How Did Growth Begin?
The Industrial Revolution and its Antecedents

Jón Steinsson*

University of California, Berkeley

March 25, 2024

We live in an era of economic growth. Over our lifetimes, the lifetimes of our parents, and grandparents, output per person in North America and much of Western Europe has grown on average by roughly 2% per year. This steady growth has led to a staggering transformation of material well-being. Ordinary workers in North America and Western Europe earn about 15 times more than they did two hundred years ago.

After several generations of steady progress, it may seem inevitable that economic growth will continue throughout our life-times and the life-times of our children and grandchildren. It is important, however, to realize that the era of rapid economic growth that we live in is a very recent phenomenon. Before 1750, economic growth was less than one tenth as rapid as it is today; and before 1500, economic growth proceeded at a truly glacial pace (as far as we can tell using current historical knowledge).

Our species has dominated the earth for thousands of years. Massive empires have risen and fallen. But over the millenia before 1500, the material well-being of ordinary workers changed very slowly if at all. Then, in a blink of an eye (from a long-term historical perspective), economic growth increased from close to zero to

*I would like to thank Sungmin An, Zhaosheng Li, Sergio Nascimento, Shen Qui, and Daniel Reuter for excellent research assistance. I would like to thank Charles Angelucci, Robert Allen, Gregory Clark, Anton Howes, Ethan Ilzetzki, Reka Juhasz, Alain Naef, Suresh Naidu, Emi Nakamura, Nuno Palma, Gary Richardson, Paul Schmelzing, Hans-Joachim Voth, David Weinstein, and Ludger Woessmann for valuable comments and discussions. I would like to thank Robert Allen, Gregory Clark, Paul Schmelzing, and Ludger Woessmann for generously sharing data and replication programs. I would like to thank Mikel Acha, Joshua Archer, Caroline Dority, Jesse Flowers, Jocob Goldstein, Chazel Hakim, Brian Hill, Henry Isselbacher, Benjamin Nhan, Kevin Wang, and David Zhou for finding errors and typos. First posted in April 2020.

a modern rate of 2% per year. This dramatic change is what we call the *Industrial Revolution*.

The Industrial Revolution first occurred in Britain in the 18th and 19th centuries, but its roots extend back considerably in time and to other parts of the world. In this chapter, I seek to shed light on why the Industrial Revolution occurred, why it occurred in Britain, and why it occurred in the 18th and 19th centuries. Unfortunately, there is no consensus among scholars regarding these important questions. To the contrary, this is a topic of active, lively, and sometimes rather contentious debate.

Given this state of scholarship, I will not try to present a single narrative about the causes of the Industrial Revolution. Instead, I will discuss four major strands of thought: changes in institutions, the Enlightenment, advances in agriculture, and the role of high wages and cheap coal. While the advocates of each of these strands of thought are sometimes quite critical regarding the importance of the others, my view is that all four of these strands (and others) emphasize important developments that likely played an important role in bringing about the Industrial Revolution.

When it comes to big questions such as why the Industrial Revolution occurred, there are proximate causes and deeper causes. Proximate causes are typically easier to identify than deeper causes. In the case of the Industrial Revolution, important innovations such as the spinning jenny, the steam engine, and improvements in agricultural yields are obvious proximate causes. Once the proximate causes have been identified, they typically beg the deeper question of why they occurred: Why did all these innovations occur in Britain at this time? Perhaps they occurred because of favorable institutions in England or because literacy was widespread. This then begs the question of why England had those characteristics. In this way, the quest to understand why the Industrial Revolution occurred is like peeling an onion: there are countless deeper layers to uncover. In this chapter, I discuss both proximate and deeper causes. I therefore try to peel the onion at least a few layers in. Inevitably, the deeper one gets the more speculative the discussion will become.

I will not provide a narrative history of the Industrial Revolution and other historical develpments discussed in this chapter such as the Enlightenment and the political upheaval in 17th century England. Several of the books that I cite in the chapter provide excellent narrative accounts of these events. Let me in particular recommend *The Industrial Revolution* by T.S. Ashton (1948/1997), *The Lever of Riches* by Joel Mokyr (1990), *The English Revolution of 1640* by Christopher Hill (1940), and *The Reformation* by Peter Marshall (2009).

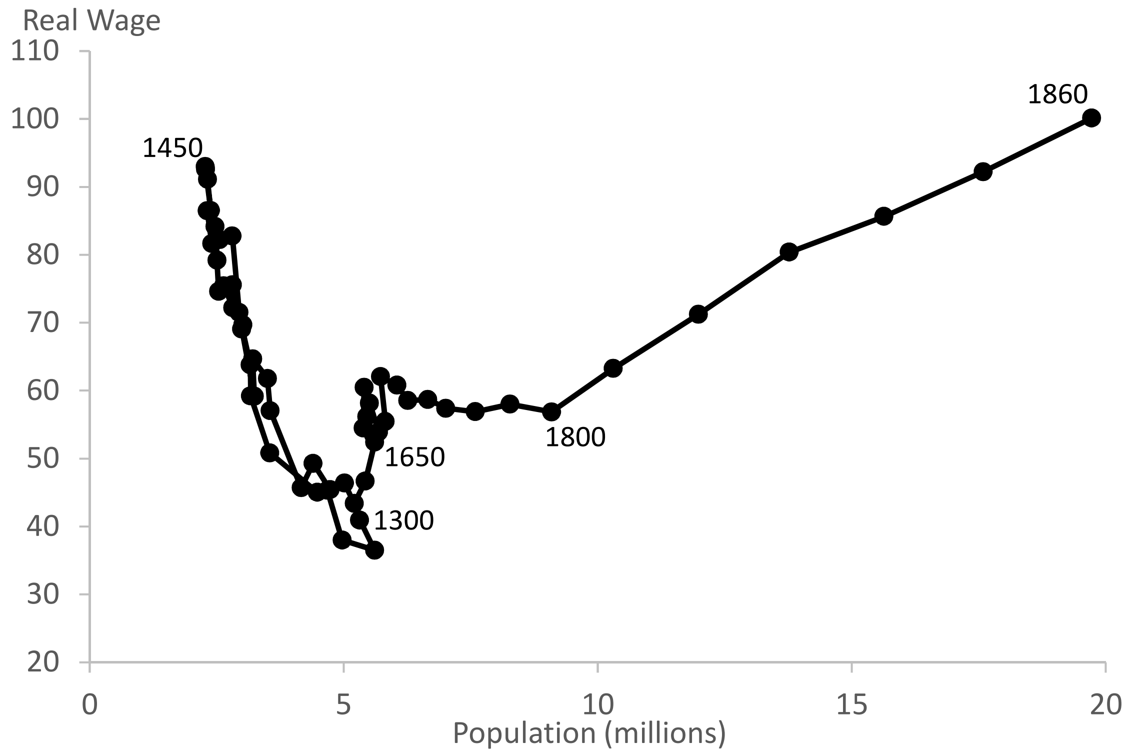

Figure 1: Real Wages and Population in England from 1300 to 1860

Note: This figure replicates Figure 5 in Clark (2005). The real wage and population series are updated series from Clark (2010). The real wage series is an index scaled to be equal to 100 in the 1860s.

1 Institutions

Figure 1 plots the evolution of real wages and the size of the population in England over the period 1300 to 1860. In 1300, the population of England was about five and a half million. Over the next 150 years, the population dropped by more than 50% to about two and a half million as plagues ravaged England. The largest of these plagues was the Black Death of 1348. As the population shrank, real wages more than doubled. Over the subsequent 200 years, the population of England recovered, and, as it did, real wages fell. In 1640, both real wages and the size of the population in England were almost identical to their levels 340 years earlier.

As I explain in greater detail in chapter XX [Malthus chapter], the evolution of real wages and the population in England between 1300 and 1640 strongly suggests that there was no productivity growth in England over this period. The English economy seems to have moved up and down a stable labor demand curve over this period with the movements along the curve being caused first by plagues and then by recovery of the population as the plagues subsided.[1] The Malthus model discussed in chapter XX [Malthus chapter] captures these dynamics well. The easiest way to see why productivity growth was close to zero over this period is that real wages and the population returned to virtually the same point in 1640 as they had been at in 1300. If productivity had increased over the intervening 340 years, the labor demand curve in England would have shifted outward. In this case, the economy would not have been able to return to its earlier state since that state would no longer have been on the labor demand curve (it would have been below and behind the new curve). The fact that the economy returned to the same point in 1640 as it was at in 1300, thus, implies that the labor demand curve did not shift over this period. This, in turn, implies that there was no productivity growth over this 340 year period.

Figure 1 suggests that something important changed in the English economy in or around the 1640s. The point for the 1650s is the first point that is clearly off the labor demand curve that the earlier points seem to be moving along. After 1650, the points keep moving further and further off the earlier curve. This indicates that around 1650 productivity started to rise in England. At first, this growth was slow. But in the late 18th century and especially after 1800, growth seems to accelerate a great deal. Figure 1, thus, suggests that the 150 year period between 1640 and 1800 was a period of transition from stagnation to modern growth in England. By 1800, the Industrial Revolution was in full swing and the growth rate of the economy was staggeringly large relative to anything experienced before.

The timing of the onset of growth suggested by Figure 1 is quite intriguing. The 1640s are not just some random decade in the history of England. This is the decade of the English Civil War between Parliamentarians and Royalists. This armed conflict ended with the overthrow (and execution) of Charles I and the establishment of a Commonwealth in 1649. The monarchy was later restored in 1661, but it was then overthrown again in the "Glorious Revolution" of 1688.

On the one hand, it is perhaps surprising that the onset of growth seems to have occurred at a time of armed conflict. But the English Civil War and Glorious Revolution brought about sweeping institutional change in England that may have played an important role in creating conditions conducive to economic growth. An influential argument along these lines was put forth by Douglas North and Barry Weingast (1989).

1.1 The Sovereign's Commitment Problem

Risk of expropriation has been an important barrier to growth throughout history. Most societies have had weak states and therefore faced a constant risk of invasion. Members of such societies have also faced risk of coercion and expropriation from stronger members within the society. A strong state can reduce the risk of invasion. It can also enforce order within its borders and thereby reduce general lawlessness. In a strong state, however, a different problem arises in the form of predatory behavior by the state itself—usually the sovereign and members of his elite (see Acemoglu and Robinson (2019) for an extensive discussion).

An important downside for a sovereign of acting in a predatory manner is that the risk of predation by the sovereign reduces the incentives of his subjects to make investments that foster economic growth. No one will invest in improving their land, building a waterwheel, or inventing a new technology if they expect that these investments will be expropriated by the sovereign or members of his elite. Even the most self-interested ruler may, therefore, desire to restrain his predatory powers and instead establish secure property rights accompanied by moderate and predictable rates of taxation. This strategy will encourage investment, help foster economic growth and growth in the sovereign's tax base. It may therefore yield higher tax revenue for the sovereign in the long run than unrestrained despotism.

North and Weingast point out a fundamental difficulty faced by a sovereign who would like to establish secure property rights in order to encourage growth: "the sovereign or government must not merely establish the relevant set of rights, but must make a credible commitment to them." In other words, it is not enough to respect property rights in the short run. The sovereign must be able to get his subjects to believe that he is committed to respecting their property rights in the future as well. Why might this be particularly difficult? Suppose the sovereign promises to respect the property rights of his subjects and the subjects believe that the sovereign will honor this promise both in the near term and further in the future. In this case, the subjects will think that the returns that they can earn from making investments are high. They will therefore invest and this will result in increased wealth creation. The problem is that, as the subjects amass wealth, the sovereign's temptation to renege on his earlier promise and confiscate their wealth grows. Eventually this temptation may become irresistible (especially in a time of war when fiscal needs are acute).

This problem is called a *commitment problem*. In a nutshell, it arises from a conflict between what is in the sovereign's interest at one point in time and what is in his interest at a later point in time. For simplicity, let's consider a sovereign that governs for two periods. We will refer to the first period as "ex ante" ("ex ante" is Latin for "before") and the second period a "ex post" ("ex post" is Latin for "after"). We choose these names for the periods in reference to the fact that the sovereign's subjects will make an important choice (invest or not invest) between these two periods. Ex ante, it is in the interest of the sovereign to promise that he will respect property rights for the entire time he is in power. If this promise is believed, it provides his subjects with incentives to invest, increase production, and thereby increase the sovereign's tax base. However, ex post, once the subjects have amassed wealth, it is in the sovereign's interest to renege on his earlier promise and expropriate his subjects' wealth.

The important consequence of this logic is that the subjects should not believe the sovereign's initial promise. If the subjects can understand the incentives of their sovereign, they should be able to anticipate that the sovereign will have an incentive to renege on his earlier promise once they have amassed wealth. This means that they should not believe the initial promise. In this case we say that the sovereign's promise is not *credible*.

Commitment problems are common in a variety of settings. For example, parents face a never ending string of commitment problems. A parent may say to their child: "If you don't finish your dinner, you won't get dessert." But the child may understand that this is not a credible threat. Once dinner is over, the parent's willpower to enforce the punishment as their child screams and cries and others at the table get dessert may be limited. An important macroeconomic example of a commitment problem is the problem of a government trying to regulate its banking system. The government may say to the banks: "If you get into trouble, we won't bail you out." But the banks may understand that this is not a credible threat. If the banks do get into trouble, the prospect of letting them go bust may be too costly for the government to follow through on its earlier threat.

Commitment problems can be solved in two ways. The first is through good *reputation* in cases where interactions are repeated. The sovereign may build a reputation over time for honoring property rights. This may eventually convince his subjects that his promises are not empty. In these cases, the sovereign has an incentive to honor his earlier promises so as to maintain his reputation because having a good reputation benefits him in the future. Conversely, when interactions are repeated, reneging on earlier promises has the cost that it tarnishes the sovereign's reputation and thereby leads to lower investment and tax revenue in the future. These long-run costs may imply that expropriation is not worth it even in cases where the short-run benefits are substantial. (Similarly, parents may endure a painful evening in the hope that this will convince their children to listen to them in the future.) For reputation to be an effective check on the sovereign's temptation to renege, the sovereign must value future tax revenue substantially (i.e., he can't be too impatient and he can't be too worried about being overthrown). These ideas have been formalized using game theory (see, e.g., chapter 2.3 of Gibbons, 1992, on repeated games). North and Weingast argue, however, that sovereigns have in practice seldom been able to make credible commitments to refrain from arbitrary wealth expropriation through reputation alone.

The second way to solve commitment problems is to create *commitment devices*. Odysseus famously tied himself to the mast. Cortez began his campaign to conquer Mexico for Spain in 1519 by ordering his men to burn the ships they arrived on. In both cases, these acts were meant to constrain potentially destructive future behavior. In other words, they were devices meant to commit Odysseus and Cortez' army to act a certain way which they might not see as in their best interest later on. The most common form of a commitment device in modern societies is a contract. Contracts commit the parties involved to perform actions at later dates that they may not see as being in their interest at that point (e.g., to pay for a car that a company has already delivered). The enforcement of private contracts is an important function of modern governments.

Constitutions are perhaps the most important commitment devices of our times. The founders of the United States worried about possible tyranny by a strong executive. To guard against this risk, they wrote the basic structure of the government, including the separation of powers and Congress's control over the power of taxation, into a constitution and made this constitution quite difficult to amend. Calls for even greater constitutional protection of liberties then led to the passing and ratification the Bill of Rights: ten amendments to the U.S. Constitution that guarantee certain basic liberties and place specific limits on government powers. The fact that it is so difficult to change the U.S. Constitution is an important check on U.S. Presidents that have authoritarian tendencies.

1.2 Institutional Change in 17th Century England

North and Weingast argue that the Glorious Revolution of 1688 created an institutional structure in England that credibly constrained the sovereign from arbitrary and confiscatory expropriation of his or her subjects' wealth. They describe several ways in which the English Crown expropriated its subjects in the first few decades of the 17th century. These included "forced loans" (i.e., loans granted under duress) that were not repaid on time or in full, the sale of monopolies in settled industries (with serious adverse consequences for incumbents in the industry), the sale of public offices such as seats in the House of Lords (which reduced the value of existing seats), the use of purveyance power (seizure of goods for "public purposes" without full compensation), the sale of "dispensation" (i.e., threatening subjects with arbitrary enforcement of otherwise lightly enforced regulations unless they paid a bribe), and outright seizure of property. The English Crown was by no means alone in employing these tactics. For example, the Kings of France employed many of the same tactics to raise revenue.

At this time, the House of Commons was mainly made up of the wealthiest segment of the English population. It was exactly these segments of the population that bore the brunt of the Crown's expropriation. Parliament sought to constrain the Crown in various ways, but with little success. Eventually, disputes between Parliament and the Crown lead to the outbreak of civil war in England in the 1640s in which the Parliamentarians eventually prevailed. The economic disputes I have described were one major cause of the English Civil War. But other disputes also played an important role, such as religious disputes between Catholics and Protestants. North and Weingast argue that the Parliamentarians in England would have been unlikely to prevail in the Civil War had the Crown had a standing army like its counterparts on the continent.

The English Civil War lead to dramatic institutional change. Parliament abolished the Star Chamber and passed an act requiring all legal disputes involving property to be tried in common law courts. Before the Civil War, the Star Chamber had various powers including final adjudication of disputes concerning royal prerogative (i.e., royal decrees). The fact that the Crown controlled the Star Chamber meant that Crown's subjects had no independent court to turn to in the event of disputes with the Crown. A second important institutional change was that the Statute of Monopolies—which had been passed by Parliament in 1624—was now enforced. This statute prohibited the creation of monopolies by sale of patents to existing businesses. Third, institutional changes having to do with labor mobility and land tenure overthrew the feudal organization of rural society in England. I will discuss this last set of institutional changes in more detail in section 3.

A major risk in revolutions is that one tyrannical government will simply be supplanted by another. History is replete with examples of this. Arriving at a form of government that is inclusive in terms of political and economic rights is evidently a delicate balance. In the half-century or so before and after the outbreak of the Civil War, England experienced several different forms of government. It was not at all clear during this period that the final outcome of the revolution would be a government that could make a credible commitment to secure property rights for more than a small elite. First, there was the period of autocratic personal rule by Charles I after he dissolved Parliament in 1629. Then Parliament revolted and civil war ensued in the 1640s. Parliament prevailed in this war and set up a Commonwealth which eventually devolved into a military dictatorship of Oliver Cromwell—the Protectorate. During this period, radical elements of the Parliamentary coalition (the Levellers) were violently suppressed. After Cromwell's death in 1658, the Parliamentary coalition sought stability by restoring Charles II to the throne in 1660. Charles II ruled largely with the consent of Parliament. But in the 1680s his successor James II, sought to reestablish the Crown's supremacy in part by issuing new charters for seats in Parliament and thereby packing Parliament with members loyal to him. His first such move was against his traditional opponents, the Whigs. But he then overstepped decisively by moving against his traditional allies, the Tories. This drove the Tories into opposition and lead to generalized revolt of Parliament against the Crown.[2]

In the Glorious Revolution of 1688, Parliament overthrew King James II and replaced him with William of Orange (with the help of an invading Dutch army). More importantly, though, the Revolutionary Settlement and the Declaration of Rights that followed clearly established Parliamentary supremacy in England. The Crown's powers were severely limited after this point and would subsequently further dwindle until the Kings and Queens of England became mere figureheads in the 20th century. After the Glorious Revolution, Parliament had exclusive authority to raise taxes and to approve the government's budget. The Crown no longer had royal prerogative powers, implying that legislative powers were solely in the hands of Parliament. Judges no longer served at the pleasure of the Crown. Rather, they could only be removed by act of Parliament. This meant that the judiciary was independent of the Crown.

The Glorious Revolution (and the institutional changes going back to the Civil War) dramatically increased the political power of Parliament in England. Parliament was dominated by the rich and the nobility. Why did this relatively small elite group not establish its own form of tyranny over the rest of the population as had happened, for example, in Hungary (Fukuyama, 2011, ch. 25)? North and Weingast provide two potential explanations for this. The first is that the institutional structure in England after the Glorious Revolution consisted of a delicate balance between the Crown, Parliament, and the common law courts. Acts of expropriation by either the Crown or Parliament were checked by the other two actors. The second potential explanation discussed by North and Weingast is that the interests represented in Parliament were sufficiently diverse to place severe limits on schemes meant to enrich one group at the expense of the rest of the population. In particular, Whigs represented the newly emergent merchant class as well as the more progressive parts of the gentry. These groups opposed widespread regulation of the economy. On the other hand, Tories represented the more conservative segments of the landowning class, as well as the nobility. Together, the Whigs and the Tories in Parliament, therefore, represented a relatively diverse set of interests, which may have meant that it was difficult for Parliament to create extractive institutions.

1.3 The Financial Revolution

One dramatic development that North and Weingast point to as evidence of a change in the British government's ability to make credible commitments after the Glorious Revolution is the sharp increase in government borrowing and the improvement in the terms at which the government could borrow. These developments are often referred to as a *Financial Revolution* (Dickson, 1967). Figure 2 plots the evolution of the ratio of British government debt to British GDP over the period 1691 to 2015. Before the Glorious Revolution, the English Crown was able to borrow only small amounts (less than 5% of English GDP). By 1700, British government debt had risen to 25% of GDP, mostly due to military spending relating to the Nine Years War with France. After 1700, British government debt continued to rise dramatically, to well over 100% of GDP in the early 19th century as Britain fought several large wars.

