11-27-2016

Working Paper Number 16010

Malthusian Stagnation is Efficient

Juan C. Cordoba
Iowa State University, cordoba@iastate.edu

Xiying Liu
Wuhan University

Malthusian Stagnation is Efficient

Abstract
Malthusian economies are generally deemed inefficient: stagnated, highly unequal, and densely populated by a labouring class prone to high fertility. This article defines and characterizes efficient allocations in Malthusian environments of fixed resources and endogenous fertility. We show, that under general conditions, efficient allocations exhibit stagnation in standards of living, inequality, differential fertility, and a high population density of poorer individuals.

Keywords
Efficiency, Endogenous fertility, stochastic abilities

Disciplines
Growth and Development | Macroeconomics

Malthusian Stagnation is Efficient

Juan Carlos Cordoba*

Xiying Liu†

First version: 02/2016. This version: 11/2016.

Malthusian economies are generally deemed inefficient: stagnated, highly unequal, and densely populated by a labouring class prone to high fertility. This article defines and characterizes efficient allocations in Malthusian environments of fixed resources and endogenous fertility. We show, that under general conditions, efficient allocations exhibit stagnation in standards of living, inequality, differential fertility, and a high population density of poorer individuals.

Keywords: Efficiency, endogenous fertility, stochastic abilities.

JEL Classification: D04, D10, D63, D64, D80, D91, E10, E60, I30, J13, N00, 011, 040, Q01.

I. Introduction

Our understanding of the preindustrial era, as well as the issue of underdevelopment, is strongly influenced by Malthus' ideas. Reacting to the idealistic

*Department of Economics, Iowa State University, 279 Heady Hall, Ames IA 50010, e-mail: cordoba@iastate.edu.

†Economics and Management School of Wuhan University, Wuhan, Hubei, China, 430072, e-mail: liuxiyingbly@gmail.com.

writings of Thomas Godwin on the perfectibility of man, Malthus formulated a pessimistic view of the world where widespread poverty is an unavoidable consequence of a fundamental law of nature: the passion between the sexes. Malthus convincingly argued that any abundance of resources ultimately resolves into a larger but not richer population.[1] While Malthus himself advocated for moral constraints and the dismantling of existing Poor Laws in England as ways to mitigate the problem, his ideas naturally lead to other sometimes radical solutions: the birth control movement of the 1800s, eugenics and Social Darwinism in the late 1800s and early 1900s, modern family planning policies and the one-child policy among many others. Family planning remains central to development efforts of institutions such as the World Bank or the Gates Foundation.

Malthusian models have recently gained renewing interest as part of a literature that seeks to provide a unified theory of economic growth, from prehistoric to modern times (Becker, Murphy, and Tamura 1990, Jones 1999, Galor and Weil 2000, Lucas 2002, Hansen and Prescott 2002, and Doepke 2004 are just some few examples). The focus of this literature has been mostly positive rather than normative: to describe mechanisms for the stagnation of living standards even in the presence of technological progress. But the fundamental issue of efficiency in Malthusian economies, which is key to formulating policy recommendations and better understand the extent to which the "Malthusian trap" could have been avoided, has received scarce attention.

This article defines and characterizes efficient allocations in Malthusian economies using modern tools of welfare economics. The focus is solely on the Malthusian era, before the onset of modern economic growth. This is not a standard exercise. Although welfare and efficiency properties of models with exogenous population, such as the Neoclassical Growth model, are well understood, wel-

1. Ashraf and Galor (2011) recently provide evidence supporting the validity of Malthus' predictions for a number of pre-industrial economies.

fare and efficiency properties of models with endogenous population, such as the Malthus model, remains largely unexplored, with some few exceptions as we discuss below.

Our model economy is populated by a large number of finitely-lived fully rational individuals who are altruistic toward their descendants. Individuals are of different types and a type determines characteristics such as labor skills, the rate of time preference, and ability to raise children. Types are stochastic and determined at birth. We formulate the problem faced by a benevolent social planner who cares about the welfare of all potential individuals, present and future. The planner directly allocates consumption and children to individuals in all generations and states subject to aggregate resource constraints, promise keeping constraints, population dynamics and a fixed amount of land. The economy is closed, there is no capital accumulation nor migration. Furthermore, there are no underlying frictions such as private information or moral hazard so that the focus is on first-best allocations.

We consider the extent to which efficient allocations can rationalize three key aspects of Malthusian economies: (i) stagnation of individual consumption in the presence of technological progress and/or improvements in the availability of land; (2) social classes, inequality, and widespread poverty; (3) differential fertilities. The core of the paper focuses on steady state, or stationary, allocations while issues of stability are left for the Appendix.

The following are the main findings. First, we show that stagnation is efficient. Specifically, steady-state consumption is independent of the amount of land, and under general conditions, the level of technology. As a result, land discoveries, as the ones discussed by Malthus, lead to more steady-state population but no additional consumption. An identical prediction holds true for technological advancements as long as the production technology is Cobb-Douglas or technological progress is land-augmenting, the type of progress that is more needed because land is the limiting factor.

The source of the stagnation is a well-known prediction of endogenous fertility models, according to which optimal consumption is proportional to the net costs of raising a child. For example, Becker and Barro find that "when people are more costly to produce, it is optimal to endow each person produced with a higher level of consumption. In effect, it pays to raise the 'utilization rate' (in the sense of a higher c) when costs of production of descendants are greater (Becker and Barro, 1988, pg. 10)." We show that this link between optimal consumption and the net cost of raising children also holds true for a benevolent planner and under more general conditions. The crux of the proof of stagnation is to show that neither land discoveries nor technological progress alter the steady-state net cost of raising a child, and in particular, the marginal product of labor, which is required to value both the parental time costs of children and children's marginal output.

Second, we show that efficient allocations exhibit social classes. Only types with the highest rate of time preference have positive population shares and consumption in steady state. Furthermore, and unlike the exogenous fertility case, it is generally not efficient to equalize consumption among types, even if their Pareto weights are identical, nor to eliminate consumption risk. Efficient consumption is stochastic even in the absence of aggregate risk. These results are further implications of consumption being a function of the net cost of raising children. Poor individuals in an efficient allocation are the ones with the lowest net costs of raising children.

Third, there is an inverse relationship between consumption and population size: the lower the consumption of a type the larger its share in the population. As a result, there are more poor individuals than rich individuals in any efficient allocation. Furthermore, population inequality is larger than consumption inequality by a factor that depends positively on the elasticity of parental altruism to the number of children and negatively on the intergenerational elasticity of substitution.[2]

Fourth, fertility differs among types. Optimal fertility depends on parental types but also on grandparent types. Given grandparent types, parents with low consumption have more children than parents with high consumption. Moreover, given parent types, consumption rich grandparents have more grandchildren than consumption poor grandparents.

Fifth, steady state allocations, and in particular the land-labor ratio, generally depends on initial conditions. The efficient steady state depends on the initial distribution of population and on Pareto weights. This is unlike the neoclassical growth model in which the efficient capital-labor ratio, or modified golden rule level of capital, is independent of initial conditions and Pareto weights. Malthusian economies thus do not exhibit a clear separation between efficiency and distribution.

Our results help explain why the so-called Malthusian trap was so pervasive in pre-industrial societies. We show that even in the best case scenario of an economy populated by loving rational parents, and governed by an all powerful benevolent rational planner, stagnation would still naturally arise, as well as social classes and differential fertility. Our results also show that is not the irrational animal spirit of human beings, as suggested by Malthus, what ultimately explains the stagnation. Stagnation can be the result of an optimal choice between the quality and quantity of life in the presence of limited natural resources.

Our paper is related to Golosov, Jones and Tertilt (2007) who have shown

2. The intergenerational elasticity of substitution is analogous to the intertemporal elasticity of substitution but applied to different generations rather than different periods. See Cordoba and Ripoll (2014).

that population is efficient in dynastic altruistic models of endogenous fertility and fixed land. The focus of their paper is not the Malthusian era and therefore they do not derive results about stagnation, the distribution of consumption and population, nor differential fertility. Moreover, they elaborate on the Pareto concept of efficiency while we study efficiency from the point of view of utilitarian social planners. Lucas (2002) studies equilibrium in Malthusian economies populated by altruistic fully rational parents and shows that stagnation arises under certain conditions. His focus is on simple representative economies where fertility is equal across groups in steady state. Lucas discusses the difficulties in generating social classes, and is able to generate classes by assuming heterogeneity in the degree of time preference and binding saving constraints. As a result, the equilibrium with social classes is not efficient in his model. We are able to generate efficient social classes and differential fertility by allowing individuals to differ in their labor skills and costs of raising children.

Our paper also relates to Dasgupta (2005) who studies optimal population in an endowment economy with fixed resources. He does not consider the cost of raising children and focuses on the special case of generation-relative utilitarianism. Our model is richer in production, altruism, and the technology of raising children. Nerlove, Razin and Sadka (1986) shows that the population in the competitive equilibrium is efficient under two possible externalities. First, a larger population helps to provide more public goods such as national defense. Second, larger population reduces wage rate if there is a fixed amount of land. Eckstein, Stern, and Wolpin (1988) show that population can stabilize and non-subsistence consumption arises in the equilibrium when fertility choices is endogenously introduced to a model with fixed amount of land. Parents exhibit warm glow altruism while our paper builds on pure altruism. Peretto and Valente (2011) also use warm glow altruism to perform positive analysis on the interaction between resources, technology and population. De la Croix (2012) studies sustainable population by proposing non-cooperative bargaining between clans living on an island with limited resources. Children in his model act like an investment good for parents' old-age support.

The rest of the paper is organized as follows. Section 2 sets up the general stochastic model. Section 3 studies the deterministic representative agent version of the general model and derives the main stagnation results. Section 4 considers deterministic heterogeneity and derives key results regarding the distribution of population and consumption across types, as well as the importance of initial conditions for the steady state. Section 5 studies the full stochastic model and derives the key result for differential fertility, consumption, and population. Section 6 concludes. All proofs are provided in the Appendix.

II. The model

The production technology is described by the function $F(\bar{K}, L; A)$ where $\bar{K}$ is a fixed amount of land, L is labor and A is a technological parameter. F is constant returns to scale in $\bar{K}$ and L. Let $\alpha\left(\frac{\bar{K}}{L}, A\right) \equiv \frac{F_K(\bar{K},L,A)\bar{K}}{F(\bar{K},L,A)}$ be the land share of output. The economy is populated by large numbers of dynastic altruistic individuals who live for two periods, one as a child and one as an adult. Children do not consume. Individuals are heterogeneous in terms of their labor skills, rate of type preferences, and ability to raise children. In particular, individuals draw a random signal, or type, $\omega \in \Omega \equiv \{\omega_1, \omega_2, ..., \omega_K\}$, upon birth which defines his or her type. Effective labor supply, $l(\omega)$, degree of altruism, $\Phi(n, \omega)$, and the goods and time costs of raising a child, $\eta(\omega)$ and $\lambda(\omega)$, are then functions of an individual's type. n is the number of children. Signals are drawn from the Markov chain $\pi(\omega', \omega) = \Pr(\omega_{t+1} = \omega' | \omega_t = \omega)$ where ω_t is parent's type and ω_{t+1} is child's type. Assume π is irreducible. Let

$\omega^t = [\omega_0, \omega_1, ..., \omega_t] \in \Omega^{t+1}$ represent a particular family history of signals up to time t while $c_t(\omega^t)$ and $n_t(\omega^t)$ denote consumption and fertility of an individual with that family history.

