Speed of Convergence in a Malthusian World: Weak or Strong *Homeostasis*?*

Arnaud Deseau†

November 23, 2023

Abstract

The Malthusian trap is a well recognized source of stagnation in per capita income prior to industrialization. However, previous studies have found mixed evidence about its exact strength. This article contributes to this ongoing debate, by estimating the speed of convergence for a wide range of economies and a large part of the Malthusian era. I build a simple Malthusian growth model and derive the speed of convergence to the steady state. A calibration exercise for the English Malthusian economy reveals a relatively weak Malthusian trap, or weak *homeostasis*, with a half-life of 112 years. I then use β-convergence regressions and historical panel data on per capita income and population to empirically estimate the speed of convergence for a large set of countries. I find consistent evidence of weak *homeostasis*, with the mode of half-lives around 120 years. The weak *homeostasis* pattern is stable from the 11th to the 18th century. However, I highlight significant differences in the strength of the Malthusian trap, with some economies converging significantly faster or slower than others.

Keywords: Convergence, Homeostasis, Malthusian trap, Preventive checks, Marriage, Fertility, Malthusian model, Beta-convergence

JEL Codes: J1, N1, N3, O1, O47

---

*I am grateful to Hugues Annoye, Thomas Baudin, Pierre de Callataÿ, Greg Clark, David de la Croix, Frédéric Docquier, Oded Galor, Marc Goñi, Andreas Irmen, Nils-Petter Lagerlöf, Hélène Latzer, Anastasia Litina, Jakob Madsen, Fabio Mariani, Stelios Michalopoulos, Luca Pensieroso and David Weil for their valuable comments and suggestions. I also thank participants at the 2018 University of Luxembourg CREA Workshop on Culture and Comparative Development; the 2017 Université Saint-Louis – Bruxelles CEREC Workshop on Macroeconomics and Growth; and seminars at Brown University and the Catholic University of Louvain.

†AMSE, Aix-Marseille Université, 5-9 Boulevard Maurice Bourdet, 13001 Marseille, France (e-mail: arnaud.deseau@univ-amu.fr).

In four centuries [1300-1700], the [French] population only increased by 2 million persons in all! And some say less! [...] Thus, an extraordinary ecological equilibrium is revealed. Of course, it did not exclude possibly prodigious, but always temporary, upheavals and negative fluctuations in its time like those experienced by animal population.

Emmanuel Le Roy Ladurie (1977), Motionless History.

1 Introduction

One of the most central prediction of the Malthusian theory is that standards of living were *stagnant* before the onset of industrialization. Stagnation however does not literally mean *constant*, or flat, per capita income. In fact, any shock striking a Malthusian economy generates fluctuations in the standards of living, namely *temporary* or *non*-sustained economic growth. Indeed, a simple Malthusian model predicts that a positive shock on the technology level – say the introduction of better cultivation techniques – increases income per capita in the short run only; in the long run, population increases and the economy returns to its initial level of income per capita. This is the so-called "Malthusian trap" mechanism, that has been recognized as one of the major obstacles to achieve sustained economic growth during millennia (Kremer, 1993; Galor and Weil, 2000; Hansen and Prescott, 2002; Clark, 2007; Ashraf and Galor, 2011; Galor, 2011).

While the existence of the Malthusian trap is widely established empirically, previous literature has found mixed evidence about its exact strength. A first group of studies, mainly focusing on England, finds evidence of a weak Malthusian trap, known as weak *homeostasis*[1] (Lee, 1993; Lee and Anderson, 2002; Crafts and Mills, 2009; Fernihough, 2013; Bouscasse et al., 2023). In these studies, the half-life of adjustment to shocks is typically about one century, and can be as long as four centuries. On the other hand, Madsen et al. (2019) find the first evidence of a

[1] *Homeostasis* comes from the Greek *homoios* "similar" and *stasis* "steady", meaning "staying the same". In demography, it refers to a population equilibrium maintained by density-dependent checks (Lee, 1987).

strong and widespread Malthusian trap, or strong *homeostasis*, with estimated half-lives between one and three decades.

In this article, I reinvestigate the question of the strength of the Malthusian trap by examining the speed of convergence of Malthusian economies – i.e. how quickly they tend to return to their steady state following a shock. I argue that the speed of convergence in Malthusian times should be relatively *slow*, and thus reflects weak *homeostasis*. The main reason is that the Malthusian trap involves demographic fluctuations which, by definition, are long and take generations to unfold. As argued by Malthus (1798) himself, the channels through which population adjusts to the amount of resources per capita are the age at marriage, fertility and mortality (the so-called *preventive* and *positive* checks). This is confirmed by empirical studies finding a significant but small response of demographic variables to changes in per capita incomes in pre-industrial times.[2]

To investigate this conjecture, I first build an overlapping-generations Malthusian growth model including both preventive and positive checks as means of population adjustment. In particular, agents first choose to marry (or not), influencing the extensive margin of fertility, and then choose the number of children within marriage, influencing the intensive margin of fertility. Both choices depend on income per capita, in a Malthusian fashion. I show that the speed of convergence of a Malthusian economy to its steady state depends on four parameters: the land share of output and the elasticities of fertility, marriage and survival with respect to income per capita. I calibrate the model for England and show that, under plausible parameter values, the speed of convergence indicates weak *homeostasis*, with a half-life of 112 years. Alternative calibration scenarios using the 10th percentile and 90th percentile of the long-run elasticities estimated in the literature for England also indicate weak *homeostasis*, with half-lives between 64 and 230 years. I also provide a quantitative analysis, showing that weak *homeostasis* is consistent with the centuries long reaction of the English Malthusian economy to the Black

[2] Empirical estimates of these elasticities can be found in Lagerlöf (2015) and Klemp and Møller (2016), among others. See Section 3.1 and Section B of the Appendix for further details.

Death.

Second, I employ β-convergence regressions *à la* Barro and Sala-i Martin (1992) to provide empirical estimates of the speed of convergence for a large set of economies and a large part of the Malthusian period. I first use the data compiled in the latest update of the Maddison Project, which offers the most comprehensive GDP per capita data available to study Malthusian economies. I also run the same regressions using historical population levels from McEvedy et al. (1978). To gain in precision, and explore the temporal and spatial heterogeneity of the speed of convergence in a more comprehensive way than previous studies, I employ two additional datasets that possess a much higher cross-sectional and time dimension than the two aforementioned sources. The first one, coming from Lagerlöf (2019), gives simulated GDP per capita series based on the same empirical moments as the Maddison Project data. The second dataset, coming from Reba et al. (2016), compiles the historical urban population series originally produced by Chandler (1987) and Modelski (2003). In all cases, I find consistent evidence of a relatively weak *homeostasis*, with the mode of the estimated half-lives around 120 years. I find evidence of a stable pattern of weak *homeostasis* throughout much of the Malthusian era, from the 11th century to the end of the 18th century. On the other hand, I find significant differences in the speed of convergence between countries, with some Malthusian economies converging significantly slower or faster than others. In particular, some economies are found compatible with strong *homeostasis*, in the same magnitudes as found by Madsen et al. (2019).

There is one main concern in my empirical analysis: weak Malthusian dynamics may be the result of a serious omitted variable bias in the β-convergence regressions. I employ several strategies to mitigate that concern. My empirical analysis includes country and time fixed effects, which respectively account for unobserved time-invariant characteristics at the country level (e.g. geography), and common trends (e.g. technology diffusion) that could simultaneously affect growth and initial development levels. In addition, I also include a time-varying control variable (*Statehist*), developed by Borcan et al. (2018), capturing state presence during the Malthusian period. In principle, this variable captures general institutional changes likely to affect the steady-state position of Malthusian economies, thus reducing the omitted variable bias. Finally, to address remaining endogeneity concerns, I employ an instrumental variable approach (GMM), which uses the lagged values of the endogenous regressors as instruments. Typically, I find that GMM estimates confirm the weak *homeostasis* pattern of my fixed-effects regressions.

This article contributes to the growing literature examining the existence and strength of the Malthusian trap (Lee and Anderson, 2002; Nicolini, 2007; Crafts and Mills, 2009; Kelly and Gráda, 2012; Fernihough, 2013; Møller and Sharp, 2014; Lagerlöf, 2015; Madsen et al., 2019; Cummins, 2020; Jensen et al., 2021; Attar, 2023). Typically, the literature finds evidence of the existence of Malthusian dynamics in a particular country, albeit small in magnitude. For instance, Crafts and Mills (2009) study Malthusian dynamics in England (1540-1870) using structural modelling, and conclude that there is a very weak Malthusian trap. Similarly, Fernihough (2013) finds evidence of weak *homeostasis* in Northern Italy (1650-1881), using VAR methods. I contribute to this literature in two main respects. First, rather than focusing on a specific country, this article is the first, to my knowledge, to provide evidence of weak *homeostasis* across a wide range of Malthusian economies and for a large part of the Malthusian era. Second, I am able to characterize, for the first time, the full distribution of convergence speed during the Malthusian period. I show that most countries are characterized by weak *homeostasis* of around a century, while highlighting significantly stronger or weaker Malthusian traps for some countries. In particular, I find that the Spanish Malthusian trap is close to strong *homeostasis*, with a half-life of less than 50 years; whereas in other countries, such as England, the half-life is of the order of a century or more. The article closest to mine is Madsen et al. (2019), which find evidence of widespread strong *homeostasis* in a panel of 17 countries (900-1870). The main difference between the two articles lies in the approach to the data and the estimation method. Whereas Madsen et al. (2019) rely on data interpolated from heterogeneous historical sources and use a SUR model, I employ the data as they appear in their original sources and use the standard techniques developed in the empirical growth literature to estimate the speed of convergence, such as fixed-effects models and GMM.

This article also adds to the literature studying Malthusian dynamics in an overlapping-generations frameworks. The existing overlapping-generations Malthusian frameworks consider the intensive margin of fertility as the only channel through which population adjusts (Ashraf and Galor, 2011; Lagerlöf, 2019). I build on these previous models by incorporating, for the first time, marriage as an explicit channel through which the population adjusts, as originally argued by Malthus (1798). The marriage channel allows me to incorporate the extensive margin of fertility, since unmarried people typically had no children in the Malthusian era, and therefore allows me to model richer population dynamics.

Finally, this article relates to the literature deriving the speed of convergence in growth models. Working in continuous time, Irmen (2004) and Szulga (2012) find that the speed of convergence of a Malthusian economy depends on the land share of output and the elasticities of the birth rate and death rate to income per capita. I contribute to this literature by showing that the elasticity of the marriage rate to income per capita also matters to characterize the speed of convergence. In a modern context, this article relates also to the seminal work of Barro (1991) and Barro and Sala-i Martin (1992).

The rest of this article is organized as follows. Section 2 presents my Malthusian growth model. Section 3 presents my calibration exercise, discussing the parameters I use and presenting my simulations. Section 4 derives the speed of convergence implied by my model and discuss it in relation to the literature. Section 5 describes my empirical strategy and the data I use to estimate the speed of convergence. Section 6 presents and discusses my empirical results. Section 7 concludes.

2 Theoretical Framework

In this section, I present the core elements of the Malthusian model I use to study the dynamics of GDP per capita and population in the Malthusian era. I consider an overlapping-generations economy with time modelled as discrete and going from zero to infinity, and where agents live two periods. In the first period of their life, they are inactive children entirely supported by their parents; they make no decisions. In the second period of their life, they work, earn an income and make decisions about consumption, marriage and fertility.

I deviate from textbook Malthusian models by modelling explicitly marriage, celibacy and childlessness decisions. In brief, that means that I am considering both the *extensive* margin of fertility, i.e. whether or not an individual marries and can have children, and the *intensive* margin of fertility, i.e. variations in individual's number of surviving children within marriage. These two elements are crucial as they directly affect the response of fertility to income per capita, and therefore the speed with which a Malthusian economy returns to its steady state after a shock. Both are consistent with empirical studies showing the importance of the so called *preventive* checks, advocated by Malthus (1798) himself, in affecting fertility. Indeed, Cinnirella et al. (2017) show that real wages affect negatively birth spacing within marriage and the time of marriage and first child in England for the period 1540-1850. Cummins (2020) finds similar results with a negative effect of living standards on the age at first marriage in France between 1650 and 1820. de la Croix et al. (2019) show that singleness and childlessness are key elements to take into account when estimating reproductive success in pre-industrial times. Therefore, modelling both the *extensive* and *intensive* margins of fertility appears crucial to a rigorous analysis of population dynamics during the Malthusian era.

I model childlessness and celibacy together, leaving the possibility to procreate only to married agents. This is fully consistent with historical studies showing very low illegitimate birth rates in pre-industrial Europe (Hajnal, 1965; Segalen and Fine, 1988; Wrigley et al., 1989). Marriage offers the opportunity for agents to gain utility from another source than just pure consumption.[3] On the other hand, the disutility of marriage is represented by a search cost that agents need to pay in order to match with a partner.[4] Agents are assumed to be heterogeneous in their search cost, which is exogenously given. At the beginning of their adult life, agents draw a search cost λ_i with $\lambda_i \sim \mathcal{U}(1, b)$ and b being the maximum of the uniform distribution. Agents maximize their utility and therefore a marriage occurs only if the utility of being married is superior to the utility of being single. Within marriage, I let the agent's fertility depend on his income per capita, according to the standard Malthusian theory and empirical evidence (Cinnirella et al., 2017; de la Croix et al., 2019; Cummins, 2020).

Preferences and Budget Constraints.— The utility of a married agent i of generation t is defined *à la* Baudin et al. (2015):

$$U_{i,t}^{M} = \ln c_t + \gamma \ln \left(n_t + \nu\right) - \ln \lambda_i \ , \tag{1}$$

where c_t denotes consumption, $\gamma > 0$ is a child preference parameter, n_t is the number of surviving children, $\nu > 0$ allows for childlessness as the individual utility remains defined when $n_t = 0$, and λ_i is the utility cost of marriage.

It follows that the utility of an unmarried agent of generation t is given by:

$$U_{i,t}^{S} = \ln c_t + \gamma \ln \left(\nu\right) \ . \tag{2}$$

Agents allocate their income between consumption and child rearing such that we have the following budget constraint:

$$c_t = y_t - f(n_t) \ , \tag{3}$$

[3]This means that parents only care about the quantity of surviving children, as in a standard Malthusian model.
[4]Alternatively, one can think the cost as representing a dowry that agents need to pay in order to marry.

where y_t is agent's income, and $f(n_t)$ is the cost of having n_t children in terms of goods.

A convenient functional form for $f(\cdot)$ capturing both the idea of childlessness ($f(0) = 0$) and allowing for different types of returns to scale in the production of children is the following one:

$$f(n_t) = q(n_t + \nu)^{1/\delta} - q\,\nu^{1/\delta} \,, \tag{4}$$

with $q > 0$ being the unit cost of a child, and $\delta > 0$ a parameter influencing the degree of return to scale in child production.

Fertility.— Maximizing (1) subject to (3), I obtain the optimal fertility behaviour of a married agent of generation t:

$$n_t = \kappa \cdot \left(y_t + q\,\nu^{1/\delta}\right)^{\delta} \; - \nu \;\equiv n_t(y_t), \tag{5}$$

where $\kappa = \left(\frac{q}{\gamma\delta} + q\right)^{-\delta}$. Thus, in accordance with Malthusian theory, the number of surviving children within marriage depends positively on income per capita ($\partial n_t / \partial y_t > 0$).

Marriage.— An agent is indifferent between being married and single if utility is the same in both situations. I define $\overline{\lambda}$ as the draw from the search cost distribution that makes an agent indifferent between being married and single. The condition for an agent to be married is: $\lambda_i < \overline{\lambda}$ with $\lambda_i \sim \mathcal{U}(1, b)$. I can therefore compute the probability for an agent of generation t to be married as:

$$p_t = P(\lambda_i < \overline{\lambda}) = \frac{\overline{\lambda}(y_t) - 1}{b - 1} \equiv p_t(y_t) \,, \tag{6}$$

where b is the maximum of a uniform distribution and the threshold draw $\overline{\lambda}$ depends on an individual's income.[5] Since I work at the generation level, p_t is also equivalent to the marriage rate in that Malthusian economy. In the rest of the article, I will use p_t as the marriage rate.

[5]The full expression of $\overline{\lambda}$ is available in Section A of the Appendix

Thus, in line with the idea of Malthus (1798), an increase in income lowers the age of marriage, resulting in a higher marriage rate at the generation level in the model ($\partial p_t / \partial y_t > 0$).