Figure 3 plots an interest rate series for government debt in England from 1530 to 1850 recently compiled by Schmelzing (2020). This figure shows a relatively steady fall in the interest rates the English Crown faced throughout the 17th century. In-

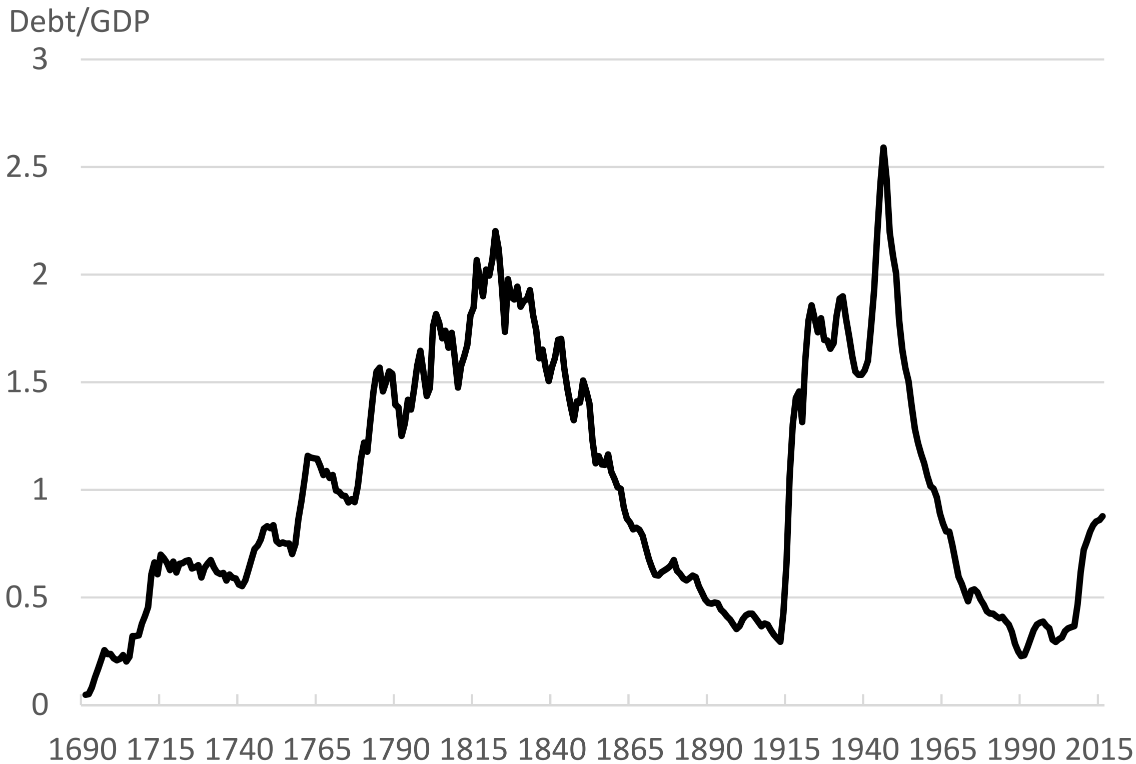

Figure 2: Government Debt to GDP for Britain from 1691 to 2015

Note: This figure plots the nominal value of British government debt divided by nominal GDP. Both series are taken from Thomas and Dimsdale (2017). The nominal GDP figures are based on the work of Broadberry et al. (2015).

terest rates spiked temporarily during the turmoil of the English Civil War and the Nine Years War. But other than that, they are on a steady downward trend. The Glorious Revolution is not an obvious break point. Yet, it is impressive that interest rates continued to fall even as the quantity of British government debt outstanding expanded massively in the 18th century. The interest rates the British government paid on this massive debt in the middle part of the 18th century were remarkably low by historical standard at around 3%. Even when the British government became extremely indebted in the late 18th and early 19th centuries, the interest rates it faced only rose to about 5%.

A key reason why the British government was able to raise so much debt at such favorable rates was its ability to collect a steady and increasing stream of taxes (O'Brien, 1988). The ability of the British government to finance large wars with debt in the 18th and 19th centuries was a major advantage vis-a-vis its main adversaries, notably France. This meant that it could avoid imposing ruinous taxes during war time and instead pay for wars over a longer period of time. Arguably, the ability

Figure 3: Interest Rates on Government Debt in England

Note: This figure plots decadal averages of interest rates on long-term government debt in England. This series is constructed based on annual data from Schmelzing (2020). The interest rates are for voluntary loans to the government. Prior to the Glorious Revolution, the crown would at times coerce its subjects to lend it money. The terms on such forced loans are excluded.

to finance wars by issuing large amounts of debt was a key contributor to Britain's success in the wars of the 18th and 19th centuries, and, therefore, its emergence as a dominant naval power and major empire.

Developments in private capital markets in the decades after the Glorious Revolution were also dramatic. The Bank of England was founded in 1694 as a profit-oriented private bank. Early on, its main activities were to assist the government in finding buyers for its massively increasing debt and to issue bank notes that were used as a means of payment for large transactions in London. As the 18th century wore on, the Bank took on more and more roles of a modern central banks, such as that of a lender of last resort during panics (see Chapter XX [Money and Banking]).

More generally, the 18th century saw an explosion of banking in England, initially in London, but later on throughout the country. These banks greatly improved the English payment system. They played a fundamental role in financing enterprise. In particular, they channeled funds from agrarian area in the South that had surplus funds to industrializing areas in the North that had profitable investment opportunities. As Bray Hammond memorably put it "banks provide credit, and credit is indispensable to enterprise" (Hammond, 1957, p. viii).

The 18th century saw in England the development of a banking system much more advanced than the world had ever seen. This may have been an important contributor to the emergence of the Industrial Revolution in England. (But it is hard to rule out reverse causation.) Kindleberger (1993) argues that the political revolutions in England in the 17th century contributed importantly to these financial developments by reducing the risk of arbitrary seizure by the state of concentrated assets such as occurred in 1640 and 1672.

1.4 Critiques of North and Weingast

North and Weingast's argument is controversial. Clark (1996) argues that if the Glorious Revolution substantially increased the security of property rights in England, rates of return on capital should have fallen substantially after the Revolution and asset prices should have increased. He presents data on rates of return on farmland and rental charges, and the real price of land from the records of English charities. Clark points out that nothing much happened to these series around the time of the Glorious Revolution. Figure 4 presents data compiled by Clark for the rate of return on farmland in England. Clearly, there is no dramatic change in the return on farmland around the Glorious Revolution. Clark, furthermore, points out that periods of political instability during the 16th and 17th centuries (such as the Civil War) did not lead to much variation in these series. Clark contrasts this with data from the town of Zele in Flanders over a similar time period. Both rates of return on land and real land prices fluctuate a great deal in Zele in response to wars affecting that area. Clark concludes that property rights in England were relatively secure as early as 1600 and did not become substantially more secure as a result of the Glorious Revolution.

Clark's critique applies to the security of the property rights of landowners. It may, however, have been that other types of property rights, or the rights of other groups of people, were less secure prior to the 17th century revolutions in England. It may, for example, have been that merchants faced less secure property rights than landowners, and that individuals or groups that the Crown disapproved of were more likely to be expropriated. Jha (2015) provides interesting evidence in support of this thesis. He shows that among members of Parliament, merchants and investors in overseas joint stock companies were more likely to rebel, while proxies for domestic wealth and ownership in domestic joint stock companies do not pre-

Figure 4: Rate of Return on Farmland in England

Note: This figure plots an estimate of the average rate of return on farmland in England from Clark (2002, 2010). It is calculated as the ratio of rental payments on land to the price of the land.

dict support for the rebellion.[3] The Crown had much more scope for expropriation of overseas activities than domestic activities because foreign trade was governed by civil law courts, which were administered by the Crown itself, while domestic property rights were governed by common law. Prior to the Civil War, the Crown repeatedly engaged in expropriation of overseas companies by raising customs duties on imports and revoking charters of overseas companies.

A related but quite different line of argument is that perhaps the Glorious Revolution and the shift of power to Parliament ignited economic growth by allowing property rights to be more easily reorganized and rationalized. Property rights in medieval England (and elsewhere) involved a complex maze of interwoven rights which complicated their reorganization in response to changing circumstances. I discuss the nature and evolution of these rights prior to the Glorious Revolution in greater detail in section 3. Bogart and Richardson (2011) discuss how in the 18th century "Parliament embraced novel ideas concerning property and established procedures for processing petitions from groups hoping to reorganize rights to land and resources." Most prominently, Parliament passed several thousand enclosure acts. But it also passed a great number of statutory authority acts that allowed for the construction of infrastructure such as turnpikes and canals. Britain experienced a transport revolution during the 18th and 19th centuries which dramatically lowered transaportation costs and therefore allowed for expanded trade (Bogart, 2014). This transport revolution may have in part resulted from Parliament's willingness to reorganize property rights and use its power of eminent domain. Rosenthal (1990) provides an interesting counterpoint. He argues that highly profitable irrigation projects in the French region Provence were not undertaken before the French Revolution due to fragmented political authority over the right of eminent domain. Interestingly, this line of argument, if correct, implies that the Glorious Revolution (and the French Revolution) ignited growth by making ancient (inefficiently organized) property rights less secure, or at least more easily renegotiable.

The Glorious Revolution and its antecedents may have been important for other reasons as well. "Rent seeking" has arguably been a serious impediment to growth and prosperity in most societies. Rent seeking refers to the pursuit of various forms of privilege by favored and powerful individuals in society—such as monopoly rights, guild protection, or other kinds of protection from competition, as well as exception from taxation. Parliament struck down many monopolies and other types of economic privileges in Britain during the 17th and 18th centuries. More generally, a gradual shift in policy occurred towards free-market policies, where, as economic historian Joel Mokyr has put it, "freedom" came to mean "the freedom to *enter* a branch of economic activity rather than the freedom to exclude others" from that activity (Mokyr, 2009, p. 27). These shifts likely created conditions that were more favorable to innovation and entry by startups and, therefore, may have been an important factor in igniting economic growth.

1.5 The Role of Atlantic Trade

Was there something special about England in the 17th century that led—through a series of revolutions and counter-revolutions—to the establishment of secure property rights and relatively inclusive economic and political institutions? Acemoglu, Johnson, and Robinson (2005) argue that the rise of Atlantic trade following the Great Discoveries of Christopher Columbus and Vasco de Gama played an important role in sparking these institutional changes. In this view, the Great Discoveries of the 1490s were an important "root" cause of the Industrial Revolution. (The Great Discoveries were, of course, themselves made possible by innovations in shipbuilding prior to 1500.)

The starting point of Acemoglu, Johnson, and Robinson's analysis is the notion that "between 1500 and 1800, Western Europe experienced a historically unprecedented period of sustained growth, perhaps the 'First Great Divergence"'. They then argue that this First Great Divergence was concentrated among a small group of countries involved in Atlantic trade. Since high quality measures of output per person are not available for most of the countries they seek to analyze, Acemoglu, Johnson, and Robinson use the urbanization rate as a proxy measure of output per person. More specifically, the measure they use is the fraction of the population that lived in cities and towns that at some point between 800 and 1800 had at least 5000 inhabitants. The logic for why the urbanization rate might be a reasonably good proxy for output per person is that only areas with high agricultural productivity—i.e., high agricultural output per person—could support large urban populations.

Figure 5 plots the urbanization rate for three groups of European countries from 1300 to 1800. The first group is the countries Acemoglu, Johnson, and Robinson refer to as *Atlantic traders*. These are modern day Britain, France, the Netherlands, Portugal, and Spain. The second group is other countries in Western Europe. Acemoglu, Johnson, and Robinson define Western Europe as the area west of the Elbe river. The second group of countries thus includes modern day Austria, Belgium, Denmark, Finland, Germany, Ireland, Italy, Norway, Sweden, and Switzerland. The third group is Eastern Europe, which Acemoglu, Johnson, and Robinson define as all countries east of the Elbe including Russia, but excluding modern day Turkey. Figure 5 shows clearly that urbanization rates rose much more in the countries engaged in Atlantic trade than in the rest of Europe. The urbanization rate of the Atlantic traders rose by 150% from 8% to 20% between 1300 and 1800, while the urbanization rate in the rest of Western Europe and in Eastern Europe only rose by 70% over this same period.

Atlantic trade—i.e., trade with the New World, the West Coast of Africa, and Asia—yielded large profits for those skillful and lucky enough to be successful. Table 1 summarizes Acemoglu, Johnson, and Robinson's estimates of annual profits from Atlantic trade in England from 1500 to 1800. The profits of £200,000 in the first half of the 17th century were enough to make quite a few merchants and privateers very wealthy (GDP per person at the time was roughly £7). However, these profits were not large enough to directly explain a sizable portion of the growth in English output in the 17th and 18th centuries. In the first half of the 17th century, these profits amounted to only 0.7% of GDP. In the first half of the 18th century, they amounted to 4% of GDP. In the second half of the 18th century, they amounted to

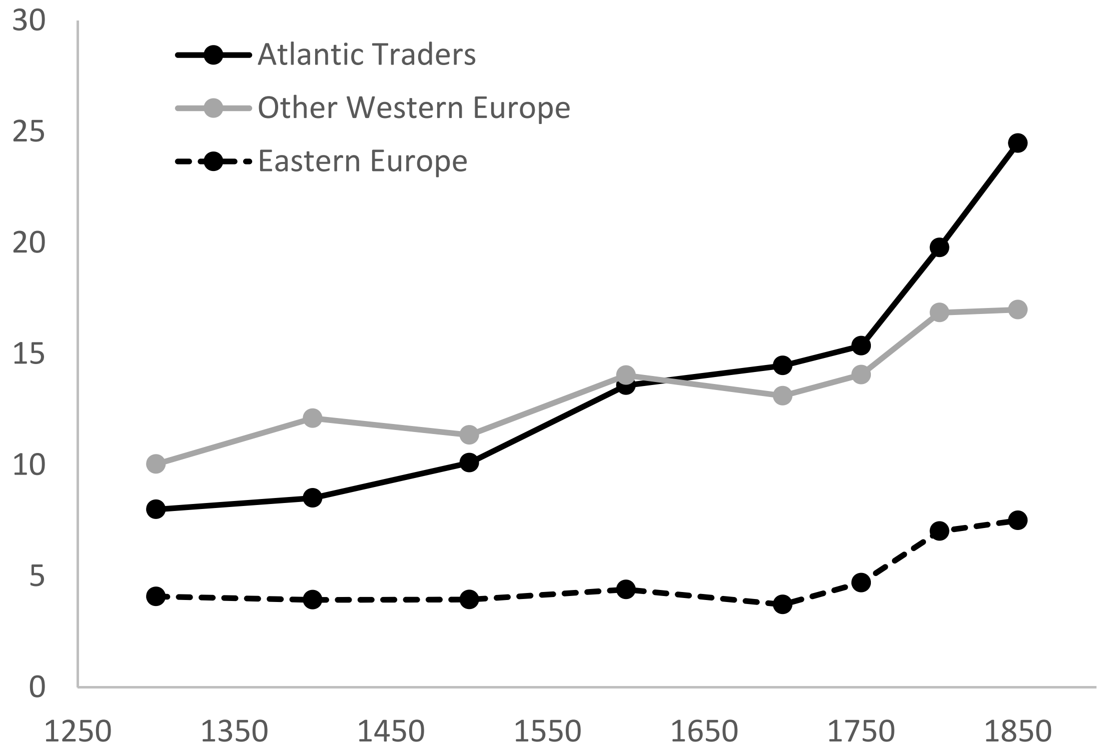

Figure 5: Urbanization Rate in European Regions

Note: These data series were constructed by Acemoglu, Johnson, and Robinson (2005) from data on urban populations compiled by Bairoch, Batou, and Chèvre (1988) and population estimates from McEvedy and Jones (1978).

7.5% of GDP. Over this 200 year period, GDP per person in Britain grew by roughly 75%.[4]

While the direct effect of Atlantic trade on growth was modest, Acemoglu, Johnson, and Robinson argue that Atlantic trade had a substantial indirect effect on growth by helping create conditions conducive to fundamental institutional change. For such institutional change to occur, however, a second condition was also important: Atlantic trade only resulted in fundamental institutional change in countries that had relatively non-absolutist institutions to begin with. Acemoglu, Johnson, and Robinson's argument goes as follows: Atlantic trade created large profits. In countries with absolutist initial institutions (e.g., Spain and Portugal), the sovereign and his elite were able to dominate the trade and exclude others from it. However, in countries with less absolutist initial institutions (e.g., Britain and the Netherlands), Atlantic trade greatly expanded and enriched commercial interests outside the royal circle. These groups eventually became powerful enough to demand institutional change that constrained the power of the sovereign and protected property

Table 1: Estimates of Annual Profits from Atlantic Trade in England

Before 1575	Negligible
1576-1600	40,000
1601-1650	200,000
1651-1675	500,000
1676-1700	900,000
1701-1750	1,700,000
1751-1800	5,000,000

Note: Estimates reported in Acemoglu, Johnson, and Robinson (2005). All figures are in pounds sterling adjusted to 1600 prices using the index of craftsmen's wages from Phelps Brown and Hopkins (1955).

rights. These institutional changes were, in turn, crucial preconditions for economic growth.

Acemoglu, Johnson, and Robinson's theory helps us make sense of the different growth rates experienced by different parts of Europe. The first group of countries is Britain and the Netherlands. These countries had access to the Atlantic and had relatively non-absolutist initial institutions. They experienced liberal revolutions in the 16th and 17th centuries, which established more inclusive economic and political institutions than other countries had (e.g., secure property rights for a larger portion of the population and relatively free entry into profitable economic activities). These countries grew, by far, the most between 1500 and 1800. Figure 6 illustrates this by comparing the evolution of the urbanization rate in England and the Netherlands with that of Spain, Portugal, and France.

The second group is Spain and Portugal. These countries had access to the Atlantic and, in fact, were first-movers when it comes to Atlantic trade and colonization. However, they had highly absolutist monarchies at the time. This implied that the sovereign and his elite were able to dominate the Atlantic trade. As a result, these countries did not experience liberal revolutions until several hundred years later (in the 20th century) and did not experience the high growth rates of the first group. France's experience is intermediate between the first and the second groups. The French monarchy was sufficiently absolutist to largely dominate the Atlantic trade early on and prevent a liberal revolution from occurring until the late 18th century. Growth in France took off later than in Britain and the Netherlands, but earlier than in Spain and Portugal.

The third group is the city states of Northern Italy and the Baltic Sea. These

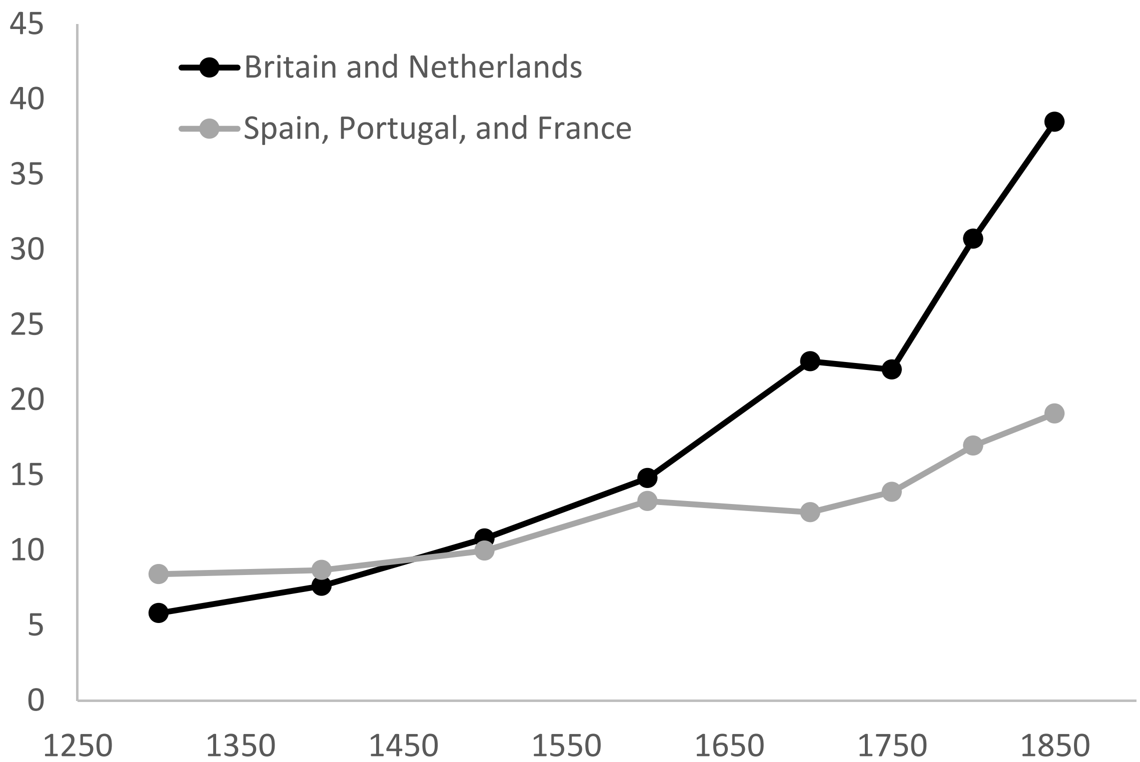

Figure 6: Urbanization Rates Among Atlantic Traders

Note: These data series were constructed by Acemoglu, Johnson, and Robinson (2005) from data on urban populations compiled by Bairoch, Batou, and Chèvre (1988) and population estimates from McEvedy and Jones (1978).

states had relatively non-absolutist initial institutions. However, they lacked access to the Atlantic and therefore did not benefit from the Atlantic trade. They saw their relative wealth and stature decline markedly after 1500. Finally, most other parts of Europe both lacked access to the Atlantic and had highly absolutist initial institutions. They grew very slowly between 1500 and 1800.

One thing that the ideas of both North and Weingast and Acemoglu, Johnson, and Robinson highlight is the notion that socially beneficial institutional change may not occur because it is resisted by groups that stand to lose privileges and power. In principle, the aggregate gains resulting from the institutional changes should make it possible for an agreement to be reached that compensates those groups that lose from the change—this is what the Coase Theorem suggests. In practice, however, such schemes appear difficult to negotiate (sometimes because the privileges and power of the elite are not considered just by the rest of the population). For this reason, the process of institutional change often involves substantial amounts of conflict. And in many cases, beneficial institutional change does not occur or occurs only after long delay. Resistance to change by special interest groups with power and privilege is surely one of the most important barriers to growth both today and throughout history.

2 Enlightenment

The Industrial Revolution involved a surge in the rate of technical innovation. The most famous innovations occurred in textiles (the spinning jenny, water frame, mule, etc.) and steam power (the Newcomen engine, the separate condenser, the rotary engine, the high pressure engine, etc.). But technical innovation was much more widespread with important innovations appearing in many industries. In the words of economic historian T.S. Ashton, "a wave of gadgets swept over England" (Ashton, 1948/1997, p. 48). The roughly simultaneous quickening of technological innovation across many parts of the economy suggests that the Industrial Revolution was not due to one or two lucky inventions in the 18th century. It suggests, instead, that something changed that improved the general conditions for technological innovation.

Joel Mokyr and Deidre McCloskey have argued that the crucial change was one of ideas and culture associated with the Enlightenment (Mokyr, 2009; McCloskey, 2006, 2010, 2016). According to this view, a culture of progress and growth emerged in Northwestern Europe in the 18th century based on the idea that mankind can improve its condition by scientific inquiry and rational thought. A central element of the Enlightenment was an attack on religious orthodoxy and the notion that God's word was the final answer to all questions (but not religion per se). Instead of religious dogma, leaders of the Enlightenment advocated empiricism and reason. In the words of Francis Bacon, "Man, being the servant and interpreter of Nature, can do and understand so much and so much only as he has observed in fact or in thought of the course of nature" (Bacon, 1620/1999). Immanuel Kant famously declared "Sapere Aude! [e. Dare to Know] Have courage to use your own reason!" (Kant, 1784/1986). These notions may seem commonplace today. They were, however, radical in a society dominated in intellectual matters by religion. The Catholic Church did not like people questioning assumptions and fiercely resisted the heretical writings of Enlightenment thinkers.

The idea that the Enlightenment was a crucial precursor to the Industrial Revolution because it created conditions favorable to technological innovation begs the question: Why did the Enlightenment occur? One possible answer is that perhaps an unusual cluster of brilliant and freethinking people happened to live in Northwestern Europe in the 17th and 18th centuries. Perhaps it was simply a stroke of luck that the likes of Francis Bacon, Isaac Newton, John Locke, Thomas Hobbes, David Hume, Adam Smith, René Descartes, Voltaire, Jean-Jacques Rousseau, Antoine Lavoisier, Gottfried Wilhelm Leibniz, Leonhard Euler, Immanuel Kant, and Christiaan Huygens, just to name a few important thinkers, lived and worked in Northwestern Europe in the 17th and 18th centuries.

