State capacity, urbanization and the onset of modern economic growth*

Liam Brunt a
NHH-Bergen and CEPR

and

Cecilia García-Peñalosa b
Aix-Marseille University

May 2015

Abstract: We present a theory of the onset of the modern growth inspired by recent developments in economic geography. The existing literature suggests that technological change and industrialization prompted urbanization. We maintain that causality ran the opposite way. The development of the modern state and the subsequent increase in taxation triggered an unprecedented flow of labour into cities. In turn, high urbanization led to the creation and diffusion of knowledge within urban communities, thereby generating sustained technological change. We argue that our model can provide an explanation of why the First Industrial Revolution took place in England in the 1700s, and provide some evidence that is consistent with the model.

* This paper has benefited from comments at seminars at Universities of Alicante and Zurich, Bocconi University, Brown University, the 2010 Workshop in Honor of Steve Turnovsky in Vienna, and the Conference on Development at the Paris School of Economics, the 2nd AMSE Workshop in Growth and Development, as well as from discussions with Raouf Boucekkine, Fidel Pérez Sebastián, Patrick Pintus, Romain Rancier, Hans-Joachim Voth, and Tanguy van Ypersele. García-Peñalosa is a CNRS researcher at Greqam and her work was partly supported by the French National Research Agency Grant ANR-08-BLAN-0245-01.

a Email: Liam.Brunt@nhh.no
b GREQAM, Centre de la Vieille Charité, 2 rue de la Charité, 13002 Marseille, France. Email: cecilia.garcia-peñalosa@univ-amu.fr
1. Introduction

Urbanization and technological progress are closely linked. For economic historians and development economists, the creation and adoption of new technologies induce a flow of rural labour into cities as workers leave the traditional sector in order to use modern technologies. In contrast, a vast literature on economic geography maintains that urbanization itself is a cause of productivity gains. Going back to Marshall, economists have explored the idea that high density in cities results in learning and thus in both skill upgrading and innovation. Despite substantial evidence of the technological gains induced by agglomerations, models of long-run growth tend to see urbanization as a consequence and not as a cause of growth. The aim of this paper is to model the two-way relationship between urbanization and innovation in order to understand how they interacted at the onset of modern economic growth.

Our analysis is motivated by the observation that prior to the Industrial Revolution Europe was unusually urbanized. China, which is regarded by many historians as the most technologically advanced country in the Middle Ages, had much higher urbanization rates than Europe around the year 1000, with about 3 per cent of the population living in cities in the former and virtually zero in the latter. However, over the next 700 years Western Europe experienced a massive increase in urbanization rates, reaching almost 10 per cent by the year 1700, whilst no substantial change took place in China (see Maddison, 2001). This evidence raises the question of to what extent urbanization was the cause and to what extent the consequence of economic growth, and the aim of our model is to show that urbanization itself can be a trigger of development.

There are two central elements in our analysis. The first is our concept of development as a shift of labour resources from agriculture into manufacturing. Specifically, we follow Crafts and define an ‘industrial revolution’ as a shift of labour resources from agriculture into industry. Growth will then be partly driven by the flow of labour from farming to industry and can be measured by the increase in the consumption of manufactures. The second key aspect is the idea that cities are ‘special’ because they create frequent human encounters that are the basic means for the diffusion and creation of knowledge. We follow Marshall, and the formalization of his ideas by Glaeser (1999), in supposing that cities allow for both the creation of ideas and their transmission across agents, thus creating an environment that leads to innovation and growth. Our model hence makes urbanization -as opposed to industrialization- the key element in a theory of development.

---

1 See, for example, Harris and Todaro (1970) for a model of rural-urban migration and development and Rosenberg and Trajtenberg (2004) for a recent analysis of the effect of technological adoption on urbanization in the 19th century.
2 See Marshall (1890) as well as recent work such as Glaeser (1999) or Glaeser and Maré (2001), and Combes, Mayer and Thisse (2008) for an overview.
Our framework shares several features with Hansen and Prescott (2002), who examine how an economy that initially exhibits no growth can move to a Solow-type equilibrium. In their setting, industrialization occurs because productivity in manufacturing is sufficiently high, at some point, for workers to start moving into that sector. However, their model has a number of unsatisfactory features. First, technological change is exogenous, à la Solow. Second, technological change increases manufacturing productivity even if the manufacturing sector is not operational. Third, the agricultural sector disappears in the long-run, as the two sectors produce a single homogeneous good. We share with their analysis the assumption that there are two sectors (agriculture and manufacturing), that land exhibits diminishing returns, and that equilibrium at any point in time is driven by the allocation of labour across the two sectors. However, in our setting the process of industrialization starts from an increase in agricultural labour productivity, which drives labour away from agriculture and into cities and generates endogenous technological change in manufacturing.

The importance of ideas for the growth process has been recently emphasized by Lucas (2009), who maintains that “all knowledge resides in the head of some individual person” and hence the productive capacity of an economy will be the sum of the knowledge of its members. This concept of knowledge has two important implications. The first is that intellectual activity results from social interactions, hence meeting people is an essential aspect of the transmission and creation of knowledge. The second is the fact that heterogeneity in experiences and ideas is necessary in order to create new or better ideas, as maintained by Jovanovic and Rob (1989). We incorporate these two aspects in our analysis and argue that urbanization is a key element in the process of creation of knowledge because it determines the frequency of social interactions and the number of different experiences that an individual can observe.

There are three steps in our analysis. First, there is an increase in urbanization. There are several possible explanations for this event, such as the impact of the Black Death on relative wages and the structure of consumer demand; see Voigtländer and Voth (2006, 2011). We argue that an alternative candidate is the increase in state capacity that took place in Europe in early modern times.3 The creation of modern states was associated with increased extractions from agriculture, which caused rural-urban migration. As cities grew, the mechanisms discussed by the recent literature on economic geography started to apply. High urban density created more encounters and favoured the creation and diffusion of knowledge, resulting in productivity gains in manufacturing.

The second element in our model is the creation of knowledge in cities. We abstract from the use of machinery (capital) and suppose that an individual’s productivity depends solely on the

---

3 See Dincecco (2011) for evidence of the increase in taxation that occurred in early modern Europe.
number of ideas that he holds. There are then two ways in which he may increase his productivity.
On the one hand, agents in cities may acquire an idea by meeting someone with that idea and
imitating him. We follow Glaeser (1999) and suppose that the number of meetings is an increasing
function of the number of individuals in the city, *i.e.* of urban density. Higher density implies more
meetings and hence more imitation of ideas. In contrast, the low density in rural areas implies that
there are no meetings and hence no transmission of ideas. On the other, we follow Jovanovic and
Rob (1989) and suppose that an individual can also create new knowledge by observing the ideas or
experience of others, thus endogenously inventing new ways of production. For example, an agent
meets someone who is using a “boat”; then he meets another agent using a “steam engine”. He can
then combine their ideas to create a “steamboat”. This second aspect is crucial because it will be the
driver of long-run growth and implies that the rate of technological change in manufacturing will be
a function of urban density. Lastly, as manufacturing productivity increases the incentives to leave
the country-side rise, leading to higher urban density and hence a further increase in manufacturing
productivity, thus setting of a process of increasing urbanization and productivity.

In sum, we argue that increased state capacity leads to higher extraction rates in agriculture
and this generates a flow of labour into cities. Cities create frequent encounters amongst individuals
and become ‘wells of knowledge’ that raise productivity through imitation and innovation. This
increase in manufacturing productivity induces further rural-urban migration that accelerates the
process of knowledge diffusion and creation. As a result growth appears endogenously. Furthermore,
allowing for both innovation and imitation implies that those individuals who have not managed to
imitate others will not be able to use the more recent technology, and hence the model
simultaneously deliver fast technological change and slow productivity growth in manufacturing, in
line with existing estimates for the 18th century; see Crafts (2004).

We then confront our model with evidence for England. We argue that the First Industrial
Revolution occurred in England in the middle of the 18th century precisely because it was uniquely
urbanized *before* 1750. Data for the 17th and 18th century indicate that England was both more
urbanized and had higher agricultural labour productivity than other European countries.4 Moreover,
the period also exhibited rising fiscal revenues and a sharp increase in extraction rates out of
agriculture, implying that agricultural taxation is a possible cause for the observed increase in
urbanization, which in turn resulted in the explosion of knowledge that gave rise to the First
Industrial Revolution. In other words, the expansion of state capacity and the subsequent increase in

---

4 See section 4 below.
urbanization made England experience an unprecedented environment for the creation and diffusion of knowledge and gave rise to the beginning of modern growth.