II.A. Resource constraints

Let $N_t(\omega^t)$ be the population with history ω^t and $N_t \equiv \sum_{\omega^t} N_t(\omega^t)$ be total population at time t. Initial levels of population of each type, $N_0(\omega_i)$, $\omega_i \in \Omega$, are given. Assuming a law of large numbers, the population with history $\omega^t \in \Omega^{t+1}$ is described by

$$N_t\left(\omega^t\right) = N_{t-1}\left(\omega^{t-1}\right) n_{t-1}\left(\omega^{t-1}\right) \pi\left(\omega_t, \omega_{t-1}\right) \text{ for } t \geq 0. \tag{1}$$

Fertility rates are assumed to be subject to a biological maximum $\overline{n}$. The potential population at history ω^t is therefore $\overline{N}_t(\omega^t) = \overline{N}_{t-1}\left(\omega^{t-1}\right)\overline{n}\pi(\omega_t, \omega_{t-1})$ with $\overline{N}_0\left(\omega^0\right) = N_0(\omega_0)$. Aggregate labor supply satisfies

$$L_t = \sum_{\omega^t} N_t\left(\omega^t\right) l\left(\omega_t\right)\left[1 - \lambda_t\left(\omega_t\right) n_t\left(\omega^t\right)\right] \text{ for } t \geq 0, \tag{2}$$

where $l(\omega_t)\left[1 - \lambda_t(\omega_t) n_t(\omega^t)\right]$ is effective individual labor supply of a particular type once time costs of raising children and individual's ability are taken into account. Finally, aggregate resource constraints are given by

$$F(\bar{K}, L_t; A) = \sum_{\omega^t} N_t\left(\omega^t\right)\left[c_t\left(\omega^t\right) + \eta\left(\omega_t\right) n_t\left(\omega^t\right)\right] \text{ for } t \geq 0. \tag{3}$$

II.B. Individual welfare

Parents are assumed to be altruistic toward their children. The lifetime utility of an individual born at time $t \geq 0$, history ω^t, $U_t(\omega^t)$, is of the expected-utility type:

$$U_t\left(\omega^t\right)=\begin{array}{c}u\left(c_t\left(\omega^t\right)\right)+\Phi\left(n_t\left(\omega^t\right),\omega\right)E\left[U_{t+1}\left(\omega^{t+1}\right)|\omega^t\right]\\+\left(\Phi\left(\overline{n},\omega\right)-\Phi\left(n_t\left(\omega^t\right),\omega\right)\right)\underline{U},\end{array}\tag{4}$$

where $u\left(\cdot\right)$ is the utility flow from consumption, $\Phi\left(\cdot,\omega\right)$ is the weight that a parent of type ω attaches to the welfare of her n born children, $\Phi\left(\overline{n},\omega\right)-\Phi\left(n,\omega\right)$ is the weight attached to the unborn children, $E\left[U_{t+1}\left(\omega^{t+1}\right)|\omega^t\right]$ is the expected utility of a born child conditional on parental history and $\underline{U}$ is the utility of an unborn child as perceived by the parent. Function u satisfies $u'>0$ and $u''<0$. The population ethics literature refers to $\underline{U}$ as the "neutral" utility level, a level above which a life is worth living (Blackorby *et al.* 2005, pg. 25).

Equation (4) describes parents as social planners at the family level. This is particularly clear in the special case $\Phi\left(n,\omega\right)=n$. The more general function $\Phi\left(\cdot,\omega\right)$ allows for flexible weights and time discounting. While $\Phi\left(n,\omega\right)$ is the total weight of the n born children, $\Phi_n\left(n,\omega\right)$ is the marginal weight assigned to child $n\in\left[0,\overline{n}\right]$. We assume $\Phi_n\left(n,,\omega\right)>0$ and $\Phi_{nn}\left(n,\omega\right)\leq 0$ so that parents are altruistic toward each child and altruism is non-increasing. These preferences are discussed in Cordoba and Ripoll (2011) who show that (4) satisfies a fundamental axiom of altruism. Specifically, parental utility increases with the number of born children if and only if children are better off born than unborn in expected value, that is, $E\left[U_{t+1}\left(\omega^{t+1}\right)|\omega^t\right]>\underline{U}$.

Let $\beta\left(\omega\right)\equiv\Phi\left(1,\omega\right)$ be the discount factor, $\xi\left(c\right)\equiv\frac{u'(c)c}{u(c)}$ be the elasticity of the utility flow and $\psi\left(n,\omega\right)\equiv\frac{\Phi'(n,\omega)n}{\Phi(n,\omega)}$ be the elasticity of the altruistic function. Barro-Becker preferences are an special case obtained when $u(c)=\frac{c^{\xi}}{\xi}$, $\Phi\left(n,\omega\right)=\beta n^{\psi}$, $\underline{U}=0$, $\xi\in\left(0,1\right)$ and $\psi\in\left(\xi,1\right)$.

II.C. Social Welfare

The planner is envisioned as the ultimate parent, someone who cares about the welfare of all potential individuals in the society. Consistent with (4), it is natural to consider a social welfare function that takes the following generalized total utilitarian form:

(5)
$$\sum_{t=0}^{\infty} \delta^t \left[\sum_{\omega^t} \Psi\left(N_t\left(\omega^t\right)\right) U_t\left(\omega^t\right) + \left(\sum_{\omega^t} \Psi\left(\overline{N}_t\left(\omega^t\right)\right) - \sum_{\omega^t} \Psi\left(N_t\left(\omega^t\right)\right) \right) \underline{U} \right].$$

The parameter $0 < \delta < 1$ reflects time discounting while the function $\Psi\left(N_t\left(\omega^t\right)\right)$, satisfying $\Psi'(\cdot) \geq 0$ and $\Psi''(\cdot) \leq 0$, is the planner's weight of group $N_t\left(\omega^t\right)$. The special case $\Psi\left(N_t\left(\omega^t\right)\right) = N_t\left(\omega^t\right)$ describes a classical total utilitarian planner while $\Psi\left(N_t\left(\omega^t\right)\right) = 1$ describes an average utilitarian. The function $\Psi(N) = N^{\psi_p}$ is the natural counterpart of Barro-Becker's altruism but applied to the planner. The welfare function (5) is a version of NG's (1986) *number-dampened total utility* generalized to include multiple periods and time discounting. Although our main results hold for a standard total utilitarian, we extent our results to the number-dampening case for two reasons: (i) it is natural given that parents in our model exhibit such behavior; and (ii) it turns out to be important for time consistency and uniqueness of the steady state.

The case $\delta = 0$ is defined as

$$\sum_{\omega} \Psi\left(N_0(\omega)\right) U_0(\omega). \tag{6}$$

It refers to a planner who cares only about the initial generation but also future generations to the extent that the initial generation does. In this case social discounting equals private discounting. $\delta > 0$ refers to a planner who is more patient than individuals, as in Farhi and Werning (2007). The following assumption bounds the extent to which the planner cares about future generations.

Assumption 1. $\delta < \beta(\omega)$ for all ω.

The role of Assumption 1 is tractability. The assumption is not particularly restrictive because it still allows for the planner to care about future generations more than parents do. We leave the more complicated case $\delta \geq \beta(\omega)$ for the Appendix. There we show that Malthusian stagnation still holds.

The standard reasoning for considering number dampening, or alternative social criteria such as the "critical-level utilitarianism" of Blackorby and Donalson (1984), is to avoid the Repugnant Conclusion. An allocation is "repugnant" when it entails maximum population and minimum utility, or immiseration. The Repugnant Conclusion is avoided in our environment, as we show below, because parental rights are explicitly considered and children are costly to raise (Hammond 1988).

We can now define the planner's problem.

Definition 1. Given an initial distribution of population $\{N_0(\omega)\}_{\omega\in\Omega}$, the planner chooses sequences $\left\{U_t\left(\omega^t\right), c_t\left(\omega^t\right), n_t\left(\omega^t\right), N_{t+1}\left(\omega^{t+1}\right), L_t\right\}_{\omega^t\in\Omega^{t+1}, t\geq 0}$ to maximize social welfare (5) subject to sequences of resource constraints (3), labor supply (2), laws of motions for population (1) and individual welfare (4).

We assume throughout that the planner's problem is well defined and refer to its solution as the optimal or efficient allocation. We also follow the standard practice in population ethics of normalizing the utility level $\underline{U}$ to zero (e.g., Blackorby *et al.* 2005, pg. 25). This means that, in the mind of parents and the planner, a life is worth living if and only if $U_t\left(\omega^t\right) \geq 0$. Since U_t can be written as a discounted sum of utility flows, then the normalization requires $u(c) \geq 0$.

For clarity, it is convenient to write the Lagrangian corresponding to the planner's problem:

$$\begin{aligned}
L &= \sum_{t=0} \delta^t \sum_{\omega^t} \Psi\left(N_t\left(\omega^t\right)\right) U_t\left(\omega^t\right) \\
&+ \sum_{t=0}^{\infty} \sum_{\omega^t} \theta_t\left(\omega^t\right) N_t\left(\omega^t\right) \left[u(c_t\left(\omega^t\right)) + \Phi\left(n_t\left(\omega^t\right), \omega\right) E_t U_{t+1}\left(\omega^{t+1}\right) - U_t\left(\omega^t\right)\right] \\
&+ \sum_{t=0}^{\infty} \sum_{\omega^{t+1}} \gamma_{t+1}\left(\omega^{t+1}\right) \left[N_{t+1}\left(\omega^{t+1}\right) - n_t\left(\omega^t\right) \pi\left(\omega_{t+1}, \omega_t\right) N_t\left(\omega^t\right)\right] \\
&+ \sum_{t=0}^{\infty} \mu_t \left[F(\bar{K}, L_t; A) - \sum_{\omega^t} N_t\left(\omega^t\right) \left[c_t\left(\omega^t\right) + \eta\left(\omega_t\right) n_t\left(\omega^t\right)\right]\right] \\
&+ \sum_{t=0}^{\infty} \kappa_t \left[\sum_{\omega^t} N_t\left(\omega^t\right) l\left(\omega_t\right) \left[1 - \lambda_t\left(\omega_t\right) n_t\left(\omega^t\right)\right] - L_t\right],
\end{aligned}$$

where $\left\{\theta_t\left(\omega^t\right), \gamma_{t+1}\left(\omega^{t+1}\right), \mu_t, \kappa_t\right\}_{\omega^t \in \Omega^{t+1}, t \geq 0}$ are non-negative multipliers. We assume parameters values are such that solutions are interior.[3] The first restriction of the problem resembles a promise keeping constraint while the remaining restrictions are resource constraints. The first order conditions with respect to $\{U_0\left(\omega_0\right), U_{t+1}\left(\omega^{t+1}\right), N_{t+1}\left(\omega^{t+1}\right), n_t\left(\omega^t\right), c_t\left(\omega^t\right), L_t\}_{\omega^t \in \Omega^{t+1}, t \geq 0}$ are:[4]

$$\theta_0\left(\omega_0\right) N_0\left(\omega_0\right) = \Psi\left(N_0\left(\omega_0\right)\right), \tag{7}$$

$$\begin{aligned}
&\theta_{t+1}\left(\omega^{t+1}\right) N_{t+1}\left(\omega^{t+1}\right) \\
&= \delta^{t+1} \Psi\left(N_{t+1}\left(\omega^{t+1}\right)\right) + \theta_t\left(\omega^t\right) N_t\left(\omega^t\right) \Phi\left(n_t\left(\omega^t\right), \omega\right) \pi\left(\omega_{t+1}, \omega_t\right),
\end{aligned} \tag{8}$$

3. For example, the Barro-Becker model possesses an interior solution under certain parameter restrictions.

4. To avoid cumbersome notation, we do not introduce new notation to identify optimal allocations. Allocations from now on should be regarded as optimal.