This theory seems unlikely to be true. It suggests that essentially no person of equal brilliance and freethinking ability to the people listed above lived on the entire continent of Europe for hundreds and hundreds of years during the "Dark Ages" (say 500 to 1200 CE), while over a dozen such people lived in a small part of the continent over the short span of the Enlightenment. If extreme talent is random, such a huge disparity in the frequency of extreme talent would, statistically speaking, be astronomically unlikely to occur given the large size of the population of Europe.

But if many brilliant and freethinking individuals have lived at all times and in all places, why have so few important ideas survived from other times and places? There are likely several reasons for this. First, the technology available to maintain and transmit knowledge throughout much of history was primitive and expensive. Paper was invented in China about 2000 years ago but took over 1000 years to reach Europe. Before paper, documents were carved in stone or written on animal skin at great expense. Most things were simply never written down. The primary method for knowledge transmission was the spoken word. "The Memory of individuals and of communities carried knowledge through time and space," as Boorstin (1985, p. 480) put it. This meant that knowledge spread much more slowly than it can in modern times and was more vulnerable to being lost forever. Mokyr (1990, ch. 9) discusses important technologies that the Chinese mastered but subsequently lost altogether, such as technologies for time measurement and for building oceangoing ships. Clark (2007, p. 142-144) discusses evidence for technological regress among Aboriginals in Australia and especially Tasmania as well as among Inuits in the Arctic. When first encountered by Europeans, these groups had a level of technical knowledge much more primitive than their own ancestors are known to have had.

Many of the important inventions of the Industrial Revolution were quite simple and did not rely on any advanced scientific knowledge. They could therefore have been invented by nearly anyone. In textiles, the spinning jenny and the flying shuttle, for example, are simple inventions that yielded dramatic increases in productivity. Millions of people spun and wove fabric for millenia before the Industrial Revolution. How could it be that none of them thought of these simple improvements? It seems likely that some of them actually did, but that knowledge of the improvements did not manage to spread.

Another important reason why so little knowledge survives from most times and places is that huge amounts of written material have undoubtedly been lost to the destruction of war and fire. A famous case is the Great Library of Alexandria, one of the largest libraries of the ancient world. Although sources differ, many say it was seriously damaged due to specific episodes of war and fire. Some say the armies of Julius Caesar accidentally burned down part of the library in 48 BCE. Others say that the library was badly damaged when Emperor Aurelian suppressed a revolt by Queen Zenobia in the third century CE.

But perhaps the most important reason why so little knowledge survives from most times and places is that powerful actors in society actively seek to suppress ideas and destroy knowledge that harms their interests. The Inquisition of the Catholic Church is a famous example. Another prominent example is the censorship and ideological repression of the Soviet Union and its satellite states during the 20th century. But censorship, the burning of books, and the imprisonment and murder of reformers and intellectuals has been endemic throughout history.

Whether the ideas of brilliant and freethinking individuals survive or not depends on the relative strength of the forces and technologies that destroy knowledge and suppress new ideas and the forces and technologies that preserve knowledge and spread new ideas. The fact that the state of human knowledge was relatively stagnant throughout most of history suggests that the forces and technologies that sought to suppress new ideas were powerful enough to completely counterbalance the forces and technologies that sought to spread new ideas.

2.1 A Watershed Moment: The Invention of the Printing Press

Then something changed. But what? I contend that a watershed moment was the invention of the movable-type *printing press* by Johannes Gutenberg around 1450.[5] The idea is that the printing press massively altered the balance of power between those seeking to suppress knowledge and those seeking to spread it: after its invention, the forces seeking to create and spread knowledge decisively had the upper hand and knowledge began to expand at a rapid rate. If this hypothesis is correct, the invention of the printing press is probably the most important invention of all time.

The hypothesis that the invention of the movable-type printing press was the watershed moment that "changed everything" faces at least two important challenges. First, movable-type printing had been invented in China long before Gutenberg—at least as early as the 11th century (Boorstin, 1985, ch. 62). Why didn't movable-type printing revolutionize knowledge accumulation in China? An important difference between the Chinese and European experience with movable-type printing is the fact that the Chinese language has no alphabet. Rather, it has more than thirty thousand characters. To print books using movable type, a Chinese printer needed type pieces for each of several thousand characters. To assemble a page, the printer needed to retrieve each character from this vast collection and then put them back in the right place after use. This was a laborious process that meant that movable-type printing was not nearly as useful an invention in China as in Europe.

Second, the Industrial Revolution occurred a full 300 years after Gutenberg's invention. If movable type was the crucial element, why did the Industrial Revolution not occur earlier? A plausible explanation for this is that European society needed to undergo a colossal transformation before conditions were ripe for the Industrial Revolution. The forces that opposed new ideas were very strong in Europe around 1450. It took quite some time to weaken these forces. The invention of movable-type printing arguably set in motion a sequence of transformational events in Europe: the Reformation in the 16th and 17th centuries, then the Enlightenment and Scientific Revolution in the 17th and 18th centuries. The explosive economic growth of the Industrial Revolution was a culmination of these changes.

Consider first the *Reformation*. Martin Luther is said to have nailed his Ninety-Five Theses to the door of the All Saints' Church in Wittenberg on October 31, 1517. This event is widely regarded as the starting point of the Reformation. In truth, Luther's revolution started off as an obscure academic debate mostly about indulgences—the sale by the Catholic Church of certificates remitting some of the punishment a person was allegedly due in purgatory because of his or her sins on Earth. It, however, escalated, and in 1520 Luther was excommunicated by Pope Leo X. Luther reacted by burning the papal bull of excommunication in Wittenberg, refusing to recant his errors at the imperial diet at Worms—famously proclaiming "here I stand, I can do no other"—translating the New Testament into powerful and idiomatic German, and publishing a series of pamphlets calling for major ecclesiastic reform. These events made Luther a celebrity and national hero in Germany. Marshall (2009) provides an excellent account of these events and the Reformation more generally, which my account draws on heavily.

There had been attempts at radical ecclesiastic reform before Luther. Famous examples include the movements of John Wycliffe in England and Jan Hus in Bohemia in the 14th and early 15th centuries. In these prior cases, the Catholic Church was able to beat the movements back and maintain control of the dogma. Jan Hus was burned at the stake in 1415. In the case of Luther's movement, however, the Catholic Church totally lost control. As Marshall (2009) points out, unlike the writings of Wycliffe and Hus, Luther's writings were printed. Luther's pamphlets were, in fact, enormously popular, helping to spread his ideas wider and faster than anything the Catholic Church had ever had to deal with before. Other ecclesiastic reform leaders of the 14th and 15th centuries were arguably deeper thinkers on doctrinal matters than Luther—including Calvin, Zwingli, and Melanchthon. Luther, however, made extraordinarily effective use of the printed pamphlet as a means of spreading his ideas, writing copious numbers of pamphlets in a language that appealed to a broad audience.

The Reformation loosened the grip of religious authorities on intellectual life in Europe, which helped lay the groundwork for the Enlightenment and Scientific Revolution (Enlightenment for short). Censorship was much less severe in some Protestant countries such as England and the Netherlands than in Catholic countries such as France in the 17th and 18th centuries. The existence of these Protestant pockets of relative liberty was arguably important for the spread of the Enlightenment. Some of the most prominent Enlightenment thinkers were French. But in many cases they published their work in the Netherlands, England, and other Protestant regions due to censorship in France. Voltaire is a good example. Early in his life, French authorities twice sent him to the Bastille for his writings. He fled to England and the Netherlands and published much of his most controversial early work in those countries (much of it was banned in France). Later in life, Voltaire settled close to the Swiss border in Southern France and used Swiss printers to publish much of his late controversial work—such as Candide.

Another channel through which the Reformation likely contributed to the emergence of the Enlightenment was by encouraging literacy. Luther advocated universal schooling to enable Protestants to read the Bible on their own (Becker and Woessmann, 2009). Of course, once people could read the Bible, they could also read other things. Historical data on literacy are, unfortunately, rather limited. However, several researchers have gathered data on the fraction of people able to sign their names on various formal records such as court depositions. The idea is that being able to

Figure 7: Literacy Rate of Husbandmen in London and Middlesex, 1560-1740

Note: Results from a regression described in notes to the text based on data from Cressy (1980) and Houston (1982). Cressy reports the fraction of witnesses who sign ecclesiastical records, while Houston reports the fraction of witnesses who sign court depositions.

sign one's name is a useful proxy for basic literacy. Figure 7 presents estimates of the evolution of the literacy rate of husbandmen—farmers of relatively modest means—in London and Middlesex from 1560-1740 based on data of this sort from Cressy (1980) and Houston (1982).[6] According to these estimates, literacy skyrocketed in the 200 year period leading up to the Industrial Revolution. In 1560, only 27% of husbandmen in London and Middlesex could sign their names. By 1740, this fraction had risen to 84%.

More fragmentary evidence seems to indicate that literacy rates were very low in medieval Europe and in Roman times (Clark, 2007, p. 175-181). The rise of mass literacy in the lead up to the Industrial Revolution in Northwestern Europe is to my knowledge a unique event in history. It seems likely that this rise of mass literacy may have played an important role in enabling increased innovation and productivity in the leadup to the Industrial Revolution.

The most important channel through which the printing press affected literacy was, however, almost certainly the enormous fall in the price of books that it caused. Clark and Levin (2011) have gathered data on the price of books in England from

1350 to 1835. They use these data to construct a price index that provides an estimate of the change in the price of a book of constant quality over this period. According to their estimate, the price of books fell by a factor of 10 between the second half of the 14th century and the first half of the 16th century. It then fell by another factor of two over the next 50 years. These estimates suggest that a book that cost one pound sterling around 1400 cost only one shilling (5% of a pound sterling) around 1600.

Despite being huge, this estimate probably understates the rise in accessibility of books over this period. As I note above, Clark and Levin are attempting to estimate the price of books of constant quality. Early printed books were of very high quality because they were still luxury goods being marketed mainly to wealthy individuals (and institutions) who were used to buying and reading books that had been transcribed on vellum. Gutenberg, in particular, put a great deal of effort into making his Bible have a very similar look and feel to contemporary Bibles transcribed on vellum. Over time, however, people got used to the new medium and it became less important to try to replicate the old medium. This allowed for smaller print (more words per page), cheaper paper, less calligraphy, cheaper binding, and perhaps most importantly much cheaper forms of written material such as pamphlets and newspapers. All these things together meant that by the 18th century written materials (pamphlets, newspapers, magazines, and books) were affordable to a much larger portion of the population than a few hundred years earlier (when the price of a book was similar to the price of a horse). The fact that people could afford books surely massively increased the demand for being able to read.

2.2 Protestant Prosperity

The previous section puts forth a rather complex theory of causal links from the invention of the printing press to the Industrial Revolution running through the Reformation, the Enlightenment, and the rise of mass literacy. I supported this theory with some narrative history as well as data on literacy and the price of books. My hope is that you find this theory plausible and intriguing. However, there is surely a great deal of variation among you—the readers of this book—as to how plausible you find this theory and how important a role you think it played in bringing about the Industrial Revolution. After all, the amount of evidence I have presented supporting the theory is relatively small. In truth, I don't think the amount of existing evidence on this topic is close to conclusive. However, there does exist some clever and intriguing evidence that supports different parts of this theory. Below, I describe in some detail an interesting recent paper by Becker and Woessmann (2009) that provides evidence on the role of Protestantism and literacy.

Writing at the turn of the 20th century, Max Weber observed that Protestant areas were more prosperous than Catholic areas (Weber, 1905/2002). He then famously argued that this difference was due to a "Protestant ethic" that resulted in a "spirit of capitalism" among Protestants, i.e., Protestants working harder, saving more, and pursuing economic gain for its own sake. The Reformation had introduced the notion of a "calling" which suggested a religious sanctification of labor. According to Weber, this created a strong work ethic specific to Protestants. The Protestant "recognizes, as the *only* means of living a life pleasing to God ... the fulfillment of innerworldly duties which arise from the individual's station in life. This then becomes one's "calling"" (p. 29). Weber contrasted this with other religions where religious devotion typically involved rejecting worldly affairs for "morality through the pursuit of monastic asceticism."

Weber's ideas have been hugely influential within social science over the past 100 years. The premise—that Protestant areas were more prosperous around 1900—is true. However, it does not necessarily follow from this that Protestantism causes prosperity. (A correlation between two variables does not prove that one causes the other.) The complicating issue is that areas that were Protestant in 1900 may differ in other ways from Catholic areas—e.g., they may have better weather, more natural resources, or better institutions—and it may have been these other differences that led the Protestant areas to become more prosperous.

This problem is called the *omitted variables* problem: There may be other unobserved variables that cause prosperity and are correlated with Protestantism. If this is the case, the correlation between Protestantism and prosperity may simply reflect the influence of these other variables. For example, suppose that the true determinant of differences in prosperity around 1900 was some hard-to-measure aspect of the weather: Places with good weather were more prosperous than places with bad weather. Suppose furthermore that Protestantism happened to have spread in places with good weather (or that good weather helped spread Protestantism). In this case, prosperity would be correlated with Protestantism without there being any true causal link between the variables. The true cause of prosperity would be weather, which just happened to be correlated with Protestantism (or also caused Protestantism).

How can we tell whether Weber was right about Protestantism causing prosperity? As with all theoretical hypotheses about how the world works, we need to look for empirical evidence that either supports or contradicts Weber's thesis. The scientific ideal would be if we could run a randomized controlled experiment. In this particular context that would mean: 1) having a large sample of regions, 2) randomly picking some of them and exposing these to Protestantism, 3) waiting 250 years and seeing if the regions that were randomly exposed to Protestantism (but otherwise the same on average) ended up becoming more prosperous. Obviously, this type of randomized controlled experiment is not feasible in practice.

Since true experiments are seldom feasible in economics, economists have developed alternative methods to provide evidence on questions of causation. The key idea that underlies these methods is the idea of a *natural experiment*. For the question we are discussing—whether Protestantism causes prosperity—a natural experiment is some fluke of nature or fluke of history that yields variation in the propensity of regions to be Protestant but is otherwise unrelated to prosperity. More specifically, we are searching for a variable that has two characteristics:

1. *Relevance*: It is correlated with variation in Protestantism,
2. *Exogeneity*: It is uncorrelated with all other determinants of prosperity.

A variable that satisfies these two conditions is called a valid *instrumental variable* (or *instrument* for short). The idea is to focus on a component of the variation in Protestantism that is uncorrelated with other determinants of prosperity (uncorrelated with all omitted variables). The instrument is a tool to capture only this component of the variation in Protestantism. In other words, the instrument is a tool that allows us to vary only a single potential determinant of prosperity (Protestantism). If the instrument is valid (i.e., uncorrelated with other determinants of prosperity), then that means that on average all things other than Protestantism are equal as we vary the instrument. As a consequence, if the instrument is correlated with prosperity, it must be that this correlation arises because Protestantism causes prosperity. This is a subtle idea. The next several pages will illustrate it through example.

First, let us consider what kind of variable might satisfy the relevance and exogenity conditions for being a valid instrument. To get some practice thinking about this, it is useful to consider an example of a variable that is not a valid instrument (i.e., doesn't satisfy the conditions described above). Consider, for example, the level of literacy at the time of the Reformation as a candidate instrumental variable. It is plausible that the level of literacy of different regions at the time of the Reformation predicts their propensity to adopt the new religion. Perhaps more literate regions were more likely to adopt Lutheranism because Luther's emphasis on reading the

Bible in vernacular appealed to them more strongly. If this was the case, literacy at the time of the Reformation satisfies the relevance condition of a valid instrument. It seems less likely to be true, however, that literacy at the time of the Reformation satisfies the exogeneity condition. To satisfy this condition, it must be the case that literacy at the time of the Reformation does not affect the subsequent prosperity of the regions in question through any other channel than potentially the adoption of Protestantism. It is easy to think of plausible violations of this idea. Literacy may in fact affect subsequent prosperity through many channels. For example, literacy allows farmers to more easily learn about new crop rotations, literacy facilitates the use of contracts which may be crucial for many forms of investment, and literacy allows traders to keep better records and learn about basic accounting practices. For all of these reasons (and many more) it is likely that literacy affects prosperity. This means that literacy at the time of the Reformation is not a good candidate for an instrumental variable for Protestantism.

Finding a variable that simultaneously is relevant and exogenous can be quite difficult. A large part of the ingenuity involved in empirical economics has to do with thinking of clever instruments and gathering the data needed to measure them. For the question we are interested in, Becker and Woessmann (2009) propose a very clever instrument: distance to Wittenberg. Wittenberg was the city where Luther lived and worked. Becker and Woessmann's idea is that places closer to Wittenberg are more likely to be Protestant because they were more strongly exposed to Luther's ideas. At the time, travel and the transmission of information was slow and costly. Luther's ideas were spread partly by thousands of students and scholars who came to Wittenberg to hear Luther's sermons and speeches. Many priests also came to Wittenberg to be officially ordained. High cost of travel made these trips much less feasible for those living further away from Wittenberg. As a consequence, the Protestant Revolution diffused from Wittenberg. This logic makes it likely that distance to Wittenberg satisfies the relevance condition for being a valid instrument, at least when applied to regions within Germany.

On the other hand, it is generally accepted by historians that, before Luther, Wittenberg was an unimportant place. This makes it plausible that distance to Wittenberg did not affect subsequent prosperity through other channels than Protestantism and therefore satisfies the exogeneity condition for being a valid instrument. It is generally not possible to prove empirically that an instrument satisfies the exogeneity condition. Researchers can make arguments—such as the one Becker and Woessmann make and I repeat above—that a variable is likely to be exogenous.

Researchers can also present evidence suggesting that—in our context—regions further away from Wittenberg are not systematically different along various observable dimensions than places closer to Wittenberg. But there are always many unobservable dimensions that are left unexplored for reasons of data limitation: It is rarely possible to make sure everything else is held constant in a natural experiment. This means that ultimately the notion that a variable satisfies the exogeneity condition must always to some degree be assumed.

Given a candidate instrumental variable, the basic structure of the empirical argument usually involves three important steps: a first stage regression, balance tests, and a reduced form regression. In our case, these steps are:

1. *First Stage*: Is distance to Wittenberg correlated with Protestantism?
2. *Balance Tests*: Were areas closer to Wittenberg different in other ways?
3. *Reduced Form*: Is distance to Wittenberg correlated with prosperity?

Becker and Woessmann present empirical evidence that sheds light on these questions using data on 452 counties of Prussia, the dominant German state around 1900. Let's start by discussing their evidence for the first stage relationship. Figure 8 shows a rough depiction of the share of the population that was Protestant across Prussian counties according to the Prussian census of 1871. The location of Wittenberg is depicted by a dot close to the center of the figure. The figure reveals a striking pattern. The vast majority of counties close to Wittenberg have a population share of Protestants that is above 85%, while counties further away from Wittenberg are much more likely to have a lower Protestant share of the population. The figure, therefore, strongly suggests that distance to Wittenberg is correlated with Protestantism within Prussia in 1871.

We can formally test whether distance to Wittenberg is correlated with Protestantism by running the following (first-stage) regression

$$\text{PROT}_i = \alpha + \beta \text{DIST}_i + \mathbf{X}_i' \gamma + \epsilon_i, \tag{1}$$

where PROT_i denotes the population share of Protestants in county i, DIST_i denotes county i's distance to Wittenberg in kilometers, $\mathbf{X}_i$ denotes a vector of control variables, and ϵ_i is an error term. The coefficient of interest in this regression is β, which measures the extent to which distance to Wittenberg is correlated with the Protestant population share conditional on the controls in $\mathbf{X}_i$. Table 2 presents Becker and Woessmann's estimate of this regression. Their estimate of β is -0.095. This means that on average the share of Protestants drops by 9.5 percentage points for each 100

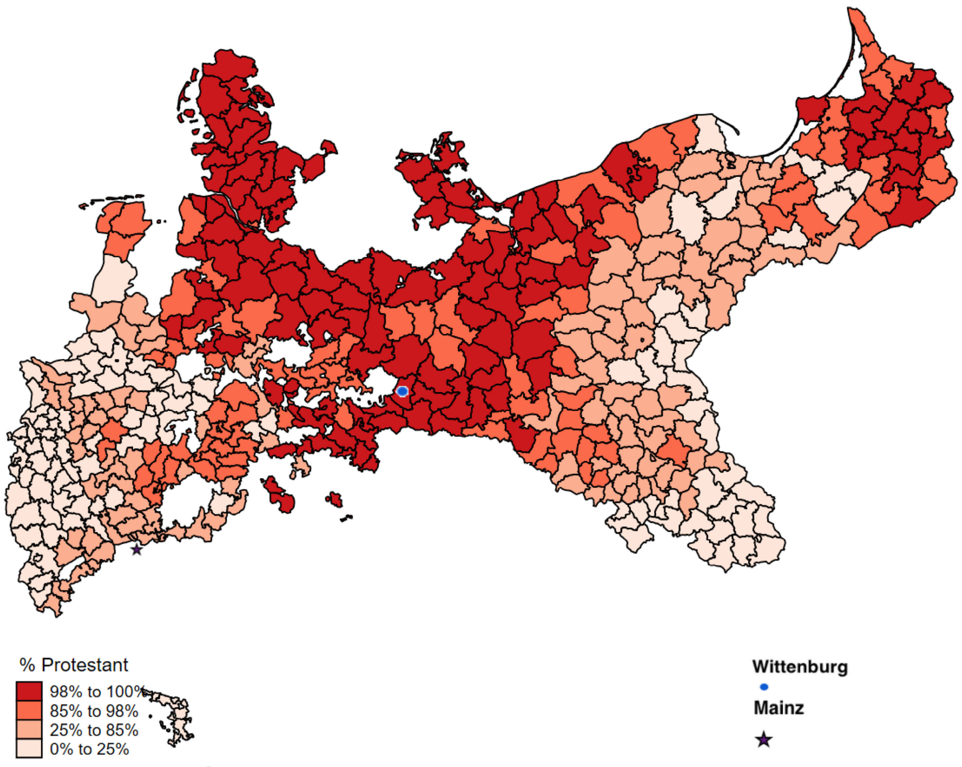

Figure 8: Protestant Population Share in Prussia in 1871

Note: County-level depiction based on 1871 Population Census. See Appendix A of Becker and Woessmann (2009) for further information about the data.

km one moves away from Wittenberg. The standard error on Becker and Woessmann's estimate is 0.011. This is an estimate of the statistical uncertainty associated with Becker and Woessmann's estimate of β. A common convention (rule of thumb) in economics is to consider values of β that are more than roughly two standard errors away from the estimate rejected, but values within roughly two standard errors not rejected. Adopting this rule of thumb, we can say that Becker and Woessmann's data reject values of β outside of the interval $[-0.073, -0.117]$. In particular, the notion that there is no relationship between the population share of Protestants and distance to Wittenberg—i.e., that $\beta = 0$—is soundly rejected by Becker and Woessmann's data. When a variable is being used as an instrument, a stricter statistical standard is usually applied. In these cases, a common rule of thumb is that the F-statistic for the hypothesis that $\beta = 0$ should be larger than 10 (which roughly corresponds to β being more than three standard deviations away from zero). The F-statistic in Becker and Woessmann's first stage is 74.2. Since the F-statistic is above

Table 2: Explaining Protestantism

Dependent Variable:	Share Protestants
Distance to Wittenberg in km	-0.095
	(0.011)
Number of Observations	452
1st-Stage F-Statistic	74.2

Notes: These estimates are taken from Table III in Becker and Woessmann (2009). Standard errors are in parentheses. The control variables included in the regression are: % age below 10, % Jews, % females, % born in municipality, % of Prussian origin, average household size, ln(population size), population growth from 1867-1871 in %, % missing education info, % blind, % deaf-mute, % insane.