Production.— Total output in period t is given by:

$$Y_t = (A_t T)^{\alpha} L_t^{1-\alpha} , \tag{7}$$

where A_t is a land-augmenting technology factor, T is total land area, L_t is the size of the labour force that is equivalent to the adult population in my analysis and $\alpha \in (0, 1)$ is the land share of output.

I assume that workers are self-employed and earn an income equal to the output per worker in t. Using (7) and normalizing land area to unity ($T = 1$), we obtain:

$$y_t = \left(\frac{A_t}{L_t} \right)^{\alpha} . \tag{8}$$

Following Lagerlöf (2019), I consider sustained but constant growth in land productivity. The technological level in period t is given by:

$$A_t = A_0 (1 + g)^t , \tag{9}$$

where A_0 is the initial technological level and g is an exogenously given and constant rate of technological progress.

Mortality.— Malthus (1798) and the Malthusian theory assert that population adjusts via the so called *positive* and *preventive* checks. My model includes the two types of Malthusian population adjustment: (i) *preventive* checks, as both the decision to marry and the number of kids within marriage result from agents' optimization, and (ii) *positive* checks as I model the survival rate of adult agents as directly depending on their income in the following way:

$$s_t = \underline{s} \, y_t^{\phi} , \tag{10}$$

where $\underline{s}$ is a parameter calibrated to target an initial survival rate and ϕ is the elasticity of the survival rate to income per capita. Thus, in accordance with the Malthusian theory, adult's survival is increasing along income since $\underline{s} > 0$ and $\phi > 0$.

Population Dynamics.— The size of the population of the next generation $t+1$ is given by:

$$L_{t+1} = n_t \, p_t \, s_t \, L_t \; . \tag{11}$$

Income per capita Dynamics.— Forwarding (8) to period $t+1$ and using (8), (9) and (11), I obtain a first-order difference equation giving the income per capita of the next generation:

$$y_{t+1} = \left(\frac{1+g}{n_t(y_t) \; p_t(y_t) \; s_t(y_t)} \right)^{\alpha} \cdot y_t \equiv \psi(y_t) \; . \tag{12}$$

Steady State.— The steady state of the economy is defined by a situation in which:

$$y^* \equiv \left(\frac{1+g}{n(y^*) \; p(y^*) \; s(y^*)} \right)^{\alpha} = 1 \; . \tag{13}$$

At the steady state, the rate of population growth equals the rate of technological progress, such that income per capita remains constant period after period.[6]

3 Quantitative Analysis

In this section, I simulate the reaction of the English Malthusian economy to the Black Death in order to illustrate the convergence process of a Malthusian economy after a shock. I start by discussing the identification of the parameters that I use to calibrate the English Malthusian economy. I then discuss the simulation results of the calibration exercise and compare them with existing data.

[6]Section A of the Appendix shows that $\psi(y_t)$ has a unique and locally stable steady state $y^* > 0$, provided that y_t is not too low.

3.1 Identification of the Parameters and Initial Conditions

In order to simulate the evolution of a Malthusian economy and study its speed of convergence, I first set the value of some parameters *a priori*, while some others are set to match some target following an exact identification procedure. I focus on England as the literature already provides a rich array of parameter values for that economy during the Malthusian period. Table 1 summarizes and explains my calibration strategy.

Table 1: Benchmark Parameter Values

Parameter	Value	Interpretation and comments
t	25	Number of years per generation. Fixed *a priori*
γ	1	Preference for children. Fixed *a priori*
q	1	Unitary cost of a child. Fixed *a priori*
δ	0.074	Gives preventive checks-income per capita elasticity of 0.21. Fixed *a priori*
ϕ	0.1	Gives positive checks-income per capita elasticity of 0.1. Fixed *a priori*
α	0.5	Land share of output. Fixed *a priori*
g	0.023	Rate of technological progress per generation. Fixed *a priori*
$\underline{s}$	0.178	Minimum of the survival rate. To match $s^* = 0.71$
ν	0.662	Child quantity preference parameter. To match $n^* = 1.62$
b	3.48	Maximum of the search cost distribution. To match $p^* = 0.89$

Notes: See text for more details on the sources.

First, the length of a period or generation t is fixed at 25 years, meaning that an agent is living at most 50 years in my model.[7] This is in line with life expectancy figures in pre-industrial England as reported by Wrigley et al. (1997). Life expectancy at the age of 20 was as high as 33-34 years on the period 1550-1799. Conditional on their survival until the age of 20, Malthusian agents have therefore good chances to reach the age of 50. This is also in line with the evidence on the so-called European Marriage Pattern (EMP) from Hajnal (1965). Indeed, the EMP is characterized by a late age of first marriage for women (between age of 24 and 26) and low illegitimacy birth rates. In my setting, agents marry and procreate only in the second period of their life, that is to say between age of 25 and 50 as indicated by the EMP.

Next, I normalize γ and q, respectively the agent's preference for children and the cost of

[7] de la Croix and Gobbi (2017) make a similar assumption in a modern context with developing economies.

raising a child, to one.

Elasticity parameters δ and ϕ are particularly important in my setting, as they directly affect the speed of convergence (see Section 4). Since I am working at the generation level, I consider these parameters as representing respectively the long-run elasticity of the preventive checks (fertility and marriage) and the long-run elasticity of the positive checks (survival) to income per capita.[8] The empirical literature testing the Malthusian model in England provides various estimates of these long-run elasticities based on wage, Crude Birth Rate (CBR), Crude Marriage Rate (CMR) and Crude Death Rate (CDR) time-series (Lee, 1981; Lee and Anderson, 2002; Nicolini, 2007; Crafts and Mills, 2009; Klemp, 2012; Møller and Sharp, 2014). I set $\delta = 0.074$ and $\phi = 0.1$ in my benchmark specification to match the median of the long-run elasticities provided by the aforementioned literature. This corresponds to a long-run elasticity of 0.21 for the preventive checks and 0.1 for the positive checks. Table B-1 in the Appendix provides a complete list of studies, elasticity values, and details the method used to calibrate δ and ϕ.

Setting $\delta < 1$ means that my model consider decreasing returns to scale in the production of children, while most standard Malthusian models assume constant returns to scale ($\delta = 1$).[9] As pointed out by Lagerlöf (2019), we may interpret decreasing returns to scale in the production of children as stemming from an implicit production function for child survival featuring two inputs: parental time devoted to each child and each child's food intake. More children automatically yields less time per child, leading to an increase in the per-child amount of the consumption good necessary to ensure the survival of each child. Furthermore, the aforementioned empirical literature consistently finds values well below unity for the long-term elasticities of the preventive and positive checks. For instance, using exogenous cross-county variations in Swedish harvest between 1816 and 1856, Lagerlöf (2015) finds long-run elasticities of fertility, marriage and mortality of 0.1, 0.16 and -0.09, respectively.

The land share of output α for England is set at 0.5, corresponding to its estimated long-run

[8]The long-run elasticity is the sum of elasticities at various time lags.

[9]See, for instance, Ashraf and Galor (2011).

value for the Malthusian period (Federico et al., 2020).

In standard Malthusian models with constant technological progress, total population at the steady state is not constant. In fact, (13) shows that population grows at the same pace as technology; this is a necessary condition to keep income per capita constant at the steady state. Consequently, g is calibrated using 25-years average population growth using Broadberry et al. (2015) data for the period 1270-1675.

Consider next the three remaining parameters, $\underline{s}$, ν and b that are calibrated to match respectively the steady-state survival rate for adults (s^*), agent's steady-state fertility (n^*) and the steady-state marriage rate (p^*) following an exact identification procedure. The first target s^* is set to 0.71 as in Wrigley (1968). This corresponds to the survival rate of population of 25 years old until the age of 50 for the period 1538-1624 in England. The second target p^* is set to 0.89, which corresponds to a percentage of never married women of 11% as reported by Dennison and Ogilvie (2014) for England. This figure is the average of the percentage of never married women for England across 45 historical studies and is also very close to the value reported in the seminal study of Wrigley et al. (1989). Knowing the two first targets, the third target n^* is given by the steady-state condition in (13). To find the value of these three remaining parameters, I also set the steady-state level of income per capita y^* to an arbitrarily high initial level, by adjusting the initial level of technology A_0.

3.2 Simulation Results

This section shows the overall ability of my model to reproduce Malthusian dynamics and match some of the long-run dynamics of the English economy after the Black Death. To do so, I simulate a Black Death alike shock killing 60% of the population at $t = 5$. This is in line with Benedictow et al. (2004), who finds an overall mortality of 62.5% for England.

Figure 1 shows the evolution of income per capita (y_t), fertility (n_t), the marriage rate (p_t) and the survival rate (s_t) under my benchmark parametrization and under two alternative specifications, across 20 generations. The two alternative specifications are identical to the benchmark, with the exception of the long-run elasticity values used to calibrate δ and ϕ. Whereas I calibrate the benchmark using the median of the long-run elasticity values found in the literature for the preventive and positive checks, I calibrate the two alternative specifications using the 10th and 90th percentiles of the long-run elasticities (see Table B-1 for an overview of the long-run elasticity values I consider).

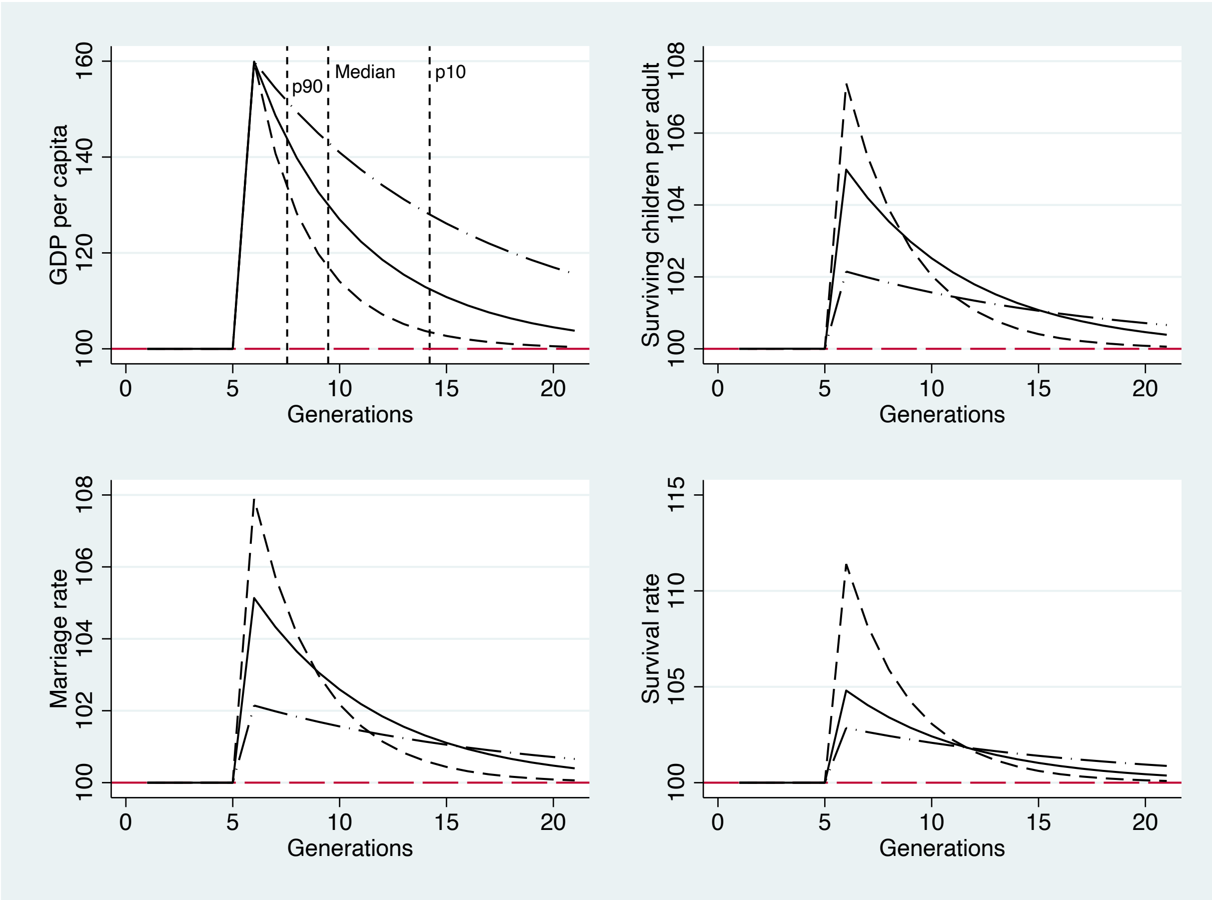
Standard Malthusian theory predicts that an exogenous negative shock on the population level (or Black Death) increases income per capita in the short run only.[10] After the shock, population increases and the economy gradually converges back to its steady state such that, at the long-run, the income per capita is constant. This is, by construction, what I observe in my model. Figure 1 shows that, right after the plague onset, the surviving agents enjoy indeed a temporarily higher level of income per capita. These better material conditions mean that agents have better chances to survive, they marry more and are able to raise more surviving children inside marriage. This translates into faster population growth, which in turn triggers the convergence process of income per capita to its steady state.

The top left panel of Figure 1 also display the half-life of convergence for the benchmark and the two alternative specifications. In the three cases, the elasticities imply long adjustments to shocks, indicating weak *homeostasis*. The half-life is about 112 years (4.47 generations $\times$ 25 years) in the benchmark scenario, 64 years (2.55 generations) in the 90th percentile scenario and 230 years (9.21 generations) in the 10th percentile scenario. It implies that any shock striking the Malthusian English economy is persistent across several generations. It takes, at least, 2.5 generations to fill half of the gap with respect to the steady-state.

As a complementary and illustrative exercise, Figure 2 evaluates the ability of the model to replicate the dynamic of income per capita after the Black Death, using English historical GDP per capita data from Broadberry et al. (2015). To do so, I first extract the cyclical component

[10] Jedwab et al. (2022) find evidence that the Black Death was indeed a plausibly exogenous shock to the European economy.

Figure 1: Responses of the English Malthusian Economy to a Black Death



Notes: This figure plots the response of per capita income (top-left panel), fertility (top-right panel), marriage (bottom-left panel) and survival (bottom-right panel) to a Black Death alike shock, killing 60% of the population at $t = 5$. The solid line indicates the benchmark scenario, using the median of the long-run elasticities of the preventive and the positive checks provided by the literature to calibrate the model (see Section B of the Appendix for more details). The longdashed and doted line indicates an alternative calibration, using the 10th percentile of the long-run elasticities. The dashed line indicates another alternative calibration, using the 90th percentile of the long-run elasticities. Vertical lines in the top-left panel indicate half-lives of the shock.

in the data using an Hodrick–Prescott filter.[11] This is necessary step, as my model analyses the dynamic of convergence to a unique and fixed steady state. On the contrary, fluctuations in the data might reflect changes in the position of the Malthusian steady state, as well as the transition to a fixed steady state. As argued by North and Thomas (1973) and Acemoglu and Robinson (2012), the Black Death might have affected the steady state of the English economy

[11] I set the smoothing parameter to 100 given that I use yearly data.

