BETA
Bureau
d'économie
théorique
et appliquée

Documents
de travail

«Counting the Missing Poor in Pre-Industrial Societies»

Auteurs

Mathieu Lefebvre, Pierre Pestieau, Gregory Ponthiere

Document de Travail n° 2021 – 29

Juin 2021

Bureau d'Économie Théorique et Appliquée BETA

www.beta-umr7522.fr

@beta_economics

Contact :
jaoulgrammare@beta-cnrs.unistra.fr

Counting the Missing Poor in Pre-Industrial Societies

Mathieu Lefebvre* Pierre Pestieau† Gregory Ponthiere‡

June 16, 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 aims at estimating 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, we show that the sign and size of the missing poor bias is sensitive to the degree of downward mobility for the non-poor.

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

JEL classification codes: I32.

*BETA University of Strasbourg.

†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 among historians that the prevalence of poverty was high in Pre-Industrial societies. (Lindert and Williamson 1982, Clark 2007, Ravallion 2016). Whereas standards of living differed across social classes and were fluctuating with wars and epidemics, the overall tendency during the Pre-Industrial period involved a large prevalence of poverty.

A classical reference on poverty in Pre-Industrial times is the work of Gregory King. In his social tables (1688), King (in Barnett 1936) provided a picture of England and Wales in the late 17th century. As shown on Table 1, which presents some figures from King's social tables (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 family income.[1] While those figures must be adjusted for the varying size of families across households, they give a raw idea of the large prevalence of poverty in Pre-Industrial England.[2]

	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.

Besides the large prevalence of poverty, 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. As documented by Clark and Hamilton (2006) on the basis of data from wills on reproductive success, social status and income in England (1585-1638), the richest testators had about twice more surviving children than the poorest. Other works confirmed the existence of an evolutionary advantage for the rich, such as Clark and Cummins (2015).

[1] Besides 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.

[2] The revised table proposed by Lindert and Williamson (1982) involves a smaller number of households belonging to the last three categories shown on Table 1. Lindert and Williamson argue that King overestimated the size of those categories, and provided accurate corrections.

Figure 2, which is based on Clark and Cummins (2015, Table 8), illustrates the existence of the evolutionary advantage of the rich, by comparing the number of surviving children per women by asset income tercile for two periods: 1500-1779 and 1780-1879. For the former period, the evolutionary advantage of the third tercile is substantial: net fertility equals more than 4 children, against 3 children for the second tercile, and less than 3 for the first tercile. However, net fertility outcomes are closer over 1780-1879, which suggests that the evolutionary advantage of the rich has weakened once industrialization has started.

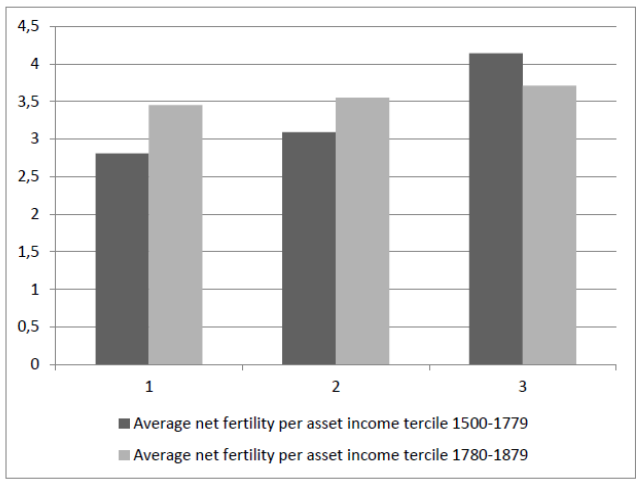

Figure 1: Net fertility by asset income tercile in England,
Clark and Cummins (2015).