10, we say that distance to Wittenberg is a "strong" instrument.

To corroborate their assumption that distance to Wittenberg is exogenous, Becker and Woessmann present a number of balance tests. The balance tests take the form of regressions of various characteristics of Prussian counties prior to the Reformation on distance to Wittenberg. The idea is that if distance to Wittenberg is to be considered exogenous in this context, it should not be correlated with observable economic characteristics of Prussian counties prior to the Reformation that may themselves have affected the subsequent prosperity of these counties. Formally, the balance test regressions take the form

$$Y_i = \alpha + \beta \text{DIST}_i + \mathbf{X}_i'\gamma + \epsilon_i, \tag{2}$$

where Y_i is an economic characteristic of county i.

Becker and Woessmann present nine such balance tests. Table 3 presents the results for four of these. The first column presents results for a regression where the dependent variable Y_i is an indicator for whether a county was a free imperial city (Reichsstädte) in 1517—i.e., the variable is equal to one if the county was a free imperial city at this time and zero otherwise. The results presented in the second column are for a regression where the dependent variable is a measure of urbanization in 1500: urban population per km^2. The dependent variable for the third column is an indicator for whether there was a university in the county and the dependent variable for the fourth column is an indicator for whether there was a school in the county. All four of these measures are chosen as proxies for economic or educational development.

In all four cases, the estimated coefficient on distance to Wittenberg is small in magnitude and also small relative to the coefficient estimate's standard error. In

Table 3: Was Wittenberg a Random Place?

Dependent Variable:	Imperial City in 1517	Urbanization in 1500	University in 1517	School in 1517
Distance to Wittenberg	0.0034	0.00006	-0.0019	-0.0073
(in 100 km)	(0.0071)	(0.00013)	(0.0047)	(0.0099)
Number of Observations	452	452	452	333
R^2	0.0005	0.0004	0.0004	0.002

Notes: These estimates are taken from Table IV in Becker and Woessmann (2009). Standard errors are in parentheses. The dependent variables are: an indicator for whether a county was a free imperial city, urban population per km^2, indicator for whether county had a university, and indicator for whether the county had a school.

none of these four cases (or the other five reported in Becker and Woessmann) is the coefficient estimate larger in absolute value than two times the standard error. As we discussed above, economists typically consider values further than roughly two times the standard error away from an estimated coefficient to be rejected by the data at hand. Values closer than this cannot be rejected due to statistical uncertainty. The fact that the estimated coefficients in the four balance tests reported in Table 3 are smaller in absolute value than two times their standard errors means that in none of these four cases can the value zero be rejected. In cases like these, we say that the coefficients are not statistically significantly different from zero. This means that none of these balance tests provide statistically significant evidence against the idea that distance to Wittenberg is exogenous. Distance to Wittenberg, therefore, passed these balance tests, which corroborates Becker and Woessmann's claim that it is a valid instrument.

We now turn to the third step of the argument: is distance to Wittenberg correlated with prosperity in the 19th century? We can formally test this by running the following (reduced-form) regression

$$\text{PROSP}_i = \alpha + \beta \text{DIST}_i + \mathbf{X}_i' \gamma + \epsilon_i, \tag{3}$$

where PROSP_i is a proxy for prosperity of county i in the late 19th century. Becker and Woessmann (2009) present results for three different proxies for prosperity. Their main proxy is income tax revenue per capita in 1877. The other two measures they use are the logarithm of average annual incomes for male elementary school teachers in 1886 and the share of the labor force working in manufacturing and services in 1882. Table 4 presents estimates for all three of these measures. The

Table 4: Does Distance to Wittenberg Predict Prosperity

Dependent Variable:	Income Tax per capita	Log Teacher Income	Share Manuf. and Services
Distance to Wittenberg	-6.0	-1.00	-0.78
(in 100 km)	(2.3)	(0.48)	(0.36)
Number of Observations	426	452	452

Notes: These estimates are produced using replication code and data provided by Ludger Woessmann. They are the reduced form estimates that correspond to the instrumental variables (IV) estimates presented in Table V of Becker and Woessmann (2009) except that I have included "% missing eduction info" as an additional control for consistency with the first stage reported in Table 2. Income tax per capita is measured in pfennig. Standard errors are in parentheses. The control variables included in the regression are: % age below 10, % Jews, % females, % born in municipality, % of Prussian origin, average household size, ln(population size), population growth from 1867-1871 in %, % missing eduction info, % blind, % deaf-mute, % insane.

estimates for all three proxies indicate that distance to Wittenberg is indeed correlated with prosperity. Specifically, counties that are further away from Wittenberg on average have lower income tax revenue per capita, lower teacher salaries, and a smaller share of the workforce working outside of agriculture.

Let's consider the estimate for income tax revenue in more detail. The point estimate in this case is -6.0. This means that income tax revenue per capita falls on average by 6 pfennig (0.06 Marks) for each 100km one moves away from Wittenberg. The standard error on this coefficient estimate is 2.3. The estimate is, therefore, considered statistically significantly different from zero (since it is further away from zero than two times the standard error). But is the relationship quantitatively meaningful? The average distance to Wittenberg across counties in Becker and Woessmann's data is 326km and the standard deviation of the distribution of distance to Wittenberg is 149km. Moving one standard deviation further away from Wittenberg, therefore, on average lowers per capita tax revenue by roughly 9 pfennig. Average per capita tax revenue across counties was 198 pfennig and the standard deviation of per capita tax revenue 70 pfennig. Moving one standard deviation further away from Wittenberg, therefore, on average lowers per capita tax revenue by roughly 0.13 standard deviations. While this is not huge, it is certainly a meaningful amount. Another way to gauge quantitative magnitude is to ask how much income tax revenue per capita falls on average when one moves from the 10th percentile of the distribution of distance to Wittenberg to the 90th percentile. This turns out to be roughly 22 pfennig, or about 11% of average per capita tax revenue.

The results in Tables 2 and 4 show that distance to Wittenberg is correlated with Protestantism and is also correlated with prosperity. If distance to Wittenberg is a valid instrument, then it is not correlated with any other determinant of prosperity (i.e., it is uncorrelated with all omitted variables). This means that when we move from cities that are close to Wittenberg to cities that are far away from Wittenberg, the only potential determinant of prosperity that is changing on average is Protestantism. The fact that prosperity is changing as we move away from Wittenberg must therefore be due to the change in Protestantism (everything else is being held constant on average). We can therefore conclude that Becker and Woessmann's data support Weber's theoretical conjecture that Protestantism causes prosperity. Obviously, this conclusion relies heavily on Becker and Woessmann's claim that distance to Wittenberg is a valid instrument.

By combining information from the first stage regression and the reduced form regression, we can produce a quantitative estimate of the effect of Protestantism on prosperity. In other words, we can answer the question: How much does prosperity increase if Protestantism increases by (say) ten percentage points? The easiest way to see this is to consider the causal chain from distance to Wittenberg to prosperity. The first link in the chain is from distance to Wittenberg to Protestantism. The second link in the chain is from Protestantism to prosperity. The assumption that distance to Wittenberg is exogenous—i.e., uncorrelated with all determinants of prosperity other than Protestantism—then implies that

$$\begin{aligned}&\text{Effect of DIST on PROSP}\\&\quad = (\text{Effect of DIST on PROT}) \times (\text{Effect of PROT on PROSP}).\end{aligned}$$

Rearranging yields

$$\text{Effect of PROT on PROSP} = \frac{\text{Effect of DIST on PROSP}}{\text{Effect of DIST on PROT}}. \tag{4}$$

The numerator on the right-hand-side of this equation is the regression coefficient in the reduced form regression, while the denominator is the regression coefficient in the first stage. This tells us that dividing the reduced form regression coefficient by the first stage regression coefficient yields an estimate of the causal effect of Protestantism on Prosperity.

There is a more direct way to estimate the effect of Protestantism on prosperity. This is a procedure called an instrumental variables (IV) regression. Using this one-step method makes it easier to construct appropriate standard errors for the estimate. Becker and Woessmann report such IV estimates, which I reproduce in Table

Table 5: IV Estimates of the Effect of Protestantism on Prosperity

Dependent Variable:	Income Tax per capita	Log Teacher Income	Share Manuf. and Services
Share Protestant	0.62	0.11	0.08
	(0.24)	(0.05)	(0.04)
Number of Observations	426	452	452

Notes: These estimates are taken from Table V in Becker and Woessmann (2009). Income tax per capita is measured in pfennig. Coefficients in the second column are multiplied by 100. Standard errors are in parentheses. The control variables included in the regression are: % age below 10, % Jews, % females, % born in municipality, % of Prussian origin, average household size, ln(population size), population growth from 1867-1871 in %, % blind, % deaf-mute, % insane.

5. We see that raising the share of Protestants by 10 percentage points increases income tax revenue per capita by 6.2 pfennig. This estimate is statistically significantly different from zero.

My discussion of Becker and Woessmann's work to this point has focused on establishing a causal link between Protestantism and prosperity. This is not, however, the main point of Becker and Woessmann's paper. Their central argument is not to confirm Weber's theory but rather to critique it and propose an alternative. Weber did not merely argue that Protestantism caused prosperity. He argued that the reason why Protestantism caused prosperity was that it fostered a "Protestant ethic" that resulted in a "spirit of capitalism" among Protestants. It is this part of Weber's theory that Becker and Woessmann take issue with. They instead propose that the channel through which Protestantism causes prosperity is that Protestantism encouraged literacy.

To support this idea, Becker and Woessmann report an IV estimate of the causal effect of Protestantism on literacy in Prussia in 1886. I reproduce this estimate in Table 6. The finding is that Protestantism did in fact have a large causal effect on literacy in late 19th century Prussia. The IV regression coefficient is 0.19 and is highly statistically significant. This estimate implies that conversion of the entire population of a region from Catholicism to Protestantism resulted in a 19 percentage point increase in literacy in the late 19th century.

Becker and Woessmann argue that this effect on literacy is large enough to explain the bulk of the difference in prosperity between Catholic and Protestant areas, leaving little room for other explanations such as Weber's notion of the Protestant ethic. They point out that modern causal estimates of the effect of education on in-

Table 6: Effect of Protestantism on Literacy

Dependent Variable:	Share Literate
Share Protestant	0.19
	(0.03)
Number of Observations	452

Notes: This IV estimate is taken from Table III in Becker and Woessmann (2009). Standard errors are in parentheses. The control variables included in the regression are: % age below 10, % Jews, % females, % born in municipality, % of Prussian origin, average household size, ln(population size), population growth from 1867-1871 in %, % blind, % deaf-mute, % insane, % missing education info.

come (based on IV regressions) turn out to yield estimates that are very similar to regular (OLS) regression estimates of this relationship (Card, 1999). In other words, the omitted variables bias in modern OLS estimates of the effect of education on income turns out to be close to zero. Becker and Woessmann argue that if the same is true in their setting, they can adjust their income measures for the effect that Protestantism has on income through literacy by estimating the effect of literacy on earnings using a simple (OLS) regression and subtracting the estimated effect of literacy from their income measures. Once they have done this, they can then estimate the causal effect of Protestantism a second time using this adjusted measure of income. When Becker and Woessmann do this, they find that Protestantism has very little (additional) effect on income. In other words, once they have adjusted for the effect of Protestantism that flows through literacy, they find that there is very little effect left over for other potential explanations. While interesting, this argument relies on the rather strong assumption that the causal effect of literacy can be estimated using an OLS regression in their setting.

To summarize, Becker and Woessmann (2009) provide causal evidence for two of the causal links in the theory that I presented in section 2.1. They show that the Protestant revolution caused a large increase in literacy and a more modest increase in prosperity. To be precise, they estimate the *relative* effect of Protestantism in regions that were more strongly treated by the Reformation. How large an *aggregate* effect the Reformation had is actually not clear from their evidence alone. The effects of the Reformation almost surely also affected areas that remained Catholic, implying that the aggregate effect may be much larger than the relative effect.

As I have emphasized above, Becker and Woessmann's conclusions rely heavily on the assumption that distance to Wittenberg is a valid instrument for the effect of Protestantism on prosperity and literacy. Earlier in this section, I discussed how the printing press likely played an important role in the success and spread of the Reformation. Rubin (2014) presents systematic causal evidence supporting this notion. Furthermore, Dittmar (2011) presents causal evidence that cities with printing presses grew substantially faster in the 16th century. Together these results suggest that the printing press may be an important omitted variable in Becker and Woessmann's analysis. In other words, perhaps it is greater exposure to the printing press that caused Protestant areas to be more prosperous, but Protestantism itself had no effect on growth.

Dittmar and Rubin use distance to Mainz—the city where Gutenberg established the first printing press—as an instrument for the printing press in their analyses. One simple way to assess whether the printing press is an important omitted variable in Becker and Woessmann's analysis is to include distance to Mainz as a control in the IV regressions that Becker and Woessman run. I have rerun the regressions presented in Tables 2, 4, and 5 with this extra control. The first stage regression is virtually unchanged. However, the reduced form and IV regressions are considerably weakened. Focusing on the results for income tax per capita, including distance to Mainz as a control leads the coefficient on distance to Wittenberg in the reduced form to fall from -6.0 to -1.2 and become insignificantly different from zero. Likewise, the IV estimate of the causal effect of Protestantism on income tax per capita falls from 0.62 to 0.11 and becomes insignificantly different from zero. In the reduced form, the coefficient on distance to Mainz is actually much more statistically significant than the coefficient on distance to Wittenberg when they are both included in the regression.

These results illustrate that it is difficult to disentangle the impact of the printing press and the impact of Protestantism on growth. It may be that the printing press affected growth through its influence on the spread of Protestantism. But it may also be that the printing press influenced growth through other channels and the results of Becker and Woessmann are driven by the fact that the printing press caused both growth and Protestantism. In this case, Protestantism itself may not have had any direct influence on growth.

3 The Agricultural Revolution

The population of England more than tripled between 1700 and 1850 from about 5.5 million to about 17.5 million. Throughout much of this period, imports of food in excess of exports made up a very small fraction of total food consumption in England. Given this, the explosion of the population could not have occurred without an explosion in agricultural output. This explosion in agricultural output was driven by large increases in both crop yields and agricultural labor productivity. Average crop yields in wheat, barley, and oats roughly doubled over this period (Broadberry et al., 2015, Table 3.06). The fraction of the English population engaged in agriculture fell substantially from about 55% to well below one third (Allen, 2000), implying that labor productivity in English agriculture must have risen substantially. The large increase in crop yields and agricultural labor productivity that occurred over the 150 year period after 1700 is referred to as the *Agricultural Revolution*.

Why did the Agricultural Revolution occur? Let's begin by considering the proximate causes and then work our way towards deeper causes. The proximate causes of the Agricultural Revolution are not particularly controversial. Several developments contributed. The most important of these are likely to have been improvement of land, more intensive crop rotations, and the breeding of higher yield animals and crops. During the 18th century, 8 million acres of 'waste' were converted to 'pasture' more than doubling the area of improved pasture in England (Allen, 2009, p. 62-63). This conversion involved substantial improvement of the land as fields were enclosed by walls and cleared of surface stone. In addition, large areas were drained so as to make them suitable for crops or to improve yields.

The traditional crop rotation in England was a three field rotation. Each field rotated between a winter grain such as wheat or rye, a summer grain such as barley or oat, and lying fallow. Thus, a third of the arable land lay fallow at any given point in time. Letting land lay fallow on a regular basis is one way to allow the land to maintain its fertility over long periods of time. During the Agricultural Revolution, the three field rotation was replaced by a more intense system of "convertible husbandry" (sometimes also called "alternative husbandry" or the "Norfolk system"). This system involved alternating fields between arable and pasture, the sowing of grass seeds, as well as the introduction of new crops, most notably roots (such as turnips) and legumes (such as clover). A prototypical example of this new rotation was a four-course rotation of wheat, turnips, barley, and clover (Timmer, 1969).

This system had two primary benefits. First, it allowed farmers to keep much larger stocks of animals resulting in more meat, dairy, and wool (turnips and clover were used mainly as fodder for the animals). Second, it allowed farmers to drastically reduce the amount of land that was left fallow (or completely eliminate the fallow) without this resulting in diminished soil fertility. Legumes have nitrogen-fixing qualities and therefore improve soil fertility. In addition, the fact that large herds of animals were grazed on the land as part of the rotation (and kept through the winter) yielded large amounts of natural fertilizer in the form of animal droppings. The new system actually did more than just maintain the fertility of the soil. It slowly increased the nitrogen content of the the soil, which improved grain yields over time (Allen, 2008).

But why did these large improvements only occur in the 18th and 19th centuries, and not earlier? Narrative evidence as well as data on crop yields and labor productivity indicate that improvements in English agriculture were much slower prior to the 18th century. Economic historian Jan de Vries has pointed out that the key principles employed to intensify the crop rotation were known long before the 18th century and used in centers of advanced agriculture such as Flanders and Lombardy. In fact, these principles were described in Roman agricultural handbooks (de Vries, 1976, p. 39-40). For some reason, however, these methods were not widely employed in early modern England.

One possibility is that the knowledge of these principles was not widespread. As we discuss in section 2, the technology for maintaining and spreading knowledge was quite primitive prior to the invention of movable-type printing around 1450. De Vries discusses how many books were published about improved agricultural productivity in the 17th and 18th centuries, the most famous perhaps being Jethro Tull's 1731 book *Horse-Hoeing Husbandry*. Spreading innovative ideas about efficient crop rotations in the 18th century took time and a great deal of effort by agricultural reformers. Furthermore, adapting these ideas to local conditions required significant experimentation in each location (Timmer, 1969).

An influential traditional view among economic historians holds that the slow growth in agriculture before the 18th century was due to institutional inflexibility resulting from the open field system of farming that was widespread in England (and much of Northern Europe). According to this view, the *Enclosure Movement* was a crucial institutional development that unleashed the productive potential of English agriculture and, according to some, caused the Industrial Revolution by freeing up labor for use in other activities than food production.

3.1 The Enclosure Movement

Before 1500 much of the land of England and Northern Europe was farmed under the ancient system of open fields. In this system, the unit of settlement was a village as opposed to a farm. Around each village there were typically three large fields. These fields were divided into a large number of small strips of land. Each farmer in the village farmed several of these strips. The three fields served as crop rotation units in the traditional three field rotation system. In addition to these three fields, the village had some meadow, pasture, woodlands, and 'wastes.' This land was managed communally by the villagers. The villagers grazed their livestock on these "commons" as well as on the field that lay fallow and the other fields after harvest time. The system of small strips within larger fields and the practice of opening the fields to grazing after harvest meant that all farmers in the village must in practice follow the same crop rotation and the same schedule regarding sowing and harvesting the field (McCloskey, 1972; Allen, 2001a).

By 1850, this ancient system had been largely swept away in England by some 5,000 acts of Parliament and a similar number of voluntary agreements. It was replaced by a system of privately owned enclosed plots of land "free of village direction" (McCloskey, 1972). This process is commonly referred to as the Enclosure Movement. The shift from ambiguous communal property rights in the open field system to a system of well-defined private property rights spread from England to other countries in the late 18th century and the 19th century.

In England, enclosures had been taking place at least since the thirteenth century. However, judging by the experience of the South Midlands, as documented by Allen (1992), most enclosures occurred in two waves: There was an initial wave between 1550 and 1650 and a second, much bigger wave between 1750 and 1850. Enclosures before 1750 were mostly by voluntary agreement. Landowners would sell each other strips so that they ended up with consolidated holdings. They would then divide up the village commons and renounce common grazing rights on each others land. An important complication was that all those that had grazing rights in the village needed to be party to the agreement (as did the tithe owner). Often enclosure coincided with increases in the concentration of land ownership. Large landowners or men who had made their fortune in trade would buy out all others in a village, bar common grazing, and disband collective decision making.

The large wave of enclosures that occurred between 1750 and 1850 were overwhelmingly by acts of Parliament. In this case, unanimity was not required. Agreement by the owners of roughly 75% of the land in question was sufficient. In these cases, an appointed commissioner would survey the land, redraw boundaries, and allocate the new parcels to all those that owned land or grazing rights in the village in proportion to the value of their prior holdings. Common grazing and collective control was abolished.

The causes and consequences of the Enclosure Movement have been hotly debated ever since the events transpired. Some see the enclosure movement as a crucial institutional change where property rights were 'rationalized,' which improved incentives for innovation and thereby caused the Agricultural Revolution. Others see the enclosure movement as a massive act of expropriation by large landowners of a complex system of ancient peasant rights: the poorest classes of people—cottagers and the landless—typically received no land or very small parcels, but lost formal or informal rights to graze a cow, keep some fowl, and gather firewood on the common. In many cases, this meant that life was no longer viable in the village for these people.

3.2 Did Enclosures Cause the Agricultural Revolution?

Contemporary accounts of agrarian reformers and early economic historians viewed the open field system as being woefully inefficient. The system of scattered strips meant that each farmer's diligence or lack thereof had spillover effects onto his neighbors' strips (externalities). Also, the communal system of grazing led to over-use of the commons (the commons problem). Furthermore, the open field system was thought to be resistant to progress. According to Ernle (1912/1961), "open-field farmers were impervious to new methods." The main reason given for this was the cumbersome nature of decision-making in an open-field village where unanimity was required to make any change to the crop rotation or land use.

Proponents of this traditional view see the Enclosure Movement as a crucial cause of the Agricultural Revolution. If this view is correct, the success of the Enclosure Movement in dramatically increasing productivity in English agriculture is an important piece of evidence supporting the importance of capitalism (i.e., clearly defined property rights) as a precondition for innovation and economic growth.

This traditional view has faced several challenges over the past 50 years. Economic historian Robert Allen gathered evidence on the adoption of new crops and on crop yields in the South Midlands over the period 1450-1850 (Allen, 1992, 2009). Allen's evidence indicates that open field farmers were far from 'impervious to new methods.' They made changes to land use—e.g., introduced new crop rotations—where this was appropriate. Yet, enclosed farms seem to have done so more completely, especially when it came to conversion of land from arable to pasture. Furthermore, Allen's data indicates that while crop yields were higher on enclosed farms than open farms, this difference was a modest fraction of the increase in yields since the middle ages. In other words, open farms managed to attain a large fraction (something like 80%) of the increase in crop yields that enclosed farms attained.