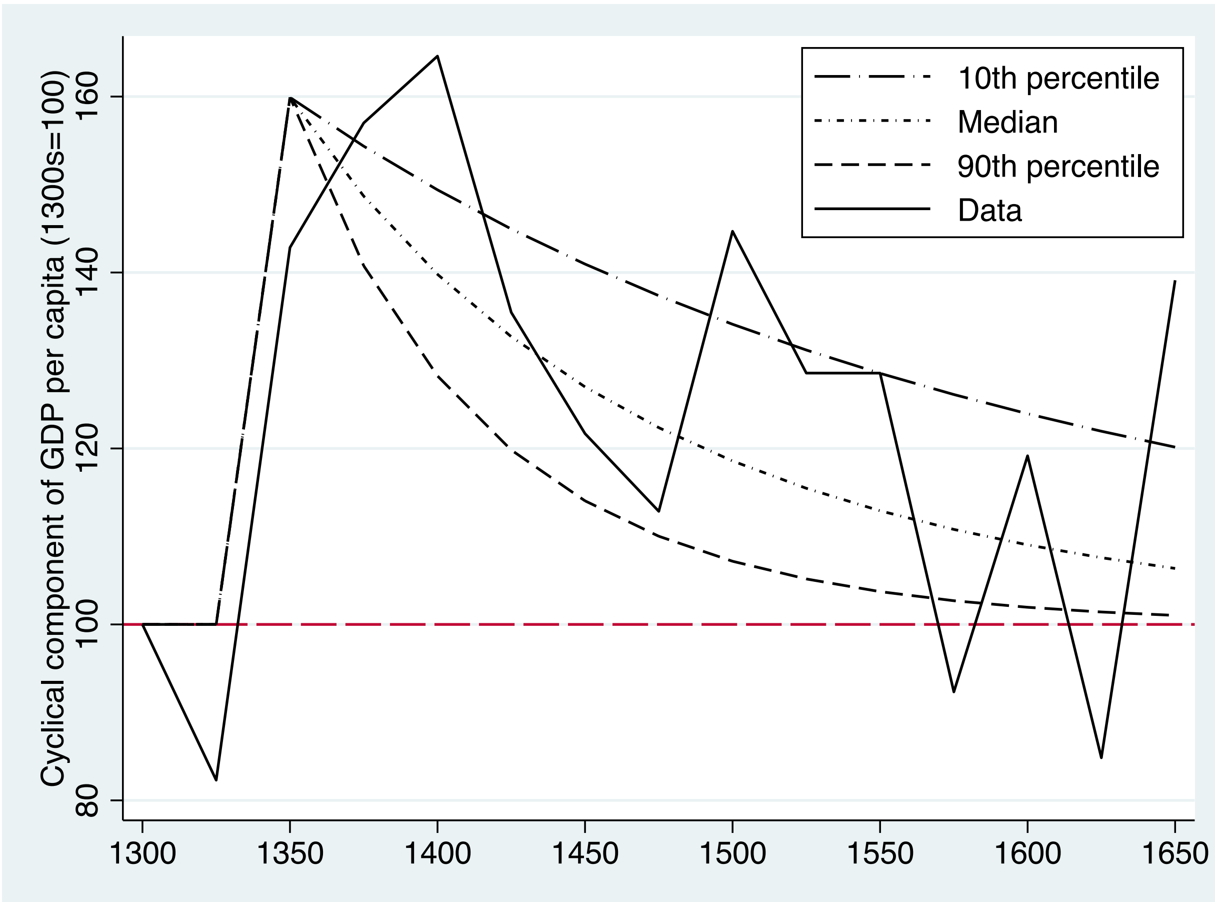
itself, through institutional changes.[12]

Figure 2 shows that a simple calibrated Malthusian model is able to generates a path for GDP per capita similar to the cyclical component of the data in the years following the Black Death for the English economy. This result is remarkable because the path predicted by the model is governed only by the initial demographic shock and the long-run elasticities provided by the empirical literature.

As my model has no stochastic components, deviations from the predicted trajectory reflect subsequent shocks hitting the Malthusian economy. The important point is that the overall trend remains within the limits of the three scenarios, all of which reflect a relativity weak Malthusian trap.

[12]For example, institutional changes allowing for an increase in the rate of technological progress g, would modify the position of the Malthusian steady state.

Figure 2: GDP per capita Dynamic after the Black Death: Simulated Paths vs. Data



Notes: This figure plots the cyclical component of GDP per capita from Broadberry et al. (2015) (solid line), and the simulated post-Black Death GDP per capita paths from the benchmark calibration using the median of the long-run elasticities (dashed and short dotted line) and two alternative calibrations using the 10th percentile (dashed and long dotted line) and the 90th percentile (dashed line) of the long-run elasticities. Data are normalized on the period 1300-1325, the last period before the occurrence of the Black Death in England (1348).

4 The Speed of Convergence in a Malthusian World

In this section, I start by deriving the speed of convergence of a Malthusian economy to its steady state, as implied by my model. Next, I use the derived formula and parameter values found in the literature to calculate the speed of convergence for various Malthusian economies, and compare it with the literature.

The speed at which GDP per capita converges to its steady state in a Malthusian economy is given by:

$$\beta^* = \alpha(\epsilon_{n_t} + \epsilon_{p_t} + \epsilon_{s_t}) \ , \tag{14}$$

where ϵ_{n_t}, ϵ_{p_t} and ϵ_{s_t} are the elasticities of fertility, marriage and survival with respect to income per capita. Section C of the Appendix provides further details on the derivation of the speed of convergence. The speed of convergence is therefore determined by the product of the land share of output α and the sum of the elasticities representing the preventive checks ϵ_{n_t} and ϵ_{p_t}, and the positive checks ϵ_{s_t}. Similar results are found by Irmen (2004) and Szulga (2012) in continuous time.

In Table 2, I compare the half-life obtained from my calibration of the English Malthusian economy with the speed of convergence found in the literature for other Malthusian and developing economies. In particular, I use equation (14) and long-run elasticity values provided by Galloway (1988), Lagerlöf (2015), Klemp and Møller (2016) and Pfister and Fertig (2020) to calculate the speed of convergence implied by my model for Denmark, Norway, Sweden, Germany and the median European Malthusian economy. I also report the half-life directly estimated by other studies for comparison purposes.

Despite the differences in period and context, the half-lives obtained for England appear to be in line with much of the literature. In particular, my benchmark result is very close to the half-life estimated by Fernihough (2013) for Northern Italy (112 years), or calculated using Galloway's (1988) long-run elasticity values for the median European Malthusian economy (115 years). My benchmark falls also close to a half-life of one century as found by de la Croix and Gobbi (2017) for Sub-Saharan Africa, or calculated using Lagerlöf's (2015) long-run elasticity values for Sweden. However, my estimations appear to be substantially higher than the half-lives found by Madsen et al. (2019) in a panel of 17 Malthusian economies.

Table 2: Speed of Convergence in my Calibrations and in the Literature

Country	Period	Authors	Half-Life (years)	Comments
England		Present study	112	Benchmark specification, calibrated using the median of the long-run elasticities reported in Table B-1 ($\delta = 0.074$; $\phi = 0.1$; $\alpha = 0.5$)
England		Present study	230	Alternative specification, calibrated using the 10th percentile of the long-run elasticities reported in Table B-1 ($\delta = 0.045$, $\phi = 0.06$ and $\alpha = 0.5$)
England		Present study	64	Alternative specification, calibrated using the 90th percentile of the long-run elasticities reported in Table B-1 ($\delta = 0.09$, $\phi = 0.23$ and $\alpha = 0.5$)
Sub-Saharan Africa	1990	de la Croix and Gobbi (2022)	198	Table 3, population regression
Europe	1540-1870	Galloway (1988)	115	Using (14), $\alpha = 0.5$ and reported elasticities
Northern Italy	1650-1881	Fernihough (2013)	112	Table 2, VAR estimates
Developing countries	1990	de la Croix and Gobbi (2017)	100	Table 7, population regression
Sweden	1816-1870	Lagerlöf (2015)	100	Using (14), $\alpha = 0.5$ and reported elasticities
Norway	1775-1853	Klemp and Møller (2016)	91	Using (14), $\alpha = 0.5$ and reported elasticities
Denmark	1821-1890	Klemp and Møller (2016)	84	Using (14), $\alpha = 0.5$ and reported elasticities
Germany	1730-1830	Pfister and Fertig (2020)	58	Using (14), $\alpha = 0.5$ and reported elasticities
Sweden	1775-1873	Klemp and Møller (2016)	48	Using (14), $\alpha = 0.5$ and reported elasticities
17 countries	1470-1870	Madsen et al. (2019)	29	Table 2, income regression
17 countries	1470-1870	Madsen et al. (2019)	12	Table 3, population regression
17 countries	1470-1870	Madsen et al. (2019)	11	Table 1, wage regression

5 Empirical Framework

In this section, I first present the data I use to empirically estimate the speed of convergence for a wide range of Malthusian economies. Then, I detail my main estimating equation and discuss potential threats to my identification strategy.

5.1 Data

In the empirical analysis that follows, I use two main types of datasets: (i) panel data on GDP per capita (historical or simulated) and (ii) panel data on historical population levels (total or urban population). The historical GDP per capita series come from the Maddison Project Database (Bolt and Van Zanden, 2020). Building on the pioneering work of Maddison (2003), the Maddison Project provides standardized historical GDP per capita series spanning several centuries. These series are regularly updated and enriched by researchers in the field of historical national accounting. However, as discussed in more detail in the next section, the uncertainty associated with past economic fluctuations is one of the concerns associated with the use of these sources. To limit measurement error issues, I focus on the period 1000-1800, and consider only countries with good data availability – i.e. countries for which GDP per capita data are available annually or every ten years before 1800. Following these two criteria, I consider a panel of twelve countries, including core (e.g. Italy, England, China) and more peripheric (e.g. Mexico, Poland, Sweden) Malthusian economies.

To complete my analysis, I also use simulated GDP per capita series from Lagerlöf (2019). Lagerlöf (2019) shows that a Malthusian model with stochastic and accelerating growth in land productivity is able to match the moments of historical GDP per capita series presented in Fouquet and Broadberry (2015). Simulations are available for 1,000 model economies and 501 years, making it very useful to circumvent the lack of GDP per capita data inherent to the pre-industrial period. From an econometric point of view, it corresponds to an ideal setting where both the cross-sectional and the time dimensions are large, limiting the bias of the different estimators on the speed of convergence.

For historical population series, I first use McEvedy et al.'s (1978) data. Population figures from this source have been widely used to answer various questions in the comparative development literature, with most of the contributions exploiting cross-country variations over a few years (Acemoglu et al., 2001; Nunn, 2008; Nunn and Qian, 2011; Ashraf and Galor, 2011, 2013).[13] My aim, on the other hand, is to exploit population changes within a country, and so I have coded McEvedy et al.'s (1978) data in their panel dimension. Although widely used in the literature, these data are also highly criticized, mainly for measurement error issues (Guinnane, 2021). To mitigate this problem, I use only a specific time frame and set of countries. First, I consider only the period between the years 1000 and 1750, which avoids the sizeable uncertainty surrounding population figures at the end of the Roman Empire and the beginning of the Middle Ages. Second, within that selected period, I keep only countries for which population figures are reported with maximum frequency – i.e. every century before 1600 and every half-century after 1600. Following these two criteria, I consider a panel of eighteen countries from this source for my empirical analysis.

To complement my analysis with historical population series, I am also using data from Reba et al. (2016), who compiled and geocoded urban population figures from Chandler (1987) and Modelski (2003). In particular, the database provides population level for cities worldwide from 3700 BC to 2000 AD. I apply the same procedure as for the other datasets, namely I first select urban population levels during the period 1000-1800.[14] Next, I focus on cities with a good data availability – i.e. cities for which a population figure is available for at least seven half-centuries (out of the seventeen potentially available) between the years 1000 and 1800.[15]

[13] For example, Ashraf and Galor (2011) use McEvedy et al. (1978) data as dependent variable, and exploit its cross-sectional variation in year 1, 1000 and 1500.

[14] When both Chandler (1987) and Modelski (2003) data are available for the same city and year, I take the average between the two figures. This was the case for 20 cities, only for year 1000.

[15] That threshold corresponds to the median of data availability.

5.2 Empirical Strategy

To empirically assess the speed of convergence of Malthusian economies, I rely on a standard β-convergence model. Such models have been extensively used in the growth literature to quantify the speed at which modern economies converge to their steady state (Barro, 1991; Barro and Sala-i Martin, 1992; Islam, 1995; Caselli et al., 1996; Barro, 2015). More recently, this framework has also been used in the Malthusian context by Madsen et al. (2019).

My main specification is the following dynamic panel:

$$\frac{\ln(y_{i,t}) - \ln(y_{i,t-\tau})}{\tau} = \beta \ln(y_{i,t-\tau}) + \gamma' \mathbf{X}_{\mathrm{i,t}} + \delta_t + \alpha_i + \varepsilon_{i,t} \ , \tag{15}$$

where $i = 1, ..., N$ indicates my unit of analysis which can be either a country or a city and $t = 1, ..., T$ corresponds to a given year. The left-hand side of equation (15) corresponds to the growth rate of my variable of interest y, which can be either GDP per capita or population levels, depending on the specification. The parameter τ indicates the number of years between two available data points, so that the dependent variable is always the average annual growth rate of y between period $t - \tau$ and t.

My coefficient of interest is β, which gives the average annual speed at which Malthusian economies converge to their steady state. Obtaining unbiased estimates of the speed of convergence is challenging in many ways. First, endogeneity is a concern, as past levels of economic development and current economic growth may be jointly determined by omitted factors. To mitigate that issue, equation (15) includes fixed effects α_i that control for time-invariant determinants of economic development, such as geography, climate and, to some extent, culture.

While partially solving the problem of omitted variables, country fixed effects are themselves recognized as a source of upward bias in the measurement of convergence speed in dynamic panels, known as the Hurwicz-Nickell bias (Hurwicz, 1950; Nickell, 1981). This is a potential problem, as it would constitute a systematic bias against weak *homeostasis* in my analysis. However, as highlighted by Barro (2015), the Hurwicz-Nickell bias tends towards zero when the overall sample length in years tends towards infinity. This means that the risk of a sizeable Hurwicz-Nickell bias is strongly mitigated in my analysis by the length of the overall sample, which spans several centuries.[16]

To address the endogeneity issue arising from time-varying omitted factors, the vector $\mathbf{X}_{\mathrm{i,t}}$ includes *Statehist* and its squared level as control variables (Borcan et al., 2018). *Statehist* is an index retracing state development every half-century from 3500 BC until today. I use it to proxy broad institutional changes that can affect the steady-state position of Malthusian economies. In equation (15), I also include time fixed effects δ_t to control for global changes in the steady-state determinants, such as the spread of new technologies or global climatic changes.[17]

To further address the endogeneity concerns, I provide results using an instrumental variable approach. In particular, I use the Arellano and Bond (1991) and the Blundell and Bond (1998) GMM estimators (hereafter referred to as AB and BB, respectively). These estimators have long been used in the context of growth regressions, either to estimate the speed of convergence of modern economies or to measure the effect of steady-state determinants.[18] Their advantage over the fixed effects estimator is the ability to instrument endogenous regressors, while controlling for country and time fixed effects.[19] However, one recognized potential issue using AB is the weakness of its instruments, which is known to bias β estimates towards their fixed effects counterparts. BB is more robust to that issue, but requires a stationarity assumption to deliver consistent results, which is found to not necessary hold in practice (Hauk and Wacziarg, 2009). Given the merits and drawbacks of each method, throughout the article I systematically present

[16] On the contrary, Barro (2015) finds that the Hurwicz-Nickell bias on the speed of convergence coefficient is sizeable in the modern growth context, where the analysis runs typically over 50 years.

[17] For instance, my analysis spans from the 11th to the 19th century, the period during which certain global climatic events, such as the Medieval Warm Period or the Little Ice Age, occurred. Time fixed effects can control for these events, provided that they affected a large part of the sample.

[18] This procedure was first used by Caselli et al. (1996) in the growth context to address both the Hurwicz-Nickell bias and the endogeneity of regressors.

[19] In particular, the AB estimator takes the first-difference of the regression equation and uses the lagged levels of the endogenous variables as instruments. The BB estimator complements AB, also using the lagged first differences of the endogenous variables as instruments for their levels.

estimates based on both estimators for comparison purposes.

Another source of concern is the measurement error of the lagged dependent variable. In presence of classical measurement error, i.e. random errors in the measurement of an explanatory variable, β will suffer from an attenuation bias, increasing the estimated speed of convergence. To limit this possibility, I implement several strategies. First, as detailed in Section 5.1, I systematically avoid using the most uncertain data on population or GDP per capita, excluding figures prior to the year 1000. Indeed, as pointed recently by Guinnane (2021), we simply "do not know the population" going that far back in the past where standardized and systematic censuses were not operated. Population and output measures between the years 1000 and 1800 also contain a sizeable part of uncertainty. However, local censuses, parish registers or proxy variables such as urbanization are increasingly available on that period, reducing measurement error. I also only consider countries or cities with the best, or at least above median, data coverage in each source. Second, I follow the usual practice in the empirical macroeconomic literature and calculate 50-year averages of the explanatory variables when the data is available at a lower frequency.[20] This allows me to avoid spurious changes and focus on long-term dynamics. Third, AB and BB estimators would also mitigate this source of bias, as instrumental variables can in principle deal with classical measurement error.

