



HAL
open science

Counting the missing poor in pre-industrial societies

Mathieu Lefebvre, Pierre Pestieau, Gregory Ponthiere

► **To cite this version:**

Mathieu Lefebvre, Pierre Pestieau, Gregory Ponthiere. Counting the missing poor in pre-industrial societies. *Cliometrica*, 2023, 17 (1), pp.155-183. 10.1007/s11698-022-00243-y . hal-03620370

HAL Id: hal-03620370

<https://amu.hal.science/hal-03620370>

Submitted on 5 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Counting the Missing Poor in Pre-Industrial Societies*

Mathieu Lefebvre[†] Pierre Pestieau[‡] Gregory Ponthiere[§]

November 18, 2021

Abstract

Under income-differentiated mortality, poverty measures suffer from a selection bias: they do not count the missing poor (i.e. persons who would have been counted as poor provided they did not die prematurely). The Pre-Industrial period being characterized by an evolutionary advantage (i.e. a higher number of surviving children per household) of the non-poor over the poor, one may expect that the missing poor bias is substantial during that period. This paper quantifies the missing poor bias in Pre-Industrial societies, by computing the hypothetical headcount poverty rates that would have prevailed provided the non-poor did not benefit from an evolutionary advantage over the poor. Using data on Pre-Industrial England and France, we show that the sign and size of the missing poor bias is sensitive to the degree of downward social mobility.

Keywords: poverty, measurement, selection effects, missing poor.

JEL classification codes: I32.

*The authors would like to thank Claude Diebolt, Oded Galor and two anonymous reviewers, as well as participants of seminar at Université Saint Louis (Brussels), for their comments on this paper.

[†]Aix-Marseille Université, CNRS, AMSE, Marseille, France.

[‡]University of Liege, CORE and PSE.

[§]UCLouvain, Hoover Chair in Economic and Social Ethics. Address: Collège Dupriez, 3 Place Montesquieu, 1348 Louvain-La-Neuve, Belgium. E-mail: gregory.ponthiere@uclouvain.be

1 Introduction

Although debates persist about the measurement of standards of living across long periods, it is widely acknowledged that the prevalence of poverty was high in Pre-Industrial societies (Lindert and Williamson 1982, Clark 2007, Ravallion 2016). Pre-Industrial standards of living were fluctuating with wars and epidemics, but the overall tendency involved a large prevalence of poverty. As shown on Table 1, which presents figures from King's social tables on England and Wales (King 1688, in Barnett 1936) as well as corrections by Lindert and Williamson (1982), about half of households had, in 1688, an average income representing less than 50 % of the average household income.¹

	King's social table for England and Wales (1688)		Revised table (Lindert and Williamson 1982)	
	No of families	Av. family income (£)	No of families	Av. family income (£)
Total population including	1,390,586	31,29	1,390,586	39,18
Common soldiers	35,000	14	35,000	14
Miners	-	-	14,240	15
Laboring people/outservants	364,000	15	284,997	15
Cottagers/paupers	400,000	6,5	313,183	6,5
Vagrants	30,000	2	23,489	2
% of families with income < 50 % of average income	59,6 %		48,2 %	

Table 1: Poverty in Pre-Industrial England based on King's social tables.

Another key aspect of Pre-Industrial societies lies in the existence of what can be called an "evolutionary advantage" for the rich with respect to the poor, in the sense that the former have a larger number of surviving offspring than the latter.² As shown by Clark and Hamilton (2006) on the basis of data from wills on reproductive success in England (1585-1638), the richest testators had about twice as many surviving children as the poorest. Figure 1, which is based on Clark and Cummins (2015), illustrates the evolutionary advantage of the rich: over 1500-1779, the average number of surviving children per women (i.e., the "net fertility") equals 4.1 children for the top asset income tercile, against 2.8 children for the bottom tercile. These findings support a thesis defended by

¹Besides the discussion of King's table (1688), Lindert and Williamson (1982) provide also a critical presentation of other social tables, such as the ones of Joseph Massie (1759) and Patrick Colquhoun (1801-1803). Lindert and Williamson (1983) revisits the social table of Dudley Baxter (1868), and provide a picture of England's development during two centuries.

²Several mechanisms may explain why the rich had, at that time, a higher number of surviving children than the poor. These mechanisms include the various channels through which material living conditions affect mortality and fertility outcomes. Throughout this paper, we will take that evolutionary advantage *as given* and examine its impact on poverty measurement, without trying to identify the particular mechanisms at work behind this advantage.

Malthus (1798): the existence of positive and preventive "population checks" adjusting the population size to the means of subsistence. Preventive population checks are fertility reductions aimed at avoiding famines, while positive population checks consist of a worsening of survival conditions of the poor due to the lack of means of subsistence.³ The smaller number of surviving children among the poor is compatible with the existence of population checks.⁴

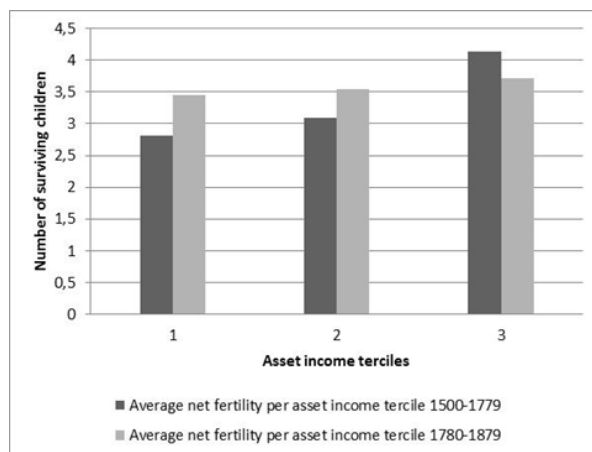


Figure 1: Net fertility by asset income tercile in England, Clark and Cummins (2015, Table 8).

The existence of an evolutionary advantage for the non-poor with respect to the poor raises some paradoxes for the measurement of poverty. As stressed by Kanbur and Mukherjee (2007), standard poverty measures suffer, under income-differentiated mortality, from a *selection bias*: poor individuals, facing worse survival conditions than non-poor ones, are under-represented in the studied populations, which pushes poverty measures downwards. Following Sen's (1998) emphasis on "missing women", this selection bias can be called the "missing poor bias". When the strength of selection bias varies across countries, the missing poor bias tends to create noise in international comparisons of poverty.

In an intergenerational context, selection biases have *cumulative* effects. The missing poor, i.e., the unborn or prematurely dead persons who would have been counted as poor provided these were alive, would, provided they were born or had been long-lived, have had children (or more children), who would have had their own children, etc. All those missing persons ought to be taken into account to avoid selection biases. The problem of cumulative measurement errors due to repeated selection biases may be particularly acute for the Pre-Industrial era, during which strong selection effects were cumulated across centuries.

³See Malthus (1798), chapter 1, p. 71-72 and chapter 5, p. 93.

⁴Note that other works evaluating the Malthusian doctrine yield mixed results. See Wrigley (1969), Mokyr (1980), Clark (2007), Crafts and Mills (2009) and Cummins (2020).

The goal of this paper is to explore the consequences of the evolutionary advantage of the non-poor over the poor for the measurement of poverty during the Pre-Industrial period, and to provide a quantification of the size of the missing poor bias in Pre-Industrial societies. For that purpose, we develop a simple matrix population model, where the population is partitioned into poor and non-poor individuals, who differ in terms of fertility, mortality and social mobility. This model allows us to characterize the long-run headcount poverty rate, and to study analytically how the degree of evolutionary advantage of the non-poor over the poor affects the measurement of poverty. We then calibrate that model and compare the actual prevalence of poverty in Pre-Industrial England and France with the hypothetical prevalence that would have obtained provided the evolutionary advantage of the non-poor over the poor is set to 0.

Anticipating our results, a first finding lies on the theoretical side: it is not necessarily the case that a higher evolutionary advantage for the non-poor over the poor pushes measured poverty down. When downward social mobility is high, a stronger evolutionary advantage for the non-poor can *increase* measured poverty. Under these conditions, the evolutionary advantage of the non-poor leads to an overestimation of poverty. Second, the comparison of the standard poverty rate and the hypothetical poverty rate (where the evolutionary advantage of the non-poor is set to zero) for Pre-Industrial England reveals that the size of the missing poor bias varies from *minus* 1 percentage point to *plus* 50 percentage points, depending on the degree of downward social mobility. Regarding Pre-Industrial France, the gaps between actual and hypothetical headcount poverty rates are smaller, because of weaker income-based selection effects in that country. The varying size of income-based selection effects across countries can affect international comparisons in terms of poverty, to an extent that depends on downward social mobility.

The paper is organized as follows. Section 2 reviews the literature. Section 3 presents our matrix population model. The long-run headcount poverty rate is derived in Section 4. Section 5 studies the effect of the evolutionary advantage of the non-poor over the poor on the measurement of poverty in the long-run. Section 6 presents our method for the adjustment of the headcount poverty rate so as to count all the missing poor individuals and their descendants. Section 7 uses data on poverty in Pre-Industrial England to compare the actual poverty rate with the hypothetical poverty that would have prevailed provided no sub-population had enjoyed an evolutionary advantage. Section 8 makes similar comparisons for Pre-Industrial France. Conclusions are left to Section 9.

2 Literature review

This paper is related to several branches of the literature.

The missing poor Our work is related to the literature on the measurement of poverty under income-differentiated mortality. That literature took off with the article of Kanbur and Mukherjee (2007), to which we referred above.