This evidence certainly undermines the traditional view that institutional rigidity of the open field farms made improvements impossible to implement. However, it does not necessarily undermine the argument that enclosure was a crucial driver of increased productivity in agriculture. Imitation is easier than innovation. It may have been that enclosed farms led the way in terms of innovative practices in agriculture. The institutional rigidities of open farms may have been severe enough to prevent innovation but not severe enough to prevent imitation of successful changes implemented by nearby (enclosed) neighbors.

A different empirical argument focuses on the timing of increases in labor productivity in agriculture. To measure labor productivity, one must be able to measure both labor input and production. An important challenge in estimating labor productivity in English agriculture over this period is lack of data on agricultural production at the national level. Since direct data on agricultural production is scant, economic historians have sought to estimate agricultural production indirectly by first estimating food consumption and then making an adjustment for net imports of food. Consider the following simple model for agricultural production:

$$Y_{at} = \phi_t C_{at} N_t, \tag{5}$$

where Y_{at} denotes agricultural output at time t, ϕ_t (the Greek letter phi) denotes the ratio of agricultural production to agricultural consumption at time t, C_{at} denotes food consumption per capita at time t, and N_t denotes the size of the population at time t. This equation states that agricultural output in England (the left hand side) must equal agricultural consumption ($C_{at}N_t$) adjusted for net imports (ϕ_t).

To get an expression for labor productivity in agriculture, we divide both sides of equation (5) by the number of people working in agriculture:

$$\frac{Y_{at}}{N_{at}} = \phi_t C_{at} \frac{N_t}{N_{at}}, \tag{6}$$

where N_{at} denotes the number of people working in agriculture at time t. The left hand side of this equation is labor productivity in agriculture.

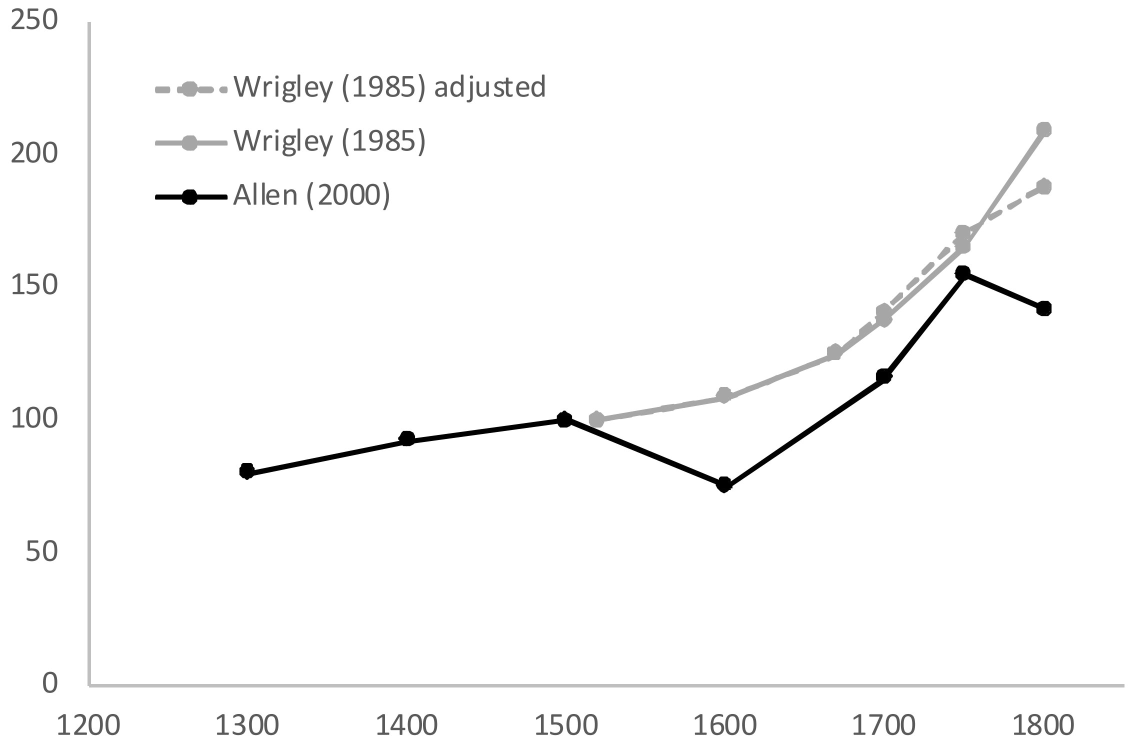

Figure 9: Labor Productivity in Agriculture in England

Note: The solid gray line is from Wrigley (1985). The dashed gray line adjusts the solid gray line to account for net imports and exports of food using estimates from Allen (2000). The solid black line is from Allen (2000).

Wrigley (1985) uses this model to estimate the change in labor productivity in English agriculture from 1520 to 1800. He makes two simplifying assumptions. First, he assumes that net imports of food into England are zero (i.e., $\phi_t = 1$). Second, he assumes that food consumption per person in England was constant over this period. Given these two assumptions, the evolution of labor productivity is simply given by the evolution of the inverse of the fraction of the English population engaged in agriculture.

Figure 9 plots Wrigley's estimates of labor productivity in agriculture. Given Wrigley's assumptions, the sharp drop in the fraction of the agricultural population in England over the 18th century implies that labor productivity in English agriculture rose substantially. The increase from 1700 to 1750 is 20% and from 1750 to 1800 the increase is a further 26%. The large increase over the second half of the 18th century provides indirect support for the importance of the Enclosure Movement since this is when the large wave of Parliamentary enclosures occurred.

An important weakness of Wrigley's estimates is that they rely on assumptions that are unlikely to be true. The eighteenth century was a period of rapid change and this change is likely to have affected per capita food consumption. In particular, per capita food consumption is likely to be significantly affected by changes in real wages and changes in the relative price of food. In addition to this, towards the end of the 18th century, England started to import a significant amount of food on net.

Allen (2000) relaxes the two simplifying assumptions that Wrigley makes. He proposes a simple model of food demand and uses data on real wages and the relative price of food to estimate how per capita food consumption changed over time. He also incorporates estimates of net imports of food. The black line in Figure 9 plots Allen's estimates of labor productivity in agriculture. Allen's estimates turn out to be substantially different than Wrigley's, especially during the 18th century. According to Allen's estimates, labor productivity rose by 34% in the first half of the 18th century (before the largest wave of enclosures), but then fell by 8% in the second half of the 18th century (during the enclosure wave). This time pattern is clearly less consistent with the Enclosure Movement having played a crucial role in the Agricultural Revolution.

The main reason behind Allen's lower estimated growth of labor productivity in the second half of the 18th century is the fact that the relative price of food rose during this period. This indicates that the supply of food was not able to keep up with increases in demand resulting from the sharp increase in the English population over this period. Increased net exports also play a (relatively modest) role. The dashed gray line in Figure 9 adjusts Wrigley's estimates for net imports of food but maintain the assumption of constant per capita food consumption.

While Allen's estimates are preferable in that they allow the demand for food to vary with prices and income, a potential drawback of these estimates is that they are sensitive to errors in the estimation of these variables. There is considerable debate among economic historians about the evolution of real wages in England in the 18th century. Furthermore, recent evidence suggests that an "industrious revolution" occurred in the lead-up to the Industrial Revolution, where hours worked per day and days worked per year increased (Humphries and Weisdorf, 2019). If this is the case, measures of real wages understate the change in real income over this period and therefore understate the increase in agricultural labor productivity.

3.3 Rural Growth and Urban Growth: Did One Cause the Other?

A second aspect of the traditional view about the Agricultural Revolution is the notion that the Agricultural Revolution caused the Industrial Revolution. Ragnar Nurkse stated this view in a particularly colorful manner in the following passage:

> Consider what happened in the original home of industrial development, in England in the eighteenth century. Everyone knows that the spectacular industrial revolution would not have been possible without the agricultural revolution that preceded it. And what was this agricultural revolution? It was based mainly on the introduction of the turnip. The lowly turnip made possible a change in crop rotation which did not require much capital, but which brought about a tremendous rise in agricultural productivity. As a result, more food could be grown with much less manpower. Manpower was released for capital construction. The growth of industry would not have been possible without the turnip and other improvements in agriculture. (Nurkse, 1953, p. 52-53)

It is clearly true that industrial development cannot occur on a large scale without improvements in agricultural productivity. If each farmer grows only enough food to feed his family, there is no surplus of food to feed people in cities. In such a circumstance, there can be no cities and no industry. This notion does not, however, imply that productivity growth in agriculture is necessarily the key *cause* of industrial development. It is quite possible that causation runs the other way.

Let's spell this out in a little more detail. One possibility is that the chain of events is as follows: Productivity in agriculture increases. This pushes some people out of agriculture because the agricultural sector can now feed more people than only those that work in that sector. Equivalently, less farmers are needed to feed a given population. This means that a larger fraction of the population is free to pursue other activities. To the extent that some of these extra people devote themselves to trade, industry, and innovation, this may initiate industrial growth.

Another possibility, however, is that the chain of events is quite different: Returns to urban activities (trade and industry) rise for some reason. This generates an incentive for people in rural areas to move to cities. In other words, the cities pull people in from the country-side. The increased wealth in the cities results in increased demand and higher prices for agricultural products. The increased demand and higher prices of agricultural products strengthens the incentives of farmers to improve the land and innovate, which in turn leads to improvements in agricultural productivity.

In the first of these two stories, the Agricultural Revolution causes the Industrial Revolution. In the second, however, the Industrial Revolution (or the rapid increase in overseas trade we discuss in section 1) causes the Agricultural Revolution. A third story is that both revolutions have a common external cause. For example, the dramatic fall in interest rates in England after the Glorious Revolution (see Figure 3) encouraged investment both in urban and rural improvements and may thereby have jump-started both the Agricultural and Industrial Revolutions.

How can we tell which of these stories is likely to have been most important? Allen (2009) points out that the first two of these stories differ sharply in their implications for urban wages. In the first story, agriculture pushes people into the cities. If this was the dominant story in 18th century England, wages in English cities should have been pushed down due to a glut of people arriving from the countryside. In the second story, however, it is the boom in the cities that pulls people in from the country-side. If this was the case, we should have seen high wages in the cities. (The third story also implies high urban wages.)

To assess whether real wages were high or low in English cities during the Agricultural Revolution, it is useful to compare them with real wages in other areas that did not experience an agricultural revolution. To this end, Figures 10 and 11 present estimates of real wages and labor productivity in agriculture, respectively, in several European countries over the period 1300 to 1850. These estimates were developed by Robert Allen. The estimates of labor productivity in Figure 11 are from Allen (2000) and are based on the same methods as those Allen used to produce the agricultural labor productivity series for England that we plot in Figure 9.

The methods that underlie the estimates of real wages in Figure 10 were first developed in Allen (2001b) and subsequently refined in Allen (2009). Comparing real wages across space is tricky because the lowest cost way to attain a certain level of well-being is different in different places. Recall that measures of real wages are calculated by dividing a measure of nominal wages (e.g., wages in pounds sterling or in grams of silver) by a measure of the price level. The price level is meant to measure the cost of attaining a certain level of well-being. The thing that is tricky about this is that attaining a certain level of well-being in two different places may entail purchasing quite different baskets of goods. Suppose, for example, that we are interested in calculating the cost of subsistence in two places. This involves calculating the cheapest way to get enough calories and other vital nutrients to live

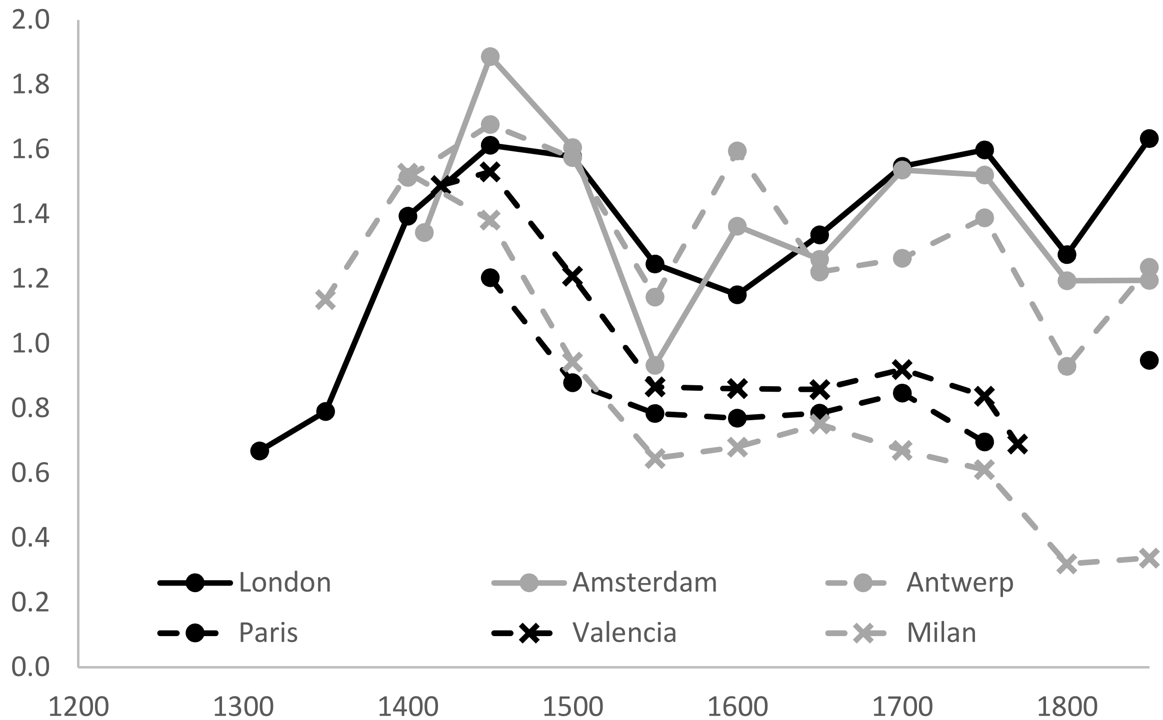

Figure 10: Real Wages in European Cities

Note: These data are estimates of respectability ratios constructed by Robert Allen. The respectability ratios divide nominal income of a laborer that works 250 days per year by the cost of 3.15 respectability baskets. See the text and Allen (2001b, 2009) for more details. Each point is a decadal average.

in these two places. This basket may differ a great deal depending on the relative price of different types of food. In areas that are well suited for growing potatoes, the cheapest way to subsist may involve buying a large amount of potatoes, while in areas that are better suited for growing oats (say) the subsistence basket will involve more oats and less potatoes.

Allen has constructed measures of the cost of subsistence and the cost of a "respectable" standard of living for quite a number of European (and non-European) cities. Figure 10 plots his estimates of "respectability ratios" for six European cities. The respectability ratio is constructed by first calculating the nominal income of a laborer that works for 250 days in a year. This level of income is then compared with the cost of purchasing a basket of goods that yields a "respectable" standard of living. The details of this basket differ across locations due to differences in climate and cuisine. Table 7 presents the respectability basket for northwestern Europe. For areas further south in Europe certain items (such as beer and butter) are replaced by

Table 7: Allen's Respectability Basket for Northwestern Europe

	Quantity per person per year	Spending Share	Nutrients Per Day Calories	Nutrients Per Day Grams of Protein
Bread	234 kg	36.0%	1,571	64
Beans/Peas	52 l	5.5%	370	28
Meat	26 kg	12.8%	178	14
Butter	5.2 kg	4.0%	104	0
Cheese	5.2 kg	3.3%	54	3
Eggs	52	1.1%	11	1
Beer	182 l	20.0%	212	2
Soap	2.6 kg	1.7%	–	–
Linen	5 m	4.8%	–	–
Candles	2.6 kg	2.9%	–	–
Lamp Oil	2.6 l	4.3%	–	–
Fuel	5.0 M BTU	4.6%	–	–
Total			2,500	112

Note: The information in this table is taken from Table 2.1 in Allen (2009). The spending shares are for the prices in Strasbourg from 1745-54. Spending shares in other places and times may be somewhat different. Wine was substituted for beer and olive oil for butter depending on local custom. The fuel ration was set at 2 M BTU in Spain and Italy due to warmer climate. Wheat bread was used in London, Paris, Valencia, and Milan, while rye bread was used in Antwerp and Amsterdam.

other items (wine and olive oil).

The respectability basket in Table 7 is for a single adult male. However, a typical laborer's income supported not only the laborer himself but also his wife and children. To take account of this Allen assumes that three respectability baskets were needed to support a family. Finally, to account for the cost of housing, Allen adds a further 5% to the cost of the respectability basket. The respectability ratio is therefore equal to 250 times the nominal daily wage in a city divided by the cost of 3.15 respectability baskets in the city.[7]

Figure 10 reveals that the 15th century was a high point for living standards across Europe. At this time, laborers could comfortably afford Allen's respectability basket (the respectability ratio was above one). Over the course of the 16th century, living standard fell throughout Europe. But starting around 1600, a marked divergence occured. Living standards remained low or continued to fall in Paris,

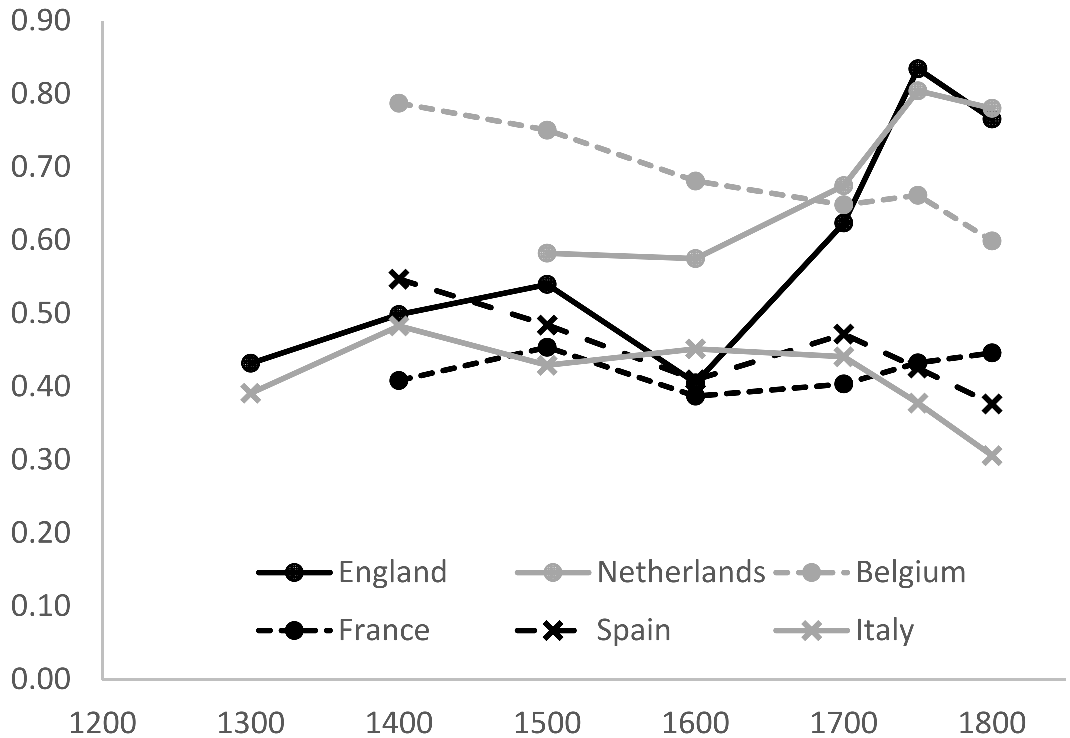

Figure 11: Labor Productivity in Agriculture in Europe Regions

Note: These data are from Allen (2000).

Valencia, and Milan, while they rose substantially in London and Amsterdam. The pattern of low living standards in Paris, Valencia, and Milan is representative of many other European cities. It is England and the Low Countries that are the outliers in terms of high living standards in the 17th and 18th centuries.

Let's now compare these estimates of living standards with the estimates of labor productivity in agriculture reported in Figure 11. This figure reveals that England and the Netherlands were the only two countries in Western Europe to experience an Agricultural Revolution in the 18th century. Other countries experienced no growth or even a decline in agricultural labor productivity during this period. Modern day Belgium is an interesting outlier. It managed much higher productivity in agriculture prior to the 18th century than other countries, but was then overtaken by England and the Netherlands in the 18th century. Labor productivity in other countries remained close to medieval levels.

Comparing Figures 10 and 11, we see that the Agricultural Revolution occurred in exactly the places where real wages were the highest. Allen (2009) argues that these facts are more consistent with the view that urban growth caused the agricultural revolution than with the alternative view that the agricultural revolution caused urban growth. The high wages in London and Amsterdam were pulling people into these cities from the countryside and increasing demand for agricultural products. Allen's argument suggests that it may have been the high demand for agricultural products that induced increased innovation in agriculture and that this may have helped bring about the Agricultural Revolution. We will discuss the idea of such "induced technical change" in more detail in section 4.

3.4 The Decline of Feudalism and the Rise of Capitalism

Medieval England was a feudal society. Formally, all land was owned by the Crown. In practice, however, the Crown granted most of the land to noble lords in exchange for periodic fees and an obligation to provide soldiers in times of war. The land was organized into manors. In most cases, a manor consisted of a village and the surrounding fields and commons. In terms of ownership, the manor consisted of three types of land: the demesne, freeholds, and villein land. The demesne was under the direct control of the lord. Freeholds were farmed by free peasant farmers, while the villein land was farmed by serfs (also referred to as villeins). The lord owned the villein land and to a large extent owned the labor power of the serfs that farmed that land, although the lords were bound by certain customs in their treatment of the serfs.

While medieval freeholds and villein farms were usually small, the demesne was often much larger. On manors with many serfs, the demesne was largely farmed by the serfs (for the benefit of the lord). However, on manors with large demesne and few serfs, the lord hired wage labor. A royal inquisition from the 13th century indicates that in the South Midlands 32% of land was demesne, 40% was villein, while 28% consisted of freeholds (Allen, 1992, ch. 4). Clearly a large fraction of the population of medieval England was shackled in serfdom.

The 600 year period between the Norman Conquest of 1066 and the English Civil War was a period of gradual but ultimately radical change in the structure of property rights and the extent of rule of law in England, as well as the degree of personal freedom of the English population. Early in this period, property rights were limited at all levels. At the top of the hierarchy, the lord could not sell or bequeath his manor without consent of his superior lord (often the Crown), and in some cases such consent required the payment of a fee. The same was true of freeholds. These could not be sold, leased, or bequeathed without the consent of the lord of the manor.

Over time, however, the property rights of lords vis-a-vis the Crown and freehold farmers vis-a-vis the lord strengthened. For freehold farmers, the gradual increase in the importance of royal courts was a crucial development. Prior to the rise of royal courts, the only source of justice available was seigneurial or manor courts (i.e., courts administered by the lord). This obviously meant that individuals had no recourse when it came to disputes with the lord, which gave the lord a great deal of power. Fukuyama (2011, ch. 17) discusses how after the Norman Conquest, the Kings of England gradually established courts of appeal: the royal courts. This was advantageous to the King for several reasons: it helped establish his prestige and legitimacy; and it generated a non-trivial amount of revenue for the Crown. Over time, the rulings of the royal courts established the English Common Law; common because it applied universally across all of England in contrast to precedents from seignueral courts.