Our paper is related to several strands of literature. The first are the recent developments in growth theory that explain how the world moved from stagnation to growth, best represented by Unified Growth Theory (UGT); see Galor and Weil (2000), Galor and Moav (2002) and Galor (2005). UGT is at the core of our understanding of the appearance of modern growth and of the close relationship between population dynamics and economic development, yet specific differences in the timing and speed of the transition are still not fully understood. For example, Bar and Leukhina (2010) combine population dynamics with a mode of structural change and explore how demographics can help us understand the economic development of England. However, they follow Hansen and Prescott (2002) and assume exogenous technological change in manufacturing. Boucekkine, Peeters and de la Croix (2007) examine the role of population growth and density in the transition to modern growth. The key element in their analysis is the endogeneity of both the decision to attend school and the decision to create a school. An increase in population density raises the number of schools, which reduces the time cost of getting to school and hence makes education cheaper, thus increasing educational investments by agents. Their analysis helps understand educational patterns in early modern England, yet abstracts from technological change which is assumed to be exogenous. We consider an additional aspect that would complement UGT and which provides a possible explanation for why England was first in terms of technological change.

A number of recent papers have focussed on the unique features of Europe and, particularly, England, before the Industrial Revolution. For example, Desmet and Parente (2012) emphasize the role of market size and argue that by the early 18th century markets in England were more integrated, and thus larger, than elsewhere and this allowed firms to implement cost-reducing production technologies. Our analysis is closely related to Voigtländer and Voth (2006, 2011) who examine the factors which resulted in the increase in per capita output experienced in Western Europe in early modern times. Both models share with ours an approach based on the migration of workers from agriculture into manufacturing. Voigtländer and Voth (2006) examine how weather shocks and demographic conditions could have pulled some economies out of the Malthusian trap, and identify the importance of demographic patterns in countries like France and England for the prospects of industrialization. Voigtländer and Voth (2011) maintain, like we do, that cities played a crucial role in

---

5 The literature by economic historians on the First Industrial Revolution is obviously extensive; see for example Mokyr (2008, 2009) for a recent discussion. In contrast, growth models focusing on that period are rare, although there are some exceptions such as Voigtländer and Voth (2006, 2011).
early modern times. In their setup, the Black Death is the exogenous trigger that initially raises wages and results in migration towards cities; city growth then prevented a population increase due to the poor health environment and associated increase in mortality which allowed Europe to escape from the Malthusian trap. Our analysis complements theirs by maintaining that the frequent interactions occurring in cities not only resulted in the transmission of viruses but also of ideas.

Lastly, the paper is also related to the literature on economic geography that has examined to what extent agglomeration results in higher productivity. Following the argument first put forward by Marshall (1980) and the seminal work of Jacobs (1969), numerous authors have modelled how the environment offered by cities improves the prospects for generating and diffusing new ideas. Cities affect the productivity of firms and of workers, as they generate both technology spillovers across the former and also learning and skill-acquisition by the latter. Various mechanisms can give rise to innovation and learning in cities; see Duranton and Puga (2004) for a review. For example, Duranton and Puga (2001) show that the diversification of workers and firms present in urban environments facilitates search and experimentation, thus increasing innovation. One of the most influential arguments has been the idea that proximity to individuals with greater skills or knowledge facilitates the exchange of and diffusion of knowledge. In this context, since higher urban density results in more frequent encounters it leads to greater transmission of skills and ideas amongst workers; see Glaeser (1999), Glaeser and Maré (2001), and Jovanovic and Rob (1989). There is also a substantial body of empirical evidence regarding the advantages of cities for innovation. Recent empirical work indicates that location has become a key factor in explaining innovation and technological change, and one of the findings of this literature is that knowledge spillovers are large and important in increasing the productivity of economic actors; see the review by Audretsch and Feldman (2004). Our paper uses the insights of this vast literature and models the process of imitation and innovation that takes place in cities. In line with empirical findings we assume a process of knowledge creation in which innovation and productivity increase with urban density, and this will be one of the cornerstones of our theory.

The rest of the paper is organised as follows. Section 2 presents the basic model and examines the determinants of urbanization. In Section 3 we examine the mechanisms for the creation and transmission of ideas, and incorporate these in our macroeconomic model in order to obtain the joint dynamics of urbanization and knowledge. Section 4 discusses the historical evidence that supports our arguments. A numerical example is then provided, while section 6 concludes.
2. A model of urbanization

We begin by setting out the components of the model. This section considers the determinants of urbanization, taking as given productivity in manufacturing. The endogenous evolution of this productivity is discussed in section 4.

2.1 Population and preferences

We consider an overlapping-generations setup. Agents live for two periods. In the first they are apprentices; in the second they work and have a single offspring so that the number of dynasties remains constant over time. There are $N$ dynasties of workers and $\varepsilon N$ dynasties of landlords, with $\varepsilon < 1$. Landlords extract rents from the farmers to whom they let their land. Workers have two possible occupations, working as a farmer or in manufacturing. We suppose that these occupations require agents to live either in the countryside or in cities, respectively. Hence agents will live in one of two locations, cities or the countryside, depending on their occupational choice. Young agents are born in the location chosen by their parents and make their occupational/location choice in the first period of their life. That is, they migrate as young adults and the next period they work. During the first period of their life, they may or may not acquire “ideas” that make them more productive, according to a process defined below.\(^6\)

We follow Voigtläder and Voth (2008) and suppose that preferences are identical for all agents and take the form

$$U(c_{at},c_{mt}) = \begin{cases} 
\beta (c_{at}-\underline{c})^\alpha & \text{if } c_{at} < \underline{c} \\
(c_{at}-\underline{c})^\alpha c_{mt}^{1-\alpha} & \text{if } c_{at} \geq \underline{c}
\end{cases} \tag{1}$$

where $c_{at}$ and $c_{mt}$ are respectively consumption of the agricultural and the manufacturing good in the second period of life, and $\underline{c}$ a minimum food requirement. This yields the demand functions

$$c_{at} = \alpha y_i + (1-\alpha)\underline{c} \tag{2a}$$

$$c_{mt} = \frac{(1-\alpha)}{p_i}(y_i - \underline{c}) \tag{2b}$$

where $y_i$ is the income of the individual, and the prices of the agricultural and the manufacturing good are, respectively, 1 and $p_i$. The resulting indirect utility function can be expressed as

\(^6\) The fact that individuals acquire skills when young, i.e. before they start working, fits well with a structure of apprenticeships whereby young individuals chose their ‘trade’ at a relatively early age and acquired specific skills.
where \( a \equiv \alpha^a (1 - \alpha)^{1-a} \).

### 3.2 Technology and factor payments

**Agriculture**

Agricultural output is produced with a constant-returns-to-scale technology using labour and land, \( T\),

\[
Y_{at} = T^{\gamma} \left( AL_{at} \right)^{1-\gamma},
\]

(4)

where \( A \) denotes productivity in the agricultural sector, \( L_{at} \) agricultural labour, and \( 0 < \gamma \leq 1 \). Farmers rent the land which is the property of landowners. The wage in the agricultural sector is supposed to be proportional to output per worker. That is,

\[
w_{at} = \omega A^{\gamma} \left( \frac{T}{L_{at}} \right)^{1-\gamma},
\]

(5)

where \( \omega \leq 1 \). The parameter \( \omega \) is a distributional parameter that tells us what fraction of average agricultural output farmers keep. If farmers receive the marginal product of labour we have \( w_{at} = (1 - \alpha) A^{\gamma} \left( T / L_{at} \right)^{1-\gamma} \). We want to be more general and suppose that landlords and the authorities extract a fraction \( \tau \) of the revenue, so that \( \omega = 1 - \tau \). In what follows we will term \( \tau \) the tax rate, even if not all of it is reaped by the state. Define \( A_{c} = A^{\gamma} \left( T / N \right)^{1-\gamma} \) which is simply output per worker when the entire labour force works in agriculture. Then, the wage can be expressed as \( w_{at} = (1 - \tau) A_{c} n_{at}^{1-\gamma} \), where \( n_{at} \) is the share of population employed in farming. The wage in agriculture will be higher when the rents extracted by landlords and when employment in farming are lower.