Clark and Hamilton's findings can be interpreted as providing some (indirect) empirical support for an old thesis defended by Malthus (1798): the existence of positive and preventive "population checks" adjusting the population size to the means of subsistence. According to the Principle of Population, there exists a fundamental imbalance between the capacity of a society to produce humans, and its capacity to produce means of subsistence. As a consequence, the population is "checked" downwards, through preventive population checks (i.e. individuals reducing fertility to avoid famines) and positive population checks (i.e. a worsening of the survival conditions of the poorest, due to the lack of means of subsistence).[3] Poor classes are the main victims of population checks.[4] The smaller number of surviving children among poorer classes is compatible with the implications of the Malthusian doctrine.[5]

[3] See Malthus (1798), chapter 1, p. 71-72.
[4] See Malthus (1798), chapter 5, p. 93).
[5] Note that other works evaluating the Malthusian doctrine yield mixed results. See Wrigley (1969), Mokyr (1980), Clark (2007) and Crafts and Mills (2009).

The existence of an evolutionary advantage for the non-poor with respect to the poor raises some paradoxes for the measurement of poverty. Since poor individuals face worse survival conditions than the non-poor, poor individuals tend to be *under-represented* in the population under study. By dying earlier, poor persons tend to disappear more quickly from social statistics, pushing the measured poverty rate down. Thus, as stressed by Kanbur and Mukherjee (2007) and Lefebvre et al (2013), 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.

The missing poor bias has been largely documented in the recent years. Lefebvre et al (2018) showed that international comparisons of poverty can be affected by the existence of income-mortality gradients of various sizes across countries. For instance, while poverty above age 60 is higher in Portugal than in Estonia on the basis of standard headcount poverty rates, Lefebvre et al (2018) show that, once one takes the missing poor into account, by computing the hypothetical poverty measure that would prevail provided all individuals faced the same survival conditions (independently from their income), the ranking of countries in terms of poverty is reversed: corrected headcount poverty rates for the old are lower in Portugal than in Estonia.[6] Thus the missing poor bias introduces a noise in international comparisons of poverty.

Implications of the missing poor bias for *intertemporal* comparisons of poverty are less well-documented. The reason lies in the high complexity raised by the missing poor problem in an intergenerational perspective. 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 missing poor bias generates cumulative errors. The problem of cumulative measurement errors due to repeated selection biases may be acute for the Pre-Industrial era. The evolutionary advantage of the rich with respect to the poor is likely to have led to an under-estimation of the prevalence of poverty. In the light of cumulative errors due to repeated selection biases, measures of poverty during the Pre-Industrial period may be subject to substantial measurement errors.

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 model of population dynamics, 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

[6] The missing poor bias is also studied in Lefebvre et al (2019).

calibrate that model and compare the actual prevalence of poverty during Pre-Industrial England with the hypothetical prevalence that would have obtained provided all individuals - poor and non-poor - had benefited from the same demographic conditions, the evolutionary advantage of the latter being 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 the downward mobility of the non-poor is high, a higher evolutionary advantage for the non-poor can *increase* measured poverty. The consequences of repeated selection effects on measured poverty are such that, under some conditions, the evolutionary advantage of the non-poor leads to an overestimation of the extent of poverty. A second result lies on the quantitative side. The comparison of the standard headcount poverty rate and the hypothetical headcount 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 mobility for the non-poor.

This paper is related to several branches of the literature. First, this is related to the economic history literature on poverty during the Pre-Industrial period, such as Lindert and Williamson (1982, 1983) and Clark (2007).[7] This is also related to the contributions of Clark and Hamilton (2007) and Clark and Cummins (2015) on the existence of an evolutionary advantage for the rich during the Pre-Industrial era. Our work is also related to the literature on the measurement of poverty under income-differentiated mortality, such as Kanbur and Mukherjee (2007), Lefebvre et al (2013, 2018, 2019). Finally, our results can also be interpreted in the light of the increasingly large literature on evolutionary growth theory (Galor and Moav, 2002, 2005, Galor 2010, 2011).

The rest of the paper is organized as follows. Section 2 presents a simple matrix population model. The long-run headcount poverty rate is derived in Section 3. Section 4 studies, from a theoretical perspective, the effect of the evolutionary advantage of the non-poor over the poor on the measurement of poverty in the long-run. Section 5 presents our method for the adjustment of the headcount poverty rate so as to count all the missing poor individuals and their descendants. Section 6 uses data on poverty in Pre-Industrial England to provide a comparison of the actual prevalence of poverty with the hypothetical poverty that would have prevailed provided non subpopulation had enjoyed an evolutionary advantage. Conclusions are left to Section 7.