Since poor individuals face excess mortality in comparison to the non-poor, they are under-represented in the population under study. In order to quantify the "missing poor" bias, Lefebvre et al (2013) compare actual old-age poverty rates with hypothetical old-age poverty rates computed by assuming that all individuals enjoy the survival conditions of the top income class, and by assigning to prematurely dead persons fictitious incomes equal to the last income enjoyed when being alive. Using the same method, Lefebvre et al (2018) showed that international comparisons of poverty are affected by the existence of income-mortality gradients of various sizes across countries.⁵

Note that Kanbur and Mukherjee (2007) and Lefebvre et al (2013, 2018) ignore income mobility, i.e., the fact that prematurely dead poor individuals could have escaped from poverty in the hypothetical case of their survival. This may lead to overestimate or underestimate the size of the missing poor bias. To deal with that issue, Lefebvre et al (2019a) compare, for 12 European countries, standard old-age poverty rates with the hypothetical old-age poverty rates that would have prevailed if (i) all individuals, whatever their income, had enjoyed the same survival conditions, and if (ii) all individuals within the same income class had been subject to the same income mobility process. They find that, even when taking income mobility into account, it is still the case that the missing poor bias introduces noise in international comparisons of poverty.

Whereas these studies are static, Lefebvre et al (2019b) reexamined the missing poor problem in a dynamic model. That paper compares actual steady-state headcount poverty rates with hypothetical steady-state headcount poverty rates that would have prevailed in the case where survival conditions would have been the same for all individuals, and equal to the ones of the top income class.⁶

Like these previous works, the present paper studies the missing poor bias, but with important differences. Unlike Kanbur and Mukherjee (2007) and Lefebvre et al (2013, 2018, 2019a)'s static models, we adopt here a dynamic perspective, and examine how the accumulation of selection biases across generations affects the measurement of poverty. Moreover, whereas Lefebvre et al (2019b) assumed equal fertility for all income classes, and, thus, focused only on the "survival advantage" of the non-poor, we consider here a more general evolutionary framework where selection takes place through two channels: income-differentiated mortality and income-differentiated fertility. Examining all sources of (repeated) income-based selection allows us to provide a more complete study of the missing poor phenomenon.

The Pre-Industrial period The present paper is also related to economic history literature revisiting the Pre-Industrial period in various directions.

A first important topic concerns the identification of factors at work during the Pre-Industrial period, and which favored the emergence of the Industrial

⁵For instance, while poverty above age 60 is higher in Portugal than in Estonia on the basis of standard headcount poverty rates, the ranking of poverty is reversed once the missing poor are taken into account.

⁶Given that this paper assumes that skills are redistributed randomly in each new cohort, it underestimates the persistence of cumulative selection effects, unlike the present paper.

Revolution in England. Allen (2009) argued that the Industrial Revolution took place in 18th century Britain as a response to the global economy of the 17th and 18th centuries, during which Britain benefited from high wages and cheap capital and energy in comparison to other countries in Europe and Asia. An alternative explanation was provided by Clark (2007), who argued that the Industrial Revolution first took place in Britain because the evolutionary advantage of the rich over the poor was larger in that country than in other countries, leading to a broader diffusion of the elite in all spheres of the economy. The existence of that evolutionary advantage - to which we referred above - was documented by Clark and Hamilton (2007) and Clark and Cummins (2015) for Pre-Industrial England.⁷ In a recent paper, Cummins (2020) showed that the evolutionary advantage of the rich during the Pre-Industrial era was less substantial in France than in England, thus confirming the potential role of evolutionary forces in the emergence of the British Industrial Revolution.

Another key issue concerns the measurement of standards of living. Angeles (2008) shows that, while the series of real GDP per capita and real wage rates grow closely during the first part of the 18th century in England, they diverge in the second part of the 18th century: while real GDP per capita kept on growing, real wages declined. Broadberry et al (2011) provide output-based estimates of GDP per capita for Britain over 1270-1870. In contrast to the long-run stagnation of daily real wage rates, their series of output-based GDP per capita exhibits modest but positive trend growth. Broadberry et al (2015) provide a new account of Britain's economic evolution over 1270-1870. They show how the transition to the modern growth regime built on the earlier foundations of a persistent upward trend in GDP per capita. These findings question the thesis - supported by daily real wage data - of a long stagnation during the Pre-Industrial era.⁸ One way of reconciling these pieces of evidence is through an increase in the number of hours worked, what De Vries (2008) called the "Industrious Revolution", i.e., a period over 1600-1800, during which productivity and consumer demand increased despite the absence of major technological innovation. Using data based on payments made to workers employed by the year rather than by the day, Humphries and Weisdorf (2019) find also that modern growth began in England two centuries earlier than commonly thought.⁹

The present paper contributes to the literature on the Pre-Industrial era by focusing on living standards at the bottom of the distribution. More precisely, we explore how the evolutionary advantage of the non-poor over the poor, documented by Clark and Hamilton (2007), Clark and Cummins (2015) and Cummins (2020), affects the measurement of poverty in Pre-Industrial times.¹⁰

⁷In a recent work, de la Croix et al (2019) show that the reproductive success of the middle class exceeded the one of the poor and the very rich classes in Pre-Industrial England.

⁸Campbell (2013) showed that per capita income increased in Italy during its 12th and 13th century commercial revolution, in Holland during its 15th and 16th century golden age, and in England during the 17th and 18th century runup to its industrial revolution. During these episodes, expanding trade sustained output and employment growth in the manufacturing and service sectors, developments that were not reflected in real wage data.

⁹The effects of that pattern on family welfare are studied by Horrell et al (2021).

¹⁰However, this paper does not address the issue of the transition to the Industrial period

Evolutionary growth theory Finally, this paper is also related to the increasingly large literature on evolutionary growth theory (Galor and Moav, 2002, 2005, Galor 2010, 2011). That literature highlights an alternative driving force for economic development over the long period: besides the standard driving factors (technological progress, physical capital accumulation, human capital accumulation and institutions), economic growth could also, over the long-run, be influenced by selection effects. For instance, individuals may differ in the weight they assign in their preferences to the human capital of their children. If the subpopulation assigning a higher weight to the human capital of the child has an evolutionary advantage, this can become relatively more and more numerous in the population, explaining, at some point in time, the economic take-off, and, hence, the transition from a stagnation regime to a growth regime. Thus evolutionary growth theory relates the dynamics of poverty and prosperity to the existence of an evolutionary advantage for some subpopulations.

Although the present paper focuses on a measurement issue, it brings nonetheless some contribution to evolutionary growth theory. Indeed, we analyze here the effect of the strength of an evolutionary advantage of the non-poor over the poor on the prevalence of poverty in the long-run. This allows us to identify the conditions under which a rise in the evolutionary advantage of the non-poor reduces the prevalence of long-run poverty. Our work thus also casts some light on the relation between evolutionary forces and long-run prosperity.

In sum, the present paper provides a threefold contribution to the literature. First, by quantifying the missing poor bias in Pre-Industrial societies, it contributes to the literature on the measurement of living standards before the Industrial Revolution. Secondly, at the methodological level, this paper complements the literature on poverty measurement in the context of unequal lifetime, by developing an indicator neutralizing the cumulative effects of repeated income-based selection across generations. Thirdly, by examining how income-based selection effects influence long-run poverty, this paper casts also some light on the role of evolutionary forces in long-run economic development.

3 The model

Let us consider a reduced-form economy whose adult population is partitioned into poor and non-poor individuals, with numbers given, respectively, by N_{pt} and N_{nt} . The adult population at period t can be represented by the vector:

$$\mathbf{N}_t = \begin{pmatrix} N_{pt} \\ N_{nt} \end{pmatrix}$$

The partition of the adult population into the poor and the non-poor subpopulations varies over time, depending on (i) fertility behaviors; (ii) survival

(Polak and Williamson 1991). Due to the demographic transition (Lee 2003), our model would have to be amended to account for variations in survival and fertility conditions.

conditions; (iii) social mobility.¹¹ Given that elements (i) to (iii) are likely to differ within the population depending on whether individuals are poor or not, some extra notations are needed. Let us denote by $f_p > 0$ (resp. $f_n > 0$) the average number of children born from a poor (resp. non-poor) adult. Let us denote by $s_p \in]0, 1[$ (resp. $s_n \in]0, 1[$) the probability of survival of a child born from a poor adult (resp. non-poor adult) to adulthood. Let us denote by $m_p \in]0, 1[$ the probability for a child born in a poor family and surviving to adulthood to escape from poverty at the adult age, and by \bar{m}_p the probability that a child born in a poor family and surviving to adulthood remains poor at the adult age.¹² Finally, let us denote by $m_n \in]0, 1[$ the probability for a child born in a non-poor family and surviving to adulthood to fall into poverty at the adult age, and by \bar{m}_n the probability that a child born in a non-poor family and surviving to adulthood remains non-poor once adult.¹³ The life cycle graph associated to this model is shown in Figure 2.

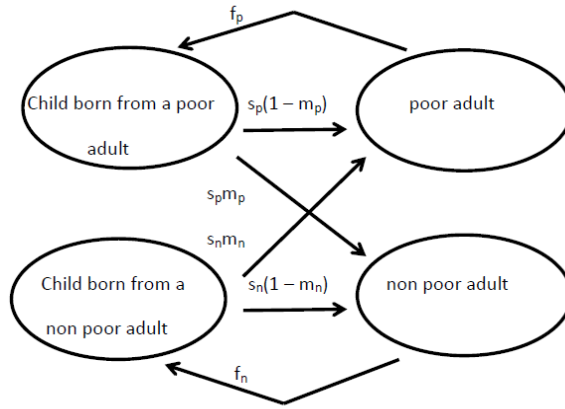


Figure 2: Life cycle graph of the model

Taken together, the parameters $\{f_p, f_n, s_p, s_n, m_p, m_n\}$ determine the dynamics of the structure of the population, in terms of its proportion living in poverty or not. To see this, let us define the (time-invariant) matrix \mathbf{M} as:

$$\mathbf{M} = \begin{pmatrix} f_p s_p \bar{m}_p & f_n s_n m_n \\ f_p s_p m_p & f_n s_n \bar{m}_n \end{pmatrix}$$

The matrix \mathbf{M} can be used to obtain the structure of the adult population at

¹¹We consider here a closed economy and we abstract from migrations. Abstracting from migrations is a simplification. However, this assumption makes sense for the Pre-Industrial epoch, during which most movements of populations took place *within* a country rather than *between* countries (see Wrigley 1967).