Lastly, note that the presence of the a minimum food requirement, \( \zeta \), can be interpreted as a Malthusian effect if we suppose that agricultural per capita output below this level would lead to death and hence a reduction in the population.\(^7\) We can then define \( n_{c} = \left( \zeta (1 + \varepsilon) / A_{c} \right)^{1-\gamma} \) which is the level of agricultural employment required in order to provide the minimum food requirement, \( \zeta \),

\(^7\) We do not consider the other element in Malthusian theories, namely that increases in agricultural output result in higher fertility and hence a larger population. See Galor (2005).
to the entire population and assume that \( c(1+\varepsilon) < A_a \) so that \( n_a < 1 \). This assumption simply says that agricultural TFP, \( A \), and/or land per capita, \( T/N \), have to be sufficiently high for the minimum food requirement to be produced. Otherwise there will be no demand for manufactures. Agricultural employment cannot fall below the lower bound \( n_a \), which in turn defines an upper bound for the share of manufacturing employment, namely \( \bar{n}_m \equiv (1-n_a) \). Lastly, the minimum food requirement also imposes a constraint on the tax rate, as the wage of farmers must be sufficiently high for them to consume at least \( c \). That is, \( \tau < 1 - c n_a / A_a \). A sufficient condition for this is \( \tau < \tau = 1 - c / A_a \), as the lowest possible productivity per worker is obtained when all workers are employed in farming, i.e. \( n_{at} = 1 \). In what follows we suppose that this condition holds.

Manufacturing

Manufacturing takes place in cities. It is characterized by constant returns and the only input is labour. Artisans without an idea produce one unit of output, while those with \( i \) ideas produce \((1+hi)\). Aggregate manufacturing output is then

\[
Y_{mt} = B_t L_{mt}
\]

where \( L_{mt} \) is employment in manufacturing and \( B_t \) is average productivity in that sector, which depends on the average number of ideas that those who work in manufacturing have, denoted \( I_t \). That is, \( B_t = 1 + hi \), where \( I_t \) will be endogenously determined in section 4.

We suppose that skills are sector specific. They are acquired when young and hence individuals will not move when old. Since the price of the manufacturing good is \( p_t \), the wage in manufacturing will then be \((1+hi)p_t \) for those with ideas and \( p_t \) for those without. Hence the expected wage in manufacturing is simply \( w_{mt} = p_t B_t \), and is independent of employment in the sector at time \( t \).

There is a disutility associated with living in cities, denoted \( v_t \), so that the expected utility of a worker living in the city and working in manufacturing is

\[
EU_{mt} = a(w_{mt} - c)p_t^{-\alpha - 1} - v_t
\]

This cost creates a wedge between rural and urban utilities that, in equilibrium, will results in higher wages in
There are various possible justifications for this cost. Murphy, Shleifer and Vishny (1989) argue that manufacturing work itself generates a disutility for which workers have to be compensated. Voigtländer and Voth (2008) emphasize the health costs of living in cities, given that they had a worse disease environment than that found in rural areas. Although either interpretation is possible, we suppose that $v_t$ captures the cost due to a higher probability of disease associated with city life. We further suppose that the cost takes the form $v_t = adp_t^\alpha$, where $d$ is a positive constant. Assuming that the health cost is proportional to the price of manufactures will imply that, as this price falls over time due to technological improvements, the disutility associated with city life will fall (or equivalently, that the probability of disease falls). In other words, we are assuming that increases in manufacturing know-how will result in improvements in medical know-how that reduce the health cost associated with urban dwelling.

2.3. Static equilibrium

Employment in agriculture and manufacturing

Workers can freely move between agriculture and manufacturing, and will do so in order to equalize the expected utility in the two sectors. The decision to migrate is taken by young individuals at time $t-1$ on the basis of their (rational) expectations of the income they would get at $t$ in each of the two sectors.

Utility in farming is simply $U_{at} = a(w_{at} - c)p_t^{\alpha-1}$. Equating this to $EU_{mt}$, and using the expressions above for the wages in agriculture and manufacturing we obtain the following relation between the price of manufactures and employment in agriculture

$$p_t = \frac{(1-\tau)A_n^{\alpha-1}}{B_t - d}$$

(7)

This equation implies a negative relationship between $p_t$ and $n_{at}$ since a higher price increases the manufacturing wage and hence results in migration away from farming. Higher productivity, $B_t$, or

---

8 Evidence of a gap between rural and urban wages is provided by Clark (2001). His data shows that an urban labourer earned 4% more than a farm labourer in 1500 and that this gap rose over time reaching 40% by 1980 (Clark, 2001, Table 1).

9 Note that some of the young in the city get ideas and others do not. This means that ex post, those in the city who got a below-average number of ideas will have lower utility than those in agriculture. Because skills are sector-specific and need to be acquired when young, these individuals will not move into farming.
higher rents, i.e. a higher value of $\tau$, result in a lower share of agricultural employment for any given price, since the former increases the manufacturing wage and the latter reduces the agricultural wage.\footnote{Other analyses of occupational decisions during the Industrial Revolution have been proposed, such as that of Doepke and Zilibotti (2008) who focus on the appearance of middle-class values. We view their approach as focusing on social developments that took place during the Second industrial Revolution.}

**Goods market equilibrium**

To obtain the goods market equilibrium, we need to consider the demands of workers and landlords. Those of workers are given by equations (2) above. There are $\varepsilon N$ landlords, each of whom gets rents equal to $\tau Y_a / \varepsilon N$. Since they have the same utility function as workers their demands are also given by (2).\footnote{We could allow landlords to demand imported ‘luxury goods’ as well as agricultural goods and manufactures, so that a fraction of output is not spent domestically. This would change slightly the expression for the relative price but would have no qualitative impact.}

Equating the supply and the demand for agricultural goods we have

$$NA_p n'_{aw} = N\left(\zeta(1+\varepsilon) + \alpha(\bar{w}-\zeta) + \alpha(\tau A_n n'_{aw} - \varepsilon \zeta)\right)$$

(8)

where $\bar{w}$ is the average wage. The first term in brackets is the minimum food requirement that the population consumes, the second is the additional demand from workers, while the third is the additional demand from landlords. Similarly, equilibrium in the manufacturing good sector is given by

$$p_t B_t (1 - n_{at}) N = (1 - \alpha) N (\bar{w} + \tau A_n n'_{aw} - \varepsilon (1+\varepsilon))$$

(9)

These two equations imply

$$(1 - \alpha) A_n n'_{aw} = (1 - \alpha)(1+\varepsilon)\zeta + \alpha p_t B_t (1 - n_{at})$$

(10)

which gives the goods market equilibrium relationship between agricultural employment and $p_t$.

The higher employment in agriculture is, the higher the price of manufactures.

**The equilibrium allocation of labour**

Equations (7) and (10) jointly determine the price and agricultural employment. Since equation (7) implies a negative relation and (10) a positive one, there will be a unique pair of equilibrium price and employment for each set of parameter values. Using equation (7) to substitute for the price, we have the equilibrium share of agricultural employment,
We can now define the functions
\[ g(n_{at}) = (1 - \alpha)A_a n_{at}^{\gamma} \]  
\[ f(n_{at}; \tau, B_t) = (1 - \alpha)(1 + \varepsilon)C + \alpha A_a \frac{1 - \tau}{1 - d / B_t} \frac{1 - n_{at}}{n_{at}^{1 - \gamma}} \]  

It is possible to show that \( g(n_{at}) \) is increasing and concave and \( g(0) = 0 \), while \( f(n_{at}) \) is decreasing and convex. As long as \( A_a > (1 + \varepsilon)C \) and \( \tau < \tau^* \), the two functions intersect once and there is a unique equilibrium level of agricultural employment \( n_{at}^* \). Note that these conditions capture the Malthusian trap, since \( A_a > (1 + \varepsilon)C \) is the requirement that when the entire workforce is employed in agriculture they produce enough to satisfy the minimum food requirement for the entire population and \( \tau < \tau^* \) that farmers can consume at least \( C \). \(^{12}\)

---

\(^{12}\) If \( A_a < (1 + \varepsilon)C \), even with the entire population working in farming there would not be enough food to satisfy the minimum consumption requirement, resulting in a reduction in the population as in the UGT. We abstract from these considerations, and simply assume that \( A_a \) is sufficiently high.
Figure 2 represents graphically the equilibrium by plotting the functions $g(n_a)$ and $f(n_a)$. Under the assumption that $A_a > (1 + \varepsilon)\xi$ and $\tau < \bar{\tau}$, there will always be a fraction of workers employed in manufacturing and living in cities. From equation (11) we can see that there are two key variables that affect the share of agricultural employment: $\tau$ and $B_r$. Higher values of these tilt downwards the $f(n_a)$ function and result in a lower value of $n_a$, as depicted in figure 3 by the red schedule. The intuition is simple. A higher rent reduces the wage in farming, while a higher productivity in manufacturing increases the wage in cities, and both induce a shift of labour away from agriculture.

The other variable of interest is agricultural productivity, $A$, which affects the static equilibrium through $A_a$. A higher value of $A$ shifts both $g(n_a)$ and $f(n_a)$ upwards, implying that agricultural employment may increase or fall. This ambiguous effect is standard and is due to the fact that higher agricultural TPF increases the marginal product of labour in farming but reduces the price of agricultural goods, so that the marginal value product of labour may be higher or lower than before the TFP increase. Differentiating (11) we can see that, as long as $\xi$ is strictly positive, prices will fall sufficiently for higher TFP to result in a reduction in agricultural employment.