Nevertheless, there remains the possibility of non-classical measurement error, such as systematic and persistent differences over time in the measure of explanatory variables between countries. If this type of measurement error is highly persistent over time, it will be treated by the country fixed effects.[21] To account for less persistent measurement error across countries, I also systematically run fixed effects regressions with year-interacted lagged dependent variables. In this case, any varying differences in measurement correlated with initial population or initial GDP per capita levels will be taken into account. Typically, I find that this approach do not

[20]This means that I take a minimal τ of 50 years in equation (15).

[21]Similarly, time fixed effects can deal with measurement errors that vary over time and are common to the countries in the sample, such as the gradual improvement in population figures as we approach year 1800.

differ significantly from the results of my baseline fixed effects regressions.

In my estimation strategy, I consider only the "within country" class of estimators (fixed effects, AB and BB), while growth regressions have also been estimated using the between or a random effects estimator. Monte Carlo simulations on β-convergence regressions in the context of modern growth have found mixed evidence about the ability of the two classes of estimators to accurately estimate the speed of convergence. Hauk (2017) finds that the speed of convergence is best estimated with the within-country class of estimators when endogeneity bias on the steady-state determinants is the main concern. On the contrary, Hauk and Wacziarg (2009) find that the speed of convergence is best estimated using the between or random effects estimator when regressor measurement error is the dominant issue. In the present case, I consider the endogeneity bias to be the most serious threat and therefore use the within-country type of estimator for two main reasons. First, there are very few control variables available for a large sample and a long period of analysis in the Malthusian context, opening the possibility of a substantial endogeneity bias stemming from omitted variables. Second, measurement error is dealt to a certain extent by the various strategies described in the previous two paragraphs. Furthermore, Hauk and Wacziarg (2009) show that the within-county estimators imply a higher speed of convergence. This means that measurement error will ultimately constitute a bias against the weak *homeostasis* hypothesis. In my results, I show consistent evidence of weak *homeostasis*, suggesting that measurement error is indeed a second-order concern.

6 Results

In this section, I present my empirical estimates of the speed of convergence for various Malthusian economies. I start by presenting my results using historical and simulated per capita income data. Then, I present my results using historical data on total and urban population.

6.1 Speed of Convergence using GDP per capita Data

In Table 3, I report the estimations of specification (15) using OLS and fixed effects. The dependant variable is the average annual growth rate of GDP per capita calculated from Maddison Project's data (Bolt and Van Zanden, 2020).[22] I first present the relationship between the dependent variable and the initial level of GDP per capita, controlling for time fixed effects (columns 1 and 4). Then, I add country fixed effects (column 2 and 5). Finally, I add *Statehist* and its squared level as control variables (columns 3 and 6).

Table 3: Speed of Convergence using GDP per capita Data from the Maddison Project

Sample Used:	Full			Europe		
	OLS (1)	FE (2)	FE (3)	OLS (4)	FE (5)	FE (6)
log(GDPpc)	-0.0006	-0.0057**	-0.0057***	0.0000	-0.0046**	-0.0046**
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	Yes	No	Yes	Yes
Statehist	No	No	Yes	No	No	Yes
Observations	85	85	85	69	69	69
adj. R-sq	-0.01	0.16	0.14	-0.05	0.11	0.08
Half-Life	1197	122	121	-18766	150	152
Half-Life 95% C.I.	[-434,252]	[422,71]	[391,72]	[-356,370]	[587,86]	[663,86]

Notes: This table presents OLS estimates of the speed of convergence using GDP per capita data from the Maddison Project at the country level. Columns 1-3 present results obtained from the full sample of countries considered from the Maddison Project data, and columns 4-6 show results obtained by focusing on European countries. For each sample, I first display the relationship controlling for time fixed effects in column 1, then include country fixed effects and finally add *Statehist* as control. Standard errors clustered at the country level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Starting with the most parsimonious specification, with only time fixed effects as controls, column 1 reveals that the lagged dependent variable coefficient is not statistically different from zero. This is not really surprising as the omitted variable bias is substantial in this case, driving the lagged dependent coefficient toward zero. Moreover, as my theoretical model suggests,

[22] In this case, GMM estimates are not reported due to the lack of observation units. Indeed, as Roodman (2009) advises, a useful rule of thumb to avoid weak instrument problems in GMM estimations is to keep the total number of instruments below the number of observation units. This is not possible with the current sample from the Maddison Project, as we have eleven countries and fifteen instruments in the most parsimonious case, resulting in unitary Hansen test p-values.

Malthusian economies should display conditional convergence rather than absolute convergence, as the steady-state position of each economy depends on its characteristics.[23]

Adding country fixed effects, column 2 reveals a negative and significant relationship between GDP per capita growth and the initial level of GDP per capita, indicating conditional convergence of Malthusian economies. The estimated coefficient implies a half-life of 122 years (ln(2)/0.0057), with a 95% confidence interval giving half-lives between 422 years and 71 years. Therefore, the most comprehensive and up-to-date historical GDP per capita series are consistent with weak *homeostasis* of Malthusian economies, as it takes at least several generations to absorb half of a shock. Compared to other studies, the results in column 2 are close to Fernihough (2013), who find a half-life of 112 years for Northern Italy (1650-1881) using VAR methods. However, this result is in great contrast with Madsen et al. (2019), who find a half-life of 29 years for income per capita and conclude in favor of strong *homeostasis* of Malthusian economies.[24]

One possible reason for the distortion of the estimated convergence speed in favor of weak *homeostasis* is the presence of a severe omitted variable bias. In particular, column 2 does not control for time-varying determinants of GDP per capita growth at the country level, as it includes only time and country fixed effects. To limit that concern, column 3 adds *Statehist* and its squared level as controls. The speed of convergence is almost unaffected, as the reported half-life is now slightly higher at 121 years.

As a robustness check, columns 4 to 6 replicate the analysis, restricting the sample to European countries, giving similar results. In particular, column 6 indicates an even slower speed of convergence on average, with a half-life of 152 years, confirming the weak *homeostasis* pattern found in the previous columns. However, I find no significant differences in the estimated

[23]From the steady-state condition in (13), it is clear that two economies, with for instance different rates of technological progress g, will not converge to the same steady state.

[24]Note that my article has several methodological differences with respect to Madsen et al. (2019). First, they rely on interpolated data coming from heterogeneous sources for GDP per capita and population data, while I take the data as given from each source. Second, they use seemingly unrelated regression (SUR) models, a random effects family estimator, while I use within-country estimators (LSDV, AB and BB).

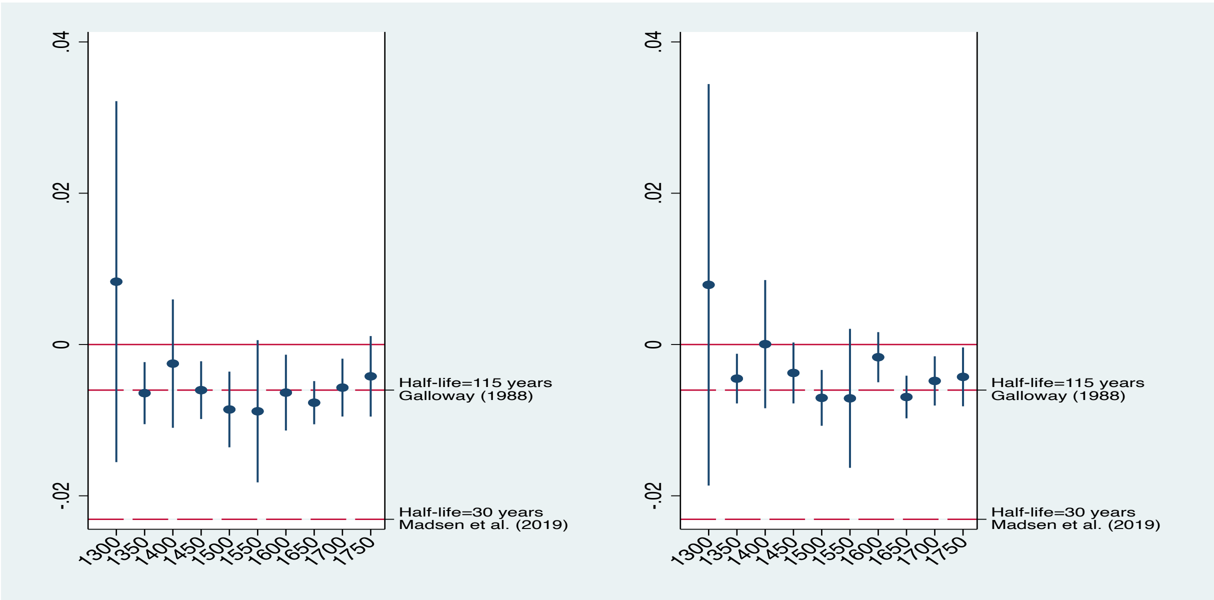
speed of convergence between the two country samples.

Figure 3 displays the fixed effects estimations of columns 3 and 6, adding an interaction term between time fixed effects and the initial level of GDP per capita. This allows me to examine the heterogeneity of the speed of convergence through time, and to check the possible influence of non-classical measurement errors. Overall, the point estimates are negative and statistically different from zero at the 5% level. Whether considering the full or the European sample of countries, the vast majority of the estimated coefficients are not statistically different from a half-life of 115 years, as found for Europe using the long-run elasticities of Galloway (1988) in Section 4. This indicates a clear and stable pattern of weak *homeostasis* during a large part of the Malthusian period. On the contrary, strong *homeostasis*, as represented by the highest half-life (about 30 years) found in Madsen et al. (2019), is always rejected at the 5% level.

Turning to the heterogeneity of the speed of convergence by country, Figure 4 displays point estimates of the fixed effects estimations of column 3 and 6, adding an interaction term between the country fixed effects and the initial level of GDP per capita. Figure 4 reveals mixed results as some countries are found compatible with weak *homeostasis* (e.g. the Netherlands), and some other countries rather lean towards strong *homeostasis* (e.g. Poland). Some countries, like France or Spain, are even found to be compatible with both types of *homeostasis*. However, precision of estimates is clearly an issue in that specification. Indeed, as shown in Figure 4, confidence intervals are generally large.

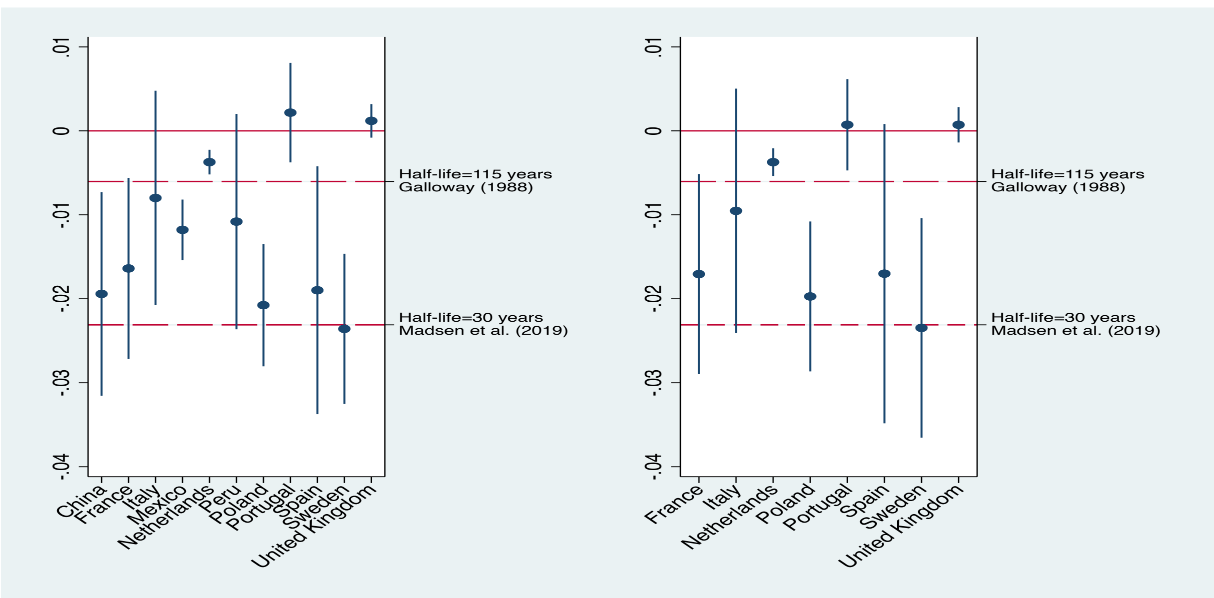
In addition to the above results, Table 4 reports OLS, fixed effects and GMM estimates of specification (15), where the dependent variable is the average annual growth rate of GDP per capita calculated using Lagerlöf's (2019) simulated data. In particular, the simulated GDP per capita series are produced from a Malthusian model with stochastic and accelerating growth in land productivity. Under plausible parameter values, Lagerlöf (2019) shows that the model is able to accurately reproduce the empirical moments of the historical GDP per capita series presented in Fouquet and Broadberry (2015) for several European economies between 1300 and

Figure 3: Speed of Convergence per period using the Maddison Project Data



Notes: This figure plots estimates of the speed of convergence using GDP per capita data from the Maddison Project. It corresponds to the FE estimations in column 3, Table 3 (left panel) and in column 6, Table 3 (right panel), adding year-interacted lagged GDP per capita levels as controls. 95% confidence intervals are reported.

Figure 4: Speed of Convergence per country using the Maddison Project Data



Notes: This figure plots estimates of the speed of convergence using GDP per capita data from the Maddison Project. It corresponds to the FE estimations in column 3, Table 3 (left panel) and in column 6, Table 3 (right panel), adding country-interacted lagged GDP per capita levels as controls. 95% confidence intervals are reported.

1800. The original series presented in Fouquet and Broadberry (2015) are still part of the latest Maddison Project database for some countries (e.g. Holland and Italy), or are updated versions using the same methodology (e.g. England and Sweden). Consequently, the main advantage of using this simulated series to estimate convergence speed is to gain in precision, since the simulated data correspond to the same moments while possessing a much greater temporal and cross-sectional dimension.

Table 4: Speed of Convergence using simulated GDP per capita Data from Lagerlöf (2019)

	OLS (1)	FE (2)	GMM-AB (3)	GMM-BB (4)
log(GDPpc)	-0.0019***	-0.0052***	-0.0063***	-0.0047***
	(0.000)	(0.000)	(0.002)	(0.001)
Time FE	Yes	Yes	Yes	Yes
Country FE	No	Yes	Yes	Yes
Observations	10000	10000	9000	10000
adj. R-sq	0.09	0.18	.	.
AR(7)			0.17	0.18
Hansen			0.22	0.23
Diff. Hansen			.	0.21
Instruments			13	15
Half-Life	363	133	110	146
Half-Life 95% C.I.	[403,330]	[141,126]	[212,75]	[250,103]

Notes: This table presents OLS and GMM estimates of the speed of convergence using simulated GDP per capita data from Lagerlöf (2019) at the country level. Column 1 controls for time fixed effects, and the subsequent columns add country fixed effects. The GMM estimations in columns 3 and 4 use the seventh and further lagged levels of GDP per capita as instruments. I use a collapsed matrix of instruments and report the number of instruments. The AR(7) row reports the p-value of a test for the absence of seventh-order correlation in the residuals. Standard errors clustered at the country level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

As expected, the speed of convergence is now estimated with much more precision. The fixed effects estimation in column 2 shows a half-life of 133 years, with a 95% confidence interval giving half-lives between 141 and 126 years. These results lie within the wide confidence intervals of the previous results in Table 3 using Maddison Project's data. The Hurwicz-Nickell bias is very unlikely to affect the estimates, as this is a setting where the time dimension is very large ($T = 500$).