¹²We have $\bar{m}_p = 1 - m_p$.

¹³We have $\bar{m}_n = 1 - m_n$.

period $t + 1$ from the structure of the adult population at period t :

$$\mathbf{M}\mathbf{N}_t = \mathbf{N}_{t+1} \quad (1)$$

This expression can be rewritten in detailed form as:

$$\begin{pmatrix} f_p s_p \bar{m}_p & f_n s_n m_n \\ f_p s_p m_p & f_n s_n \bar{m}_n \end{pmatrix} \begin{pmatrix} N_{pt} \\ N_{nt} \end{pmatrix} = \begin{pmatrix} N_{pt+1} \\ N_{nt+1} \end{pmatrix} \quad (2)$$

Let us write the number of poor adults at time $t + 1$ as follows:

$$N_{pt+1} = N_{pt} f_p s_p \bar{m}_p + N_{nt} f_n s_n m_n \quad (3)$$

The above expression captures the fact that the size of the adult population that is poor at period $t + 1$ has two components, which are the two terms of the right-hand side (RHS) of that expression: on the one hand, the number of children who were born in poor families at t , survived to adulthood and remained poor once adult (first term of the RHS), and, on the other hand, the number of children who were born in non-poor families at t , survived to adulthood and fell into poverty once adult (second term of the RHS).

4 Long-run poverty

The model developed in Section 3 can be used to study the dynamics of the prevalence of poverty over time. For that purpose, let us assume that the poverty phenomenon is measured by the headcount ratio (HC_t), the ratio of the number of poor adults over the total adult population at period t , i.e.

$$HC_t = \frac{N_{pt}}{N_{pt} + N_{nt}} \quad (4)$$

The headcount poverty rate is a basic measure of poverty, which suffers from some shortcomings, such as its insensitivity to income transfers between individuals remaining either above or below the poverty threshold (Sen 1976) and its invariance to various dimensions of the income distribution (Foster et al 1984). However, the headcount poverty rate is most relevant for our framework, where the adult population is partitioned in only two groups - the poor and the non-poor -, without extra information on the income distribution. The poverty headcount ratio is also adequate for the study of poverty in the Pre-Industrial epoch, for which we have few data on the income distribution.

The level of measured poverty HC_t is likely to vary over time, depending on the prevalence of poverty at the previous period, and on the matrix \mathbf{M} and its components, which depend on parameters $\{f_p, f_n, s_p, s_n, m_p, m_n\}$. Studying the dynamics of poverty across long periods of time is not trivial, but some features of our model are worth noticing, since these will allow us to use fundamental theorems of population analysis (Caswell 2001).

In order to study the long-run prevalence of poverty as measured by the headcount ratio, let us first notice the following property of matrix \mathbf{M} .

Proposition 1 *The matrix \mathbf{M} is irreducible and primitive.*

Proof. See the Appendix. ■

Proposition 1 provides a simple, but important result, which allows us to use both the Perron-Frobenius Theorem and the Strong Ergodic Theorem (Caswell 2001) for the analysis of the long-run prevalence of poverty.

The Perron-Frobenius Theorem states that, under conditions of irreducibility and primitivity of a non-negative matrix, there exists in general one eigenvalue that is greater than or equal to any of the other eigenvalue of that matrix.¹⁴ This is called the "dominant eigenvalue". According to the Strong Ergodic Theorem, that dominant eigenvalue determines the ergodic properties of population growth.¹⁵ To be more accurate, the Strong Ergodic Theorem states that if the matrix is primitive, then, regardless of the initial population, the population will, in the long-run, grow at a rate given by the dominant eigenvalue, with a stable population structure proportional to the eigen vector associated to that eigenvalue (the influence of other eigenvalues being negligible).

Proposition 2 gives us the long-run partition of the population, as well as the associated headcount poverty rate.

Proposition 2 *The long-run population structure is defined, up to a constant $c > 0$, by:*

$$\begin{pmatrix} N_p \\ N_n \end{pmatrix} = \begin{pmatrix} c \frac{f_p s_p \bar{m}_p - f_n s_n \bar{m}_n + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}{2f_p s_p - (f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}} \\ c \frac{2f_p s_p (1 - \bar{m}_p)}{2f_p s_p - (f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}} \end{pmatrix}$$

while the associated long-run headcount poverty rate is

$$HC = \frac{f_p s_p \bar{m}_p - f_n s_n \bar{m}_n + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}{f_p s_p (2 - \bar{m}_p) - f_n s_n \bar{m}_n + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}$$

Proof. See the Appendix. ■

Proposition 2 provides a closed-form solution for the long-run prevalence of poverty, as measured by the headcount ratio HC . The long-run prevalence of poverty does not depend on the level of initial conditions. Whatever the economy considered involves initially a small or a large fraction of the population living in poverty, this has no effect on the long-run poverty rate. The long-run headcount ratio depends only on the parameters $\{f_p, f_n, s_p, s_n, m_p, m_n\}$ describing group-specific fertility, survival and social mobility.

The fact that the long-run level of the poverty headcount ratio does not depend on initial conditions may seem surprising. But this result follows from the application of the Strong Ergodic Theorem to the particular context under study. The Strong Ergodic Theorem provides conditions on a population

¹⁴See Caswell (2001), p. 83-84.

¹⁵See Caswell (2001), p. 84-85.

process under which the structure of the population stabilizes asymptotically independently of the initial structure of the population. That theorem was widely used, for instance, in theoretical demography, to study the stabilization of the age-structure of populations independently from the initial age-structure. Proposition 2 follows from the application of the Strong Ergodic Theorem to the study of the structure of the population in terms of prevalence of poverty.

5 Evolutionary forces and poverty measurement

Recent empirical evidence, such as Clark and Hamilton (2006), Clark and Cummins (2015), de la Croix et al (2019) and Cummins (2020), identified differentials in the number of surviving children per adult in Pre-Industrial societies, the poor having a lower number of surviving offspring in comparison to the non-poor. In terms of our model, this implies that the numbers of children surviving to adulthood satisfy:

$$f_n s_n = \mu f_p s_p \tag{5}$$

where $\mu > 1$. The parameter $\mu \equiv \frac{f_n s_n}{f_p s_p}$ is the ratio of the number of surviving children of the non-poor over the number of surviving children of the poor. As such, this parameter measures what we can call the "evolutionary advantage" of the non-poor, that is, the advantage that they have over the poor in terms of the number of offspring surviving to adulthood.¹⁶

The precise level of the parameter μ is an empirical issue: it has been shown to vary across societies and epochs, but there is nonetheless solid evidence that, during the Pre-Industrial period, μ is strictly larger than 1, that is, that non-poor individuals had a larger number of surviving children than poorer individuals.¹⁷ One can regard this as empirical evidence supporting the existence of an "evolutionary advantage" for the non-poor in Pre-Industrial societies.

In this paper, we will not examine the empirical issue of the determinants of the "evolutionary advantage" of the non-poor over the poor in Pre-Industrial societies, nor the effect of policies on this advantage. Using empirical evidence from Clark and Hamilton (2006), Clark and Cummins (2015), de la Croix et al (2019) and Cummins (2020), we will take the existence of an advantage of the non-poor in terms of the size of surviving offspring as given, and examine the consequences of that "evolutionary advantage" for the measurement of poverty across long periods of time. The question is: how does the existence of an evolutionary advantage for the non-poor over the poor affect the measurement of poverty in Pre-Industrial societies?

¹⁶Various mechanisms may explain why $\mu > 1$. On the mortality side, bad material living conditions can lead to higher mortality of the poor because of low quality/quantity of food or lack of health care. On the fertility side, the effects of poverty on fertility may take distinct forms. For instance, poverty may cause physiological stress that reduces fertility. However, it is also possible that humans, like other living beings, react to a poor environment by adopting a particular life history strategy involving a larger number of descendants (r-strategy), which may counterbalance the previous effect. We will not try here to quantify all these effects. Rather, we will take μ as given and consider its effect on the measurement of poverty.

¹⁷On the evolutionary advantage of the rich in Pre-Industrial times, see Clark (2007).

To answer that question, a first step consists of rewriting the level of the long-run headcount poverty rate of Proposition 2 while substituting for the parameter μ measuring the strength of the evolutionary advantage of the non-poor over the poor. We can then examine how variations in μ affect the level of the long-run poverty rate.¹⁸ Our results are summarized in Proposition 3.

Proposition 3 *The long-run headcount poverty rate is:*

$$HC = \frac{\bar{n}_p - \mu\bar{m}_n + \sqrt{(\bar{n}_p + \mu\bar{m}_n)^2 - 4\mu(1 - m_n - m_p)}}{2 - \bar{m}_p - \mu\bar{m}_n + \sqrt{(\bar{m}_p + \mu\bar{m}_n)^2 - 4\mu(1 - m_n - m_p)}}$$

where $\mu > 1$ measures the strength of the evolutionary advantage of the non-poor over the poor.