3. Cities and the transmission of ideas

3.1 The creation and transmission of ideas

We consider now the evolution of ideas. We suppose that at each period in time at most one idea can be invented according to a process defined below. Ideas are ordered from $i=1, 2, \ldots t$ according to the period in which they were invented. An individual may hold all or only a fraction of these ideas, but always in a sequential order. Hence if two individuals hold, respectively, $j$ and $j+1$ ideas, the second will hold all the ideas that the former holds and one additional one. We denote by $P(i,t)$ the probability that an individual has at least $i$ ideas at time $t$, and by $p(i,t)$ the probability that an individual has exactly $i$ ideas at $t$. Since there are a large number of individuals, these probabilities are, respectively, the fraction of the urban population that holds at least $i$ ideas and that which holds $i$ ideas. The sequentiality of ideas also implies that $p(i,t) = P(i-1,t) - P(i,t)$. As we will see, at any
There are two ways in which a young individual may increase his productivity (become skilled). First, those in cities may acquire existing ideas by meeting someone with ideas. Second, young agents may also innovate and acquire one new idea during the period. We follow Glaeser (1999) and suppose that the number of meetings, $D_t$, is a function of urban density. In particular we suppose that the number of meetings is an increasing and convex function of the number of old individuals in the city, and for analytical tractability assume that it takes the form $D(n_{old}) = \rho n_{old}^2$. Following Glaeser (1999), we suppose that a young agent who meets someone with at least $i$ ideas, acquires these ideas with probability $z$. This means that in each meeting the probability of acquiring at least $i$ ideas is $zP(i,t)$, the probability of imitating times the probability that the individual of the previous generation that has been met has at least those skills. A young person will not imitate if he fails to learn in every meeting. The total probability of not acquiring at least $i$ ideas is $(1-zP(i,t))^\rho$, and the probability of acquiring at least $i$ ideas is one minus this amount. That is, the fraction of the population that has imitated at least $i$ ideas is given by

$$M(i,t+1) = 1 - (1-zP(i,t))^\rho, \quad (13)$$

An individual may also come up with one new idea, i.e. may innovate. This may happen if he meets a sufficiently large number of people. To capture this idea, we suppose that only a fraction of the meetings that take place can potentially give rise to an innovation and assume that only $\sqrt{D_t}$ of meetings provide interactions that may result in innovation. The different effect of density on innovation and imitation captures the idea that the former requires more interactions, and hence much greater density, than the latter. For those meetings were potentially useful information is transmitted, we suppose that the probability that the individual innovates after the meeting is $\delta$. Hence, an individual innovates with probability $q_{i+1} = 1 - (1-\delta)^\rho$, where $\rho = \sqrt{\rho}$. 

There are two crucial differences between innovation and imitation. First, the amount of imitation depends on how skilled the previous generation was, while innovation is independent of this. Second, imitation increases rapidly with the number of meetings, since all meetings can result in a transmission of information. In contrast, innovation requires a ‘sufficiently large’ number of meetings because the thought process of the individual that leads to innovation is the result of the
complementarity between the experiences of two individuals, and only a fraction of meetings can yield it. Note also that the assumption that agents can at most have one new idea during their youth implies that, at any point in time \( t \), there is an upper bound to the number of ideas held by an agent, namely \( t \). This assumption is not needed for our results but yields a tractable analytical expression for the number of ideas and its distribution at any point in time.

In this context there are two ways in which average productivity will grow: individuals copy the ideas of others that they meet—the diffusion of knowledge- and individuals observe the experiences of others in order to create new knowledge. The importance of these two aspects for the growth process is examined by Jovanovic and Rob (1989). In their setup, agents possess ideas of different ‘quality’. When two agents meet, there is both imitation by the less informed agent and invention of new knowledge by both agents, and the amount of imitation and innovation depends on the distance between the two. We simplify their setup by separating the two processes.

Let us now consider the distribution of the number of ideas. The average number of ideas is simply \( I_t = 1 \cdot p(1,t) + 2 \cdot p(2,t) + \ldots + (t-1) \cdot p(t-1,t) + t \cdot p(t,t) \). Expressing the number of individuals with \( i \) ideas as \( p(i,t) = P(i,t) - P(i+1,t) \) and since \( p(t,t) = P(t,t) \), we can write

\[
I_t = P(1,t) + P(2,t) + \ldots + P(t-1,t) + P(t,t) = \sum_{i=1}^{t} P(i,t)
\]

and it follows that \( I_{t+1} = \sum_{i=1}^{t} P(i,t+1) \).

Who may have a certain number of ideas? The fraction of individuals that have at least \( i \) ideas is given by

\[
P(i,t+1) = M(i,t+1) + m(i-1,t+1)\left[1 - (1 - \delta)^{m(i-1,t+1)}\right]
\]

where \( m(i-1,t+1) \) is the fraction of the population that has imitated exactly \( i-1 \) ideas at \( t+1 \), and

\[
m(i-1,t+1) = M(i-1,t+1) - M(i,t+1)
\]

Note also that \( P(1,t+1) = M(1,t+1)(1 - q_{t+1}) + q_{t+1} \) and \( P(t+1,t+1) = M(t+1,t+1)q_{t+1} \).

Using (15), the dynamic equation for the average number of ideas in the population is then

\[
I_{t+1} = \sum_{i=1}^{t} M(i,t+1) + \left[1 - (1 - \delta)^{m(i-1,t+1)}\right]\sum_{i=1}^{t} m(i-1,t+1),
\]

which rearranging and using (16) can be rewritten as

\[
I_{t+1} = (t+1) - (1 - \delta)^{\sum_{i=1}^{t} m(i-1,t+1)} - \sum_{i=1}^{t} \left[1 - zP(i,t)\right]^{\sum_{j=1}^{\infty} q_{t+1}}
\]
The first term captures the fact that a new idea is invented each period. This aspect, the “creation of new ideas” is crucial, as it will be what drives long-run growth. The second term captures the extent to which innovation is prevalent in the population and depends on urbanization. The last term captures the extent of imitation, and again depends on \( n_{mt} \). We can see from equation (17) that the current average number of ideas is increasing in the number of ideas held by the population last period and in the number of individuals that were living in cities.

### 3.2 The dynamics of urbanization

Equations (11) and (17) together with the labour market clearing condition, \( n_{mt} = 1 - n_m \), and the expression for average manufacturing productivity, \( B_t = 1 + hI_t \), determine the dynamics of skills and urbanization. We can express these as the pair of equations

\[
\frac{1 - n_{mt}}{n_{mt}} \left( 1 - \left( \frac{1 - n_{mt}}{1 - n_m} \right)^\gamma \right) = \frac{\alpha}{1 - \alpha} \left( 1 - \tau \right) \frac{1 + hI_t}{1 + hI_t - d} \tag{18a}
\]

\[
I_{t+1} = \left( 1 - (1 - \delta) \pi_{mt} \right) + \sum_{i=1} \left( 1 - zP(i,t) \right) \pi_{it} \]

where \( \pi_{mt} = (1 - n_m) \) is the highest possible share of manufacturing employment. Equation (18a) implicitly defines manufacturing employment as a function of the average number of ideas \( I_t \), that is \( n_{mt} = F(I_t; \tau) \). Then, we have that the dynamics of the model are given by

\[
I_{t+1} = \left( 1 - (1 - \delta) \pi_{mt}(i) \right) + \sum_{i=1} \left( 1 - zP(i,t) \right) \pi_{it}(i) \]. \tag{19}

This equation describes the dynamics of the number of ideas in the economy, from which we obtain the dynamics of urbanization. Since \( F'(I_t; \tau) > 0 \), then \( dI_{t+1} / dI_t > 0 \) implying that ideas grow over time, and, by equation (18a), so does urbanization.

### 3.3 The dynamic behaviour of the economy

The dynamics of the economy are given by

\[
\frac{1 - n_{mt}}{n_{mt}} \left( 1 - \left( \frac{n_m}{1 - n_{mt}} \right)^\gamma \right) = \frac{\alpha}{1 - \alpha} \left( 1 - \tau \right) \frac{1 + hI_t}{1 + hI_t - d} \tag{E.1}
\]
\[ I_t = \sum_{i=1}^{t} P(i,t) \]  
(E.2)

\[ P(i,t+1) = 1 - (1 - zP(i,t))^{n_{\mu}} + \left(1 - (1 - \delta)^{n_{\mu}}\right)\left[(1 - zP(i,t))^{n_{\mu}} - (1 - zP(i-1,t))^{n_{\mu}}\right] \]  
(E.3)

The first equation yields urbanization, \( n_{\mu} \), given average knowledge, which by (E.2) is determined by the distribution of ideas at \( t \), i.e. \( P(i,t) \). Then the initial distribution of ideas and urbanization determine, by equation (E.3), next period’s average ideas and their distribution. This will in turn determine next period’s average productivity and the level of urbanization, \( n_{\mu+1} \).

