INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH

ICER

WORKING PAPER SERIES

Fabio Privileggi and Guido Cozzi

WEALTH POLARIZATION AND PULVERIZATION
IN FRACTAL SOCIETIES

Working Paper no. 39/2002
September 2002

APPLIED MATHEMATICS
WORKING PAPER SERIES

Wealth Polarization and Pulverization in Fractal Societies*

Guido Cozzi† Fabio Privileggi‡

September 2002

Abstract

In this paper we study the geometrical properties of the support of the limit distributions of income/wealth in economies with uninsurable individual risk, and how they are affected by technology and preference parameters and by policy variables. We work out two simple successive generation models with stochastic human capital accumulation and with R&D and we prove that intense technological progress makes the support of the wealth distribution converge to a fractal Cantor-like set. Such limit distribution implies the disappearance of the middle class, with a "gap" between two polarized wealth clusters that widens as the growth rate becomes higher. Hence, we claim that in a highly meritocratic world in which the payoff of the successful individuals is high enough, and in which social mobility is strong, societies tend to look highly "fractalized". We also show that a redistribution scheme financed by proportional taxation does not help cure society's disconnection/polarization; on the contrary, it might increase it. Finally we show that these results are not confined to our analytically worked out examples but are easily extended to a widely used class of macroeconomic and growth models.

Keywords: Inequality and Growth, Education, Technological Change, Wealth Polarization/Pulverization, Iterated Function System, Attractor, Fractal, Cantor Set, Invariant Distribution.

JEL Classification Numbers: C61, O41

*We wish to thank Karl Shell, Neil Wallace, Manuel Santos, Randall Wright, Salvador Ortigueira and all the participants of the *Cornell-PSU Macro Workshop* at Pennsylvania State University (spring 2002) for insightful and constructive comments. We are especially indebted to Luigi Montrucchio for inspiration and encouragement, and to Tapan Mitra for giving the second author the opportunity to be visiting at the Department of Economics, Cornell University, in spring 2000 and in spring 2002. All the typical caveats apply.

†Dipartimento di Economia Pubblica, Università di Roma *La Sapienza*; e-mail: gcozzi@dep.eco.uniroma1.it.

‡Dipartimento di Politiche Pubbliche e Scelte Collettive - *POLIS*, Università del Piemonte Ortientale, and ICER; e-mail: fabio.privileggi@unipmn.it.

Contents

1 Introduction

How do we predict a growing and unequal society's wealth distribution to look like? In this paper we show how easily the support of the limit distribution of individual relative wealth levels can easily look like a peculiar geometric object called a Cantor set: a fractal, that is, a totally disconnected set with self-similar structure. We construct two variants of a simple competitive economy with successive generations and uninsurable individual risk in which a high exogenous growth rate renders the general equilibrium wealth distribution converge to a support that is a Cantor set. Hence, we argue that a high growth rate is a source of socioeconomic disconnection.

An emerging phenomenon of income or wealth polarization has been lately observed in many economies. For some empirical contributions see, for example, Alesina and Rodrik [4], Perotti [38], Benabou [12], Benhabib and Spiegel [14], Barro [10] and [11], Forbes [26]. From the theoretical point of view, the literature on income inequality and polarization appears to be already rich enough, both from the perspective of the possible consequences that polarization may have on growth rates and from the perspective of analyzing what aspects of growth may generate inequality and polarization. See, for example, Loury [35], Banerjee and Newmann [7], Galor and Zeira [27], Alesina and Rodrik [4], Persson and Tabellini [39], Benhabib and Rustichini [15], Aghion and Bolton [1], Benabou [12], Piketty [40], Aghion, Caroli and Garcia-Penalosa [3], Benabou [13].

In this paper, economies with (possibly) polarized wealth distribution in the long run are analyzed by means of a new approach, which makes use of Iterated Function Systems (IFS) to describe their dynamics and their limit distribution. Such methodology seems to be new in the literature. Up to our knowledge, only recently very few works appeared focussing on this type of dynamics with applications to economics, and none of them with the aim to explain wealth or income inequality. Some examples are Bhattacharya and Majumdar [16], [17], [18], who deal with IFS with random monotone maps, Montrucchio and Privileggi [37] and Mitra, Montrucchio and Privileggi [36], who studied stochastic economies converging to invariant probabilities supported on fractal sets.

Our analysis departs from the whole models on wealth inequality tackled by the authors cited above, where some imperfections in capital markets are being assumed in order to obtain persistent income or wealth inequality. Our framework is characterized by markets with a strong mobility engine associated to growth, and it is this strong mobility pattern itself that generates a fractal societies. Mobility is introduced through stochastic labor income heterogeneity, which represents the ability of the individuals to adopt better and better technologies. If better technologies entail some adoption uncertainty at the individual level and if such risk is uninsurable, due to the unobservable or unverifiable individual commitment in a learning effort, income heterogeneity becomes a natural consequence of aggregate growth, and the faster aggregate growth the relatively stronger the weight of the uncertain part of the individual resources.

A faster growing environment implies stronger family mobility prospects, because a successful individual from a poor family can more easily overtake the unsuccessful individuals of a richer family, but it means a tendency for the middle class to disappear as well. Hence a "hole" in the middle of the support of the wealth distribution is more likely to appear the faster the pace of technological growth: the wealth distribution becomes "polarized" in a high and a low wealth class. However, the random dynamical system that governs the individual assignment across the ever expanding social wealth distribution is not only polarizing the wealth distribution, but will mirror the central "hole" everywhere through the wealth distribution itself: the absence of a middle class at the social level implies the absence of "middle subclasses" at all levels, due to the diversity of the destinies of the different individuals who travel stochastically through society's wealth distribution. It follows that the same individual income stochasticity that generates a wealth distribution that is disconnected in the middle multiplies such disconnection at infinity - in all subintervals of it - generating a totally disconnected support of the wealth distribution in which no matter how small the interval considered it always lacks a "middle class". Therefore we reach what we can call a "pulverized" society. Such a "fractal society" is an intriguing mix of polarization and pulverization.

This kind of "polarization/ pulverization" of the aggregate wealth distribution has never been analyzed by the literature on inequality and growth, and it differs in kind from the traditional idea of "polarization". As a matter of fact, though if we photograph the wealth distribution at each point in time we get a highly "polarized" picture, when we track the processes for the successive wealth levels of any individual we observe a strong mobility. Hence, dynamically, such societies are not polarized in durable "classes", but they show a tremendous degree of mobility. Indeed it is mobility itself that generates polarization of the limit wealth distribution. The very fact that the gains of a lucky poor can make her richer than an unlucky rich is at the same time an important mobility aspect and the generator of a fractal society.

We will obtain fractalized wealth distributions from two versions of a simple macroeconomic model with no aggregate uncertainty and individual idiosyncratic income risk. This will deliver our main results in a very transparent way. We choose specifications to generate enough linearity in the random dynamical system and immediately translate the dynamics into well known properties of the Barnsley IFS used to generate the Cantor set.

An exercise that we provide in the paper regards the effect of a fiscal policy aimed at eliminating polarization/ pulverization through income taxation of those who are successful and redistribution to the unlucky individuals. Intuitively, since such policy is directly attacking the social mobility mechanism responsible - through dynamic general equilibrium effects - for the fractalization of society, one would expect that this would easily reach its target. We show that this is not the case. In fact, such a redistribution scheme can never eliminate the polarization/ pulverization of society. What's more, even if the free workings of the private economy itself did not imply socioeconomic disconnection, the introduction of taxation of stochastic incomes of all individuals may be able to generate socioeconomic disconnection and induce the polarization/ pulverization of society. Also the adoption of a random taxation scheme, which has in principle the potential of creating an artificial middle class in a polarized economy, proves essentially ineffective provided that the incentive compatibility constraint is sufficiently tight.

Finally, we show how easily the framework of generalized IFS may be used to study polarization/ pulverization phenomena in more general economic models, possibly with non-linear dynamics and state-dependent probabilities. We provide sufficient conditions for polarization/ pulverization and extend the result on inefficacy of redistributive policies also to the case of non-linear dynamics. Moreover, an extreme example of *a.s.* polarization due to state-dependent probabilities, and independent of the (possibly non fractal) properties of the attractor of the system, is presented. In a future paper we shall present a full characterization of neoclassical models with infinitely lived agents exhibiting polarization/ pulverization features.

The paper is organized as follows. In Section 2 we briefly review some of the basic mathematical methods we use to analyze the possibly fractal support of the limit distribution for a random dynamical system. A brief survey on the main results available from the IFS literature pertaining the linear (affine maps) case will be reported. The discussion will proceed at a simple illustrative level, while the reader will be referred to the literature for rigorous proofs and extensions. We will make use of diagrams to focus on the intuition about aspects that might be interesting for other potential economic applications. In Sections 3 and 4 we apply the methodology to two simple successive generation models of technological adoption and will describe the conditions necessary and sufficient for the limit distribution to admit a Cantor support. We will perform comparative statics exercises obtaining our main implications for the link between inequality and growth and the inefficacy of fractal-eliminating policies. Section 5 is devoted to a closer examination of the interplay between a more refined definition of polarization, in the spirit of Esteban and Ray [23], and what we have somewhat tentatively called "pulverization". Our conclusion is that a fractal society is compatible with (a deeper notion of) polarization also in the long run, *i.e.*, pulverization must not necessarily thwart polarization of the invariant distribution. Section 6 deals with non-linear IFS, to which the results on polarization/pulverization obtained for the previous models are substantially extended and translated into properties of generalized Cantor sets. The aim here is to show how an approach based on Iterated Function Systems may be helpful in exploring some particular aspects of dynamics in economics, like wealth polarization, usually hardly grasped by the traditional analysis. Finally, in Section 7 we shortly recall some facts on IFS with state-dependent probabilities which illustrate that, even in this setting, essentially our results remain unchanged under standard assumptions. A pathological counterexample of a (with probability 1) totally polarized economy is discussed, mainly to illustrate the role of the conditions imposed on the probability function. Section 8 concludes with some comments.

2 Some Mathematical Preliminaries

Since we shall study several variants of the same random system, here we describe general results pertaining to a class of contractive random dynamical systems that can be all normalized so that its possible values converge to some compact subset of the interval $[0, 1]$.

Consider the linear random dynamical system

$$w_t = \alpha w_{t-1} + z \tag{1}$$

where $0 < \alpha < 1$, z is a random variable taking one of the two values z_1 and z_2 with fixed probabilities $1-p$ and p respectively, $0 < p < 1$, for all $t \geq 1$, and $w_0 \in \mathbb{R}$. System (1) belongs to a larger family known in the literature as *(Hyperbolic) Iterated Function Systems* and can be expressively rewritten as

$$w_t = \begin{cases} \alpha w_{t-1} + z_1 & \text{with probability } 1-p \\ \alpha w_{t-1} + z_2 & \text{with probability } p \end{cases} \tag{2}$$

Proposition 1 *Suppose that $z_1 < z_2$. Then system (2), is* similar *to the system*

$$y_t = \begin{cases} \alpha y_{t-1} & \textit{with probability } 1-p \\ \alpha y_{t-1} + (1-\alpha) & \textit{with probability } p \end{cases} \tag{3}$$

which has the interval $[0, 1]$ as trapping region*. The (similar) map that converts (2) into (3) is given by*

$$y_t = \frac{1-\alpha}{z_2 - z_1} w_t - \frac{z_1}{z_2 - z_1} \tag{4}$$

Proof. Direct application of (4) to system (2). ■

In Section 6.1 it will be shown that such normalization to the interval $[0,1]$ is possible for any Iterated Function System with maps not necessarily linear and not necessarily increasing. System (3) can be actually obtained from system (2) by a direct application of formula (44) discussed there. In Section 6.1 it will be also argued that the analysis on polarization/pulverization carried out in the following sections can be easily generalized to a non-linear environment.

The concept of *similarity* used here means that there exist a transformation $S:\mathbb{R}\rightarrow\mathbb{R}$ such that $|S(x)-S(y)|=c\,|x-y|$ for some constant c and all $x,y\in\mathbb{R}$. A similarity transforms sets into geometrically similar ones, in the sense that preserves relative distances between points of the original set. This is peculiar in our analysis because similarities preserve the main features of the *attractor* of system (2), which, as we shall see, is the subset of the trapping region (to which the system eventually converges) that exhibits the polarization properties of the economy we want to discuss.

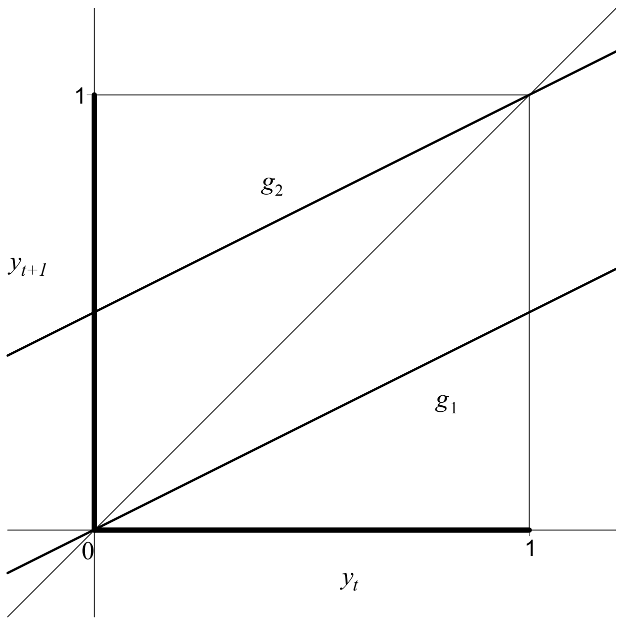


Figure 1: the interval $[0,1]$ is the trapping region of system (3), where $g_1(y)=\alpha y+(1-\alpha)$ and $g_2(y)=\alpha y$.

Figure 1 illustrate why interval $[0,1]$ is the *trapping region* of the contractive system (3): 0 is the fixed point of the map $g_1(y)=\alpha y$ and 1 is the fixed point of the map $g_2(y)=\alpha y+(1-\alpha)$; since, at each period, the system "jumps" from one map to the other with probabilities $1-p$ and p respectively, it must eventually remain "trapped" between 0 and 1.

Proposition 1 has an important consequence that will provide our results on ineffectiveness of redistributive policies in the following sections: the similarity between systems (2) and (3) implies that the geometrical properties of the former depend only on the contraction factor α, and not on the additive constants z_1 and z_2. Thus, any modification of z_1 and z_2 will have, through (4), only a "re-scaling effect" on the dynamics, without affecting the normalized system (3). More general results extended to non-linear Lipschitz maps are further discussed in Section 6.3.