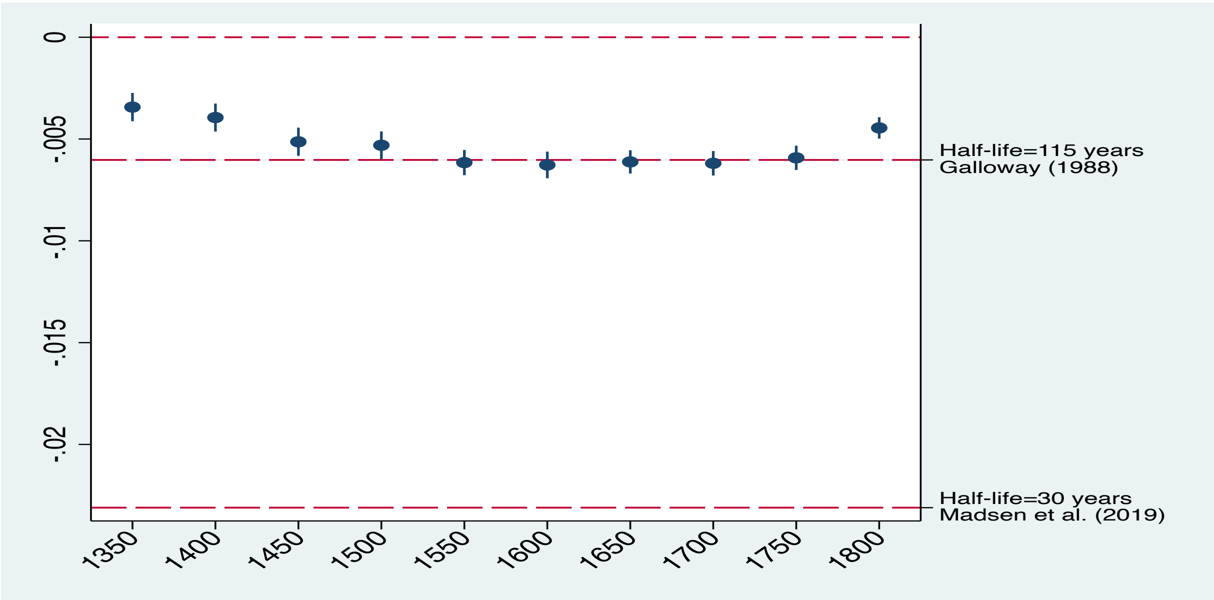
Columns 3 and 4 present AB and BB GMM results. In both cases, the estimated speed of convergence is highly significant and consistent with weak *homeostasis*. In particular, I find that both GMM estimates are not statistically different from the speed of convergence estimated by the fixed effects model in column 2. This may seem worrying, as it is generally considered to be a a sign of the weak instrument problem in the literature. However, as mentioned above, the fixed effect estimation of column 2 takes place in an ideal setting where its main source of bias – i.e. the Hurwicz-Nickell bias – is expected to be small. Under these conditions, it is plausible that GMM and fixed effects estimations give similar results.

The classical GMM post-estimation tests give also clear signs that the moment conditions are globally satisfied. In particular, I reject the null hypothesis of seventh-order serial correlation in the residuals (AR(7) test), meaning that using the seventh (and greater) lag of GDP per capita as instruments does not violate the exclusion restriction. Second, I reject both the null hypothesis of the Hansen test and the difference in Hansen test for all GMM instruments, indicating that the set of used instruments are plausibly exogenous. Overall, I consider that the GMM estimates provide converging evidence of weak *homeostasis*.

Figure 5 investigates the time heterogeneity of the speed of convergence. All the coefficients are statistically different from zero and very precisely estimated, thanks to the large time and sample size. The speed of convergence is fairly stable over time. Half of the estimated coefficients cannot reject a half-life of 115 years at the 5% level, as found for Europe using the long-run elasticities of Galloway (1988). In addition, all the remaining coefficients show a slower speed of convergence, again indicating a weak *homeostasis* of Malthusian economies.

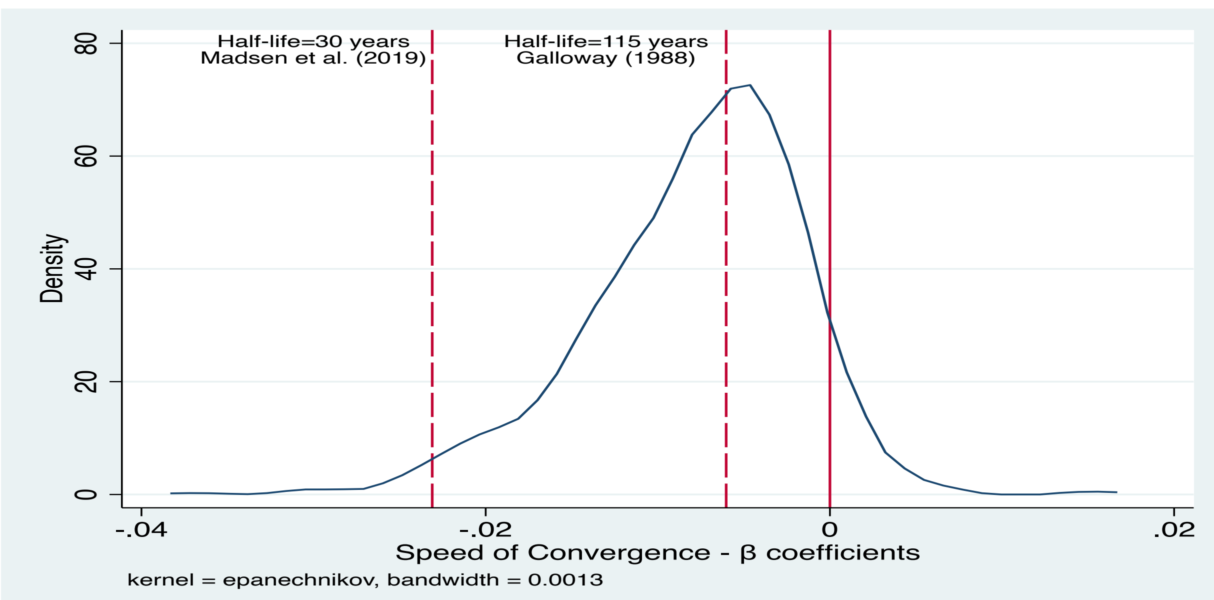
The large cross-sectional dimension of Lagerlöf's (2019) data allows me to study the range of plausible half-lives in Malthusian economies with greater consistency than with Maddison Project's data. To do so, I perform the fixed effects estimation in column 2, adding an interaction term between the country fixed effects and the initial level of GDP per capita to estimate the speed of convergence for each Malthusian economy. Figure 6 displays the kernel density of

Figure 5: Speed of Convergence per period using Lagerlöf (2019) Data



Notes: This figure plots estimates of the speed of convergence by period using simulated GDP per capita data from Lagerlöf (2019). It corresponds to the FE estimation in column 2, Table 4, adding year-interacted lagged GDP per capita levels as controls. 95% confidence intervals are reported.

Figure 6: Speed of Convergence per country using Lagerlöf (2019) Data



Notes: This figure plots the kernel density of the estimated speed of convergence by country using simulated GDP per capita data from Lagerlöf (2019). It corresponds to the LSDV estimation in column 2, Table 4, adding country-interacted lagged GDP per capita levels as controls.

the estimated speed for the 1000 simulated Malthusian economies in Lagerlöf (2019).[25] Consistent with the previous country-level evidence and my results, it appears that the mode of the distribution is very close to a half-life of 115 years, as found for Europe using the long-run elasticities of Galloway (1988). As a result, most pre-industrial economies were in a moderate Malthusian trap or weak *homeostasis*. Interestingly, some Malthusian economies appear to have lived under a strong Malthusian trap, with half-lives of 30 years or less, as found by Madsen et al. (2019).

6.2 Speed of Convergence using Population Data

In Section C of the Appendix, I show that the speed of convergence of population to its steady state in my Malthusian model is the same as for GDP per capita. Therefore, in this section, I use the same β-convergence models and population data to provide additional estimates of the speed of convergence during Malthusian times.

In Table 5, I present my results based on OLS, fixed effects and GMM estimations of equation (15). The dependent variable is the average annual population growth rate, calculated from McEvedy et al. (1978) population figures. I first present the relationship between the dependent variable and the initial population levels, controlling for time fixed effects (column 1). Then, I add country fixed effects (column 2) and *Statehist* and its squared level as control variables (column 3). Finally, I perform AB and BB GMM estimations (columns 4 and 5).

Controlling for time and country fixed effects, column 2 reveals a negative and highly significant relationship between population growth and its initial level. The implied half-life is about 147 years, which is in line with my previous results using historical GDP per capita series (see Table 3, column 3, and Table 4, column 2). The 95% confidence interval indicates half-lives between 224 and 109 years, which stays clearly in the range of weak *homeostasis*.

[25] Figure D-1 in the Appendix delivers the point estimates along with their 95% confidence intervals for the 200 first simulated economies in Lagerlöf (2019).

Table 5: Speed of Convergence using Population Data from McEvedy et al. (1978)

	OLS (1)	FE (2)	FE (3)	GMM-AB (4)	GMM-BB (5)
log(Population)	-0.000***	-0.005***	-0.006***	-0.009***	-0.004*
	(0.000)	(0.001)	(0.001)	(0.003)	(0.002)
Time FE	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	Yes	Yes	Yes
Statehist	No	No	Yes	Yes	Yes
Observations	180	180	180	162	180
adj. R-sq	0.48	0.60	0.61	.	.
AR(2)				0.69	0.36
Hansen				0.94	0.99
Diff. Hansen				.	0.87
Instruments				18	22
Half-Life	4414	147	125	73	167
Half-Life 95% C.I.	[12873,2663]	[224,109]	[231,86]	[215,44]	[-1176,78]

Notes: This table presents OLS and GMM estimates of the speed of convergence using population data from McEvedy et al. (1978) at the country level. Column 1 controls for time fixed effects, the subsequent columns add country fixed effects and *Statehist*. The GMM estimations in columns 4 and 5 use the second to fourth lagged levels of population as instruments. *Statehist* is and its squared level are treated as endogenous and instrumented with the same set of lags as population. I use a collapsed matrix of instruments and report the number of instruments. The AR(2) row reports the p-value of a test for the absence of second-order correlation in the residuals. Standard errors clustered at the country level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Dealing further with the omitted variable issue, column 3 adds *Statehist* and its squared level as controls. Convergence tends to be faster on average, with a half-life of 125 years. However, I do not find significant differences in the speed of convergence between columns 2 and 3.

Columns 4 and 5 use GMM estimation procedures. Starting with the AB estimation, column 4 shows a faster average speed of convergence than the fixed effects results, with a half-life of 73 years. This is potentially problematic, as it could reflect the influence of weak instruments. Keeping the same set of instruments, the BB estimation indicates weak *homeostasis*, but is only weakly significant. The difficulties associated with using GMM in this case stem from the fact that the cross-sectional dimension is small using McEvedy et al.'s (1978) data relative to the number of instruments used. This is the well-known problem of "too many instruments", highlighted by Roodman (2009).[26] Under these conditions, my preferred specification is the

[26]Symptomatic of this problem, column 4 and 5 of Table 5 reveal Hansen's test p-value very close to one. This

fixed effects model with controls in column 3, assuming the Hurwicz-Nickell bias is small. As mentioned in Section 5.2, this is all the more plausible given that the time dimension spans over several centuries in this Malthusian context.

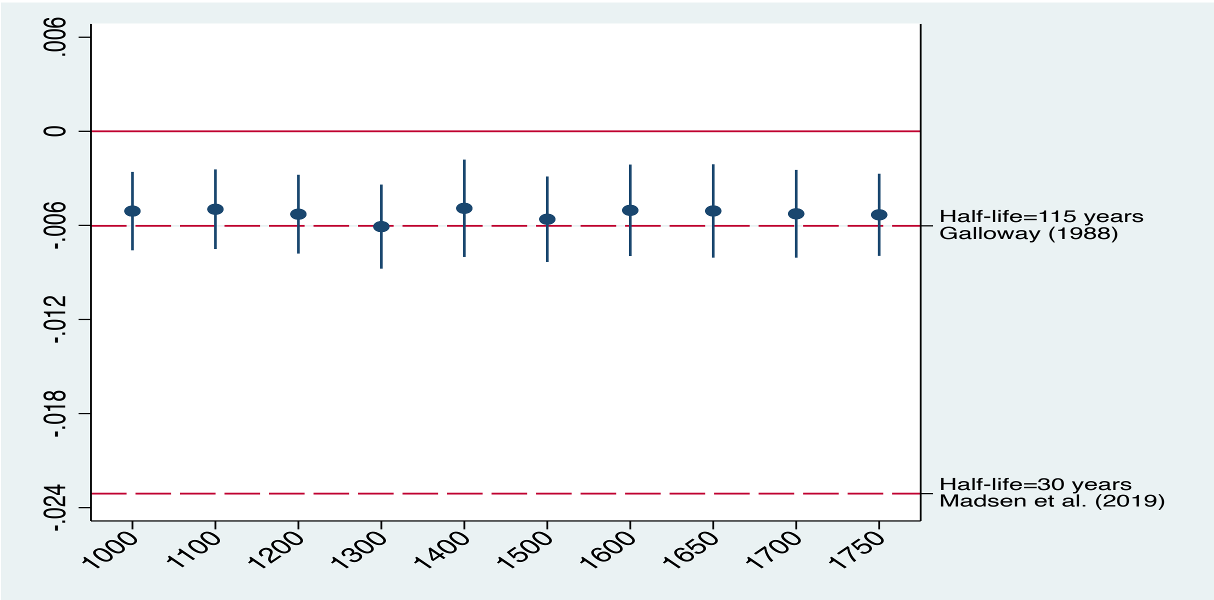
Figure 7 displays the point estimates for the fixed effects estimation in column 3, adding year-interacted initial population levels. All estimated coefficients are statistically different from zero and consistent with a half-life of 115 years, as found for Europe using the long-run elasticities of Galloway (1988). The point estimates are fairly stable in terms of magnitude, and within a range compatible with weak *homeostasis*, confirming my previous results using GDP per capita data.

Figure 8 investigates the cross-country heterogeneity of the speed of convergence. Confidence intervals are narrower than for the Maddison Project data, highlighting significant differences in the speed of convergence between Malthusian economies. The strongest Malthusian trap is in Spain, with an half-life of 48 years and a 95% confidence interval between 40 and 61 years. On the other side of the spectrum, the weakest Malthusian trap is in Japan, with an half-life of 118 years. The estimated half-life of the English Malthusian economy is 85 years, with a 95% confidence interval between 71 and 106 years. This figure is lower than the half-life obtained from the calibration of my benchmark Malthusian model in Section 3.2 (112 years). However, the order of magnitude remains similar, as the estimated half-life lies between the two alternative calibration scenarios, which give half-lifes of 64 and 230 years. These significant differences among Malthusian economies suggest that a common shock could persist substantially longer in England than in Spain, which is closer to strong *homeostasis*. Despite these significant differences, the overall pattern remains compatible with a relatively weak *homeostasis*, since it takes at least several generations to absorb half of a shock.

In Table 6, I present my results based on OLS, fixed effects and GMM estimations of equation (15). The dependent variable is the average annual urban population growth rate,

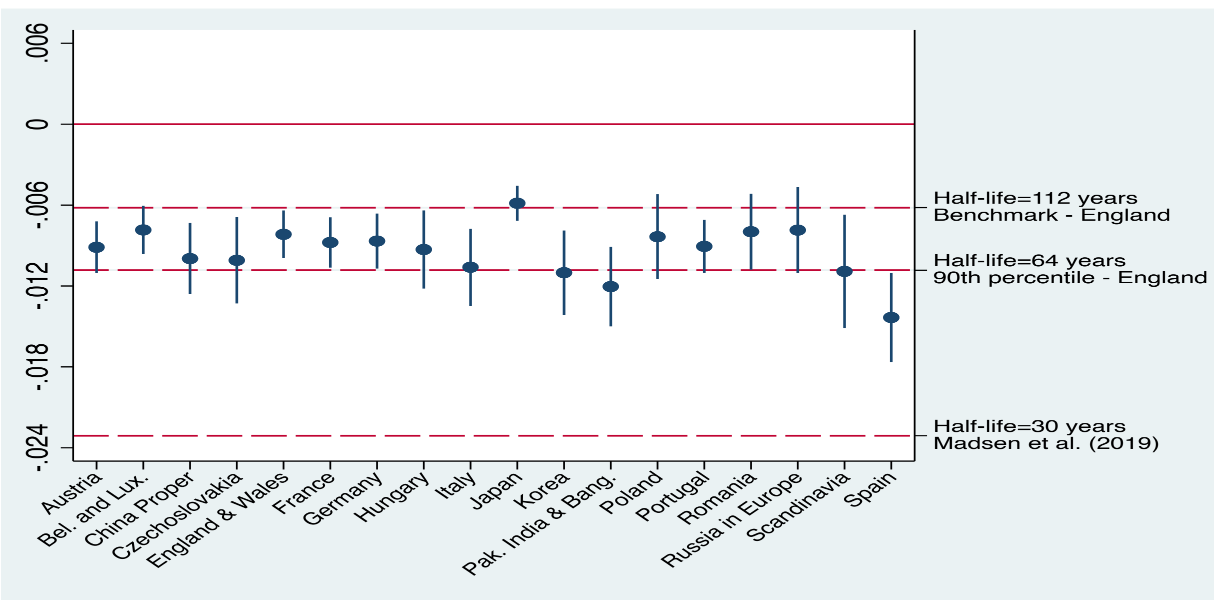
is due to the fact that the number of countries in the sample (18 in this case) is very close to or less than the number of instruments, even when considering a parsimonious instrumentation.