2 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}

[7] Note that this paper does not address the Industrial period (Polak and Williamson 1991). Due to the demographic transition (Lee 2003), the theoretical setting that we use would have to be amended to account for variations in survival and fertility conditions. Our analysis thus focuses only on the Pre-Industrial period.

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 conditions; (iii) social mobility.[8] 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 young adulthood. Let us denote by $m_p \in]0,1[$ the probability for a child born in a poor family to escape from poverty, and by $\bar{m}_p$ the probability that a child born in a poor family remains poor.[9] Finally, let us denote by $m_n \in]0,1[$ the probability for a child born in a non-poor family to fall into poverty, and by $\bar{m}_n$ the probability that a child born in a non-poor family remains non-poor.[10]

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, on the contrary, escaping from poverty. To see this, let us define the matrix $\mathbf{M}$ as follows:

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

By construction, the matrix $\mathbf{M}$ can be used to obtain the structure of the population in terms of income at period $t+1$ from the structure of the population in terms of income at period t, by premultiplying the latter by the matrix $\mathbf{M}$:

$$\mathbf{M}.\mathbf{N}_t = \mathbf{N}_{t+1} \tag{1}$$

This expression can be rewritten in detailed form as:

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

In order to understand matrix $\mathbf{M}$, let us write the number of poor adults at time $t+1$ as follows:

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

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 (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 (second term of the RHS).

[8] We consider here a closed economy and we abstract from migrations.
[9] We have $\bar{m}_p = 1 - m_p$.
[10] We have $\bar{m}_n = 1 - m_n$.

3 Long-run poverty

The model developed in Section 2 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.

$$HC_t = \frac{N_{pt}}{N_{pt} + N_{nt}} \tag{4}$$

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. 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 clearly the case for matrix $\mathbf{M}$.

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. ■

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.[11] This is called the "dominant eigenvalue". According to the Strong Ergodic Theorem, that dominant eigenvalue determines the ergodic properties of population growth.[12] 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.

---

[11] See Caswell (2001), p. 83-84.

[12] See Caswell (2001), p. 84-85.

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

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

while the associated long-run headcount poverty rate is

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

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. Quite interestingly, 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 at all on the level of 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.

4 Malthusian economies

Recent empirical evidence, such as Clark and Hamilton (2006) and Clark and Cummins (2015) identified sizeable differentials in the number of surviving children per adult in Pre-Industrial England. Those authors used data on wills to demonstrate that rich individuals had a larger number of surviving children than poorer ones. In terms of our model, this empirical evidence implies that the numbers of children surviving to adulthood satisfy:

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

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

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.[13] One can regard this as empirical evidence supporting the existence of an "evolutionary advantage" for the rich in Pre-Industrial societies.

[13] On variations of the evolutionary advantage of the rich in Pre-Industrial times, see Clark (2007). Clark argues that it is in England that this advantage was the largest, and that this may explain why the Industrial Revolution took place in England and not in other countries such as the Netherlands or France.

Note that the term "evolutionary advantage" must be used with caution, since the size of the gap, in terms of number of surviving children, between the poor and the non-poor has nothing natural, but depends on existing institutions, and, as such, can vary across epochs and societies, as this was already illustrated in Section 1 (Figure 1).[14] We are thus far from what Darwin (1859) called a "natural selection" process: there is nothing "natural" in having societies where the number of surviving children varies a lot or not across income groups.

In Pre-Industrial England, there existed, since the late 16th century, a parish-level system of social assistance: the Poor Laws. The effects of that system of social assistance have been the object of lots of debates. In his *Essay on the Principle of Population*, Malthus (1798) argued that the Poor Laws were not reducing, but were extending the scope of poverty, on the grounds of (i) encouraging fertility among low-income classes, hence extending the population size quicker than the available means of subsistence (thus reinforcing the imbalance stated in his Principle of Population), and (ii) discouraging the spirit of entrepreneurship, leading to a reduction of the available means of subsistence.[15] According to Malthus, those two perverse effects reinforced positive population checks, leading to a larger overmortality among poor classes. Those arguments were criticized by Marx (1858), who questioned the reliance on the general concepts of "means of subsistence" and of "overpopulation". Marx argued that the "means of subsistence" can only be defined *relatively* to the distribution of earnings, whereas the concept of "overpopulation" is relative to a particular mode of production.[16] Hence, according to Marx, the overmortality of the poor has no relation with the population size, and is mainly due to problems of distribution.