We now recall some well known facts on Iterated Function System (3).

2.1 Iterated Function Systems

There is a huge literature available on Iterated Function Systems, which has grown very fast since, a few decades ago, it proved useful in techniques for generating approximated images of fractals on computer screens. Exhaustive treatment can be found, among others, in [30], [8], [21], [24], [50], [44], [33] and [25]. For a simplified exposition, focused on discussing an optimal growth model exhibiting the same dynamics as in (3), see also [36]. Here we shall contain ourselves by providing only the notions strictly necessary for our analysis.

For a given $0 < \alpha < 1$, define the maps from $[0, 1]$ to $[0, 1]$ by

$$\begin{cases} g_1(y) = \alpha y \\ g_2(y) = \alpha y + (1 - \alpha) \end{cases} \tag{5}$$

Then $\{g_1, g_2, p\}$ is a (contractive) *Iterated Function System* (IFS) over the interval $[0, 1]$; it summarizes the dynamics expressed in (3), whose asymptotic behavior is the focus of the present subsection. System (5) induces an operator T on $\mathbb{R}$, called *Barnsley operator*, defined by

$$T(B) = g_1(B) \cup g_2(B), \qquad B \subset \mathbb{R}, \tag{6}$$

where $g_j(B)$ denotes the image of the set B through g_j, $j = 1, 2$. Successive iterations of T transform the original set $B \subset \mathbb{R}$ into a sequence of sets $B_t = T\left[T^{t-1}(B)\right]$ through time. We are interested in properties of the limiting set, if it exists, to which the sequence B_t might eventually converge. A set $A \subset \mathbb{R}$ is called an *invariant set* or *attractor* for (5) if it is compact and satisfies

$$T(A) = A.$$

It is a set such that, once entered by the IFS, successive iterations of T keep the system inside it.

Since (3) describes a stochastic dynamical system, another important aspect of the IFS is the evolution through time of marginal probability distributions. Given any initial distribution μ_0 over $\mathbb{R}$, it is interesting to study how this probability evolves following the IFS. Let $\mathcal{B}$ be the σ-algebra of Borel measurable subsets of $\mathbb{R}$ and $\mathcal{P}$ the space of probability measures on $(\mathbb{R}, \mathcal{B})$. Define the Markov operator $M : \mathcal{P} \to \mathcal{P}$ as

$$M\mu(B) = (1 - p)\,\mu\left[g_1^{-1}(B)\right] + p\mu\left[g_2^{-1}(B)\right], \qquad \text{for all } B \in \mathcal{B} \tag{7}$$

where $\mu \in \mathcal{P}$ and $g_j^{-1}(B)$ denotes the set $\{x \in \mathbb{R} : g_j(x) \in B\}$, $j = 1, 2$. Operator M is often called *Foias operator*. As we did for operator T, we want to study successive iterations of M starting from some initial probability μ_0: $\mu_t(B) = M\left[M^{t-1}\mu_0(B)\right]$, which yields the evolution of marginal probabilities of the system as time elapses. A probability distribution $\mu^* \in \mathcal{P}$ is said to be invariant with respect to M if

$$\mu^* = M\mu^*. \tag{8}$$

An invariant probability distribution is usually interpreted in economics as the stochastic steady state to which the economy eventually might converge starting from some initial distribution μ_0 (see for example [48] and [37]).

Here is the main result available on the "fixed point" of our IFS. Recall that the *support* of a probability distribution μ is the smallest closed set S such that $\mu(S) = 1$, and that a sequence μ_t of probabilities *converges weakly* to μ^* if $\lim_{t\to\infty} \int f d\mu_t = \int f d\mu^*$ for every bounded continuous function $f : \mathbb{R} \to \mathbb{R}$.

Theorem 1 *Consider the IFS described by* $\{g_1, g_2, p\}$.

i) *There is a unique attractor for the IFS; that is, a unique compact set* $A \subseteq [0,1]$, *such that* $g_1(A) \cup g_2(A) = A$.

ii) *There is a unique probability distribution* μ^* *on* $([0,1], \mathcal{B}([0,1]))$ *satisfying the functional equation (8), that is,*

$$\mu^*(B) = (1-p)\,\mu^*\left[g_1^{-1}(B)\right] + p\mu^*\left[g_2^{-1}(B)\right] \qquad \text{for all } B \in \mathcal{B}([0,1]). \tag{9}$$

iii) *A is the support of* μ^* *and, for any probability*[1] μ_0 *on* $([0,1], \mathcal{B}([0,1]))$, *the sequence* $\mu_t = M^t\mu_0$ *for* $t = 0,1,2,\ldots$, *converges weakly to* μ^*.

The original proof dates back to Hutchinson [30]. See also Falconer [24], Lasota and Mackey [33] and Mitra, Montrucchio and Privileggi [36] for further discussion.

Theorem 1 and the definition of weak convergence immediately provides some information on the limiting distribution, which will be useful later. Denote by $y^* \in [0,1]$ the random variable associated to the invariant distribution μ^*, that is, y^* is the *random fixed point*[2] of system (3). Then, functional equation (9) can be rewritten as

$$\mu^*(y^* \in B) = (1-p)\,\mu^*\left(\frac{y^*}{\alpha} \in B\right) + p\mu^*\left(\frac{y^*}{\alpha} - \frac{1-\alpha}{\alpha} \in B\right),$$

which allows for a direct computation of expectation and variance of y^*:

$$\mathbb{E}(y^*) = p \tag{10}$$

$$Var(y^*) = \frac{1-\alpha}{1+\alpha}p\,(1-p). \tag{11}$$

Note that these computations are justified thanks to weak convergence, since expectation and variance are the integrals of the identity function $f(y) = y$ and the function $f(y) = (y - \mathbb{E}_g(y))^2$ respectively, which are both bounded and continuous over $[0,1]$.

2.2 The Support of the Limiting Distribution

It is important to recall here some features of the support A of the invariant distribution (the attractor of system (5)) which depend only on contraction factor α of the IFS $\{g_1, g_2, p\}$ and are independent of p. This will be the key ingredient in explaining polarization phenomena in the following sections.

A quick glance at Figure 1 makes clear that the support of our IFS will be the whole interval $[0,1]$ whenever $1/2 \le \alpha < 1$. This is because $T([0,1]) = g_1([0,1]) \cup g_2([0,1]) = [0,1]$ if the images of g_1 and g_2 overlap, that is, if $1/2 \le \alpha < 1$, as Figure 2 shows. In this case we shall say that the distribution has "full support".

More interesting is the case when images $g_1([0,1])$ and $g_2([0,1])$ do not overlap: this happens for $0 < \alpha < 1/2$, since $g_1([0,1]) \cup g_2([0,1]) = [0,\alpha] \cup [1-\alpha, 1]$. As $\alpha < 1/2$, there is a "gap" between the two image sets, with amplitude $h(\alpha)$ equals to

$$h(\alpha) = 1 - 2\alpha > 0. \tag{12}$$

[1] To be precise, weak convergence holds for any initial probability μ such that $\int |x-a|\,d\mu < \infty$ for some constant a. See Section 2.1.2 in [36] for more details.

[2] See Arnold [6] for a detailed treatment of random dynamical systems and random fixed points.

Note that this amplitude is a decreasing function of α, the common slope of g_1 and g_2. Moreover, this gap will "spread" through the unit interval by successive applications of the maps (5), reproducing itself, scaled down by a factor $1/\alpha$, in the middle of each subinterval born after each step t. Figure 3 reproduces the first 3 iterations of (5) starting from $[0, 1]$, obtaining a union of 8 $(= 2^3)$ intervals of length α^3.

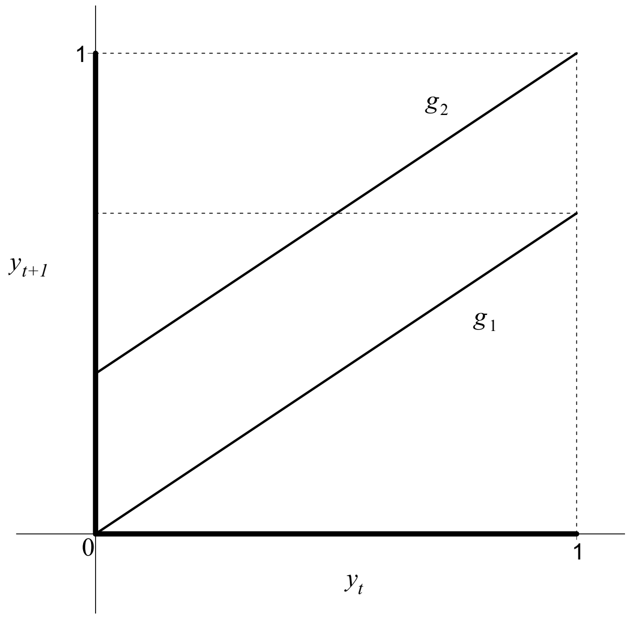


Figure 2: $g_1([0, 1]) \cup g_2([0, 1]) = [0, 1]$ when $1/2 \leq \alpha < 1$.

By pushing these iterations to the limit, we eventually find an attractor with features of the usual Cantor ternary set; in fact, for $\alpha = 1/3$, the support is precisely the Cantor ternary set. Cantor-like sets of the kind constructed by computing $\lim_{t\to\infty} T^t([0, 1])$ for $0 < \alpha < 1/2$ exhibit several geometrical properties that are typical of *fractals*.

The most bewildering - and intriguing - feature of fractals is the need of a more sophisticated tool than the topological dimension - which allows only for integer values - to measure the "consistency" of their structure. Several dimensions has been constructed for this purpose, like, among others, the Hausdorff dimension, the Box-counting dimension and the Similarity dimension. These dimensions consider the infimum partition made up of some "regular sets" (like balls or squares) that contains the fractal, and measure how fast the number of sets in the partition grows as finer and finer partitions are applied (for a discussion on dimensions see, for example, [24]). All fractals have the peculiarity that their dimension is a "fraction", from which the name "fractal"; for instance, Cantor-like sets which are the attractors of IFS (5), for $0 < \alpha < 1/2$, have Hausdorff dimension $-\ln 2/\ln \alpha$, which, in this case, is the same as the Box-counting and the Similarity dimensions.

Since $\alpha < 1/2$, the attractors we are studying have all dimension less than 1, which implies that they are *totally disconnected*, that is, between any two points in them, there are "holes" (points laying outside the attractor). To see this, note that if a Cantor-like set were containing some interval, its dimension[3] would be 1. Conversely, even if dimensions less than 1 denote sets with very "disperse" points, by means of a standard Cantor diagonal argument, it can

[3]This follows also from the fact that the topological dimension is not larger than the Hausdorff dimension. For a proof of this property see Theorem 6.2.9 in [21].

be easily shown that Cantor-like sets contain uncountably many points. In other words, our attractor has the same number of points as the whole interval [0, 1], but they are all pulverized across the interval itself (in the mathematical literature it is often referred as “Cantor dust”). However none of these points are isolated. Hence, since it is a closed set, any Cantor-like set is *compact, perfect and totally disconnected.* A terse and accessible discussion of the Cantor ternary set and its properties can be found in Chapter 11 in [49]. Also [19] is a good reference for an introductory approach.

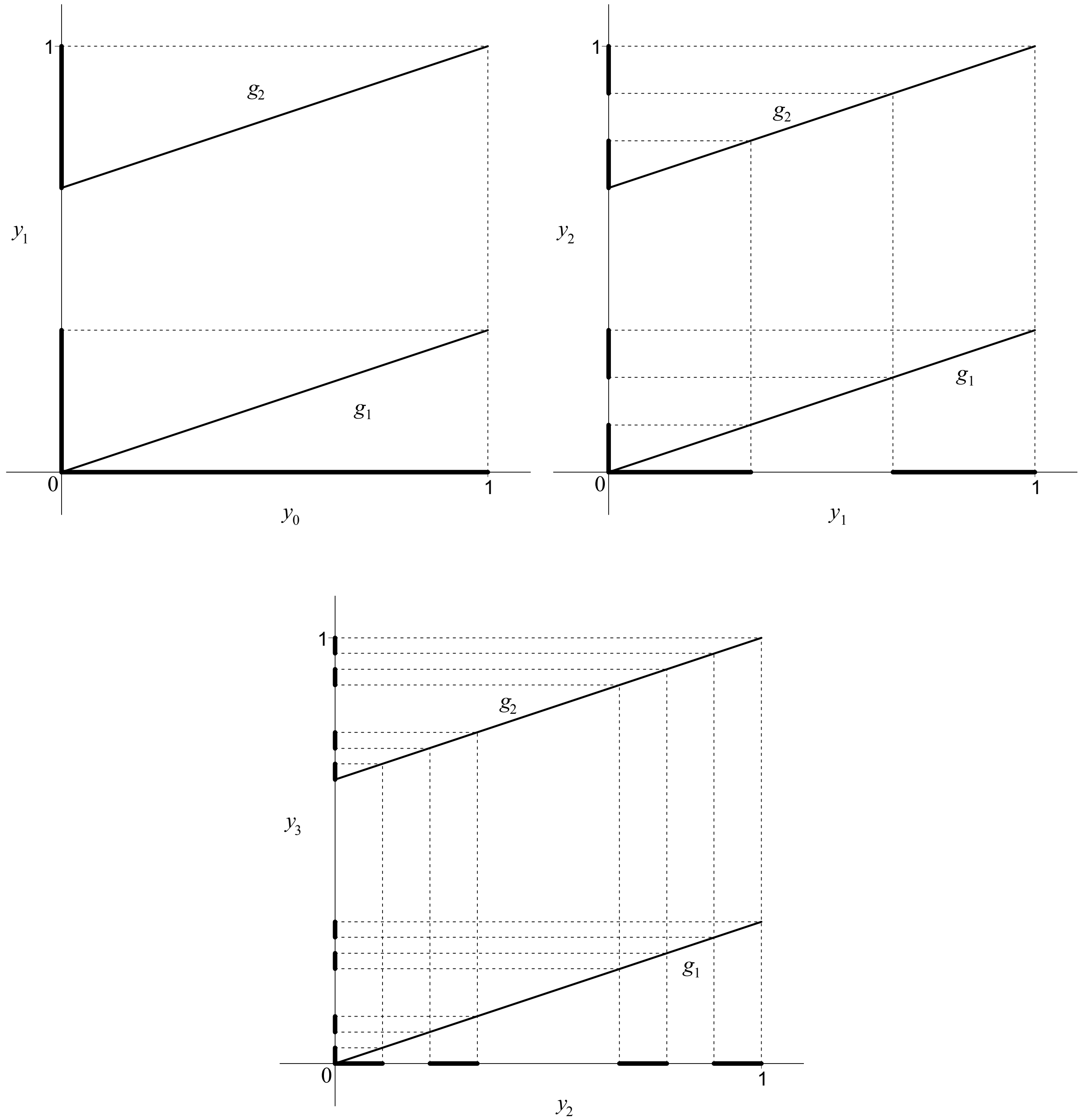


Figure 3: first 3 iterations of our IFS for $\alpha < 1/2$ starting from $[0, 1]$. The third iteration gives a union of 8 intarvals of length α^3, as can be seen on the vertical axis of the last figure.