Figure 7: Speed of Convergence per period using McEvedy et al. (1978) Data



Notes: This figure plots estimates of the speed of convergence by period using population data from McEvedy et al. (1978). It corresponds to the FE estimation in column 3, Table 5, adding year-interacted lagged population levels as controls. 95% confidence intervals are reported.

Figure 8: Speed of Convergence per country using McEvedy et al. (1978) Data



*Notes:*This figure plots estimates of the speed of convergence by country using population data from McEvedy et al. (1978). It corresponds to the FE estimation in column 3, Table 5, adding country-interacted lagged total population levels as controls. 95% confidence intervals are reported.

calculated using Reba et al. (2016) data. Using urban population data to estimate the speed of convergence is interesting because the frequency of observations and the sample size are higher than for the country-level population data of McEvedy et al. (1978), which increases precision. I perform city-level estimations in columns 1-4 and estimations with urban population data aggregated at the country level in columns 5-9. In each case, I first present the relationship between the dependent variable and the initial population level, controlling for time fixed effects (columns 1 and 5). Then, I add respectively city and country fixed effects (columns 2 and 6). When possible, I add *Statehist* and its squared level as control variables (column 7). Finally, I provide GMM estimation results (columns 3, 4, 8 and 9).

Starting with the city-level estimations, column 2 reveals a negative and highly significant relationship between urban population growth and the initial level of urban population, conditional on time and city fixed effects. The corresponding half-life is 95 years, with a 95% confidence interval indicating half-lives between 155 and 68 years.

The GMM estimates in columns 3 and 4 confirm the results of the fixed effects estimation, with half-lives of 97 and 104 years respectively, and similar confidence intervals. Both GMM estimates reject the presence of second-order correlation in the residuals (AR(2) test), demonstrating the validity of the set of instruments used. It is worth noticing that the AB estimation fails to satisfy the Hansen test at the usual confidence levels. Reassuringly, using the same set of instruments, the Hansen test of overindentifying restrictions and the difference in Hansen test indicate that the moment conditions are satisfied in the BB estimation in column 4.

Turning to the country-level estimations, column 6 reveals a negative and highly significant relationship between urban population growth and its initial level, conditional on time and country fixed effects. The half-life is almost identical to the previous fixed effects estimate using city-level data in column 2, but is now estimated with greater precision.

Aggregating at the country level, I am now able to control further for time varying determinants of population growth. Column 7 adds *Statehist* and its squared level as control variables.

Table 6: Speed of Convergence using Urban Population Data from Reba et al. (2016)

Observational Unit:	City				Country				
	OLS (1)	FE (2)	GMM-AB (3)	GMM-BB (4)	OLS (5)	FE (6)	FE (7)	GMM-AB (8)	GMM-BB (9)
log(Population)	-0.003***	-0.007***	-0.007***	-0.007***	-0.004***	-0.007***	-0.008***	-0.009**	-0.006**
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.002)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes	No	No	No	No	No
Country FE	No	No	No	No	No	Yes	Yes	Yes	Yes
Statehist	No	No	No	No	No	No	Yes	Yes	Yes
Observations	1706	1706	1239	1706	509	509	509	411	509
adj. R-sq	0.08	0.22	.	.	0.16	0.24	0.26	.	.
AR(2)			0.49	0.50					
AR(3)								0.10	0.10
Hansen			0.07	0.13				0.51	0.46
Diff. Hansen				0.92					0.37
Instruments			30	32				24	28
Half-Life	258	95	97	104	192	94	84	80	109
Half-Life 95% C.I.	[500,174]	[155,68]	[212,63]	[171,75]	[270,149]	[133,73]	[118,64]	[481,44]	[480,62]

Notes: This table presents OLS and GMM estimates of the speed of convergence using urban population data from Reba et al. (2016) at the city and country level. Columns 1-4 present results using city-level data and columns 5-9 show results using urban population aggregated at the country level. For each sample, I first display the relationship controlling for time fixed effects, then include city or country fixed effects and finally add *Statehist* and its squared level as controls (this was not possible for the city-level estimates). The GMM estimations in columns 3 and 4 use the second and further lagged levels of urban population at the city level as instruments. Columns 8 and 9 use the third to fifth lagged levels of urban population aggregated at the country level as instruments. *Statehist* and its squared level are treated as endogenous and instrumented with the same set of lags as urban population in columns 8 and 9. I use a collapsed matrix of instruments and report the number of instruments. The AR(2) row reports the *p*-value of a test for the absence of second-order correlation in the residuals for the city-level GMM estimations and the AR(3) row reports the *p*-value of a test for the absence of third-order correlation in the residuals for the country-level GMM estimations. Standard errors clustered at the city level in columns 1-4 and at the country level in columns 5-9 are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

The estimated speed of convergence is now faster with a half-life of 84 years, but remains consistent with weak *homeostasis*. In particular, the 95% confidence interval indicates half-lives between 118 and 64 years.

Columns 8 and 9 present AB and BB GMM estimates of the speed of convergence at the country level. Both estimates confirm the weak *homeostasis* pattern found in the previous column, with half-lives estimated at 80 and 109 years, respectively. In both cases, the GMM estimates appear to be much less precise than the fixed effect estimate in column 7, as the confidence intervals for the half-lives are now larger, while remaining compatible with weak *homeostasis*.

Overall, my results using historical urban population data clearly confirm the weak *homeostasis* pattern found in the previous sections. Most of the estimated half-lives in Table 6 are close to one century, and the smallest half-life found is 80 years.

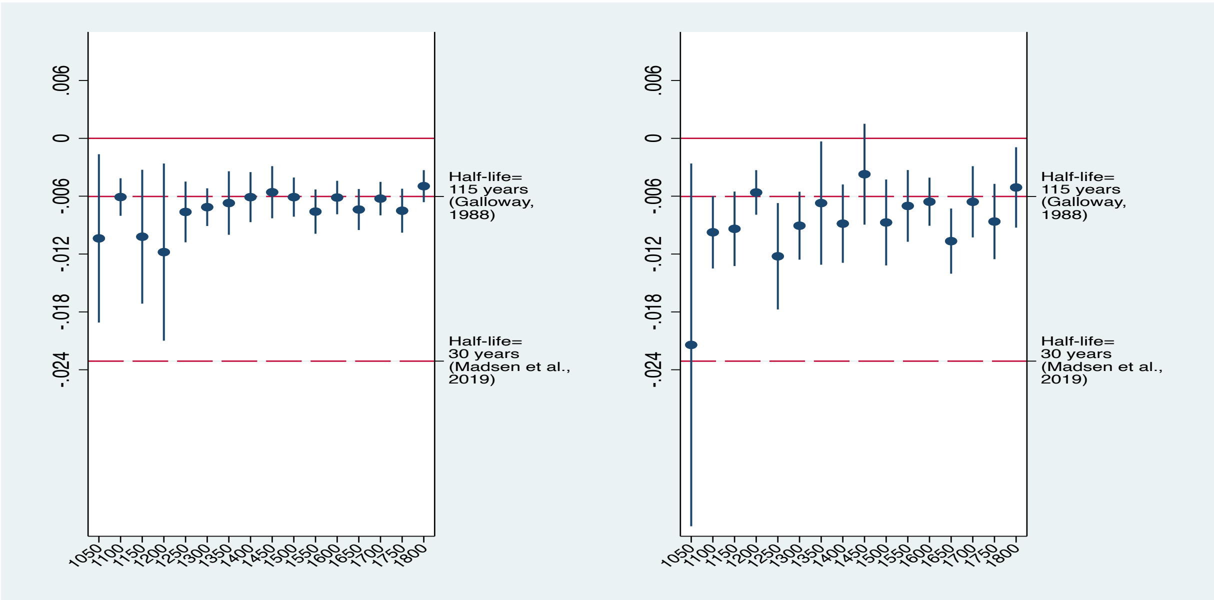
Figure 9 explores the time heterogeneity of the speed of convergence, both for the city-level and country-level estimations. In both cases, a stable pattern of weak *homeostasis* over time is confirmed. This is particularly striking for the city-level data, where all the point estimates starting from the year 1250 onwards cannot reject a half-life of 115 years at the 5% level.

Figure 10 plots the kernel density of the estimated speed of convergence for a sample of 185 cities.[27] It reveals a pattern similar to my previous findings using Lagerlöf's (2019) simulated data, with the mode of the distribution very close to a half-life of 115 years. Moreover, the distribution is also more concentrated around that value than my previous estimates (Figure 6), giving additional support to the widespread of weak *homeostasis* across Malthusian societies. Finally, as in Figure 6, a strong Malthusian trap cannot be rejected for some cities.

Figure 11 shows the heterogeneity of the speed of convergence at the country level, using urban population data. Alongside the previous estimates using total population data (Figure 4) and the above results at the city level (Figure 10), Figure 11 highlights significant differences

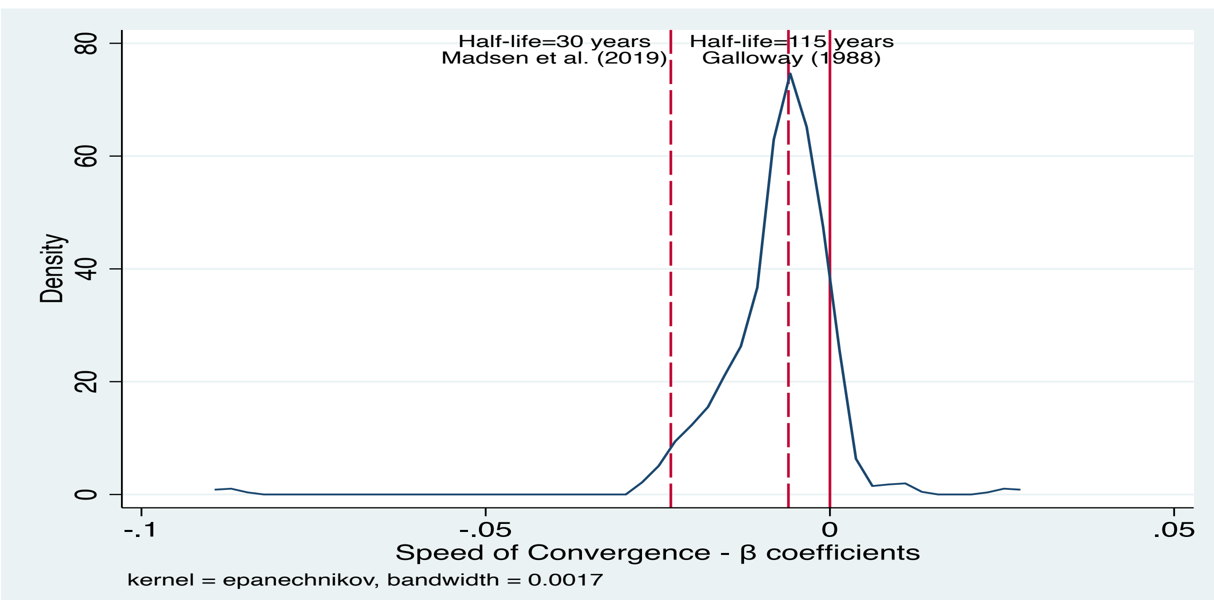
[27] Figure D-2 of the Appendix shows the point estimates along with their 95% confidence intervals for the 185 cities in the sample.

Figure 9: Speed of Convergence per period using Reba et al. (2016) Data



Notes: This figure plots estimates of the speed of convergence by period using urban population data from Reba et al. (2016). It corresponds to the FE estimations in column 2, Table 6 (left panel) and in column 7, Table 6 (right panel), adding year-interacted lagged urban population levels as controls. 95% confidence intervals reported.

Figure 10: Speed of Convergence per city using Reba et al. (2016) Data



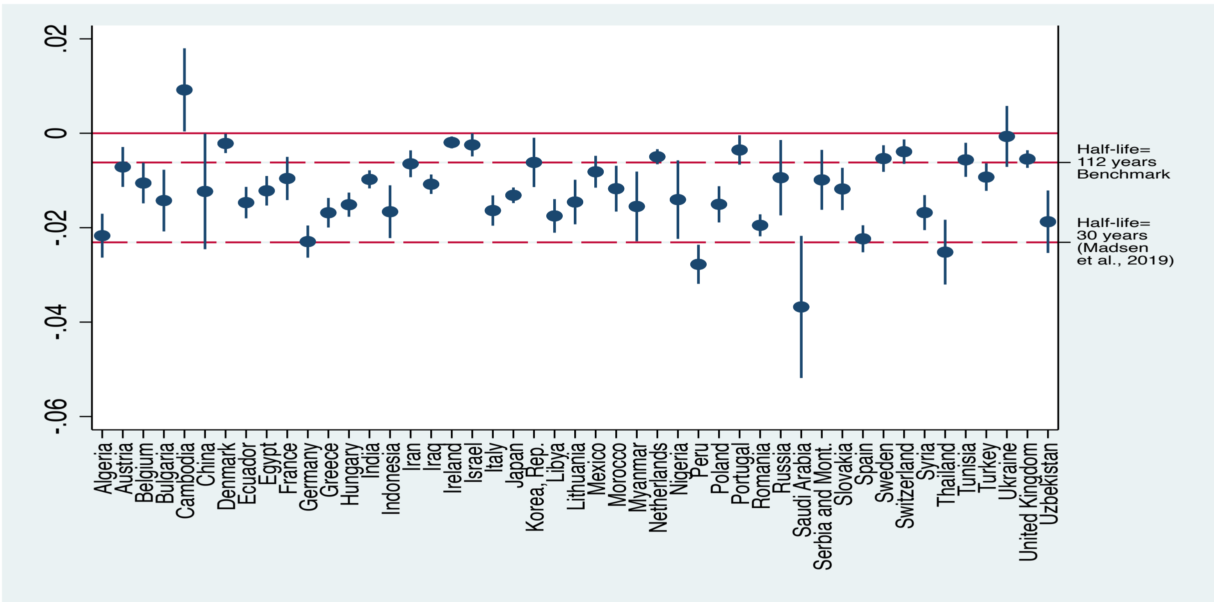
Notes: This figure plots the kernel density of the estimated speed of convergence by city using urban population data from Reba et al. (2016). It corresponds to the FE estimation in column 2, Table 6, adding city-interacted lagged urban population levels as controls.

in the strength of the Malthusian trap across countries. Some Malthusian economies, such as Algeria, Germany, Peru and Thailand are found under a strong Malthusian trap regime, with half-lives close to 30 years. The results also confirm that Spain has one of the strongest Malthusian trap, with an half-life of 31 years, and a 95% confidence interval between 28 and 35 years. This result is even stronger than the previous estimate based on total population figures of McEvedy et al. (1978).

On the other hand, countries such as Denmark, Israel and Portugal are found with the weakest estimated Malthusian trap. For instance, the estimated half-life for Denmark is 322 years. Contrary to the previous findings using population data of McEvedy et al. (1978), I find that Japan converges at a much faster pace in the present estimation, with an half-life of 53 years.

The estimated half-life for the United Kingdom is 127 years, with a 95% confidence interval between 95 and 190 years. This is very similar to the result of my benchmark calibration (112 years) in Section 3.2, giving further evidence about the relative weakness of the Malthusian trap in England.

Figure 11: Speed of Convergence per country using Reba et al. (2016) Data



Notes: This figure plots estimates of the speed of convergence by country urban population data from Reba et al. (2016). It corresponds to the FE estimation in column 7, Table 6, adding country-interacted lagged urban population levels as controls. 95% confidence intervals are reported.

7 Conclusion

The Malthusian trap has been recognized as one of the main obstacles to sustained economic growth before industrialization (Hansen and Prescott, 2002; Clark, 2007; Galor, 2011). Despite its wide acceptance, little consensus exists on the exact strength and widespread of the Malthusian trap. A first group of studies, mostly focused on England, find evidence of a weak Malthusian trap, referred to as weak *homeostasis* (Lee, 1993; Lee and Anderson, 2002; Crafts and Mills, 2009; Fernihough, 2013; Bouscasse et al., 2023). On the other hand, Madsen et al. (2019) find evidence of a strong and widespread Malthusian trap, or strong *homeostasis*.