In this paper, we will not examine the empirical issue of the determinants of the "evolutionary advantage" of the rich in Pre-Industrial societies, nor explore whether or not Malthus's arguments against the Poor Laws were correct. We will take the existence of an "evolutionary advantage" for the non-poor as given, and examine the consequences of that evolutionary advantage for the measurement of poverty across long periods of time. The question that we raise here 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? Proposition 3 provides the first elements of an answer.

Proposition 3 *In Malthusian economies, the long-run headcount poverty rate is:*

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

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

[14] Clearly, the modern Welfare State, by providing social insurance to citizens against many risks of life (disease, unemployment, etc.) is likely, under some conditions, to decrease the "evolutionary advantage" of the rich in comparison to a pure laissez-faire world

[15] See Malthus (1798), chapter 5, p. 97.

[16] See Marx (1858), p. 563-565.

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:

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

Proof. See the Appendix. ■

Proposition 3 states that the long-run level of the headcount poverty rate in Malthusian economies 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 $\{m_p, m_n\}$. Quite interestingly, the long-run prevalence of poverty in Malthusian economies depends on survival conditions and fertility *only insofar* as they determine the strength of the evolutionary advantage of the non-poor, but not otherwise.

The second part of Proposition 3 states a quite interesting result, which could not have been anticipated without a rigorous 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. Actually, based on the condition stated in Proposition 3, 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 $1 - HC$, then 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.

That result is original, since one might be tempted to believe, at first glance, that a stronger evolutionary advantage of the non-poor over the poor would mechanically reduce the long-run prevalence of poverty, by making the non-poor more and more numerous within the population. But this belief can be misleading. True, when there is little downward mobility, a stronger evolutionary advantage of the non-poor will push the long-run poverty rate down. But this is not necessarily true when there is a larger downward mobility.

5 Counting the missing poor

Under income-differentiated mortality, standard measures of poverty suffer from a selection bias: the poor being subject to harsher survival conditions than the non-poor, the poor tends to be under-represented in the populations under study, leading to a downward bias in poverty measures (see Kanbur and Mukherjee 2007, Lefebvre et al 2013).

When considering static economies, a simple way to quantify the extent of the selection bias due to income-differentiated mortality 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. 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 more prematurely than the non-poor - and to add these missing poor persons to the population under study. Quite interestingly, in the presence of income mobility, the exercise of counting the missing poor is complex, since in case of survival a poor person may have moved upwards in the income scale, and may have escaped from poverty (see Lefebvre et al 2019).

Counting the missing poor becomes even more difficult when considering an intergenerational context. The reason is 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.

In order to quantify the selection bias due to income-differentiated mortality 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 that would have prevailed provided μ is fixed to unity. Let us define that hypothetical long-run poverty rate as:

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

The hypothetical poverty headcount ratio HC^H can be interpreted as follows. 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 provided all missing poor individuals (and their descendants) had been added to the population and had been properly counted as poor. Thus the hypothetical headcount poverty rate HC^H does not depend at all on differences of survival conditions or fertility conditions across income groups. By construction, 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 next section uses data on the Pre-Industrial period in order to examine which case prevails. Focusing on the Pre-Industrial period makes a lot of sense when considering the comparison of long-run poverty rates, since those measures of poverty (corrected or not) are conditional on the existence of a mobility matrix $\mathbf{M}$ whose elements are *constant* over time (at least in trend). Thus one cannot make any comparison for periods where there were large changes in the fertility or mortality trends by groups. This fixity 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 demographic parameters $\{f_n, f_p, s_n, s_p\}$.