2.3 The Invariant Distribution

Properties of the attractor discussed before shed some light also on the limiting distribution defined over the attractor itself. As a matter of fact, a set $A \subset \mathbb{R}$ with Hausdorff dimension less than 1 have Lebesgue measure zero[4]. This can be also shown through a direct argument following the construction in Figure 3: the Lebesgue measure of $T([0,1])$ is 2α, that of $T^2([0,1])$ is $2^2\alpha^2$ and that of $T^3([0,1])$ is $2^3\alpha^3$. By induction, it is immediately seen that $T^t([0,1])$ consists of 2^t intervals of Lebesgue measure (length) α^t, so that its Lebesgue measure $2^t\alpha^t \to 0$ as $t \to \infty$. Since $A = \lim_{t\to\infty} T^t([0,1])$ is the support of the invariant distribution μ^*, $\mu^*(A) = 1$, from which follows that μ^* turns out to be singular with respect to Lebesgue measure.

More results on singularity versus absolute continuity of μ^* are available and are widely discussed in [36]. Here we briefly recall some of them.

First, for all values of parameters α and p, the invariant distribution is of the "pure" type, *i.e.*, it is either singular or absolutely continuous on its whole support.

Secondly, it is highly sensitive with respect to changes in the one-period probability p. In fact, when $0 < \alpha < 1/2$, two invariant distributions generated by probabilities p and p' respectively, are mutually singular. This means that measures with the same support A (which depends on α and not on p) concentrate over subsets of A with empty intersection for different probabilities p and p'.

Moreover, μ^* can be singular also for values $\alpha \geq 1/2$, i.e. when images of g_1 and g_2 overlap. In particular, this happens for values of probability p bounded away from 1/2 (i.e., when p is close enough to 0 or 1). However, also for $p = 1/2$, some examples of singularity for some special values of α exist; Erdös [22] constructed some of them. These results, widely surveyed in [36], show that even if μ^* has the whole interval $[0,1]$ as its support, it may be concentrated over a subset of $[0,1]$ with zero Lebesgue measure.

To have a flavor of what such an invariant distribution might look like, one may draw some iterations of Foias operator M defined as in (7) starting from the uniform distribution over $[0,1]$. This, in the case $0 < \alpha < 1/2$, is equivalent to the following construction. Split a unit mass so that the right interval of $T([0,1])$ has mass p and the left interval has mass $1-p$. Then, divide the mass on each interval of $T([0,1])$ between the two subintervals of $T^2([0,1])$ in the ratio $p : 1-p$. Continue in this way, so that the mass on each interval of $T^t([0,1])$ is divided in the ratio $p : 1-p$ between its two subintervals in $T^{t+1}([0,1])$ (see also Example 17.1 in [24]). Figure 4 depicts some iterations of M using the construction of the Cantor ternary set ($\alpha = 1/3$) starting from the uniform distribution for $p = 1/3$.

Figure 5 shows two examples of 8 iterations of M in the overlapping case, *i.e.* for values of $\alpha \geq 1/2$, when the invariant distribution μ^* has full support. Note that for values of α close to 1 (high "degree of overlapping" of the images $g_1([0,1])$ and $g_2([0,1])$) and p sufficiently close to 1/2, as in case (a), the figure suggests that μ^* will be "smooth" (absolutely continuous); while, whenever α gets closer to 1/2 and p gets closer to the extrema 0 or 1, as in case (b), the approximation resembles the traits observed in the last approximations of Figure 4, where the limiting distribution is known to be singular.

[4] A rigorous proof of this fact, which uses the notion of *Hausdorff measure*, can be found in [21].

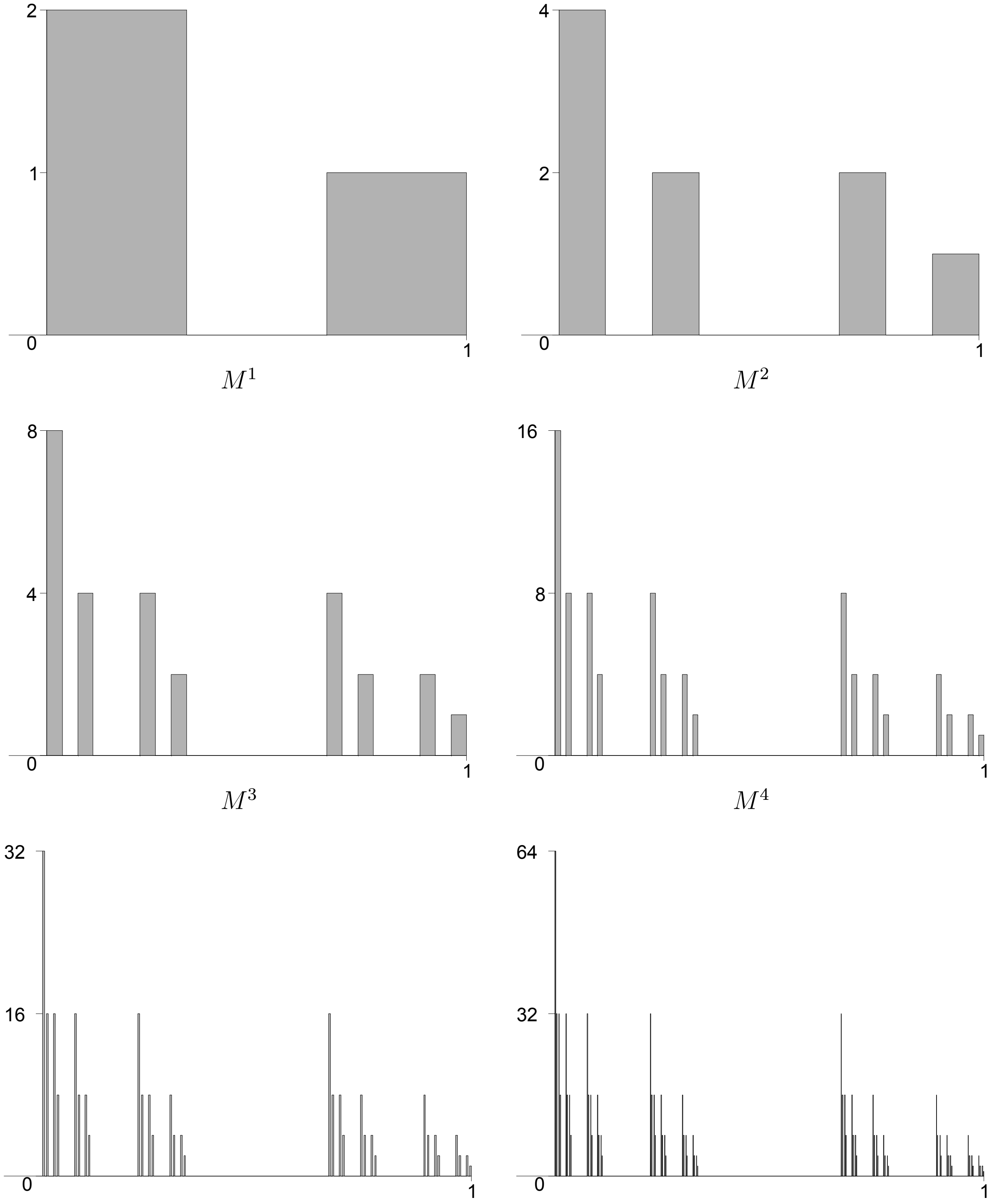
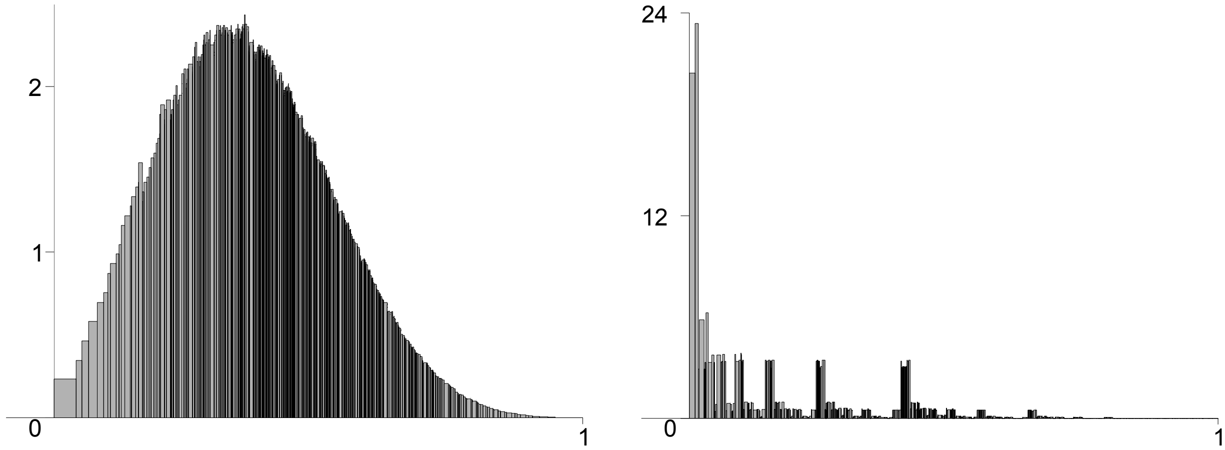


Figure 4: first 6 iterations of Foias operator starting from the uniform probability for $\alpha = 1/3$ and $p = 1/3$.



a) $\alpha = 4/5$ and $p = 1/3$; b) $\alpha = 3/5$ and $p = 1/8$.

Figure 5: two examples of the first 8 iterations of Foias operator starting from the uniform probability in the overlapping case, that is, for values $\alpha \geq 1/2$.

3 Adoption of New Technologies

In this section and in the next one we will apply the previously described methodology to a couple of models with inequality and growth. Here we assume a sequence of successive generations of altruistic individuals who take a consumption and bequest decision on their wealth accumulated out of a stochastic income acquired at the utility cost of learning a technology that is new at every generation. It may be interpreted as a simple macroeconomic model of technological adoption of exogenously arriving General Purpose Technologies.

Consider an infinite horizon discrete time economy with a continuum of infinitely lived families that will be indexed by i. With no loss of generality we shall normalize population over the unit interval, *i.e.*, $i \in [0, 1]$. Each family is formed by a one-period lived altruistic individuals whose preferences are represented by the following "warm glow" (see Andreoni [5]) utility function

$$u(c, b, e) = c^{1-\beta} b^{\beta} - e$$

where $c > 0$ denotes end-of-life consumption, $b > 0$ the bequest left to the unique heir, $e \geq 0$ a learning effort[5], and $0 < \beta < 1$ the degree of intergenerational altruism. As, for example, in Banerjee and Newman [7], Galor and Zeira [27], or Piketty [40], such Cobb-Douglas altruistic preferences imply that a fraction β of each individual's end of life wealth will be passed over to her child. Hence, the indirect utility of end-of-life wealth W is linear (risk neutral preferences) and equal to

$$U(W) = (1 - \beta)^{1-\beta} \beta^{\beta} W - e.$$

The end-of-life wealth W of each family is uncertain at the beginning of each generation: it depends on the wealth level inherited from the past, that is on the bequest left by the ancestor, and on individual success in learning the technology that become available during her lifetime.

[5] As will become clear later, each agent chooses to exert effort e between two values: zero and a strictly positive fixed amount which depends on time.

Individuals of generation t are endowed with one unit of labor time which they will inelastically use to produce a perishable consumption good at the common productivity level $A_t > 0$. At the beginning of period t, a new General Purpose Technology (see Helpman [28]) appears exogenously and every individual has to learn it in order to successfully enter production. Learning technology A_t requires an effort that entails a certain utility cost $e_t > 0$. Whether an individual exerts the required effort for learning such technology is something that cannot be observed by anybody but the individual[6]. Moreover "success" in the adoption of the technology is not sure, but it occurs to each individual with probability $0 < p < 1$ constant through time, independently of all other individuals. Since the (exertion of) learning effort is unobservable, borrower-creditor interaction lasts one period only and individual's offspring cannot be sanctioned no idiosyncratic risk can be insured.

Technology is assumed to evolve exogenously: $A_t = \gamma A_{t-1}$, where $\gamma > 1$. Consistently, we will assume that $e_t = \gamma e_{t-1}$, that is, learning a more advanced technology requires more effort.

Provided that individual $i \in [0, 1]$ alive in period t undertakes the learning effort e_t at the beginning of her life, her end-of-period *income* Y_t will be:

$$Y_t^i = \begin{cases} 0 & \text{with probability } 1-p \\ A_t & \text{with probability } p \end{cases}$$

Notice that in this model income derives from the "ability" in the use of current technologies and entails no utility loss.

3.1 The Wealth Distribution Dynamics

The evolution of technology yields $A_t = \gamma^t A_0$ and that of effort $e_t = \gamma^t e_0$, with both A_0 and e_0 strictly positive. Individual i wealth at the beginning of her life in period t is given by the bequest inherited from period $t-1$:

$$b_t^i = \beta W_{t-1}^i,$$

where W_{t-1}^i represents the wealth accumulated by her ancestor at the end of time $t-1$. Provided that individual i will perform effort e_t in order to learn technology A_t, her expected indirect utility conditional to the past wealth and the performed effort is given by

$$\begin{aligned} \mathbb{E}\left[U\left(W_t^i\right) \mid \left(W_{t-1}^i, e_t\right)\right] &= (1-\beta)^{1-\beta}\beta^\beta \mathbb{E}\left[W_t^i | W_{t-1}^i\right] - e_t \\ &= (1-\beta)^{1-\beta}\beta^\beta \left[p\left(\beta W_{t-1}^i + A_t\right) + (1-p)\beta W_{t-1}^i\right] - e_t \\ &= (1-\beta)^{1-\beta}\beta^\beta \left(\beta W_{t-1}^i + pA_t\right) - e_t \end{aligned} \tag{13}$$

where the probability of success p in adopting technology A_t does not depend on time. We shall assume the following.