This article brings new answers to that ongoing debate, providing the first evidence on the widespread of weak *homeostasis* over time and in a large number of Malthusian economies. I provide two type of analysis. First, I build an overlapping-generations Malthusian growth model and derive the speed of convergence to the steady-state. Compared to the existing literature, the model adds marriage (the extensive margin of fertility) as an additional channel through which population adjusts to economic shocks, as originally advanced by Malthus (1798). The speed at which a Malthusian economy adjusts to shocks is governed by four elasticity parameters: the land share of output and the elasticities of fertility, marriage and survival with respect to income per capita. I calibrate the model for the English Malthusian economy and show that the Malthusian trap was relatively weak in England, with a half-life of 112 years. I also provide a quantitative analysis showing that this pattern is compatible with the centuries long reaction of the English Malthusian economy to the Black Death.

Second, I provide empirical estimates of the speed of convergence, using the familiar concept of β-convergence and historical panel data on per capita incomes and population. The data employed cover a wide range of the Malthusian period and a large set of economies. I first use two standard source of data on per capita incomes and population levels for the Malthusian period: the Maddison Project database and McEvedy et al. (1978). In addition, I employ simulated GDP per capita series from Lagerlöf (2019), and historical urban population from Chandler (1987) and Modelski (2003). These sources have a much greater cross-sectional and time dimension, enabling me to increase the precision of convergence speed estimates and explore spatial differences in a more comprehensive way than other datasets. Across my estimations, I find consistent evidence of weak *homeostasis*, with the mode of the half-lives distribution close to 120 years. While I find a relative stability of the weak *homeostasis* pattern over time, from the 11th to the 18th century, I also find significant differences in the strength of the Malthusian trap across countries, with some economies compatible with a strong *homeostasis* and some countries compatible with weaker *homeostasis* compared to England. Using two-way fixed effects, a time-varying control variable that captures institutional changes, and instrumental variable techniques (GMM), the analysis tackles the omitted variable bias that could result in an estimated weak Malthusian trap.

Overall, my results contribute to a better understanding of pre-industrial economic fluctuations. Reconstructed historical series of GDP per capita for the Malthusian period often exhibit persistence, attested by long cycles of expansion and contraction following shocks. A relatively weak Malthusian trap, or weak *homeostasis*, is required to reproduce this persistence, while remaining compatible with the fact that the Malthusian trap ultimately prevented per capita income growth in the long run. This study also provides new evidence on the differences in the strength of the Malthusian trap between countries. The study of the cultural or institutional determinants of these differences and their implications for the transition of economies out of the Malthusian trap is a fruitful area for future research.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D. and Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity and Poverty*. Crown, New York, 1st edition.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277–297.

Ashraf, Q. and Galor, O. (2011). Dynamics and Stagnation in the Malthusian Epoch. *American Economic Review*, 101(5):2003–41.

Ashraf, Q. and Galor, O. (2013). The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review*, 103(1):1–46.

Attar, M. A. (2023). Technology and survival in preindustrial england: a malthusian view: A malthusian view. *Journal of Population Economics*, pages 1–40.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, 106(2):407–443.

Barro, R. J. (2015). Convergence and Modernisation. *The Economic Journal*, 125(585):911–942.

Barro, R. J. and Sala-i Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2):223–251.

Baudin, T., de la Croix, D., and Gobbi, P. E. (2015). Fertility and Childlessness in the United States. *American Economic Review*, 105(6):1852–82.

Benedictow, O. J., Benedictow, O. L., et al. (2004). *The Black Death, 1346-1353: The Complete History*. Boydell & Brewer.

Blundell, R. and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 87(1):115–143.

Bolt, J. and Van Zanden, J. L. (2020). Maddison Project Database, version 2020. *Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update*.

Borcan, O., Olsson, O., and Putterman, L. (2018). State History and Economic Development:

Evidence from Six Millennia. *Journal of Economic Growth*, 23(1):1–40.

Bouscasse, P., Nakamura, E., and Steinsson, J. (2023). When Did Growth Begin? New Estimates of Productivity Growth in England from 1250 to 1870. Unpublished manuscript.

Broadberry, S., Campbell, B. M., Klein, A., Overton, M., and Van Leeuwen, B. (2015). *British economic growth, 1270–1870*. Cambridge University Press.

Caselli, F., Esquivel, G., and Lefort, F. (1996). Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics. *Journal of Economic Growth*, 1(3):363–389.

Chandler, T. (1987). *Four Thousand Years of Urban Growth: An Historical Census*. Mellen.

Cinnirella, F., Klemp, M., and Weisdorf, J. (2017). Malthus in the Bedroom: Birth Spacing as Birth control in Pre-Transition England. *Demography*, 54(2):413–436.

Clark, G. (2007). *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press.

Crafts, N. and Mills, T. C. (2009). From Malthus to Solow: How Did the Malthusian Economy Really Evolve? *Journal of Macroeconomics*, 31(1):68–93.

Cummins, N. (2020). The Micro-Evidence for the Malthusian System. France, 1670–1840. *European Economic Review*, 129:103544.

de la Croix, D. and Gobbi, P. E. (2017). Population Density, Fertility, and Demographic Convergence in Developing Countries. *Journal of Development Economics*, 127:13–24.

de la Croix, D. and Gobbi, P. E. (2022). Population Homeostasis in Sub-Saharan Africa. *Economics & Human Biology*, 45:101102.

de la Croix, D., Schneider, E. B., and Weisdorf, J. (2019). Childlessness, Celibacy and Net Fertility in Pre-Industrial England: The Middle-Class Evolutionary Advantage. *Journal of Economic Growth*, 24(3):223–256.

Dennison, T. and Ogilvie, S. (2014). Does the European Marriage Pattern Explain Economic Growth? *The Journal of Economic History*, 74(3):651–693.

Federico, G., Nuvolari, A., and Vasta, M. (2020). Inequality in Pre-Industrial Europe (1260–1850): New Evidence From the Labor Share. *Review of Income and Wealth*.

Fernihough, A. (2013). Malthusian Dynamics in a Diverging Europe: Northern Italy, 1650–1881. *Demography*, 50(1):311–332.

Fouquet, R. and Broadberry, S. (2015). Seven Centuries of European Economic Growth and Decline. *Journal of Economic Perspectives*, 29(4):227–44.

Galloway, P. R. (1988). Basic Patterns in Annual Variations in Fertility, Nuptiality, Mortality, and Prices in Pre-Industrial Europe. *Population Studies*, 42(2):275–303.

Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.

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

Guinnane, T. W. (2021). We Do Not Know the Population of Every Country in the World for the past Two Thousand Years. CESifo Working Paper, No. 9242, Center for Economic Studies and ifo Institute (CESifo), Munich.

Hajnal, J. (1965). "European Marriage Pattern in Perspective". W: Population in History: Essays in Historical Demography, red. David V. Glass, David EC Eversley, 101–143. Chicago: Aldine Publishing Company.

Hansen, G. D. and Prescott, E. C. (2002). Malthus to Solow. *American Economic Review*, 92(4):1205–1217.

Hauk, W. R. (2017). Endogeneity Bias and Growth Regressions. *Journal of Macroeconomics*, 51:143–161.

Hauk, W. R. and Wacziarg, R. (2009). A Monte Carlo Study of Growth Regressions. *Journal of Economic Growth*, 14(2):103–147.

Hurwicz, L. (1950). Least Squares Bias in Time Series. *Statistical Inference in Dynamic Economic Models*, 10:365–383.

Irmen, A. (2004). Malthus and Solow-a note on Closed-Form Solutions. *Economics Bulletin*, 10(6):1–6.

Islam, N. (1995). Growth Empirics: A Panel Data Approach. *The Quarterly Journal of Economics*, 110(4):1127–1170.

Jedwab, R., Johnson, N. D., and Koyama, M. (2022). The Economic Impact of the Black Death. *Journal of Economic Literature*, 60(1):132–78.

Jensen, P. S., Pedersen, M. U., Radu, C. V., and Sharp, P. R. (2021). Arresting the Sword of Damocles: The Transition to the Post-Malthusian Era in Denmark. *Explorations in Economic*

History, page 101437.

Kelly, M. and Gráda, C. Ó. (2012). The Preventive Check in Medieval and Preindustrial England. *The Journal of Economic History*, 72(4):1015–1035.

Klemp, M. and Møller, N. F. (2016). Post-Malthusian Dynamics in Pre-Industrial Scandinavia. *The Scandinavian Journal of Economics*, 118(4):841–867.

Klemp, M. P. (2012). Prices, Wages and Fertility in Pre-Industrial England. *Cliometrica*, 6(1):63–77.

Kremer, M. (1993). Population Growth and Technological Change: One Million BC to 1990. *The Quarterly Journal of Economics*, 108(3):681–716.

Ladurie, E. L. R. (1977). Motionless History. *Social Science History*, 1(2):115–136.

Lagerlöf, N.-P. (2015). Malthus in Sweden. *The Scandinavian Journal of Economics*, 117(4):1091–1133.

Lagerlöf, N.-P. (2019). Understanding Per-Capita Income Growth In Preindustrial Europe. *International Economic Review*, 60(1):219–240.

Lee, R. (1993). Accidental and Systematic Change in Population History: Homeostasis in a Stochastic Setting. *Explorations in Economic History*, 30(1):1–30.

Lee, R. and Anderson, M. (2002). Malthus in State Space: Macro Economic-Demographic Relations in English History, 1540 to 1870. *Journal of Population Economics*, 15(2):195–220.

Lee, R. D. (1981). Short-Term Variation: Vital Rates, Prices and Weather. *The Population History of England*, 1541(1871):356–401.

Lee, R. D. (1987). Population Dynamics of Humans and other Animals. *Demography*, pages 443–465.

Maddison, A. (2003). *Development Centre Studies the World Economy Historical Statistics: Historical Statistics*. OECD publishing.

Madsen, J. B., Robertson, P. E., and Ye, L. (2019). Malthus Was Right: Explaining a Millennium of Stagnation. *European Economic Review*, 118:51–68.

Malthus, T. (1798). *An Essay on the Principle of Population*. J. Johnson, London.

McEvedy, C., Jones, R., et al. (1978). *Atlas of World Population History*. Penguin Books Ltd, Harmondsworth, Middlesex, England.

Modelski, G. (2003). *World cities:-3000 to 2000*. Washington, DC Faros.

Møller, N. F. and Sharp, P. (2014). Malthus in Cointegration Space: Evidence of a Post-Malthusian Pre-Industrial England. *Journal of Economic Growth*, 19(1):105–140.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Nicolini, E. A. (2007). Was Malthus Right? A VAR Analysis of Economic and Demographic Interactions in Pre-Industrial England. *European Review of Economic History*, 11(1):99–121.

North, D. C. and Thomas, R. P. (1973). *The Rise of the Western World: A New Economic History*. Cambridge University Press.

Nunn, N. (2008). The Long-Term Effects of Africa's Slave Trades. *The Quarterly Journal of Economics*, 123(1):139–176.

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

Pfister, U. and Fertig, G. (2020). From Malthusian Disequilibrium to the Post-Malthusian Era: The Evolution of the Preventive and Positive Checks in Germany, 1730–1870. *Demography*, 57(3):1145–1170.

Reba, M., Reitsma, F., and Seto, K. C. (2016). Spatializing 6,000 years of global urbanization from 3700 bc to ad 2000. *Scientific Data*, 3(1):1–16.

Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.

Segalen, M. and Fine, A. (1988). VIII-Famille et Mariage. In *Histoire de la Population Française (3)*, pages 403–469. Presses Universitaires de France.

Szulga, R. S. (2012). Malthus vs. History: Speed of Convergence in the Malthusian World. *SSRN Working Paper*.

Wrigley, E. A. (1968). Mortality in Pre-Industrial England: The Example of Colyton, Devon, over Three Centuries. *Daedalus*, pages 546–580.

Wrigley, E. A., Davies, R. S., Oeppen, J. E., and Schofield, R. S. (1997). *English Population History from Family Reconstitution 1580-1837*. Cambridge University Press.

Wrigley, E. A., Schofield, R. S., and Schofield, R. (1989). *The Population History of England 1541-1871*. Cambridge University Press.

Appendix

A Uniqueness and Stability of the Steady State

In this section, I analyze in more depth the dynamic of the Malthusian economy described in Section 2. In particular, I show that the economy has a unique, locally stable, and positive steady state.

The dynamic of the economy is given by the first-order difference equation (12), that I further develop here:

$$\psi(y_t) = \left(\frac{1+g}{\left(\kappa^{-\delta} \ \left(y_t + q \ \nu^{1/\delta}\right)^{\delta} \ - \nu\right) \cdot \frac{(\overline{\lambda}(y_t)-1)}{(b-1)} \cdot \underline{s} \ y_t^{\phi}} \right)^{\alpha} \cdot y_t \ ,$$

with $\overline{\lambda}(y_t) = \exp\left(\ln(1-\frac{q}{\kappa}) + (1+\delta\gamma)\ln(y_t + q\ \nu^{1/\delta}) - \delta\gamma\ln(\kappa) - \ln(y_t) - \gamma\ln(\nu)\right).$

Uniqueness of the Steady State. — The model admits a unique state state $y^* > 0$ for all $y_t \in [0, +\infty[$. As stated by equation (13), the steady-state level of income per capita y^* is attained when:

$$y^* \equiv \left(\frac{1+g}{n(y^*)\ p(y^*)\ s(y^*)} \right)^{\alpha} = 1 \ .$$

There is only one steady state for the economy because $n_t(y_t)$, $p_t(y_t)$ and $s_t(y_t)$ are increasing and montononeous functions of y_t, given $y_t > \underline{y} > 0$. Therefore, there is only one value $y_t > 0$ that satisfies the steady-state condition (13) and equates $(1+g) > 0$.

To see this, lets first define $\underline{y}$, the minimum income per capita level for which $n_t(y_t)$, $p_t(y_t)$ and $s_t(y_t)$ start to have meaningful values – i.e. $(n_t, p_t, s_t) \in \mathbb{R}_{++}$. For $n_t(y_t)$, we have that:

$$n_t(\underline{y}) = 0 \Leftrightarrow \underline{y} = \frac{q}{\gamma\delta} \cdot \nu^{1/\delta} > 0 \ ,$$

since q, δ, γ and ν are strictly positive.

Similarly, for $p_t(y_t)$, we have:

$$p_t(\underline{y}) = \frac{\lambda(\underline{y}) - 1}{b-1} = \frac{1-1}{b-1} = 0 \ .$$

For $s_t(y_t)$, we have:

$$s_t(\underline{y}) = \underline{s}(\frac{q}{\gamma\delta})^{\phi} \cdot \nu^{\phi/\delta} > 0 \ ,$$

since $\phi > 0$.

These functions are strictly increasing for all $y_t \in [0, +\infty[$, except p_t that is strictly increasing for all $y_t > \underline{y}$:

$$n'_t(y_t) = \delta\kappa(y_t + q\nu^{1/\delta})^{\delta-1} > 0 \ .$$

$$p'_t(y_t) = \left(\frac{1+\delta\gamma}{(y_t + q\nu^{1/\delta})} - \frac{1}{y_t} \right) \cdot \overline{\lambda}(y_t) > 0 \text{ if } y_t > \underline{y} \ .$$

$$s'_t(y_t) = \phi \underline{s} y_t^{\phi-1} > 0 \ .$$

Therefore, it exists one and only one value $y_t > \underline{y}$ that solves equation (13), and that value is a steady state $y^* > 0$.

Levels of income per capita $y_t \in [0, \underline{y}]$ result in unbounded dynamics. To prevent such cases in the quantitative exercise of Section 3, I make the economy starts at an arbitrarily large steady-state level $y^* = 1 \cdot 10^6$. This is well above the value of $\underline{y}$ in my simulations, which is $\underline{y} = 0.051$.

Stability of the Steady State. — The steady state y^* is locally stable. The steady state y^* is stable if the absolute value of the derivative of $\psi(y_t)$ evaluated at the steady state is in the unit circle.

The first derivative of $\psi(y_t)$ with respect to y_t gives:

$$\psi'(y_t) = \left(\frac{(1+g)}{n_t(y_t)\ p_t(y_t)\ s_t(y_t)} \right)^{\alpha} \cdot \ (1 - \alpha\,(\epsilon_{n_t} + \epsilon_{p_t} + \epsilon_{s_t})) \ ,$$

with $\epsilon_{n_t} = n'_t(y_t)/n_t \cdot y_t$, $\epsilon_{p_t} = p'_t(y_t)/p_t \cdot y_t$ and $\epsilon_{s_t} = s'_t(y_t)/s_t \cdot y_t$ elasticities of the income per capita to fertility, marriage and survival respectively.