Knowledge, defined as the number of ideas available in the economy will grow without bound, inducing a flow workers into the city until the lower bound of agricultural employment, \( n_a \), is reached. At this point migration will stop. Note that even if migration stops, both knowledge and productivity will keep growing. Knowledge grows because, as long as \( n_{\mu} > 0 \), each period a new idea will be invented with positive probability. Since productivity in manufacturing is \( B = 1 + hi \), we can see from equation (18b) that it grows due to both the diffusion of existing knowledge and the creation of new ideas. In general, the rate of growth of productivity will be lower than the rate of growth of knowledge because not all workers will acquire the latest idea. In the early stages of development, when urbanization is low, imitation will be moderate and the gap between knowledge and productivity will be large. As urbanization increases, the gap between the two will narrow due to a greater diffusion of knowledge.

Per capita real output in the economy is given by
\[
\frac{Y_i}{(1 + \varepsilon)N} = \frac{A_i n_{\mu}^{1/\alpha} + p_i B n_{\mu}}{(1 + \varepsilon)P_i^{1-\alpha}}. 
\]  
(20)

In the short-run, there are four effects of increased productivity on output. First, the allocation of labour across sectors has an impact on output, as the reduction in agricultural employment raises output per worker in his sector. Second, output per worker in manufacturing grows due to the increase in average productivity, \( B_i \). Lastly, the price of manufactures, and hence the price index, \( p_i^{1-\alpha} \), falls as productivity in industry increases, having a further effect on real output growth. Using equations (7) and (18a), real per capita output can be expressed as
\[ \frac{Y_t}{(1 + \varepsilon)N} = \left( \frac{B_t - d}{1 - \tau} \right)^{1-a} A_{\tau}^{\alpha} \left( \frac{n_{aat}}{n_{aat}} \right)^{1-a(1-\gamma)} \left( 1 - (1 - \alpha) \left( \frac{n_{aat}}{n_{aat}} \right)^{\gamma} \right) \]  

(21)

In the long-run, the allocation of labour will be constrained by the number of farmers needed to produce the food requirement, yielding the following expression for per capita output:

\[ \frac{Y_t}{(1 + \varepsilon)N} = \left( \frac{B_t - d}{1 - \tau} \right)^{1-a} A_{\tau}^{\alpha} \left( \frac{n_{aat}}{n_{aat}} \right)^{1-a(1-\gamma)} \left( 1 - (1 - \alpha) \left( \frac{n_{aat}}{n_{aat}} \right)^{\gamma} \right) \]  

(22)

which grows with the level of productivity.

### 3.3 An example

Section 5 below provides numerical solutions for the model in order to examine the joint evolution of urbanization and productivity. In order to obtain an intuition of this relationship we finish this section by examining an analytically tractable example. We make two simplifications. First, we consider a situation in which there are constant returns to labour in agriculture, that is \( \gamma = 1 \). Second, we suppose that there is no imitation and hence the only source of ideas is innovation by agents. Since, there is no transmission of ideas young agents can at most have one idea. As a result, the average number of ideas at any point in time is simply equal to the fraction of the population that holds one idea, denoted \( s_{t+1} \).

The dynamic equations governing the evolution of urbanization and the fraction of the population with an idea are now

\[ s_{t+1} = 1 - (1 - \delta)^{\phi_{aat}} \]  

(23a)

\[ n_{aat+1} = \frac{(1 - \alpha) \bar{n}_a}{1 - \alpha \bar{\tau}} \left( 1 - \frac{ad(1 - \tau)}{(1 - \alpha \bar{\tau})(1 + h s_{t+1}) - (1 - \alpha)d} \right) \]  

(23b)

These two equations together yield the interior solution

\[ n_{aat+1} = \bar{n}_a \frac{(1 - \alpha)\left(1 + h - h(1 - \delta)^{\phi_{aat}}\right) - (1 - \alpha)d}{(1 - \alpha \bar{\tau})\left(1 + h - h(1 - \delta)^{\phi_{aat}}\right) - (1 - \alpha)d} \]  

(24)

which gives the dynamics of urbanization. Let us define the function

\[ G(n_{aat}; \tau, A) = \bar{n}_a \frac{(1 - \alpha)\left(1 + h - h(1 - \delta)^{\phi_{aat}}\right) - (1 - \alpha)d}{(1 - \alpha \bar{\tau})\left(1 + h - h(1 - \delta)^{\phi_{aat}}\right) - (1 - \alpha)d} \]  

(25)

where \( G(n_{aat}; \tau, A) \) is less than \( \bar{n}_a \) for all \( n_{aat} \), and agricultural productivity has an impact through its effect on \( \bar{n}_a \) (with \( \partial \bar{n}_a / \partial A > 0 \)). Differentiating with respect to \( n_{aat} \) we have that \( G(n_{aat}; \tau, A) \) is
increasing and concave, while \( G(0; \tau, A) > 0 \). Moreover, \( \partial G / \partial \tau > 0 \) and \( \partial G / \partial A > 0 \), implying that the function \( G(.) \) shifts upwards with \( \tau \) and with \( A \), and \( G(n_{m+1}, 1, A) = \bar{n}_m \).

The dynamics of urbanization are depicted in figure 2. Note first that since there is no diffusion of ideas, the economy will always converge to a steady state with a constant rate of urbanization which is below \( \bar{n}_m \). The \( n_{m+1} \) schedule is increasing and concave, implying that for any initial rate of urbanization \( n_{m0} < n_m^* \), the fraction of the population living in cities will increase over time until it reaches its steady state value \( n_m^* \).

Figure 3: Taxes and the dynamics of \( n_{mt} \)

Consider now an increase in the tax rate that shifts the \( n_{m+1} \) schedule upwards from the continuous line to the dotted one. The economy will move to a new steady state with higher urbanization, denoted \( n_{m}^{**} \). A shock to agricultural productivity would have a similar effect, since it would increase \( \bar{n}_m \), shifting the urbanization schedule upwards.
We can now obtain an expression for per capita output in the economy, which from (21) is

\[
\frac{Y_t}{(1+\epsilon)N} = \left(1 + h - h(1-\delta)^{\pi_{mc}} - d\right)^{1-\alpha} \alpha^{\alpha \epsilon} \left(1 - \frac{1-\pi_{mc}}{1-n_{mt}}\right)
\]

which implies that as the economy becomes more urbanized per capita income grows. As manufacturing employment converges to its steady state value, \( n_{mt}^* \), output per capita converges to its steady state and the growth rate to zero.

4. Historical evidence

Our model explains the international and temporal variation in four key variables: agricultural extractions, agricultural labour productivity, urbanization levels, and output growth. In this section we discuss the salient features of these variables across space and through time. We argue that 18th century Europe, and in particular England, fit well the dynamics of our model. Before the First Industrial Revolution, England experienced a period of high and growing urbanization. Moreover, urbanization was accompanied by high agricultural output and a sharp increase in the rate of extraction out of the agricultural sector. Our model predicts that it was the increase in extraction that induced a flow of labour into cities and led to higher output per worker in agriculture, thus bringing about an increase in knowledge and manufacturing productivity.

This section discusses the available data on urbanization, agricultural productivity and taxation. The evidence we present indicates that and argues that by 1750 England was unique in all these dimensions. We then use some of these data to simulate the dynamics of the model in section 5.

Patterns and changes in urbanization

As we have argued in the introduction, by early modern times Europe exhibited unprecedented rates of urbanization. Amongst the western European nations, England and, to a lesser extent, Scotland and Ireland had some of the highest urbanization rates, as can be seen from figure 1. Notably, England was highly urbanized when compared to other large nations such as France or Spain. Urbanization was also high in Belgium and Netherlands. Belgium was the first country in continental Europe in which the industrial revolution took place, while the fact that the Netherlands experienced a relatively late industrialization has been seen a puzzle by historians; see Mokyr (1999b).
Why was the urbanization rate high and rising in England and Scotland during the 18th century? The reason is simply that there were many more large cities. This aspect is crucial because, in our model, it is not the urbanization rate of a country per se that matters but the absolute size of cities. The frequency of meetings (and hence the diffusion and creation of ideas) is determined by the population of the city, so it is more effective to herd people into a small number of larger cities than a large number of small cities. Table 1 reports the urban population living in the largest cities in the western European economies at benchmark dates. England has the greatest population living in large cities already by the middle of the 18th century, pulls dramatically ahead of other countries between 1750 and 1800 and accelerates thereafter.