A. 1

$$0 < e_0 < (1-\beta)^{1-\beta}\beta^\beta p A_0.$$

Assumption A.1 implies that the expected indirect utility obtained by exerting effort e_t is greater than the certain effort for all $t \geq 0$, thus rational individuals will always put the required effort into learning the new technology. It follows that the intergenerational motion of the wealth of family $i \in [0, 1]$ is described by

$$W_t^i = \begin{cases} \beta W_{t-1}^i & \text{with probability } 1-p \\ \beta W_{t-1}^i + A_t & \text{with probability } p. \end{cases} \tag{14}$$

[6]Specifically, it is not the amount of learning effort which is not observable, but whether an individual undertakes such effort at all.

To fix ideas, assume, for now, that the "original" bequest available at the beginning of period $t = 0$ is zero, that is,

$$W_0^i = \begin{cases} 0 & \text{with probability } 1-p \\ A_0 & \text{with probability } p \end{cases}$$

Since A_t grows exogenously through time, the random dynamical system (14) described by the two maps $f_1(W) = \beta W$ and $f_2(W) = \beta W + A_t$ evolves along increasing sets of possible wealths. In particular, at the end of period t generation i will be endowed with some wealth W_t^i in the interval

$$\left[0\ ,\ \left(\frac{1-(\beta/\gamma)^{t+1}}{\gamma-\beta}\right)\gamma^{t+1}A_0\right] \tag{15}$$

which, since $\gamma > 1$, $0 < \beta < 1$ and $A_0 > 0$, diverges to the interval $[0, +\infty]$ as $t \to +\infty$.

However, notice that, since $0 < \beta < 1$, both f_1 and f_2 in (14) are contractions in the variable W, that is, wealth grows only thanks to technological parameter A_t as time elapses. This allows a better highlighting of the features of this dynamics by transforming system (14) into an equivalent law of motion adjusted by the productivity level A_t, which turns out to be contractive and thus having some compact set as trapping region. In other words, (14) can be transformed into a system of the type of (2), so that the analysis of Section 2 can be directly applied.

Dividing (14) by A_t we get the equivalent system in terms of $w_t^i = W_t^i/A_t$:

$$w_t^i = \begin{cases} (\beta/\gamma)\, w_{t-1}^i & \text{with probability } 1-p \\ (\beta/\gamma)\, w_{t-1}^i + 1 & \text{with probability } p \end{cases} \tag{16}$$

whose trapping region is the interval $\left[0, (1-(\beta/\gamma))^{-1}\right]$. In view of Proposition 1, let

$$\alpha = \frac{\beta}{\gamma}, \tag{17}$$

which implies $0 < \alpha < 1$, and consider the transformation $y_t^i = (1-\alpha)\, w_t^i$ of (16). Hence we obtain the following productivity-adjusted normalized dynamic conjugate to (14):

$$y_t^i = \begin{cases} \alpha y_{t-1}^i & \text{with probability } 1-p \\ \alpha y_{t-1}^i + (1-\alpha) & \text{with probability } p \end{cases} \tag{18}$$

which has the same general form as in (3). Hence, Theorem 1 applied to process (18) guarantees that the productivity-adjusted normalized dynamic of wealths in our model converges weakly to some unique distribution μ^* with support in $[0, 1]$.

3.2 Growth and Fractalization

The stochastic dynamic model expressed by (18), or more generally by (3) in Section 2, turns out to be especially useful for a slightly different interpretation, which is the main focus of this paper. One-period probability p of individual i of successfully adopting technology A_t at the end of period t, can be seen, by the law of large numbers, as the "average proportion of the whole population" that in the long run is eventually able to catch the opportunity of benefitting from the new technology. In this scenario, in its steady state our economy tends to grow at a constant exogenous rate $\gamma - 1$, with social mobility due to uninsurable success and failure to exploit new technologies making each individual family travel over time across a constant distribution of relative wealths.

From this aggregate perspective, expectation (10) can be read as the aggregate productivity-adjusted wealth in the steady state, and variance (11) as the dispersion of individual wealths. From both expressions (10) and (11) it is immediately seen that the higher the individual probability p of exploiting technology A_t, the "richer" the economy on average, and the lower parameter $\alpha = \beta/\gamma$ (*i.e.*, the lower the altruism rate β or the higher the exogenous growth rate γ), or the more parameter p is bounded away from 1/2, the more the invariant distribution μ^* is concentrated around its mean.

However, in view of Section 2, we are in the position of saying much more on the steady state of such kind of economy. Specifically, we are interested in the possibility that it may be more or less polarized, that is, we are interested in the existence of a strong "middle class", which is often considered important for growth itself, for democracy, for sociopolitical stability, and for the law and order, as quantified, among others, in the empirical analyses of Alesina and Rodrik [4], Perotti [38] and Barro [10].

A strong "middle class" in this economy is represented by an invariant distribution that gathers a proportionally larger fraction of the population around 1/2 than close to the extrema 0 and 1 of the interval $[0, 1]$. Viceversa, we will call "polarized/ pulverized" a limiting distribution with Cantor-like support. Notice that the absence of a middle class is a necessary condition for the latter distribution, and in this sense we refer to *polarization*. However, the fractal nature of such support, that is its self-similar structure that infinitely replicates a "hole", creates a form of social disconnection that has not been studied in the existing theoretical and empirical literature on polarization, and that we somewhat tentatively will label *pulverization*.

The following result is an immediate consequence of the discussion in Section 2.2.

Proposition 2 *Under A.1, if* $1 < \gamma \leq 2\beta$*, there exists a middle class. Whenever* $\gamma > 2\beta$ *the economy becomes polarized/ pulverized in the long run, and the support* A *of its limit distribution* μ^* *is a Cantor-like set. Moreover, the larger* γ *(and the smaller* β*), the larger the gap between the fractions of the population near the extremes of the support.*

Proof. Apply (17) and (12) to system (18). ■

The previous proposition shows that a high economic growth rate, by rewarding the successful individuals and penalizing in relative terms those who are not ready to catch the opportunities associated with the new technologies, make the middle class disappear and polarize society in two very different wealth classes. Polarization becomes dramatic the larger the jump in productivity γ and the smaller the individual degree of altruism β (or, equivalently, the more selfish the individuals).

It is important to note that polarized wealth distribution does not mean that wealth classes are trapping the individuals: these become rich and poor in this economy and it is precisely the amplitude of the *social mobility* - and not the frequency, that is the individual probability p of catching the technological opportunity - that generates wealth polarization.

3.3 Redistribution and Social Cohesion

In this section we will show that any redistribution scheme aimed at doing away with social polarization/ pulverization is not capable of achieving its goal. By applying results from

Section 2, we shall see that lump-sum transfers from the rich generations to the poor do not have any effect on the polarization/ pulverization features of our economy. In fact, the "hole" that generates a "fractal society" depends only on parameters β (preferences) and γ (growth rate). We will not assume that polarization/ pulverization implies productivity losses, though it would be natural to motivate such political need for more socioeconomic cohesion.

Let us assume that the gains from success are taxed at the end of each period a proportion $0 \leq \tau < 1$ and that proceeds are redistributed lump-sum to the unluckies[7]. If all individuals exert effort e_t in order to learn technology A_t, the steady state proportion of rich families in the economy will still be p. Hence, the government in the long run will be able to collect tax revenues equals to $p\tau A_t$, which - assuming a balanced government budget every period - equals the aggregate lump sum transfer received at the end of period t by the whole poor.

Since taxation further reduces the expected benefit derived from having the opportunity of adopting technology A_t, in order to let all individuals keep putting effort e_t even under taxation and thus obtain a dynamic similar to that in (14), an upper bound on tax rate τ is needed. Let us discuss in detail how Assumption A.1 needs to be modified to avoid free riding behavior due to the possibility of receiving, out of nothing, a transfer that generates a higher utility than the expected utility gain produced by putting effort e_t.

Let $0 \leq l \leq 1$ denote the fraction of the population who decides to put effort e_t in learning technology A_t. Then, at the steady state, the total amount of tax revenues is $pl\tau A_t$, and each non-successful individual i - which are both the unlucky ones who exerted effort e_t and the lazy ones who did not exert any effort, that amount to a proportion $1-pl$ of families - receives a transfer given by

$$T_t^i = \frac{pl}{1-pl}\tau A_t. \tag{19}$$

In view of (13), the individual i expected utility gain conditional to effort e_t is given by

$$\begin{aligned}\mathbb{E}\left[U\left(Y_t^i\right)|e_t\right] &= \rho\left[p\left(1-\tau\right)A_t + \left(1-p\right)T_t^i\right] - e_t \\ &= \rho\left[p\left(1-\tau\right)A_t + \left(1-p\right)\frac{pl}{1-pl}\tau A_t\right] - e_t,\end{aligned}$$

where $\rho = \left(1-\beta\right)^{1-\beta}\beta^{\beta}$, while the individual i certain utility gain obtained by exerting zero effort is given by

$$U\left(T_t^i\right) = \rho\frac{pl}{1-pl}\tau A_t.$$

In order to let all the families put the effort $e_t = \gamma^t e_0$ required to learn technology A_t, we need

$$\mathbb{E}\left[U\left(Y_t^i\right)|e_t\right] > U\left(T_t^i\right)$$

to hold for all $0 \leq l \leq 1$, which leads to

$$\left(1 - \frac{\tau}{1-pl}\right)p\rho A_0 > e_0.$$

Since the minimum of the left hand side is reached for $l = 1$, then, for each given e_0 satisfying A.1, the following restriction on parameter τ guarantees that all families will always put effort e_t in learning technology A_t also under government taxation.

[7] Assuming lumps sum redistribution to all individuals - not only to the unluckies - would not alter the qualitative results of our analysis, as it will become clear later. See also Appendix 6.3.

A. 2 *Assumption A.1 holds and*

$$0 \leq \tau < (1-p)\left(1 - \frac{e_0}{p\rho A_0}\right), \tag{20}$$

where $\rho = (1-\beta)^{1-\beta}\beta^{\beta}$.

Hence, in view of (14), the dynamics of individual i wealth becomes:

$$W_t^i = \begin{cases} \beta W_{t-1}^i + p(1-p)^{-1}\tau A_t & \text{with probability } 1-p \\ \beta W_{t-1}^i + (1-\tau) A_t & \text{with probability } p, \end{cases} \tag{21}$$

where, in the first line, $p(1-p)^{-1}\tau A_t$ represents the transfer received by a single unlucky family, *i.e.*, T_t^i in (19), with $l = 1$.

Here is the main result of this section.

Proposition 3 *If* $\gamma > 2\beta$, *polarization/ pulverization never disappears for all income tax rates* τ *satisfying A.2.*

Proof. Divide both equations in (21) by A_t to get the productivity-adjusted dynamic

$$w_t^i = \begin{cases} (\beta/\gamma)\, w_{t-1}^i + p(1-p)^{-1}\tau & \text{with probability } 1-p \\ (\beta/\gamma)\, w_{t-1}^i + (1-\tau) & \text{with probability } p. \end{cases} \tag{22}$$

Since A.1 holds, the right hand side in (20) implies $\tau < 1-p$, which, in turn, implies $p(1-p)^{-1}\tau < (1-\tau)$. Therefore, under A.2, Proposition 1 holds, and thus system (22) is similar to system (18). Hence, Proposition 2 applies and polarization/ pulverization is completely determined by condition $\gamma > 2\beta$. ■

The comments following Corollary 1 in Section 6.3 provide a geometrical interpretation of this result: only the slopes of the two maps constituting the IFS affect polarization/ pulverization; constants have no effect in shaping the attractor.

There is, however, a main difference with respect to the previous section, which is not evident from the normalization provided by Proposition 1. Observing the evolution through time of the feasible sets of both systems (21) or (22), it is immediately clear that, while zero was included in all supports of the marginal distributions of the model in Section 3.2, the standard of living of the poor under wealth redistribution will be bounded away from zero in the long run, that is, nobody will end up with a zero wealth in the steady state. As a matter of fact, the feasible wealths of system (21) at time t lay in some subset of the interval

$$\left[\left(\frac{1-(\beta/\gamma)^{t+1}}{\gamma-\beta}\right)\gamma^{t+1} p(1-p)^{-1}\tau A_0\ ,\ \left(\frac{1-(\beta/\gamma)^{t+1}}{\gamma-\beta}\right)\gamma^{t+1}(1-\tau)A_0\right],$$

with lower extremum strictly positive and increasing over time. Therefore, although government redistribution does not affect polarization/ pulverization, it still proves effective in sustaining the wealth of the poor. Clearly also the "rich side" of the population is being affected by having a reduced - by factor $1-\tau$ - maximum possible wealth compared to that of the original feasible region (15). Thus, the overall effect of a redistributive policy by the government is to narrow the whole absolute wealth around its mean, without changing possible polarization/ pulverization phenomena in relative terms.

Policy inefficacy of income tax in eliminating polarization/ pulverization is a counterintuitive result. However we can prove an even stronger result by introducing capital taxation (not redistributed lump-sum). As a matter of fact, if final wealth is taxed at a rate $0 < \tau_w < 1$, the dynamical system for productivity adjusted wealth becomes:

$$w_t^i = \begin{cases} (1-\tau_w)(\beta/\gamma)\, w_{t-1}^i & \text{with probability } 1-p \\ (1-\tau_w)(\beta/\gamma)\, w_{t-1}^i + (1-\tau_w) & \text{with probability } p \end{cases}$$

which implies the following result, as can be easily established through the same techniques used before.

Proposition 4 *Suppose A.1 holds and* $0 < \tau_w < 1 - (p\rho A_0)^{-1} e_0$. *Then, if* $\gamma > (1-\tau_w)2\beta$, *polarization/ pulverization emerges.*

In this case, government intervention proves effective (for the worse) in modifying polarization/ pulverization as it is capable of affecting the slope of the maps of the IFS (16) - and so of the IFS (18) - rather than additive constants (see Section 6.3).

