as the *natural income* for a given society, given a CDR and CBR schedule, rather than the actual income needed for survival for that specific country/region.

The bottom graph (sometimes referred to as the *Technology schedule*) shows the negative relationship between income and population, which follows from the binding land constraint and an assumption of slow technological growth. It is not clear *how* fixed the supply of land was in the industrial era as a whole, and different scholars yield different views (e.g. Clark 2007 argues that land was inherently fixed in supply, while we in the course have discussed the possibility of the constraint not being as strong [2]). However, there is no doubt that land during specific eras was a fixed input in production, thus implying diminishing returns of labor during these periods.

[2] More specifically, we discussed the post-plague society as an example of a period, where land was most definitely not a binding constraint, as large parts of the country were abandoned

Combining the two graphs (which are drawn for a given technology level), it becomes eminent that society will converge to subsistence income, y^* and a given population, P^*. Assume e.g. that the initial equilibrium is distorted due to the plague, shifting population from P^* to P_2 and thus increasing income from y^* to y_2. Society is now at a point, where $CBR > CDR$ implying population growth. However, as population grows, income falls, until society again reaches the Malthusian equilibrium. The same argument can be made from the *other side* of equilibrium, i.e. a starting point to the left of y^*.

In fact even an increase in the technology level will not increase living standards in the long run. Here's why; assume farmers around the world find a new crop that increases efficiency (e.g. yields more calories and is easier to grow [3]). Graphically speaking we can interpret that as an outward shift of the technology schedule, as shown below.

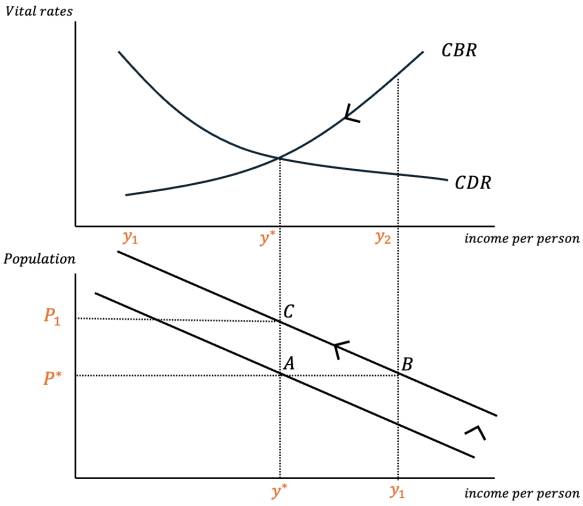


Figure 2: Technology shift in the Malthusian model

Note: The figure is inspired by Clark, 2007 and Persson, et al. 2023. The convex and concave CBR/CDR schedules do not immediately follow from Malthus' primary texts.

The initial effect from the shift can be thought of as going from point A to point B, implying higher income for the same population size (P^*). However, as seen from the top graph, society is now to the left of equilibrium, meaning that births exceed deaths and the population will increase until society eventually ends up in point C, where population has increased from P^* to P_1, but income remains at subsistence level, y^*. The fact that societies are bound to their subsistence income is coined *The Malthusian Trap*.

Malthus himself proposed possible solutions to *The Malthusian Trap*, ranging from changing human behaviour, to an increased level of the quality needed for subsistence. The first proposition is closely tied to the so-called *quantity-quality tradeoff* of children that families are facing when choosing to reproduce. In short it argues that families — due to their own resources being finite — may choose between the quantity of children they wish to have, and the quality (i.e. education, health, etc.) of the very same. If society favours quantity, e.g. through an economy based on agriculture relying on unskilled labor, parents would have a relative preference towards quantity, whereas the opposite

[3] Nunn and Qian (2011) find that the potato was exactly this type of technology shift and simply increased the birth rate during the Malthusian period.

is true for a society relying on high skilled labor. The paradox here, however, seems to be that a society based on high skilled labor is a society with high population densities as specialization can only emerge here. A possible answer, in line with the so-called *Unified Growth Theory* therefore may be, that mankind simply had to live through the Malthusian era to increase welfare at some point in the future. A preference towards quality was something Malthus referred to as *Consumer Behaviour*, and saw a possible way of escaping the trap (Rutherford, 2007) . However, Malthus was sceptical about increased specialization (which would lead to *consumer behaviour*), as it create an imbalanced economy, relying on manufacturing and commerce (Rutherford, 2007). History has though shown that specialization, trade, and the efficiency gains associated herewith, combined with a preference towards quality in children, effectively lead to humankind escaping the Malthusian trap.

During the course, we have numerous times discussed the validity of Malthus' theory and the short answer is that results are inconclusive. However, we can safely dismiss the *strong version* of the Malthusian model as presented in Persson et al. 2023, arguing for zero population and income growth. Work done by Galor and Weil (2000) e.g. suggests growth in population of around 0,1 to 0,4 pct. yearly from 500 to 1820, whereas growth in output per head is between 0 and 0,2 pct. yearly.

Empirical results furthermore show little evidence of the positive check, which itself is in line with Malthus' theory that suggested that the preventive check, by reducing population growth, *prevents* the positive check. The preventive check itself, is well documented until around 1800 in e.g. Clark, 2007 and Møller & Sharp (2008), however, John Hajnal, hypothesized that this simply reflect peculiarities of fertility rates in Western Europe, even though the empirical evidence for his theory remains weak. Evidence for the diminishing returns remains somewhat inconclusive, however, Pedersen et al. 2021, have recently found some evidence in North Italy. Lastly, researches agree that none of the Malthusian characteristics can be observed after 1900 in Europe, giving rise to the *Unified Growth theory...*

Both Dittmar (2011) and Nunn and Qian (2011) find results of technology growth in the preindustrial era, and while the former does not deal with Malthusian forces, Nunn and Qian find no positive relationship between population growth and urbanization until 1750/1800, implying Malthusian dominance.

2.6 Concluding Remarks on Pre-industrial growth

During the course, we have worked with two diverging views on pre-industrial growth: Smith and Malthus. Whereas the former argues that sustained growth in output per head is indeed possible, Malthus held a more dismal view, arguing that society was bound to live at subsistence income [4].

History suggests small increases in welfare during the pre-industrial era, implying smithian dominance. However, we shall keep in mind that data on real wages in this period is associated with some degree of uncertainty, and often only exists on day-basis leading to potential issues described in later chapters. We shall, furthermore, remember, that national income is composed of both wages and capital income (rent), why a stagnating real wage may not necessarily imply Malthusian dominance. To make matters even more complicated, work done by Voigtländer and Voth (2009) has suggested, that urbanization has shifted the CDR curve outwards permanently leading to a higher income, why the conclusion about smithian dominance seems increasingly ambiguous.

Empirical work by e.g. Thirsk (1988) , has tried to estimate the TFP growth during the period

[4] Which, as we might remember, does not in fact mean that they are living on the verge of famine, but rather describes a *steady state*, where the CBR and CDR schedules intersect, and $CBR - CDR = 0$

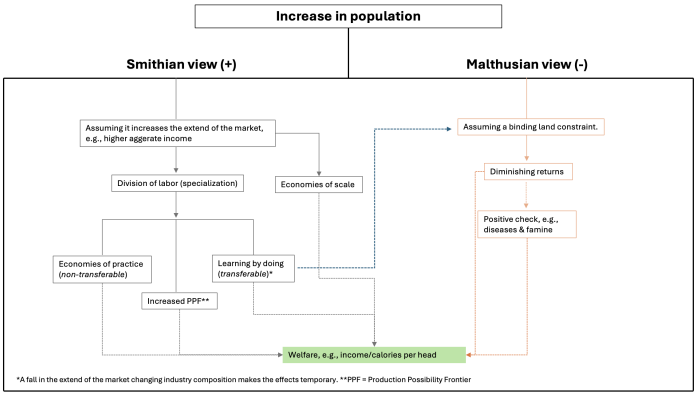


Figure 3: Smithian and Malthusian preindustrial forces

Note: The figure is inspired by Persson, et al. 2023.

to validate the above made conclusion, and find a growth of around 0,2 percent year year in England before the plague. Their work is based on the assumption, that societies were agrarian and the sum of wages and rents thus is equal to output, and if we see an increase this necessarily implies that either land and/or labor is used more efficiently. Other estimates, based on urbanization as a proxy for welfare, yield similar results (e.g. Persson et al. 2023). To see why urbanization is a helpful proxy, consider a closed economy, where we see a relative increase in the on-food-producing population. As the agrarian sector can now feed a higher amount of people, we can safely conclude that there has been a productivity increase in the agrarian sector. In reality, economies are, however, not closed, why we need to control for imports, which is done in the specific studies. It is important to add that the empirical work on TFP growth finds considerable heterogeneity among different parts of Europe, whereas areas with poor institutions, such as bad trading networks, or social disorder, seem to have lacked behind. We may conclude that there seems to be strong evidence pointing towards TFP growth in the preindustrial era, while assessing whether Malthusian or Smithian forces dominated seems to be less clear.

2.7 Unified Growth Theory

Unified Growth Theory attempts to divide economic history into three distinct periods; (1) the Malthusian era, (2) the post-Malthusian era, and (3) the modern era, each associated with different characteristics. In the Malthusian era, growth in income per head was slow, due to the Malthusian forces pushing societies back to subsistence income, as technological progress was slow and in the long run simply increased population. Another way of phrasing the prior is the lack of a positive relationship between population growth and income per head. Eventually, Galor and Weil (2000) argue, the population growth will, however, lead to specialization and rapid technological progress that is faster than the population growth, implying a positive relationship between population growth and income per head. A period with this relationship is coined *the post-Malthusian era.* Nunn and Qian (2011) find empirical evidence for both the Malthusian (approx. 1700 - 1800) and post-Malthusian (approx. 1800-1900) era using urbanization as a rate of output per head, as it is assumed that urbanization is positively correlated to income per head. The era is followed by a period of sustained growth, but where households experienced a change in their relative preferences favoring the *quality of children*, relative to the quantity. This epoch is referred to as the modern era, whereas the change in fertility preferences is coined the demographic transition.

While it is clear that the demographic transition reduced family size of four children to around two children in Western Europe, the forces behind the transition seem to be more nuanced. First and foremost, it is important to understand that the decline in fertility occurred before the widespread use of contraceptives, and rather coincided with early urbanization, industrialization and higher household income. We can thus isolate both an income- and substitution effect, in regards to family size, that both point in opposite directions. While income increases the amounts of children one can have (income effect), the income forgone in giving birth is rising as income increases, implying a negative effect (substitution effect). As we have been witness to decades of falling fertility we might conclude that the substitution effect dominates. However, we there some to be other factors at play aswell. First and foremost, we might add that that parents are altruistic and care about the success of their children. Furthermore, it seems plausible that a fall in child mortality influences fertility (families have a target size), but that the latter comes with a lag, as families want to make sure that the decline is permanent. In societies where the child mortality thus has fallen earlier, we might expect the demographic transition to step in earlier. Finally, the fertility strategy is dependent on whether or not households can externalise the cost of children through institutions such as the welfare state. Here we, however, run into opposite effects once again, as the welfare state increases the extend to which parents need economic care from their children when they are elderly (negative effect), while simultaneously decreasing the cost of having children. While it seems ambiguous how strong each of the forces mentioned above may be, it is clear that the negative effects have dominated, resulting in the demographic transition.

2.8 The Great Divergence

As Europe — economically speaking – lagged behind the leading civilizations well into the second millennium, it seems mysterious why the industrial revolution happened in Western Europe. The so-called California School gives an intuitive explanation, arguing that the location of coal deposits in Western Europe eased the energy constraint since the alternative at that time was to eradicate scarce land to grow fuel from timber. Europa thus managed to escape the Malthusian trap by concentrating on energy-intensive manufacturing and thus leveraging their comparative advantages, increasingly importing food and agricultural raw materials from the Americas. However, other scholars argue that the divergence started long before the industrial revolution, as income differences were small around 1500, but increased as income fell in China and India over the following centuries. The questions as the why Europe managed to keep the relatively high income, remains unanswered, however, we Persson et al. 2023, argues that the preventive check, through late marriages, in Europe may have balanced the relationship between diminishing returns and technological progress in an economically more sustainable fashion. He further argues, that European households were *better* at practicing family planning within marriage, whereas Europe as a whole regulated nature and agriculture more *systemically* — unbounded by religious and political authorities. Thus we might add that institutions in Europe (both formal and informal), such as openness to trade and a culture for specializing in agriculture, contributed to a more efficient adoption of new useful knowledge (a non rival good), effectively improving the technological level (A), hereby escaping the Malthusian Trap.

3 Institutions

3.1 Introducing Institutional Explanations

During the course, we followed North's (2005) framework of describing institutions as *the rules of the game*. They range from formal, such as private property rights, to informal, such as trust, while

defining the way the game is played, with organizations being the players (North, 2005). North (2005), defines organizations as groups of individuals held together by some common objectives, such as trade unions, universities, firms, political organization, with the ultimate objective of survival, since all organizations live in a world of competition, as resources are scarce (North, 2005).

North (1991) derives this explanation from a game theoretic context, arguing that wealth maximising individuals (hence organizations) will find maintaining cooperation (needed for survival) difficult when the time horizon on a game is unknown, when information on the other players (organizations) is lacking, and when there is a large number of players (organizations). It turns out that individuals (organizations) often engage in games, with similar characteristics as the above, why institutions in part can be understood as *rules* that permit low transaction costs (in a broad sense), given the game outlined above. North (1991) argues that efficient institutions do this by raising the benefit of cooperative solutions (provide mechanisms that reward individuals or groups for working together), while simultaneously raising the costs of defection (i.e. raising the costs of going against the agreed-upon rules or norms). An example of an early institution, that we worked with in the course, are the champagne fairs, where trade between a large number of individuals was facilitated in a specified region, raising the benefit of cooperative solution by e.g. reducing transaction costs and making gains from trade realizable, while raising the costs of defection, as fraudster were fined and banned from attending the fairs, which in and of itself was costly.

When talking about institutions, it is worth stressing that they are rarely Pareto efficient in the sense that they will increase welfare without leaving anybody worse off. In fact, most institutions will create both winners and looser, why we may refer to institutions as efficiency-enhancing, as long they lead to a net increase in welfare [5] The fact that institutional change creates losers is, furthermore, an important explanation as to why inefficient institutions may survive, as long as the losing organization has enough political or military power to maintain the specific rule of the game — a mechanism coined *incuded preferences* by Acemoglu (2006). Examples of the latter include colonial trading companies that lasted for decades, without being efficiency enhancing, as the institution served the interest of a powerful elite, and the abolition was not Pareto improving either, as the slave trades were not fully compensated for their loss in welfare.

However, even the existence of efficient institutions requires a peculiar type of causality, why we might not always be able to find a reason for an institution's existence beyond the fact that it serves certain interests or simply arose by chance. To appreciate the peculiar nature of institutional explanations consider the following statement:

LLC companies emerged at t_0, because they stimulate investment by reducing the risk for the individual investor at t_1

and compare it with:

Higher interest rates at t_0 decrease spending at t_1.

What you notice is the fact that the time structure is reversed, in the sense that LLC companies emerge before the beneficial effects arise, whereas the second statement describes an action in t_0 that happens before the effect in t_1. The structure of the former statement is called a a consequence explanation, and why it may be appropriate to describe certain institutions, it is important to understand that intentional creation of institutions has been (and is being) practiced throughout history.

[5] While we have not explicitly introduced it during the course, this efficiency criterion is closely tied to the *Kaldor-Hicks efficiency criterion*, which states that something is efficient, if the winners could — in theory — compensate the loser and still be better off, thus creating the possibility, for Pareto efficiency without the compensation actually taking place.

3.2 The institutional characteristics of a modern economy

We define the institutional characteristics of a modern economy as follows; free access to functioning markets, advanced division of labour, and a government that respected and enforced (intellectual)property rights. Vries and Woude (1997) have declared Holland the *first modern economy*, as it possessed the three characteristics described above. Political stability, trade and colonialism [6] seem to be some of the reasons, as to why Holland the institutions emerged, as they were needed to support a urbanized society and associated division of labour cf. North's explanation on game theory.

The understanding that functioning markets foster growth, is based on the assumption that factors of production seek the most productive use, when it is allowed to be trade freely, as the person with the highest marginal benefit will acquire the factor. As the most productive use promotes growth, we can easily validate the claim. However, we must keep in mind the game proposed by North (1991) and how markets, if not accompanied with other institutions, may not be efficient. Examples of mechanisms challenging the efficiency include the *fundamental problem of exchange*, and slow information flow, leading to markets not clearing, thus inhibiting trade.

The idea that advanced division of labour has a positive effect on the economy is based on Smiths work, discussed earlier, highlighting how it can (1) lead to economies of practice, (2) increase the PPF and (3) increase technology levels by learning by doing. It, furthermore, touches upon the game theoretic setting described by North (1991), as it creates a demand for more sophisticated institutions, lowering the transaction costs of the inter-sector/country trades.

The relationship between of (intellectual)property rights and growth is likewise central, and touches upon two central arguments. Firstly, private property solve the *tragedy of the commons*, where common resources are over exploited, as they are freely available and the externalises associated with ones use are not internalized. Thus the social costs (MSB) are higher than the private costs (MB), leading to market inefficiencies. While property rights solve the above outlined problem, they may — if new *rules* are not put into play — lead to distributional effects, such as monopolies on certain resources, which may not be efficiency enhancing. We might, furthermore, add that private property rights stimulate investment, as they provide security and incentives for organizations (broadly speaking, using Norths terminology).