By the 14th century, freehold farmers had become the full owners of their property: they could use it as they pleased, lease it, bequeath it, and sell it (with relatively minor restrictions). The situation of serf farmers, however, was much less favorable. Their only protection was custom. They could not seek justice from royal courts if disputes arose with the lord. The lord could expel them from their land as he pleased. The serf was subject to arbitrary fines (tallage), restrictions on marrying outside the manor (merchet), arbitrary requests for labor service on the demesne, corporal punishment, and imprisonment by the lord. Finally, serfs could be sold as chattel. See Allen (1992, ch. 4) for more detail.

Allen (1992) argues that this sorry state of affairs for serfs in England "collapsed under the impact of the Black Death of 1348-9 and the subsequent century of population decline." At first after the Black Death, lords sought to fill vacant farms by promoting landless cottagers. However, as depopulation continued, the grip of lords on their vassals loosened. A peculiar feature of serfdom in England was that a serf's subordinate position only applied vis-a-vis his own lord. In other fiefs, the serf was no different from a free person. Furthermore, there was no 'Fugitive Serf Law' in England to facilitate the return of serfs that ran away. Finally, other lords were quite keen to attract and retain any tenant that was able to pay rent. As a consequence, serfs that managed to flee their lords left their servitude behind and by 1500 feudal serfdom was almost completely eliminated in the midlands.

The labor scarcity of the fifteenth century also led lords to switch from labor intensive arable farming to less labor intensive pastoral farming. This was often done through enclosure. In some cases, areas were enclosed that had been voluntarily abandoned. But in others, enclosures involved eviction and subsequent depopulation of the manor. These 15th and early 16th century enclosures involved significant expropriation of peasant property and customary rights, destruction of villages, and expulsion of their inhabitants. They therefore line up well with the Marxist view of enclosures as exploitation of the poor by the rich.

Allen (1992) argues that the exploitative and depopulating nature of these early enclosures "so alarmed official opinion that the Crown began protecting peasant farmers." This was done through several means. The most effective was the gradual extension of secure property rights to tenant farmers. Copyholds and beneficial leases—the main forms "rental agreements" of this era—became secure forms of tenure. This process limited greatly the ability of lords to expropriate their vassals through enclosure. With enforceable titles, enclosures could only proceed with the agreement of peasant tenants.

Traditionally, copyholders were free tenants that held villein land 'at the will of the lord according to the custom of the manor.' Generally, copyholders paid a modest annual rent and a more substantial fee to the lord when the land was sold or passed to an heir. The determination of these fees was somewhat ambiguous. In many cases, they were governed by custom. But lords had some ability to alter them arbitrarily and could block sales by stipulating a high enough fee. More importantly, villein tenure was traditionally at will. Lords could therefore evict copyhold tenants at any time even in violation of manoral custom without the tenant having any redress.

Over the course of the quarter millennium following the Black Death, the royal courts gradually extended the property rights of tenant farmers in England. They came to enforce the right of tenants to sell their copyhold in exchange for a customary fee. They also eliminated the right of lords to evict tenants in violation of custom. This meant that copyholds either became completely inheritable (copyhold of inheritance) or a lease 'for lives' (copyhold for lives). Copyholds for lives were usually granted for three lives—the farmer's, his wife's, and his son's. When the son took over the farm, the copyhold would be renewed and new lives added (his wife and son) in exchange for a fee. By 1600, copyholds had therefore acquired a substantial proprietary interest in the soil they tilled.

Allen (1992) refers to this process as the 'rise of the yeoman.' He argues that the rise of the yeoman with a substantial proprietary interest in the soil was a key cause of the Agricultural Revolution. Before this development, most land was farmed by serfs. Since the serfs didn't have secure property rights, they had very limited incentives to innovate. If they did innovate, the lord was in a position to rob them of the entire fruits of their innovation. The lord might want to promise to respect property rights. But in this regard, he faced the commitment problem we discussed in section 1.

The Civil War and subsequent period of turmoil marks the formal end of feudalism in England. The first act of Parliament after the Restoration in 1660 abolished the feudal duties of lords vis-a-vis the Crown (Hill, 1961, p. 127). This same act, however, affirmed and secured the remaining rights of the lords vis-a-vis their copyholders and leasees. It is therefore a significant step in the 'modernization' of property rights in England.

The Civil War marked the high point of the 'yeoman farmer' in England. Over the subsequent two hundred years, the yeoman farmer gradually disappeared as the ownership of land in England was radically transformed. This process involved both a massive increase in the concentration of land ownership—the rise of the 'Great Estates'—and also a substantial increase in average farm size. The classic yeoman farmer of the 17th century farmed a moderately sized farm (on the order of 60 acres) and used mostly family labor. By 1800, most farmland was organized into farms that were much larger (on the order of 200 acres) and worked mostly by wage labor. This transformation resulted in a rural society consisting of three classes: wealthy landowners, large-scale tenant farmers, and landless laborers (a rural proletariat). The shift from a feudal to a capitalist agrarian society therefore involved both a transformation of property rights and a transformation of labor relations.

The causes of the massive increase in the concentration of landownership that occurred in England in the 18th and 19th centuries have generated much debate among economic historians. Allen (1992, ch. 5) proposes an interesting explanation: the development of the modern mortgage. He argues that the traditional system of modest rents and periodic fees used by lords in England can be thought of as a way for the lords to borrow from their tenants: the lords received a large up-front payment instead of a higher annual rent. This system was obviously inefficient in that poor tenants were not natural sources of financing. But the lords had no alternative absent a modern mortgage market.

The 17th century saw the gradual development of the modern mortgage. Before that time, mortgages were very short term (six months) and extremely onerous for debtors: if a debtor defaulted even by one day he forfeited his land to the creditor but still owed the principal of the loan! Gradually, the royal courts allowed debtors who had defaulted to recover their land upon repayment of the loan. The most significant shift occurred after the Restoration. Many Royalists had at that point lost their land to creditors (it had been confiscated by Parliament in the Civil War and the Royalists had purchased it back with the help of mortgages that they subsequently defaulted on). The royal courts allowed these Royalists to repossess their land upon repayment. By the end of the 17th century, it had become an established part of the common law that debtors could automatically and indefinitely extend their mortgages as long as they paid interest.

This financial innovation allowed lords to finance the purchase of small freeholds. More importantly, it allowed them to cease renewing copyhold agreements and instead let them 'run out.' The lord could then reorganize the land into rent-maximizing (large) farms and lend them to tenants at market rates. This process involved the lord forgoing the copyhold renewal fees that were otherwise a significant source of his income. His ability to mortgage parts of his land allowed him to maintain his prior level of consumption until the higher rents restored his income (and paid off the mortgages).

The mortgage market no doubt significantly facilitated the concentration of land in England and its reorganization into large capitalist farms. However, if this type of transformation had always been profitable, it seems likely that it would have occurred gradually through self-financed purchases and copyhold conversions by lords even before the development of mortgages. The absence of a gradual process of this type prior to the Civil War suggests that other institutional barriers were important. The feudal rights of the Crown vis-a-vis lords and the threat of arbitrary confiscatory actions by the Crown against successful lords along the lines discussed in section 1 were likely important in this respect.

The idea that the Black Death caused the collapse of feudalism in England was challenged by historian Robert Brenner in a wonderfully insightful and important paper (Brenner, 1976). Brenner points out that while feudal serfdom declined in Western Europe after the Black Death, it intensified in Eastern Europe. If labor scarcity led to the collapse of serfdom in Western Europe, why did the same not happen in Eastern Europe?

Acemoglu and Wolitzky (2011) present an interesting response to this argument. They develop a theoretical model of labor coercion. In their model, lords use force to coerce serfs into accepting unfavorable terms of 'employment.' The coercion is modeled as the lords expending resources to reduce the outside option of the serfs. Acemoglu and Wolitzky show that in their model an increase in the scarcity of labor (e.g., due to a large plague) can either increase coercion or reduce it depending on the relative importance of two effects. On the one hand, labor scarcity increases the value of output which encourages increased coercion by lords. On the other hand, labor scarcity may increase the outside option of the serfs which discourages coercion by lords.

Acemoglu and Wolitzky's model suggests that the divergent paths of serfdom in Western and Eastern Europe after the Black Death are due to different relative effects of this shock on the value of output and the outside option of serfs. As we discuss above, certain peculiar aspects of feudalism in England led to a large improvement in the outside option of serfs in that country after the Black Death: lords found themselves competing for tenants in an environment without a 'Fugitive Serf Law.' This however begs the question why the lords in England did not collectively manage to enact and enforce a Fugitive Serf Law and other repressive collective measures. In Eastern Europe, collective action among the lords succeeded along these lines limiting the increase in the outside option of serfs (see discussion in Brenner, 1976, and Aston and Philpin, 1987). The concurrent expansion of Baltic and North Sea trade may also have played an important role in these divergent paths by increasing the demand for grain in Eastern Europe. Finally, recent palaeoecological research that uses radiocarbon-dated pollen data from roughly 250 coring sites (lakes and wetlands) across Europe to assess changes in land use suggests that mortality dues to the Black Death was lower in Eastern Europe than in Western Europe (Izdebski et al., 2022).

Our discussion in this section has focused on the shift in England from feudal to capitalist economic institutions. To understand why this shift occurred in Europe as opposed to somewhere else in the world, it is useful to push the discussion further back in time. Early human societies were tribal. Fukuyama (2011, ch. 16) discusses the decline of tribal kinship based organization of society and the rise of individualism in Europe. In tribal societies, property rights and personal freedom (e.g., to marry) are typically severely constrained by a person's status and various obligations towards family members. The object of these customs is often to guarantee that the land and other assets of a kinship group remain within the group.

Fukuyama argues that the early decline of these kinship obligations and the rise of individualism is unique to Europe. He furthermore argues—citing Goody (1983)—that this shift was driven by the material interests of the Catholic Church. In the sixth century, the church took a strong stand against various practices that helped maintain property within families in Europe: marriage of close kin, marriage of widows to relatives of their dead husbands, the adoption of children, and divorce. The resulting decline in these practices lead to a large increase in widows and spinsters without male heirs. At the same time the church strongly promoted donations of land and property to itself. Within a short period, the church had become the owner of a substantial fraction of productive land in Europe.

According to Fukuyama, a byproduct of these policies of the church was to weaken kinship ties in Europe and strengthen the rights of individuals. Feudalism arose to fill the void left by the decline in kinship based systems of protection against violence. Much later, the rise of individual property rights and the decline of feudalism was aided by the relative absence of a complex set of kinship based entails. The feudal relationship of vassalage between lord and serf was a relationship between individuals, not between a lord and a kin group.

Acemoglu and Robinson (2019) argue that Europe benefited from a fortuitous fusion of two sets of institutions: state institutions inherited from the Roman Empire and participatory norms and institutions inherited from the Germanic tribes that overthrew the Western Roman Empire. They argue that the key to political development is a Shackled Leviathan: a strong state that is held in check by a strong society. This relies on a balance of power between the state and the people, which is hard to obtain. In most places, the state is either too strong and becomes despotic or social norms and customs limiting the emergence of centralized power—what they call "the cage of norms"—are too strong in which case no state emerges.

4 High Wages and Cheap Coal

Perhaps the most curious aspect of the Industrial Revolution is how geographically concentrated it was. With relatively minor exceptions, all the important industrial innovations of the 18th and early 19th century were British inventions. These include the original atmospheric steam engine invented by Thomas Newcomen, John Smeaton's efficiency improvements of that engine, the separate condenser (James Watt), the high pressure steam engine (Richard Trevithick), the flying shuttle (John Kay), the spinning jenny (James Hargreaves), the water frame (Richard Arkwright), the mule (Samuel Crompton), and the coke smelting furnace (Abraham Darby) to name only a few of the most important innovations. Some of these inventions (most notably the steam engine) relied on scientific discoveries that had mainly been made on the continent. But industrial engineering innovation overwhelmingly occurred in Britain. In short, the Industrial Revolution occurred in Britain.

Why did an explosion of industrial innovation all of a sudden occur in Britain in the 18th century? In 1500, Britain was a relatively unimportant island off the coast of Northern Europe with a very modest population in comparison to, say, France, Northern Italy, the Lower Yangzi Delta in China, the Kanto plain in Japan, or Gujarat in India. By 1800, however, Britain was in the midst of transforming itself into an industrial powerhouse like no other the world had every seen, leaving all other regions of the world far behind. Explaining the concentrated burst of industrial innovation in Britain in the 18th and early 19th centuries is a key element of explaining the Industrial Revolution.

We have already discussed several potential contributors to the Industrial Revolution: changes to institutions, the Enlightenment, and the Agricultural Revolution. Britain's scorecard on being a "special place" when it comes to these three potential contributors is uneven. Consider agriculture. Figure 11 shows that within Europe, England was a typical place when it comes to agricultural productivity until the 17th or 18th century. Belgium held a large lead in agricultural productivity within Europe before 1700. But agricultural productivity was much higher still in the Yangzi Delta (as evidenced by its huge population density). The Enlightenment set Europe apart from other regions of the world intellectually. But it was experienced by a large part of Europe as opposed to just Britain. Of the three potential contributors, the one that most favors Britain is institutional change. Institutional change in Britain over the centuries proceeding the Industrial Revolution—but particularly in the 17th century—resulted in much greater liberty and a more inclusive and capitalist-oriented government in the 18th century than in most other times and places. But is this enough to explain why the Industrial Revolution occurred in Britain?

In a provocative and highly influential book, Robert Allen argues that while increased liberty and a more inclusive and capitalist-oriented government may have been an important contributor to the emergence of the Industrial Revolution, this was not the whole story (Allen, 2009). According to Allen, a crucial additional factor that uniquely favored Britain was the fact that wages in Britain in the 18th century were higher than in most other places in Europe and the price of energy was lower. This combination of wages and prices gave entrepreneurs in Britain a strong incentive to produce labor-saving innovations, and, in particular, to invent ways to replace labor with machinery driven by cheap coal. Allen writes: "The Industrial Revolution, in short, was invented in Britain in the eighteenth century because it paid to invent it there, while it would not have been profitable to invent in other times and places."

Another influential idea, emphasized by historian Kenneth Pomeranz, is that the huge relaxation of ecological constraints on growth that resulted from the discovery of the New World and steam power where a crucial contributor to industrialization in Britain (Pomeranz, 2000). According to Pomeranz, "the British story ... is unimaginable without two crucial discontinuities—one created by coal and one by colonies." A weakness of this idea is that it does not help explain the explosion of innovation—growth in productivity—which is arguably a crucial feature of the Industrial Revolution. In a world constrained by ecology (i.e., a Malthusian world), a relaxation of ecological constraints results in population growth, but not growth in productivity and real wages. We will discuss ways in which the ecological boon associated with the discovery of the New World and the exploitation of fossil fuels where extremely important. But since they don't help explain the sudden acceleration of technical progress, it is hard to see them as providing an explanation for the Industrial Revolution.

4.1 Directed Technical Change

Allen's (2009) argument discussed above may seem curious in that high wages are said to have been a key advantage for British industry. How can this be? Shouldn't high wages, if anything, make British industry less competitive and thereby discourage investment in Britain? Simple economic models indeed have this prediction. Those models, however, ignore the effects of high wages on innovation. Allen's argument is that high wages spurred labor-saving innovation. But if high wages spur labor-saving innovation, doesn't that mean that low wages spur innovation that economizes on other factors of production (e.g., land and capital)? Is there any reason to believe that the overall level of innovation should be higher when wages are high than when wages are low?

These are difficult questions that are not fully addressed in Allen's book. An interesting paper by the economist Daron Acemoglu presents a model of "directed technical change" that sheds some light on these questions (Acemoglu, 2002). Acemoglu's model indeed has the implication that more innovation will be directed to factors of production that are more expensive. High wages therefore induce labor-saving innovation, while high land prices induce land-saving innovation. Whether the overall level of technical progress is higher when wages are high or when land prices are high, however, depends on a number of other aspects of reality. In particular, this depends on the "innovation possibilities frontier." This is a fancy way of saying that the cost and technical difficulty of making different types of innovations matters for the direction and overall level of technical progress. For example, if innovation that economizes on land or capital is for some reason very costly or technically difficult, then conditions that favor economizing on land and capital but discourage labor-saving innovation (i.e., low wages) will result in a low overall level of technical progress.

One way to make sense of Allen's argument is that industrialization had to follow a certain particular path that began with labor-saving innovations. The pre-industrial world featured very few machines. Innovation that economized on machines was therefore not really possible (there were no machines to economize on). The initial phase of industrialization must therefore be labor saving: the invention of machines. Once machines had been invented, capital-saving technical progress became a possibility and indeed a great deal of such innovation occurred (e.g., the massive efficiency improvements of steam engines and textiles machinery that we discuss in more detail later in this section). But to get the whole thing started, machines needed to be invented and this could only happen in a high-wage economy because it was only in a high-wage economy that it was profitable to invent the inital (highly innefficient) machines.

What about land? Land certainly featured prominently in the pre-industrial economy. Furthermore, our discussion of the Agricultural Revolution in section 3 indicates that innovation that economized on land was possible (e.g., intensive crop rotations). An interesting feature of the cross-country data that we reviewed in that section is that a subset of European countries underwent a substantial Agricultural Revolution, while others did not. It so happens that the countries in which agriculture was revolutionized were exactly the countries that were most exposed to the growth in trade and commerce associated with the massive expansion of Atlantic trade in the early modern period. One way to make sense of this cross-country variation in the extent to which different countries experienced an Agricultural Revolution is that it reflects the importance of directed technical change: The countries that saw a rapid expansion of wealth and population associated with Atlantic trade experienced a sharp increase in food demand. This increase in food demand and the associated increase in food prices increased the profitability of technical change in agriculture and therefore spurred on an Agricultural Revolution in these countries.

But if high food demand in England and the Netherlands in the 18th century spurred the Agricultural Revolution, why didn't high demand for food associated with Malthusian population pressure in earlier times spur land-saving innovation? It seems that such innovation did occur to a greater extent in Asia than in Europe. Agricultural yields in certain regions of Asia (e.g., the lower Yangzi Delta) were high enough to support dramatically higher population densities than anywhere in Europe. Furthermore, substantial innovation was possible in Europe as evidenced by the Agricultural Revolution. Why didn't it occur earlier? Something must have held back land-saving innovation in Europe. It is not well understood what this was.

4.2 The Steam Engine, Coal, and the Energy Transformation

Before the Industrial Revolution, the possibility of sustained increases in the living standards of ordinary workers was seriously hampered by ecological constraints: Land was (and is) in virtually fixed supply; production was very land intensive; and this meant that the economy faced sharply diminishing returns to labor. The Malthus model discussed in chapter XX [Malthus chapter] explains this logic in detail. In such an environment, increases in technology lead to a larger population in the long run, but not higher wages.

The invention of the steam engine radically altered this state of affairs by allowing production to become much less land intensive than before. Before the Industrial Revolution, it was not only food production that was land intensive. The production of both heat and mechanical energy was also highly land intensive. Table 8 contains estimates of energy consumption in England and Wales between the 1560s and the 1850s. In the 16th century, a very substantial portion of energy consumption was firewood, mostly for heating. The source of a large majority of mechanical energy was human and animal power, which required food as an input. Expansion of industry with the associated increase in demand for energy would therefore have required a great deal of extra land. The advent of steam engines powered by coal dramatically changed this. Table 8 shows how the Industrial Revolution caused a 30-fold increase in energy consumption in England and Wales and an increase in the share of coal in energy consumption from about 10% in the 1560s to more than 90% in the 1850s. Clearly, this massive increase in energy consumption would not have been possible without the use of coal.

Wrigley (2010) argues that the shift from an organic economy constrained by limited land to an energy-rich economy powered by coal (and later other fossil fuels) was the most fundamental change that occurred over the course of the Industrial

Table 8: Energy Consumption in England and Wales

	1560s	1700s	1750s	1800s	1850s
Farm Animals	21.1	32.8	33.6	34.3	50.1
Population	14.9	27.3	29.7	41.8	67.8
Firewood	21.5	22.5	22.6	18.5	2.2
Wind	0.2	1.4	2.8	12.7	24.4
Water	0.6	1.0	1.3	1.1	1.7
Coal	6.9	84.0	140.8	408.7	1,689.1
Total	65.1	168.9	230.9	517.1	1,835.3
Total less coal	58.2	84.9	90.1	108.4	146.2

Notes: Energy is measured in petajoules. The source of these estimates is Table 2.1 of Wrigley (2010). The numbers for farm animals and the population are estimates of calories consumed. Wind energy includes sail ships. The firewood number for 1850s is actually for the 1840s.

Revolution. He furthermore argues that the key puzzle regarding the Industrial Revolution is not how it got started but rather why it did not end quickly. The answer to this is that the shift to fossil fuels as a source of energy dramatically relaxed the ecological constraint that had previously held back growth in living standards.

From this point of view, the steam engine is by far the most transformative technological innovation of the Industrial Revolution, perhaps the innovation that rid the economy of the crucial bottleneck holding back growth, and that thereby made all the difference. For this reason, the story of how the steam engine was developed is of particular interest. Was there something special about 18th century England that resulted in the steam engine being invented then and there? Or was this just a stroke of good luck that could have occurred elsewhere or at another time?

Thomas Newcomen's first successful steam engine was put into service in Dudley in 1712, where it drained a mine. It was an atmospheric engine in which steam was condensed to form a vacuum under a piston. The piston was then pushed down by the pressure of the atmosphere above it. This design was based on the scientific discovery by Torricelli in the 17th century that the atmosphere had weight and the subsequent discovery of Papin that condensing steam can produce a vacuum. Enlightenment advances in science are therefore one of the crucial precursors to the invention of the steam engine. One important reason the Romans (for example) didn't invent the steam engine was that they didn't have access to knowledge of 17th century science.

But as Allen (2009) describes, Newcomen's engine was not a simple application of earlier scientific ideas. It was a highly complex machine (by the standards of the time) and its invention required a major project of research and development. It took Newcomen over a decade of experimentation and considerable expense to arrive at a successful design. This large investment of time and money was worthwhile in Britain only because Britain had a large coal industry which created a substantial demand for drainage and because the huge amount of fuel needed to operate the highly inefficient early steam engines was virtually free in the coal fields.

In the 17th and 18th centuries, the British coal industry grew enormously—by a factor of roughly 60—to become by far the largest in the world (Allen, 2009, p. 81-82). Allen argues that the cause of this huge growth was the rapid expansion of London during this period that resulted from Britain's success in international trade. The growth of London led to a large increase in the demand for fuel for heating purposes. Firewood was the traditional fuel used to heat houses. The large increase in the demand for firewood caused a 'timber crisis' around London in the 17th century (Nef, 1932; Hatcher, 1993).

Figure 12 presents the evolution of the real price of wood and coal in London from 1400 to 1830—i.e., the price of wood and coal divided by a general consumer price index. The price of wood roughly doubled in real terms over the course of the 17th century and stayed high thereafter. This large increase resulted from the explosive growth in London's population, which rose from 200,000 in 1600, to 575,000 in 1700, and 960,000 in 1800 (Wrigley, 2010).