$$
\begin{aligned}
& \delta^{t+1}\Psi'\left(N_{t+1}\left(\omega^{t+1}\right)\right)U_{t+1}\left(\omega^{t+1}\right) \\
& +\kappa_{t+1}l\left(\omega_{t+1}\right)\left[1-\lambda_{t+1}\left(\omega_{t+1}\right)n_{t+1}\left(\omega^{t+1}\right)\right]+\gamma_{t+1}\left(\omega^{t+1}\right) \\
= \;& \mu_{t+1}\left[c_{t+1}\left(\omega^{t+1}\right)+\eta\left(\omega_{t+1}\right)n_{t+1}\left(\omega^{t+1}\right)\right] \\
& +n_{t+1}\left(\omega^{t+1}\right)\sum_{\omega^{t+2}|\omega^{t+1}}\gamma_{t+2}\left(\omega^{t+2}\right)\pi\left(\omega^{t+2}|\omega^{t+1}\right),
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
& \theta_t\left(\omega^t\right)\Phi_n\left(n_t\left(\omega^t\right),\omega\right)E_tU_{t+1}\left(\omega^{t+1}\right) \\
& =\mu_t\eta\left(\omega_t\right)+\kappa_tl\left(\omega_t\right)\lambda_t\left(\omega_t\right)+\sum_{\omega^{t+1}|\omega^t}\gamma_{t+1}\left(\omega^{t+1}\right)\pi\left(\omega^{t+1}|\omega^t\right),
\end{aligned} \tag{10}
$$

$$
\theta_t\left(\omega^t\right)u'(c_t\left(\omega^t\right))=\mu_t, \tag{11}
$$

$$
\mu_tF_{L,t}=\kappa_t. \tag{12}
$$

This system of equations together with (1), (2), (3), (4) and proper transversality conditions fully describe interior efficient allocations. Equation (7) states that the initial social value of providing utility to a particular group, $\theta_0\left(\omega_0\right)N_0\left(\omega_0\right)$, depends on the exogenous Pareto weight of group $N_0\left(\omega_0\right)$. Equation (8) then allows to trace the dynamics of this value. The right hand side of the equation is the marginal benefit of promising utility $U_{t+1}\left(\omega^{t+1}\right)$ while the left hand side is its marginal cost. Notice that if the Markov chain π is irreducible, the planner eventually assigns social value, and therefore provides utility, to individuals of all types just because all dynasties eventually have descendants of every type. This would not be the case if π is reducible.

Equation (9) equates marginal benefits to marginal costs of population. To better understand this expression, assume for a moment that population is not constrained by (1), for example, because the planner have access to an infinite pool of immigrants. In that case $\gamma_{t+1}\left(\omega^{t+1}\right) = 0$ for all t and ω^{t+1}. In other words, γ is the value of an immigrant. The marginal benefit of an additional individual of type ω^{t+1} includes her direct effect in social welfare, $\delta^{t+1}\Psi'\left(N_{t+1}\left(\omega^{t+1}\right)\right)U_{t+1}$ plus her effect in the labor supply, $\kappa_{t+1}l\left(\omega_{t+1}\right) \times\left[1-\lambda_{t+1}\left(\omega_{t+1}\right)n_{t+1}\left(\omega^{t+1}\right)\right]$, while the marginal cost includes the cost of providing consumption and fertility to the individual, $\mu_{t+1}[c_{t+1}\left(\omega^{t+1}\right)+\eta\left(\omega_{t+1}\right)n_{t+1}\left(\omega^{t+1}\right)]$. Adding restriction (1) makes the individual more valuable in the amount $\gamma_{t+1}\left(\omega^{t+1}\right)$ because it relaxes the population constraint at $t+1$, but also increases marginal costs because the planner needs to endow the individual with children at $t+2$.

The condition for optimal fertility is Equation (10). The marginal benefit of a child for an altruistic parent with history ω^t is the expected utility of the child, E_tU_{t+1}, times the weight that the parent attaches to the child, $\Phi_n\left(n_t\left(\omega^t\right),\omega\right)$. The marginal benefit for the planner is this amount times $\theta_t\left(\omega^t\right)$. The corresponding marginal cost of the child for the planner includes good costs, $\mu_t\eta\left(\omega_t\right)$, time costs, $\kappa_tl\left(\omega_t\right)\lambda\left(\omega_t\right)$, and the shadow costs of the descendants, $\sum_{\omega^{t+1}|\omega^t}\gamma_{t+1}\left(\omega^{t+1}\right)\pi\left(\omega^{t+1}|\omega^t\right)$.

To characterize the solution of this system of equations, we focus primarily on the steady state and proceed in three steps. First we characterize the deterministic case with only one type (Section 3), then the case with multiple but deterministic types (Section 4) and finally the stationary solution with stochastic types. We show that Malthusian stagnation generally arises when technological progress is of the land augmenting type meaning that steady state optimal consumption allocations and fertility choices are independent of $\bar{K}$ and A. We also characterize the optimal composition of population, the potential dependence of the steady state land-labor ratio on initial conditions, and fertility differentials among types.

III. Deterministic case with one type

This section considers the representative agent case with only one type. We show that stagnation is efficient if technological progress is of the land augmenting type or the production function is Cobb-Douglas. Let $n(\omega) = n$, $\lambda(\omega) = \lambda$, and $l(\omega) = 1$ for simplicity. In this case the resource constraint (3) reduces to:

$$F\left(\frac{\bar{K}}{N_t}, 1 - \lambda n_t; A\right) = c_t + \eta n_t. \tag{13}$$

Moreover, using (13), (11), and (12), equations (7) to (10) simplify to:

$$\Psi(N_0) = \theta_0 N_0, \tag{14}$$

$$\delta^{t+1}\Psi(N_{t+1}) + \theta_t N_t \Phi(n_t) = \theta_{t+1} N_{t+1}, \tag{15}$$

$$\delta^{t+1}\Psi'(N_{t+1}) U_{t+1} + \gamma_{t+1} = \mu_{t+1} F_{K,t+1} \frac{\bar{K}}{N_{t+1}} + n_{t+1}\gamma_{t+2}, \text{ and} \tag{16}$$

$$\Phi'(n_t) \frac{U_{t+1}}{u'(c_t)} = \eta + F_{L,t}\lambda + \frac{\gamma_{t+1}}{\gamma_t}\frac{\gamma_t}{\mu_t}. \tag{17}$$

Equation (16) is obtained from (9) after using (12), (13) and the constant returns to scale assumption. Equation (17) is obtained from (10), (11), and (12).

III.A. Steady state

Consider a steady state situation in which N and c are constant while the present value Lagrange multipliers grow at a constant rate, possibly zero. In that case $n = 1$ and Equation (15) can be written as $\theta_{t+1}/\theta_t = \beta + \frac{\delta^{t+1}\Psi(N)}{N\theta_t} \geq \beta$. Under Assumption 1, the ratio $\frac{\delta^{t+1}\Psi(N)}{N\theta_t}$ goes to zero in the limit. Therefore $\theta_{t+1}/\theta_t = \beta$ in a steady state. Moreover, Equation (11) becomes $\theta_t u'(c) = \mu_t$ so that $\theta_{t+1}/\theta_t = \mu_{t+1}/\mu_t = \beta$. Equation (16) can be written as:

$$\delta^{t+1}\Psi'(N)\,U/\gamma_{t+1} + 1 = \frac{\mu_{t+1}}{\gamma_{t+1}} F_K \frac{\bar{K}}{N} + \frac{\gamma_{t+2}}{\gamma_{t+1}}.$$

It is easy to see that $\frac{\gamma_{t+1}}{\gamma_t} = \delta$ or $\frac{\gamma_{t+1}}{\gamma_t} = \beta$ are the only two potential steady state growth rates satisfying this equation. However, if $\frac{\gamma_{t+1}}{\gamma_t} = \delta$ then the first term would be constant in a steady state but the term $\frac{\mu_{t+1}}{\gamma_{t+1}}$ would be exploding because μ grows at the rate $\beta > \delta$. Thus $\frac{\gamma_{t+1}}{\gamma_t} = \delta$ is not a steady state solution. Hence, the only solution is $\frac{\gamma_{t+1}}{\gamma_t} = \beta$ so that the first term in the equation above becomes zero in steady state simplifying the expression to[5]

(18) $$\frac{\gamma_t}{\mu_t} = \frac{F_K \frac{\bar{K}}{N}}{1-\beta} = \frac{F\left(\frac{\bar{K}}{N}, 1-\lambda; A\right) - F_L\,(1-\lambda)}{1-\beta}.$$

Equation (18) states that the steady state value of an immigrant in units of goods, $\frac{\gamma}{\mu}$, is the present value of "land rents". On the other hand, equation (4) simplifies in steady state to $U = \frac{u(c)}{1-\beta}$. Therefore, Equation (17) can be written, using the results obtained for U, $\frac{\gamma_t}{\mu_t}$, $\frac{\gamma_{t+1}}{\gamma_t}$ and the definitions of ψ and ξ as:

(19) $$\beta\psi(1)\,c/\xi(c) = (1-\beta)\,\eta + (\lambda-\beta)\,F_L + \beta F\left(\frac{\bar{K}}{N}, 1-\lambda; A\right).$$

This equation together with the resource constraint

(20) $$F\left(\frac{\bar{K}}{N}, 1-\lambda; A\right) = c + \eta$$

5. If $\delta \geq \beta$ then the steady state system is more complex because the terms $\frac{\delta^{t+1}\Psi(N)}{N\theta_t}$ and $\frac{\delta^{t+1}\Psi'(N)U}{\gamma_{t+1}}$ do not vanish. We study this case in the Appendix.

form a system of two equations in two unknowns: c and N. They can be used to write consumption as:

$$c = \frac{\xi(c)/\beta}{\psi(1) - \xi(c)} \left[\eta + F_L (\lambda - \beta)\right]. \tag{21}$$

This expression is analogous to the one obtained by Barro and Becker (1989) who show that consumption is proportional to the net costs of raising a child: $\eta + \lambda F_L - \beta F_L$. For example, if $\lambda = \beta$ and $\xi(c) = \xi$ then consumption is proportional to η. The parametric restriction $\psi(1) > \xi(c)$ is needed for consumption to be positive. An implication is that immiseration and the Repugnant conclusion, $c = 0$ and $N = \infty$, is not optimal unless the net cost of children is zero. Equation (21) is not a final solution for consumption because F_L still needs to be solved for. For this purpose, write (19), using the constant returns to scale assumption for F and the definition of α, the land share of output, to obtain:

$$c = \frac{\xi(c)/\beta}{\psi(1)} \left[\eta(1-\beta) + \left(\frac{(\lambda-\beta)(1-\alpha)}{1-\lambda} + \beta\right) F\left(\frac{\bar{K}}{N}, 1-\lambda, A\right)\right].$$

Finally, using (20) and collecting terms, consumption can be solved as

$$c = \xi(c)\eta \frac{1 - \beta - \alpha(\lambda - \beta)}{(\psi(1) - \xi(c))\beta(1-\lambda) - \xi(c)(1-\alpha)(\lambda-\beta)}. \tag{22}$$

This expression describes efficient steady state consumption as long as the expression is non-negative.[6] A sufficient, but not necessary, condition for this to be the case is $\lambda \leq \beta$. We can now state our first main result: Malthusian stagnation is efficient under general conditions.

PROPOSITION 2. Suppose the steady state is interior. Then, in a steady state efficient consumption is independent of the amount of land while efficient

6. If the denominator is negative the efficient allocation is not interior.

population increases proportionally with the amount of land. Furthermore, if technological progress is land augmenting then efficient consumption is independent of the level of technology while efficient population increases proportionally with the level of technology.

Intuitively, land discoveries do not affect the land share because population increases to exactly match the extra land leaving consumption and land-labor ratio, the solutions to equations (20) and (22) unchanged. Furthermore, land augmenting technological progress acts analogously to land discoveries. Notice that the proposition hold for general functions F, u and Φ. A particular case is the Barro-Becker formulation.

IV. Deterministic case with multiple types

Consider now the case of multiple deterministic types. Specifically, suppose $\omega^t = [\omega, \omega, \omega ...]$ or just $\omega^t = \omega$ for short. We assume in this section $\delta = 0$. This restriction is without much loss in generality since similar steady state results would be obtained as long as $\delta < \beta(\omega)$, as shown in the previous section. For tractability, we also restrict altruism to be of the Barro-Becker form, $\Phi(n, \omega) = \beta(\omega) n^{\psi}$ but still allow more general formulations for u and F.