However, the Dutch case shows that inclusive institutions are necessary but not sufficient conditions for sustained growth, as the British economy surpassed the dutch in the early nineteenth century. The change, furthermore, highlights how knowledge (a non rival good) on institutions may lead to sigma convergence between two countries, as Britain simply implemented the progressive dutch institutional framework. However, we need to keep in mind that Holland and Britain both had constraint on the political elite, leading to a certainty in tax assessment, whereas other European countries were ruled by absolutist monarchies, where the future taxes were uncertain, harming growth. Schumpeter further argues that innovation inevitably *destroys* what preceded it, e.g. streaming media has *destroyed* physical media. Acemoglu and Robinson (2012) continue the argument, and add that this process is only possible in markets free from significant monopoly and oligopoly constraints (such as Britain or Holland).

[6] While colonialism explains plays a role when assessing how Holland got wealthy, it is not a feasible explanation as to why Holland got ahead of e.g. Africa. This is due to the fact that the colonization itself was only possible, because Europe was *ahead*, economically speaking, in the first place.

3.3 Business firms

Workers in factories and farming households historically differed in the sense that peasants usually would own or lease their land and thus be the residual claimant, i.e. receive the profit, whereas workers would be employed at a fixed wage by a factory owner, who acts as the residual claimant. The explanation as to why this difference arose, seems to touch upon different economic aspects.

Firstly, agriculture is associated with risk, as climate variability affects the harvest, thus giving landowners incitement to rent out their land. The risk will thus be spread between the landowner and the tenant (peasant). Furthermore, monitoring work effort is difficult in agriculture, but since the peasants work for themselves, they can by definition not *cheat* and thus monitoring is not necessary. The opposite goes for business firms; monitoring is needed, but often easy in e.g. factories. Furthermore, there economies of scale are more prevalent in industry sectors. However, this does not explain why capital owners employ labour, and not the other way around. It is tempting to argue that imperfect capital and social mobility markets constrained workers from borrowing capital to set up labour-managed firms, however, the argument seems to forget the fact that the two company types have varying interests, affecting their productivity and adeptness. Whereas capitalist, business firms, simply maximize returns to owners, labour-managed firms have a more diverse set of goals, including limiting redundant activity. We may thus conclude that the rise of the business firm can ultimately be explained as the *survival of the fittest organizational type*. In some cases, capitalist firms were challenged by co-operative, however these did typically not dominate, except in Denmark, were e.g. Arla crowded out capitalist firms within a couple of decades. It seems difficult to point out the exact reasons why Arla was able to penetrate the market, however, Persson et al. (2023) highlight technological capabilities, the perishability of goods, legal frameworks, and the traditional structure of ownership and employment in the industry.

Up until the First Industrial Revolution [7], big business firms were, however, relatively uncommon, beyond the trading monopolies. The main exception were so called *manufactories* that functioned as place where labourers could work *under the same roof*, rather than modern factories dividing production parts into small parts and allowing for specialization. However, after the first industrial revolution (1750/1760 - 1820/1840) factories began to emerge, while the putting-out system [8] declined. Manufacturing and factory production, however, coexisted for a long time, highlighting how the transition between the two was rather complex.

During the second industrial revolution (1870-1914), firms starting becoming multiunit, i.e. multifunctional, multiproduct, and multinational entities. This was made possible by new new transportation and communication systems, highlighting how technology and institutions shaped the world economy. With increasing units, firms needed managers with specific technical skills to include the newest technology into their production. Firms started investing heavily in R&D (Research and Development), while innovation at universities etc. was likewise rapid. As knowledge that is not patented is a non rival good, the public innovations increased the possibilities to create new sectors and products, thus increasing both supply and demand [9]. The Second Industrial Revolution, with its focus on chemicals, was followed by a Third Industrial Revolution founded in physics, leading to the microelectronics, nuclear energy, jet engines, mass air transportation, and the internet. The latter has

[7] We have during the course challenged the view of a Industrial Revolution, as it seems to have been more of a continuous process, see more in the relevant section in this note.

[8] Loosely speaking, the putting-out system served a pre-industrial production method where merchant-employers distributed raw materials to rural workers who processed them into finished goods at home, which were then collected and sold by the merchants.

[9] An example of conscious demand increases relates to Henry Ford, who paid a *high* salary to his employees, in part, to stimulate consumption.

allowed organizations to reduce controlling costs and becoming more decentralized, while products have become more *modular*, as they source technologies from a wide variety of different firms [10]. It is clear that the increased outsourcing and trade, has lead to more complex *games*, where the number of players has increased dramatically, while their respective utility function (they aim maximise), may, too, be increasingly different, highlighting and understanding both the need and emergence of new *rules for the game* (institutions) to ensure a welfare maximising outcome.

3.4 Course Literature on Institutions

Cantoni and Yuchtman (2014) investigate the causal link between universities and market establishment in medieval Germany. They argue that the papal schism of 1378 created a natural experiment by causing the establishment of the first universities in Germany, which significantly reduced the distance German students had to travel to receive university education. Using a DD-model they then compare the rate of market establishments before and after 1386 (the year the first German university was founded) in cities that experienced a significant reduction in distance to a university versus those that did not. Cantoni and Yuchtman find a statistically significant increase in the trend of market establishments after 1368, with the trend increasing the most, where the distance to a university shrank most. They explain the development with the fact that universities mainly provided training in Roman and Canon law, which fostered the development of legal and administrative institutions. The latter in turn had a positive effect on economic activity, as it reduced transaction costs (in a broad sense), and thus provided efficient *rules for the game*, as described by North (1991). The study implies a positive feedback loop between different institutions, as more sophisticated legal institutions increase economic activity and thus increase the demand for legal professionals — schooled from university — even more.

Acemogulu et al. (2005) examine the impact of Atlantic trade on institutional development and economic growth in Western Europe between 1500-1800. Using a DD-model they compare the urbanization rate (as a proxy for GDP per capita) [11], between countries that were included in the Atlantic trade, and those that weren't, before and after the rise of Atlantic trade. However, as being a part of Atlantic trade is clearly endogenous, they use the Atlantic coastline-to-area ratio as a measure of potential access to the Atlantic trade, as this is clearly exogenous. They argue that Atlantic trade contributed to European growth through a interplay two different effects, in countries that met to preconditions: easy access to the Atlantic and nonabsolutist initial institutions. Firstly (1), it created large profits that enriched and empowered commercial interests. These profits gave the Monarch relatively less power, leading to reforms in political institutions, such as more secure property rights — one of three preconditions for a modern economy. The process was particularly effective in Britain and the Netherlands, where initial non-absolutist institutions facilitated greater economic benefits from trade, whereas countries with more absolutist regimes, like Spain and Portugal, saw limited institutional change and economic growth. Finally, the study implies a positive feedback loop between trade, institutional development, and economic growth. As institutions improved and protected property rights, economic activity increased, essentially creating a more complex *game*, thus demanding better *rules* (institutions) to maintain an efficient outcome.

Angelucci et al. (2022) study the causal relationship between urban self-governance and parliamentary representation, and its subsequent effect on national institutions in England after the Norman

[10] It has e.g. been estimated that around 250,000 patents are involved in a modern smartphone.
[11] They later validate their results with Maddison GDP per capita data.

Conquest of 1066. Using a both a DD and IV-approach they find that towns with Farm Grants (the right of self-governed tax collection and law enforcement) were significantly more likely to be summoned to Parliament, suggesting that municipal autonomy was a stepping stone to parliamentary representation. The authors, furthermore, argue that the process strengthened national institutions over time, as self-governing towns supported Parliament's role both checking and constraining the monarchy, while protecting local autonomy. The study thus implies a positive feedback loop between division of power and economic growth, as further economic growth leads to an even higher degree of division of power, and ultimately pave the way for democracy.

Michalopoulos and Papaioannou (2013) examine the influence of national institutions on subnational development in Africa, leveraging the fact that over 200 ethnic groups across Africa were partitioned during the colonial era, which subjected similar cultures in homogeneous geographic areas to different national institutions. By comparing regions within the same ethnic group but across different countries, they employ a spatial regression discontinuity (RD) approach to estimate the causal impact of national institutions on economic performance, measured by satellite images of light density. The authors find no systematic correlation between variations in national institutions and economic performance within partitioned ethnic groups. This lack of correlation is especially evident in areas further from capital cities, suggesting a diminishing role of national institutions in more remote regions. However, in regions close to capitals, a positive association between national institutions and economic development is observed. This finding highlights the limited penetration of national institutions in the hinterland and underscores the importance of proximity to centers of political power. Furthermore, the study finds significant heterogeneity in the impact of national institutions. For approximately 60% of partitioned ethnicities, national institutions have an insignificant effect on development, while a positive relationship can only be found for around 20-25% of the ethnic groups. The remaining groups show a negative correlation. These findings challenge the conventional wisdom that national institutions are the primary determinants of economic development, suggesting that local factors and historical contingencies play a more significant role in shaping subnational outcomes in lesser developed countries.

4 Knowledge, Revolutions and Convergence

4.1 Knowledge

The shift towards a more scientific understanding of the word was, initially, a pan-European phenomenon, which relied heavily on a strong institutional framework (what Moses Abramovitz called *social capabilities*), and had lead to three different impacts on per capita growth. Firstly, (1) TFP increased, as science based knowledge became an increasingly strong force in production, while secondly (2), the invention of new products and production processes stimulated investments, leading to higher capital per worker, thus erasing a potential diminishing returns to labor due to a capital constraint. Finally, increasingly sophisticated technologies changed the *game*, increasing demand for education, human capital investment and stronger institutions.

However, even though the scientific discoveries were largely made in Europe, this does not imply that the growth was centered there. Firstly (1), as knowledge is a non rival good it can be *copied* freely by developing economies, given they have somewhat inclusive initial institutions. Furthermore (2), knowledge is transformed into working technology by means of blueprints, which likewise flows into less developed economies. This hypothesis suggests beta convergence, i.e. that relatively poor

economies — given a degree of initial inclusive institutions — are expected to grow faster than more advanced economies, why the have *the advantage of backwardness*, as suggested by Gerschenkron.

4.2 The Industrial Revolution

While the traditional view of The Industrial Revolution conveys the idea of a sudden, radical shift, the reality was much different. While it is true that new technologies in the period, (1) the adaption was much slower than initially expected and (2) the technologies were, mostly, sector specific and thus not *general-purpose*. Even the steam engine, which eventually turned into a GPT, was for most of the eighteenth century simply used to pump water from coal mines, as the early engines were extremely energy inefficient and thus needed constant refills of cheap fossil fuel — something only possible at a coal mine, as transportation costs were high. With time, however, the steam engine became increasingly energy efficient, and in the middle of the nineteenth century, steam engines began to have an impact on transportation and only in the late nineteenth century, did steam engines replace long distance sea transport. Furthermore, most industry factories were still powered by waterpower during the industrial revolution, hence the name *mills*. We may thus remark the period has a continuous development process, rather than a revolutionary impact from 1770 to 1830. The latest empirical work, furthermore, finds that the technological that did occur — specifically in the textile sector — were the result of trial and error (or learning by doing, in Smithian terms), rather than scientific discovery. Crafts (2001) thus estimates yearly TFP growth of around 0,14 from 1760 - 1801, which certainly is not much higher than in medieval times, and not at all close to Britain from around the middle of the nineteenth century, where yearly TFP growth was between 1 and 1,5 pct. The main reason, why the new results challenge the old belief of a revolution come down to the fact, that the previous estimates gave too much weight on new and fast-growing industries (such as cotton) when deriving total output, while the size of the industrial sector as such was exaggerated aswell.

Even though it may not have been a *revolution* in the quite literal sense, the question still remains as to why the development started in Britain. In short, there is no clear answer, however, different scholars, yield different views. Allen (2009) for example argues that high wages, driven up by trade, gave an incentive to develop labour-saving technology, he thus adds that capital intensive technology is not profitable in other places of Europe, as wages were simply too low. However, this does not seem to align with reality, as it is not clear, weather unit costs were in fact higher in Britain. Furthermore, it is not generally true that industrial technologies save labour, rather than other factor inputs. Kelly et. al (2015) furthermore find that industrialization actually happened in places with relatively low wages and high mechanical skills. They instead argue, that the quality of British labour , due to better nutritional standards, was simply better than in the rest of Europe and thus had better cognitive and physical skills develop and adapt new technologies. Clark (2007) argues that the spread of *middle class values* from the elite to the rest of the society leads to a economically more successful individuals [12], however his claim seems to miss empirical support. Hodgson (2015) argues that the post 1688 Financial and Administrative Revolutions, which lead to a new financial system, was the main determinant, as it allowed for the growth of collateralizable property and saleable debt, and thus fueled innovation, whereas North and Weingast (1989) argue that the Glorious Revolution in 1688 allowed for property rights and thus led to The Industrial Revolution. Finally, Vries (2008) focuses on the demand side, as he coins the century before the Industrial Revolution as the Industrious Revolution. Here, market involvement increases for all household members, leading to more days

[12] The argument builds upon the preventive check assuming that higher income leads to higher fertility.

of work and an increased income. This spills over into a demand for new products, highlighting how a strong demand side prior to the industrial revolution may have created opportunities for new technologies, contrary to the traditional view focusing on the supply side.

4.3 The Industrial Enlightenment

The strong focus on the Industrial Revolution, tends to suppress the importance of the Industrial Enlightenment, a term coined by Mokyr (2009). Mokyr argues that there was a growing pan-European intellectual enlightenment from 1700 to 1850, which was fundamental for the ground-breaking scientific revolutions of the nineteenth century, which continued to be an integral part of production technologies most of the twentieth century. During the period, scientific societies with a focus on open access knowledge formed, and they actively discouraged their innovators from seeking patents, as they believed knowledge sharing would lead to further scientific discoveries. While it today seems clear that patents are needed to incentivize R&D, it the empirical results on whether patent rights were decisive in early industrializing innovations, remain inconclusive. In the second half of the nineteenth century, however, patent applications were increasing, as the previous knowledge allowed for many ground-breaking innovations to developed simultaneously and independently. Generally speaking, we can identify four general characteristics of modern technological progress;

1. Resource augmenting, thus lessening the constraints of nature.
2. Improved quality of commodities.
3. Development of new products and services.
4. Widening the resource base for industrial use.

While the characteristics seem somewhat self-explanatory, it may be important to comment on the difference between the first and the fourth point. When referring to something as resource augmenting (1), we are essentially saying that the existing, finite, resources we are using (e.g.) are being augmented trough a more efficient use, whereas the widening of the resource base (2) refers to new resources being used (e.g. wood in paper), thus likewise lessening the constraints of nature, but in a somewhat different matter. Furthermore, plenty of the products that were developed in the second half of the nineteenth century, where further improved during the twentieth century and are still articles of mass consumption, e.g.; the telephone, the gramophone and sound recording, cameras and movies, wireless communication including radio, chemical fertilizers which dramatically increased yields in twentieth-century agriculture, durable plastics, dynamite, viscose and the bicycle. We might, however, add that the GPT technology of the twentieth century is electronic computing, i.e. the computer.

4.4 Technology Catch-Up and Convergence

As explained earlier, non-patented technology is a non-rival good, allowing for convergence in living standards between countries, if the initial institutional framework is inclusive enough to support it. There are, however, other reasons for convergence, too. Firstly, (1) sectors dominated by large-scale firms tend to me more efficient than those with many small scale firms, such as e.g. agriculture at the end of the nineteenth century. However, as the market dynamics *work*, the least efficient sectors tend to be crowded out, allocating labour to more efficient uses and thus increasing TFP- and economic growth. Furthermore, the sectors who are losing labour may be forced to increase productive in a Boserupian way to remain *alive*. We may refer to this as convergence due to *structural change*.

Secondly, traditional growth theory argues that poor countries with lower capital-to-labour ratios, will begin saving more and eventually converge to the level of rich countries, as diminishing returns to capital set in. Modern scholars, however, argue that the actions of governments in providing e.g. R&D may in fact counteract the diminishing returns, thus implying a weaker convergence.

Empirically, we see that sometimes countries converge, sometimes they don't, and sometimes they even diverge. After the first world war, institutional differences between capitalist and socialist societies are furthermore important factors, when explaining divergence across Europa. Figure 6.1 in Persson et al. (2023) e.g. shows a strong convergence for earlier *socialist* countries after 1990, whereas the convergence is extremely slow during 1950 to 1990 before. We furthermore need to remember, that beta convergence (e.g. faster growth for initially poor countries, compared with rich countries) does not necessarily imply sigma convergence (e.g. smaller variance between the income per head in countries), as beta convergence for one country may in fact increase variance on a global scale.

Finally, we need to remember, that GDP per capita is at best a weak proxy for welfare, as the latter, broadly speaking, is determined by utility functions that differ between regions. So while hours worked fell about 35 pct in Europe between 1960 and 2010 and income dropped aswell, it does not necessarily imply lower welfare, if Europeans have a strong preference towards leisure. Some might argue, that the high higher marginal income taxes in Europe distort incentives for work, thus having nothing to do with a preference for leisure. Here, however, one might add that working hours in the Nordic countries (those with the highest marginal tax rate) exceed the European average, why the argument is flawed. Moreover, some argue that the gap between the US and Europe is due to the rigidities in the adoption of new technology. Finally, the sectors — and the growth of that sector in the world economy — in which different economies leverage their comparative advantages has effects on how strong convergence we might expect, as shown in later sections.

4.5 Course Literature on Technology and Convergence

Dittmar (2011) examines the causal link between technology and economic growth in preindustrial times, by studying the impact of the printing press on city size (as a proxy for economic growth) between mainly 1500 and 1600. By using running a simple OLS regression, he initially establishes a positive relationship between the adoption of print technology and city growth. To address potential endogeneity concerns, such as e.g. that printers selected cities with grow potential, he uses an IV approach, with distance to Mainz as an instrument, as it is exogenous, and correlates with the diffusion of the printing press, but does not correlate with the outcome variable (city growth). His estimates, based on the method outlined above, imply that printing press accounted for between 18% and as much as 68% of European city growth between 1500 and 1600. The paper thus challenges the conventional wisdom that technological innovations in pre-industrial economies had limited macroeconomic impacts, while it does not necessarily argue against Malthusian forces, as it does not consider the effect on GDP per head. It is therefore entirely possible — and indeed likely — that the technology to some extend simply increased fertility, as other data sources suggest limited growth in welfare per head (e.g. Clark, 2007).