6 Counterfactuals in History

To what extent did the evolutionary advantage of the non-poor over the poor affect the prevalence of poverty under 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 all components of the measured headcount poverty ratio $\{m_p, m_n, \mu\}$. Then, in a second stage, we use those calibrated parameters to compute the hypothetical long-run poverty measure HC^H in Pre-Industrial England, and we compare this hypothetical poverty measure with the standard one.

As far as the long-run prevalence of poverty in Pre-Industrial England is concerned, it is difficult to come with a single number: the extent of poverty has been fluctuating year after year, due to events such as wars and epidemics. Moreover, there is also a more conceptual problem: how can one fix a plausible poverty line for long periods of time? We will not try here to provide an answer to that question, which go far beyond the scope of this paper. Our analysis does not aim at estimating the "true" or "correct" poverty measure for Pre-Industrial England, but only aims at examining how the evolutionary advantage of the non-poor over the poor affected the measurement of poverty in that society. For that particular purpose, our calculations will be based on two benchmark values, which are related to the figures in King's social tables (1688), potentially 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 in pre-Industrial England. 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.[17]

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 $\mu = 1,286$.

Finally, 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, especially 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. For our computations, we will take two values for m_n: a lower bound equal to $m_n = 0.100$ (low downward mobility) and an upper bound equal to $m_n = 0.300$ (high downward mobility).

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. When $HC = 24.2$ %, we have that, when $m_n = 0.100$, $m_p = 0.186$, and that, when $m_n = 0.300$, $m_p = 0.992$. When $HC = 48.2$ %, we have that, when $m_n = 0.100$, $m_p = 0.001.$, and that, when $m_n = 0.300$, $m_p = 0.266$. Those combinations of structural parameters consists of various ways to replicate or "rationalize" 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 rationalize a higher headcount poverty rate, one needs to assume, for a given degree of downward

[17] It 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 (1992) and Ravallion (2006), which lie around 40-45 %.

mobility for the non-poor (i.e. a given parameter m_n) a lower degree of upward income mobility for the poor (i.e. a lower parameter m_p). For instance, under a low downward mobility for the non-poor ($m_n = 0.100$), a probability of leaving poverty of 0.186 can rationalize a headcount poverty rate of 24.2 %, but to rationalize $HC = 48.2$ % one needs, under $m_n = 0.100$, a lower upward mobility for the poor, m_p being extremely low (0.001). Second, it should also be stressed that, for a given headcount poverty rate, 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 the headcount poverty rate remains unchanged.

Let us now 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. Actually, the condition of Proposition 3 regarding the sign of the derivative of HC with respect to μ tells us that reducing μ reduces measured poverty when $1-HC > \bar{m}_n$, and that reducing μ raises measured poverty when $1 - HC < \bar{m}_n$. Our calculations involve four distinct calibrations of $\{1 - HC, \bar{m}_n\}$: $(0.758, 0.900)$, $(0.758, 0, 700)$, $(0.518, 0.900)$ and $(0.518, 0.700)$. From Proposition 3, we can deduce that the condition $1 - HC > \bar{m}_n$ holds in case $(0.758, 0, 700)$, and that in the other three cases the condition $1-HC < \bar{m}_n$ holds. In the first case, the hypothetical headcount ratio HC^H is lower than the standard headcount HC, whereas the opposite holds in the three other cases.

Let us now examine to what extent the hypothetical headcount ratio differs from the standard one, that is, to what extent did the evolutionary advantage of the non-poor affect poverty measurement. Figures 2 and 3 summarize our results for the case where the headcount poverty rate takes, respectively, its lower bound value (24.2 %) and its upper bound value (48.2 %).

On Figure 2, 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 mobility for the non-poor, that is, on the probability for the children of the non-poor to fall into poverty. When the postulated 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 for the non-poor, 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 rationalize 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.