We show the following results in this section. First, if $\beta(\omega)$ is different for different types then their population sizes grow at different rates and in steady state only the most patient groups, the ones with highest $\beta(\omega)$, survive. This result implies that efficient social classes cannot be sustained by persistent differences in rates of time preference. Lucas (2002) is able to generate social classes using such a mechanism in a competitive equilibrium with savings constraints which suggests that social classes are not efficient, in the first best sense, in his model.

As an alternative to Lucas (2002), we are able to generate multiple social classes using a more standard mechanism based on heterogeneity in labor skills, $l(\omega)$, and the cost of raising children, $\eta(\omega)$ and $\lambda(\omega)$. This is the second main result of the section. Efficiency requires to provide more consumption to individuals with higher costs of raising children. Consumption also increases with labor ability, $l(\omega)$, but only if $\lambda(\omega) > \beta$, that is, only if the time costs of raising children are sufficiently high. Otherwise, the efficient allocation involves the high skilled having lower consumption.[7]

Third, we show that relative population sizes of various types are inversely related to their relative consumption. Therefore, the population of the poor is larger than the population of the middle class and so on. The planner thus faces a quantity-quality trade-off: she can deliver certain level of welfare by allocating children and/or consumption. If children are particularly costly to raise for a certain group, then the planner optimally delivers welfare more through consumption than through children and vice versa.

Fourth, in the deterministic steady state of this section all types have one child and therefore steady state welfare differences among types only arise from differences in consumption. As a result, types with lower consumption are worse-off than types with higher consumption. All benefits from a larger population accrue only to early members of the dynasty at the expense of later members.

IV.A. Dynamics

The following lemma characterizes the evolution of efficient population sizes of different types over time.

7. This result could rationalize, for example, why high skilled women may end up having more children and low consumption compared with an equally skilled man. Extending the model to introduce gender differences is a promising agenda for future research.

LEMMA 3. Let $(\omega, \omega') \in \Omega$. Efficient population sizes satisfy:

$$\frac{N_t(\omega)}{N_t(\omega')} = \left(\frac{N_0(\omega)}{N_0(\omega')}\right)^{-\frac{\psi}{1-\psi}} \left[\frac{\Psi(N_0(\omega))}{\Psi(N_0(\omega'))} \frac{\beta(\omega)^t u'(c_t(\omega))}{\beta(\omega')^t u'(c_t(\omega'))}\right]^{\frac{1}{1-\psi}}. \tag{23}$$

We characterize next the steady state.

IV.B. *Steady state*

IV.B..1 Distribution of population

Consider now a steady state in which consumption, population shares and population are constant. This requires $n(\omega) = 1$ for all types. In that case, equation (23) simplifies to:

$$\frac{N(\omega)}{N(\omega')} = \left(\frac{\beta(\omega)}{\beta(\omega')}\right)^{\frac{t}{1-\psi}} \left(\frac{N_0(\omega)}{N_0(\omega')}\right)^{-\frac{\psi}{1-\psi}} \left[\frac{\Psi(N_0(\omega))}{\Psi(N_0(\omega'))} \frac{u'(c(\omega))}{u'(c(\omega'))}\right]^{\frac{1}{1-\psi}}. \tag{24}$$

We can now state our second main result which is apparent from this equation.

PROPOSITION 4. In an interior steady state: (i) Only the most patient types, the ones with the highest $\beta(\omega)$, have positive mass; (ii) The distribution of population depends on the initial distribution unless $\Psi(N_0) = N^{\psi}$; In particular, it depends on the initial distribution in the classical utilitarian case; (iii) The relative population size of a particular type is inversely related to its per-capita consumption.

The first part of the proposition states that impatient types eventually disappear from the economy. Children are like an investment for altruistic parents as they deliver a stream of future utility flows. Impatient individuals discount future streams more heavily and therefore value children less than patient individuals do. As a result, it is efficient for the planner to provide more consumption to current individuals in exchange for fewer future family members.[8]

The second part of Proposition 4 states that the steady state distribution of population depends on initial conditions, a result that is analogous to the dependence of the steady state wealth distribution on initial conditions in the neoclassical growth model (Chatterjee 1994). However, as we see below, this dependence has more profound implications in Malthusian economics because the steady state aggregate land-labor ratio and steady state population depends on initial conditions, and Pareto weights, as well. This is in contrast to the neoclassical growth model where the golden rule level of capital is independent of initial conditions and Pareto weights. Efficiency and distribution are interdependent in Malthusian economies unless Pareto weights are of the form $\Psi(N_0) = N^{\psi}$, that is, Pareto weights resemble parental weights.

The third part of Proposition 4 shows a fundamental prediction of endogenous population models: an inverse relationship between population size and per-capita consumption. The lower the consumption of a type the larger its share of the total population. The reason is that the planner needs to deliver welfare by providing consumption and children to parents. Whenever the planner chooses to use one channel then it downplays the other.

We still need to solve for consumption to fully derive the consequences of this inverse relationship. For the rest of this section it is convenient to assume a specific functional form for $\Psi(N)$, $\Psi(N) = N^{\psi_p}$, and restrict attention to the set of most patient types, $\Omega_p \subseteq \Omega$. That is, $\beta(\omega) = \beta$ for all $\omega \in \Omega_p$ and $\beta \geq \beta(\omega)$ for all $\omega \in \Omega$. Equation (24) thus simplifies to:

$$\frac{N(\omega)}{N(\omega')} = \left(\frac{N_0(\omega)}{N_0(\omega')}\right)^{\frac{\psi_p - \psi}{1-\psi}} \left(\frac{u'(c(\omega))}{u'(c(\omega'))}\right)^{\frac{1}{1-\psi}}, \ \omega \in \Omega_p. \tag{25}$$

8. This result also helps qualify a common view that the poor are inherently more impatient, less willing to save, and that their large families somehow reflects their impatience. According to our model, if the poor were really impatient, they would have fewer children and their type would eventually disappear from the population.

Thus, the long term composition of the population depends on the initial distribution unless $\psi_p = \psi$. Moreover, the initial distribution of population tend to persist if $\psi_p > \psi$. Classical utilitarianism is represented by $\psi_p = 1$. In this case, the steady state distribution never resembles the initial distribution unless consumptions are equal across types which is not the case in general, as we show below.[9]

The following lemma characterizes the steady state distribution of population in terms of consumptions.

LEMMA 5. Let $p(\omega) \equiv \frac{N(\omega)}{N}$. Then

$$p(\omega) = \frac{N_0(\omega)^{\frac{\psi_p-\psi}{1-\psi}} u'(c(\omega))^{1/(1-\psi)}}{\sum_{\omega'} N_0(\omega')^{\frac{\psi_p-\psi}{1-\psi}} u'(c(\omega'))^{1/(1-\psi)}}. \tag{26}$$

The lemma is important because it provides a simple description of the steady state distribution of population in terms of the given initial distribution and steady state consumptions.

IV.B..2 Consumption

One can show, similarly to the first part of Proposition 4, that only the most patient types have positive consumption in steady state. According to (11), for consumption to be constant $\frac{\theta_{t+1}(\omega)}{\theta_t(\omega)} = \frac{\mu_{t+1}}{\mu_t}$ is required. Otherwise, $\frac{\theta_{t+1}(\omega)}{\theta_t(\omega)} < \frac{\mu_{t+1}}{\mu_t}$ refers to a type for which consumption falls, and vice versa. Therefore, only the types with the highest ratio $\frac{\theta_{t+1}(\omega)}{\theta_t(\omega)}$ have positive steady state consumption. Moreover, according to (8), $\frac{\theta_{t+1}(\omega)}{\theta_t(\omega)} = \Phi(1,\omega) = \beta(\omega)$ at steady state. Therefore, $\frac{\theta_{t+1}(\omega)}{\theta_t(\omega)}$ is the highest for all $\omega \in \Omega_p$.

The following Lemma provides the solution for consumptions in terms of

9. In the utilitarian case, efficient allocations are not time consistent because re-optimizing starting with an initial steady state distribution of population results in a different steady state distribution.

population shares and the marginal product of labor.

LEMMA 6. Efficient consumption satisfies:

$$c(\omega) = \frac{\xi(c(\omega))/\beta}{\psi - \xi(c(\omega))}\left[\eta(\omega) + (\lambda(\omega) - \beta) F_L l(\omega)\right] \text{ for } \omega \in \Omega_p. \tag{27}$$

Equation (27), analogous to (21), shows that consumption is proportional to the net financial cost of a child. In particular consumption is larger for types with higher cost of raising children, either higher $\eta(\omega)$ and/or higher $\lambda(\omega)$. The relationship between skills, $l(\omega)$, and consumption is slightly more complicated. If $\lambda(\omega) > \beta$ then efficient consumption is higher for high skilled individuals. But if $\lambda(\omega) < \beta$, then efficient consumption is actually lower for the high skilled.

We can now state the third main result of the paper which follows from (24) and (27).

PROPOSITION 7. The steady state efficient allocation exhibits inequality of consumptions and populations. Types with low consumption have larger population.

Proposition 7 is important for at least three reasons. First, as is discussed by Lucas (2002), obtaining an efficient allocation with heterogeneous social classes in Malthusian economies is not trivial yet important. Lucas's solution, which relies on differences in time discounting, generates inefficient social classes in presence of binding constraints. Different discount factors would still lead to only one social group surviving in steady state in an efficient allocation. Second, the efficient allocation can rationalize a distribution of social classes in which the poor are a larger fraction of the population. Third, the proposition also states that, in a world where the planner can choose which types survive and which types disappear, it is not optimal to end a lineage just because it is of lower skill or poorer. This is in contrast to a literature that argues in favor of limiting the fertility of the poor (e.g., Chu and Koo, 1990). Only impatient types disappear from an efficient allocation.

It is possible to find final solution for consumptions and relative population sizes without knowing the marginal product of labor in the following special Barro-Becker case.

Example 8. Suppose $u(c) = c^{\xi}/\xi$ with $\xi \in (0,1)$, $\Phi(n) = \beta n^{\psi}$, $\psi \in (\xi, 1)$, $\Psi(N) = N^{\psi}$ and $\lambda(\omega) = \beta$. Then

$$c(\omega) = \frac{\xi}{\psi - \xi} \frac{\eta(\omega)}{\beta} \text{ and } \frac{N(\omega)}{N(\omega')} = \left(\frac{\eta(\omega')}{\eta(\omega)} \right)^{\frac{1-\xi}{1-\psi}}.$$

In this example, consumption is proportional to the goods cost of raising a child, $\eta(\omega)$, while the exponent $\frac{1-\xi}{1-\psi} \in (1, \infty)$ controls the extent to which consumption inequality translates into population inequality. Since the restriction $\psi > \xi$ is needed for an interior solution, the exponent is larger than 1. Therefore, population inequality is larger than consumption inequality. For example, if consumption of the rich is 5 times that of the poor, $\frac{\eta(\omega')}{\eta(\omega)} = 5$, and $\frac{1-\xi}{1-\psi} = 2$ then the population of the poor is 25 times that of the rich. The planner in this example is more willing to accept a large share of poor individuals when intergenerational substitution of consumption is particularly low (ξ is low) and/or parental altruism does not decrease sharply with family size (ψ is high).

IV.B..3 Average output

A full solution requires to find the marginal product of labor which itself requires a solution for the land-labor ratio. For this purpose, rewrite the steady-state resource constraint as

$$LF\left(\bar{K}/L, 1\right) = N \sum_{\omega} p(\omega) \left[c(\omega) + \eta(\omega)\right].$$

Furthermore, total labor supply relative to population is expressed, at steady state, by

$$\frac{L}{N} = \sum_{\omega} p(\omega)\, l(\omega)\, [1 - \lambda(\omega)]\,. \tag{28}$$

Dividing these two equations yields

$$\frac{F}{L} = F\left(\frac{\bar{K}}{L}, 1\right) = \frac{\sum_{\omega} p(\omega)\, [c(\omega) + \eta(\omega)]}{\sum_{\omega} p(\omega)\, l(\omega)\, [1 - \lambda(\omega)]}. \tag{29}$$

The system of three set of equations, (26), (27) and (29), can then be used to solve for the following unknowns: $p(\omega)\,, c(\omega)$ and L.