Nunn and Qian (2011) examine relationship between technology and economic growth in preindustrial times, by analyzing the impact of introduction of the potato [13] from the New World to the Old World on population, urbanization and height. Using a DD approach, they compare the changes in

[13] As it yielded more calories, nutrients and vitamins per area of land, compared to other crops, potatoes allowed for more *efficient* diets, thus effectively functioning as a technology.

population and urbanization/height between countries with less or more potato growth. To address the issue of reverse causality (i.e population pressure and the associated demand for food might have caused the adoption of potatoes), they control with a suitability measure for potato cultivation based on geoclimatic conditions (exogenous). Their estimates suggest that the introduction of potatoes accounted for approximately 25% to 26% of the growth in Old World population and 27% to 34% of the increase in urbanization between 1700 and 1900. To validate the fact that urbanization is a proxy for welfare, they include height data aswell, and find a positive effect on the latter aswell. Their research, in line with Dittmar (2011), challenges the traditional view that pre-industrial technology had limited effects on growth, while it finds that the Malthusian forces seem to have dominated until 1750/1800, as they find no positive relationship between population growth and urbanization rates (proxy for GDP per head).

Pascali (2017) estimates the effect of the introduction a new technology (the steamship [14]) on economic growth during the first era of trade globalization (1870 - 1913). Using a DD approach, he compares the changes in trade and economic outcomes between countries before and after the introduction of the steamship, using the asymmetric reduction in shipping times across countries as an exogenous source of variation. Based on a rough calculation, Pascali suggests that the introduction of the steamship might have accounted for at least half of the increase in international trade during the second half of the nineteenth century. However, he argues, that the increased trade has only benefited a small number of countries, namely those with strong institutions. He thus finds that countries with strong constraints on executive power and better investment environments were more likely to benefit from trade. The fact that institutions are an important factor for growth, is in line with other literature from the course (i.e. Cantoni and Yuchtman (2014), Acemogulu et al. (2005) and Angelucci et al. (2022)). Finally, the paper challenges the traditional Ricardian view, which argues that trade benefits all participating countries equally, and instead shows empirical results highlighting that technological advancements, such as the steamship, can lead to significant economic divergence between countries.

5 Money, Credit and Banking

5.1 Defining Money

Money — understood as a common account that all *payments* can be made in — emerged in with the use of grain equivalents in Babylon. The introduction of money thus essentially solved the problem of noncoincidence [15] of wants that was present in a barter economy, by providing a standardized unit of account that serves as a means of payment. Furthermore, money can be exchanged in varying quantities, and allows for a sophisticated financial sector, allowing e.g. loans and investments. Modern economy is thus highly depended on money, as growth enhancing developments throughout history, such as division of labour and international trade, would simply not be possible without money. Money was — and is — thus a prerequisite for long term sustained economic growth.

Following Mankiw (2009) we can define three fundamental purposes that money serves; (1) a store of value, (2) a unit of account, and (3) a medium of exchange. The first purpose (1) refers to the fact that money allows individuals to transfer purchasing power from the present to the future (however with imperfection due to inflation), while the second purpose (2) simply states that money provides

[14] The steamship is specifically used, as it had a significant impact on global trade patterns, with shipping times by steam becoming a dominant factor influencing trade flows post-1865

[15] I.e. the fact that two parties in a barter system cannot find a mutually beneficial exchange because each party does not have what the other desires.

the terms in which prices are set and debts are recorded. Finally (3), money is *what* we use to buy goods and services, pay our debt and finally, pay our taxes.

Throughout most of history, money was bound to a commodity with some intrinsic value [16], leading to so called *commodity money* that could always be exchanged for the underlying commodity. The most prominent example is the gold standard, where money was pegged to gold (which we will come back to in later sections), however on the island of Yap a large stone named *fei* was used, highlighting the possible diversity in the commodity itself. Today, the money we use often has no intrinsic value, and is thus coined *fiat money*, as it is established by government fiat.

5.2 The History of Money and Emergence of Banks

With the decline of the Roman Empire, Europe lost an systematized monetary system that only revived in 794 CE with the Carolingian Empire where the principle of a hierarchy of denominations was introduced [17]. The latter had a strong impact on European money and survived in Britain into the 1970s, when the decimal system was introduced. During the Carolingian Empire, minting was uniformed, however the mint was associated with a seigniorage fee of around 5 to 10 pct. of the face value of the coin. The fee was in part due to the actual minting costs, but also served as an indirect tax. The extrinsic value of a coin was thus the sum of the (intrinsic) value of the raw material plus the seigniorage fee. It is, furthermore, worth noting that governments in times of fiscal needs debased their coins by lowering the amount of gold or silver, which in the short run, normally, would go unnoticed. But, as the older money with higher gold content was driven out of circulation in the long run, and only the new coins were left, an inflationary process would start. The times of uniform minting would, however, come to end, when the Carolingian Empire collapsed, leading to coins and silver contents varying a throughout Europe. With money markets got increasingly integrated leading to the law of one prices, i.e. that money can be exchanged so that the same silver/gold price ratio prevailed in all markets. This initial integration did, however, not lead to the creation of Fiat money, as as the technology (and the institutions) at the time did not allow for counterfeit proof money.

Early developments hinting towards the need for more sophisticated *money-instrument*, include the promissory notes (i.e. one party promises in writing to pay a specific sum of money to another party at a future date or on demand), and the bill of exchange. The latter served as a written financial instrument, where one party (the drawer) orders another party (the drawee) to pay a specific amount to a third party (the payee) at a certain date. Assume e.g. a coffee trader in Amsterdam, Anna (the remitter) wants to buy coffee beans from a coffee exporter in Rio de Janeiro, Maria (the payee). Instead of shipping gold to Maria, which would be costly and dangerous, Anna pays for the coffee by buying a bill of exchange from a merchant banker in Amsterdam, John (the drawer). John orders a merchant banker in Rio de Janeiro, Pedro (the drawee or acceptor), to pay the specified amount to Maria upon presentation of the bill. Pedro accepts the bill, creating a debt to John and credit to Maria. The system thus represents an early form of credit, as it allows Anna to obtain goods (coffee beans) immediately while delaying the actual payment until a later date when the bill of exchange is settled (typically 2-6 months).

The bill of exchange was used from the thirteenth or fourteenth century until the early century. The risks associated with the debtor or his agent defaulting on their payments was mitigated through the use of branch offices and and correspondents of identical ethnic origin, as penalization was thus easier. During the fifteenth and sixteenth century the transferability and negotiability was, further-

[16] Furthermore, the commodity shall not be perishable and have a low volatility

[17] A pound of silver was divided into 240 pennies (denarii), each containing approximately 1.7g of silver, and later a dozen pennies were called a shilling or sou (solidus); 20 shillings consequently made a pound (libra).

more, improved, making it a liquid asset for many banks as buying, selling and discounting was not possible. With the bill-market emerging, there also came other forms of promissory notes, such as a receipt from a goldsmith or a moneychanger giving a promise to redeem the coins.

The first fiat money was issued by a Swedish bank in 1661, but as it was victim to a bank run in 1668 and hence went bankrupt, England became the leading force in developing note-issuing banks during the 1690s and into the eighteenth century. As note-issuing banks practise so called fractional reserve banking (i.e. they only hold a fraction of their deposits as reserve), the reputation of the bank that issued the bank was now of crucial importance, as the note was only accepted if the public trusted the bank. Furthermore, fractional reserve banking implies that (1) banks are vulnerable to bank runs, as the Swedish case shows, and (2) banks can increase the money supply, thereby contributing to the monetization of the economy. A simply equation for the money-supply-enhancing effect of banks, from a previous macroeconomics course (Mankiw, 2009), is given below;

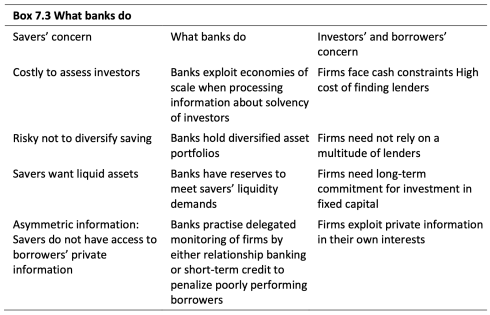
$$\mu = \frac{cr + 1}{cr + rr} \tag{2}$$

Where μ describes the money-supply-multiplier, cr describes the ratio between currency held by the public and demand deposits, and rr describes the ratio of reserves to demand deposits. Assuming $cr = rr = 0,1$ the multiplier is equal to 5,5, increasing as rr falls and decreasing when cr falls.

An important implication of the banking system is, however, the fact that the *game* (following Norths notion of applying game theory on institutional explanations) that each bank is playing has the characteristics of public goods game with a free rider problem. That is due to the fact that each bank (player) benefits from the collective reputation of the banking system (public good) but faces the temptation to deviate from conservative risk-taking to maximize individual profits. This increases the chances of default, potentially imposing a negative externality on the other banks by worsening the collective reputation (the public good). As North (1991) argues, it is difficult to maintain cooperation (which would mitigate the issue) in a game with many players, and an uncertain end date, why institutions such as supervisory agencies and a lender of last resort mechanism emerged. Institutions are thus — in this specific context — to be seen as a *forced* cooperation strategy ensuring stability. The lender of last resort mechanism, however, may lead to a moral hazard, as banks know that they will always be bailed out. Furthermore, both the the lender and the supervising agency needs to be independent, as banks otherwise may have incentives to sustain from helping competition. As governments became accountable, and independent central banks emerge, these tasks could increasingly conducted by government entities leading to a stronger fiat system. The fact that central banks work as lenders of last result is sometimes referred to as Bagehot's rule, which states that during a financial crisis, central banks should lend freely (meaning without limit) to solvent institutions, against good collateral, and at a high interest rate. Bank runs have, however, been remarkably cut down since deposit insurance was introduced, e.g. in 1934 in the US, since customs know that their deposits are safe and there don't *run* on the bank. This does, however, not imply that bank failures are a phenomena of the past, as the crash of 2008 showed. The latter furthermore showed, how a liquidity crisis may lead to a *crises of solvency*, as fire sales affect the prices of assets negatively leading to lover solvency. Solvency crises, generally, have more severe economic consequences and are difficult to manage. They often emerge after periods of excessive risk-taking, which are typically driven by low interest rates, rapid money supply growth, and rising asset prices. A way of assessing weather the economy is *overheating* is to look at the so-called natural interest rate gap (interest rate minus the rate of real GDP growth), which is typically very low in

years before crisis, implying *overheating*.

Finally, as banks emerged, interest rates dropped from 10-20 percent in the twelfth to fourteenth centuries to below 5 per cent by the eighteenth and nineteenth centuries [18]. An overview over the main tasks of a bank, is shown below.



Box 7.3 What banks do		
Savers' concern	What banks do	Investors' and borrowers' concern
Costly to assess investors	Banks exploit economies of scale when processing information about solvency of investors	Firms face cash constraints High cost of finding lenders
Risky not to diversify saving	Banks hold diversified asset portfolios	Firms need not rely on a multitude of lenders
Savers want liquid assets	Banks have reserves to meet savers' liquidity demands	Firms need long-term commitment for investment in fixed capital
Asymmetric information: Savers do not have access to borrowers' private information	Banks practise delegated monitoring of firms by either relationship banking or short-term credit to penalize poorly performing borrowers	Firms exploit private information in their own interests

Figure 4: A banks primary tasks

Note: The figure is found in Persson, et al. 2023.

5.3 The Impact of Banks on Economic Growth

We can summarize the impact of banks on economic growth into three mechanisms: (1) a positive impact on the savings ratio (thereby increasing capital), (2) savings are channelled into the most productive uses and (3) the effect of increased monetization of the economy. The first (1) follows, as bank simply increased the opportunity cost of hoarding, thereby attracting more capital. The second (2) follows from the fact that banks need to monitor their potential debtors, as there is an asymmetry of information. This specialization in information-gathering and monitoring, will lead to borrowers' savings being used more efficiently. Finally (3), higher monetization inherently leads to more economic activity, as e.g. Rousseau (2003) finds that a 1 per cent increase in monetization led to a 2 per cent increase in industrial output after five years, in the Dutch economy during the seventeenth and eighteenth century. Furthermore, there is empirical evidence showing how stock market depth increases investments and productivity in the same way as bank depth. The intuition is simple to follow, as the stock market consists of a large number of uncoordinated individual traders and investment fund managers who do the same job as bank manager (i.e. picking the most promising investments).

Banks, however, emerged first, as they deal with non-marketed assets, and stock markets simply trade firms that are large enough to have an incentive to issue stocks. Within Europe, the importance of banks and stock markets in providing finance and the monitoring of firms has differed historically, and still differs today. Gerschenkron (1962) argues that is due to path dependence, as large banks in Germany were initially effective at financing and thus crowded out parts of the stock market with lasting effects. In the end, any any combination of the two institutional solutions is better than just one, as banks may fail, whereas stock markets, on the other hand, may experience crashes, but still function, when banks crash.

[18] There is, furthermore, an interesting relationship between interest rates and religion, where Christianity increasingly allowed the payment for interest payments, as credit became a bigger part of the economy.

5.4 Course literature on Money, Credit and Banking

Lehmann-Hasemayer and Wahl (2021) study the link between banking and economic growth, by examining the influence savings banks (Sparkassen), on Prussia's industrial development during the early nineteenth century. Using a DD approach, the authors compare city growth and industrialization before and after the establishment of savings banks. To address potential endogeneity (e.g. savings banks emerging in areas with growth potential), they take advantage of a quasi-natural experiment (a 1854 decree) that mandated the creation of at least one savings bank per county and offered financial support to poorer communities. This decree led to the establishment of numerous savings banks between 1854 and 1865, significantly increasing the geographical spread of financial services for exogenous reasons. Using the latter, they estimate the their regression for two sup-samples, firstly (1) excluding savings banks founded before 1838 and secondly (2) excluding savings banks that were founded before the decree was issued and for more than 10 years after it (1865). The paper finds that the presence of a savings bank increases city growth with about 1,5 - 2,2 %, depending on the sample. It, furthermore, led to higher levels of industrialization, as measured by the number of steam engines per factory. The impact was more pronounced in small and medium-sized cities, which benefited from the localized investment strategies of these banks. The paper aligns with other literature from the course, such as Cantoni and Yuchtman (2014), Acemoglu et al. (2005), and Angelucci et al. (2022), which emphasize the role of institutions in economic growth. Finally, the the authors challenge the perspective that banks emerge as a result of economic growth (which we discussed in the seminar classes), as they — with their results — argue that financial institutions are not merely passive entities that emerge in response to economic growth, but rather play an important role in creating the growth itself.

Temin and Voth (2008) investigate the effect of interest rate restrictions on loan allocation (through Banks) on economic growth, by utilizing a quasi natural experiments (the 1714 usury law), which reduced the maximum permissible interest rate from 6% to 5%. Using an OLS regression, they find that the usury limit led to Hoare's Bank (the bank they have data for) simply shifting to more collateralized lending, with the proportion of loans secured by collateral rising from 12.1% before 1714 to 43.0%. Simultaneously, Hoare's favored wealthier and well-connected borrowers. A likely explanation to this shift is the fact that the bank could no longer give out loans to risky clients (i.e. low collateral), as they were not *allowed* to charge their risk premium. To remain profitable in the new market, the bank furthermore shifted towards lending out higher amounts, leading to an increase in average loan size from £640 to £1,259. The paper thus argues, that there is no evidence that the usury laws led a Pareto-improving effect, as the usury laws may simply have led to financial repression (i.e. measures by which governments channel funds to themselves as a form of debt reduction by e.g. capping the interest rate)[19]. Furthermore, the laws may have hindered private-sector-led financial deepening, thus having negative effects on productivity due to diminishing returns to labour.

Hodgson (2016) argues that the financial system played a crucial role leading up to to the Industrial Revolution, more so than the enhancement of property rights, resulting from the Glorious Revolution of 1688. Hodgson highlights the significance of the Financial and Administrative Revolutions that followed, driven by the pressures of war and Britain's expanding global role. He thus argues that the establishment of the Bank of England in 1694, facilitated economic growth by providing new mechanisms for credit and investment. These financial innovations allowed land to be used as

[19] They base this on the fact that government spending — and debt — increased during the period to facilitate war spending. In 1815, the ratio of debt to GDP surpassed 200%, mostly in the form of government debt, while capital in the private sector was scarce.

collateral and expanded the market for negotiable debt, which were critical in financing industrial and infrastructural projects.

6 Trade and Globalization

6.1 The Different Eras of Globalization

In this course, we have defined globalization as market integration (i.e. increasing dependency between domestic and international markets) on a world scale. This implies that factor rewards, and hence prices (in a market economy), will, generally, reflect global rather than local supply and demand. We thus introduce the concept of the moderate *law of one price* for integrated markets, implying that the absolute value of the price difference between identical tradeable goods in two geographically separated markets is equal to or less than the transport and transaction costs (in a broad sense) of moving the commodity from one market to the other. If this is not the case (i.e. the price difference is higher) markets can not be integrated, as profitable arbitrage would be possible. Furthermore, we may measure the degree of globalization as the speed at which a e.g. a local supply shock is *absorbed*, as integrated countries deal at the world price, and hence should experience short lived effects.