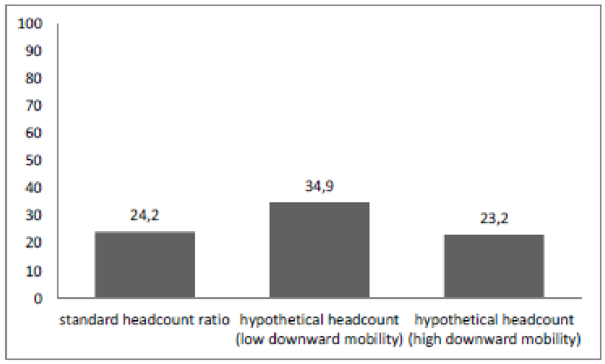

Figure 2: Standard headcount poverty rate and hypothetical headcount poverty rates under the low benchmark for poverty in Pre-Industrial England.

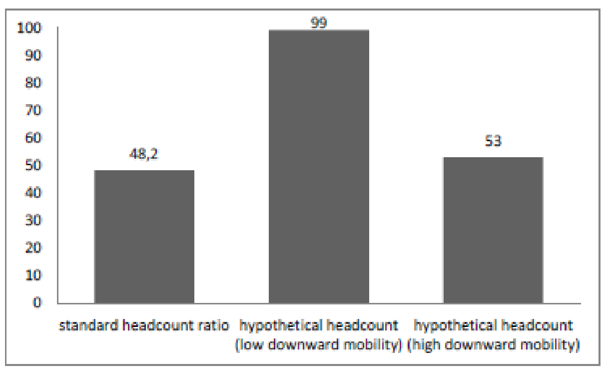

Figure 3: Standard headcount poverty rate and hypothetical headcount poverty rates under the high benchmark for poverty in Pre-Industrial England.

Let us now compare those results with the ones obtained under a high standard headcount poverty rate (Figure 3). In that case, the hypothetical headcount ratio is always higher than the standard one. Hence, if one takes as a benchmark a high prevalence of poverty during the Pre-Industrial period, the evolutionary advantage of the non-poor over the poor has pushed measured poverty down. It should be stressed, however, that the extent to which standard poverty measures have been biased downwards because of the evolutionary advantage of the non-poor depends on the degree of downward mobility for the non-poor. If the postulated downward mobility for the non-poor is low, HC^H is about 50 percentage points higher than HC, whereas if the postulated downward mobility for the non-poor is high, the hypothetical poverty rate HC^H is about 5 percentage point higher than HC.

How can one interpret this extreme lack of robustness? The intuition behind those results is that, in order to rationalize 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 mobility of the non-poor, the selection bias induced by the evolutionary advantage of the non-poor had a high effect on the measurement of poverty, and tended to reduce it by one half. 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, the hypothetical poverty rate HC^H is only slightly higher than the standard HC.

All in all, our exploratory calculations suggest that the extent to which the evolutionary advantage of the non-poor over the poor in Pre-Industrial England has led to large or small selection bias in poverty measurement is sensitive to (i) the postulated degree of downward mobility for the non-poor (who enjoyed the evolutionary advantage); (ii) the postulated prevalence of poverty in Pre-Industrial times. Under a low prevalence of poverty, the correction of the selection bias in poverty measurement leads to either decrease measured poverty under a high downward mobility, and to increase it by 10 percentage points under a low downward mobility, whereas, under a high prevalence of poverty, the correction varies from + 5 percentage points (under a high downward mobility) to + 50 percentage points (under a low downward mobility).

7 Conclusions

Pre-Industrial societies being characterized by a large prevalence of poverty and by a significant evolutionary advantage of the non-poor over the poor (Clark and Hamilton 2006, Clark and Cummins 2015), one may expect that measures of poverty during that 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 the selection bias across generations, one may also expect that the impact of the evolutionary advantage of the non-poor on measures of poverty turns out to be substantial.

In order to quantify the size of the missing poor bias for 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 prevalence of poverty as the eigen vector associated to the dominant eigenvalue of the population matrix.

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

A second finding is of quantitative nature: the comparison of the standard headcount poverty rates in Pre-Industrial England with the hypothetical measures of poverty that would have prevailed provided there had been no evolutionary advantage for the non-poor over the poor suggests that the sign and size of the missing poor bias is sensitive to the postulated degree of downward mobility for the non-poor. Under a low downward mobility for the non-poor, the missing poor bias lies between 10 and 50 percentage points, whereas under a high downward mobility for the non-poor, the missing poor bias may be either slightly negative, or of about 5 percentage points. Those findings suggest that the extent to which the evolutionary advantage of the non-poor over the poor affected the measurement of poverty in Pre-Industrial societies depends on the patterns of social mobility.