---

13 Historical evidence on the geographical size of cities is scarce, and certainly as population size grew so did the physical size of cities. However, the latter is likely to have grown more slowly, and cities tended not to increase the size of ‘public spaces’ such as markets were commercial (and hence intellectual) exchanges took place.

14 We define “large” cities as having population of 200 000 or above. Of course, one could argue about the appropriate definition of “large” cities. But the result that England pulled rapidly ahead of other countries in the late eighteenth century is robust to alternative definitions and the data that we report here are broadly representative of the situation.
Table 1. Population living in the largest western European cities, 1750-1850.

<table>
<thead>
<tr>
<th></th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>0</td>
<td>247,000</td>
<td>431,000</td>
</tr>
<tr>
<td>France</td>
<td>570,000</td>
<td>550,000</td>
<td>1,053,000</td>
</tr>
<tr>
<td>Italy</td>
<td>339,000</td>
<td>430,000</td>
<td>618,000</td>
</tr>
<tr>
<td>Netherlands</td>
<td>210,000</td>
<td>217,000</td>
<td>225,000</td>
</tr>
<tr>
<td>Spain</td>
<td>0</td>
<td>0</td>
<td>501,000</td>
</tr>
<tr>
<td>England</td>
<td>675,000</td>
<td>948,000</td>
<td>3,148,000</td>
</tr>
<tr>
<td>Ireland</td>
<td>0</td>
<td>200,000</td>
<td>262,000</td>
</tr>
<tr>
<td>Scotland</td>
<td>0</td>
<td>0</td>
<td>345,000</td>
</tr>
</tbody>
</table>

Source: Bairoch et al. (1988).

Table 2 considers the four largest cities in each of the major western European countries in 1700 and tracks their expansion. The fast rate of urbanization of the UK stands out. The rise of Birmingham, Liverpool, and Manchester after 1750 is especially marked, as is the expansion of Dublin and Glasgow (both in the United Kingdom, but not England); by 1850 they were all larger than most European capital cities. Some initially large cities, such as Amsterdam, grew by only 12 per cent over a century and a half, and while the large Italian cities (Naples and Milan) experienced faster growth it remains well below that observed in the United Kingdom. Table 2 also indicates that England experienced a growth spur before 1750, i.e. before the First Industrial Revolution, with the three cities other than London at least doubling their populations in the first half of the 18th century. By 1850 there were in Western Europe 15 ‘large’ cities, i.e. of more than 200,000 inhabitants, and 6 of them were in the British Isles, accounting for half of the urban population living in large cities in 1850.
Table 2. Expansion over time of the largest European cities in 1850 (population in `000s).

<table>
<thead>
<tr>
<th>City</th>
<th>Country</th>
<th>1700</th>
<th>1750</th>
<th>1800</th>
<th>1850</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amsterdam</td>
<td>Netherlands</td>
<td>200</td>
<td>210</td>
<td>217</td>
<td>225</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Netherlands</td>
<td>51</td>
<td>44</td>
<td>53</td>
<td>90</td>
</tr>
<tr>
<td>'S Gravenhague</td>
<td>Netherlands</td>
<td>30</td>
<td>38</td>
<td>39</td>
<td>72</td>
</tr>
<tr>
<td>Utrecht</td>
<td>Netherlands</td>
<td>30</td>
<td>25</td>
<td>32</td>
<td>48</td>
</tr>
<tr>
<td>Madrid</td>
<td>Spain</td>
<td>140</td>
<td>160</td>
<td>168</td>
<td>281</td>
</tr>
<tr>
<td>Barcelona</td>
<td>Spain</td>
<td>34</td>
<td>50</td>
<td>100</td>
<td>220</td>
</tr>
<tr>
<td>Sevilla</td>
<td>Spain</td>
<td>72</td>
<td>66</td>
<td>96</td>
<td>113</td>
</tr>
<tr>
<td>Malaga</td>
<td>Spain</td>
<td>30</td>
<td>36</td>
<td>49</td>
<td>93</td>
</tr>
<tr>
<td>Berlin</td>
<td>Germany</td>
<td>55</td>
<td>113</td>
<td>172</td>
<td>437</td>
</tr>
<tr>
<td>Hamburg</td>
<td>Germany</td>
<td>70</td>
<td>90</td>
<td>130</td>
<td>149</td>
</tr>
<tr>
<td>Konigsberg</td>
<td>Germany</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breslau</td>
<td>Germany</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lisbon</td>
<td>Portugal</td>
<td>180</td>
<td>185</td>
<td>195</td>
<td>240</td>
</tr>
<tr>
<td>Porto</td>
<td>Portugal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milan</td>
<td>Italy</td>
<td>125</td>
<td>124</td>
<td>135</td>
<td>209</td>
</tr>
<tr>
<td>Naples</td>
<td>Italy</td>
<td>300</td>
<td>339</td>
<td>430</td>
<td>409</td>
</tr>
<tr>
<td>Rome</td>
<td>Italy</td>
<td>135</td>
<td>158</td>
<td>153</td>
<td>175</td>
</tr>
<tr>
<td>Palermo</td>
<td>Italy</td>
<td>100</td>
<td>124</td>
<td>139</td>
<td>168</td>
</tr>
<tr>
<td>Paris</td>
<td>France</td>
<td>500</td>
<td>570</td>
<td>550</td>
<td>1 053</td>
</tr>
<tr>
<td>Marseille</td>
<td>France</td>
<td>90</td>
<td>68</td>
<td>101</td>
<td>195</td>
</tr>
<tr>
<td>Lyon</td>
<td>France</td>
<td>97</td>
<td>114</td>
<td>109</td>
<td>177</td>
</tr>
<tr>
<td>Bordeaux</td>
<td>France</td>
<td>45</td>
<td>62</td>
<td>96</td>
<td>130</td>
</tr>
<tr>
<td>Vienna</td>
<td>Austria</td>
<td>114</td>
<td>175</td>
<td>247</td>
<td>431</td>
</tr>
<tr>
<td>Graz</td>
<td>Austria</td>
<td>22</td>
<td>20</td>
<td>31</td>
<td>55</td>
</tr>
<tr>
<td>Linz</td>
<td>Austria</td>
<td>10</td>
<td>10</td>
<td>17</td>
<td>27</td>
</tr>
<tr>
<td>Salzburg</td>
<td>Austria</td>
<td>13</td>
<td>15</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>London</td>
<td>UK (England)</td>
<td>575</td>
<td>675</td>
<td>948</td>
<td>2 236</td>
</tr>
<tr>
<td>Liverpool</td>
<td>UK (England)</td>
<td>6</td>
<td>22</td>
<td>83</td>
<td>376</td>
</tr>
<tr>
<td>Manchester</td>
<td>UK (England)</td>
<td>8</td>
<td>18</td>
<td>84</td>
<td>303</td>
</tr>
<tr>
<td>Birmingham</td>
<td>UK (England)</td>
<td>7</td>
<td>24</td>
<td>71</td>
<td>233</td>
</tr>
<tr>
<td>Dublin</td>
<td>UK (Ireland)</td>
<td>60</td>
<td>129</td>
<td>200</td>
<td>262</td>
</tr>
<tr>
<td>Belfast</td>
<td>UK (Ireland)</td>
<td>2</td>
<td>9</td>
<td>20</td>
<td>87</td>
</tr>
<tr>
<td>Cork</td>
<td>UK (Ireland)</td>
<td>25</td>
<td>58</td>
<td>75</td>
<td>85</td>
</tr>
<tr>
<td>Limerick</td>
<td>UK (Ireland)</td>
<td>11</td>
<td>16</td>
<td>39</td>
<td>53</td>
</tr>
<tr>
<td>Glasgow</td>
<td>UK (Scotland)</td>
<td>13</td>
<td>25</td>
<td>70</td>
<td>345</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>UK (Scotland)</td>
<td>36</td>
<td>57</td>
<td>83</td>
<td>194</td>
</tr>
<tr>
<td>Dundee</td>
<td>UK (Scotland)</td>
<td>10</td>
<td>12</td>
<td>26</td>
<td>79</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>UK (Scotland)</td>
<td>12</td>
<td>15</td>
<td>27</td>
<td>72</td>
</tr>
</tbody>
</table>

Note: Includes all western European cities with population >200 000 in 1850.
Source: Bairoch et al. (1988).
Agricultural labour productivity