Throughout history, we have seen three eras of globalization. The first one starting in the mid-nineteenth century when the trade/income ratio (exports as a share of GDP) rose significantly: from 3–5% to 20–25% for small economies and the UK, and to 10–15% for larger economies. By 1900, markets were thus just as globalized as they are today, with even less restricted labor mobility until 1914. However, historical events in the early twentieth century such as two World Wars, the Great Depression and central planning in Eastern Europe, limited international trade, leading to an anti globalization era. The result throughout the world, were more domestic commodity-, labor- and capital markets, essentially putting globalization levels back to before the first era. Only in 1978's and 1980's — during the second era of globalization — did the capitalist countries return to late nineteenth-century globalization levels. The second era was, in part, made possible due to a round of tariff reductions in the first round of GATT (General Agreement on Tariffs and Trade) in 1947, which an increasing number of nations subsequently joining the negotiations, ultimately leading to the a major breakthrough in the 1960s with the 'Kennedy' round, resulting in substantial tariff cuts on non agriculture goods, averaging about 35% (Riedel, 1976) . Additionally, the European Economic Community (later the EU) was formed reducing internal trade barriers and increasing intra-European trade, while imposing tariff barriers on countries outside the union. The former socialist countries saw a development towards increased market integration in the 1990's. Since then, globalization has, generally, increased, with sanctions against Russia and exogenous shocks such as COVID, being the exception. Tariffs on goods have thus,on world average, fallen from 26 in 1986 % to 8,8% in 2007, while trade has grown twice as fast on average, as world output. Furthermore, developing countries have almost 17 doubled their share of exports, to 37 per cent in 2007, which is in part due to the transformation of the former socialist countries in Europe, but also from the opening and market liberalization of China that was marked with their access to WTO in 2001, even though their trade already starting rising in the 1990's. While it remains uncertain, to which extend Indian growth will match Chinese levels, there seems to be a potential for increased world trade with India.

In mid-seventeenth century Europe, we see the first signs of regional price convergence for ordinary goods, such as grain, while price convergence for long-haul trade between US and Europe didn't occur until the nineteenth century, where European market integration likewise spored. During the

twentieth century, prices seem to have diverged, as transport costs stagnated and tariffs rose on e.g. agriculture, whereas the began to converge again with the second era of globalization, post WW2. From the former we can thus derive the two main determinants for globalization: technology and politics.

Technological factors contributing to globalization often reduce transaction- and transport costs, while also allowing information to flow faster, and more freely. Politics that promote globalization have a positive effect, whereas mercantilism (i.e. a strong focus a surplus on the current account) harms globalization and disintegrates markets, as seen in the the interwar period, with a strong peak during the Great Depression [20]. Some degree of protectionism, in line with the infant industry argument, can, however, increase globalization in the long run, as it can allow for larger industries to form (e.g. dairy production in Denmark, see Henriksen et. al 2012).

6.2 The Welfare Implications of Trade

The traditional view on international trade was first put forward by David Ricardo, who argued that trading allowing countries to leverage their comparative advantages, thus increasing the aggregate production possibility frontier. By trading at the goods' relative prices, the trading partners can thus enjoy more goods and thus higher aggregate welfare [21]. His theorem has later been developed further by e.g. Hekscher and Olin (1993) , who argued that comparative advantages are based on the relative abundance of factors of production, rather than on differences in labour productivity. They argue that this will lead to a factor and commodity prices converge, as factors of production move to where they are most efficiently used and goods are traded based on comparative advantage. However, the wage dimension of the claim has found little empirical support. Another important elaboration of the Ricardian view is the Rybczynski theorem that states that the production of goods produced intensively using the factor, where the country has a comparative advantage, will also increase. Empirical support of this claim is found historically, as economies in times of mass migration have increase their supply of labour intensive goods, as they suddenly had a comparative advantage in labour. Finally, the Stolper-Samuelson theorem adds that the relative prices of factors are driven by the relative prices of output goods. As Stolper and Samuelson found that the relative prices of export goods will increase, while the relative prices of import goods will decrease, we see how the group owning the export factor will be better off, while the group owning the import factor will experience a welfare loss, as their relative *rent* on either labour or capital will fall. If the losing group is strong enough, it may be able to restrict trade in certain areas, even though the net effect would have been welfare increasing (e.g. if the winning group consists of more, less powerful people).

However, what the above mentioned theories seem to underestimate, is the fact that the industry in which countries chose to specialize in, has large effects on demographic and educational development, leading to a lasting divergence between rich and poor countries (Galor and Mountford, 2008), as outlined in the figure below.

Consider e.g. the trade between Sub Saharan Africa, where the comparative factor advantage (using the Hekscher and Olin framework (1993)) is in unskilled labour, and Germany, where the comparative factor advantage lies in capital. In accordance to the framework outlined above, Africa will export goods in sectors relying on unskilled labour, and import goods from capital intensive sector (vice versa for Germany). While this, in the short run, increases welfare, Galor and Mountford (2008)

[20] The Smoot-Hawley Act (1930) marked a shift in American trade policy and was an important force in disintegrating markets, as it increased tariffs to such a degree that any country pursuing an independent free trade policy would be very unlikely to experience a beneficial impact on growth, partly due to trade volume falling drastically.

[21] It does, however, not necessarily say something about growth as such, as trade theory does not concern the matter of growth rates.

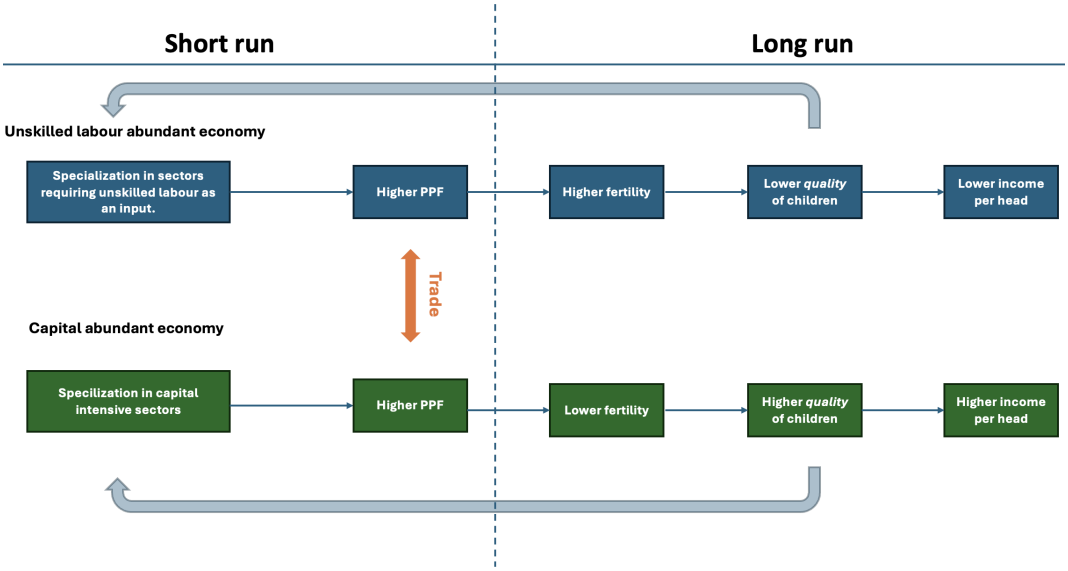


Figure 5: Static short run and dynamic long run effects of trade

Note: Own illustration based on arguments in Galor and Mountford (2008)

argue that it keeps them poverty in the long run. This is due to the fact, that the trade alters the children quantity-quality tradeoff for families in such a way, that quantity becomes relatively more attractive, leading to higher birth rates and lower educational levels — a modern Malthusian force — similarly to empirical reality in e.g. African developing countries. Galor and Mountford (2008) thus find empirical evidence showing that that trade positively affects fertility rates and negatively affects education in non-OECD economies, while the opposite is true for OECD economies. It, furthermore, seems reasonable to assume some sort of positive feedback loop, i.e. that the initial impacts of trade on human capital and fertility further reinforce the specialization patterns and comparative advantages of these economies. The result is that a cycle of higher human capital investment and economic growth in developed countries and continued population growth and lower human capital investment in less-developed countries. This suggests that trade, even though it is traditionally thought of as a unifying force, spreading knowledge (a non rival good) and thus leading to convergence (under the right institutional circumstances), may in fact lead to divergence, given that we don't see initial factor advantages as an institution. It is not clear, weather Asia serves as a counterexample, as they *did* specialize in utilizing their abundant low-skilled labour, but *did not* do it in stagnating industries, and rather chose capital intensive sectors such as manufacturing.

Intra-European trade, however, often remain intra-sectional. During the course we have partly explained this with Paul Krugman's New Trade theory, that states that trade does not need to be based on comparative advantage if there are economies of scale, resulting in monopolistic competition (i.e. a situation where a number of firms have monopolies in a differentiated good, but where monopoly profits are competed away as additional firms, each supplying other varieties of the product, enter the market). If the market for the good is large enough (e.g. cars), free trade will allow for cheaper goods and more variety, which consumers value, thus increasing intra-sectional trade.

6.3 The Infant Industry Argument for protectionism

The infant industry argument postulates that emerging domestic industries require protection from international competition until they have matured and become competitive. The assumption behind

the argument concern economies of scale, and goes as follows; if an industry is protected at an early stage, where it is not competitive at the world market, it can increase its production and thus lower unit costs, until it eventually can compete at the world market. If the protection is, however, not offered, the industry will not be able to *mature*, as it will be out-competed by international competition that already has economies of scale. The argument, furthermore, assumes some sort of market failure, as a company with true potential would be able to attract private investors to allow for economies of scale through capital investments. Especially in emerging countries, where investors are uncertain of their return, and thus might not be prepared to invest, the assumptions seems plausible. However, while the argument has indeed allowed industries to mature, and remain competitive afterwards (e.g. Henriksen et al. 2012, and Juhász, 2018), protecting incorrectly will lead to industries being dependent on government and thus not becoming comparative. Finally, it is worth noting that the vast amount of tariffs historically have prevailed due to either power from the *losing* industry or a political focus on mercantilism (e.g. surplus on the current account.

Research has fund that the structure of protection which matters, rather than protection itself, matters for growth. Thus industrial protection might be positively correlated with growth, as it fosters industrialization (and thus shifts the quantity-quality tradeoff towards quality, improving education) and takes advantage of economies of scale, whereas agricultural tariffs (such as those of the EU) might harm growth, if they protect an uncompetitive agricultural sector (like in Denmark), while possibly also shifting the quantity-quality tradeoff towards quantity (in developing countries). On the other hand, Pascali (2017) argues that increased trade in the late nineteenth century had a positive effect on growth in countries with a small number of countries with strong and 'inclusive' institutions, but not countries with a lack of the former.

6.4 Course Literature on Trade and Globalization

In their paper, Henriksen et al. (2012), provide empirical support for the infant industry argument by studying the growth of the Danish dairy industry in the late nineteenth century. Using micro-level data from individual dairies, they find that the protectionist policy on agriculture (i.e. import taxes on beef, pork, butter & cheese) provided danish diaries with an implicit subsidy, accounting to about 10% of total income per cow, before 1864. After 1864 protection was removed from beef, pork and butter, however the implicit subsidy on cheese remained. However, even after the subsidies on cheese were removed, the positive results lasted, as Denmark e.g. increased their butter exports to the UK with 25 pp, and their bacon exports with 49 pp, from 1870 to the beginning of the twentieth century (seen from the British side). The article thus challenges the widespread view that Denmark was a liberal *bastion* in the late nineteenth century globalization, as it argues that the protectionist measures played a central role in fostering a diary sector that could compete later on the international market. However, they add that other institutional factors played a central role aswell, such as land reforms, peasant emancipation, a high level of farmer education, the absence of grain tariffs allowing for cheap animal feed, and public support for research and research institutions. They thus highlight how the infant industry argument may be dependent on certain institutional frameworks being in place, prior to the protection.

Juhász (2018) — following e.g. Henriksen et al. 2012 — provides empirical support for the infant industry argument. By utilizing a quasi natural experiment during the Napoleonic Wars (1803-1815), where some regions were more effectively protected from British trade due to exogenous blockades, she finds that the protected regions saw a significant short term increase in mechanized cotton spin-

ning capacity, compared to regions that remained more exposed. She, furthermore, finds that the protected regions were able to develop higher long-term industrial activity and value-added per capita in the industry, going beyond the protected period and lasting until the second half of the nineteenth century. However, Juhász also highlights the importance of other institutional factors, such as access to markets, availability of human capital, and regional economic conditions, which significantly contributed to the industry's success. She thus argues that, if these *underlying conditions are not in place, infant industry protection can turn out to be an extremely blunt tool* (Juhász, 2018).

Acemogulu et al. (2005) examine the impact of Atlantic trade on institutional development and economic growth in Western Europe between 1500-1800. Using a DD-model they compare the urbanization rate (as a proxy for GDP per capita) [22], between countries that were included in the Atlantic trade, and those that weren't, before and after the rise of Atlantic trade. However, as being a part of Atlantic trade is clearly endogenous, they use the Atlantic coastline-to-area ratio as a measure of potential access to the Atlantic trade, as this is clearly exogenous. They argue that Atlantic trade contributed to European growth through a interplay two different effects, in countries that met to preconditions: easy access to the Atlantic and nonabsolutist initial institutions. Firstly (1), it created large profits that enriched and empowered commercial interests. These profits gave the Monarch relatively less power, leading to reforms in political institutions, such as more secure property rights — one of three preconditions for a modern economy. The process was particularly effective in Britain and the Netherlands, where initial non-absolutist institutions facilitated greater economic benefits from trade, whereas countries with more absolutist regimes, like Spain and Portugal, saw limited institutional change and economic growth. Finally, the study implies a positive feedback loop between trade, institutional development, and economic growth. As institutions improved and protected property rights, economic activity increased, essentially creating a more complex *game*, thus demanding better *rules* (institutions) to maintain an efficient outcome.

In her paper, Waldinger (2022), analyses the economic effects of long-term climate change by studying the relationship between temperature changes during the Little Ice Age (1600-1850), and city size. Using a panel data set for 2,120 European cities, she finds that the cooler temperatures during this period had substantial negative effects on economic activity (city size), primarily through reduced agricultural productivity and increased mortality. Specifically, a 1°C decrease in long-term mean temperature decreased city size by around 70% [23]. She, furthermore, finds that increased trade played a central role in mitigating the negative effects, as cities with greater access to long-distance maritime trade and larger inland trade opportunities experienced smaller negative impacts from the temperature decline. The main reason being that trade allowed these cities to import e.g. food or crops that became scarce due to reduced agricultural yields. Waldinger thus adds a geopolitical dimension to our analysis of trade, as she shows how trading essential goods - and the lack thereof - has severe consequences for economic growth. As a contemporary example of how this dependency increased fragility, we may consider the African dependency on Ukrainian food supplies, that led to severe consequence when trade was disrupted due to Russia's attack on Ukraine.

Pascali (2017) estimates the effect of the introduction a new technology (the steamship [24]) on economic

[22] They later validate their results with Maddison GDP per capita data.

[23] It is important to note that long-term changes in temperature are substantially smaller than the difference in temperature between the warmest and the coldest year. The city with the largest decrease in long-term temperature thus experienced a decrease of 0,287 1°C, i.e. a city size reduction of about 20 pct. over a 100 year period (sixteenth to seventeenth century)

[24] The steamship is specifically used, as it had a significant impact on global trade patterns, with shipping times by steam becoming a dominant factor influencing trade flows post-1865

growth during the first era of trade globalization (1870 - 1913). Using a DD approach, he compares the changes in trade and economic outcomes between countries before and after the introduction of the steamship, using the asymmetric reduction in shipping times across countries as an exogenous source of variation. Based on a rough calculation, Pascali suggests that the introduction of the steamship might have accounted for at least half of the increase in international trade during the second half of the nineteenth century. However, he argues, that the increased trade has only benefited a small number of countries, namely those with strong institutions. He thus finds that countries with strong constraints on executive power and better investment environments were more likely to benefit from trade. The fact that institutions are an important factor for growth, is in line with other literature from the course (i.e. Cantoni and Yuchtman (2014), Acemogulu et al. (2005) and Angelucci et al. (2022)). Finally, the paper challenges the traditional Ricardian view, which argues that trade benefits all participating countries equally, and instead shows empirical results highlighting that technological advancements, such as the steamship, can lead to significant economic divergence between countries.

In his paper, Autor (2018), explores the impact of China's rapid economic rise, commonly referred to as the "China Shock," on labor markets in developed countries. Autor finds that regions in the United States heavily exposed to import competition from China experienced substantial declines in manufacturing employment and wages, particularly affecting non-college-educated workers. The paper thus empirically shows that while international trade raises GDP in all participating countries (i.e. it lives up to the Kaldor-Hicks efficiency criterion [25]), the benefits (i.e. lower prices) are broadly distributed, whereas the costs for the out-competed sectors are highly concentrated among specific worker groups and locations. Autor therefore argues that policy interventions should aim to share the benefits of trade so that the winners to a larger degree compensate the losers, through e.g. trade adjustment programs and wage insurance.

7 Globalization of Labour Markets, Immigration and Monetary Regimes

7.1 Capital flows

Defining an international monetary system as a set of rules (broadly speaking), that regulates the flow international capital (e.g. the exchange of currencies).It allows for international investment and capital markets to form, as international trade, without such a system, would be reduced to balanced balanced bilateral trade. To see why, consider the national identity, stating that $Y = C+I+G+NX$, where $S \equiv Y-C-G$. Substituting and rearranging we get $NX = S-I$, implying that current account deficit (surplus) leads to the country borrowing (investing) capital abroad. However, if there is no system regulating the flow of such investment, and it is e.g. not possible to exchange currencies, trade is balanced and $NX = 0$, leading to $S = I$, where domestic savings equal domestic investments, i.e. there being to international capital markets. The latter is economically inefficient because it restricts capital from being used where its most efficient. The downside to this being that history provides plenty of examples of governments who borrow instead of taxing their citizens, possibly leading to them defaulting on their debt.