The derivative of the long-run headcount ratio with respect to the strength of the evolutionary advantage of the non-poor μ has an ambiguous sign, which depends on the following condition:

$$\frac{\partial HC}{\partial \mu} \geq 0 \iff 1 - HC \geq \bar{m}_n$$

Proof. See the Appendix. ■

Proposition 3 states that the long-run level of the headcount poverty rate is entirely determined by (i) the strength of the evolutionary advantage of the non-poor over the poor, i.e. the parameter μ ; (ii) the patterns of social mobility, i.e. parameters $\{m_p, m_n\}$. The long-run prevalence of poverty depends on survival conditions and fertility *only insofar* as they determine the strength of the evolutionary advantage of the non-poor, but not otherwise.¹⁹

The second part of Proposition 3 states a result that could not have been anticipated without a modeling of the measurement problem at stake: the long-run headcount poverty rate may be increasing or decreasing with respect to the strength of the evolutionary advantage of the non-poor over the poor μ . Two distinct cases can arise. On the one hand, if the downward mobility for the non-poor is low (i.e. \bar{m}_n is high) with respect to $1 - HC$, a rise of the strength of the evolutionary advantage of the non-poor over the poor contributes to decrease the long-run poverty headcount ratio. On the other hand, if the downward mobility for the non-poor is high (i.e. \bar{m}_n is low) with respect to $1 - HC$, an increase in the strength of the evolutionary advantage of the non-poor contributes to increase the long-run poverty headcount ratio.

¹⁸The headcount poverty rate HC is fully determined by parameters $\{m_n, m_p, \mu\}$. Hence the derivative $\frac{\partial HC}{\partial \mu}$ measures the full effect of a marginal variation of the strength of the evolutionary advantage of the non-poor on the prevalence of poverty in the long-run.

¹⁹The intuition behind that result goes as follows. If we were considering the *size* of the population in the long-run, the absolute levels of group-specific fertility rates and survival rates would matter. However, from the perspective of the *structure* of the population in the long-run, only the *relative* levels of fertility and survival rates across groups matter. The long-run structure of the population depends on whether one group has a larger number of surviving offspring than another.

The condition $1 - HC \geq \bar{m}_n$ can be interpreted while using marginalist reasoning. A marginal rise in μ can be interpreted as an infinitely small increase in the number of individuals born from non-poor adults and surviving to adulthood. $1 - HC$ is the long-run proportion of non-poor individuals in the total (adult) population, while \bar{m}_n is the probability, for a child born in a non-poor family who survived to adulthood, to avoid poverty at adulthood. The effect of a marginal rise of μ on HC depends on whether the prevalence of poverty in the *added* (adult) population is larger, equal or smaller than the prevalence of poverty in the *total* (adult) population. If we had $1 - HC = \bar{m}_n$, the proportion of the added population who escapes from poverty at adulthood would be equal to the proportion of the non-poor in the total (adult) population, so that the addition of that extra population would not affect the proportion of the non-poor in the total (adult) population. If we have $1 - HC < \bar{m}_n$ (resp. $1 - HC > \bar{m}_n$), the proportion of the added population who escapes from poverty at adulthood would be superior (resp. inferior) to the proportion of the non-poor in the total (adult) population, so that the addition of those persons increases (resp. decreases) the proportion of the non-poor in the total (adult) population, implying a fall (resp. rise) of the long-run headcount poverty rate.

6 Counting the missing poor

When considering static economies, a simple way to quantify the extent of income-based selection effects consists of (i) computing a hypothetical poverty measure that would have prevailed provided all income classes considered had faced exactly the same survival conditions, and (ii) comparing that hypothetical measure of poverty with the standard one (see Kanbur and Mukherjee 2007, Lefebvre et al 2013). The difference between the two measures quantifies the extent to which standard poverty measures are subject to a selection bias. The stage (i) amounts to count the "missing poor"- i.e. the persons who would have been counted as poor provided these did not die earlier than the non-poor - and to add these missing poor persons to the population under study.

In an intergenerational context, counting the missing poor raises extra difficulties. The reason is twofold. First, in an intergenerational perspective, a selection bias arises not only because of income-differentiated mortality, but, also, because of *income-differentiated fertility*. Harsh living conditions may prevent the poor from having children. Those "missing" children are also likely to become "missing poor" adults in the future. Income-differentiated fertility thus leads to some form of income-based selection effects. Second, the additional complexity lies also in the fact that repeated selection biases may lead to *cumulative* measurement errors that add up over time, generations after generations. When computing hypothetical poverty measures corrected for the selection bias, one needs, in an intergenerational perspective, to add not only the poor persons themselves, but, also, all their descendants, for all successive generations. Counting the missing poor becomes then more complex.

To quantify the missing poor bias in an intergenerational context, one method

consists of comparing the *actual* long-run poverty rate with the *hypothetical* long-run poverty rate that would have prevailed provided the evolutionary advantage of the non-poor over the poor were hypothetically set to zero. This amounts to compute the long-run poverty rate when μ is fixed to unity. That hypothetical long-run poverty rate is written as:

$$HC^H \equiv \frac{\bar{m}_p - \bar{m}_n + \sqrt[2]{(\bar{m}_p + \bar{m}_n)^2 - 4(1 - m_n - m_p)}}{2 - \bar{m}_p - \bar{m}_n + \sqrt[2]{(\bar{m}_p + \bar{m}_n)^2 - 4(1 - m_n - m_p)}} \quad (6)$$

The hypothetical poverty headcount ratio HC^H measures the poverty that would have prevailed provided no income class did benefit from any evolutionary advantage. Hence HC^H measures the poverty that would have prevailed if all missing poor individuals (and their descendants) had been added to the population and had been properly counted as poor. Evolutionary forces being neutralized, HC^H does not depend on differences of survival conditions or fertility across income groups. HC^H depends only on the degree of upward income mobility and downward income mobility within the society under study.

The size of the selection bias in poverty measurement can be quantified by comparing the standard poverty measure HC with the hypothetical poverty measure HC^H . Three cases can arise:

- If $HC^H > HC$, adding the missing poor and their descendants contributes to increase the measured poverty. In that case, the evolutionary advantage of the non-poor over the poor has pushed poverty rates down, and the selection bias exhibits a positive sign;
- If $HC^H = HC$, adding the missing poor and their descendants does not affect the measured poverty. In that case, the evolutionary advantage of the non-poor over the poor did not affect poverty measurement, because selection effects had benign effects;
- If $HC^H < HC$, adding the missing poor and their descendants contributes to lower the measured poverty. In that case, the evolutionary advantage of the non-poor over the poor has pushed poverty rates up, and the selection bias exhibits a negative sign.

The sign and the extent of the gap between the actual headcount poverty rate and the hypothetical headcount poverty rate are likely to vary with the structural parameters of the economy, i.e., demographic conditions $\{f_n, f_p, s_n, s_p\}$ as well as social mobility parameters $\{m_n, m_p\}$. The next section uses data on the Pre-Industrial period in order to examine which case prevails.

7 The missing poor in Pre-Industrial England

To what extent did the evolutionary advantage of the non-poor over the poor affect the prevalence of poverty during the Pre-Industrial era ?

In order to answer that question, this section proceeds as follows. We first use empirical evidence on poverty in Pre-Industrial England to calibrate the components of the long-run headcount poverty ratio $\{m_p, m_n, \mu\}$. Then, we use these calibrated parameters to compute the hypothetical long-run poverty measure HC^H in Pre-Industrial England, and we compare this hypothetical measure with the standard one.

Time scale In order to quantify the impact of the evolutionary advantage of the non-poor over the poor on the measurement of poverty, this section focuses on measures of poverty in the late 17th century, as obtained from King's social tables (1688). Our analysis takes the level of poverty prevalence in the late 17th century as a proxy indicator of the long-run poverty during the Pre-Industrial period. However, we do not examine the dynamics of poverty on a yearly basis, and we ignore also the transition towards the Industrial Revolution. These choices relative to the time horizon of our analysis require some explanations.

The reason why we do not study the yearly dynamics of poverty prevalence lies in the fact that our theoretical framework provides a basis for the comparison of *long-run* measures of poverty, i.e., measures of the prevalence of poverty once all adjustments related to the demographic dynamics and social mobility have taken place. Focusing on long-run measures of poverty allows us to avoid complex reconstructions of hypothetical populations on a year-by-year basis. Moreover, ignoring transitional issues allows us to focus on the long-run accumulation of selection biases across generations, the topic of the present study.

Regarding the transition from the Pre-Industrial period to the Industrial Revolution, the reason why we restrict our horizon to the Pre-Industrial period lies in the fact that our long-run measures of poverty (actual and hypothetical) are conditional on the existence of a mobility matrix \mathbf{M} whose elements $\{f_n, f_p, s_n, s_p, m_n, m_p\}$ are *constant* over time (at least in trend). This constancy condition prevents us from using our framework to study the Industrial Revolution, which is temporally close to the Demographic Transition (associated with large changes in mortality and fertility patterns).

But even under those restrictions, our analysis faces difficult challenges. As far as the long-run prevalence of poverty in Pre-Industrial England is concerned, it is difficult to come with a single number. Our calculations rely on the figures in King's social tables (1688) amended by the corrections of Lindert and Williamson (1982, 1983). The first figure for the prevalence of poverty in Pre-Industrial England comes from Lindert and Williamson (1983), who estimate that poverty in England (1688) was about 24.2 percents. That number corresponds to the ratio of two numbers: at the numerator, the number of "paupers" (i.e. recipients of the Poor Laws benefits) and the (corrected) number of "vagrants" in 1688: 336,672 (see Table 1); at the denominator, the total number of "able-bodied" income recipients plus "paupers": 1,390,586. Note that this poverty rate of 24.2 % should be regarded as a lower bound for the prevalence of poverty. The reason is that the paupers were only one segment of the population living in poverty. Table 1 shows that if we add to the paupers all workers

earning less than 50 % of the average income, we obtain a poverty rate that lies between 48.2 % (corrected figures) and 59.6 % (uncorrected figures). We will thus take 48.2 % as an upper bound.²⁰ We will also consider an intermediate value for the headcount poverty rate, equal to 36.2 %.