IV.B..4 Stagnation

Combining (29) and (27), and using the Cobb-Douglas production function one obtains:

(30)

$$c(\omega) = \frac{\xi(c(\omega))/\beta}{\psi - \xi(c(\omega))}\left[\eta(\omega) + (\lambda(\omega) - \beta)(1-\alpha)\frac{\sum_{\omega} p(\omega)\, [c(\omega) + \eta(\omega)]}{\sum_{\omega} p(\omega)\, l(\omega)\, [1 - \lambda(\omega)]} l(\omega)\right]$$

Equations (26) and (30) can be use to solve for $c(\omega)$ and $p(\omega)$. Notice that as long as α is independent of $\bar{K}$ and A, so are $c(\omega)$ and $p(\omega)$. Once these two variables are solved for then (29) can be used to solve for L and (28) for N. The following Proposition summarizes these results. The proof is similar to that of Proposition 2 and hence omitted.

Proposition 9. Suppose $\delta = 0$, $\Phi(n, \omega) = \beta n^{\psi}$ and the steady state is interior. Then: (i) in steady state optimal consumption is independent of the amount of land and optimal population is proportional to the amount of land; (ii) if technological progress is land augmenting then optimal consumption is independent of the level of technology and population increases proportionally with the level of technology; and (iii) optimal allocations depend on the initial distribution of population unless $\Psi(N_0) = N_0^{\psi}$.

V. Stochastic case

The deterministic version of the model considered so far counterfactually predicts equal fertility among different social groups. Malthus, however observed that fertility rates were higher among the poor. We now show that a version of the model with stochastic types can generate differential fertility. For tractability we once again assume $\delta = 0$ and use the Barro-Becker functional forms: $\Phi(n, \omega) = \beta n^{\psi}$ and $u(c) = c^{\xi}/\xi$. Equation (8) can be simplified, using equation (11) and the law of motion for population, equation (1), as:

$$\frac{\Phi\left(n_t\left(\omega^t\right)\right)}{n_t\left(\omega^t\right)} = \frac{\mu_{t+1}}{\mu_t}\frac{u'(c_t\left(\omega^t\right))}{u'(c_{t+1}\left(\omega^{t+1}\right))}. \tag{31}$$

An implication of this equation is that all children within a family have the same consumption:

$$c_{t+1}\left(\omega^t;\omega_{t+1}\right) = c_{t+1}\left(\omega^t\right) \text{ for all } \omega_{t+1} \in \Omega. \tag{32}$$

The following Lemma shows that optimal consumption allocations are history independent and satisfy a formulation similar to that of Equations (21) or (27). In particular, the consumption of a child is proportional to the expected net costs of raising that child.

LEMMA 10. Optimal consumption satisfies:

(33)
$$c_{t+1}(\omega_t) = \frac{\xi}{\psi - \xi}\left[\frac{\mu_t}{\mu_{t+1}}\eta(\omega_t) + \frac{\mu_t}{\mu_{t+1}}F_{L,t}l(\omega_t)\lambda(\omega_t) - F_{L,t+1}E_t(l(\omega_{t+1}))\right].$$

Notice that according to the lemma $c_{t+1}\left(\omega^{t+1}\right) = c_{t+1}(\omega_t)$ so that efficient consumption is not history dependent. Similarly, substituting (33) into (31), it follows that $n_t(\omega^t) = n_t(\omega_{t-1}, \omega_t)$ so that the number of children only depends on the types of the parent and grandparent.

V.A. Steady state

Consider now stationary steady state allocations in which $n_t(\omega_{t-1}, \omega_t) = n(\omega_{t-1}, \omega_t)$, $c_t(\omega_{t-1}) = c(\omega_{t-1})$, $N_t(\omega^t) = N(\omega_{t-1}, \omega_t)$ and $N_t = N_{t+1}$. Let $R \equiv \frac{\mu_t}{\mu_{t+1}}$ be the planner's shadow gross return and with a little bit abuse of notation let $p(\omega_{t-1}, \omega_t) \equiv \frac{N(\omega_{t-1}, \omega_t)}{N_t}$ be the population share with recent history (ω_t, ω_{t-1}). The following Lemma summarizes the system of equations and unknowns describing stationary steady state.

LEMMA 11. Steady state allocations, $c(\omega)$, $n(\omega_{-1}, \omega)$, $p(\omega_{-1}, \omega)$, R, L and N are solved from the following systems of equations:

$$c(\omega) = \frac{\xi R}{\psi - \xi}\left[\eta(\omega) + F_L l(\omega)\lambda(\omega) - F_L E\left[l(\omega_{+1})|\omega\right]/R\right], \tag{34}$$

$$n(\omega_{-1}, \omega) = \left[\beta R \frac{u'(c(\omega))}{u'(c(\omega_{-1}))}\right]^{\frac{1}{1-\psi}}, \tag{35}$$

$$p(\omega, \omega_{+1}) = \sum_{\omega_{-1}} n(\omega_{-1}, \omega)\pi(\omega_{+1}, \omega)p(\omega_{-1}, \omega), \tag{36}$$

(37) $$\sum_{\omega_{-1}} \sum_{\omega} p(\omega_{-1}, \omega)\, n(\omega_{-1}, \omega_t) = 1,$$

(38) $$F\left(\frac{\bar{K}}{L}, 1; A\right) = \frac{N}{L} \sum_{\omega_{-1}} \sum_{\omega} p(\omega_{-1}, \omega) \left[c(\omega_{-1}) + \eta(\omega)\, n(\omega_{-1}, \omega)\right],$$

(39) $$\frac{L}{N} = \sum_{\omega_{-1}} \sum_{\omega} p(\omega_{-1}, \omega)\, l(\omega) \left(1 - \lambda(\omega)\, n(\omega_{-1}, \omega)\right).$$

Equation (34) shows the consumption of an individual whose parent is of type ω. Consumption is positively associated with the parental costs of raising children and parental skills and negatively associated with the expected skills of the child.

Equation (35) shows fertility differentials among different types. Optimal fertility depends on parental and grandparent types. Given grandparent types, parents with low consumption have more children than parents with high consumption. Also, given parent types, consumption rich grandparents have more grandchildren than consumption poor grandparents. Equation (37), which in principle serves to solve R, restricts fertility to be one on average. Equations (38) and (39) are resource constraints of goods and labor.

The next Proposition shows that the stagnation property still holds in the stochastic case.

Proposition 12. Suppose the steady state is interior. Then, steady state optimal consumption is independent of the amount of land and optimal population increases proportionally with the amount of land. Furthermore, if technological progress is land augmenting then optimal consumption is independent of the level of technology and population increases proportionally with the level of technology.

To summarize, in addition to stagnation, the key properties of the stochastic steady state are differential fertility and heterogeneous social groups. Moreover, all types, or social groups, are represented in a steady state even if their initial population is zero as long as π is non-reducible.

VI. Concluding comments

The pre-industrial world was to a large extent Malthusian. As documented by Ashraf and Galor (2011), periods characterized by improvements in technology or in the availability of land eventually lead to a larger but not richer population. This is remarkable given the diversity of political, social, religious, geographical, cultural, and economic environments they considered, some arguably more advanced than others. Why were no particular systems able or willing to control population size to avoid the Malthusian trap?

We show that even in the best case scenario of a Malthusian economy populated by loving rational parents and governed by an all powerful benevolent rational planner, stagnation, inequality, high population of the poor and differential fertility could still naturally arise as an optimal choice. Our findings thus help explain why the Malthusian trap was so pervasive in pre-industrial societies. We also show that is not the irrational animal spirit of human beings, as suggested by Malthus, what ultimately explains the stagnation. Stagnation can be the result of an optimal choice between the quality and quantity of life in the presence of limited natural resources.

Finally, this article proposes and implements a novel approach to study issues of efficiency when the population is endogenous. We solve the planner's problem in a novel and tractable way. We expect this methodology to further facilitate the integration of demographics and macroeconomics.

References

Ashraf, Q., and O. Galor (2011): "Dynamics and Stagnation in the Malthusian Epoch," *American Economic Review*, 101(5), 2003-2041.

Barro, R. J. and G. S. Becker (1989): "Fertility Choice in a Model of Economic Growth," *Econometrica*, 57, 481-501.

Becker, G. S. and R. J. Barro (1988): "A Reformulation of the Economic Theory of Fertility," *The Quarterly Journal of Economics*, 103, 1-25.

Becker, G. S., K. M. Murphy, and R. Tamura (1990): "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy,* 98, S12-S37.

Blackorby, C., W. Bossert and D. Donalson (2005): "Population Issues in Social Choice Theory, Welfare Economics, and Ethics". Cambridge University Press, 23(2), 256-260.

Blackorby, C. and D. Donalson (1984): "Social Criteria for Evaluating Social Change." *Journal of Public Economics,* 25, 13-33.

Chatterjee, S. (1994): "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model," *Journal of Public Economics,* 54, 97-119.

Chu, C. Y. C. and H.-W. Koo (1990): "Intergenerational Income-Group Mobility and Differential Fertility", *American Economic Review*, 80, 1125-1138.

Cordoba, J. C. and M. Ripoll (2011): "A Contribution to the Economic Theory of Fertility," Economics Working Papers 11011, Iowa State University, Department of Economics.

Cordoba, J. C. and M. Ripoll (2011): "The Intergenerational Elasticity of Substitution, Parental Altruism, and Fertility Choice," Economics Working Papers 14016, Iowa State University, Department of Economics.

Dasgupta, P. (2005): "Regarding Optimum Population," *The Journal of Political Philosophy*, 13, 4, 414-442.

De la Croix, D. (2012): "Fertility, Education, Growth and Sustainability," Cambridge University Press.

Doepke, M. (2004): "Accounting for Fertility Decline During the Transition to Growth," *Journal of Economic Growth*, 9(3), 347-383.

Eckstein, Z., Stern, S., and K.I., Wolpin (1988): "Fertility Choices, Land, and the Malthusian Hypothesis", *International Economic Review*, 29 (2), 353-361.

Farhi, E. and I. Werning (2007): "Inequality and Social Discounting," *Journal of Political Economy*, 115 (3), 365-402.

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

Golosov, M., L. Jones and M. Tertilt (2007): "Efficiency with Endogenous Population Growth," *Econometrica*, 75 (4), 1039-1071.

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

Hammond, P. J. (1988): "Consequentialist Demographic Norms and Parenting Rights," *Social Choice Welfare*, 5, 127-145.

Jones, Charles I. (1999): "Was the Industrial Revolution Inevitable? Economic Growth Over the Very Long Run," NBER Working Paper No. 7375.

Lucas, R. (2002): "The Industrial Revolution: Past and Future.: in Lectures on Economic Growth," Harvard University Press, Cambridge, MA.

Nerlove, M., Razin, A. and E., Sadka (1986): "Endogenous Population with Public Goods and Malthusian Fixed Resources: Efficiency or Market Failure," *International Economic Review*, 27 (3), 601-609.

Ng, Y.-K. (1986): "Social Criteria for evaluating population change: an alternative to the Blackkorby-Donalson criterion," *Journal of Public Economics,* 29, 375-381.

Peretto, P. F. and S. Valente (2015): "Growth on a Finite Planet: Resources, Technology and Population in the Long Run," *Journal of Economic Growth*, 20, 315-331.