Another consequence of an international monetary system, is the interest rate parity, describing how globalization of capital markets imply a convergence of interest rates. However, this does not mean, that the domestic interest rates are equal to world interest rates, as buying bonds in a

[25] This efficiency criterion states that an outcome is an improvement if those that are made better off could in principle compensate those that are made worse off, i.e. a Pareto improving outcome could be achieved.

certain countries requires you to invest in their currency, and the currency can fluctuate. Following Blanchards (2021) notation we might expect an interest rate parity of the form;

$$r = r^* + (e - \bar{e}^e_{+1}) \tag{3}$$

where r represents the domestic interest rate, r^* represents the foreign interest rate, e represents the current exchange rate, and $\bar{e}^e_{+1}$ represents the expected future exchange rate (which we assume to be constant). Note that even though this gives a more nuanced view at the interest rate parity, it still does not account for differences in risk between government bonds and the risk premium associated with that. Thus, capital market integration does not necessarily imply $r = r^*$, however if the risk in two countries is the same and the currency does not fluctuate, it can be expected.

As with commodity markets, capital market integration depends on the technology and institutions, however since the telegraph enabled instant information transmission, capital market integration has been primarily driven by institutional factors, particularly capital mobility restrictions. However, we furthermore need to add that there is a home bias in investment [26], leading to domestic saving and investment remaining closely linked. Empirical work by Obstfeld and Taylor (2010) shows that investors were more globally oriented pre-1930, whereas the Great Depression strengthened the domestic saving-investment link. The link was been gradually weakened, leading to about 25% of additional savings now going abroad, compared to over 40% before the Great Depression.

7.2 The Two Eras of Capital Market Integration

Generally speaking, we may divide history into two distinct periods of capital market integration. The first (1) one being associated with the International Gold Standard, implying unregulated capital markets in the fifty years before 1914, and the second (2) starting after the breakdown of the the Bretton Woods System in the 1970s, which placed strict restrictions on capital mobility. Empirically, we — not surprisingly — find that the efficiency of capital markets [27] is positively related with the *openness* of the international monetary system. We thus find the same U-shaped pattern for capital markets, that we found for commodity markets; one period from about 1870 to 1914, with efficient capital markets, the Interwar years with low capital mobility, followed by the Bretton Woods System, which was ultimately followed by the free capital mobility that is still prevalent today.

When choosing an international monetary regime, policy makers are confronted with the so called *open economy trilemma*, as describes by Obstfeld and Taylor, stating that policy makers not both have a fixed exchange rate, free capital flows and independent monetary policy. They are restricted to two choices, as shown below.
To see why this is the case, consider an economy with fixed exchange rates, and unrestricted capital mobility. This economy can not have monetary autonomy, as monetary policy is *reserved* to fix the exchange rate. If that would not be the case, and the government in the given countries raises the interest rates, demand after the currency would increase, putting an upward pressure on the exchange rate, which is not allowed and can now only be countered by giving up e.g. unrestricted capital mobility and simply stopping capital from *entering* the market.

The policy makers must thus choose between fixed exchange rates, which (if credible) may increase trade as unexpected prices changes don't occur, monetary autonomy, which gives the option to use fiscal policy to stimulate the economy, and unrestricted capital mobility, which likewise increases

[26] Probably due to information asymmetries, giving national investors an advantage.
[27] Measured by price differences between similar assets, that would be minimal with free capital flows, due to arbitrage

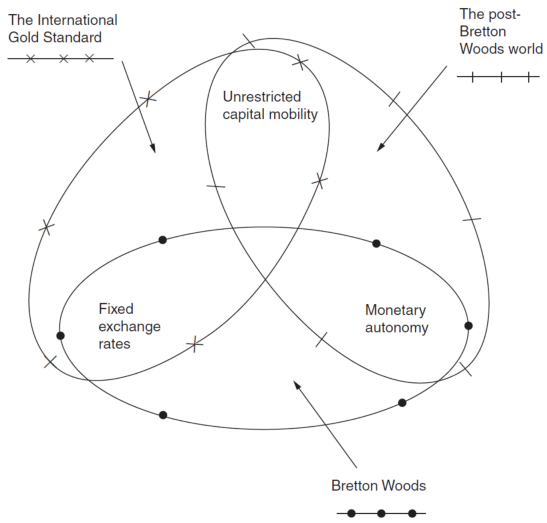


Figure 6: The open economy trilemma visualized

Note: Illustration taken from Persson et al., 2023

trade. Historically, different combinations have been used, as seen on the figure, which will be described further in the next section.

7.3 The International Gold Standard, 1870–1914

The gold standard has its origin in Britain's Resumption Act of 1819, which institutionalized the practice of exchanging currency notes for gold on demand at a fixed rate. During that time, other countries practiced bimetallism, where their currency was pegged to both gold and silver, as gold seemed to valuable. However, as Britain emerged as an economic power, other countries followed her lead, with the US finally abandoning bimetallism in 1873. During the time there, generally was a spirit towards internationalism with the emergence of the Scandinavian Monetary Union from 1875 and the Latin Monetary Union (LMU) from 1865 between a number of European countries, including France, Belgium, Italy, Switzerland, Spain. However, even though these unions were political significant, they played a limited role in a world of a gold standard, as the latter itself implied a fixed exchange rate. To see why, consider the *rules of the game* of the gold standard that emerged through practice (rather than institutional structure) of individual countries.

1. The currency should be freely convertible to gold at a set price or 'mint parity'
2. There should be no barriers to capital and gold flows between countries.
3. Currency must be convertible to gold on demand, backed by gold reserves.

As capital could flow freely (2) and be converted on demand (3), differences in gold prices set by (1) would be minimal, as arbitrage would effectively exchange currencies for gold and afterwards the gold to currency in a different country (or vice versa), until fixed exchange rates are again resorted. However, we may add that *quasi fixed* may be a better terminology, as small differences could emerge due to costs associated with shipping gold and the associated transaction costs.

The rationale for the gold standard was largely based on a laissez faire attitude economic policy from the government, why they did not see it as an issue that their monetary autonomy was being

limited by the system. However, while Hume (1752) argued that the gold standard served as an equilibrium mechanism, economically, central banks were often worried about gold losses. They thus often sterilized gold inflows (i.e. preventing them from entering the money supply), which was direct contravention of the rules of the game. The rationale behind this was to avoid a *bank run*, where large amounts of savers would convert their money into gold and export it. In theory this outflow should have been been met with a reduction in the money supply by selling domestic assets, and thus increasing interest rates, leading to higher inflows and a higher money supply. However, e.g. the Bank of England instead lent freely to banks at higher interest rates, also leading to an expansion of the money supply, however without the initial reduction in the domestic money supply. While actions like these were risky, as it e.g. may have induced excessive risk-taking by banks, the gold standard lasted for decades. The reasons for this are often summarized as commitment (1), confidence (2) and symmetry (3). Commitment implied that if a country faced economic difficulties and temporarily devalued its currency to increase the money supply, and increase exports, it would later revalue the currency back to its original gold parity once conditions improved. Confidence (2) meant that people (including financial investors) actually believed that a currency pegged to gold would maintain its value, thus preventing destabilizing speculation. Finally, (3) symmetry implied that no single country had disproportionate control over global price levels, as these were were determined by the global supply and demand for gold. If one country e.g. experienced inflation, it would eventually result in gold outflows, which would reduce the money supply and bring prices down again, and vice versa for deflation. However, when a new method was found for extracting gold from ore, prices increased with the new supply. This would lead to shorter periods with uniform rates of inflation in the world economy. Generally, countries turned a blind eye as they did not have to focus on combating unemployment, however as the US was noticeably more democratic than many other countries, and domestic issues played a larger role in political debate.

The beginning of the end for the gold standard was marked by the First World War, where the gold standard was suspended, as government expenditure needed to be financed by expanding the monetary supply through issuing bonds and in some cases printing money. When the war was over, some countries continued to finance post war-related costs with printing more money, leading to inflation, sometimes on a massive scale, as in Germany where hyperinflation reached the staggering heights of 3,25 million pct. per month. Governments, however, viewed the world before WW1 as a golden age, and wanted to return to the gold standard. Led by the US, who were relatively unscathed by the war, and came back on the gold standard in 1919. According to the rules of the game, all countries should have returned to gold at the original parities, but inflation during the war would require deflation in many countries, as the increased prices and expanded money supply caused by wartime spending would need to be reversed to restore the pre-war gold value of their currencies. This deflationary process would involve reducing the money supply and lowering price levels, often through austerity measures and tight monetary policies, which would e.g. increase unemployment. As democracy spread, this was no longer possible, and some countries, such as France, returned at lower parities. The UK government, however, wished to wished to restore confidence in the gold standard (that they build and managed) and thus accepted deflation, with increased unemployment being the price.

Around 1929 both France and the USA began sterilizing gold inflows to reduce domestic inflation as the large inflows of gold threatened to expand their money supplies and increase price levels. They thus absorbed the world's gold, to the extent that they ended up holding 70 per cent of the world's supply of gold. This meant that other countries were forced to restrict their money supply. The worldwide monetary contraction and the Wall Street Crash in 1929 contributed to the Great

Depression of the 1930s, which had a worldwide effect through unemployment and bank failures. As Barry Eichengreen (1996) argues, the gold standard, contributed significantly to the prolongation and worsening of the Great Depression, as countries were reluctant to provide necessary liquidity, prioritizing the protection of their gold reserves instead. Moreover, countries that were previously been forced to rely on US lending, were now forced to restrict capital mobility, as these loans could not be given anymore, leading to a widespread use of capital and currency controls, that resulted in a decline in world trade in the 1930s. Eichengreen and Douglas A. (2010) show that the latter was not due to a general rush to raise trade barriers during the Great Depression. Instead, it was primarily a result of the few countries that remained on the gold standard for a longer period and increasingly resorted to measures like tariffs, import quotas, and exchange controls. The constraints imposed by the gold standard thus significantly influenced the trade policies of these countries.

In 1931, the UK finally faced a severe outflow of gold, which it could not control, forcing them to leave the gold standard. Other countries followed, including the US in 1993, and empirical work has shown that nations leaving gold early and devaluing to a fraction of their 1929 exchange rate recovered faster from the Great Depression, as they were not bound to an overvalued currency. Furthermore, devaluation led to increased inflation, which lowered real wages, and real interest rates since nominal wages remained stable while devaluation allowed producers to raise prices slightly. Finally, monetary policy could now be used to stimulate the economy, further fostering speedy recovery. The opposite was true for France, who did not leave the gold standard until 1936, causing them to miss the revival of the international economy in the mid 1930s. Concluding, gold standard collapsed, in part because an era emerged, where monetary policy became critically important for managing economies. The strong preference towards a fixed rate, however, survived, ultimately leading to policy makers giving up free capital mobility in accordance to the trilemma, leading to the so-called Bretton Woods System.

7.4 The Bretton Woods System, 1944-1971

As world leaders began to recognize the socioeconomic consequences of the protectionism that emerged after the Great Depression — such as higher unemployment, a decline in global trade, and the rise of fascism — the need for a new monetary system emerge. In July 1944, 44 countries thus signed the Articles of Agreement of the International Monetary Fund (IMF), designing a monetary system, which allowed for a fixed exchange rate and independent monetary policy, giving up free capital mobility. It was agreed, that;

1. The dollar was fixed against the price of gold: $35 an ounce.
2. Member countries hold reserves in gold or dollar assets, with the right to sell dollars to the US Federal Reserve for gold at the official price. These reserves were to guard against short run swings in the exchange rates, with no link to the monetary supply as they had under the gold standard.
3. All currencies were fixed in value against the dollar, giving $n-1$ exchange rates [28]

The system effectively implied that all currencies were pegged to the dollar, which was in turn pegged to gold. Thus, the system was similar to the gold standard, with the exception that the US was now solely responsible for maintaining the dollar price of gold. This allowed other countries to use monetary policy to stimulate their economy, as they did not face the constraints of maintaining their

[28] This follows from the fact that all currencies are fixed against each other, why you only need $n-1$ exchange rates, as the last exchange rate can be determined based on the others.

own gold reserves. Furthermore, the system had incorporated two mechanisms to ensure flexibility, in case member countries were experiencing currency account deficits. Firstly (1), the IMF would control a pool of gold and currencies from member countries, which it could lend to members who were experiencing current account deficits, but where contractionary policy would cause unemployment. Members who borrowed from the IMF would be supervised by the IMF. Secondly (2), parities (the value against the dollar) was adjustable, if the balance of payments (i.e. the record over aggregate financial transactions made by residents of a country) was in *fundamental disequilibrium*, which was not further defined. The UK and France e.g. chose to devalue, as their inflation was higher than other countries. This saved them from using deflation to bring back the fixed exchange rate, which is consider to have more severe domestic economic consequences. However, the fact that this was possible caused instability for the system, as financial speculators began to bypass the capital mobility by either borrowing from abroad by delaying payments for goods (i.e., speculating on currency depreciation), or lending by forwarding money in advance (i.e. speculating on currency appreciation). Downward speculative pressure would thus emerge if countries had large and persistent current account deficits, as a devaluation seemed plausible due to the country being in a *fundamental disequilibrium*. Such speculative pressure was somewhat unique, as financial investors during the gold standard held confidence in the fact that a currency pegged to gold would maintain its value. Furthermore, there two fundamental asymmetries in the system, as (1) the dollar could not be devalued and (2) the US was the only country that was able to set an independent interest rate and use monetary policy. This meant that other countries had to follow US monetary policy, including their inflation. This was acceptable, until the 1960s democratic administrations expanded welfare spending and got involved in the Vietnam War, leading to budget deficits, expansionary monetary policy and inflation doubling. As the European countries did not want to import the high inflation, they revaluated, which solved the issue temporally. However, as the US became more and more overvalued they themselves devalued against gold in 1971, and later abandoned gold convertibility in 1973. The latter lead to the system collapsing, as trust was fully lost. We may thus summarize the failure as a lack of commitment, confidence and symmetry.

7.5 Floating Exchange Rates

Today, floating exchange rates dominate, and even though they emerged by *accident* and initially were sen as temporary measures, it seems unlikely that the world will return to a worldwide fixed exchange rate regime. The reason for this is twofold. Firstly (1), theorist such as Mundell have argue, that fixed exchange rate regimes that are very diverse in their membership (such as a worldwide regime) have little chance of surviving because economic shocks are asymmetric (partly due to the difference level of institutions) and political preferences differ across countries. Secondly (2), countries have simply lost interest in a worldwide international fixed exchange rate system, since floating exchange rates are efficient and compatible with both free capital and trade flows and allow for monetary autonomy. However, that does not rule out the possibility of fixed exchange rates at all. In fact, Mundells (1991) theory of Optimum Currency Areas (OCA) argues that regions or countries with high labor and capital mobility, price and wage flexibility, similar business cycles, and a degree of political integration have the prerequisites to successfully share a common currency or maintain fixed exchange rates. An example of such an area is given by the European Union, which employs a common currency, the most extreme form of fixed exchange rates, or Denmark, where the local currency is pegged to the euro. Critics will, however, point out that it is far from obvious that the Eurozone constitutes an optimal currency area (OCA), as member countries have found it difficult to coordinate fiscal policy, both before and after the financial crisis. Furthermore, labour mobility

restrictions and the asymmetric consequences of economic shocks such as the crisis in 2008, highlight the difficulties of the EU as an OCA. With reference to the euro crisis, affecting countries in southern Europe, Krugman argues for the so called *euro trap*, where countries within the Eurozone experienced high prices and wages after periods of low interest rates, and instead of devaluing (under a floating exchange rate), to remain competitive, had to deflate their economies, through austerity measures, leading to economic hardship and higher unemployment.

7.6 Labour Market Globalization and Migration

Generally speaking, the convergence mechanism on labour markets, is much weaker than on both commodity- and capital markets. There are, however, forces which induce convergence of real wages. Firstly (1), differences in wage levels and labour supply, may lead migration from countries with low wages and excess labour, to countries with high wages and excess demand. However, there are transportation costs, political boundaries and a home bias (i.e. workers prefer to work in their *own* country), that challenge the mechanism described above. Nevertheless, the first wave of mass migration started after the disturbing harvest shocks of the 1840s, at about 300,000 per year, progressively increasing until WW1. Initially, migration was coming from Isles, Germany and Scandinavia, however around 1900-1910, migrants from eastern and southern Europe became the majority. At that time, the transportation costs and travel time, had even fallen to a level, where seasonal migration was both possible and attractive. The extent of globalization in labour markets from before the First World War was, however, never reached again. Reasons include political barriers, such as a literacy test in the US (1920's), and the fact that real wages converged due to the earlier migration and a fall in grain prices, lowering prices in Europe and thus increasing real wages. Taylor and Williamson (1997) further argue, a considerable amount of the economic convergence in real wages and labour productivity between 1870 and 1913 was due to mass migration, as labour would travel to where it was most productive. They further argue, that the the lack of convergence in the interwar period might be due to the imposition of quotas and other barriers to migration. It is, however, adding that while this is true for a number of countries, who already trade, it seems ambiguous to assume that the same goes for a wider sample of economies including developing economies, as there are simply limits to what migration can do to cause wage convergence. The latter is true, as wages (in a market economy) are determined by the marginal product of labour, which, in turn is determined by e.g. the technological level, which, in turn, is partly depended on the institutions of a country. As developing countries, generally speaking, have poorer institutions, the technology level is lower, thus *naturally* leading to a lower wage, no matter how much migration will flow [29].

Generally speaking, migration has the potential to improve the aggregate demand in an economy, while simultaneously increasing technological through knowledge transfer. However, the causal link between immigration and technological progress and growth is difficult to measure empirically, since there is a twofold endogeneity concern. Firstly (1), it is difficult to asses whether immigration directly causes technological advancement or if countries with advanced technology naturally attract more immigrants. Secondly, (2) immigration policies and economic conditions that attract skilled migrants might also independently increase technological progress. Furthermore, the subjectivity in choosing a time horizon may lead to different results across studies, as some of the benefits (or disadvantages) might come with a lack. Finally, the impacts of immigration can vary across different sectors and regions, making it hard to generalize findings.