Urbanization was accompanied by high agricultural output per worker. By the beginning of the 18th century, agricultural yields were higher in Western Europe, and particularly England, than elsewhere in the world. Brunt and Fidalgo (2009) have collected new data on agricultural labour productivity that show massive differentials in output per agricultural worker around the world in 1700, 1775, 1845 and 1870. They are reported in Table 3. Brunt and Fidalgo use geographical units of similar sizes, so the data are mostly for countries but by province for India and China (for which we report the most and least productive regions only). The sizes of the differentials may seem surprising but note that very similar magnitudes are reported in Maddison (2000). We prefer the data reported here because they refer explicitly to agriculture, whereas Maddison constructs data for the whole economy of each country. However, since agriculture was by far the largest sector in each economy throughout this period, the two data sets inevitably turn out to be quite similar. 15

Table 3. Real output per worker in agriculture (England in 1870=100)

<table>
<thead>
<tr>
<th></th>
<th>1705</th>
<th>1775</th>
<th>1845</th>
<th>1870</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>59</td>
<td>84</td>
<td>71</td>
<td>100</td>
</tr>
<tr>
<td>Scotland</td>
<td>NA</td>
<td>57</td>
<td>109</td>
<td>157</td>
</tr>
<tr>
<td>Netherlands</td>
<td>NA</td>
<td>45</td>
<td>48</td>
<td>36</td>
</tr>
<tr>
<td>France</td>
<td>22</td>
<td>15</td>
<td>19</td>
<td>27</td>
</tr>
<tr>
<td>Prussia</td>
<td>1</td>
<td>2</td>
<td>23</td>
<td>50</td>
</tr>
<tr>
<td>USA</td>
<td>17</td>
<td>26</td>
<td>58</td>
<td>73</td>
</tr>
<tr>
<td>China (Hupei, most productive)</td>
<td>61</td>
<td>48</td>
<td>39</td>
<td>44</td>
</tr>
<tr>
<td>China (Kweichow, least productive)</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>India (Bengal, most productive)</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>India (Punjab, least productive)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

Source: Brunt and Fidalgo (2009).

Table 3 highlights not only the high level of output per worker in England but also its rapid growth during the 18th century; with an increase of 42 per cent over the period 1705-1775. Alternative estimates for output per worker in agriculture in England yield similar growth rates. For

15 Another reason to prefer the Brunt-Fidalgo data is that they report the allocation of labour between agriculture and other sectors. Hence their agricultural labour productivity data are necessarily consistent with their measure of structural change towards non-agricultural production. This is useful because, except for England, there are basically no data on the size of the manufacturing workforce in each country.
example Clark (2002b; table 4) finds an increase in output per worker of 44 per cent between 1700-09 and 1770-79. Moreover, Clark documents that output per worker already rose sharply in the 16th and 17th century, increasing fivefold between 1500 and 1650.

Agricultural taxes in early modern times

The data we have just discussed raises the question of what caused the increase in agricultural labour productivity that we observe in England and Scotland. The increase was partly due to improvements in agricultural productivity, as documented by a large literature. Productivity growth was nevertheless small. According to Clark (2002b, table 5), agricultural TFP increased by 55 percent between 1550 and 1800, amounting to an average annual growth rate of 0.3%. An alternative candidate explanation is the increase in extractions from English farmers levied by landlords and the government. In table 4 we report the tax burden on agricultural workers in England and China in c. 1775, which was almost 100 times greater in England.16 Since output per worker was around ten times higher in England than in the typical Chinese province, the tax rate was around ten times higher in England than China. How is this possible? In a Malthusian world the answer is straightforward. In equilibrium, output per worker at the margin will be equal to subsistence (by assumption). Now impose a tax equal to 50 per cent of output. Then the output of the marginal worker must double, so that he can pay the tax and still continue to subsist. Workers who are unable to meet this level of output will die (or never be born). So high-tax countries – such as England – will be characterized by low agricultural population densities, whereas low-tax countries – such as China – will be characterized by high agricultural population densities. This is confirmed in the last line of table 4, which indicates that average acreage per worker was 20 in England and only 3 in China.

16 In fact, land was mostly owner-occupied in China, so there was virtually no extraction through land rents.
Table 4. Extractions per adult male agricultural worker in c. 1775 (English d).

<table>
<thead>
<tr>
<th>Extraction</th>
<th>England</th>
<th>China (average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land rent</td>
<td>3389</td>
<td>0</td>
</tr>
<tr>
<td>Tithe</td>
<td>471</td>
<td>0</td>
</tr>
<tr>
<td>Poor rate</td>
<td>272</td>
<td>0</td>
</tr>
<tr>
<td>Land tax</td>
<td>354</td>
<td>47</td>
</tr>
<tr>
<td>TOTAL</td>
<td>4487</td>
<td>47</td>
</tr>
</tbody>
</table>

Acres per ag. worker 20 3

Source: Brunt and Fidalgo (2010).

Of course, there could be other factors driving the apparent relationship between taxes, land area and output per worker. Social, political, military and economic institutions were very different in England and China. To control for some of the possible confounding factors, it would be interesting to see how taxes were related to agricultural population density, and ideally we would like to use time-series data to examine the impact of extractions from agriculture. Unfortunately, neither time-series nor cross-country comparable data exist for the period. However, Brunt and Fidalgo (2010) have data for Chinese provinces, where, as we saw in Table 3, there was substantial variability in agricultural productivity per worker. Moreover, taxes varied across Chinese provinces allowing us to run a regression of population per acre of arable land on the region’s tax burden. The results are reported in table 5 below. We partially control for the possibility of endogeneity by regressing the population density in 1775 on tax rates in 1753, and find that higher taxes in a province reduced the population density. This is consistent with the idea that higher taxes in agriculture induce the flow of labour away from the country-side and into cities, although it could also be due to higher taxes leading to a lower population.
Table 5. Population densities and per capita tax burdens across Chinese provinces.

<table>
<thead>
<tr>
<th>Dependent variable: Population density in 1775</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.00*</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
</tr>
<tr>
<td>Per capita tax burden in 1753</td>
<td>-1.18**</td>
</tr>
<tr>
<td></td>
<td>(0.33)</td>
</tr>
<tr>
<td>r-squared</td>
<td>43.94</td>
</tr>
<tr>
<td>N</td>
<td>18</td>
</tr>
</tbody>
</table>

Notes and sources. All variables are in natural logarithms. Dependent variable is the population per acre in 1775, drawn from Brunt and Fidalgo (2010); independent variable is the tax burden per capita in 1753, drawn from Wang, *Land taxation*, table 4.1. ** denotes statistical significance at the one percent level and * denotes statistical significance at the five percent level.

As we have just established, the level of taxation in England by the late 18th century was particularly high. This was the result of a sustained increase in extraction from agriculture over the previous 250 years. Figure 4 presents the evolution of an index for real wages and real taxes plus rents. In 1550, taxes and rents paid by farmers represented only 7 percent of total output; by 1600 their share had increased to 19 percent. They fluctuated between 17 and 19 percent during the 17th century and then increased steadily over the next hundred years reaching 25 percent by 1800.

**Figure 4: Wages and rents plus taxes in agriculture, England (1775 English d).**

![Figure 4: Wages and rents plus taxes in agriculture, England (1775 English d).](image)

Source: Authors’ calculations from Clark (2002b) and Brunt (2000); see text.

17 Clark (2002b; table 1) provides an index of real wages and of the real value of taxes and rents. In order to get tax rates, we have used data from Brunt (2000) on the value of output per worker and rents and taxes for 1775. Output is 975d/acre, which with 20 acres per worker yields an output per worker of 19 500d. This generates an extraction rate from gross output of 23 percent (=4 487/19 500).
To sum up, during the 17th and 18th century extractions out of agriculture increased more than threefold in England. Prior to the Industrial Revolution, England also exhibited high and growing urbanization, while agriculture exhibited both high acreage per worker and high and growing labour productivity. These stylized facts are consistent with our model’s prediction that the high extraction rate forced farmers to leave the countryside and was hence responsible for the changes in urbanization and agricultural productivity.