Calibration For a given value of the long-run headcount poverty rate, it is possible, by using the formula of Proposition 3, to calibrate jointly the structural parameters $\{\mu, m_n, m_p\}$.

Regarding the parameter μ , which captures the strength of the evolutionary advantage of the non-poor over the poor, we rely on measures of net fertility by asset income terciles in England (1500-1779) provided in Clark and Cummins (2015). Defining the first tercile as the poor and the second and third terciles as the non-poor, we obtain that $\mu = 1.286$.

Concerning the mobility parameters m_p, m_n , it is difficult to have precise estimates. The measurement of social mobility has been widely debated in the recent years, following Clark's (2014) study. Using original data on dynasties based on family names, Clark (2014) argued that standard measures of social mobility (based on pairs "parent-children") tend to *overestimate* social mobility, and to underestimate social inertia. The reason is that standard estimates are sensitive to all shocks that take place over time and weaken the strength of the link between the social position of the parent and the one of his children, unlike estimates that cover the life of a dynasty (the longer time horizon allowing the cancellation of random terms). Concerning Pre-Industrial England, Clark and Hamilton (2006, p. 26) argue, on the basis of their data, that "Nearly half of the sons of higher class testators would end up in a lower asset class at death." But one cannot take this as evidence that $m_n = 0.500$. The reason is that suffering from downward intergenerational income mobility does not imply falling in poverty. Thus $m_n = 0.500$ is too large in magnitude, and can only be taken as an upper bound. Our computations will rely on three values for m_n : a lower bound equal to $m_n = 0.100$, an intermediate value equal to $m_n = 0.300$ and an upper bound equal to $m_n = 0.500$.

Having values for HC , μ and m_n , it is possible, using the formula for HC , to calibrate the parameter m_p in a way consistent with other calibrations. Table 2 shows the calibrated values for m_p under each value of HC and m_n .

	HC		
m_n	0.242	0.362	0.482
0.100	$m_p = 0.186$	$m_p = 0.044$	$m_p = 0.001$
0.300	$m_p = 0.992$	$m_p = 0.497$	$m_p = 0.266$
0.500	$\nexists m_p \in]0, 1[$	$m_p = 0.951$	$m_p = 0.543$

Table 2: Calibration of structural parameters $\{m_n, m_p\}$ under $\mu = 1.286$.

²⁰It should be stressed that the latter number is higher, but still of a close magnitude to estimates for poverty in the UK around 1820 in Bourguignon and Morrisson (2002) and Ravallion (2006), which lie around 40-45 %.

These combinations of structural parameters consist of various ways to replicate the measured long-run prevalence of poverty during the Pre-Industrial period. Several remarks are in order. First, given that the non-poor has the evolutionary advantage over the poor, it is no surprise that, in order to replicate a higher headcount poverty rate, one needs to assume, for a given degree of downward mobility (i.e. a given m_n) a lower degree of upward income mobility for the poor (i.e. a lower m_p).²¹ Second, for a given HC , assuming a higher downward mobility for the non-poor (i.e. shifting from $m_n = 0.100$ to $m_n = 0.300$) must be followed by a rise in the postulated upward mobility for the poor, so that HC remains unchanged. Third, it should be stressed that, under a high downward mobility ($m_n = 0.500$), there exists no value of m_p in the unit interval that is compatible with $HC = 24.2\%$. This shows that our model imposes some restrictions on the set of values for parameters $\{\mu, m_n, m_p\}$ compatible with plausible values of the poverty rate HC .

Results Let us compare the standard headcount poverty rates with the hypothetical ones obtained under the postulate of no evolutionary advantage for the non-poor ($\mu = 1$). From our theoretical findings, we can anticipate some qualitative results. The condition of Proposition 3 tells us that reducing μ reduces measured poverty HC when $1 - HC > \bar{m}_n$, and that reducing μ raises HC when $1 - HC < \bar{m}_n$. Our calculations involve eight distinct calibrations of $\{1 - HC, \bar{m}_n\}$ (see Table 2). From Proposition 3, we can deduce that the condition $1 - HC > \bar{m}_n$ holds in cases (0.758, 0.700), (0.638, 0.500) and (0.518, 0.500) and that in the other five cases the condition $1 - HC < \bar{m}_n$ holds. In the first three cases, the hypothetical headcount ratio HC^H is lower than the standard headcount ratio HC , whereas the opposite holds in the five other cases.

Let us now examine to what extent the hypothetical headcount ratio differs from the standard one, that is, to what extent the evolutionary advantage of the non-poor affects poverty measurement. Figures 3 to 5 summarize our results for the case where HC takes, respectively, its lower bound value (24.2 %), its intermediate value (36.2 %) and its upper bound value (48.2 %).

In Figure 3, we can see that whether the hypothetical headcount poverty rate is superior or inferior to the standard headcount poverty rate depends on the postulated downward social mobility m_n . When downward mobility is low, HC^H exceeds HC by about 10 percentage points. In that case, we can say that provided the non-poor had no evolutionary advantage over the poor, the measured poverty in Pre-Industrial England would have been higher by 10 percentage points. The evolutionary advantage of the non-poor over the poor has thus contributed to push measured poverty down. However, if one assumes a high downward mobility, HC^H is slightly *below* HC . In that case, counting the missing poor does not raise the measured poverty, but tends to reduce it. The intuition behind that result goes as follows. Remember that, in order to replicate

²¹For instance, under a low downward mobility ($m_n = 0.100$), a probability of leaving poverty of 0.186 can replicate a headcount poverty rate of 24.2 %, but to reproduce $HC = 48.2\%$, one needs, under $m_n = 0.100$, a lower upward mobility for the poor, m_p being extremely low (0.001).

a low prevalence of poverty when there is a high probability for the children of the non-poor (who have the evolutionary advantage) to fall into poverty, one needs a high upward mobility for the poor. But once we (hypothetically) cancel the evolutionary advantage of the non-poor, we remain with this high upward mobility for the poor, which pushes measured poverty down.

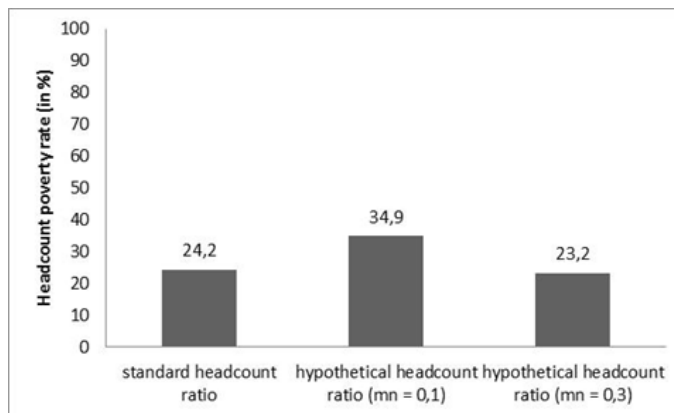


Figure 3: Standard headcount poverty rate and hypothetical headcount poverty rates under the low benchmark for poverty in Pre-Industrial England.

Let us now compare those results with the ones obtained under higher values for the standard headcount poverty rate (Figures 4 and 5). In that case, the hypothetical headcount ratio is higher than the standard one, except when one assumes a high downward social mobility (i.e., $m_n = 0.500$). The extent to which standard poverty measures are biased downwards depends on the degree of downward mobility for the non-poor. In Figure 5, if the postulated downward mobility for the non-poor is low ($m_n = 0.100$), HC^H is about 50 percentage points higher than HC , whereas if the postulated downward mobility for the non-poor is intermediate ($m_n = 0.300$), the hypothetical poverty rate HC^H is about 5 percentage point higher than HC . However, under a high downward social mobility ($m_n = 0.500$), the hypothetical headcount ratio becomes very close to the standard headcount ratio.

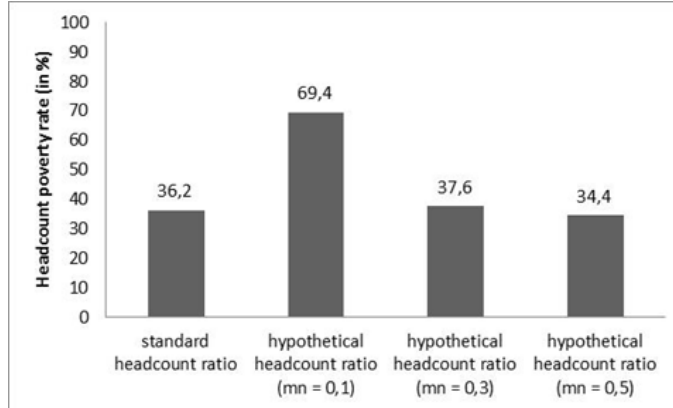


Figure 4: Standard headcount poverty rate and hypothetical headcount poverty rates under the intermediate benchmark for poverty in Pre-Industrial England.