As a result of the last proposition, a high enough wealth tax rate can generate a fractal wealth distribution even if $\gamma < 2\beta$, that is, even if growth and altruism are such that that the private sector let alone does not generate social disconnection.

Notice that here, to isolate the pure effect of taxation, we have not assumed any transfer from the government. However, Proposition 3 states that any lump sum transfer would not affect the results of the previous proposition, for fractalization only depends on the slope of the maps f_1 and f_2 in system (14), and not on their additive constants. Thus, we can say that, somewhat paradoxically, in this model the middle class may disappear and the economy become polarized/ pulverized as a result of an active redistributive policy.

3.4 Random Taxation

We here show that a redistribution scheme based on random taxation may reduce and, in some cases, even eliminate polarization. The idea is to increase the uncertainty in the model so that the two-maps IFS (21) is replaced by a three-maps IFS in which the image set of the second map might fill the hole left by the other two images set in case of polarization.

Let us assume that the gains from success are taxed at some rate $0 < \tau < 1$ with probability $1-q$, with $0 < q < 1$. At each period, the successful individuals face a tax lottery such that they have to pay τA_t with probability $1-q$ and 0 with probability q. Probability q is constant through time and is independent of the probability of success p. The government controls parameters q and τ. The total amount of proceeds are redistributed lump-sum to the unluckies.

If all individuals exert effort e_t in order to learn technology A_t, the steady state proportion of rich families in the economy will still be p. A fraction q of this proportion will be tax exempt, while the other fraction $1-q$ will be taxed at rate τ. Hence, the government in the long run will be able to collect tax revenues equals to

$$p\,(1-q)\,\tau A_t,$$

which - assuming a balanced government budget every period - equals the aggregate lump sum transfer received at the end of period t by the whole poor.

The dynamics of individual i wealth becomes:

$$W_t^i = \begin{cases} \beta W_{t-1}^i + \dfrac{p(1-q)}{1-p}\tau A_t & \text{with probability } 1-p \\ \beta W_{t-1}^i + (1-\tau) A_t & \text{with probability } p(1-q) \\ \beta W_{t-1}^i + A_t & \text{with probability } pq, \end{cases}$$

where, in the first line, $p(1-p)^{-1}(1-q)\tau A_t$ represents the transfer received by a single unlucky family. By dividing all equations by A_t we get the productivity-adjusted dynamic

$$w_t^i = \begin{cases} f_1\left(w_{t-1}^i\right) = \alpha w_{t-1}^i + \dfrac{p(1-q)}{1-p}\tau & \text{w. pr. } 1-p \\ f_2\left(w_{t-1}^i\right) = \alpha w_{t-1}^i + (1-\tau) & \text{w. pr. } p(1-q) \\ f_3\left(w_{t-1}^i\right) = \alpha w_{t-1}^i + 1 & \text{w. pr. } pq, \end{cases} \tag{23}$$

where $\alpha = \beta/\gamma$.

Note that system (23) contains three (affine) contractive maps identified by the parameters α, p, q and τ, where the last two are decision variables for the government. We want to investigate for what values of these parameters 1) incentive compatibility holds, that is, all individuals exert effort e_t, 2) the three maps are ordered so that $f_1 < f_2 < f_3$, and 3) whether values of the parameters exist so that the image set of f_2 fills the (possible) gap left by the image sets of f_1 and f_3. The last point would mean the possibility of eliminating possible polarization through government redistribution under this random scheme.

With no loss of generality for the rest of this section we shall assume

$$\frac{1}{3} \leq \alpha < \frac{1}{2},$$

which implies that the two maps IFS (3) exhibits polarization (the images of the two normalized maps do not overlap), and by adding one third affine map with slope α exactly between the two given maps the hole left by the two images sets may be "filled". From the first plot in figure 3, it is easily understood that maps with slope $\alpha < 1/3$ have images sets which cannot fill the whole interval $[0, 1]$. Clearly, for maps with $\alpha < 1/3$, arguments similar to the one carried out in this section can be implemented for random taxation schemes that use more tax rates. For example, if $\alpha < 1/n$, $n-1$ tax rates, each with positive probability, are necessary.

In order to let all individuals keep putting effort e_t even under taxation, an upper bound on the tax rate τ similar to that in Assumption A.2 is needed. An argument similar to that in Section 3.3, with the certain tax rate suitable replaced by the expected rate tax $(1-q)\tau$, leads to the following inequality:

$$\tau < \frac{1-p}{1-q}\left(1 - \frac{e_0}{p\rho A_0}\right),$$

where $\rho = (1-\beta)^{1-\beta}\beta^{\beta}$. Moreover, in order to have $f_1 < f_2$, it is immediately seen that

$$\tau_2 < \frac{1-p}{1-pq}$$

must hold; while $f_2 < f_3$ follows from $0 < \tau < 1$. Hence, the following assumption is what we need.

A. 3 *Assumption A.1 holds and*

$$0<\tau<\min\left\{\frac{1-p}{1-q}\left(1-\frac{e_0}{p\rho A_0}\right)\;,\;\frac{1-p}{1-pq}\right\},$$

where $\rho=(1-\beta)^{1-\beta}\beta^{\beta}$.

To analyze the possibility of eliminating polarization, let us normalize system (23) to the interval $[0,1]$. We shall apply formula (44) of Section 6.1 using as fixed points a and b the fixed points of the maps f_1 and f_3, that is,

$$a=\frac{p(1-q)\tau}{(1-p)(1-\alpha)},\qquad b=\frac{1}{1-\alpha},$$

and therefore

$$k=b-a=\frac{(1-p)-p(1-q)\tau}{(1-p)(1-\alpha)}.$$

By construction, we get the normalized system

$$w_t^i=\begin{cases} g_1\left(w_{t-1}^i\right)=\alpha w_{t-1}^i & \text{w. prob. } 1-p\\ g_2\left(w_{t-1}^i\right)=\alpha w_{t-1}^i+(1-\eta)(1-\alpha) & \text{w. prob. } p(1-q)\\ g_3\left(w_{t-1}^i\right)=\alpha w_{t-1}^i+(1-\alpha) & \text{w. prob. } pq,\end{cases}\tag{24}$$

where

$$\eta=\frac{(1-p)\tau}{(1-p)-p(1-q)\tau}.$$

Note that, under A.3, $0<\eta<1$.

The overlapping condition for the three image sets is a straightforward computation that leads to

$$1-2\alpha\leq(1-\alpha)\eta\leq\alpha,$$

which, in terms of τ, boils down to

$$\frac{(1-2\alpha)(1-p)}{(1-p)(1-\alpha)+(1-2\alpha)p(1-q)}\leq\tau\leq\frac{\alpha(1-p)}{(1-p)(1-\alpha)+\alpha p(1-q)}.\tag{25}$$

Note that condition (25) is nonempty for $1/3\leq\alpha<1/2$, and coincides with a single value for τ when $\alpha=1/3$, that is when inequalities in (25) become equalities and there is only one map g_2 in (24) with image set that can fill the hole left by the other two.

The left hand side of condition (25) is the most important in our analysis: it requires τ to be large enough in order to eliminate polarization. However, in view of Assumption A.3, we observe that τ must be not too large to let the incentive compatibility be satisfied:

$$\tau\leq\frac{1-p}{1-q}\left(1-\frac{e_0}{p\rho A_0}\right)$$

must be fulfilled to let all individuals exert effort e_t. If this constraint is too tight, due, *e. g.*, to a high value of the ratio $e_0/p\rho A_0$, the left hand side in (25) might not hold, thus leaving the government with no room for applying redistributive policies against polarization[8].

[8] Note that the other component of assumption A.3 is always satisfied since

$$\frac{(1-2\alpha)(1-p)}{(1-p)(1-\alpha)+(1-2\alpha)p(1-q)}<\frac{1-p}{1-pq}$$

is always true.

Specifically, polarization is neglected if τ is chosen to be equal to the left hand side of (25) and

$$\frac{e_0}{p\rho A_0} < 1 - \frac{(1-2\alpha)(1-q)}{(1-p)(1-\alpha)+(1-2\alpha)p(1-q)} \tag{26}$$

holds.

Note that we did not discuss any restrictions for the choice of parameter q by the government so far. Since by Assumption A.1 $e_0/p\rho A_0 < 1$, there always exist values for parameter $q < 1$, possibly close to 1, such that (26) is satisfied. In other words, there is always room for the government to eliminate polarization through a random taxation and lump-sum redistribution scheme in the sense of making the support of the steady state distribution of system (24) to be the whole interval $[0,1]$. However, values of q close to 1 imply that almost the whole p fraction of the steady state successful population - that is, pq out of p - is paying no taxes, while only a negligible fraction $p(1-q)$ of the successful population - which is the "middle class" artificially created through the random taxation - is paying taxes. This means that, from the probabilistic point of view, polarization remains substantially unaltered, since the rich individuals are almost of the same size as before and the new middle class carries nearly no weight. That is, once again, a tight incentive compatibility constraint in Assumption A.3 leaves little room for government intervention and substantially reduces hopes of eliminating polarization even through random taxation.

4 Patents and Firms

In this section we show that the main results of the previous section are easily applied to slightly different models, and do not hinge on the assumption that unlucky people should earn zero income, unless sustained by redistributive policies. We will show that even if they can reap the benefits of technological advance with a one period lag, a strong enough growth rate is able to generate a fractal society.

While keeping the same framework of Section 3, let us now assume that every individual of generation t at the beginning of her economic life has the same probability $0 < p < 1$ of discovering a better production method that allows the productivity of a number $\theta \geq 1$ of individuals, including herself, to jump to the new technological frontier $A_t = \gamma A_{t-1}$, provided she undertook an indivisible innovation effort $e_t = \gamma e_{t-1}$.

To render growth endogenous we will assume that productivity growth rate γ is an increasing and bounded[9] function of the aggregate innovative effort $\int_0^L e_0^i di$, where L denotes the constant population size[10].

Inventions are immediately patented and the patents expire after a generation. We will assume that each individual can run only one research project during her life. Hence we are building a simple Schumpeterian model in which the entrepreneurs are new people (Schumpeter [42] and [43]) who try to adapt the ever-evolving society knowledge frontier (as in Aghion and Howitt [2], Ch. 3, and Howitt [29]) to their sphere of production. The parallel with Aghion and Howitt ([2], Ch. 3) and Howitt [29] cross-sector spillover is in our assumption that A_t evolves as an increasing function of social R&D adoption effort. This adds a zero growth equilibrium due to R&D coordination failure: if each individual expects nobody to exert effort

[9] With this simple assumption - that may be motivated by some kinds of congestion effects - we eliminate Jones [31] scale effects.

[10] It would not be difficult to allow for population growth. Interestingly, offspring's division of bequest would re-inforce fractalization effects in this model and/or even generate it.

she will be better off not exerting it. In the rest of the analysis we will concentrate only on the positive growth equilibrium.

Unlike usual Schumpeterian models we are here assuming a limited productive capacity per-firms and/or a limited number of patent licensees. In fact we will assume that in order to implement each successful innovation the cooperation of θ workers (including the innovator) is necessary. Hence, by the law of large numbers, in the steady state there will be a fraction p of innovators, and a fraction $p\theta$ of individuals employed in all innovative productive processes. Since we keep the whole population normalized to the interval $[0,1]$, in order to let all innovators carry on their activity, the fraction $p\theta$ of employed individuals cannot exceed $L = 1$, that is, the number of workers for each activity must be bounded by

$$1 \leq \theta \leq \frac{1}{p}. \tag{27}$$

If the right hand side of (27) holds with equality, the society is perfectly divided in a fraction p of entrepreneurs/ innovators and a fraction $1-p$ of workers. If the right hand side of (27) holds with strict inequality, then there will be a fraction of people who will be treated as self-employed in production processes that use the technology A_{t-1} available from the last period. Since patents expire after one period, the technology A_{t-1}, available only for the innovators at time $t-1$, becomes of public domain at time t. Therefore we shall assume that, at each period t, both employed workers in the innovative sectors and self-employed workers in the old sectors perceive salaries equal to their productivity under the old technology A_{t-1}. In this last case there will be a fraction $0 < p\theta < 1$ of individuals employed in the A_t technology sector and a fraction $1-p\theta$ of individuals employed in the A_{t-1} technology sector. Of these families, only a fraction p is able to reap the benefits of the innovative technology A_t (each by employing $\theta - 1$ workers) by means of patents, while the other fraction $1-p$, being them employed in the innovative sector or self-employed in the old sector, is remunerated by the productivity of the A_{t-1} technology[11].

The innovations of this model can alternatively be interpreted as the discovery of an "entrepreneurial talent" that allows the innovator to found a firm that allows a more efficient use of θ workers by making them use the best productive practices available in her firm. In this sense, the model of this section can be viewed as an education model of the firm: in the particular case $\theta = 1$ the individual is only able to privately accumulate the "state of the art" human capital. Unlike the previous example, the technology learned by generation t will be observed by everybody when it is operated, and, afterwards, every family will become able to use it at no additional educational cost. With $\theta = 1$ this model depicts an economy similar to that of the previous example, except for a perfect educational spillover which allows the wealth of the children of the unlucky generation to instantaneously reach the level of the lucky members of the previous cohort.

4.1 Fractalization of Schumpeterian Growth

To summarize, in every period p "innovators" will appear and $p\theta \leq 1$ skilled workers will be producing with the cutting-edge technology, paying their extra productivity to each successful innovator. The innovator - as a patent holder or as an entrepreneur - is able to extract the complete productivity increment for one period, thereby rendering the appropriable technology

[11] If $\theta > 1/p$, the innovators would not be able to implement their discoveries, and in a competitive equilibrium all profits would be zero, leading to a society with a unique wealth group and no polarization.

of every non-innovator equal to the same value A_{t-1}. In other words, each single innovator in period t can appropriate the productivity gains of the non-innovators, equal to

$$(A_t - A_{t-1})(\theta - 1) = (\gamma - 1)(\theta - 1) A_{t-1}. \tag{28}$$

Hence, the wealth of family i at the end of period t, provided she undertook the indivisible innovation effort e_t at the beginning of the period, will be

$$W_t^i = \begin{cases} \beta W_{t-1}^i + A_{t-1} & \text{with probability } 1-p \\ \beta W_{t-1}^i + A_t + (\gamma - 1)(\theta - 1) A_{t-1} & \text{with probability } p. \end{cases} \tag{29}$$

Notice that the successful innovator, through the patent, benefits from both the adoption of the new technology A_t and the productivity gain of the $\theta - 1$ non-innovators workers employed in her firm; that is, each non-innovator employed in her firm has to pay the full monopolistic rent to the patent holder, though she can choose between alternative patent holders. On the other side, the unlucky will get only the one-period lagged productivity A_{t-1} wealth, being her employed by some patent holder firm or self-employed. Notice that we are focussing on perfectly symmetric equilibria.