5. Numerical examples

In order to illustrate our arguments, we provide a numerical example of the dynamics of the model. The parameter values we use are given in Table 7. Each period is assumed to last 20 years. We suppose that the long-run share of agricultural goods in expenditure is 0.5, while one per cent of the population are supposed to be landlords. The productivity of labour in agriculture is \( \gamma = 0.6 \) and the land/labour ratio is assumed to stay constant at 20 acres per capita. Agricultural TFP is set at \( A = 0.136 \) in the first period and grows to reach \( A = 0.210 \) at the end, implying TFP growth of 55% over the period 1540 to 1800, consistent with the figures reported by Clark (2002b). The minimum food requirement is set at \( c = 0.1345 \), while \( d \) takes the value 0.9, both value chosen for the model to match the initial urbanization rate. These figures imply that initially the share of the population needed to produce the minimum food requirement is 45 per cent, and this falls to 60 per cent as agricultural TFP increases. It is difficult to find evidence to pin down the parameters of the imitation and innovation functions. We set \( h = 1 \) and \( \rho = 2 \). The probability of imitating in a particular meeting is 25 per cent while the probability of innovating at a particular meeting is 70 per cent. The initial value of the tax rate is 0.07 per cent, which yields a rate of urbanization of 3.8, the one observed in the data for the mid 1500s.

<table>
<thead>
<tr>
<th>Table 7: Parameter values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
</tr>
<tr>
<td>Tax rates</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
We examine the effect of various shocks, and all figures present the evolution of urbanization in the model as well as in the data. The urbanization data is from De Vries (1984), and is the fraction of the population living in cities of at least 10,000 people. The data are available every 50 years, starting in 1550 when the rate of urbanization was 3.8%. Urbanization then increases steadily, reaching 9.5% in 1650, 14% in 1700 and 22% by 1800.

Figure 5 reports the evolution of urbanization in the model in the absence of any shock, as well as that caused by agricultural TFP growth. In the absence of shocks, innovation and learning cause some increase in urbanization but this is moderate and urbanization converges to a value of around 9%. Low agricultural output implies high prices for food and thus a low return to manufacturing, resulting in low urbanization. Note that agricultural employment is well above the fraction of the population needed to produce the minimum food requirement, a result due to high food prices. The shock to agricultural TFP reduces the relative price of food and induces an increase in urbanization. Although rising urbanization is sustained throughout the period, the level is well below that observed in the data.

Figure 5: Urbanization – no shock and agricultural TFP growth

![Urbanization chart](image-url)
Figure 6 compares the effect of agricultural TFP growth with that of tax changes. We shock the economy by increasing the tax rate in line with the data, using the figures reported in Table 7. There was a small increase (from 0.07 to 0.09 per cent) early in the period and then the extraction rate rose sharply around 1600, reaching 0.19 per cent. It fluctuated around 20 per cent during the 17th century and rose again in the 18th century, reaching 0.25 per cent by 1800. The dashed line indicates that the model is successful at reproducing the early acceleration in urbanization that we observe, although from 1650 onwards the increase in urbanization is below the data. Nevertheless, tax changes generate substantially faster urbanization than agricultural TFP changes. The continuous line presents the combined effect of the two shocks, and the simulated series fits the data closely. The improvement when compared with the results obtained by having only the agricultural shock is substantial. Moreover, note that although the bulk of the tax increase occurs around 1600, urbanization keeps rising even though tax rates change little after that period. The reason for this is that the high level of taxation implies that moderate changes in agricultural TFP or in manufacturing productivity induce substantial flows of labour from rural to urban areas.
In order to understand what drives the behaviour of urbanization, figure 7 presents our core simulation, with changes in both taxes and agricultural TFP, for an economy with imitation (depicted by the continuous line) and one in which imitation is not possible (that is, $z=0$, dashed line). We can see that the increase in taxation that took place between the 1550s and 1600 resulted in a moderate increase in urbanization, but the subsequent behaviour of taxes is incapable of reproducing the sharp increase in the urban population that followed, indicating that although higher taxes could have been the trigger of increased urbanization, the creation and diffusion of ideas played a crucial role in the further increases in the urban population. The reason for this is that although the number of available ideas increases at the same rate in both cases, average productivity in manufacturing is not growing fast enough to create a sufficiently large flow of labour from agriculture into industry. As a result, by the end of the 18th century the model without imitation predicts an urbanization rate of 9.33 percent, rather than the 22.5 percent obtained in the benchmark case (the actual urbanization rate was 21.7).

**Figure 7: Urbanization**

*Simultaneous tax and agricultural TFP changes with and without imitation*
Figure 8 presents the simulated time series for real per capita output, as well as data on British per capita output from Broadberry et al. (2009, tables 19 and 22). It is interesting to see that although urbanization starts increasing rapidly in the 17th century, output growth during that period is slow and starts to accelerate from 1700 onwards. The dashed line depicts output per worker, and we can see a slow take-off followed by acceleration as the diffusion of ideas increases. However, this acceleration is substantially faster than that observed in the data. One difference between the two series is that the data refers to output per capita. Since this period witness a rapid population expansion, the continuous line depicts the series of output per capita under the assumption of an endogenous rate of growth of population, where the growth rate in increasing in the previous period’s per capita income. The resulting growth rate averages 0.37% per year over the period, while average growth for per worker output is 0.62% (and 0.25% for population).

![Figure 8: Output – data and simulation results](image)

To understand the acceleration of growth, Figure 9 presents the corresponding time series for the number of available ideas or knowledge (dotted line), the average number of ideas in the urban population (dashed line), which is equivalent to average productivity in manufacturing, and per

---

18 See the appendix (to be written) for the details.
capita output (continuous line). The first thing to note is the gap between knowledge and productivity, capturing the fact that available knowledge and the diffusion of innovations do not move together. Second, although the growth rate of knowledge is constant (one idea per period) the model delivers an acceleration of productivity. The reason for this is that, initially, low urban density implies a slow diffusion of innovations amongst city-dwellers. As urbanization increases, the rate of diffusion or knowledge rises resulting in an acceleration of productivity growth from 1750 onwards. Output growth follows this pattern too. Lastly, output growth remains moderate despite the fact that available knowledge in increasing fast, thus making a fast pace of innovation compatible with moderate per capita output and productivity growth.

**Figure 9: Simulated output, productivity and knowledge**

It is important to comment at this point on our choice of model. The key idea in our analysis is that cities are special because they allow for interactions amongst individuals which in turn permit the diffusion and creation of knowledge. We could have model this effect simply by assuming that productivity in manufacturing is an increasing function of the rate of urbanization. Note, however, that unless we had made strong assumptions about the functional form of this knowledge-generating function, it is unlikely that we would have obtained an acceleration of productivity growth. In
contrast, our micro-founded model of interactions between agents, based on the idea that urbanization determines the number of encounter and that the probability of imitation or innovation in each of this is constant, has the implication that there is such an acceleration. The intuition is simple. For a given number of encounters, the probability of imitating $i$ ideas in a particular encounter is $zP(i,t)$; if the fraction of individuals with $i$ ideas at $t$ is larger, this will increase the fraction that has $i$ ideas at $t+1$, making the probability of imitation next period higher. This mechanism creates an acceleration of the transmission of knowledge that results in slow initial growth that increases over time.

6. Conclusions

This paper has presented a theory of industrialization based on the key role played by cities in creating and diffusing knowledge. We argue that the expansion of states and the resulting increase in taxation of agricultural output created a flow of labour out of farming and into cities. Cities are special in that they generate frequent encounters amongst individuals, resulting in an exchange of knowledge that leads to both imitation and innovation. As urban density increases, interactions amongst individuals rise, and this leads to the diffusion of ideas which, on the one hand increase the manufacturing skills of those who imitate, and, on the other, allow individuals to innovate on the basis of those ideas. Higher urban density then increases productivity in manufacturing, further inducing workers to leave farming and move into the city, which raises innovation and imitation even more. As a result growth takes off.

The key features of our model—high urbanization and agricultural labour productivity prior to a take off—fit well with the evidence for mid-18th century England. We thus maintain that urbanization was the cause, and not the consequence, of the unprecedented spur of innovations that started then and there and gave rise to the First Industrial Revolution. Our hypothesis raises the question of why the English population was so concentrated in cities. Although there were certainly several causes, we maintain that a key element was the sharp increase in the extraction rate out of agriculture that took place during the 17th century. Rents and taxes paid by farmers rose sharply during that century, leading to a flow of labour out of agriculture that resulted in both the observed high levels of output per worker in agriculture and high urbanization. The evidence we discuss, although limited, lends support to this hypothesis, and our numerical examples show that the model generates dynamics that are consistent with the data.

A key aspect of our analysis is that we see innovation as the result not of market activities but of social interactions. Social interactions are in fact often called ‘non-market interactions’ to
emphasize the fact that in this context the actions of agents are not determined by the price mechanism. In our model, growth is endogenous but, in contrast with much of the recent literature, we suppose that innovation is not motivated by a profit motive but rather occurs as an externality resulting from high urbanization. One of the questions raised by our analysis is what the role of urbanization is when innovation is intentional and to what extent the mechanism explored by the literature on economic geography for the 20th century apply in early modern times.  

19 See, for example, Scheinkman (2008) and Cabrales and Zenou (2010).
References