[29] Of course, migration also brings technological level with it, however it seems ambiguous to assume that it can outweigh the initial lower level of institutions in the receiving country.

Nevertheless, there are plenty of empirical studies finding a positive relationship between immigration, and technology and growth. Sequeira et al. (2019) e.g. finds that locations in the US with more historical immigration have higher incomes, less poverty and less unemployment, whereas both Akcigit et al. (2017) and Moser et al. (2014) find that immigration fosters innovation, measured by patents) i both the US (Akcigit et al. (2017)) and Germany (Moser et al. (2014)). Finally, Hornung (2014) and Boberg-Fazlic and Sharp (2024) find long-term positive effects of both economically *desirable* (Hornung, 2014) and non *desirable* (Fazlic and Sharp, 2024) migration on textile manufacturing in Prussia and dairy production in the US, respectively. These studies all stand in contrast to the political landscape, where immigration, and the clear challenges it presents, is often seen as a burden, rather than a growth opportunity. The fact that most west-European countries face labour shortage in parts of their public sector, makes the latter even less understandable. Finally, labour market integration reduces the bargaining power of employees, as prices are dictated by the world market, and wages over the latter either encourages immigration or the import of goods with lower labour costs.

While some argue that integration of labour markets leads to a worsening of working conditions in Europe, the opposite was observed under the first era of globalization. Huberman and Lewchuk (2003) thus argue for a *race to the top* during the period, implying that labour protection became more advanced, as economies opened up. Rodrik (1998) finds the same relationship for a larger sample of economies towards the end of the twentieth century. The intuition behind the phenomena is the fact that open economies are more vulnerable to shocks from the international economy, thus having a larger demand for a tight safety net (e.g. wage insurance). To limit the use of the latter, some countries have established formal or informal rules requiring trade-exposed sectors to set the standard wage increases. This is done to prevent high public sector wage increases from putting upward pressure on wages in exposed sectors, leading to a loss of competition. However, this model only works, as long as the wage negotiations are, to some extent, centrally guided.

7.7 Course Literature on Labour Markets, Immigration and Monetary Regimes

Eichengreen and Irwin (2010) show that countries adhering to the gold standard were more likely to resort to protectionist trade policies such as tariffs, import quotas, and exchange controls, after the Great Depression. The reason for this is twofold. Firstly (1), as they were unable to use monetary policy [30] to counter the recession, they sought to protect their balance of payments and domestic economies by restricting trade. Secondly (2), as they aimed to keep their peg to the gold as their economic activity fell, the needed to counteract the falling demand for their currency fell, putting a downward pressure on the currency. As they began to devalue, letting the downward pressure *work*, France and Switzerland, part of the gold bloc, thus were able to reduce their tariffs by between 15-20% and 50%, respectively. In contrast, countries like the UK, which left the gold standard in 1931, saw less dramatic increases (and changes) in tariffs. Based on the former, the paper goes on to argue that the rigid adherence to the gold standard prolonged the global economic crisis by preventing countries from adopting more flexible and effective economic policies. They show that countries remaining on the gold standard experienced deeper and more severe economic backlashes, in part due to their protection. Conversely, countries that devalued their currencies saw a quicker recovery and less need for trade restrictions. The paper thus highlights an essential tradeoff that policymakers faced during the Great Depression: (1) abandoning the gold standard, (2) restricting trade to maintain the fixed

[30] They chose fixed exchange rates and free capital flow, effectively giving up independent monetary policy.

exchange rate, (3) devaluing the currency, or (4) creating deflation to revert to the former exchange rate. The fourth option, leading to austerity measures, was generally unacceptable to policymakers, reducing the choices to three, that — on a European scale — all came into play, with varying success, as shown by the former.

Boberg-Fazlić and Sharp (2024) estimate the causal effect of established immigration communities on economic growth, by studying how economically non *desirable* Danish-Americans (e.g. low-skilled farmers seeking better opportunities abroad) immigrants promoted growth in the American dairy industry through knowledge spillovers. Using a DD model, they compare changes over time in e.g. dairy cows between counties with high concentrations of Danish immigrants and those with fewer or no Danish immigrants. They find a positive relationship, arguing that the rapid development of Denmark's dairy industry from around 1880, increased the knowledge spillovers of the Danish immigrants, leading to spillover effects in the US. The paper also explore the mechanisms behind this knowledge transfer, and finds that danish newspapers (technology) in the US played an important role in connecting the danish communities in the US. Furthermore, Danish emigration archives show that a significant number of Danish-Americans traveled back to Denmark, facilitating the flow of information and technology. Their study thus shows, how (1) small, existing, immigrant communities may yield long term benefits by through knowledge spillovers, even though (2) they, at the time of arrival, are by no means economically *desirable*, as they e.g. lacked education, and/or wealth.

Hornung (2014) estimates the causal effect of skilled-worker immigration on long-term productivity by analyzing the Huguenot migration to Prussia. Using a (DD) model, he compares changes in productivity between towns with high concentrations of Huguenot immigrants and those with fewer or no Huguenot immigrants. To address potential endogeneity concerns, such as the fact that immigrants may choose cities which have growth potential, he uses an IV-approach, using population losses during the Thirty Years' War as an instrument for Huguenot settlement patterns, as this is clearly exogenous, and explains settlement patterns, without correlating with the outcome variable (productivity). His results show that a 1 percentage point increase in the share of Huguenots in 1700 is associated with a 1.5 percentage point increase in productivity in 1802 [31], highlighting how the migration led to positive long term effects of migration. Hornung explains the productivity gains by knowledge transfer between the high-skilled Huguenots and their ability to integrate into local economies. The latter may be explained by the fact that Huguenots were seen as *desirable* immigrants by the Prussian population, as there was religious homogeneity.

Peri (2014) provides a meta study with an overview of the literature on the impact of immigration on native workers' wages and productivity. The emperical work done so far, shows that immigration has a very small effect on the average wages of native workers, with little evidence that it lowers wages for less-educated native workers. Furthermore, the long term effects of especially of high-skilled immigration, are increased innovation and to potential positive productivity effects, through knowledge spillover. He furthermore finds that immigrants are, generally, absorbed into the economy through adjustments by firms and workers, where the firms adopt new technologies and expanding production, while the native workers are channelled into more communication-intensive jobs. Ultimately, the studies seem to suggest that *open and balanced* immigration policies, attracting a balanced mix of education and skills, are likely to have no adverse effects on native wages and *increase innovation and the skill mix, with potentially positive productivity effects* (Peri, 2014).

[31] Hornung interprets this result *as a very conservative estimate because knowledge spillovers to other towns might also have increased productivity there, diminishing differences over time.* (Hornung, 2014)

8 Political Economy

8.1 The Role of Governments

During the course, we have argued that modern governments intervene in the economy in two primary ways: (1) by establishing institutions (e.g. rules or laws) and (2) providing certain goods (often characterized as public goods, i.e. non rival and non-excludable) and services, by taxing the population. In fact, these roles are similar to what Smith saw as the three fundamental functions of the state; (1) administration of certain institutions (such as e.g. justice), (2) providing national defence (internal and external) and (3) the supply of certain public goods (Jody W and Slice, 2007). What both Smiths proposed roles, and the actual roles of modern governments, have in common, is the fact that they provide *something* that — in theory — could be provided by individual contributions, which, however, would result in a socially inefficient level of *supply*. The reason being, broadly speaking, market failures in the form of both a free rider problem (i.e. individuals with avoid paying for public goods, e.g. defence or justice, as they can hardly be excluded) and externalities (i.e. costs or benefits that affect third parties who are not involved in the economic transaction, such as pollution, education or *cheating* on a free market by creating a monopoly). Following Norths notion (1991), we can thus summarize this as by stating that wealth maximising individuals will find maintaining cooperation difficult when the time horizon on a game (e.g. maintaining a free market) is unknown, when information on the other players is lacking, and when there is a large number of players. Hence creating a need for a supervising entity such as a government.

8.2 The Fall of The Minimal State

During the seventeenth and eighteenth century, mercantalist regulations and *strong* governments — in the sense that that they were efficient at collecting tax revenue — emerged, as military expenditure rose. Using Britain as an example, O'Brien (2007) that the efficient collection of taxes and lending, allowed the UK to develop strong external security, thereby supporting trade, investments in institutions, public good provision, and ultimately playing an important role for the Industrial Revolution itself. However, in the nineteenth century mercantilist regulations and policies were abandoned across Europe, as an idea of a Laissez-faire government spread, arguing that the government should function as Night-watchman, simply enacting laws and regulation. The following years, government expenditure as a share of GDP thus fell to around 10 per cent and most public expenditure was consumed by the military, law and order and civil administration, limiting investments in efficiency enhancing institutions. Less than half of the spending was dedicated to public education, state-funded health and assistance to the poor and elderly. The gold standard can be seen as a result of this view on state intervention, as it was thought that the monetary regime ensured that the economy would equilibrate after business cycles, through changes in prices and wages.

The First World War, however, challenged the prevalent view on government intervention. During the war, countries financed war expenditure by expanding the monetary supply through issuing bonds and in some cases printing money, leading to inflation both during, and after the war. According to the rules of the game, all countries should have returned to gold at the original parities, but since full parliamentary democracy was increasingly being introduced, austerity measures to bring inflation down, became increasingly difficult to implement. Some countries devalued, but with the Great Depression in 1929, it became clear that the gold standard — and its supposed equilibrating mechanism — did not work as indented, as countries were reluctant to provide necessary liquidity, prioritizing the protection of their gold reserves instead. Moreover, countries that were previously

been forced to rely on US lending, were now forced to restrict capital mobility, as these loans could not be given anymore, leading to a widespread use of capital and currency controls, that resulted in a decline in world trade in the 1930s. Furthermore, even though nominal interest rates were still low, deflation kept real interest rates high, limiting investments. Finally, some countries increased taxes to equlibrate their balance of payments, leading to pro-cyclical fiscal policy, worsening the crisis. Ultimately, countries who left the gold standard early, e.g. the UK, experienced a *softer landing*, as they were able to use monetary policy to stimulate the economy, whereas countries who stayed on the gold standard longer, e.g. France, did not fully recover until they left gold (Eichengreen et al., 2010) .

8.3 Macroeconomic Spending

During the 1930's macroeconomics was born, as a group economists, lead by Keynes, argued that governments should assist markets by stabilizing demand to smooth out business cycles, and secure high employment. While some may view this as a social-democratic doctrine advocating for big governments, the views rather represent an insight into the fact that markets are not self-equilibrating to a social optimum. Keynes further argued, that this smoothing is effectively done by increase government spending during recessions, as this gives rise to a multiplier effect, as it increased the total income (Y), which in turn increased consumption (C), which in turn increases income, and so forth. While his multiplier effect is still debated, as it, theoretically allows expansionary fiscal policy to have no effect on long term debt, if the tax revenue from increased income compensates for the additional expenses. Empirical work has found multipliers between 0 and 2 (i.e. a 1 pct. increase in government spending leads to 0 - 2 pct. increase in national income), highlighting the uncertainty regarding his multiplier theory. It is, furthermore, worth noting that *fine tuning* demand can — and did — go wrong, as timing is extremely difficult to asses. As data often lacks behind and initial estimates may be uncertain, what initially seems like a recession, and thus demands higher spending, may turn out to be a boom, why a policy intervention based on the first estimate may lead to pro-cyclical fiscal policy, leading to *overheating*, in this context. Today, an important tool for fiscal policy is the use of automatic automatic stabilizers such as progressive taxes that increase, when national income rises, slowing down growth, and decrease, when national income falls, fostering growth. This is, however, only possible in economies where the state has sufficient size, i.e. the tax *reductions* are large enough to have a macroeconomic impact. Furthermore, emperical and theoretical results have shown the importance of consumer and producer expectations in regards to policy interventions, as they — in the most extreme case (Ricardian equivalence) — can make e.g. tax cuts *useless*, as individuals are forward looking and know that the tax cut will be financed by a tax increase in the future, thus having no effect on consumption or investments. As a consequence a lot more attention is now paid to containing the inflationary expectations of macroeconomic policies, as e.g. lower inflation expectations themselves can lead to lower inflation, as trade unions will accept nominal lower wages, as the low inflation leads to higher real wages, thus lowering the cost of labour, thereby decreasing inflation, without the need of actual government interventions. The latter explains why the trust is crucial to central banks, as it allows them to anchor inflation expectations.

While the Phillips Curve (i.e. a negative relationship between unemployment and inflation) was used to guide policy decisions in the 1960's, it *fell apart*, as both unemployment and inflation increased during the 1970's. Attempting to address rising unemployment with fiscal policy in such a situation, turned out to be difficult, as increased government spending would lead to higher inflation and have no effect on employment. The higher inflation would, in turn, challenge current balance and the

exchange rate (if it was pegged to e.g. the German Mark), as imports rose and exports fell. This led to a devaluation pressure, to which e.g. Sweden and the UK in the 1990's responded, by creating their own floating currencies, following Japan and the US. However, as Germany had a tradition of not tolerating high inflation, a number of countries pegged — and stayed pegged — to the German Mark, including France, Italy and Denmark, effectively giving up an independent monetary policy to import low German inflation.

8.4 Why Austerity Measures Might Work

When comparing the duration and severity of the financial crises in 2007/2008 with that of the Great Depression, it seem clear that world economy recovered *relatively* quickly from the recent crisis, with just 3-4 consecutive quarters of negative output growth, compared against up to 12-16 quarters during the Great Depression. It might also, seem plausible, to argue that the expansionary policies during the financial were more effective than the austerity measures during the Depression. This does, however, not imply that austerity measures, at all times, are doomed useless, as the two crises different in their characteristics, namely the public debt was much smaller during the Depression, implying that the government's ability to implement expansionary fiscal policies was less constrained by debt levels. That is, however, not always the case and during crises with high public debt, the effectiveness of fiscal stimulus can be limited.

The theoretic argument for the effectiveness of Austerity measures builds upon two mechanisms. Firstly (1), falling government spending leads to less demand for capital, putting a downward pressure on the interest rate, stimulating investments. Secondly (2) - under the assumptions of rational expectations — a decline in the growth of the government debts will increase household spending, as they realize that that they need to save less to meet future tax claims to service the government debt. Finally, it has since the 1970's been empirically observed, that the *natural rate of unemployment*, i.e. the rate compatible with a with a non-accelerating inflation, has decreased after long deep recessions. Historically, however, the efficiency of fiscal consolidations seems to somewhat ambiguous, as the danish case during 1980-1990 showed no clear results, and the Finnish and Swedish case did succeed, however, with a combination of devaluations, causing a boom and expanding the economy. The lesson from history thus seems to to be that open economies, with substantial export opportunities, can successfully use mild fiscal consolidation to foster growth, if they achieve a competitive advantage through devaluation. For countries with fixed exchange rates (or in a monetary regime) the latter is, however, only possible through so-called internal devaluation, i.e. deflation, which itself is costly and unlikely to happen out of free will.

8.5 Karl Marx's Trap

Continuing on our earlier description of institutions and the fact that they emerge to allow for a beneficial cooperative solution to a complex game, Marx argues that social systems thrive and expand only if they can develop technologies and sustain an increase in material welfare. We may therefore argue that the socialist experiment failed, because it did not deliver technological progress and material improvements on a scale comparable with the market economies of Europe. We may further add, that Marx saw the emergence of capitalism, and the material progress it would create, as an essential prerequisite to communism, why Russia — a relatively backwards agrarian economy — thus by no mean had the necessary preconditions to implement socialism/communism effectively. Generally speaking we may add that the following characteristics of early Soviet planning stand out: (1) an abolition of private ownership of resources and the means of production; (2) very high

investment ratios (i.e. lower consumption); (3) collectivization of agriculture to force structural change; (4) a strong bias towards investment in capital goods industries; (5) and a neglect of consumer goods production. However, even though investments were high, the efficiency was low, as central planning made it difficult to allocate capital to where it was most needed (in a market economy, that would be where the marginal benefit, and hence the price is the highest).

8.6 The Welfare State

While critics may argue that the Welfare State is an egalitarian, socialist system with its primary goal being redistributing wealth between classes, the reality is different. While redistribution does occurs through progressive taxes, welfare States rely more heavily on consumption taxes, which are regressive. Thus, today, welfare states are primarily about smoothing consumption, that is; *transferring* resources over individuals' lifetimes. To see why this is the case, consider a typical household, which starts at the family-formation phase as a net receiver (i.e. it receives more transfers than it pays in taxes), as e.g. the children receiver daycare, go to school or attend university. However, as children leave the educational the household starts becoming a net contributor until the parents enter old age, and become intensive public health consumers as well as beneficiaries of publicly funded pensions. The cycle is illustrated below;

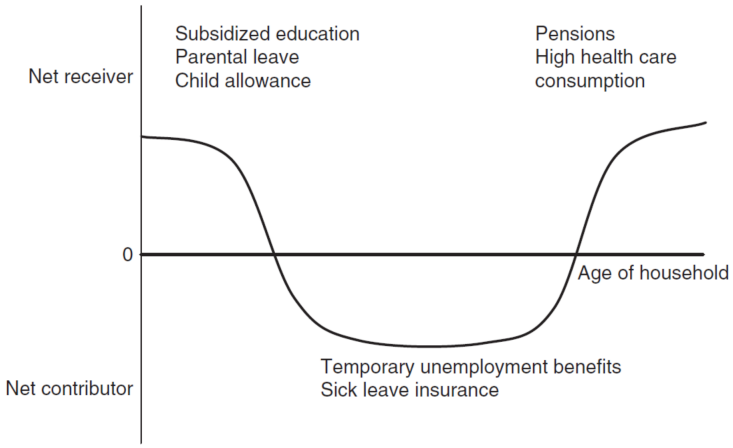


Figure 7: The welfare state as a consumption smoothing mechanism

Note: Illustration taken from Persson et al., 2023

In theory, the above mentioned consumption smoothing could be provided by the market, however, we shall introduce five explanations that challenge this view. Firstly (1), market solutions may lead distributional effects, limit social mobility and violating a common held preference for equal access to some essential services. Secondly (2), externalities associated with government provided services would make a market outcome inefficient, as the marginal social benefit exceeds the private benefit ($MSB > MB$) and private actors act after their own perceived benefit. An example is education, where there are positive spillover effects for society, for which individuals may not account. Thirdly (3), capital market imperfections do not allow for universal access to intertemporal smoothing of income, as e.g. students may not be able to borrow for certain education, which don't have a high enough *expected payoff*. Fourthly (4), time-inconsistent preferences make most people under invest in pension saving, sick leave, unemployment insurance and education. Fifthly (5), health and and unemployment insurance would be associated with adverse selection, as people with high risk would drive the price for insurance up, crowding out the low risk individuals.