At the steady state, we have:

$$\begin{aligned} \psi'(y^*) &= \left(\frac{(1+g)}{n_t(y^*)\ p_t(y^*)\ s_t(y^*)} \right)^{\alpha} \cdot \ \left(1 - \alpha\left(\epsilon^*_{n_t} + \epsilon^*_{p_t} + \epsilon^*_{s_t}\right)\right) \\ &= 1 - \alpha\left(\epsilon^*_{n_t} + \epsilon^*_{p_t} + \epsilon^*_{s_t}\right) \ . \end{aligned}$$

In the quantitative exercise of Section 3, $1 - \alpha(\epsilon^*_{n_t} + \epsilon^*_{p_t} + \epsilon^*_{s_t}) = 0.84$, ensuring that the steady state y^* is locally stable.

B Calibration of Elasticity Parameters

In this section, I detail the sources and the method used to calibrate the elasticity parameters controlling the preventive checks (δ), and the positive checks (ϕ) in the model described in Section 3.

Since my model is written at the generation level, I focused on studies estimating the long-run elasticities of the preventive and positive checks for the English Malthusian economy. I considered the six studies listed in the first column of Table B-1. When available, I collected the long-run elasticity estimates by sub-period, in order to have more variation for the Malthusian period. I have avoided including data after 1800, as they are less likely to be representative of Malthusian dynamics (especially for the English economy). This was always possible in the studies considered except for Lee and Anderson (2002), where estimates are only available for the period 1540-1870. Unfortunately, elasticity estimates are not available for England before the 16th century. This is because all studies rely on the same demographic series (birth, death and marriage rates) compiled by Wrigley et al. (1989), and starting in the year 1540.

Table B-1: Preventive and Positive Checks Elasticities considered for the Calibration

Article	Model	Data (Population/Wage)	Period	Preventive Check Elasticity	Positive Check Elasticity
Crafts and Mills (2009)	State space	WS/C	1541–1645	0.31	0.24
	State space	WS/C	1646–1799	0.22	+
	SVAR	WS/C	1541–1645	0.09	0.03
	SVAR	WS/C	1646–1799	0.23	+
Klemp (2012)	CVAR	WS/C	1701–1759	0.31	N.A.
Lee and Anderson (2002)	State space	WS/PBH	1540–1870	0.12	0.08
Lee (1981)	OLS	WS/PBH	1548–1834	0.14	0.1
Møller and Sharp (2014)	CVAR	WS/A	1564–1760	0.32	0.1
	CVAR	WS/C	1564–1760	0.21	0.22
Nicolini (2007)	SVAR	WS/A	1541–1640	0.03	0.11
	SVAR	WS/A	1641–1740	0.11	+

Notes: This table presents the source I used to calibrate the elasticities parameters in my model in Section 5. Column 3 gives the source of the population and wage data used in each paper mentioned. WS indicates that the population data comes from Wrigley et al. (1989), C indicates that the wage data comes from Clark (2007), PBH indicates that the wage data comes from Hopkins (1957), and A indicates that the wage data comes from Allen (2001). + indicates sub-period for which the positive checks were estimated with a counter-intuitive sign.

The two last columns of Table B-1 indicate the elasticity values taken into account to calibrate my model. I was able to collect 10 long-run elasticity estimates for the preventive checks and 7 for the positive checks. I have not included the positive check elasticity values if they were find with the "wrong" sign in the studies, which I indicate by a "+" in Table B-1. Table B-1 also provides information on the data and the model used to estimate the elasticities in each article, which could explain some of the differences in estimates across studies.

The value of the long-run elasticity of the positive checks is directly given by ϕ in my model, as equation (10) corresponds to the unit-elastic case. Therefore, I set ϕ directly to the median of the elasticities provided by the aforementioned studies for the benchmark specification – i.e. $\phi = 0.1$.

Concerning the preventive checks, I fix δ such that the sum of the elasticities of fertility and marriage with respect to income per capita in my model is equal to the median of the long-run elasticity of the preventive checks provided by the aforementioned literature. In Table B-1, the median elasticity for the preventive checks is 0.21. Therefore, I solve $\epsilon^*_{n_t} + \epsilon^*_{p_t} = 0.21$ for δ to calibrate my model. It implies that $\delta = 0.074$ in my benchmark specification.

Table B-2 provides the elasticities ϵ_{n_t}, ϵ_{p_t} and ϵ_{s_t} estimated using the data generated by model in the quantitative exercise of Section 3. This is a way of checking the accuracy of my calibration strategy. Table B-2 shows that the elasticities of the preventive and positive checks are successfully calibrated, both for my benchmark specification and for the two alternative calibrations considered. In particular, the sum of the elasticity of ϵ_{n_t} in column 1 and ϵ_{p_t} in column 7 gives $0.104 + 0.107 = 0.21$. This matches the the target value for preventive checks, validating my calibration strategy.

Table B-2: Estimated Long-Run Elasticities from Quantitative Analysis of Section 3

Dependent Variable:	$\log(n_t)$			$\log(s_t)$			$\log(p_t)$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(y_t)$ – Benchmark (Median)	0.1036***			0.1000***			0.1068***		
	(0.000)			(0.000)			(0.000)		
$\log(y_t)$ – Alternative Specification (90th percentile)		0.1521***			0.2300***			0.1622***	
		(0.000)			(0.000)			(0.000)	
$\log(y_t)$ – Alternative Specification (10th percentile)			0.0452***			0.0600***			0.0452***
			(0.000)			(0.000)			(0.000)
Observations	21	21	21	21	21	21	21	21	21

Notes: This table presents estimates of the long-run elasticities of fertility, survival rate and marriage rate with respect to income per capita, using the simulated data of Section 3. Standard errors robust against heteroskedasticity are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

C Derivation of the Speed of Convergence

Taking a first-order Taylor expansion of $\psi(y_t)$ around y^*, we have:

$$
\begin{aligned}
\psi(y_t) &\approx \psi(y^*) + \psi'(y^*) \cdot (y_t - y^*) \\
y_{t+1} &\approx y_t - \alpha(\epsilon_{n_t} + \epsilon_{p_t} + \epsilon_{s_t}) \cdot (y_t - y^*) \ .
\end{aligned}
$$

It follows that GDP per capita growth rate at the neighbourhood of the steady state is:

$$
g^y \equiv \frac{y_{t+1} - y_t}{y_t} \approx -\beta^* \cdot (\ln y_t - \ln y^*) \ , \tag{C-1}
$$

with $\beta^* = \alpha(\epsilon_{n_t} + \epsilon_{p_t} + \epsilon_{s_t})$ the speed of convergence to the steady-state.

In my model, population is not constant at the steady state but rather growth at the same pace as technology. To analyse the speed of convergence to the population steady state, I first need to express labour L_t in terms of effective units:

$$
\widehat{L}_t \equiv \frac{L_t}{A_t} \ .
$$

Recall equation (8), we can express effective units of labour as:

$$
\widehat{L}_t = y_t^{-1/\alpha} \ . \tag{C-2}
$$

Taking the logarithm of (C-2) and highlighting growth rates, we have:

$$
g^{\widehat{L}} = \frac{\partial \ln \widehat{L}_t}{\partial t} = -\frac{1}{\alpha} \frac{\partial \ln y_t}{\partial t} = -\frac{1}{\alpha} \, g^y \ . \tag{C-3}
$$

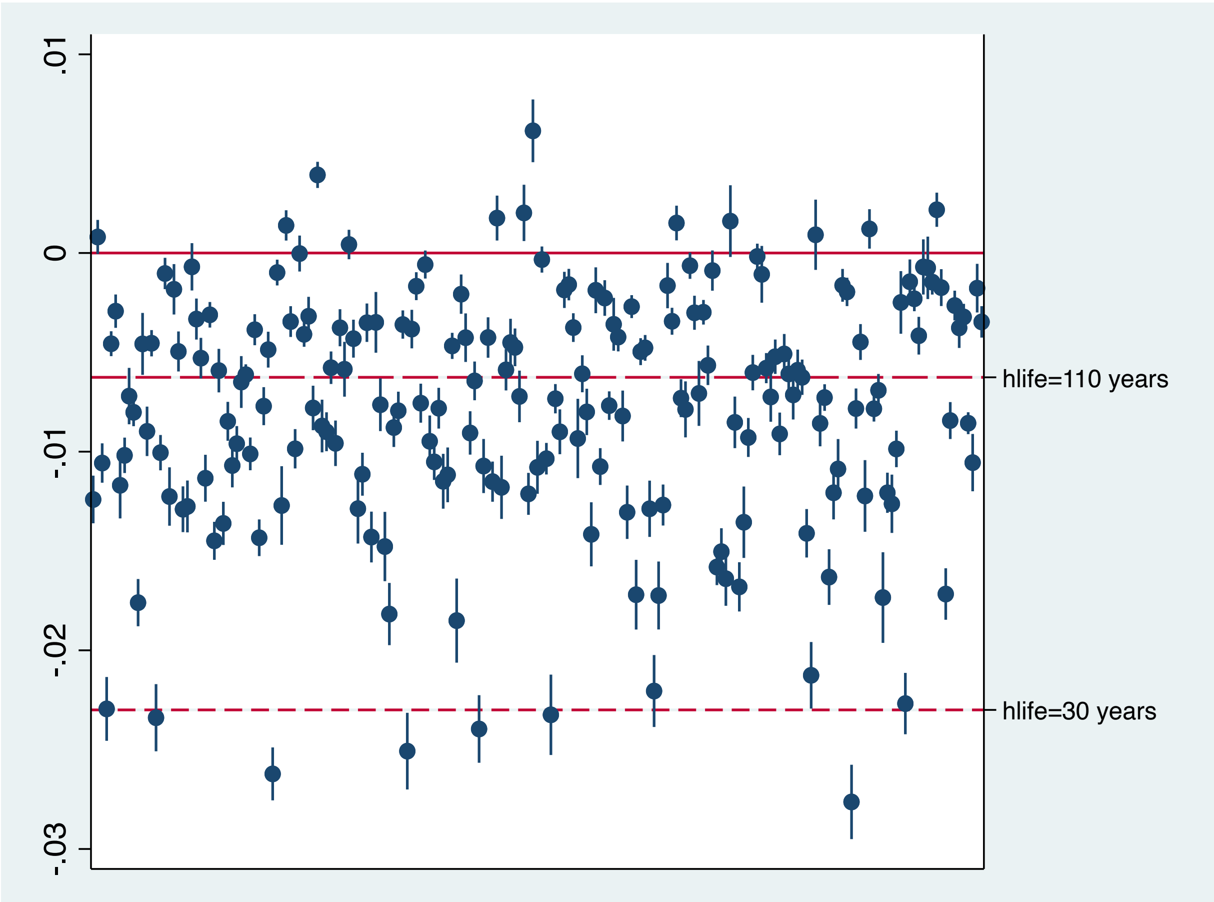
Using (C-3) and (C-1), we have:

$$
\begin{aligned}
g^{\widehat{L}} &\approx -\beta^* \ -\frac{1}{\alpha} \cdot (\ln y_t - \ln y^*) \\
g^{\widehat{L}} &\approx -\beta^* \cdot (\ln \widehat{L}_t - \ln \widehat{L}^*) \ .
\end{aligned} \tag{C-4}
$$

It means that in a Malthusian economy, effective unit of labour converges to its steady state at the same pace than GDP per capita. Consequently, once technological progress and the size of land is hold constant, population data can be used to estimate the speed of convergence of a Malthusian economy.

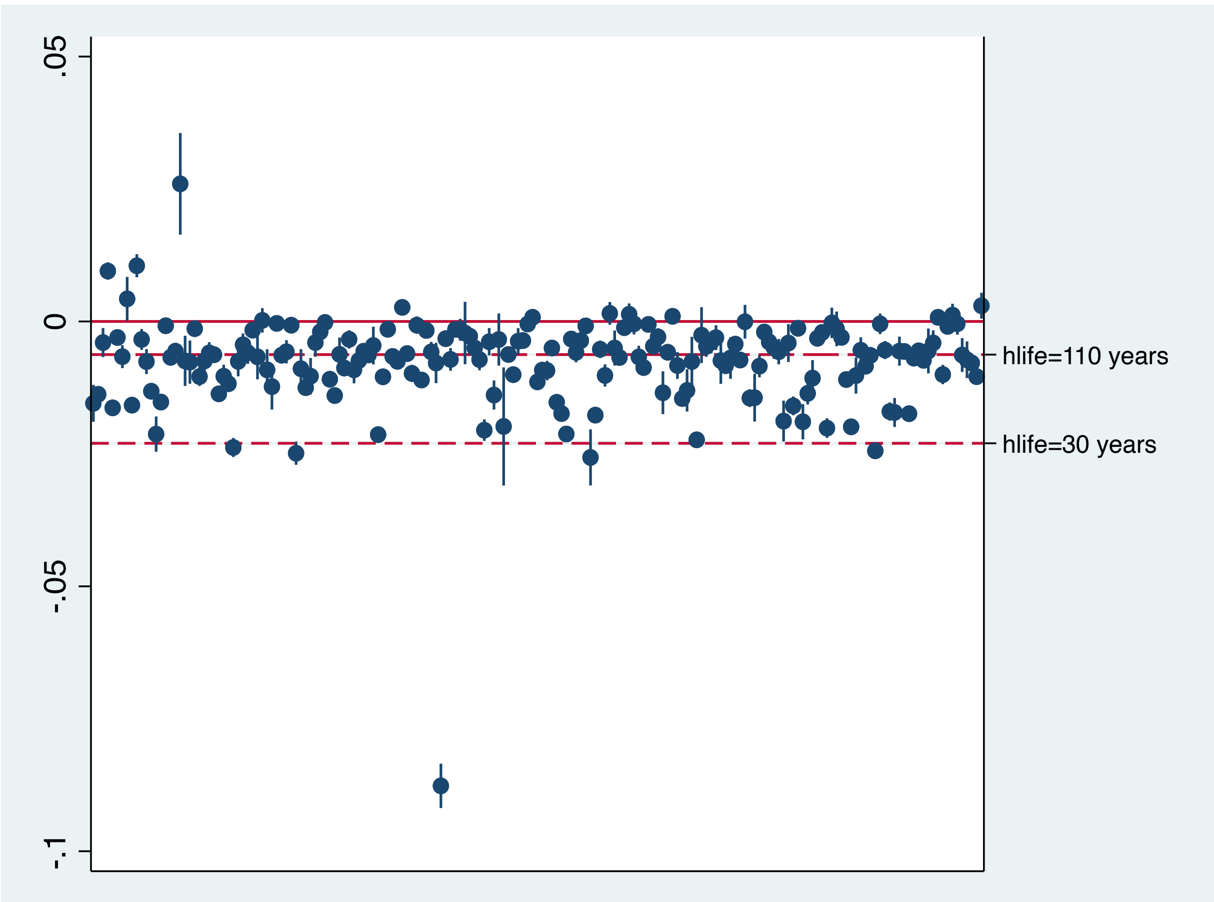
D Additional Results

Figure D-1: Speed of Convergence for the 200 first Malthusian Economies in Lagerlöf (2019)



Notes: This figure plots the estimated speed of convergence for the 200 first simulated economies in Lagerlöf (2019). It corresponds to the FE estimation in Table 4, column 2, adding country-interacted lagged GDP per capita as controls. 95% confidence intervals are reported.

Figure D-2: Speed of Convergence per city using Data from Reba et al. (2016)



Notes: This figure plots the estimated speed of convergence for the 185 cities in my sample using data from Reba et al. (2016). It corresponds to the FE estimation in Table 6, column 2, adding city-interacted lagged population levels as controls. 95% confidence intervals are reported.

Appendix references (not cited in the main text)

Allen, R. C. (2001). The Great Divergence in European Wages and Prices from the Middle Ages to the First World War. *Explorations in Economic History*, 38(4):411–447.

Clark, G. (2007). The Long March of History: Farm Wages, Population, and Economic Growth, England 1209–1869 1. *The Economic History Review*, 60(1):97–135.

Hopkins, S. V. (1957). Wage-Rates and Prices: Evidence for Population Pressure in the Sixteenth Century. *Economica*, 24(96):289–306.