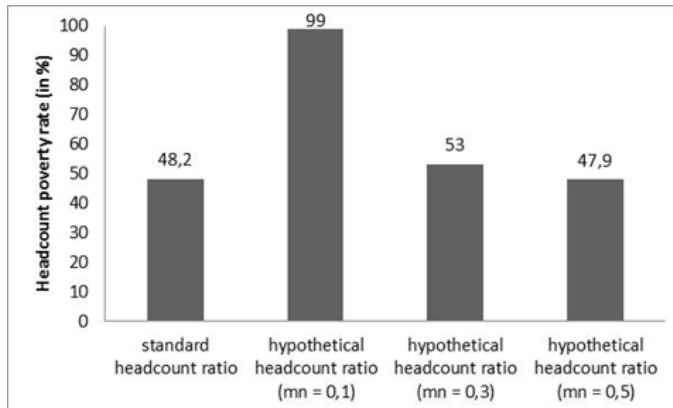


Figure 5: Standard headcount poverty rate and hypothetical headcount poverty rates under the high benchmark for poverty in Pre-Industrial England.

How can one interpret this lack of robustness? The intuition behind those results is that, in order to replicate a high poverty rate when the non-poor (who have the evolutionary advantage) have little downward mobility, the upward mobility for the poor must be extremely low, poverty being a kind of *absorbing state* across generations. But when one hypothetically cancels out the evolutionary advantage of the non-poor, we remain with poverty being an absorbing state, which explains the extremely high level of HC^H . Thus, if the observed poverty was high despite the low downward social mobility, the selection bias

induced by the evolutionary advantage of the non-poor had a high effect on the measurement of poverty. However, if, on the contrary, the observed poverty was associated to a high downward mobility of the non-poor, the upward mobility of the poor could be higher, and once the evolutionary advantage of the former is cancelled, HC^H is only slightly higher than the standard HC .

8 The missing poor in Pre-Industrial France

To check the robustness of our results, this section reexamines the measurement of poverty during the Pre-Industrial period in France, by applying the method developed in Section 7.

Calibration In order to calibrate the long-run headcount poverty rate in 18th century France, we use the social tables of Isnard (1781) and their corrections by Morrisson and Snyder (2000). Table 3 summarizes the main statistics used for the calibration of HC . Given that the total income equals 4,170 millions of livres distributed on 6035 households, the average income per household equals 690 livres. Hence, if one fixes the poverty threshold at 50 % of the average household income, one obtains that 43.1 % of households live in poverty. One can take that figure as a proxy for the headcount poverty rate prevailing in 18th century France.

Morrisson and Snyder (2000) underline that, even though Isnard's social table is quite close to the best estimates of the income distribution prevailing in France in the late 18th century, there are nonetheless some corrections to be brought to Isnard's figures. Morrisson and Snyder (2000) notice that Isnard's income distribution implies a total population of about 24 millions inhabitants, which is about 4 millions less than it is generally accepted. According to Morrisson and Snyder (2000), Isnard may have underestimated the number of persons in the low income categories, as well as the existence of a large number of poor individuals who have no income at all. Within Isnard's framework, the addition of 4 million people at the bottom of the income distribution is equivalent to adding about 1000 households in category (8). As a consequence, the proportion of households with less than 50 % of the average household income becomes now 51,4 %. This figure will be taken as the upper bound estimate for the prevalence of poverty in 18th century France.

categories of households	Isnard's income distribution (1781)		
	Heads of households (thousands)	Income per household (per cent)	Income per household (livres)
(1)	335	5.5	3,670
(2)	300	5.0	1,800
(3)	400	6.6	1,200
(4)	600	10.0	800
(5)	800	13.2	500
(6)	1,000	16.6	400
(7)	1,200	19.9	300
(8)	1,400	23.2	200
Total	6,035	100.0	
% of households with income < 50 % of average income	43,1 %		

Table 3. Isnard's income distribution for France (1781). Source: Morrisson and Snyder (2000).

Regarding the calibration of the parameter μ , we rely on the recent study of Cummins (2020) on the prevalence of Malthusian preventive and positive checks in late 18th century France. Using husband's occupation from the parish records of 41 French rural villages, Cummins (2020) finds that the number of surviving children per marriage equals 3.1 in the bottom wealth tercile, 3.4 children in the middle wealth tercile, and 3.1 children in the top wealth tercile. Defining the first tercile as the poor and the second and third terciles as the non-poor, we obtain that $\mu = 1.048$. Note that this value is much lower than the one obtained for England ($\mu = 1.286$), suggesting that the evolutionary advantage of the non-poor in France was lower than in England.

Regarding the calibration of parameters m_n and m_p , we proceed in the same way as in the previous section. Table 4 summarizes the calibrations of those parameters replicating the lower bound and the upper bound value for the headcount poverty rate HC .

m_n	HC	
	0.431	0.514
0.100	$m_p = 0.111$	$m_p = 0.076$
0.300	$m_p = 0.388$	$m_p = 0.274$
0.500	$m_p = 0.664$	$m_p = 0.472$

Table 4: Calibration of structural parameters $\{m_n, m_p\}$ under $\mu = 1.048$.

Results Let us now compare the actual headcount poverty rate in France with the hypothetical headcount poverty rate that would have prevailed in the absence of evolutionary advantage of the non-poor over the poor. As in Section 7, the hypothetical headcount poverty rate is obtained by calculating the level of HC^H , which assumes that the parameter μ is set to unity. Our results are summarized in Figures 6 and 7.

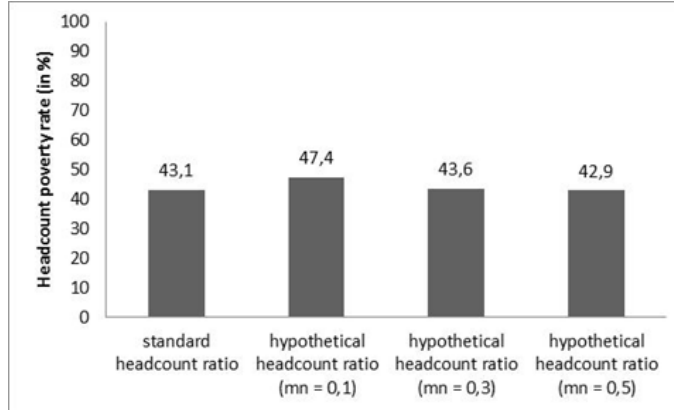


Figure 6: Standard headcount poverty rate and hypothetical headcount poverty rates under the low benchmark for poverty in Pre-Industrial France.

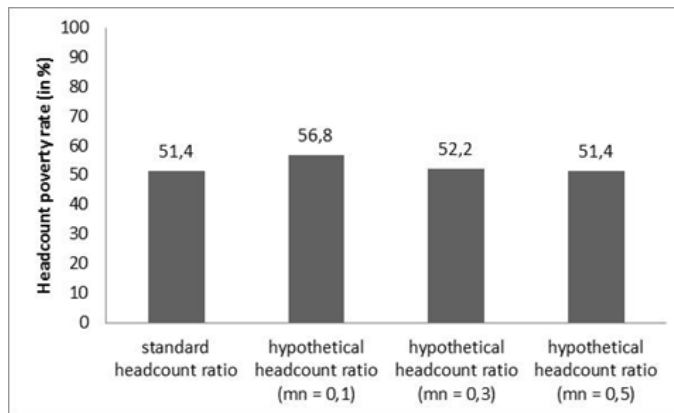


Figure 7: Standard headcount poverty rate and hypothetical headcount poverty rates under the high benchmark for poverty in Pre-Industrial France.

In comparison to our calculations concerning Pre-Industrial England, an important difference is that the gap between the standard headcount ratio and the hypothetical headcount ratios is, under all calibrations, smaller for the case of Pre-Industrial France. The intuition behind that result is that the evolutionary advantage of the non-poor over the poor has been shown to be smaller in France than in England (see Cummins, 2020), leading to a lower μ in Pre-Industrial France. As a consequence, the neutralization of income-based selection effects leads to a larger correction in the case of England than in the case of France.

Having stressed this, Figures 6 and 7 tend to confirm previous results: the size of the missing poor bias is sensitive to the postulates on downward social mobility for the non-poor. For instance, in Figure 7, the gap between the standard headcount ratio and the hypothetical one is as high as 5.4 percentage points when $m_n = 0.100$ (low downward mobility), but vanishes almost entirely when $m_n = 0.500$ (high downward mobility). That result confirms the previous findings for pre-Industrial England.

Finally, it should be stressed that the existence of income-based selection effects of unequal strength across countries can affect the international comparisons of poverty in the Pre-Industrial era. Given that our measures of poverty in England concern the late 17th century, while the ones in France concern the late 18th century, we cannot compare these numbers directly. However, since the evolutionary advantage of the non-poor over the poor is stronger in England than in France, it is possible, when considering comparable measures of poverty during the Pre-Industrial period, that the correction for the missing poor bias modifies the ranking of these countries in terms of poverty. This is likely to be the case when a low downward social mobility is assumed, since in that case the missing poor bias takes its highest level in each country.²²

9 Conclusions

Pre-Industrial societies being characterized by a large prevalence of poverty and by an evolutionary advantage of the non-poor over the poor, one may expect that measures of poverty during the Pre-Industrial period suffer from the missing poor bias, in the sense that the poor are under-represented, and, hence, not properly counted. Given the repetition of selection biases across generations, one may also expect that the impact of the evolutionary advantage of the non-poor on poverty measures turns out to be substantial in the Pre-Industrial era.

To quantify the size of the missing poor bias in Pre-Industrial societies, this paper developed a simple matrix population model, where the population is partitioned into poor and non-poor subpopulations, each subpopulation being characterized by specific mortality, fertility and social mobility. That setting allowed us to characterize the long-run partition of the population into poor and non-poor as the eigen vector associated to the dominant eigenvalue of the population matrix.²³

²²For the sake of making that point, let us assume that the measures of poverty in Sections 7 and 8 are comparable. If, for instance, one adopts the intermediate benchmark for poverty in Pre-Industrial England, and compare it with the measure of poverty in Pre-Industrial France (lower bound), one can see that HC is *higher* in France than in England. However, this may not be the case when considering hypothetical poverty rates HC^H where all selection biases are neutralized. Under a low downward social mobility, HC^H is *lower* in France than in England. Note that such an inversion of the ranking does not occur when a high downward social mobility is assumed (since in that case HC^H is close to HC in each country).