Appendix

A.1. Proofs of Propositions and Lemmas

Proof of Proposition 2. Equation (22) is the final solution for consumption if the technology is Cobb-Douglas because in that case α is a parameter. This equation shows that c is independent of land K and technology A when F is Cobb-Douglas for general functions u and Φ. Otherwise, α is not a parameter and equations (20) and (22) need to be solved simultaneously for c and N. However, one can show that c is independent of the amount of land. To see this, let $c(K_0)$ and $N(K_0)$ be the steady state solutions when K_0 is the amount of land and let K_1 be a different amount. Consider the solution $c(K_1) = c(K_0)$ and $N(K_1) = \frac{K_1}{K_0/N(K_0)}$. $L_0 = N(K_0)(1-\lambda)$ and $L_1 = N(K_1)(1-\lambda)$. Notice that the proposed solution when $K = K_1$ exhibits the same land labor ratio as the solution for K_0 and therefore it still solves equation (20). Furthermore $\alpha(K_1) = \frac{F_K(K_1,L_1,A)K_1}{F(K_1,L_1,A)} = \frac{F_K(K_1/L_1,1,A)}{F(1,L_1/K_1,A)} = \frac{F_K(K_0/L_0,1,A)}{F(1,L_0/K_0,A)}$ so that the land share is unchanged. As a result, $c(K_0)$ still solves (22) when $K = K_1$. Finally, land augmenting technological progress implies $F\left(\frac{\bar{K}}{N}, 1-\lambda, A\right) = F\left(\frac{A\bar{K}}{N}, 1-\lambda\right)$. Let $c(A_0)$ and $N(A_0)$ be the efficient steady state for the level of technology A_0 and let A_1 be a different level. Consider the solution $c(A_1) = c(A_0)$ and $N(A_1) = \frac{A_1 K}{A_0 K/N(A_0)}$. Notice that the proposed solution when $A = A_1$ exhibits the same effective land labor ratio as the solution for A_0 and therefore it still solves equation (20). Furthermore $\alpha = \frac{F_K(A_1K,L_1)A_1K}{F(A_1K,L_1)} = \frac{F_K(A_1K/L_1,1)}{F(1,L_1/(A_1K))} = \frac{F_K(A_0K/L_0,1)}{F(1,L_0/(A_0K))}$ so that the land share is unchanged. As a result, $c(A_0)$ still solves (22). ■

Proof of Lemma 3. Let $s_t(\omega) \equiv \theta_t(\omega) N_t(\omega)$. Equation (8), given that $\delta = 0$ is assumed, can then be written as:

$$s_1(\omega) = s_0(\omega)\Phi(n_0(\omega)),\ s_2(\omega) = s_0(\omega)\prod_{i=0}^{1}\Phi(n_i(\omega)).$$

More generally, $s_t(\omega) = s_0(\omega)\prod_{i=0}^{t-1}\Phi(n_i(\omega))$. Assuming $\Phi(n) = \beta(\omega)n^{\psi}$, it follows that:

$$\begin{aligned} s_t(\omega) &= s_0(\omega)\,\beta(\omega)^t\left(\prod_{i=0}^{t-1} n_i(\omega)\right)^{\psi} \\ &= \theta_0(\omega)\,N_0(\omega)\,\beta(\omega)^t\left(\frac{N_t(\omega)}{N_0(\omega)}\right)^{\psi} \\ &= \theta_0(\omega)\,(N_0(\omega))^{1-\psi}\,\beta(\omega)^t\,(N_t(\omega))^{\psi}\,. \qquad (40) \end{aligned}$$

Now, (11) can be written as $\mu_t N_t(\omega) = s_t(\omega)\,u'(c_t(\omega))$. Therefore

$$\frac{N_t(\omega)}{N_t(\omega')} = \frac{s_t(\omega)\,u'(c_t(\omega))}{s_t(\omega')\,u'(c_t(\omega'))}.$$

Substituting (40) into this equation gives

$$\frac{N_t(\omega)}{N_t(\omega')} = \frac{\theta_0(\omega)\,(N_0(\omega))^{1-\psi}\,\beta(\omega)^t\,(N_t(\omega))^{\psi}\,u'(c_t(\omega))}{\theta_0(\omega')\,(N_0(\omega'))^{1-\psi}\,\beta(\omega')^t\,(N_t(\omega'))^{\psi}\,u'(c_t(\omega'))}.$$

Finally, use (7) to substitute $\theta_0(\omega)$ and solve for $\frac{N_t(\omega)}{N_t(\omega')}$ to obtain (23). ■

Proof of Lemma 5. According to equation (25), and let $\omega' = \omega_0$,

$$N(\omega) = N(\omega_0)\left(\frac{N_0(\omega)}{N_0(\omega_0)}\right)^{\frac{\psi_p-\psi}{1-\psi}}\frac{u'(c(\omega))^{1/(1-\psi)}}{u'(c(\omega_0))^{1/(1-\psi)}}.$$

Adding $N(\omega)$ over ω,

$$N = \sum_{\omega} N(\omega) = \frac{N(\omega_0)}{N_0(\omega_0)^{\frac{\psi_p-\psi}{1-\psi}}\,u'(c(\omega_0))^{1/(1-\psi)}}\sum_{\omega} N_0(\omega)^{\frac{\psi_p-\psi}{1-\psi}}\,u'(c(\omega))^{1/(1-\psi)}$$

and therefore

$$p(\omega_0) = \frac{N(\omega_0)}{N} = \frac{N_0(\omega_0)^{\frac{\psi_p-\psi}{1-\psi}}\,u'(c(\omega_0))^{1/(1-\psi)}}{\sum_{\omega} N_0(\omega)^{\frac{\psi_p-\psi}{1-\psi}}\,u'(c(\omega))^{1/(1-\psi)}} \text{ for all } \omega_0 \in \Omega_p.$$

■

Proof of Lemma 6. Rewrite (9) using (12) as:

$$1 = \frac{\mu_{t+1}}{\gamma_{t+1}(\omega)}\left[c(\omega) + \eta(\omega) - F_L l(\omega)(1-\lambda(\omega))\right] + \frac{\gamma_{t+2}(\omega)}{\gamma_{t+1}(\omega)}.$$

Since $\frac{\gamma_{t+2}(\omega)}{\gamma_{t+1}(\omega)}$ is constant in steady state then $\frac{\mu_{t+1}}{\gamma_{t+1}(\omega)}$ needs to be constant for this equation to hold, which means that $\frac{\gamma_{t+1}(\omega)}{\gamma_t(\omega)} = \frac{\mu_{t+1}}{\mu_t} = \beta$. The last equality holds by (11) and (15). Therefore, the previous equation can be written as:

$$\frac{\gamma_t(\omega)}{\mu_t} = \frac{1}{1-\beta}\left[c(\omega) + \eta(\omega) - F_L l(\omega)(1-\lambda(\omega))\right]. \tag{41}$$

This expression states that the value of an immigrant in terms of goods, $\frac{\gamma_t(\omega)}{\mu_t}$, is the net present value of the net cost. In steady state $U(c(\omega)) = \frac{u(c(\omega))}{1-\beta}$. Use this result, (11) and (12) to rewrite (10) as:

$$\frac{\Phi'(1)}{1-\beta}\frac{u(c(\omega))}{u'(c(\omega))} = \eta(\omega) + F_L l(\omega)\lambda(\omega) + \beta\frac{\gamma_t(\omega)}{\mu_t}. \tag{42}$$

One can combine equation (41) and (42) to solve for consumption as

$$\begin{aligned} & \frac{\Phi'(1)}{\xi(c(\omega))}c(\omega) \\ = \; & (\eta(\omega) + F_L l(\omega)\lambda(\omega))(1-\Phi(1)) + \beta(c(\omega) + \eta(\omega) - F_L l(\omega)(1-\lambda(\omega))) \\ = \; & \eta(\omega) + F_L l(\omega)\lambda(\omega) + \beta(c(\omega) - F_L l(\omega)) \end{aligned}$$

or

$$c(\omega) = \frac{\xi(c(\omega))}{\Phi'(1) - \beta\xi(c(\omega))}\left[\eta(\omega) + (\lambda(\omega) - \beta)F_L l(\omega)\right].$$

Using $\Phi(n) = \beta n^{\psi}$ provides the result. ■

Proof of Lemma 10. Write Equation (9), the optimality condition for population, as:

(43)

$$\gamma_t(\omega^t) + \kappa_t l(\omega_t) - \mu_t c_t(\omega^t) = n_t(\omega^t)\begin{bmatrix} \mu_t \eta(\omega_t) + \kappa_t l(\omega_t)\lambda_t(\omega_t) \\ +\sum_{\omega^{t+1}|\omega^t}\gamma_{t+1}(\omega^{t+1})\pi(\omega^{t+1}|\omega^t) \end{bmatrix}.$$

Moreover, use (11) to rewrite the first order condition with respect to fertility, Equation (10), as

(44)

$$E_t U_{t+1} = \frac{u'(c_t(\omega^t))}{\mu_t}\frac{\mu_t\eta(\omega_t) + \kappa_t l(\omega_t)\lambda(\omega_t) + \sum_{\omega^{t+1}|\omega^t}\gamma_{t+1}(\omega^{t+1})\pi(\omega^{t+1}|\omega^t)}{\Phi_n(n_t(\omega^t))}.$$

Using (31), Equation (44) can be written as:

(45)

$$E_tU_{t+1} = \frac{u'(c_{t+1}(\omega^{t+1}))}{\mu_{t+1}} \frac{\mu_t\eta(\omega_t) + \kappa_t l(\omega_t)\lambda(\omega_t) + \sum_{\omega^{t+1}|\omega^t}\gamma_{t+1}(\omega^{t+1})\pi(\omega^{t+1}|\omega^t)}{\psi}$$

Plugging (43) into (44),

$$E_tU_{t+1} = \frac{u'(c_t(\omega^t))}{\mu_t} \frac{\gamma_t(\omega^t) + \kappa_t l(\omega_t) - \mu_t c_t(\omega^t)}{n_t(\omega^t)\Phi_n(n_t(\omega^t))}. \tag{46}$$

Plugging (46) into the individual's value function (4):

$$\begin{aligned} U_t(\omega^t) &= u(c_t(\omega^t)) + \Phi(n_t(\omega^t), \omega) \frac{u'(c_t(\omega^t))}{\mu_t} \frac{\gamma_t(\omega^t) + \kappa_t l(\omega_t) - \mu_t c_t(\omega^t)}{n_t(\omega^t)\Phi_n(n_t(\omega^t))} \\ &= u'(c_t(\omega^t))\left[\left(\frac{1}{\xi} - \frac{1}{\psi}\right)c_t(\omega^t) + \frac{1}{\psi}\frac{1}{\mu_t}\left(\gamma_t(\omega^t) + \kappa_t l(\omega_t)\right)\right] \end{aligned}$$

Forwarding this equation one period ahead and taking a conditional expected value, E_t:

$$\begin{aligned} & E_tU_{t+1} \quad (47) \\ = \quad & u'(c_{t+1}(\omega^{t+1}))\left[c_{t+1}(\omega^{t+1})\left(\frac{1}{\xi} - \frac{1}{\psi}\right) + \left(\frac{1}{\psi}\frac{1}{\mu_{t+1}}\left[E_t\gamma_{t+1}(\omega^{t+1}) + \kappa_{t+1}E_t l(\omega_{t+1})\right]\right)\right] \end{aligned}$$

As shown in equation (32) consumption of every individual depends only on the ability of his/her parent, while the aggregate terms μ_{t+1} and κ_{t+1} are deterministic since there is no aggregate risk. Finally, equating (45) and (47), using (12) and simplifying one obtains (33). ■

Proof of Proposition 11. At steady state, (33) becomes (34), (35) can be obtained using Equation (31) and the specified functional forms, the law of motion of population (1) becomes (36), total population is constant and therefore average fertility is equal to 1 as stated by (37). Equations (38) and (39) are steady state versions of (2) and (3). ■

Proof of Proposition 12. The proof is similar to that of Proposition 2 and uses guess and verify. A longer direct proof is also possible. Consider the solution for an initial amount of land, say K_0. Let $\frac{K_0}{N_0}$ and L_0 be the steady state solution of land-population ratio and labor for K_0. Then consider a different amount of land, say K_1. Guess that the solution for the new steady state is identical to the initial solution except for two changes: $N_1 = K_1/\frac{K_0}{N_0}$ so that the land-labor ratio is unchanged, and $L_1 = N_1\frac{L_0}{N_0}$ so that the labor-population ratio is unchanged. Notice that under the proposed solution the marginal product of labor is unchanged too. One can then use Lemma 11 to verify that under the proposed guess, the solutions for consumption, fertility, the distribution of population and R that solve for K_0 also solve for K_1. Similarly for land-augmenting technological progress, guess that population responds to keep $\frac{AK}{L}$ unchanged for different levels of A, while labor responds to keep the ratio $\frac{L}{N}$ unchanged, and nothing else changes. One verifies that the proposed solution satisfies all equations in Lemma 11. ■

A.2. Deterministic case with one type.