Finally, one could argue that high taxes, which are needed to fund a Welfare State, have a negative effect on labour supply, thus harming growth. In his article, Kleven (2014) argues that Scandinavian countries seemed to have *solved* the tradeoff, as they — despite having high tax rates — enjoy significantly higher labour participation rates than in the US, where taxes are lower. Kleven argues that substantial public goods and services that complement labor supply are an essential part of the *solution* (e.g. childcare, education, and elder care) as they effectively subsidize work and reduce the implicit tax burden on individuals. He furthermore adds, that third-party information reporting, minimizing tax evasion, and a broad tax base, allow for efficient tax collecting, mitigating the negative effects on employment. Finally, he adds that cultural factors, such as trust and empathy, might play a role.

Bastiani and Selin (2014) find a similar result, showing that while average hours per employed is about 10 per cent higher in the US compared to Sweden between 2000 and 2015, the effective hours per adult person, defined employment rate multiplied by average hours worked per employee, is higher in Sweden. The Welfare state thus allows more people to participate in labour force, most probably due to public goods and services described above. Barro (2009), however, finds a negative correlation between public spending and growth, although his findings remain a subject of hefty debate.

8.7 Course Literature on Political Economy

Eichengreen and Irwin (2010) show that countries adhering to the gold standard were more likely to resort to protectionist trade policies such as tariffs, import quotas, and exchange controls, after the Great Depression. The reason for this is twofold. Firstly (1), as they were unable to use monetary policy [32] to counter the recession, they sought to protect their balance of payments and domestic economies by restricting trade. Secondly (2), as they aimed to keep their peg to the gold as their economic activity fell, the needed to counteract the falling demand for their currency fell, putting a downward pressure on the currency. As they began to devalue, letting the downward pressure *work*, France and Switzerland, part of the gold bloc, thus were able to reduce their tariffs by between 15-20% and 50%, respectively. In contrast, countries like the UK, which left the gold standard in 1931, saw less dramatic increases (and changes) in tariffs. Based on the former, the paper goes on to argue that the rigid adherence to the gold standard prolonged the global economic crisis by preventing countries from adopting more flexible and effective economic policies. They show that countries remaining on the gold standard experienced deeper and more severe economic backlashes, in part due to their protection. Conversely, countries that devalued their currencies saw a quicker recovery and less need for trade restrictions. The paper thus highlights an essential tradeoff that policymakers faced during the Great Depression: (1) abandoning the gold standard, (2) restricting trade to maintain the fixed exchange rate, (3) devaluing the currency, or (4) creating deflation to revert to the former exchange rate. The fourth option, leading to austerity measures, was generally unacceptable to policymakers, reducing the choices to three, that — on a European scale — all came into play, with varying success, as shown by the former.

Boberg-Fazlić and Sharp (2017) investigate the crowding-out hypothesis, which suggests that public spending reduces charitable donations as individuals perceive taxes as a form of charity. They test the hypothesis by examining the relationship between government spending and charitable activity during the period of the Poor Laws in England, as welfare provision was decided at the parish level, allowing for heterogeneity within a single country, thus eliminating a range of potential endogeneity concerns

[32] They chose fixed exchange rates and free capital flow, effectively giving up independent monetary policy.

(e.g. difference in cultural norms between countries). However, even the parish level spending may not be entirely exogenous, why the employ an IV approach with geographical distance to London being the instrument, as they argue that the cost of migration and the attractiveness of London as a destination influenced the amount of poor relief provided by local parishes (i.e. parishes closer to London had to offer higher relief to retain laborers). As the instrument explains the amount of poor relief, but does not correlate with the outcome variable (charitable spending), it allows to isolate the causal effect even further. Their results show a positive relationship, i.e. areas with more public provision also enjoyed higher levels of charitable income. The study thus finds a complementary rather than *substitutive* (crowding out) relationship between public and private welfare provision, thereby effectively showing emperical evidence against the crowding out hypothesis.

Meissner et al. (2021) explore the relationship between fiscal austerity and the electoral success of the Nazi Party in Germany during the early 1930s. By leverages variations in austerity impacts across different regions, controlling for factors like unemployment and economic output, they find that areas more affected by austerity measures saw higher vote shares for the Nazis. More specifically, they find that spending cuts in health services was particularly influential, in driving voters toward the Nazi Party. This may explained by the fact that reduced public spending increases mortality rates, which may increase desperation and lead to Nazi support. The study thus finds that austerity policies may lead to a strong rejection of established political parties, triggering unintended political consequences with unpredictable, and potentially tragic outcomes, far outweighing the economic benefits, such as a lower natural unemployment rates.

9 Inequality

9.1 Defining Inequality

During the course, we have defined inequality as unequal access to welfare manifested in consumption, health, life expectancy and schooling. While this definition — to some extend — allows us to measure inequality in welfare by looking at income inequality, we need to keep in mind, that utility functions (broadly speaking) differ between individuals and that income shall not be used as the ultimate proxy for welfare. We shall, furthermore, remember, that even though inequality originally emerged during the neolithic revolution, when mankind began cultivating nature, we have since then experienced numerous phases of economic development, each creating winners and losing, and thus affecting the degree and nature of inequality differently. Consider e.g. the eras of globalization. As trade picked up, goods, capital and knowledge (a non rival good), were able to flow more freely, leading to sigma convergence between some countries (those with initially inclusive institutions, as we have argued), as the nature of technology and knowledge, parred with free capital flows, allowed countries to implement best practices across sectors from another. However, we have also been witnessing the so-called great divergence, where the gap between the richest and poorest countries is still widening, even though market integration, on a world scale, is deepening. While we, during the course, have argued that the lack of initially strong institutions seems to be the cause for this, the answer is not satisfactory, as it simple raises a new question; why did these countries have non-inclusive and why didn't they develop?

While the first question has roots to history, beyond Europe, and thus exceeds the scope of this course, the second question falls entirely within our reach. However, the determinants are many, and trying to isolate the primary *driver* would be ambiguous. I will, though, argue that trade itself — while allowing for convergence — may in fact foster divergence between rich and poor countries, thus

not necessarily being Kaldor Hicks efficient in *poor* trading countries (in the long run). I refer to the section about the welfare implications of trade for a thorough discussion of the argument.

9.2 Trends in Long Run Inequality

Generally speaking, income inequality today, measured by the Gini coefficient [33], is lower (0,29 in Sweden, 2000), or approximately equal (0,4 in US) to ancient (0,39, Rome 14 CE) and medieval times (0,45, England and Wales, approx. 1688). That is, even though, income (before tax, ruling out redistrubitonal effects) has increased from being 2x subsistence in Rome to 50x subsistence in Sweden. There is, in other words, no a positive causal relationship between income per head, and inequality. This is surprising, as we to some extend might expect such a relationship. Allow me to explain: in poor economies (such as ancient Rome or medieval England), income can simply not be very unequal, as most people live at or near the subsistence level. But, as average increases above, income variation becomes physically possible, allowing for a higher Gini. This might explain, why the UK has a higher Gini than Rome, as average income was around twice as high, but the poorest people probably lived under similar conditions. It also highlights that we throughout history have made improvements in resource allocation, and despite the possibility for higher income inequality, have not experienced it (on a broad scale).

As an alternative measure to Gini, Milanović, introduced the concept of the so called maximum feasible inequality an economy can experience given a rise in average income, to make comparison between societies at different income levels possible. The idea is to assume that a small elite, receives all income above the subsistence level, survives on the subsistence income. This implies that the higher the average income, the higher the maximum feasible inequality, as measured by the Gini coefficient. By taking the ratio between the actual Gini and the maximum feasible Gini, we can thus asks how close a society is to its theoretical maximum inequality given its per capita income. Using this method we find that Rome was almost as unequal as it could be, with the actual Gini being about 70 pct. ($\frac{0,39}{0,5}$) of the maximum feasible, whereas the US is at about 40 pct. ($\frac{0,4}{0,98}$). Thus, regardless of the method our initial conclusion still holds; inequality in the very long run has fallen. A possible explanation regards the fact that skills, property and capital have — in part due to globalization — been distributed more evenly, giving the broad population (in certain countries) the resources and skills to raise their income above subsistence. This mechanism is, however, a relatively new phenomena, as the ratio between the maximum feasible- and actual gini, did not decline uniformly before the twentieth century, as we e.g. see jumps of about 10 pp. between UK/Wales in 1500's and the Netherlands in the 1600's. We may add that if we were to include the distributional effects of taxes by using after tax income and transfers (which we have not), the fall in inequality witnessed in the twentieth and twenty-first century would be even greater, as the effect on Gini is estimated to be a decline of 0.05–0.10 (e.g. from 0,3 to 0,25/0,2). However, it is far from clear that the egalitarian trend will continue. A number of studies suggest that income inequality in the industrialized world reached a historically low level around 1980 and there is a tendency of inequality to increase again. Finally, we need to acknowledge that data sources for world population income do not exist, making claims about world inequality somewhat ambiguous. Our best *guesses* (including those above) regard large samples of households — not individuals.

[33] Here income inequality is measured as the ratio of the area between the Lorenz curve and the line of perfect equality to the total area under the line of perfect equality (45 degree line). A Gini coefficient of 0 indicates perfect equality, while a coefficient of 1 indicates maximum inequality.

9.3 What Explains Changes in Inequality

During the course, we have established that the main sources of come from are work, property and social transfers, thus changes in income inequality must arise from any of the three determinants. Firstly (1), rapid improvements in technology has increased the demand for highly labour, but since the supply reacts with a lag (since education takes time), wage increases have proportionally favoured high skilled workers. Furthermore, technology has replaced some low skilled workers, potentially leaving them unemployed, if the right institutional framework is not in place. This has lead to a so called race between education and technology, as coined by Goldin and Katz (2008), which has increased income inequality between socially strong and educated individuals, and unskilled labour. Some scholars argue that we should expect the effect to decrease in the future, as mass schooling and education has spread throughout the world, however, quality differences in the latter and the demand being focused on extremely high levels of education, the future in this regard seems hard to predict. Finally, a rise in executive pay, due to, in part, international competition and increasingly complicated compensation packages (consisting of financial assets) further complicate the picture. Secondly (2), while inequality in wealth was rapidly rising after the plague, the political turmoil of the First and Second World Wars, slowed down the development. However, the second era of globalization, and especially the reintegration of financial markets from the 1970s, has allowed capital abundant individuals to again increase their investments and optimize taxes, thus increasing wealth inequality. We might thus be in the middle of a new era, where capital is an increasingly inequality-generating mechanism (see e.g. Atkinson et al., 2011). Finally (3), there seems to be some emperical evidence showing how income inequality is self-reinforcing mechanism, as a negative correlation between inequality and social mobility, is found by e.g. Krueger (2012), implying that the more unequal a society is, the stronger the forces are that maintain that inequality. However, the direction of causality in this relationship remains unclear.

9.4 Gender Inequality

While gender inequality has generally been reduced, as economies have become less reliant on physical strength, it still exists. The raw gender gap thus today varies between 30 to 10 per cent, with a substantial part of it is due to differences in occupations, education and job experience. While some may argue that this is a *free* choice, emperical work has shown that institutional forces play an important role. Thus, economies where women have a high and lasting participation in the labour market, such as the Nordic countries, have lower gender gaps than economies where labour market participation is interrupted by long and frequent spells of childcare. There are, however, also other mechanisms at play, since pay differences (for comparable skills, and experience) are smaller between males at the time they entered the labour market, compared to when they exit. One explanation to this is the fact that females do not obtain the same on-the-job training as males, as their work history was previously less regular (due to them giving birth and leaving the workforce), and that even though this has changed (i.e. women take parental leave, but join the labour market again), it seems that some employees still hold old stereotypes, favoring males (Goldin, 2014). Another explanation given by Goldin concerns the so called *pollution theory*, arguing that even though technological progress limits strength requirements, men have historically lobbied for artificial entry barriers of entry to certain jobs, by highlighting that they are seemingly too hazardous, difficult or stressful for women. Seen through an institutional lens, we might argue that men aim to create rules that preserve their social status and incomes. However, the above does not explain, why studies estimating the pay difference between men and women with identical skills, job history and training, have found a 5 to 10 per cent pay disadvantage of females (a so called residual pay gap). Ultimately, it is still up to the

emerging field of research concerning economics of gender, to uncover further mechanisms that might explain the difference in pay for residual pay gap, as there may be legitimate reasons as to why these differences exist [34]

9.5 Country Inequality

The two eras of globalization both increased and decreased inequality. As trade picked up, goods, capital and knowledge (a non rival good), were able to flow more freely, leading to sigma convergence between some countries (those with initially inclusive institutions, as we have argued), as the nature of technology and knowledge, parred with free capital flows, allowed countries to implement best practices across sectors from another. However, we have also been witnessing the so-called great divergence, where the gap between the richest and poorest countries is still widening, even though market integration, on a world scale, is deepening. While we, during the course, have argued that the lack of initially strong institutions (or perhaps a legacy of non-inclusive institutions), insufficient savings and investment and internal conflict seems to be the cause for this, the answer is not satisfactory, as it simple raises a new question; why did these countries have non-inclusive institutions and why didn't they develop?

While the first question has roots to history, beyond Europe, and thus exceeds the scope of this course, the second question falls entirely within our reach. However, the determinants are many, and trying to isolate the primary *driver* would be ambiguous. I will, though, argue that trade itself — while allowing for convergence — may in fact foster divergence between rich and poor countries, thus not necessarily being Kaldor Hicks efficient in *poor* trading countries (in the long run). I refer to the section about the welfare implications of trade for a thorough discussion of the argument. Besides, one can argue that poverty in itself is a major barrier to the escape from poverty capital is needed to invest in education an infrastructure. However, this is based on an ambiguous assumption of a market failure not allowing these countries to borrow capital on the international market. Other arguments include the fact that poor nations are being, and have been, exploited by foreign investors and by unfavourable trading relations, such as colonizers, who installed dysfunctional institutions and/or foreign investors who have exploited unequal bargaining power. However, this too is ambiguous, as very poor nations don't attract much foreign investment or trade today. This, however, fits into our model described in the section about the welfare effects of trade. As countries, initially specialized in *the wrong* industries, they ended up being dependent on falling demand for a good (unskilled labour), which itself fosters high fertility and thus lowers income per head — a modern Malthusian trap, one might argue. It might also explain why investments are not channelled into these countries, as the industries available for investment cannot afford the high risk premium being demanded (for good reasons) by the lender, as their demand is falling. And because they lack higher education, a complete transformation is both difficult and costly, and would require large interventions from existing, inclusive institutions. Given the current political landscape and the upcoming US election, the latter scenario appears unlikely.

Furthermore, the so-called resource trap [35] might explain why especially certain African countries struggle to escape poverty; their resource endowments simply give rulers weaker incentives to set up inclusive institutions, leading them to instead focus on building extractive industries, which do

[34] This is due to the nature of how residual pay gaps are constructed, i.e., they account for observable factors such as education, experience, occupation, and hours worked, but the remaining unexplained portion of the gap could be attributed to a variety of factors including discrimination, differences in negotiation behavior, or, while unlikely, legitimate reasons such as differences in productivity.

[35] I.e., a phenomenon where countries with abundant natural resources tend to have less economic growth and worse development outcomes due to factors such as corruption, rent-seeking, and weak institutions

not promote sustainable economic growth for the broad population. Moreover, it might lead to the *dutch disease*, where the influx of revenue from natural resources causes the local currency to appreciate, making other export sectors less competitive and leading to a decline in the other sectors (e.g. manufacturing and agricultural). Empirical arguments for the two latter mechanisms include a meta article by Frankel (2012). Finally, Acemoglu et al. 2001 show that colonization and the institutions it created, have had persistent effects, such that countries with extractive institutions tend to have lower levels of economic development, weaker property rights, and less secure political environments compared to those with inclusive institutions. We might, however, add that some, countries who were initially a *little less very poor* such as China, Korea, Brazil, have escaped this trap, while they, however, were not colonized in the same way.

10 Epilogue

10.1 Past, Present & Future

Since the Industrial Revolution (even though we have discussed the validity of the term itself), great improvements in welfare have been made around the world. Firstly (1), productivity increases in agriculture have allowed mankind to eradicate the issue of hunger at world level, making it a question of poor distribution, rather than scarcity as such. Secondly (2), child mortality, on a world level, has decreased from over 20 pct. in 1950, to under 5 pct. in 2021. Thirdly (3), literacy rates (as a proxy for education), on a world level, have increased from around 10 pct. in 1800 to around 80 pct. in 2016, allowing more individuals to potentially escape a life in poverty. Fourthly (4), Malthus' concern about overpopulation and the positive check *kicking in*, seems — for the most part — to be one of the past, as a majority of countries are now facing birth rates under the so-called *replacement rate* (i.e. the rate needed to maintain the same population level), as increased wealth has increased the alternative cost of giving birth and the quantity, quality tradeoff has favored quality. However, as often is the the case in history, development has been unequal, leading to some countries still being trapped in poverty for various reasons. These include (but are not limited to) a persistent favoring of the quantity of children, leaving little room for the improvement of institutions, which thus often remain extractive, rather than inclusive. This, in turn, limits the ability of such countries to experience growth based on globalization and knowledge transfer (which is often regarded as a *free lunch*), limiting sigma convergence on a global scale. Furthermore, one might add that the recent stagnation of globalization and democracy (on a world scale), seem to limit convergence further, as it seems reasonable to argue that political and civil liberties are key to human development. To systematize our concerns regarding the future of global income inequality, we have, during the course, raised the following questions:

1. *Will developing countries experience rapid growth until they finally settle at the lower steady rates of the mature economies? (beta convergence)*
2. *Will all nations eventually enter a phase of modern economic growth?*
3. *Will the new industrializing nations follow the pattern of Europe and North America and become less unequal after the initial phase of increasing domestic inequality?*

If the answer is *yes* to all of the above, we may safely conclude that inequality will decline in the future, however, if the answer is *no* to any one of them, our guess about the becomes ambiguous, at best.