²³One limitation of this matrix population model is that it cannot account for potential second-order effects due to interactions between fertility, mortality and intergenerational mobility, because these are modelled as parameters, and not as variables. We leave the exploration of more general models for future research.

Our main finding is that the sign of the effect of a stronger evolutionary advantage of the non-poor over the poor on long-run poverty measures depends on the degree of downward social mobility. A stronger evolutionary advantage for the non-poor does not necessarily bias poverty measures downwards and may, under some conditions, lead to measures of poverty that are biased upwards. The latter case is especially likely when there is a high downward social mobility.

The comparison of the standard headcount poverty rates with the hypothetical measures of poverty that would have prevailed provided there had been no evolutionary advantage for the non-poor over the poor confirms these results. Under a low downward mobility, the missing poor bias lies, in the case of Pre-Industrial England, between 10 and 50 percentage points, whereas under a high downward mobility, the missing poor bias may be either slightly negative, or of about 5 percentage points. Concerning Pre-Industrial France, hypothetical poverty rates are closer to the standard one, because the evolutionary advantage of the non-poor is lower in France than in England. The varying strength of income-based selection mechanisms across countries can affect international comparisons in terms of poverty, to an extent that varies with the postulated degree of downward social mobility.

These results do not only highlight difficulties to measure and compare poverty in Pre-Industrial societies, but, also, cast some light on evolutionary growth theory (Galor and Moav 2002, 2005, Galor 2010, 2011). Evolutionary growth models emphasized that an evolutionary advantage of the most skilled individuals could favor the take-off of the economy, leading to sustained economic growth. Our findings suggest that evolutionary forces do not necessarily suffice to generate prosperity. If downward social mobility is high, a stronger evolutionary advantage for the non-poor does not lead to a lower long-run poverty rate, but to a higher long-run poverty rate. In Pre-Industrial societies, downward social mobility was substantial (Clark and Hamilton 2006, Clark 2007), so that one cannot exclude *a priori* that the evolutionary advantage of the non-poor over the poor had the effect to increase rather than to decrease the prevalence of poverty, against the claim that evolutionary forces may have driven the early economic take-off in England. Undoubtedly, further explorations are needed in order to have a more precise quantification of the complex role played by evolutionary forces in the process of long-run economic development.²⁴

10 Bibliography

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

Angeles L. (2008). GDP per capita or real wages? Making sense on conflicting views on pre-industrial Europe. *Explorations in Economic History*, 45: 147–63.

²⁴For that purpose, it would be useful to study the various mechanisms at work behind the evolutionary advantage of the non-poor over the poor in Pre-Industrial societies, by using the literature in evolutionary anthropology (see Hill and Kaplan 1999, Kaplan et al 2000).

- Barnett, G. (1936), *Two Tracts by Gregory King*. Baltimore: Johns Hopkins University Press.
- Broadberry S.N., Campbell B.M.S., Klein A., Overton M., van Leeuwen B. (2015), *British Economic Growth 1270–1870*. Cambridge University Press, Cambridge.
- Bourguignon, F., Morrisson, C. (2002). World distribution among world citizens: 1820-1992. *American Economic Review*, 92: 727-744.
- Campbell B.M.S. (2013). National incomes and economic growth in pre-industrial Europe: insights from recent research. *Quaestiones Medii Aevi Novae*, 18: 167–96.
- Caswell, H. (2001). *Matrix Population Models. Construction, Analysis, and Interpretation*. Sinauer, Massachusetts.
- Clark, G. (2007). *A Farewell to Alms. A Brief Economic History of the World*. Princeton University Press.
- Clark, G. (2014). *The Son Also Rises. Surnames and the History of Social Mobility*. Princeton University Press.
- Clark, G., Cummins, N. (2015). Malthus to modernity: wealth, status and fertility in England, 1500-1879. *Journal of Population Economics*, 28: 3-29.
- Clark, G., Hamilton, G. (2006). Survival of the richest. The Malthusian mechanism in pre-industrial England. *Journal of Economic History*, 66: 707-736.
- Crafts, N., Mills, T. (2009). From Malthus to Solow: How did the Malthusian economy really evolve? *Journal of Macroeconomics*, 31: 68-93.
- Cummins, N. (2020). The micro-evidence for the Malthusian system. France, 1670-1840. *European Economic Review*, 129 (C).
- de la Croix, D., Schneider, E, Weisdorf, J. (2019). Childlessness, celibacy and net fertility in Pre-Industrial England: the middle-class evolutionary advantage. *Journal of Economic Growth*, 24: 223-256.
- De Vries, J. (2008), *The Industrious Revolution: Consumer Behavior and the Household Economy*, Cambridge University Press, New York.
- Foster, J., Greer, J., Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 52: 761-766.
- Galor, O.(2010). The 2008 Laurence R. Klein Lecture - Comparative economic development: insights from Unified Growth Theory, *International Economic Review*, 51: 1-44.
- Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.
- Galor, O., Moav, O. (2002). Natural selection and the origin of economic growth. *Quarterly Journal of Economics*, 117: 1133-1191.
- Galor, O., Moav, O (2005). Natural selection and the evolution of life expectancy. Working Paper, Department of Economics, Brown University, Providence.
- Hill, K., Kaplan, H. (1999). Life history traits in humans: Theory and empirical studies. *Annual Review of Anthropology*, 28: 397-430.
- Horrell, S., Humphries, J., Weisdorf, J. (2021) Family Standards of Living Over the Long Run, England 1280–1850. *Past & Present*, 250: 87–134.
- Humphries, J., Weisdorf, J. (2019) Unreal Wages? Real Income and Economic Growth in England, 1260–1850. *Economic Journal*, 129: 2867–2887
- Isnard, A. (1781). *Traité des richesses*. Londres et Lausanne.
- Kanbur, R., Mukherjee, D. (2007). Premature mortality and poverty measurement. *Bulletin of Economic Research*, 59: 339-359.

- Kaplan, H., Hill, K., Lancaster, J., Hurtado, M. (2000). A theory of human life history evolution: diet, intelligence and longevity. *Evolutionary Anthropology*, 9: 156-185.
- Lee, R. (2003). The demographic transition. Three centuries of fundamental change. *Journal of Economic Perspectives*, 17: 167-190.
- Lefebvre, M., Pestieau, P., Ponthiere, G. (2013). Measuring Poverty Without the Mortality Paradox. *Social Choice and Welfare*, 40: 285-316.
- Lefebvre, M., Pestieau, P., Ponthiere, G. (2018). FGT Old-Age Poverty Measures and the Mortality Paradox: Theory and Evidence. *Review of Income and Wealth*, 64: 428-458.
- Lefebvre, M., Pestieau, P., Ponthiere, G. (2019a). Missing poor and income mobility. *Journal of Comparative Economics*, 47: 330-366.
- Lefebvre, M., Pestieau, P., Ponthiere, G. (2019b). Missing poor and income mobility. *Journal of Population Economics*, 32: 621-664.
- Lindert, P., Williamson, J. (1982). Revisiting England's Social Tables 1688-1812. *Explorations in Economic History*, 19: 385-408.
- Lindert, P., Williamson, J. (1983). Reinterpreting England's Social Tables, 1688-1913. *Explorations in Economic History*, 20: 94-109.
- Malthus, T. (1798). *An Essay on the Principle of Population*, London.
- Mokyr, J. (1980). Malthusian models and Irish history. *Journal of Economic History*, 40: 159-166.
- Morrisson, C., Snyder, W. (2000). The income inequality in France in historical perspective. *European Review of Economic History*, 4: 59-83.
- Polak, B., Williamson, J. (1991). Poverty, Policy and Industrialization. World Bank, Working Paper.
- Ravallion, M. (2016). *The Economics of Poverty: History, Measurement and Policy*. Oxford University Press.
- Sen, A.K. (1976). Poverty: an ordinal approach to measurement. *Econometrica*, 44: 219-231.
- Sen, A.K. (1998). Mortality as an indicator of economic success and failure. *Economic Journal*, 108: 1-25.
- Wrigley, E.A. (1967). *History and Population*. Weidenfeld and Nicolson, London.

11 Appendix

11.1 Proof of Proposition 1

Irreducibility prevails when the life cycle graph associated to the matrix admits at least one path from each node and towards each node. This is the case for matrix \mathbf{M} . As shown in Figure 2, the life cycle graph associated to our model includes two distinct populations, i.e., poor and non-poor, which contribute to each other through the social mobility process, which can go upwards or not for the poor, and downwards or not for the non-poor. Thus each population makes contributions to the other population, implying irreducibility of \mathbf{M} .

Primitivity arises when there exists a power k such that raising the matrix to that power makes it positive. This is clearly the case for matrix \mathbf{M} , which is a positive matrix.

11.2 Proof of Proposition 2

Let us characterize eigenvalues of the matrix \mathbf{M} . We look for solutions for the equation:

$$\mathbf{M}\mathbf{w} = \lambda\mathbf{w}$$

where λ is the eigenvalue (a scalar) while \mathbf{w} is the associated eigen vector, a vector that makes matrix multiplication and scalar multiplication equivalents. From the definition of the eigen vectors, it follows that:

$$\begin{aligned}\mathbf{M}\mathbf{w} - \lambda\mathbf{w} &= \mathbf{0} \\ (\mathbf{M} - \lambda\mathbf{I})\mathbf{w} &= \mathbf{0}\end{aligned}$$

Non-zero solutions require $(\mathbf{M} - \lambda\mathbf{I})$ to be a singular matrix, that is, that it has a zero determinant.