Once again, we need to make sure that all families find it convenient to undertake the indivisible innovation effort e_t at the beginning of each period t. The individual i expected utility gain conditional on effort e_t is given by

$$\begin{aligned} \mathbb{E}\left[U\left(Y_t^i\right)|e_t\right] &= p\rho\left[A_t + (\gamma - 1)(\theta - 1) A_{t-1}\right] + (1-p)\rho A_{t-1} - e_t \\ &= \rho\left[1 + p\theta(\gamma - 1)\right] A_{t-1} - e_t, \end{aligned}$$

where $\rho = (1-\beta)^{1-\beta}\beta^{\beta}$, while the individual i certain utility gain obtained by exerting zero effort is given by

$$U(A_{t-1}) = \rho A_{t-1}.$$

To achieve our goal,

$$\mathbb{E}\left[U\left(Y_t^i\right)|e_t\right] > U(A_{t-1})$$

must hold, which easily translates into the next assumption.

A. 4

$$0 < e_0 < \rho\theta(1 - 1/\gamma) p A_0.$$

where $\rho = (1-\beta)^{1-\beta}\beta^{\beta}$.

A result similar to the previous ones follows.

Proposition 5 *Assume condition (27) and assumption A.4 to hold. Then, if* $1 < \gamma \leq 2\beta$, *there exists a middle class. Whenever* $\gamma > 2\beta$ *the economy becomes polarized/ pulverized in the long run, and the support* A *of its limit distribution* μ^* *is a Cantor-like set. Moreover, the larger* γ *(and the smaller* β*), the larger the gap between the fractions of the population near the extremes of the support, independently of the values of parameters* p *and* θ*.*

Proof. Following the usual technique, divide both equations in (29) by A_t to get the productivity-adjusted dynamic

$$w_t^i = \begin{cases} (\beta/\gamma)\, w_{t-1}^i + 1/\gamma & \text{with probability } 1-p \\ (\beta/\gamma)\, w_{t-1}^i + 1/\gamma + (1 - 1/\gamma)\,\theta & \text{with probability } p. \end{cases} \tag{30}$$

Clearly, under A.4, Proposition 1 holds, and thus system (30) is similar to system (18). By Proposition 2, polarization/ pulverization is completely determined by condition $\gamma > 2\beta$. ■

Notice that in this case nobody ends up with a zero wealth, but instead even the "poorest" segment of the population improves its standards of living at the same steady rate $\gamma - 1$ as the richest. In particular, the true support of the marginal distribution of system (29) at time t lay in some subset of the interval

$$\left[\left(\frac{1-(\beta/\gamma)^{t+1}}{\gamma-\beta}\right)\gamma^t A_0\ ,\ \left(\frac{1-(\beta/\gamma)^{t+1}}{\gamma-\beta}\right)\gamma^t\left[1+(\gamma-1)\,\theta\right]A_0\right].$$

This is a consequence of the temporary nature of patents that allows the inventors to "exploit" the unlucky only for a limited lapse of time and, upon expiry, makes that innovation available for everybody to be freely used.

Also notice that here larger population size implies faster growth (higher γ): therefore a more populated world or a more globalized economic space can generate or enhance fractalization.

In light of Section 3.3, it would be natural to believe that a policy that "cures" polarization/ pulverization is possible, provided the effort e_t required to promote innovation is sufficiently small. The government could, in fact, reward the innovator and at the same time make the innovation immediately publicly available to everybody. We show that again, also in an economy with patented discoveries, government intervention proves ineffective in modifying possible polarization.

4.2 Government Purchase of Innovations

Under the normalization that the population is indexed in $[0, 1]$, the society as a whole will put effort e_t in the R&D for new technological projects and at the steady state there will be a fraction p of successful innovators who possess technology A_t. Suppose that the government, in order to make technology A_t publicly available in period t, buys the technological know-how from the p fraction of innovators at the lowest incentive compatible price (i.e. at[12] $p^{-1}e_t$) and allows the fraction $1-p$ of unluckies to freely use it in their own firms. At first sight one could expect such policy to eliminate wealth fractalization.

Assume that the government charge all the unluckies the whole cost e_t of research through a lump-sum tax to be fully transferred to the luckies. The law of motion of wealth becomes:

$$W_t^i = \begin{cases} \beta W_{t-1}^i + A_t - (1-p)^{-1}e_t & \text{with probability } 1-p \\ \beta W_{t-1}^i + A_t + p^{-1}e_t & \text{with probability } p, \end{cases} \tag{31}$$

where $(1-p)^{-1}e_t$ denotes the per capita cost of research charged to the unluckies and $p^{-1}e_t$ denotes the per capita compensation for the productivity gain loss (28). We will assume e_0 small enough to guarantee that the unluckies are better off under this forced purchase of the new technology than under *laissez faire*.

Observe that, at least for the case $p < 1/2$, which seems sufficiently realistic, system (31) can once again be reduced to system (18) by Proposition 1. Therefore Proposition 2 applies stating that polarization/ pulverization is completely determined by condition $\gamma > 2\beta$, thus

[12] Note that any price slightly higher than $p^{-1}e_t$ makes each individual strictly better off undertaking the R&D effort.

obtaining a result similar to Proposition 5: government financing private innovations does not affect polarization/ pulverization.

The delicate part, as usual, is enforceability of such a policy: nobody would vote a government who leaves everybody worse off under this program. The individual expected indirect utility gain is

$$\begin{aligned}\mathbb{E}\left[U\left(Y_t^i\right)\right] &= \rho\left[p\left(A_t+\frac{e_t}{p}\right)+(1-p)\left(A_t-\frac{e_t}{1-p}\right)\right]-e_t\\ &= \rho p A_t - e_t,\end{aligned}$$

and thus, the effort condition turns out to be the same as[13] in Assumption A.1:

$$e_0 < \rho p A_0. \tag{32}$$

Note that, by assuming $p < 1/2$, necessarily $\rho p < 1-p$, and thus (32) implies $e_0 < (1-p)\,A_0$, which guarantees that the infimum of the support of the marginal probabilities of process (31), which at time t is

$$\left(\frac{1-(\beta/\gamma)^{t+1}}{\gamma-\beta}\right)\gamma^{t+1}\left(A_0-\frac{e_0}{1-p}\right),$$

is strictly positive for all t. This means that, like in Proposition 5, the "poorest" segment of the population improves its standards of living at the same steady rate $\gamma-1$ as the richest.

5 Polarization Versus Pulverization

So far we have used the term polarization to generically describe the disappearance of the middle class in a distribution supported on a Cantor set. In a recent work Esteban and Ray [23] provided a more refined concept of polarization and constructed a measure of polarization based on an axiomatic approach. Their measure of polarization compares the homogeneity of a group with the overall heterogeneity of a population. If the distribution of wealth is highly gathered within groups but very diverse between groups in a population, then wealth is considered "polarized" between the groups. The fundamental difference of such measure with standard measures of inequality, such as the Gini coefficient or any Lorenz consistent index, is that the latter fails to capture the "sense of identity" within members of the same group, while the former includes a sensitivity parameter which takes it into account. The idea behind adding some identification function into a inequality measure is that polarization increases not only as inter-group heterogeneity raises, but also when the number of members within a group raises. In other words, the generation of tensions possibly evolving to rebellion, revolt, or social unrest is more likely if wealth is distributed among groups which have a strong self-identity feeling.

Formally, given a finite distribution of weights $\pi_1,\ldots,\pi_n$ on wealths[14] $W_1,\ldots,W_n$, with $\pi_i, W_i > 0$, the polarization measure is given by

$$P_n = K\sum_{i=1}^{n}\sum_{j=1}^{n}\pi_i^{1+\lambda}\pi_j\left|W_i-W_j\right|, \tag{33}$$

[13] This seems to be reasonable since utility is linear and what is taken from the unluckies goes to the luckies, leaving the expected utility gain unchanged.

[14] To be precise, the authors interpret the W_i's as the natural logarithm of wealths. This is based on the assumption that only percentage differences matter. We shall take into account log transformations of wealths at the end of this section.

where $K > 0$ is a constant used for population normalization and $0 \leq \lambda \leq 1.6$ indicates the degree of self-identity within each group i - that is, the degree of polarization sensitivity - by raising to a non-negative power the frequency associated to wealth W_i. Note that for $\lambda = 0$ P_n boils down to the Gini coefficient.

The measure defined in (33) has been constructed for statistical purposes; it is meant to measure polarization by using statistical data available for societies with finite populations. The pursue of some generalization of the measure (33) to include infinite distributions supported over fractal sets is well beyond the scope of this paper. In this section, however, we try to estimate the limit of P_n as $n \to \infty$ in some cases. Of course this is a very coarse measure of polarization for infinite distributions, and it is certainly not based on any axiomatic justification.

Our goal is actually more modest: we plan to use such a rough measure of polarization to investigate how what we called "pulverization" - that is, the replica of the "polarization" on a smaller scale within each sub-interval generated after each iteration of the IFS (14) when $2\beta < \gamma$ - affects polarization in the long run. The question is non trivial in view of formula (33), since it is not clear whether the process of removing a middle interval from each sub-interval after iterations of the Markov operator M will eventually reduce or increase the striking polarization phenomena occurring after the first iteration of the IFS (see, *e.g.*, the first plot in figure 4). Two opposite effects occur by applying formula (33) directly to our IFS after each step: on one hand new wealth groups W_i are born and, at least in the original model not normalized over the interval $[0, 1]$, the distances between wealth clusters increase, thus raising the number of terms in the sum and also raising the values $|W_i - W_j|$; on the other hands, the weights π_i decrease after each step, since, under our assumptions, the same constant population is progressively fractionalized over more and more wealth clusters. Moreover, since in our model the population is normalized on $[0, 1]$, the weights π_i are always less than 1; this further reduces the "polarization effect" due to the exponent $\lambda \geq 0$, which has been introduced to enhance the importance of large groups, but, when $\pi_i < 1$, has the opposite effect that the Gini inequality index turns out to be larger than the polarization measure. This seems to be coherent with the theoretical goals of formula (33): small groups have low self- identity, thus reducing polarization. In our specific model, all the weights π_i become eventually negligible on the support of the invariant distribution, which is nonatomic indeed.

All these considerations raise the question of whether our choice of describing polarization and pulverization as if they were different aspects of the same phenomenon so far, might turn out to be contradictory, as they may well be working one against the other. In other words, the aim of the present section is to figure out whether the growth of the terms $|W_i - W_j|$, both in number and in size, dominates the shrinking of terms $\pi_i^{1+\lambda}\pi_j$ in formula (33), as it is being applied to the finite distributions of wealth obtained after each iteration of the dynamic (14). This may shed some light on the polarization properties of the limit distribution of wealths.

This study, far from being complete, essentially shall provide some partial results for some values of the parameters. This is however enough to envisage a variety of different polarization features of the limit distribution. For values of the sensitivity parameter λ close to 0, economies have infinite polarization in the long run even for low values of the growth rate γ. However, for $\lambda > 0$, we are able to show that polarization may still be infinite, or at least positive, in the limit with growth rates γ high enough; while, if γ is small, polarization vanishes as time elapses. Finally, by following more closely Esteban and Ray [23] construction, we consider logarithmic transformations of wealths instead of wealths themselves. In this last case, we are able to show that, at least for $p = 1/2$, polarization becomes negligible for all $\gamma > 1$ and all $\lambda > 0$. This partial analysis is enough to convince the reader that pulverization does not necessarily conflict with polarization in the long run. There is enough room for (feasible) values of the parameters such that pulverization allows for increasing polarization through time.

Consider the dynamic system (14) discussed in Section 3:

$$W_t = \begin{cases} \beta W_{t-1} & \text{with probability } 1-p \\ \beta W_{t-1} + A_t & \text{with probability } p, \end{cases} \tag{34}$$

where W_t denotes some wealth amount at time t, $0 < \beta < 1$ is the degree of intergenerational altruism, $A_t = \gamma^t A_0$ is the exogenous technology with $A_0 > 0$, $\gamma > 1$, and $0 < p < 1$ represents the probability of "success" in the adoption of the technology. Since we are studying polarization, in the sequel we shall assume

$$2\beta < \gamma \tag{35}$$

as prescribed by Proposition 2.

The choice of studying system (34) instead of system (18) normalized on the interval $[0, 1]$ is made to conform ourselves with the mainstream literature, where true wealth values available from statistical data are used, instead of productivity adjusted values.

Theorem 1 cannot be applied directly to the IFS (34), which has unbounded support for $t \to \infty$, however we can refer to the invariant distribution of the conjugate system (18) as the equivalent of the unique invariant distribution of (34) defined on the positive real line[15]. The system converges to this distribution starting from any initial distribution of wealths. Thus, for convenience, we assume that the distribution at time $t = 0$ concentrates a mass $1 - p$ on some bequest $b_0 \geq 0$ inherited from the past and a mass p on $b_0 + A_0$; that is, $\mu_0 = (1-p)\,\delta_{b_0} + p\delta_{b_0+A_0}$, where δ_W is the Dirac function. We may also write

$$W_0 = \begin{cases} b_0 & \text{with probability } 1-p \\ b_0 + A_0 & \text{with probability } p. \end{cases} \tag{36}$$

Having an initial distribution concentrating masses over a finite set of points implies that also the distribution of wealths at each date $t > 0$ concentrates masses over finite sets of points. This allows a direct application of formula (33) to the distribution of wealths at each date t. By construction, it is easily seen that, for all $t \geq 0$, the are 2^{t+1} values of wealth $W_t^1, \ldots, W_t^{2^{t+1}}$ each with weight π_t^i, which can be interpreted as the frequency of wealth W_t^i, $i = 1, \ldots, 2^{t+1}$. Therefore, polarization at time t is given by

$$P_t = \sum_{i=1}^{2^{t+1}} \sum_{j=1}^{2^{t+1}} \left(\pi_t^i\right)^{1+\lambda} \pi_t^j \left|W_t^i - W_t^j\right|, \tag{37}$$

where the constant K in (33) has been set equal to 1, as population is already normalized in our model.