High wood prices spurred demand for an alternative to firewood for heating houses. The alternative that emerged was coal. But to switch from heating houses with wood to coal was not a trivial change. The coal-burning house had to be invented (Allen, 2009). Burning coal emits sulphurous fumes that render a house uninhabitable unless the fumes are vented out effectively. Furthermore, burning coal requires an enclosed space with a generous draft to maintain both high heat and substantial oxygen flow to the fire. The fire in a medieval house was typically on an open hearth in the middle of a large room. For coal to be used for residential heating, narrow (and tapered) chimneys with an enclosed fireplace and grates under the fuel were essential. Arriving at such a design involved a great deal of experimentation. The explosive growth of London over this period with the associated building boom aided the needed experimentation.

The combination of high prices for firewood and low prices for coal were an essential driver of the shift to coal-burning houses in London and therefore of the

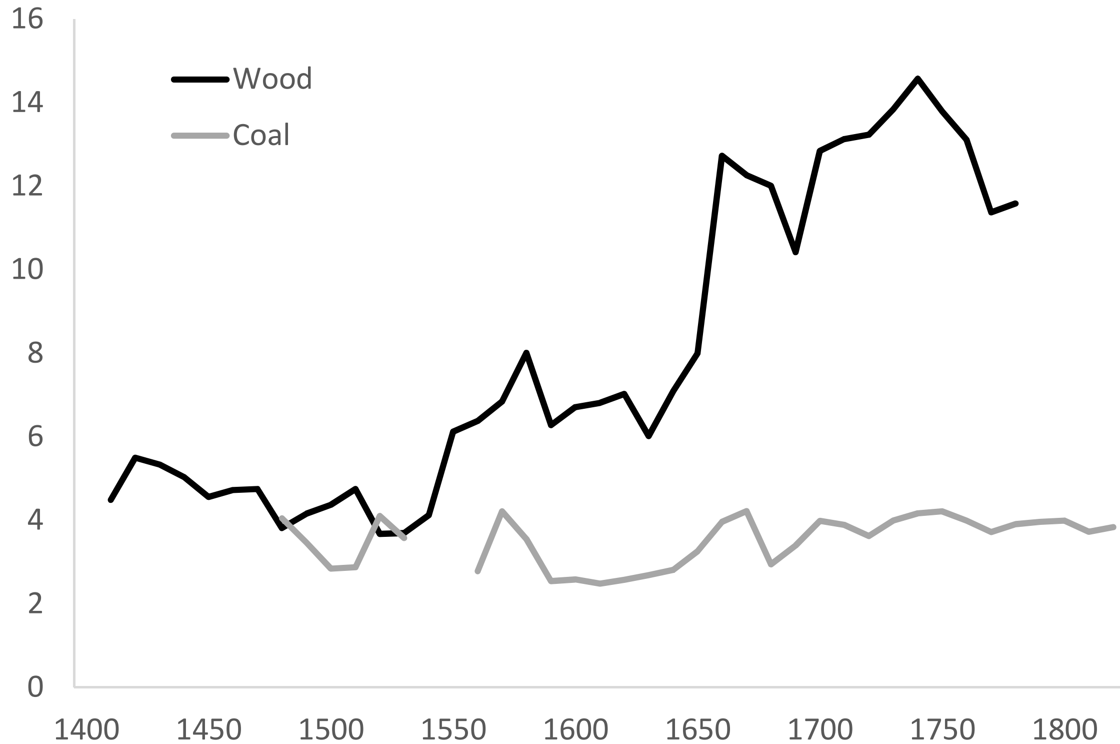

Figure 12: Real Prices of Wood and Coal in London

Note: This figure is based on Figure 4.3 in Allen (2009). Prices are decadal averages. The prices of wood and coal are deflated by the consumer price index described in Allen (2001b). Allen reports these prices in terms of the price level in Strasbourg in 1754-54. The prices in the figure should therefore be interpreted as the price of wood and coal in grams of silver per million BTU if the overall price level was that of Strasbourg in the mid 18th century.

rise of the coal industry in England. The rise of the coal industry provided large demand for drainage, which helped induce the invention of the steam engine. Another byproduct of the rise of the coal industry was that the price of energy in Northern and Western England was by far the lowest in the world. Figure 13 presents the price of energy in several European cities in the early 1700s. The price of energy in London is based on coal and was roughly 5 grams of silver per million BTUs. This price mostly reflects the cost of transport of coal from Newcastle to London. The price of energy in Newcastle was below one gram of silver per million BTUs, much lower than elsewhere in Europe. This low price of energy in Northern and Western Britain helped spur not only the invention of the steam engine, but also other energy intensive innovations including coal based metallurgy.

These developments—the rise of the coal industry and the associated low price of energy—are important factors that set Britain apart even from other Western Eu-

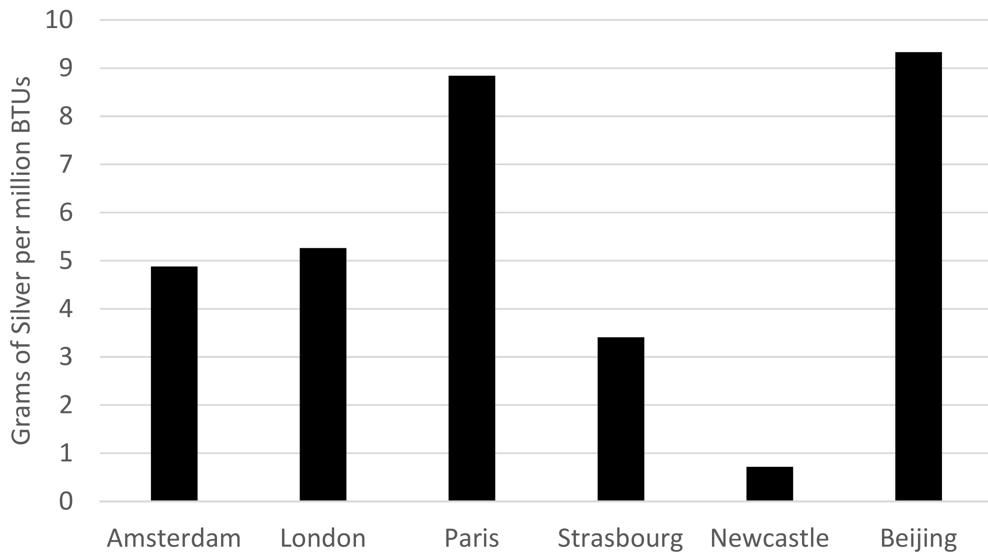

Figure 13: Price of Energy in the Early 1700s

Note: This figure replicates Figure 4.1 in Allen (2009).

ropean regions that shared many other advantages with Britain. Consider, in particular, the Dutch Republic: it was similarly exposed to the Enlightenment and to the Reformation, it benefited from the growth in international trade, and it saw liberal political reform. Together, these developments led to explosive growth of Amsterdam and high wages as in Britain (see Figure 10). So, why was there no Industrial Revolution in the Netherlands? This is a tough question and no fully satisfactory answer has yet been developed. But one important difference is that Amsterdam had a different backstop energy supply: peat. When Amsterdam grew, its population increasingly switched to using peat for heating. This kept energy prices in Amsterdam relatively low. But the increased use of peat did not have any of the auxiliary longer-term benefits for technological innovation that coal mining had in Britain.

Allen's argument is that high wages and cheap coal created strong incentives for technical change aimed at substituting energy and capital for labor in Britain during the Industrial Revolution. Critics of this view have, however, pointed out that a great many inventions in Britain over this period actually aimed to economize on energy and capital. Allen counters this point by drawing a distinction between what economic historian Joel Mokyr has called macro-inventions and micro-inventions (Mokyr, 1990). Macro-inventions are inventions of something entirely new—such as Newcomen's steam engine—that then set in motion a long train of subsequent improvements. Micro-inventions, on the other hand, to refer to the many incremental improvements on existing machines or techniques that aim to reduce cost and improve efficiency. Allen (2009) argues that the critical macro-inventions that got the Industrial Revolution going—e.g., the steam engine, the spinning jenny, and the coke smelting furnace—did indeed radically change the relative importance of different factors of production by substituting energy and capital for labor. High wages and cheap coal gave the British much stronger incentives to invent these game-changing technologies than people in other regions.

Once the new technologies had been invented, however, a process of improving them immediately began. The nature of the technical change involved in this process of improvement was quite different from that of the original macro-inventions: it involved improvements that saved all inputs to production, not the least fuel. The history of the steam engine illustrates this well. The initial Newcomen steam engine was extremely inefficient. It used an enormous amount of fuel for each horsepower-hour of mechanical energy it produced. Figure 14 plots the evolution of the coal consumption of pumping engines in the 18th and 19th century. The original Newcomen engine needed 45 pounds of coal to produce a single horsepower-hour of mechanical energy. This staggering level of inefficiency meant that the only task for which it could profitably be used was to drain coal mines (where coal was essentially free). However, as the steam engine was made more energy efficient, the scope of its use dramatically increased: first to draining other types of mines and then more widely to powering machinery. The invention of the original steam engine therefore shifted the incentives of inventors. Now they faced strong incentives to invent technologies that improved energy efficiency.

The scale of the efficiency improvements of steam engines in the 18th and 19th centuries is staggering. Coal consumption fell from 45 pounds per horsepower-hour in the early 18th century to less than one pound per horsepower-hour in the late 19th century (Allen, 2009, p. 164-165). Little is known about the early improvements in the efficiency. The drop from 30 to 17.6 pounds per horsepower-hour in 1769 is due to various improvements made by the engineer John Smeaton. The drop from 17.6 to below 10 pounds per horsepower-hour was due to James Watt's separate condenser. The drop from just below 10 pounds in the 1790s to below 2 pounds per horsepower-hour in the 1830s was due to several inventions. The most important of these was the high-pressure engine of Richard Trevithick. Other important inventions include the use of steam expansively and Arthur Woolf's invention of the drop valve and compounding.

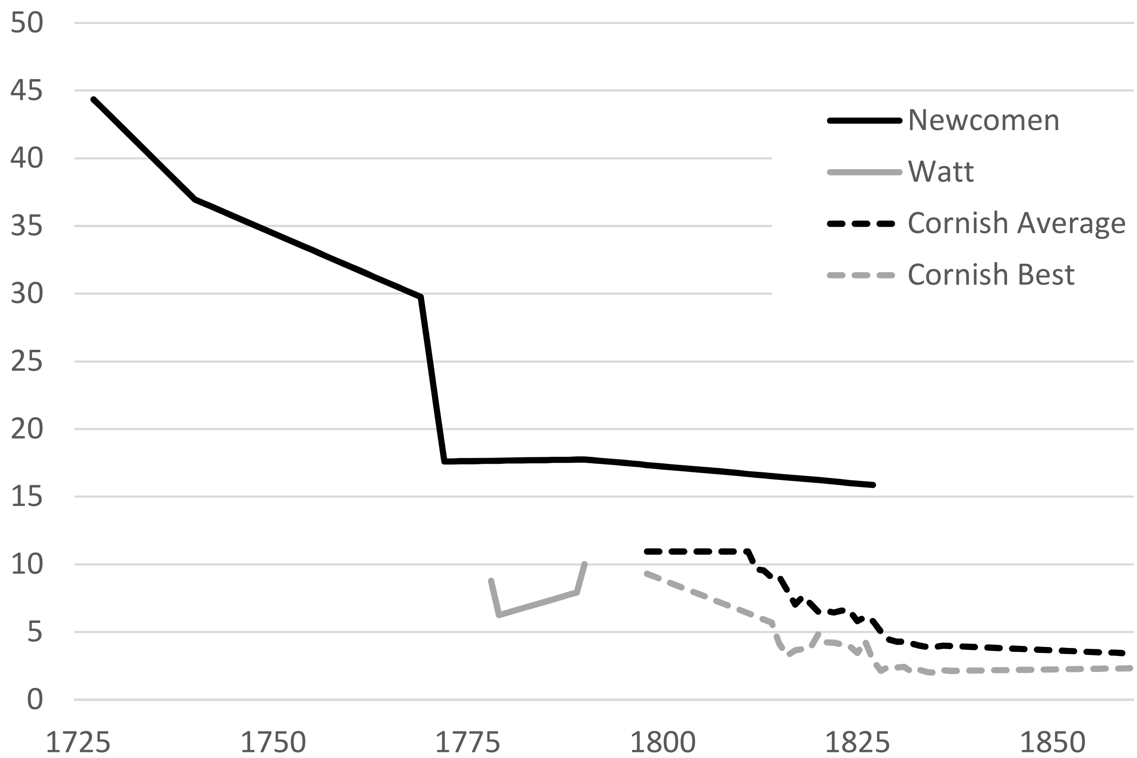

Figure 14: Coal Consumption of Pumping Engines

Note: This figure replicates Figure 7.1 in Allen (2009). Coal consumption is measured in pounds per horsepower-hour.

The original Newcomen engine was not well suited for powering machinery because its rocking motion was too irregular. The first solution to this problem was to combine the Newcomen engine with water power, i.e., use the Newcomen engine to return water to an upstream reservoir so that it could pass over a water wheel repeatedly. This technology was widely used in the second half of the 18th century. In the 1780s, however, James Watt took up the challenge of creating a rotary steam engine that produced smooth motion. One of his innovations was 'double-action': injecting steam alternatively into each end of the cylinder. This along with a reorganized system of rods allowed for 'parallel motion': the piston to both push and pull the beam. Watt's third important innovation was 'sun and planet' gears that were used to rotate the drive shaft and double the rotation speed. Finally, there was the 'centrifugal governor', which was used to stabilize the speed of the engine. The result was a rotary steam engine that could power machinery.

However, Watt's rotary steam engine was not very fuel efficient. It used 12-15 pounds of coal per horsepower-hour (Allen, 2009, p. 172). This is one reason for the

Table 9: Stationary Power Sources in Great Britain

	1760	1800	1830	1870	1907
Steam	5,000	35,000	160,000	2,060,000	9,659,000
Water	70,000	120,000	160,000	230,000	178,000
Wind	10,000	15,000	20,000	10,000	5,000
Total	85,000	170,000	340,000	2,300,000	9,842,000

Notes: This table replicates Table 7.1 in Allen (2009).

slow adoption of steam power in industry before 1830. Table 9 presents estimates of the evolution of installed power in Britain by power source. Even in 1830, less than half of installed power is steam power. Another reason for the slow adoptions of steam is the impressive improvements in the efficiency of water power in the 18th and early 19th centuries (Mokyr, 1990, p. 90-92). The major breakthroughs include the breast wheel (John Smeaton, 1750s), the sliding hatch (John Rennie, 1780s), and the water turbine (Benoit Fourneyron, 1837). In the 1840s, however, William McNaught added high pressure and compounding to Watt's rotary engine design. This cut the fuel consumption of rotary engines to 5 pounds per horsepower-hour. After this, the use of steam power expanded rapidly and "the general mechanization of British industry began" as Allen (2009) puts it.

4.3 Cotton

The cotton industry was the marvel of the early Industrial Revolution. Its growth both in terms of productivity and output was breathtaking. Between 1770 and 1815, the British cotton industry grew by 7% per year. This yielded total growth of roughly 2200% over this 45 year period (Harley, 1982; Crafts and Harley, 1992). Between 1780 and 1860, total factor productivity (TFP) in the British cotton industry grew by 2.6% per year and accounted for a quarter of all TFP growth in the British economy (Crafts, 1985). This rapid growth in productivity led to a large fall in the relative price of cotton cloth. Figure 15 plots the price of cotton (or linen) cloth relative to the price of bread in England from 1690 to 1900. From 1750 to 1820, the relative price of cotton cloth fell by a factor of ten.

The story of the growth of the cotton industry is one of mechanization of an extremely labor-intensive activity. Table 10 reports estimates constructed by Robert Allen of the breakdown of the cost of producing a pound of relatively coarse cotton yarn using different technologies. The first column provides this breakdown

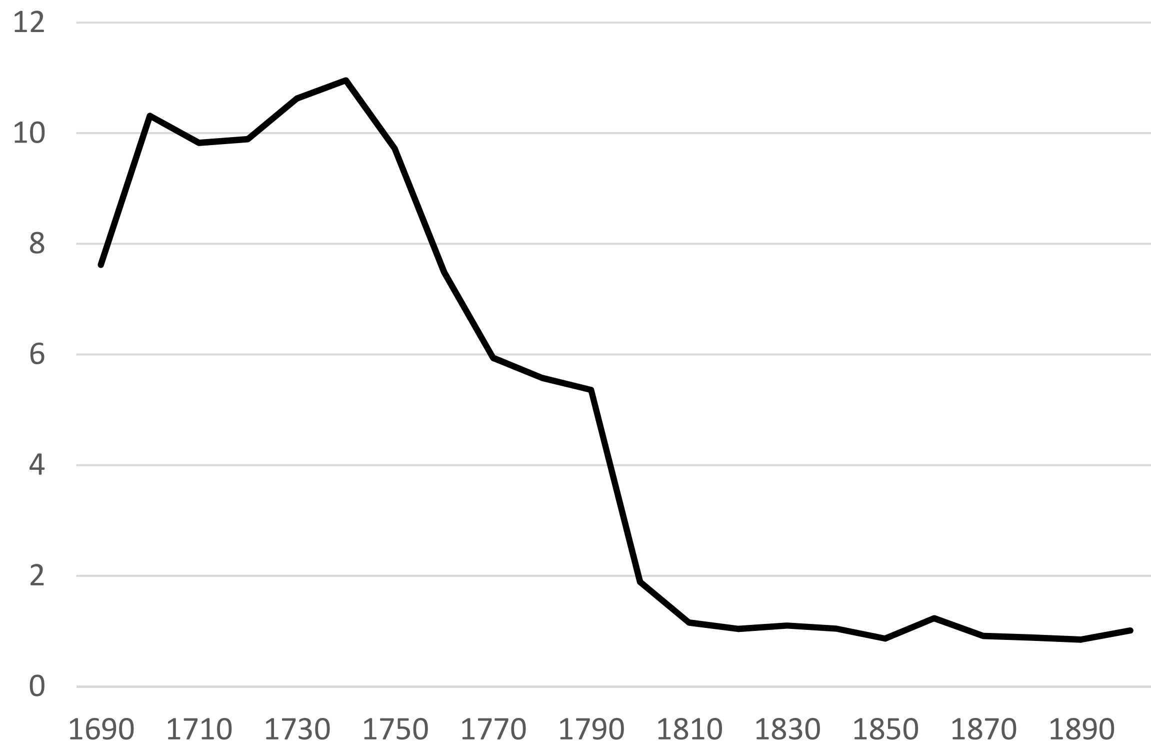

Figure 15: Prices of Cotton (or Linen) Cloth Relative to Bread in England

Note: This figure is based on Figure 4 in Allen (2009). Prices are decadal averages.

for traditional hand methods used prior to the onset of mechanization. Using these traditional methods, Allen estimates that a cotton spinner could spin roughly one pound per day and was payed 7 pence for this. Preparing the cotton for spinning by cleaning it and carding it into rovings took a comparable amount of time. In addition to this, the "putting out" system of production—where most spinning occurred in people's homes as a part-time activity alongside farming and other activities—involved a substantial amount of administration. Allen estimates the total labor costs of producing a pound of cotton yarn using these traditional methods to be 17.19 pence, a bit more than the cost of the raw cotton. In sharp contrast, these traditional methods involved a minimal amount of capital expenses: 0.93 pence per pound according to Allen's estimate. These estimates therefore imply that the labor share of value added in the production of cotton yarn before mechanization was 95% (17.19/(17.19+0.93) = 0.949).

The most famous innovations in the cotton industry were to spinning: James Hargreaves' spinning jenny (1764), Robert Arkwright's water frame (late 1760s), and Samuel Crompton's mule (1779). But many other aspects of textile produc-

Table 10: Real Cost of Cotton Yarn (16 count)

	Hand Method 1760	24-Spindle Jenny 1775	Arkwright Mill 1784	Glasgow Mill 1836
Labor:				
Cleaning and Carding	7.00	7.00	2.69	0.16
Spinning	7.00	2.33	2.57	0.34
Reeling, bundling, etc.	0.47	0.47	2.19	0
Administrative	2.72	2.72	0.41	0.02
Total Labor	17.19	12.52	7.86	0.52
Materials:				
Raw Cotton	16.88	16.88	16.88	16.70
Other	0	0	1.20	0.53
Capital	0.93	1.88	2.00	0.47
Total Cost	35.00	31.28	27.94	18.22
Labor Share of Value Added	95%	87%	80%	53%

Notes: Costs are reported pence per pound at 1784 prices. The data are from Table 8.1 of Allen (2009). The "count" of cotton yarn measures the number of hanks (770m) of yarn that make one pound of yarn. The higher is the count, the finer the yarn. 16 count yarn is relatively coarse, similar to the yarn used for modern jeans.

tion were also thoroughly transformed by mechanization. The carding machine patented by Robert Arkwright—an improvement on an earlier machine invented by Lewis Paul—improved labor productivity in the production of cotton yarn by a similar amount as Arkwright's more famous water frame. Other important innovations include the cotton gin (Eli Whitney) for separating cotton fibers from seeds, bleaching (Claude Berhollet), metal printing (Thomas Bell), and the power loom in weaving (Edmund Cartwright). Mokyr (1990, p. 96-103) and Allen (2009, ch. 8) provide a more detailed discussion of innovation in textiles during this period.

Most if not all of these innovations represented "biased" technical change in the sense that they disproportionately saved labor. This can be clearly seen for the innovations involved in spinning in Table 10. Comparing the first and second column shows the impact of the spinning jenny. According to Allen's estimates, it lowered the cost of spinning by 2/3, from 7 pence per pound to 2.33 pence per pound. The spinning jenny was a much more expensive machine than the traditional spinning wheel. The cost of capital therefore rose, but only from 0.93 to 1.88 pence per pound.

Clearly, the spinning jenny substituted capital for labor in the production of cotton yarn. The overall cost saving was 3.72 pence per pound and the labor share of value added fell from 95% to 87%.

Arkwright's water frame reduced the cost of spinning by roughly the same amount as the spinning jenny (column 3 of Table 10). Arkwright's other innovations, however, reduced the cost of cotton yarn a substantial amount further. In addition to inventing the water frame and improving the carding machine, Arkwright also invented the cotton mill, i.e., the factory. Factory production allowed for more efficient use of power sources such as water wheels and later steam engines. It eliminated the administrative labor associated with the putting-out system of domestic production. It also allowed for more discipline in the workplace (Clark, 1994). Workers that worked in a domestic setting or in workshops without discipline tended to work irregular hours and do a substantial amount of socializing at work. Arkwright imposed strict discipline in his cotton mills: workers were required to show up on time, work certain hours, work at a constant rate, and refrain from socializing. This raised their productivity and allowed for more efficient use of machinery. However, workers disliked discipline and demanded a wage premium to work under such conditions. Factories with discipline and workshops without discipline coexisted throughout the 19th century. Eventually, however, discipline came to dominate the workplace (although often in a less extreme form than in the early cotton mills).