Quite interestingly, our findings contribute not only to the - widely debated - issue of poverty measurement, but, also, cast some light on the increasingly large literature 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. According to Clark (2007), a possible reason why the Industrial Revolution took place in England - and nowhere else - may lie in the existence of a stronger evolutionary advantage for the rich in that society, which has contributed to diffuse high skills in all segments of the society, and not just in some highly-qualified jobs.

Our findings tend, to some extent, to qualify those claims. Our calculations suggest that a rise in the evolutionary advantage of the non-poor does not necessarily lead to a reduction of long-run poverty. If the downward income mobility is high, a stronger evolutionary advantage for the non-poor may not be at the origin of a lower long-run poverty rate, but may be associated to a larger long-run poverty. Whether or not a higher evolutionary advantage for the non-poor over the poor leads to a lower measured poverty depends on the degree of downward income mobility. In Pre-Industrial societies, downward income 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.

8 Bibliography

Bamett, G. (1936), *Two Tracts by Gregory King.* Baltimore: Johns Hopkins University.Press.

Bourguignon, F., Morrisson, C. (2002). World distribution among world citizens: 1820-1992. *American Economic Review*, 92: 727-744.

Caswell, H. (2001). *Matrix Population Models. Construction, Analysis, and Interpretation*. Sinauer, Massachussets.

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.

Darwin, C. (1859). *On the Origin of Species*. London.

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.

Kanbur, R., Mukherjee, D. (2007). Premature mortality and poverty measurement. *Bulletin of Economic Research*, 59: 339-359.

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. (2019). Missing poor and income mobility. *Journal of Comparative Economics*, 47: 330-366.

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.

Marx, K. (1858).*Economic Manuscripts 1857-1858*, in MEGA.

Mokyr, J. (1980). Malthusian models and Irish history. *Journal of Economic History*, 40: 159-166.

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.

Wrigley, E.A. (1967). *History and Population.* Weidenfeld and Nicolson, London.

9 Appendix

9.1 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:

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

Note that Δ can be rewritten as:

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

Hence the two eigenvalues are:

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

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:

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

Hence we have

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

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

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

From which it follows that

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

Hence the eigen vector associated to λ_1 is

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

From the Strong Ergodic Theorem, we have that

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

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:

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

9.2 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:

$$\begin{aligned} HC &= \frac{f_p s_p \left(\bar{m}_p - \mu \bar{m}_n\right) + \sqrt[2]{\left(f_p s_p \bar{m}_p + \mu f_p s_p \bar{m}_n\right)^2 - 4\mu (f_p s_p)^2 \left(1 - m_n - m_p\right)}}{f_p s_p \left(2 - \bar{m}_p - \mu \bar{m}_n\right) + \sqrt[2]{\left(f_p s_p \bar{m}_p + \mu f_p s_p \bar{m}_n\right)^2 - 4\mu (f_p s_p)^2 \left(1 - m_n - m_p\right)}} \\ &= \frac{\bar{m}_p - \mu \bar{m}_n + \sqrt[2]{\left(\bar{m}_p + \mu \bar{m}_n\right)^2 - 4\mu \left(1 - m_n - m_p\right)}}{2 - \bar{m}_p - \mu \bar{m}_n + \sqrt[2]{\left(\bar{m}_p + \mu \bar{m}_n\right)^2 - 4\mu \left(1 - m_n - m_p\right)}} \end{aligned}$$

Regarding the effect of μ on HC, let us define:

$$\phi \equiv \sqrt[2]{\left(\bar{m}_p\right)^2 + \left(\mu \bar{m}_n\right)^2 + 2 \bar{m}_p \mu \bar{m}_n - 4 \mu \bar{m}_n - 4 \mu \bar{m}_p + 4\mu}$$

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:

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

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

Hence we have:

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

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

Hence the condition can be written as:

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