Hence eigenvalues are solutions to:

$$\det \begin{pmatrix} f_p s_p \bar{m}_p - \lambda & f_n s_n m_n \\ f_p s_p m_p & f_n s_n \bar{m}_n - \lambda \end{pmatrix} = 0$$

Therefore we have:

$$(f_p s_p \bar{m}_p - \lambda)(f_n s_n \bar{m}_n - \lambda) - f_n s_n m_n f_p s_p m_p = 0$$

Hence, after some simplifications:

$$\lambda^2 - \lambda(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) + f_p f_n s_n s_p (1 - m_n - m_p) = 0$$

Eigenvalues can be found as the roots of this polynomial. We have:

$$\Delta = (f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)$$

Note that Δ can be rewritten as:

$$\begin{aligned}\Delta &= (f_p s_p (1 - m_p))^2 + (f_n s_n (1 - m_n))^2 + 2f_p s_p (1 - m_p) f_n s_n (1 - m_n) \\ &\quad - 4f_p f_n s_n s_p (1 - m_n - m_p) \\ &= (f_p s_p (1 - m_p))^2 + (f_n s_n (1 - m_n))^2 - 2f_p f_n s_n s_p (1 - m_p - m_n) + 2f_p f_n s_n s_p m_p m_n \\ &= (f_p s_p (1 - m_p))^2 + (f_n s_n (1 - m_n))^2 - 2f_p f_n s_n s_p (1 - m_p - m_n - m_p m_n)\end{aligned}$$

Hence the two eigenvalues are:

$$\begin{aligned}\lambda_1 &= \frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}{2} \\ \lambda_2 &= \frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) - \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}{2}\end{aligned}$$

We have $\lambda_1 > \lambda_2$, so that the dominant eigenvalue is λ_1 .

We can then derive the long-run population structure by calculating the eigenvector \mathbf{w}_1 associated to the dominant eigenvalue λ_1 . The associated eigenvector is such that:

$$\begin{pmatrix} f_p s_p \bar{m}_p & f_n s_n \bar{m}_n \\ f_p s_p m_p & f_n s_n \bar{m}_n \end{pmatrix} \begin{pmatrix} N_p \\ N_n \end{pmatrix} = \begin{pmatrix} f_p s_p \bar{m}_p + f_n s_n \bar{m}_n + \sqrt{\frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2}{-4f_p f_n s_n s_p (1 - m_n - m_p)}} \\ \frac{2}{2} \end{pmatrix} \begin{pmatrix} N_p \\ N_n \end{pmatrix}$$

Hence we have

$$\begin{aligned} f_p s_p \bar{m}_p N_p + f_n s_n m_n N_n &= \frac{f_p s_p \bar{m}_p + f_n s_n \bar{m}_n + \sqrt{\frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2}{-4f_p f_n s_n s_p (1 - m_n - m_p)}}}{2} N_p \\ f_p s_p m_p N_p + f_n s_n \bar{m}_n N_n &= \frac{f_p s_p \bar{m}_p + f_n s_n \bar{m}_n + \sqrt{\frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2}{-4f_p f_n s_n s_p (1 - m_n - m_p)}}}{2} N_n \end{aligned}$$

Two equations and two unknowns. Normalizing to $N_p + N_n = 1$, the second equation can be rewritten as:

$$f_p s_p m_p N_p + f_n s_n \bar{m}_n (1 - N_p) = \frac{f_p s_p \bar{m}_p + f_n s_n \bar{m}_n + \sqrt{\frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2}{-4f_p f_n s_n s_p (1 - m_n - m_p)}}}{2} (1 - N_p)$$

From which it follows that

$$N_p = \frac{f_p s_p \bar{m}_p - f_n s_n \bar{m}_n + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}{2f_p s_p - (f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}$$

Hence the eigen vector associated to λ_1 is

$$\mathbf{w}_1 = \begin{pmatrix} N_p \\ N_n \end{pmatrix} = \begin{pmatrix} N_p \\ 1 - N_p \end{pmatrix} = \begin{pmatrix} \frac{f_p s_p \bar{m}_p - f_n s_n \bar{m}_n + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}}{2f_p s_p - (f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}} \\ \frac{2f_p s_p (1 - \bar{m}_p)}{2f_p s_p - (f_p s_p \bar{m}_p + f_n s_n \bar{m}_n) + \sqrt{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2 - 4f_p f_n s_n s_p (1 - m_n - m_p)}} \end{pmatrix}$$

From the Strong Ergodic Theorem, we have that

$$\lim_{t \rightarrow \infty} \frac{\mathbf{N}_t}{\lambda_1^t} = c_1 \mathbf{w}_1$$

that is, the asymptotic population structure is given by the eigen vector \mathbf{w}_1 , while the precise size of the different population groups can always be scaled as desired, since eigenvectors are always defined up to a multiplicative constant.

Hence the long-run headcount is given by:

$$HC = \frac{N_p}{N_p + N_n} = \frac{f_p s_p \bar{m}_p - f_n s_n \bar{m}_n + \sqrt[2]{\frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2}{-4f_p f_n s_n s_p (1 - m_n - m_p)}}}{2f_p s_p - f_p s_p \bar{m}_p - f_n s_n \bar{m}_n + \sqrt[2]{\frac{(f_p s_p \bar{m}_p + f_n s_n \bar{m}_n)^2}{-4f_p f_n s_n s_p (1 - m_n - m_p)}}}$$

11.3 Proof of Proposition 3

Let us substitute for $f_n s_n = \mu f_p s_p$ in the long-run poverty rate of Proposition 2. We obtain:

$$HC = \frac{\bar{m}_p - \mu \bar{m}_n + \sqrt[2]{(\bar{m}_p + \mu \bar{m}_n)^2 - 4\mu(1 - m_n - m_p)}}{2 - \bar{m}_p - \mu \bar{m}_n + \sqrt[2]{(\bar{m}_p + \mu \bar{m}_n)^2 - 4\mu(1 - m_n - m_p)}}$$

Let us define $\phi \equiv \sqrt[2]{(\bar{m}_p)^2 + (\mu \bar{m}_n)^2 + 2\bar{m}_p \mu \bar{m}_n - 4\mu \bar{m}_n - 4\mu \bar{m}_p + 4\mu}$. Hence the headcount ratio is:

$$HC = \frac{\bar{m}_p - \mu \bar{m}_n + \phi}{2 - \bar{m}_p - \mu \bar{m}_n + \phi}$$

The derivative of the headcount with respect to μ is:

$$\frac{\partial HC}{\partial \mu} = (-\bar{m}_n + \phi') 2 \frac{(1 - \bar{m}_p)}{(2 - \bar{m}_p - \mu \bar{m}_n + \phi)^2}$$

Whose sign depends on the sign of $-\bar{m}_n + \phi'$.

Hence we have:

$$\begin{aligned} \frac{\partial HC}{\partial \phi} &\geq 0 \iff \phi' \geq \bar{m}_n \\ &\iff \frac{1}{2} \frac{[2(\mu \bar{m}_n) \bar{m}_n + 2\bar{m}_p \bar{m}_n - 4\bar{m}_n - 4\bar{m}_p + 4]}{\left[(\bar{m}_p)^2 + (\mu \bar{m}_n)^2 + 2\bar{m}_p \mu \bar{m}_n - 4\mu \bar{m}_n - 4\mu \bar{m}_p + 4\mu\right]^{\frac{1}{2}}} \geq \bar{m}_n \\ &\iff \frac{[2(\mu \bar{m}_n) \bar{m}_n + 2\bar{m}_p \bar{m}_n - 4\bar{m}_n - 4\bar{m}_p + 4]}{2\phi} \geq \bar{m}_n \\ &\iff \frac{\bar{m}_n (\mu \bar{m}_n + \bar{m}_p) - 2\bar{m}_n - 2\bar{m}_p + 2}{\phi} \geq \bar{m}_n \end{aligned}$$

Since $HC = \frac{\bar{m}_p - \mu \bar{m}_n + \phi}{2 - \bar{m}_p - \mu \bar{m}_n + \phi} \rightarrow \frac{HC(2 - \bar{m}_p - \mu \bar{m}_n) - \bar{m}_p + \mu \bar{m}_n}{(1 - HC)} = \phi$.

Hence the condition can be written as:

$$\begin{aligned}
\frac{\bar{m}_n (\mu \bar{m}_n + \bar{m}_p) - 2\bar{m}_n - 2\bar{m}_p + 2}{\frac{HC(2 - \bar{m}_p - \mu \bar{m}_n) - \bar{m}_p + \mu \bar{m}_n}{(1 - HC)}} &\geq \bar{m}_n \\
(1 - HC) [\bar{m}_n (\mu \bar{m}_n + \bar{m}_p) - 2\bar{m}_n - 2\bar{m}_p + 2] &\geq \bar{m}_n [2HC - HC (\bar{m}_p + \mu \bar{m}_n) - \bar{m}_p + \mu \bar{m}_n] \\
[\bar{m}_n (\mu \bar{m}_n + \bar{m}_p) - 2\bar{m}_n - 2\bar{m}_p + 2] - HC [-2\bar{m}_p + 2] &\geq \bar{m}_n [-\bar{m}_p + \mu \bar{m}_n] \\
\bar{m}_n \bar{m}_p - \bar{m}_n - \bar{m}_p + 1 - HC [-\bar{m}_p + 1] &\geq 0 \\
\bar{m}_n (\bar{m}_p - 1) + (1 - \bar{m}_p)(1 - HC) &\geq 0 \\
1 - HC &\geq \bar{m}_n
\end{aligned}$$