The third column in Table 10 presents a breakdown of the cost of producing a pound of cotton yarn in one of Richard Arkwright's cotton mills in the mid 1780s. The cost of spinning (using water frames) is roughly 1/3 of its value with traditional methods. The cost of cleaning and carding (using Arkwright's carding machine) is also roughly 1/3 of its value prior to mechanization. There are also large savings of administrative labor from moving production to a factory setting. Total labor costs have fallen to only 7.86 pence per pound. Capital costs are 2 pence per pound, only slightly higher than in column 2, but there are some extra materials costs. Overall, Arkwright's mill reduced the cost of producing yarn by roughly 7 pence per pound relative to traditional methods. This equals 20% of the traditional cost of producing yarn, but roughly 40% of the traditional cost of yarn less the cost of raw cotton. Arkwright's innovations were highly biased towards saving labor. The labor share in value added in his mill was 80%, much lower than the 95% labor share using traditional methods.

The early inventions of the spinning jenny, water frame, and carding machine set in motion a torrent of further innovation that drove costs lower and lower over time. The most important of these inventions was Samuel Crompton's mule, which combined the best features of the spinning jenny and the water frame. In addition to improving efficiency, the mule allowed much finer yarn to be spun using a mechanized process and thus reduced the cost of high-count yarn by enormous amounts. The final column of Table 10 presents a breakdown of the cost of producing yarn in a cotton mill in Glasgow in 1836. This mill used the self-acting mule—which had been invented by Richard Roberts in the 1820s—and was vastly more efficient than Arkwright's mill from 1784 along all dimensions. The total cost of a pound of cotton yarn was down to 18.22 pence per pound, with 16.70 of that being the cost of raw cotton. The improvements between the mid 1780s and the mid 1830s had been strongly labor saving just like the early inventions. This is evident from the fact that the labor share of value added had fallen from 80% to only 53%.

The vast majority of the innovation in spinning cotton occurred in Britain. Furthermore, even though the technology became widely known in other countries quite quickly and adoption was in some cases even subsidized by foreign governments, widespread adoption in other countries lagged behind Britain by roughly 40 years. Why did all this innovation happen in Britain and not (say) in France? Allen argues that the crucial reason is that wages were much higher in Britain. Higher wages in Britain made the operation of these more capital intensive (labor saving) technologies more profitable in Britain than in France. In short, he argues that inventing and using the spinning jenny, water frame, and mule was profitable in Britain but was not profitable in France, since wages were much lower in France. Allen (2009, p. 203) argues that the annual profit rate of an Arkwright-type mill in Britain in the 1780s was 40%, while in France it was only 9% (lower than the 15% rate of return fixed capital in business at the time).

Britain had an additional important advantage over France: a large watch-making industry. The water frame and mule required precision parts (gears) to operate. Arkwright employed a large number of watch-makers to build his machines. This may be one reason why the cotton industry arose in Lancashire: most of the world's watch movements were made in Lancashire at the time. France did not have a comparable watch-making industry. It is therefore likely that the cost of building the needed machines was substantially higher in France than in Lancashire. Landes (2000) argues that the substantial size of the watch-making industry in Britain was due to high wages in Britain that led to high demand for clocks and watches.

Allen's high wage hypothesis is controversial. Kelly, Mokyr, and O Grada (2014) argue that British wages were higher in the 18th century because British workers were more productive. If true, this implies that unit labor costs were no higher in Britain than in France, undermining Allen's thesis of high labor costs in Britain. Kelly, Mokyr, and O Grada present evidence on relative productivity in reaping and threshing—two physically demanding, low skill, agricultural activities. These estimates indicate that British workers were about 70% more productive than French workers, enough to cancel out a large part of the difference in wages. Whether this difference is due to differences in the inherent productivity of the workers (as Kelly, Mokyr, and O Grada argue) or the incentives the workers faced is not clear. Our discussion of agriculture in section 3 suggests that incentives might have mattered. The main channel Kelly, Mokyr, and O Grada suggest for higher British productivity of unskilled workers is better nutrition. But evidence from modern developing countries does not support the notion that higher wages increase the productivity of poor workers enough to prevent unit labor costs from rising (Banerjee and Duflo, 2011). Furthermore, it is not clear that these productivity differences existed in domestic cotton spinning—a sedentary occupation performed largely by women in their homes.

Gragnolati, Moschella, and Pugliese (2011) argue that Allen understates the profitability of the spinning jenny in France by assuming that output per worker remains constant and hours fall when the spinning jenny is adopted as opposed to assuming that hours stay constant and output rises (or something in between). Their calculations suggest that under the assumption that hours stay constant, the spinning jenny was also profitable in France. This critique does not, however, dispute the notion that the spinning jenny was *more* profitable in Britain. The overall level of profitability is hard to pin down. There are no doubt many factors left out of the relatively simple calculations that Gragnolati, Moschella, and Pugliese do (and also those that Allen does). The fact that the spinning jenny was (substantially) more profitable in Britain therefore continues to provide a possible reason why it was invented and adopted there and not in France.

Stephenson (2018) argues that the English wage data that Allen (and other economic historians) have used for the the 18th century overstate actual wages. The issue is that the underlying data are payments to contractors as opposed to wages earned by worker. Stephenson (2018) argues that the actual wages are lower by 20 to 30 percent due to markups by the contractors. Allen (2019) disputes this. But even if it is true, it does not necessarily imply a smaller wage premium in England relative to (say) France. This depends on whether French data also suffers from the same problem, which is unfortunately not clear. Suppose, however, that Stephenson is correct and the English wage premium is say 20% smaller than Allen's data suggest. Even in this case, Stephenson's adjustment only closes a relatively modest portion of the wage gap between London and Paris. Allen's estimates of respectability ratios in different European cities (Figure 10) indicate that the difference in real wages between London and Paris in 1750 was 120%.[8]

5 Concluding Thoughts

The Industrial Revolution and its antecedents are a complex topic about which much is still poorly understood. Scholars disagree strongly about many important issues and scholarship is rapidly evolving. It is therefore with some trepidation that I venture to provide a summary of my own views about this topic based on my study of the literature over the past few years.[9] To the extent that there is a watershed event that made the Industrial Revolution possible, I think it is the invention of movable type printing by Johannes Gutenberg around 1450. This event radically altered the balance that had existed between forces seeking to create and spread knowledge on the one hand and forces that destroyed knowledge and suppressed new ideas on the other hand. Movable type printing gave the forces seeking to create and spread knowledge a decisive upper hand and set in motion an intellectual transformation: first the Reformation, then the Enlightenment and Scientific Revolution, and ultimately the Industrial Revolution.

Changes to institutions in the leadup to the Industrial Revolution were surely very important. But my views on this differ from North and Weingast (1989) in two significant ways. First, I think North and Weingast overemphasize the Glorious Revolution. I am sympathetic to Allen's (1992) narrative regarding the importance of gradual expansion of liberty and property rights in England over the 500 year period leading up to the Glorious Revolution. Second, it is not clear to me that the primary growth enhancing characteristic of the post Glorious Revolution regime was more secure property rights. I would rather point to Parliament's policy of facilitating the reorganization and rationalization of property rights through enclosures, statutory authorities, and estate acts (Bogart and Richardson, 2011). The feudal system of property rights that existed before was likely woefully inefficient. The post Glorious Revolution Parliament can be viewed as a new technology to facilitate reorganization (lower transaction costs in the language of the Coase Theorem).

Finally, I am sympathetic to the view that directed technical change was important. First, Robert Allen's (2009) ideas about high wages and cheap coal in Britain creating unusually strong incentives for the invention of labor saving devices are in my view the most plausible existing theory for why the Industrial Revolution was so extremely geographically concentrated in Britain. Second, the fact that Agricultural Revolutions occurred exactly in the places that had the highest wages (Britain and the Netherlands) exactly when the demand for food in these places rose, suggests to me that the Agricultural Revolution was another example of directed technical change.

Notes

[1]The labor demand curve is a relationship between wages and labor supplied, not wages and the population. Changes in labor supplied will differ from changes in the population if days worked per year vary over time or if the fraction of the population that works varies. Humphries and Weisdorf (2019) estimate how days worked per year evolved over time in England from 1260 to 1850 by combining data on earning of workers on annual contracts and data on wages of day laborers. They document substantial changes in days worked per year over the period 1300 to 1640: days worked per year dropped substantially in the 14th century (especially after than Black Death) and then rose gradually after that. Days worked were quite similar in 1300 and 1640. The conclusion I reach in the text based on wage and population data is therefore robust to allowing for variation in days worked per year over this period.

[2]See Hill (1940) and Hill (1961) for a more detailed account of the revolutions of 17th century England and a Marxist perspective on their economic causes.

[3]An empirical challenge Jha must overcome is that ownership in an overseas joint stock company may be correlated with other characteristics that cause support for the rebellion. Jha (2015) shows that owners of shares are very similar to non-owners along a range of characteristics. He also exploits variation in propensity to own shares resulting from coincidence of the timing of an initial public offering of overseas shares and the time when an individual turned 21 and was able to sign a legally binding contract to purchase such shares.

[4]The calculations in this paragraph make use of estimates of nominal English GDP from Broadberry et al. (2015) and estimates of the popluation of England from Clark (2010). For compatibility with the profit numbers reported by Acemoglu, Johnson, and Robinson (2005), the GDP estimates are adjusted to 1600 prices using the index of building craftsmen's nominal wages from Clark (2005). English GDP in 1630s was £35 million. English population in the 1630s was 5.2 million. This means that GDP per person was roughly £7.

[5]Gutenberg did not invent the printing press in a single year. He finished the first copies of his famous Bible—his first major work—in approximately 1454. But sources vary even as to the exact completion date for the Gutenberg Bible. By the time Gutenberg finished printing the Bible, he had been working on his technology for many years. He had several rounds of "venture capital funding" for at least 15 years prior to publishing the Bible. In numerous cases, his investors sued him for failing to deliver a product to market or related infractions. Much of what is known about Gutenberg's life is from these lawsuits. Boorstin (1985, ch. 63) provides an excellent discussion of the technical challenges Gutenberg overcame and the long and arduous process involved.

[6]Cressy (1980) reports the fraction of witnesses who sign ecclesiastical records. He reports this data broken down by several different occupations and locations. Houston (1982) reports the fraction of witnesses who sign court depositions by gender but not by occupation or location. I pool the data from Cressy and Houston and run a regression with time fixed effects, location fixed effects, and occupation fixed effects. I use a separate location fixed effect for the data from Houston and treat the male and female data from Houston as two separate "occupations." Cressy also reports literacy rates for women, which I also treat as a separate "occupation." Figure 7 reports the time fixed effects from this regression when the occupation "husbandmen" and the location "London/Middlesex" are the omitted categories.

[7]Allen has since made some adjustments to the way he calculates these baskets. Fore a discussion, see Allen (2015).

[8]Other critiques include Humphries (2013) and Humphries and Schneider (2019). Allen's responses can be found in Allen (2015) and Allen (2020). Both responses are in my view convincing.

[9]This summary is written in the spring of 2020.

References

ACEMOGLU, D. (2002): "Directed Technical Change," *Review of Economic Studies*, 69, 781–809.

ACEMOGLU, D., S. JOHNSON, AND J. ROBINSON (2005): "The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth," *American Economic Review*, 95, 546–579.

ACEMOGLU, D. AND J. A. ROBINSON (2019): *The Narrow Corridor*, New York, NY: Penguin Press.

ACEMOGLU, D. AND A. WOLITZKY (2011): "The Economics of Labor Coercion," *Econometrica*, 79, 555–600.

ALLEN, R. C. (1992): *Enclosure and the Yeoman*, Oxford, UK: Clarendon Press.

——— (2000): "Economic Structure and Agricultural Productivity in Europe, 1300-1800," *European Review of Economic History*, 3, 1–25.

——— (2001a): "Community and Market in England: Open Fields and Enclosure Revisited," in *Communities and Markets in Economic Development*, ed. by M. Aoki and Y. Hayami, Oxford, UK: Oxford University Press, 42–69.

——— (2001b): "The Great Divergence in European Wages and Prices from the Middle Ages to the First World War," *Explorations in Economic History*, 38, 411–447.

——— (2008): "The Nitrogen Hypothesis and the English Agricultural Revolution: A Biological Analysis," *Economic Journal*, 68, 182–210.

——— (2009): *The British Industrial Revolution in Global Perspective*, Cambridge, England: Cambridge University Press.

——— (2015): "The High Wage Economy and the Industrial Revolution: A Restatement," *Economic History Review*, 68, 1–22.

——— (2019): "Real Wages Once More: A Response to Judy Stephenson," *Economic History Review*, 72, 738–754.

——— (2020): "Spinning Their Wheels: A Reply to Jane Humphries and Benjamin Scheider," *Economic History Review*, forthcoming.

ASHTON, T. S. (1948/1997): *The Industrial Revolution*, Oxford, UK: Oxford University Press.

ASTON, T. AND C. PHILPIN (1987): *The Brenner Debate: Agrarian Class Structure and Economic Development in Pre-Industrial Europe*, Cambridge, UK: Cambridge University Press.

BACON, F. (1620/1999): "The New Organon: Or, True Directions Concerning the Interpretation of Nature," in *Selected Philosophical Works*, ed. by R.-M. Sargent, Indianapolis, IN: Hackett.

BAIROCH, P., J. BATOU, AND P. CHÈVRE (1988): *La population des villes europeenees de 800 a 1850: Banque de donnees et analyse sommaire des resultats. (The Population of European Cities, 800-1850: Data Bank and Short Summary of Results.)*, Centre of International Economic History Series, No. 2, Geneva: Li: Librarie Droz.

BANERJEE, A. V. AND E. DUFLO (2011): *Poor Economics*, New York, NY: Public Affairs.

BECKER, S. O. AND L. WOESSMANN (2009): "Was Weber Wrong? A Human Capital Theory of Protestant Economic History," *Quarterly Journal of Economics*, 124, 531–596.

BOGART, D. (2014): "The Transport Revolution in Industrializing Britain: A Survey," in *Cambridge Economic History of Britain, 1700-1870*, ed. by R. Floud and J. Humphries, Cambridge, UK: Cambridge University Press.

BOGART, D. AND G. RICHARDSON (2011): "Property Rights and Parliament in Industrializing Britain," *Journal of Law and Economics*, 54, 241–274.

BOORSTIN, D. (1985): *The Discoverers*, New York, NY: Random House, Vintage Books edition.

BRENNER, R. (1976): "Agrarian Class Structure and Economic Development in Pre-Industrial Europe," *Past & Present*, 70, 30–75.

BROADBERRY, S., B. M. CAMPBELL, A. KLEIN, M. OVERTON, AND B. VAN LEEUWEN (2015): *British Economic Growth 1270-1870*, Cambridge, UK: Cambridge University Press.

CARD, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics, Voluma 3A*, ed. by O. Ashenfelter and D. Card, Amsterdam, Holland: Elsevier, 1801–1861.

CLARK, G. (1994): "Factory Discipline," *Journal of Economic History*, 54, 128–163.

——— (1996): "The Political Foundations of Modern Economic Growth: England, 1540-1800," *Journal of Interdisciplinary History*, 26, 563–588.

——— (2002): "Land Rental Values and the Agrarian Economy: England and Wales, 1500-1914," *European Review of Economic History*, 6, 281–308.

——— (2005): "The Condition of the Working Class in England, 1209-2004," *Journal of Political Economy*, 113, 1307–1340.

——— (2007): *A Farewell to Alms*, Princeton, NJ: Princeton University Press.

——— (2010): "The Macroeconomic Aggregates for England 1209-1869," *Research in Economic History*, 27, 51–140.

CLARK, G. AND P. A. LEVIN (2011): "How Different Was the Industrial Revolution? The Revolution in Printing, 1350-1835," Unpublished Manuscript, University of California, Davis.

CRAFTS, N. F. (1985): *British Economic Growth During the Industrial Revolution*, Oxford, UK: Clarendon Press.

CRAFTS, N. F. AND C. K. HARLEY (1992): "Output Growth and the British Industrial Revolution: A Restatement of the Crafts-Harley View," *Economic History Review*, 45, 703–730.

CRESSY, D. (1980): *Literacy and the Social Order: Reading and Writing in Tudor and Stuart England*, Cambridge, UK: Cambridge University Press.

DE VRIES, J. (1976): *The Economy of Europe in the Age of Crisis, 1600-1750*, Cambridge, UK: Cambridge University Press.

DICKSON, P. (1967): *The Financial Revolution in England: A Study in the Development of Public Credit, 1688-1756*, New York, NY: St. Martin's Press.

DITTMAR, J. (2011): "Information Technology and Economic Change: The Impact of the Printing Press," *Quarterly Journal of Economics*, 126, 1133–1172.

ERNLE, L. (1912/1961): *English Farming: Past and Present*, London, UK: Heinemann Educational Books and Frank Cass and Company.

FUKUYAMA, F. (2011): *The Origins of Political Order*, New York, NY: Farrar, Straus, and Giroux.

GIBBONS, R. (1992): *Game Theory for Applied Economists*, Princeton, NJ: Princeton University Press.

GOODY, J. (1983): *The Development of the Family and Marriage in Europe*, New York, NY: Cambridge University Press.

GRAGNOLATI, U., D. MOSCHELLA, AND E. PUGLIESE (2011): "The Spinning Jenny and the Industrial Revolution: A Reappraisal," *Journal of Economic History*, 71, 455–460.

HAMMOND, B. (1957): *Banks and Politics in America: From the Revolution to the Civil War*, Princeton, NJ: Princeton University Press.

HARLEY, C. K. (1982): "British Industrialization Before 1841: Evidence of Slower Growth During the Industrial Revolution," *Journal of Economic History*, 42, 267–289.

HATCHER, J. (1993): *The History of the British Coal Industry, Volume I, Before 1700: Towards the Age of Coal*, Oxford, UK: Oxford University Press.

HILL, C. (1940): *The English Revolution 1640*, London, UK: Lawrence and Wishart Ltd.

——— (1961): *The Century of Revolution*, New York, NY: Van Nostrand Reinhold (International) Co. Ltd.

HOUSTON, R. (1982): "The Development of Literacy: Northern England, 1640-1750," *Economic History Review*, 35, 199–216.

HUMPHRIES, J. (2013): "The Lure of Aggregates and the Pitfalls of the Patriarchal Perspective: Critique of the High Wage Economy Interpretation of the British Industrial Revolution," *Economic History Review*, 66, 693–714.

HUMPHRIES, J. AND B. SCHNEIDER (2019): "Spinning the Industrial Revolution," *Economic History Review*, 72, 126–155.

HUMPHRIES, J. AND J. WEISDORF (2019): "Unreal Wages? Real Income and Economic Growth in England, 1260-1850," *Economic Journal*, forthcoming.

IZDEBSKI, A., P. GUZOWSKI, R. PONIAT, AND ET AL. (2022): "Palaeoecological Data Indicates Land-Use Changes Across Europe Linked to Spatial Heterogeneity in Mortality During the Black Death Pandemic," *Nature Ecology & Evolution*, https://doi.org/10.1038/s41559-021-01652-4.

JHA, S. (2015): "Financial Asset Holdings and Political Attitudes: Evidence from Revolutionary England," *Quarterly Journal of Economics*, 130, 1485–1545.

KANT, I. (1784/1986): "An Answer to the Question: What is Enlightenment?" in *Philosphical Writings*, New York, NY: Continuum, 263–269, Lewis White Beck, translation.

KELLY, M., J. MOKYR, AND C. O GRADA (2014): "Precocious Albion: A New Interpretation of the British Industrial Revolution," *Annual Review of Economics*, 6, 363–389.

KINDLEBERGER, C. P. (1993): *A Financial History of Western Europe, Second Edition*, Oxford, UK: Oxford University Press.

LANDES, D. S. (2000): *Revolution in Time: Clocks and the Making of the Modern World*, London, UK: Viking.

MARSHALL, P. (2009): *The Reformation: A Very Short Introduction*, Oxford, UK: Oxford University Press.

MCCLOSKEY, D. (2006): *The Bourgeois Virtues: Ethics for an Age of Commerce*, Chicago, IL: University of Chicago Press.

——— (2010): *Bourgeois Dignity: Why Economics Can't Explain the Modern World*, Chicago, IL: University of Chicago Press.

——— (2016): *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*, Chicago, IL: University of Chicago Press.

MCCLOSKEY, D. M. (1972): "The Enclosure of Open Fields: Preface to a Study of the Impacts of the Efficiency of English Agriculture in the Eighteenth Century," *Journal of Economic History*, 32, 15–35.

MCEVEDY, C. AND R. JONES (1978): *Atlas of World Population History*, New York, NY: Facts on File, Inc.

MOKYR, J. (1990): *The Lever of Riches*, New York: Oxford University Press.

——— (2009): *The Enlightened Economy: An Economic History of Britain 1700-1850*, New Haven: Yale University Press.

NEF, J. U. (1932): *The Rise of the British Coal Industry*, London, UK: George Routledge & Sons.

NORTH, D. C. AND B. R. WEINGAST (1989): "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth Century England," *Journal of Economic History*, 49, 803–832.

NURKSE, R. (1953): *Problems of Capital Formation in Underdeveloped Countries*, Oxford, UK: Blackwell.

O'BRIEN, P. (1988): "The Political Economy of British Taxation, 1660-1815," *Economic History Review*, 41, 1–32.

PHELPS BROWN, H. AND S. V. HOPKINS (1955): "Seven Centuries of Building Wages," *Economica*, 22, 195–206.

POMERANZ, K. (2000): *The Great Divergence: China, Europe, and the Making of the Modern World*, Princeton, NJ: Princeton University Press.

ROSENTHAL, J.-L. (1990): "The Development of Irrigation in Provence, 1700-1860: The French Revolution and Economic Growth," *Journal of Economic History*, 50, 615–638.

RUBIN, J. (2014): "Printing and Protestant: An Empirical Test of the Role of Printing in the Reformation," *Review of Economics and Statistics*, 96, 270–286.

SCHMELZING, P. (2020): "Eight Centuries of Global Real Rates, R-G, and the 'suprasecular' decline, 1311-2018," Bank of England Working Paper 845.

STEPHENSON, J. Z. (2018): "'Real' Wages? Contractors, Workers, and Pay in London Building Trades, 1650-1800," *Economic History Review*, 71, 106–132.

THOMAS, R. AND N. DIMSDALE (2017): "A Millenium of UK Data," Bank of England OBRA dataset, http://www.bankofengland.co.uk/research/Pages/onebank/threecenturies.aspx.

TIMMER, C. P. (1969): "The Turnip, The New Husbandry, and the English Agricultural Revolution," *Quarterly Journal of Economics*, 83, 375–395.

WEBER, M. (1905/2002): *Protestant Ethic and the Spirit of Capitalism*, New York, NY: Penguin Books, Peter Baehr and Gordon C. Wells, translation.

WRIGLEY, E. A. (1985): "Urban Growth and Agricultural Change: England and the Continent in the Early Modern Period," *Journal of Interdisciplinary History*, 15, 683–728.

——— (2010): *Energy and the English Industrial Revolution*, Cambridge, UK: Cambridge University Press.