Since, by independence, for all $t \geq 0$, weights π_t^i have the form

$$\pi_t^i = p^{h_i} (1-p)^{t+1-h_i}, \qquad 0 \leq h_i \leq t+1, \qquad 0 < p < 1,$$

[15]Alternatively, since $0 < \beta < 1$, one may invoke Theorem 7.2 in Lasota [34] to prove existence and uniqueness of the invariant distribution for IFS (34).

clearly $\lim_{t\to\infty}\pi_t^i = 0$, and thus also $\lim_{t\to\infty}(\pi_t^i)^{1+\lambda}\pi_t^j = 0$. In other words, masses p and $1-p$ initially concentrated on b_0 and $b_0 + A_0$, are progressively spread over a set of points that eventually converge to the support of the invariant distribution, which is a continuum of points, thus vanishing in the limit.

Since both wealths W_t^i and weights π_t^i have a recursive formulation generated by dynamic (34), it is convenient to write formula (37) in a form more suitable for direct handling.

Lemma 1 *For each $t \geq 0$, label the set of wealths so that they are ordered: $W_t^1 < W_t^2 < \cdots < W_t^{2^{t+1}}$. Then formula (37) can be rewritten as follows:*

$$P_t = \sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} \pi_t^i \pi_t^{i-j}\left[\left(\pi_t^i\right)^\lambda + \left(\pi_t^{i-j}\right)^\lambda\right]\left(W_t^i - W_t^{i-j}\right) \tag{38}$$

Proof. Note that in the sum (37) each difference $\left|W_t^i - W_t^j\right|$ is counted twice: once with coefficient $\left(\pi_t^i\right)^{1+\lambda}\pi_t^j$ and once with coefficient $\pi_t^i\left(\pi_t^j\right)^{1+\lambda}$. By counting all possible distances $W_t^i - W_t^j$, with $W_t^i > W_t^j$, starting from those between contiguous points $W_t^i - W_t^{i-j}$ for $j = 1$ and $2 \leq i \leq 2^{t+1}$, then adding those between points more apart, for $2 \leq j \leq 2^{t+1}-1$ and $1 + j \leq i \leq 2^{t+1}$, we get

$$P_t = \sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} \left(\pi_t^i\right)^{1+\lambda}\pi_t^{i-j}\left(W_t^i - W_t^{i-j}\right) + \sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} \pi_t^i\left(\pi_t^{i-j}\right)^{1+\lambda}\left(W_t^i - W_t^{i-j}\right),$$

which is (38). ■

Formula (38) becomes especially appealing when $\lambda = 0$, since it allows for a recursive formulation of polarization P_t as t increases, as the following lemma establishes.

Lemma 2 *Let $2\beta < \gamma$. Then, if $\lambda = 0$, polarization P_t has the following recursive formulation:*

$$P_{t+1} = \left[1 - 2p(1-p)\right]\beta P_t + 2p(1-p)A_{t+1}, \qquad t \geq 0. \tag{39}$$

Proof. First note that for $\lambda = 0$ (38) in Lemma 1 boils down to

$$P_t = 2\sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} \pi_t^i\pi_t^{i-j}\left(W_t^i - W_t^{i-j}\right). \tag{40}$$

Our plan is to write P_{t+1} as a function of P_t. Recall that at time t, there are 2^{t+1} wealth points that we order as follows: $W_t^1 < W_t^2 < \cdots < W_t^{2^{t+1}}$. By following the dynamic (34), each point W_t^i, with associated weight π_t^i, at time t is being split into the following two points at time $t+1$: βW_t^i with mass $(1-p)\pi_t^i$ and $\beta W_t^i + A_{t+1}$ with mass $p\pi_t^i$. Hence, it is possible to write all 2^{t+2} wealth points and their associated mass at time $t+1$ in terms of wealths W_t^i

and weights π_t^i at time t. By using (40) for $t+1$ we then obtain

$$\begin{aligned} P_{t+1} &= 2\sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} (1-p)^2 \pi_t^i \pi_t^{i-j} \beta \left(W_t^i - W_t^{i-j}\right) \\ &+2\sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} p^2 \pi_t^i \pi_t^{i-j} \beta \left(W_t^i - W_t^{i-j}\right) \\ &+2\sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} p(1-p) \pi_t^i \pi_t^{i-j} \left(\beta W_t^{i-j} - \beta W_t^i + A_{t+1}\right) \\ &+2\sum_{j=1}^{2^{t+1}-1}\sum_{i=1+j}^{2^{t+1}} p(1-p) \pi_t^i \pi_t^{i-j} \left(\beta W_t^{i} - \beta W_t^{i-j} + A_{t+1}\right) \\ &+2\sum_{k=1}^{2^{t+1}} p(1-p) \left(\pi_t^k\right)^2 A_{t+1}, \end{aligned}$$

where in gathering the homogeneous terms in each of the 5 sums we have exploited the fact that, under the polarization assumption (35) $2\beta < \gamma$, $\beta W_t^{2^{t+1}} < \beta W_t^1 + A_{t+1}$. After some algebra and by plugging (40) into the above expression we readily get equality (39). ■

Lemma 2 is the key in the proof of the next result which provides a very striking characterization of the limit behavior of the pure inequality measure, that is when $\lambda = 0$.

Proposition 6 *If $\lambda = 0$, $P_t < P_{t+1}$ for all $t \geq 0$ and $\lim_{t\to\infty} P_t = \infty$ for all feasible values of parameters β, p, γ, b_0 and A_0 such that the polarization condition $2\beta < \gamma$ is satisfied.*

Proof. Since $P_0 = 2p(1-p)A_0 > 0$, system (39) in Lemma 2 has solution

$$P_t = 2p(1-p)A_0 \frac{1-(s/\gamma)^{t+1}}{\gamma - s}\gamma^{t+1}, \qquad t \geq 0,$$

where $s = [1-2p(1-p)]\beta$. Since $0 < s < 1$ and $\gamma > 1$, the result is established. ■

Proposition 6 states that pure inequality grows as time elapses in a polarized society[16], where polarization is characterized by Proposition 2. This is a non-trivial result since weights π_t^i vanish as $t \to \infty$ already when $\lambda = 0$. Increasing inequality means that the distances $\left|W_t^i - W_t^j\right|$ rise -both in number and in size - more rapidly than the speed at which weights π_t^i converge to 0.

Next result is aimed at investigating how the result in Proposition 6 is being modified as we let sensitivity parameter λ having positive values. Since positive values of λ further reduce the weights $\left(\pi_t^i\right)^{1+\lambda}\pi_t^j$ associated to each distance $\left|W_t^i - W_t^j\right|$ in the sum (37), we expect that some stricter condition on parameters, other than polarization condition $2\beta < \gamma$, might become necessary in order to have positive polarization in the long run. Since in this case formula (38) in Lemma 1 becomes much harder to handle, we shall focus on the special case $p = 1/2$. This example, nevertheless, confirms our intuition.

[16]In view of the result in the following Proposition 7, which is independent of polarization condition (35), our conjecture is that the Proposition 6 may hold also for non polarized economies, that is, also when $2\beta \geq \gamma$. In this case the construction in Lemma 2 becomes more complicated since the difference $\left(\beta W_t^{i-j} - \beta W_t^i + A_{t+1}\right)$ in the third term of the expression of P_{t+1} must be replaced with its absolute value. We did not investigate further this case.

Proposition 7 *Let $p = 1/2$ and $0 < \lambda \leq 1.6$. Then,*

$$\lim_{t\to\infty} P_t = \begin{cases} \infty & \text{if } \gamma > 2^\lambda \\ (2\gamma - 1)^{-1} A_0 & \text{if } \gamma = 2^\lambda \\ 0 & \text{if } \gamma < 2^\lambda \end{cases}$$

Proof. If $p = 1/2$, $\pi_t^i = \pi_t^{i-j} = (1/2)^{t+1}$ and

$$\pi_t^i \pi_t^{i-j} \left[\left(\pi_t^i\right)^\lambda + \left(\pi_t^{i-j}\right)^\lambda \right] = 2 \left(\frac{2^{-\lambda}}{4} \right)^{t+1}.$$

Then (38) in Lemma 1 becomes

$$P_t = 2 \left(\frac{2^{-\lambda}}{4} \right)^{t+1} \sum_{j=1}^{2^{t+1}-1} \sum_{i=1+j}^{2^{t+1}} \left(W_t^i - W_t^{i-j} \right), \qquad t \geq 0,$$

which boils down to

$$\begin{aligned} P_t &= 2 \left(\frac{2^{-\lambda}}{4} \right)^{t+1} \frac{(2\gamma/\beta)^{t+1} - 1}{2\gamma - 1} 2^t \beta^{t+1} A_0 \\ &= \frac{A_0}{2\gamma - 1} \left[\left(\frac{\gamma}{2^\lambda} \right)^{t+1} - \left(\frac{\beta}{2^{\lambda+1}} \right)^{t+1} \right], \end{aligned}$$

which, as $\gamma > 1$ and $0 < \beta < 1$, yields the result. ■

Proposition 7 establishes a direct link between polarization and growth rate of the economy: for all feasible values of sensitivity parameter $0 < \lambda \leq 1.6$, polarization may be strictly positive (even infinity) in the long run provided that the exogenous growth rate γ is sufficiently large. Moreover, for smaller values of λ, smaller values of γ are required for positive limit polarization, with nearly all feasible values $\gamma > 1$ as λ approaches 0. Simulations on computer support the conjecture that a generalization of Proposition 7 to cover all probability values $0 < p < 1$, providing some threshold function $f(p)$ above which polarization diverges and below which polarization vanishes, should hold. Note also that Proposition 7 holds independently of polarization condition $2\beta < \gamma$, as we do not use the construction of Lemma 2 in its proof. In other words, if λ is small enough, the limit polarization might diverge even when $2\beta > \gamma$.

Finally we turn to a slight modification of the polarization measure in (37) by considering natural logarithms of wealths in place of wealths themselves. This is based on Esteban and Ray [23] presumption that only percentage differences matter. Polarization at time t in our model then is given by

$$P_t^l = \sum_{i=1}^{2^{t+1}} \sum_{j=1}^{2^{t+1}} \left(\pi_t^i\right)^{1+\lambda} \pi_t^j \left| \ln W_t^i - \ln W_t^j \right|, \tag{41}$$

where W_t^i and W_t^j are generated by system (34). To let the expression in (41) be defined for all $t \geq 0$, we need to assume $b_0 > 0$.

Since logarithms grow much slower than wealths themselves, in view of the previous results that link positive limit polarization to high growth rates γ, we expect that the limit in (41) as $t \to \infty$ will be zero in most cases. This intuition is supported by the following partial result, which states that, at least when $p = 1/2$, polarization in the long run must be generically zero.

Proposition 8 *Let $p = 1/2$. Then, for all $0 < \lambda \leq 1.6$ $\lim_{t \to \infty} P_t^l = 0$.*

Proof. If $p = 1/2$ it is immediately seen that

$$\left(\pi_t^i\right)^{1+\lambda} \pi_t^j = \left[\left(\frac{1}{2}\right)^{t+1}\right]^{2+\lambda} = \left(\frac{1}{2}\right)^{2+\lambda} \left[\left(\frac{1}{2}\right)^{2+\lambda}\right]^t.$$

Therefore (41) becomes

$$P_t^l = \left(\frac{1}{2}\right)^{2+\lambda} \left[\left(\frac{1}{2}\right)^{2+\lambda}\right]^t \sum_{i=1}^{2^{t+1}} \sum_{j=1}^{2^{t+1}} \left|\ln W_t^i - \ln W_t^j\right|.$$

■

Proof. If $W_t^i > W_t^j$, by construction,

$$\ln W_t^i - \ln W_t^j \leq \ln \left[\beta^t b_0 + \frac{1 - (\beta/\gamma)^{t+1}}{\gamma - \beta} \gamma^{t+1} A_0\right] - \ln \left(\beta^t b_0\right),$$

and thus, at each time $t \geq 0$ and for all $1 \leq i, j \leq 2^{t+1}$,

$$\left|\ln W_t^i - \ln W_t^j\right| < \ln \left[1 + \frac{\gamma A_0}{(\gamma - \beta) b_0} \left(\frac{\gamma}{\beta}\right)^t\right].$$

We then have the following upper bound for P_t^l:

$$\begin{aligned} P_t^l &\leq \left(\frac{1}{2}\right)^{\lambda} \left[4 \left(\frac{1}{2}\right)^{2+\lambda}\right]^t \ln \left[1 + \frac{\gamma A_0}{(\gamma - \beta) b_0} \left(\frac{\gamma}{\beta}\right)^t\right] \\ &= \left(\frac{1}{2}\right)^{\lambda} \left[\left(\frac{1}{2}\right)^{\lambda}\right]^t \ln \left[1 + \frac{\gamma A_0}{(\gamma - \beta) b_0} \left(\frac{\gamma}{\beta}\right)^t\right]. \end{aligned}$$

The right hand side of the last inequality converges to zero for all $\lambda > 0$, and the proof is complete. ■

Proposition 8 provides only a sufficient condition for zero limit polarization. However, we have not been able to find any example of positive polarization in the long run with logarithms of wealths. Furthermore, computer simulations seem to bear out the conjecture that Proposition 8 might hold for all $0 < p < 1$ and for $\lambda = 0$ as well.

6 Non-Linear Iterated Function Systems

In the remaining sections we briefly sketch some possible generalizations of the theory discussed in Section 2 to non-linear IFS and to IFS with state-dependent probabilities. This is to show how widespread is the potential of these tools in economic dynamics, especially when polarization/pulverization phenomena are the main focus.

6.1 Scaling Maps

Consider a pair of continuous, strictly increasing, contractive maps f_1, f_2 so that $f_1 < f_2$, and let a and b be their fixed points respectively, that is, $f_1(a) = a$ and $f_2(b) = b$, as in figure 6. Since Theorem 1 holds for any contractive IFS (see, *e.g.*, [30], [24] or [33]), given any fixed probability p, $0 < p < 1$, the system

$$x_t = \begin{cases} f_1(x_{t-1}) & \text{with probability } 1-p \\ f_2(x_{t-1}) & \text{with probability } p \end{cases} \tag{42}$$

converges weakly to some unique invariant probability concentrated over a support which is some compact subset of the interval $[a, b]$. In other words, the portion of the maps f_1, f_2 which is relevant in the long run is included in the square T in figure 6 (where the plots of f_1 and f_2 are bold).