A.2.1.Case $\delta \geq \beta$.

LEMMA 13. Assume $\eta > 0$ and the production function $F(\cdot,\cdot)$ satisfies Inada condition and

$$\lim_{L\to\infty} F_{KL}\left(\bar{K}, L\right) = 0.$$

(i) If $\delta > \beta$, a steady state satisfies the following equations:

$$\begin{aligned}(48)\quad & \frac{N\Psi'(N)}{\Psi(N)}\frac{\delta-\beta}{(1-\beta)(\delta-1)} \\ = \quad & \frac{\xi(c)}{c}\left[\Phi'(1)\frac{c}{\xi(c)}\frac{1}{1-\beta} - \eta + F\left(\frac{\bar{K}}{N}, 1-\lambda, A\right)\left(\frac{\alpha\delta}{\delta-1} - \frac{(1-\alpha)\lambda}{1-\lambda}\right)\right].\end{aligned}$$

$$F\left(\frac{\bar{K}}{N}, 1-\lambda, A\right) = c + \eta \tag{49}$$

and

$$\frac{\theta_{t+1}}{\theta_t} = \frac{\mu_{t+1}}{\mu_t} = \frac{\gamma_{t+1}}{\gamma_t} = \delta.$$

(ii) If $\delta = \beta$, the steady state does not exist.

Proof.(i) When $\eta > 0$, N is finite. Consider first the case $\delta > \beta$. One can first show that $\theta_{t+1}/\theta_t = \delta$. Otherwise if $\theta_{t+1}/\theta_t > \delta$, then in the limit, according to equation (15),

$$\theta_{t+1}/\theta_t = \beta + \frac{\delta^{t+1}\Psi(N)}{N\theta_t} \tag{50}$$

then $\theta_{t+1}/\theta_t = \beta < \delta$, a contradiction. If $\theta_{t+1}/\theta_t < \delta$, then the right hand side of (50) explodes which also leads to a contradiction. ((11)) then implies that the growth rate of μ_t is the same as that of θ_t, which is δ. Furthermore, (16) at steady state simplifies to:

$$\Psi'(N)\,U - \frac{\mu_{t+1}}{\delta^{t+1}} F_K \frac{\bar{K}}{N} = \frac{\gamma_{t+1}}{\delta^{t+1}} \left(\frac{\gamma_{t+2}}{\gamma_{t+1}} - 1 \right). \tag{51}$$

The left hand side of this equality is constant in steady state since the growth rate of μ is δ. Then for the right hand side to converge to a constant we have the following three possibilities: γ_t grows at a rate smaller than δ, γ_t grows at the rate δ, and γ_t keeps constant over time, e.g. $\frac{\gamma_{t+2}}{\gamma_{t+1}} = 1$. Consider the first possibility when γ_t grows at a rate smaller than δ, then

$$\Psi'(N)\,U = \frac{\mu_{t+1}}{\delta^{t+1}} F_K \frac{\bar{K}}{N} \tag{52}$$

Express (15) and (11) at steady state,

$$\frac{\delta^{t+1}}{N\theta_t} \Psi(N) + \Phi(1) = \delta \Rightarrow \theta_t = \frac{\delta^{t+1}\Psi(N)}{N(\delta - \beta)}$$

$$\mu_t = \theta_t u'(c) = \frac{\delta^{t+1}\Psi(N)}{N(\delta - \beta)} u'(c) \tag{53}$$

Plug it into (52) multiplied by $\frac{N}{\Psi(N)}$ which is zero,

$$\frac{N\Psi'(N)}{\Psi(N)} U = \frac{\delta}{\delta - \beta} u'(c)\, F_K \frac{\bar{K}}{N}$$

$$\Rightarrow \frac{N\Psi'(N)}{\Psi(N)} c = \frac{\delta(1-\beta)}{\delta - \beta} \xi(c)\, F_K \frac{\bar{K}}{N}$$

By the constant return to scale assumption and the definition of α above, it can be written as

$$\frac{N\Psi'(N)}{\Psi(N)} \frac{c}{\xi(c)} = \frac{\delta(1-\beta)}{\delta - \beta} \alpha \frac{F\left(\bar{K}, L, A\right)}{N} \tag{54}$$

which together with (49) and $L = N(1-\lambda)$ can be used to solve (c, N). Express (17) at steady state as

$$\Phi'(1) \frac{c}{\xi(c)} \frac{1}{1-\beta} = \eta + (1-\alpha) \frac{F\left(\bar{K}, L, A\right)}{N} \frac{\lambda}{1-\lambda} \tag{55}$$

which uses the result that μ_t grows at the rate δ according to (53). (c, N) solved from (54) and (49) do not satisfy (55) in general. Therefore, in the case of δ bigger than β, the steady state with each multiplier growing at a constant rate, in particular γ_t growing at a constant rate smaller than δ, is not the optimal solution except for a knife-edge condition in which (c, N) satisfies (49), (54), and (55) simultaneously. Next consider the second possibility when γ_t grows at the rate of δ. Express (15), (17) and (11) at steady state, respectively, as

$$\frac{\delta^{t+1}}{N\theta_t}\Psi(N) + \Phi(1) = \delta \Rightarrow \theta_t = \frac{\delta^{t+1}\Psi(N)}{N(\delta-\beta)},$$

$$\gamma_t = \frac{1}{\delta}\mu_t\left[\Phi'(1)\frac{U}{u'(c)} - \eta - F_L\lambda\right],$$

and

$$\mu_t = \theta_t u'(c) = \frac{\delta^{t+1}\Psi(N)}{N(\delta-\beta)}u'(c).$$

Plug μ_{t+1} into the steady state formula of ((16)),

$$\frac{\delta^{t+1}}{\gamma_{t+1}}\Psi'(N)U = \frac{\delta^{t+2}\Psi(N)}{\gamma_{t+1}N(\delta-\beta)}u'(c)F_K\frac{\bar{K}}{N} + \delta - 1.$$

Plug γ_{t+1} into it,

$$\frac{\delta^{t+1}}{\gamma_{t+1}}\Psi'(N)U = \frac{\delta}{\Phi'(1)\frac{U}{u'(c)} - \eta - F_L\lambda}F_K\frac{\bar{K}}{N} + \delta - 1$$

Manipulate terms and use the formula of α, we obtain

$$\begin{aligned} & \frac{N\Psi'(N)}{\Psi(N)}\frac{1}{1-\beta} \\ = \;& \frac{\delta-1}{\delta-\beta}\frac{\xi(c)}{c}\left[\Phi'(1)c\frac{1}{\xi(c)}\frac{1}{1-\beta} - \eta + F\left(\frac{\bar{K}}{N}, 1-\lambda, A\right)\left(\frac{\alpha\delta}{\delta-1} - \frac{(1-\alpha)\lambda}{1-\lambda}\right)\right]. \end{aligned}$$

which together with (49) solves (c, N). For the third possibility $\frac{\gamma_{t+1}}{\gamma_t} = 1$ at steady state, which together with $\frac{\mu_{t+1}}{\mu_t} = \delta < 1$ contradicts with equation (17).
(ii) If $\delta = \beta$, (15) becomes

$$\frac{\theta_{t+1}}{\theta_t} = \frac{\beta^{t+1}}{\theta_t}\frac{\Psi(N)}{N} + \beta \geq \beta.$$

If $\frac{\theta_{t+1}}{\theta_t} = \beta$, then $\lim_{t\to\infty}\frac{\beta^{t+1}}{\theta_t}\frac{\Psi(N)}{N} > 0$ and $\frac{\theta_{t+1}}{\theta_t}$ converges to a number strictly bigger than β and a contradiction arises since $N > 0$ when $\eta > 0$. If $\frac{\theta_{t+1}}{\theta_t} > \beta$, then $\frac{\theta_{t+1}}{\theta_t} = \lim_{t\to\infty} \frac{\beta^{t+1}}{\theta_t} \frac{\Psi(N)}{N} = 0$, a contradiction. Hence steady state with Lagrange multipliers growing at constant rate over time does not exist if $\eta > 0$ which requires N to be finite. ■

A.2.2. Stability of the steady state

To get some insights about the stability of the steady state, in this section we also focus on the case when $\delta = 0$, in which case the social planner only cares about future generations to the extent that the initial generation does. The cost of raising children is only in terms of goods cost, which implies $\lambda = 0$. Furthermore, assume the social planner's weight and parental altruism functions take the Barro Becker's form. They are $\psi(N) = N^{\psi}$ and $\Phi(n) = \beta n^{\psi}$, respectively. We still assume $u(c) = \frac{c^{\xi}}{\xi}$. Then the initial parent's utility is

$$U_0 = u(c_0) + \beta n_0^{\psi} U_1 = \sum_{t=0}^{\infty} \beta^t \prod\nolimits_{j=0}^{t-1} n_j^{\psi} u(c_t)$$

The social planner's objective is

$$\psi(N_0) U_0 = N_0^{\psi} \sum_{t=0}^{\infty} \beta^t \prod\nolimits_{j=0}^{t-1} n_j^{\psi} \frac{1}{\xi} c_t^{\xi} = \sum_{t=0}^{\infty} \beta^t N_t^{\psi} \frac{1}{\xi} c_t^{\xi}$$

The weight on every individual living in generation t is $\beta^t N_t^{\psi}$. Then the problem becomes

$$\max_{\{C_t, N_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t N_t^{\psi} \frac{1}{\xi} c_t^{\xi}$$

subject to

$$c_t N_t = F\left(\bar{K}, N_t; A\right) - N_{t+1}\eta$$

where c_t is the per capita consumption of every individual belonging to generation t. Let the production take the Cobb-Douglas form, e.g. $F\left(\bar{K}, N_t; A\right) = A\bar{K}^{\alpha} N_t^{1-\alpha}$. The optimality choice of population in period t is

$$N_t^{\psi-\xi} C_t^{\xi-1} \eta = \beta N_{t+1}^{\psi-\xi} C_{t+1}^{\xi-1} \left[\left(\frac{\psi}{\xi} - \alpha\right) A\bar{K}^{\alpha} N_{t+1}^{-\alpha} - \eta \frac{\psi - \xi}{\xi} \frac{N_{t+2}}{N_{t+1}}\right]$$

where $C_t = c_t N_t$ is the aggregate consumption of all people of generation t. Let $X_t \equiv N_t^{-(\psi-\xi)/(1-\xi)} C_t$, the aggregate consumption times a factor that depends on population of time t. Then the following two equations characterize the dynamics of the system:

$$X_t = N_t^{\frac{\xi-\psi}{1-\xi}} \left[A\bar{K}^{\alpha} N_t^{1-\alpha} - N_{t+1}\eta \right] \tag{56}$$

$$\left(\frac{X_{t+1}}{X_t}\right)^{1-\xi} \eta = \beta \left[\left(\frac{\psi}{\xi} - \alpha\right) A\bar{K}^{\alpha} N_{t+1}^{-\alpha} - \eta \frac{\psi - \xi}{\xi} \frac{N_{t+2}}{N_{t+1}} \right] \tag{57}$$

Steady state population, N^*, can be solved as

$$A\bar{K}^{\alpha} N^{*-\alpha} = \frac{\psi - \xi + \xi/\beta}{\psi - \alpha\xi}\eta.$$

Stability of this steady state can be obtained by analyzing the log linearized the system represented by a matrix form

$$\begin{bmatrix} \eta \frac{N^{*(1-\psi)/(1-\xi)}}{X^*} & 1 \\ \beta \frac{\psi-\xi}{\xi} & 1-\xi \end{bmatrix} \begin{bmatrix} \frac{dN_{t+2}}{N^*} \\ \frac{dX_{t+1}}{X^*} \end{bmatrix} \tag{58}$$

$$= \begin{bmatrix} (\xi-\psi)/(1-\xi) + (1-\alpha)\left(1 + \eta \frac{N^{*(1-\psi)/(1-\xi)}}{X}\right) & 0 \\ (1-\alpha)\beta\frac{\psi-\xi}{\xi} - \alpha & 1-\xi \end{bmatrix} \begin{bmatrix} \frac{dN_{t+1}}{N^*} \\ \frac{dX_t}{X^*} \end{bmatrix}$$

The following proposition provides the condition under which the steady state is saddle path stable.