As the answers to questions 2 and 3, clearly dependent on weather or not 1 is possible, our assessment shall start there. Looking at the development of countries within Europe since 1914, we have indeed seen strong growth rates for, at the time, developing countries, allowing e.g. the UK, the Iberian economies, Ireland and pre-socialist economies, to catch up with the leading economies in Europe. During the process, non-socialist economies converged faster, highlighting that the initial institutions matter, however since these economies eventually did converge, increased openness was perhaps a stronger force. However, Russia, Ukraine, Moldova and Belarus highlight significant counterforces, as they still have not reached European levels. Finally, the question is, whether European history can be projected into the future of the world economy. On the one hand, Asia, India and parts of South America serve as arguments for the latter. One may highlight that the gains from modern economic growth are now so apparent, thanks to easily accessible information technologies, that political elites, whatever their political orientation, cannot ignore them. However, Sub Saharan Africa proves to be an important counterexample, with minimal catch up growth, political turmoil and prevailing exclusive institutions. The explanations for this are complicated, and span broadly, including (not limited to), a resource-trap, lasting institutional effects of colonization and divergence through trade, as most African countries had a comparative factor advantage in unskilled labour. Ultimately, the answer, on a world scale, regarding *all* developing countries remains ambiguous, however, it seems reasonable to assume that *some* developing countries will experience some kind of beta convergence. If this is this the case, the answer to question two — for those countries — is *yes*. Finally, as sustained economic growth is most likely associated with a broadening of the skills base, it seems reasonable to assume that the initial phase of increased domestic inequality, in low-income but rapidly growing economies, is offset by a long run decrease. Thus, the answer to question three is a cautious and somewhat uncertain *yes, in the long run*.

However, we need to pause for a moment and ask ourselves, if such a catch up mechanism is even sustainable. In fact, we see that no countries with an HDI of above 0,8 (i.e. very high development) are close to an ecological footprint at or below 1.7 global hectares (gha), which is the level at which resources are replaceable. The countries with the highest HDI an a footprint around 1.7 currently are Tunisia (HDI 0,8). To this we might add that the current development of the world leads to developing countries being affected the most of the consequences, while they are only to a small extend accountable for the damage itself. We might explain this by highlighting how small elites (e.g. rich countries), use their power over international institutions to preserve existing — and create new — *solutions* that favor them the most, relatively speaking [36]. However, climate change is, in the long run, a global phenomena, potentially dangering our species as a whole, if selfish behaviour continues. We might thus be experiencing a new kind of binding landing constraint (the amount of CO_2 our atmosphere can absorb is finite), possibly leading to Malthusian Trap at many times subsistence level. However, concerns about binding land constraints have, historically, been *solved* through technological progress (sometimes in the last minute, following a Boserupian logic, i.e. *necessity is the mother of invention*), thus allowing for a certain degree of optimism. Finally, some scholars (e.g. Daly) argue for *degrowth*, in the sense that we shall accept the *finitness* of nature, and reduce our impact on it, by shrinking our economies, rather than growing them. While this remains a theoretic standpoint as of now, there is no doubt that its policy implications will become more important in the future, if technological progress lacks behind.

[36] This can e.g. be seen by the fact that even though rich countries, for the most part, got rich exploiting natural resources, such as CO_2, they are hesitant to compensate poor countries for the harmful effects of their actions, while simultaneously (for good reasons) wanting to limit the use of those same resources for developing countries. This effectively limits their growth potential.

10.2 Literature on the Past, Present & Future

Bloom et al. (2017) provide emperical evidence showing that the productivity of R&D has been declining sharply across various industries, including semiconductors, agriculture, and pharmaceuticals. They e.g. find that the number of researchers needed to double the density of computer chips every two years (Moore's Law) today is more than 18 times larger than in the early 1970s. However, as the example illustrates, the decreased research productivity is, in the short term, simply being met with an increased amount of researchers, to sustain the same level of innovation add ultimately economic growth. I.e;

$$\text{Economic growth} = \underset{\downarrow \text{ (falling)}}{\text{Research productivity}} \times \underset{\uparrow \text{ (rising)}}{\text{Number of researchers}}$$

However, the paper argues that even though the increased difficulty in generating new ideas can be offset by employing more researches in the short run, the does pose significant challenges for long-term economic growth. This is due to the fact that *ideas are getting harder to find*, why the amount of researches we would need in the future to offset the productivity decline, is simply unsustainable. However, the paper does not account for the emergence of new GPT's, such as — potentially — AI, which could lead to a massive, and lasting, increase in research productivity.

Noy and Zhang (2023) provide experimental evidence showing that gen-AI, specifically ChatGPT, raises average productivity in mid-level professional writing tasks. They find that that ChatGPT *compresses* the productivity distribution by benefiting low-ability workers the most.

11 On Using Historical Data

11.1 Maddison data

As the Maddison dataset is based on the so-called extrapolation method [37], we are indirectly assuming relative prices to maintain constant throughout history. Specifically, the Maddison Dataset uses 1990 as its base year, thus assuming relative prices to be constant at that year. This leads to diminishing accuracy over time, as significant changes in a country's economic structure, such as e.g. those during and after the industrial revolution (where we, however, have challenged the notion), change relative price levels. Bolt and Van Zanten (2020), further find that initial choice of the base year has effects on the results of Maddison Project, as e.g. Chinas's historical income was significantly lower, when using 2011 as a year, compared to 1990. Further, more common, issues when using historical data — that also play a role in the Maddison project — include, data quality challenges, certain assumptions and territorial changes over time [38].

11.2 Bob Allen Wage Data

Data on nominal (day) wages and prices is often easier available (and with less assumptions than e.g. the Maddison data on GDP/capita), allowing for a somewhat simple construction of real wages, by simple dividing the nominal wage level by a CPI (i.e. an index tracking the price changes of

[37] I.e., national accounts growth rates are applied on a benchmark year and then used to estimate e.g. GDP for both subsequent and preceding years

[38] Especially Germany faces this issue, as territorial changes here have occurred often and are complicated. E.g. *from 1950 onwards, the data refer to the 1991 boundaries. From 1918 to 1946, figures refer to Germany within its 1936 boundaries, and from 1850 to 1918 to Germany within its 1913 boundaries.* (Bolt and Van Zanten, 2020)

a representative basket of consumer goods over time), which is constructed using the price data. Formally speaking, we achieve;

$$w_{real} = \frac{w_{nominal}}{CPI} \tag{4}$$

However, calculating the CPI attractively turns out to be extremely challenging, as (1) it has a tendency to overestimate inflation (and hence underestimate real growth), as it is difficult to account for quality improvements, which *nominally* do increase the prices, but *effectively* may reduce the price, since e.g. the efficiency has increased. Secondly (2), new goods are added with a lag, and since such goods often fall in prices after their initial launch, CPI again overestimates inflation, and hence underestimates real growth. Finally, there is the issue of shifting of a substitution effect (if using an index with fixed weights, i.e. a Laspeyres index [39]), where consumer change their preferences, in respond to price changes. As discussed in earlier courses this issue is, however, mitigated by using a so-called Paache index, where the weights (i.e. the quantities) are changed every period;

$$PA_{price} \equiv \frac{\sum_{i=1}^{n} p_{i,t} q_{i,t}}{\sum_{i=1}^{n} p_{i,0} q_{i,t}} \tag{5}$$

This index, on the other hand overestimates the substitution effect, as it assumes that the substitution happens instantaneously. Modern economies (such as the US) therefore often a Fisher index to calculate the CPI, being the geometric mean of the Paache- (changing weights) and Laspeyres index (fixed weights). However, the Paache (and thus Fisher) index require enormous amounts of data, which simply are not possible for historical data. We are thus often stuck with using a simple Laspeyres index, where the substitution effect is not accounted for.

Real wages, further have the disadvantage that they — in contrast to GDP/capita — do not include income from land and capital, which are important income sources if we want to measure the overall income per head. Finally, real wage data is usually only available for day wages, making assumptions about days worked per year necessary, if we want to construct yearly wages. Adding further assumptions, we can even estimate GDP per capita data using real day wages, by;

$$y = \frac{l \cdot w \cdot d}{s} \tag{6}$$

where y is GDP per capita, l is the proportion of workers in the population, w is the daily wage per worker, d is the number of days worked per worker, and s is the share of wages in national income, per worker. However, we now have to make assumptions about the both l, s and d, whereas our initial method only required us to make an assumption about d. Using nominal wage data to asses GDP per head thus come with various sources of errors, as we have to; (1) use a Laspeyres CPI with fixed weight, not accounting for a substitution effect, (2) make assumptions about numbers of days worked, (3) make assumptions about the proportion of workers in the population and (4) make assumptions about the share of wages in national income. Finally, we may add that the data sources themselves propose a certain degree of error, by e.g. not having wage data for all the different occupations during a specific time.

[39] Defined as: $LA_{price} \equiv \frac{\sum_{i=1}^{n} p_{i,t} q_{i,0}}{\sum_{i=1}^{n} p_{i,0} q_{i,0}}$

References

[1] D. Acemoglu. Modeling inefficient institutions. In *Cambridge Collections Online*, chapter 9, pages 21–30. Cambridge University Press, 2006.

[2] D. Acemoglu, S. Johnson, and J. Robinson. The rise of europe: Atlantic trade, institutional change, and economic growth. *The American Economic Review*, 95(3):546–579, 2005.

[3] D. Acemoglu and J. A. Robinson. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business, 2012.

[4] U. Akcigit, J. Grigsby, and T. Nicholas. The rise of american ingenuity: Innovation and inventors of the golden age. Working Paper 23047, National Bureau of Economic Research, 2017.

[5] R. C. Allen. *The British Industrial Revolution in Global Perspective*. Cambridge University Press, 2009.

[6] C. Angelucci, S. Meraglia, and N. Voigtländer. How merchant towns shaped parliaments: From the norman conquest of england to the great reform act. *American Economic Review*, 112(10):3441–3487, 2022.

[7] D. H. Autor. Trade and labor markets: Lessons from china's rise. *IZA World of Labor*, (431), 2018.

[8] R. J. Barro. Government spending is no free lunch. *Wall Street Journal*.

[9] S. Bastiani and H. Selin. Bunching and non-bunching at kink points of the swedish tax schedule. *Journal of Public Economics*, 109:36–49, 2014.

[10] N. Bloom, C. I. Jones, J. Van Reenen, and M. Webb. Are ideas getting harder to find? *CEP Discussion Paper No 1496*, September 2017.

[11] N. Boberg-Fazlić and P. Sharp. Does welfare spending crowd out charitable activity? evidence from historical england under the poor laws. *Economic Journal*, 127(604):50–83, 2017.

[12] J. Bolt and J. L. van Zanden. The maddison project: Maddison style estimates of the evolution of the world economy. a new 2020 update. Maddison-project working paper wp-15, Maddison Project, Oct. 2020.

[13] D. Cantoni and N. Yuchtman. Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, 129(2):823–887, 2014.

[14] G. Clark. *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press, 2008.

[15] J. de Vries and A. van der Woude. *Cambridge University Press*. Cambridge University Press, 1997.

[16] J. Diamond. *Guns, Germs, and Steel: The Fates of Human Societies*. W. W. Norton & Company, New York, 1997.

[17] J. E. Dittmar. Information technology and economic change: The impact of the printing press. *The Quarterly Journal of Economics*, 126(3):1133–1172, 2011.

[18] B. Eichengreen and D. A. Irwin. The slide to protectionism in the great depression: Who succumbed and why? *The Journal of Economic History*, 70(4):871–897, 2010.

[19] J. Frankel. The natural resource curse: A survey of diagnoses and some prescriptions. CID Working Paper 233, Center for International Development at Harvard University, April 2012.

[20] O. Galor and A. Mountford. Trading population for productivity: theory and evidence. *Review of Economic Studies*, 75(4):1143–1179, 2008. NIH Public Access Author Manuscript; available in PMC 2014 July 31.

[21] O. Galor and D. N. Weil. Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4):806–828, 2000.

[22] A. Gerschenkron. Economic backwardness in historical perspective. In *Economic Backwardness in Historical Perspective, a Book of Essays*, pages 6–30. Harvard University Press, Cambridge, MA, 1962.

[23] C. Goldin. *A Pollution Theory of Discrimination: Male and Female Differences in Occupations and Earnings*, pages 313–348. University of Chicago Press, 2014.

[24] E. F. Heckscher and B. Ohlin. *Interregional and International Trade*. Harvard University Press, 1933.

[25] G. M. Hodgson. 1688 and all that: property rights, the glorious revolution and the rise of british capitalism. *Hertfordshire Business School, University of Hertfordshire*, 2015.

[26] E. Hornung. Immigration and the diffusion of technology: The huguenot diaspora in prussia. *American Economic Review*, 104(1):84–122, 2014.

[27] M. Huberman and W. Lewchuk. European economic integration and the labour compact, 1850–1913. *European Review of Economic History*, 7(1):3–41, 2003.

[28] R. Juhász. Temporary protection and technology adoption: Evidence from the napoleonic blockade. *American Economic Review*, 108(11):3339–3376, 2018.

[29] M. Kelly, J. Mokyr, and C. Ó Gráda. Roots of the industrial revolution, 2015.

[30] H. J. Kleven. How can scandinavians tax so much? *Journal of Economic Perspectives*, 28(4):77–98, 2014.

[31] S. Lehmann-Hasemeyer and F. Wahl. The german bank–growth nexus revisited: savings banks and economic growth in prussia. *Economic History Review*, 74(1):204–222, 2021.

[32] N. G. Mankiw. *Macroeconomics*. Worth Publishers, 10th edition, 2019.

[33] C. M. Meissner, G. Galofré-Vilà, M. McKee, and D. Stuckler. Austerity and the rise of the nazi party. *Journal of Economic History*, 81(1):81–113, 2021.

[34] S. Michalopoulos and E. Papaioannou. National institutions and subnational development in africa. *The Quarterly Journal of Economics*, 129(1):151–213, 2014.

[35] D. C. North. Institutions. *Journal of Economic Perspectives*, 5(1):97–112, 1991.

[36] D. C. North. Institutions and the performance of economies over time. In C. Ménard and M. M. Shirley, editors, *Handbook of New Institutional Economics*, pages 21–30. Springer, 2005.

[37] D. C. North and R. P. Thomas. The first economic revolution. *Economic History Review*, 30(2):229–241, 1977.

[38] S. Noy and W. Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Working Paper*, March 2023.

[39] N. Nunn and N. Qian. The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics*, 126(2):593–650, 2011.

[40] M. Obstfeld and A. M. Taylor. The great depression as a watershed: International capital mobility over the long run. *Oxford Review of Economic Policy*, 26(3):285–313, 2010.

[41] O. Olsson. The rise of neolithic agriculture. Working Paper in Economics 57, University of Goteborg, 2001.

[42] L. Pascali. The wind of change: Maritime technology, trade, and economic development. *American Economic Review*, 107(9):2821–2854, 2017.

[43] K. G. Persson, M. Lampe, and P. Sharp. *An Economic History of Europe: Knowledge, Institutions and Growth*. Cambridge University Press, 3rd revised edition, 2023. Unpublished, draft chapters.

[44] J. Riedel. Tariff concessions in the kennedy round and the structure of protection in west germany: An econometric assessment. Kiel Working Paper 41, Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges, 1976.

[45] D. Rodrik. Why do more open economies have bigger governments? *Journal of Political Economy*, 106(5):997–1032, 1998.

[46] P. L. Rousseau. Historical perspectives on financial development and economic growth. *Review*, 85(4):81–106, 2003.

[47] D. Rutherford. Malthus and three approaches to solving the population problem. *Population*, 62(2):213–237, 2007.

[48] S. Sequeira, N. Nunn, and N. Qian. Migrants and the making of america: The short- and long-run effects of immigration during the age of mass migration. *The Review of Economic Studies*, 87(2):650–692, 2019.

[49] I. Sharp, I. Henriksen, and M. Lampe. The strange birth of liberal denmark: Danish trade protection and the growth of the dairy industry since the mid-nineteenth century. *Economic History Review*, 65(2):770–788, 2012.

[50] V. L. Smith. The primitive hunter culture, pleistocene extinction, and the rise of agriculture. *Journal of Political Economy*, 83:727–755, 1975.

[51] P. Temin and H.-J. Voth. Financial repression in a natural experiment: Loan allocation and the change in the usury law in 1714. *The Economic Journal*, 118(533):743–758, 2008.

[52] J. Thirsk, editor. *The Agrarian History of England and Wales*, volume 2. Cambridge University Press, Cambridge, 1988.

[53] N. Voigtländer and H.-J. Voth. Malthusian dynamism and the rise of europe: Make war, not love. *American Economic Review: Papers & Proceedings*, 99(2):248–254, 2009.

[54] M. Waldinger. The economic effects of long-term climate change: Evidence from the little ice age. *Journal of Political Economy*, 130(9):2275–2296, 2022.

[55] J. L. Weisdorf. From foraging to farming: Explaining the neolithic revolution. *Journal of Economic Surveys*, 19(4):561–586, 2005.