Proposition 14. The sufficient and necessary conditions for saddle path stability of the steady state are

$$\begin{aligned} &(2-\alpha)\frac{\frac{\psi-\xi}{\xi}(1-2\xi+(1-\alpha)\beta) + (1-\alpha)2(1-\xi)}{1/\beta - (1-\alpha)} \\ > \; &\alpha(1-2\xi) - 2(1-\psi). \end{aligned}$$

and

$$\frac{1}{\beta}(\alpha(1-\xi) - 1 + \psi) \neq (1-\alpha)\frac{\psi-\xi}{\xi},$$

*Proof.*Equation (58) can be written in the following form

$$\begin{bmatrix} \frac{dN_{t+2}}{N^*} \\ \frac{dX_{t+1}}{X^*} \end{bmatrix} = D \begin{bmatrix} \frac{dN_{t+1}}{N^*} \\ \frac{dX_t}{X^*} \end{bmatrix}.$$

where

$$D \equiv \begin{bmatrix} \eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}} & 1 \\ \beta\frac{\psi-\xi}{\xi} & 1-\xi \end{bmatrix}^{-1} \begin{bmatrix} \frac{\xi-\psi}{1-\xi}+(1-\alpha)\left(1+\eta\frac{N^{(1-\psi)/(1-\xi)}}{X}\right) & 0 \\ (1-\alpha)\beta\frac{\psi-\xi}{\xi}-\alpha & 1-\xi \end{bmatrix}$$

$$= \frac{1}{d}\begin{bmatrix} \begin{pmatrix} \xi-\psi+(1-\xi)(1-\alpha)\left(1+\eta\frac{N^{(1-\psi)/(1-\xi)}}{X}\right) \\ +\alpha-(1-\alpha)\beta\frac{\psi-\xi}{\xi} \end{pmatrix} & \xi-1 \\ -\beta\frac{\psi-\xi}{\xi}\left((1-\psi)/(1-\xi)-\alpha\right)-\alpha\eta\frac{N^{(1-\psi)/(1-\xi)}}{X} & \eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}(1-\xi) \end{bmatrix} \quad (59)$$

where $d=\eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}(1-\xi)-\beta\frac{\psi-\xi}{\xi}$. $\det(D)$ and $tr(D)$ are solved as

$$\begin{aligned} & d^{2}\det(D) \\ = & \begin{bmatrix} \xi-\psi+(1-\xi)(1-\alpha)\left(1+\eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}\right) \\ +\alpha-(1-\alpha)\beta\frac{\psi-\xi}{\xi} \end{bmatrix}\eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}(1-\xi) \\ & -(1-\xi)\left[\beta\frac{\psi-\xi}{\xi}\left((1-\psi)/(1-\xi)-\alpha\right)+\alpha\eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}\right] \end{aligned} \quad (60)$$

$$\begin{aligned} & d\cdot tr(D) \\ = & \ \xi-\psi+(1-\xi)(1-\alpha)\left(1+\eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}\right) \\ & +\alpha-(1-\alpha)\beta\frac{\psi-\xi}{\xi}+\eta\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}(1-\xi) \end{aligned} \quad (61)$$

Let λ_1 and λ_2 denote the eigenvalues of the matrix D. Assume $\lambda_1>\lambda_2$ without loss of generality, and they are determined by

$$\begin{aligned} \lambda_1 &= \frac{tr(D)+\sqrt{tr(D)^2-4\det(D)}}{2} \\ \lambda_2 &= \frac{tr(D)-\sqrt{tr(D)^2-4\det(D)}}{2} \end{aligned}$$

The necessary and sufficient condition for saddle-path stability is that $|\lambda_1|<1$ and $|\lambda_2|>1$ or $|\lambda_1|>1$ and $|\lambda_2|<1$. Since $\lambda_1>\lambda_2$, this condition can be divided into two groups: (i) $\lambda_1>1$ and $-1<\lambda_2<1$ and (ii) $\lambda_2<-1$ and $-1<\lambda_1<1$. Let us first consider case (i) in the following. In this case we have

$$\sqrt{tr(D)^2-4\det(D)}>2-tr(D)$$

and

$$tr\left(D\right)-2<\sqrt{tr\left(D\right)^{2}-4\det\left(D\right)}<tr\left(D\right)+2.$$

They can be reduced to

$$1-tr\left(D\right)<-\det\left(D\right)<1+tr\left(D\right),$$

which is equivalent with

$$\left(1-D_{11}\right)\left(1-D_{22}\right)<D_{12}D_{21}<\left(1+D_{11}\right)\left(1+D_{22}\right).$$

Next let us consider case (ii) in which

$$-2-tr\left(D\right)<\sqrt{tr\left(D\right)^{2}-4\det\left(D\right)}<2-tr\left(D\right)$$

and

$$tr\left(D\right)+2<\sqrt{tr\left(D\right)^{2}-4\det\left(D\right)}.$$

They can be reduced to

$$tr\left(D\right)+1<-\det\left(D\right)<-tr\left(D\right)+1,$$

which is equivalent with

$$\left(1+D_{11}\right)\left(1+D_{22}\right)<D_{12}D_{21}<\left(1-D_{11}\right)\left(1-D_{22}\right).$$

Writing these two set of conditions using elements of D. In case (i),

$$\left(1-D_{11}\right)\left(1-D_{22}\right)<D_{12}D_{21}$$

is equivalent with

$$\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}>\beta\frac{\psi-\xi}{\xi}$$

$$\Leftrightarrow\frac{1}{\beta}\left(-1+\psi+\alpha\left(1-\xi\right)\right)<\frac{\psi-\xi}{\xi}\left(1-\alpha\right).$$

while

$$D_{12}D_{21}<\left(1+D_{11}\right)\left(1+D_{22}\right)$$

can be written as

$$\begin{aligned}&\left(\begin{array}{c}2\left(2-\alpha\right)\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}\\-\left(2-\alpha\right)\beta\frac{\psi-\xi}{\xi}+2\left(1-\psi\right)+\alpha-2\alpha\left(1-\xi\right)\end{array}\right)\left(\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}-\beta\frac{\psi-\xi}{\xi}\right)\\ &> 0\end{aligned}$$

If $\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}>\beta\frac{\psi-\xi}{\xi}$, which is equivalent with $\left(1-D_{11}\right)\left(1-D_{22}\right)<D_{12}D_{21}$, then the above inequality holds if and only if

$$2\left(2-\alpha\right)\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}-\left(2-\alpha\right)\beta\frac{\psi-\xi}{\xi}>2\alpha\left(1-\xi\right)-2\left(1-\psi\right)-\alpha.$$

At steady state, $\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}=\frac{\psi/\xi-\alpha}{1/\beta-(1-\alpha)}$, the inequality above can be written as

$$\left(2-\alpha\right)\frac{\frac{\psi-\xi}{\xi}\left(1-2\xi+\left(1-\alpha\right)\beta\right)+\left(1-\alpha\right)2\left(1-\xi\right)}{1/\beta-\left(1-\alpha\right)}>\alpha\left(1-2\xi\right)-2\left(1-\psi\right)$$

In case (ii),

$$D_{12}D_{21}<\left(1-D_{11}\right)\left(1-D_{22}\right)$$

which is equivalent with

$$\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}<\beta\frac{\psi-\xi}{\xi}$$

$$\Leftrightarrow\frac{1}{\beta}\left(\alpha\left(1-\xi\right)-1+\psi\right)>\left(1-\alpha\right)\frac{\psi-\xi}{\xi}$$

while

$$\left(1+D_{11}\right)\left(1+D_{22}\right)<D_{12}D_{21}$$

can be written as

$$\begin{aligned}&\left[\begin{array}{c}2\left(2-\alpha\right)\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}\\-\left(2-\alpha\right)\beta\frac{\psi-\xi}{\xi}+2\left(1-\psi\right)+\alpha-2\alpha\left(1-\xi\right)\end{array}\right]\left[\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}-\beta\frac{\psi-\xi}{\xi}\right]\\ &< 0\end{aligned}$$

If $\eta\left(1-\psi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}<\beta\frac{\psi-\xi}{\xi}$, which is equivalent with $D_{12}D_{21}<\left(1-D_{11}\right)\left(1-D_{22}\right)$, then the above inequality holds if and only if

$$2\left(2-\alpha\right)\eta\left(1-\xi\right)\frac{N^{*(1-\psi)/(1-\xi)}}{X^{*}}-\left(2-\alpha\right)\beta\frac{\psi-\xi}{\xi}>2\alpha\left(1-\xi\right)-2\left(1-\psi\right)-\alpha.$$

Substitute $\frac{N^{*(1-\psi)/(1-\xi)}}{X^*}$ with the value at steady state, the inequality in this case becomes

$$\frac{1}{\beta}\left(\alpha\left(1-\xi\right)-1+\psi\right)>\left(1-\alpha\right)\frac{\psi-\xi}{\xi}$$

and

$$\begin{aligned}&\left(2-\alpha\right)\frac{\frac{\psi-\xi}{\xi}\left(1-2\xi+\left(1-\alpha\right)\beta\right)+\left(1-\alpha\right)2\left(1-\xi\right)}{1/\beta-\left(1-\alpha\right)}\\ > \quad & \alpha\left(1-2\xi\right)-2\left(1-\psi\right).\end{aligned}$$

The conditions in two cases together consist of the sufficient and necessary condition for saddle path stability, which is summarized as

$$\begin{aligned}&\left(2-\alpha\right)\frac{\frac{\psi-\xi}{\xi}\left(1-2\xi+\left(1-\alpha\right)\beta\right)+\left(1-\alpha\right)2\left(1-\xi\right)}{1/\beta-\left(1-\alpha\right)}\\ > \quad & \alpha\left(1-2\xi\right)-2\left(1-\psi\right).\end{aligned} \tag{62}$$

and

$$\frac{1}{\beta}\left(\alpha\left(1-\xi\right)-1+\psi\right)\neq\left(1-\alpha\right)\frac{\psi-\xi}{\xi}$$

■

Given $\xi<\psi$, the Barro-Becker's assumption for the concavity of the problem, this condition holds for most sets of parameters. The second condition holds except for a wide range of parameters. In particular, a nice sufficient condition guarantees saddle path stability is $\xi<\frac{1}{2}$ and $\alpha\left(1-2\xi\right)\leq2\left(1-\psi\right).$ We summarize it in the following Corollary.

Corollary 15. A sufficient condition for saddle path stability of the steady state are $\xi<\frac{1}{2}$ and $\alpha\left(1-2\xi\right)<2\left(1-\psi\right).$

Under this condition, the left hand side of (62) is positive while its right hand side is nonpositive.