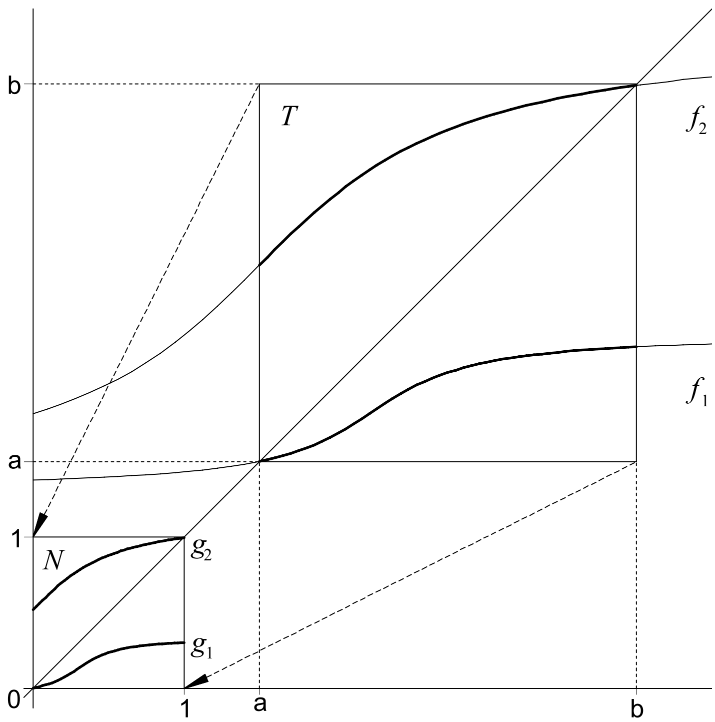


Figure 6: normalization of two contractive maps f_1, f_2 over the unit square.

For any contractive maps f_1, f_2, such relevant region can be "normalized" over the interval $[0, 1]$ (that is, the square T can be transformed into the square N in figure 6) by the following two transformations:

1. by a *rigid translation* towards the origin, so that the fixed point a becomes the origin itself, and

2. by *scaling* the whole system *by a factor* $k = b - a$.

The outcome of such transformation is a new IFS

$$x_t = \begin{cases} g_1(x_{t-1}) & \text{with probability } 1-p \\ g_2(x_{t-1}) & \text{with probability } p \end{cases} \tag{43}$$

where the maps g_j are given by

$$g_j(x) = k^{-1}[f_j(kx+a) - a], \qquad j = 1, 2, \tag{44}$$

with $k = b - a$. Figure 6 illustrates this translation/scaling procedure that transforms the original relevant region T into the new "normalized" relevant region N, which is the unit square.

Such normalization can be generalized to maps $f_1 < f_2$ that are not necessarily increasing and even not necessarily monotone[17]. The next Lemma provides a general technique for calculating the extrema of the "relevant" interval $[a, b]$ containing the trapping region of system (42) for any pair of contractive maps. Note that numbers a, b are needed to apply formula (44).

Lemma 3 *For $j = 1, 2$, let $f_j : [c, d] \to [c, d]$ (possibly with $c = -\infty$ and/or $d = +\infty$), so that $f_1(x) < f_2(x)$ and some constant $0 < \alpha_j < 1$ exists such that $|f_j(x) - f_j(y)| \leq \alpha_j |x - y|$ for all $x, y \in [c, d]$. Then, for any probability $0 < p < 1$, the relevant interval $[a, b] \subseteq [c, d]$ of the IFS $\{f_1, f_2, p\}$ has extrema satisfying*

i) $a = \min_{x\in[a,b]} f_1(x)$,

ii) $b = \max_{x\in[a,b]} f_2(x)$.

Proof. Denote by x_j the (unique) fixed points of the maps f_j, *i.e.* $x_j = f_j(x_j)$ for $j = 1, 2$. First note that $f_1 < f_2$ implies $x_1 < x_2$ and the relevant interval $[a, b]$ must contain the interval $[x_1, x_2]$. Moreover, since f_j are contractions, by starting from any point in $[x_1, x_2]$ the system cannot move too far from it; specifically, it is readily seen that the size of the relevant interval must satisfy

$$b - a \leq \frac{1+\alpha}{1-\alpha}(x_2 - x_1) \tag{45}$$

where $\alpha = \max\{\alpha_1, \alpha_2\}$.

It remains to show that numbers a, b exist so that (i) and (ii) hold. Consider the sequence $\{[a_n, b_n]\}_{n=0}^{\infty}$ of intervals with extrema defined recursively by

$$a_{n+1} = \min_{x\in[a_n,b_n]} f_1(x) \qquad \text{and} \qquad b_{n+1} = \max_{x\in[a_n,b_n]} f_2(x), \qquad \text{with } x_1 \leq a_0 < b_0 \leq x_2.$$

Since contractivity of f_j's implies continuity, the sequences $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$ are well defined provided that $a_n < b_n$ for all $n = 0, 1, \ldots$. To see this, note that $a_1 \leq a_0$, as $f_1(x) \leq x$ for all $x \geq x_1$; similarly, $b_1 \geq b_0$. But $[a_0, b_0] \subseteq [a_1, b_1]$ implies $a_2 \leq a_1$ and $b_2 \geq b_1$. Continuing by induction, we observe that $\{[a_n, b_n]\}_{n=0}^{\infty}$ is a nested non-decreasing sequence of intervals which, since each interval $[a_n, b_n]$ has length bounded by (45), converges to some interval $[a, b]$, with extrema clearly satisfying (i) and (ii). ■

[17] To be precise, at least in the study of polarization phenomena, also the contractivity property could be relaxed somewhere in the "relevant region" (the square T in figure 6). The only minimum requirement is that the graphics of f_1, f_2 do not intersect inside this area and that the maps are contractions outside such area, so that the system is being attracted to the interval $[a, b]$ as time elapses. However, by relaxing contractivity, Theorem 1 cannot be applied and a characterization of the steady state(s) of system (43) becomes harder. Existence and uniqueness of invariant limiting distribution for non-contractive IFS will be discussed more thoroughly in Section 7.

Lemma 3 applies easily when maps f_j are monotone. In this case, the maximum and the minimum in conditions (i) and (ii) are reached on the extrema of interval $[a, b]$ and the relevant interval $[a, b]$ coincides with the trapping region. For example, if they are both increasing a, b coincide with the fixed points x_1, x_2 of the f_j's; if they are both decreasing, a, b are the unique solution of the system

$$\begin{cases} f_1(b) = a \\ f_2(a) = b \end{cases};$$

if f_1 is increasing and f_2 is decreasing, $a = x_1$ and $b = f_2(a)$; while if f_1 is decreasing and f_2 is increasing, $b = x_2$ and $a = f_1(b)$.

However, if maps f_j are not monotone, the trapping region in general will be some (proper) subset of $[a, b]$, as can be easily grasped by thinking of two unimodal maps reaching the maximum and the minimum point respectively somewhere in the middle of interval $[a, b]$, but with (disjoint) image sets $f_1([a, b])$ and $f_2([a, b])$ sufficiently apart one from the other, so that no point in $[a, b]$ is being mapped into $\arg\min_{x \in [a,b]} f_1(x)$ or into $\arg\max_{x \in [a,b]} f_2(x)$.

A few examples of general contractive non-linear maps $f_1 < f_2$ will be recalled in the next section.

6.2 Cookie-Cutter Sets

As in the linear case, the attractor of a non-linear IFS of the type (43) is the whole interval $[0, 1]$ provided that the two image sets $g_1([0, 1])$ and $g_2([0, 1])$ overlap. In this case we shall say that the invariant distribution has "full support". For our purposes, the non-overlapping case looks more appealing.

When $g_1([0, 1]) \cap g_2([0, 1]) = \varnothing$, a "hole" appears after one iteration of the IFS, and this gap will spread through the unit interval after successive iterations, in a fashion similar to that in figure 3. By taking infinite iterations of the Barnsley operator associated to the IFS, the attractor of the system turns out to be a set which is totally disconnected and topologically equivalent to a Cantor-like set of the kind discussed in Section 2.2. However, now maps g_1 and g_2 are non-linear, that is, they are not similarity maps in the sense of Section 2, and relative distances between any two points in $[0, 1]$ are modified through application of both maps. The invariant set of the IFS (43) is a "non-linear perturbation" of some Cantor-like set. The IFS is termed *cookie-cutter system* and its invariant set is called *cookie-cutter set*. Such a set might be thought as a "distorted" Cantor set, which nevertheless is "approximately self-similar".

Chapter 4 in [25] makes precise the idea of a set being "approximately self-similar" by stating the principles of bounded variation and bounded distortion. Loosely speaking, these principles say that "any sufficiently small neighborhood of the attractor may be mapped onto a large part of the set by a transformation that is not unduly distorting, and, conversely, the whole set may be mapped into small neighborhoods of itself without too much distortion". The intuitive interpretation of these invariant sets as "distorted" Cantor sets follows naturally, and is well illustrated in figure 4.3, p. 65 in [25]. Chapter 5 in [25] further explores properties of such invariant sets by establishing a formula to calculate their (fraction) Hausdorff dimension.

Examples of non-linear IFS generating cookie-cutter sets are easy to find in dynamical systems textbooks. The natural approach is to consider the "inverse" of a contractive IFS, which is an expanding system, for it can usually be described by a unique map f. Let I_1 and I_2 be disjoint closed subintervals of $[0, 1]$ and $f : I_1 \cup I_2 \to [0, 1]$ such that both I_1 and I_2 are mapped bijectively onto $[0, 1]$. Assume that f is continuous and is expanding, *i.e.*, there is some constant $\alpha > 1$ such that $|f(x) - f(y)| \geq \alpha |x - y|$ for all $x, y \in I_j$, $j = 1, 2$. Examples of this type are often constructed by taking the restriction to $[0, 1]$ of a unimodal function, like the logistic map $f(x) = rx(1-x)$, for $r > 2 + \sqrt{5}$, or the tent map $f(x) = r(1 - |2x - 1|)$, for $r > 1$, as in figure 7.

Since the system is expanding, we are interested in the invariant set A of the system: the set of points that remain in $I_1 \cup I_2$ under iteration by f. Such set turns out to be compact and non-empty and satisfies

$$f(A) = A = f^{-1}(A), \tag{46}$$

where $f^{-1}(X)$ denotes the pre-image of any set X under f. The invariant set A is a *repeller*, in the sense that $x \in A$ if and only if $x \in f(A)$ and points outside A eventually move outside $I_1 \cup I_2$ under iteration by f.

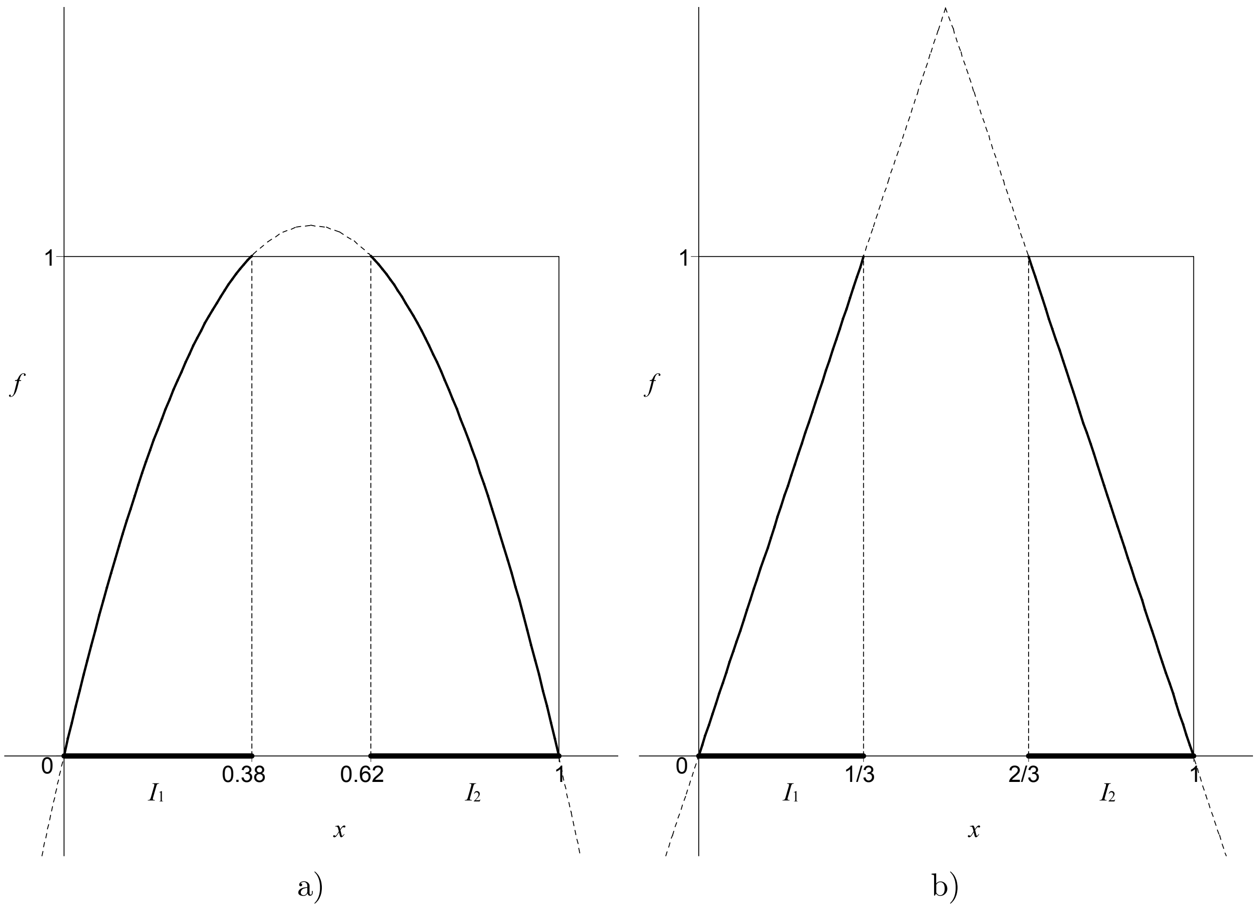


Figure 7: examples of expanding systems: a) the logistic map $f(x) = (17/4)x(1-x)$; b) the tent map $f(x) = (3/2)(1 - |2x - 1|)$.

As a matter of fact, the invariant set A can be seen as the *attractor* of the contractive IFS which is the inverse of the expanding system represented by f. Define $g_j : [0,1] \to [0,1]$, $j = 1, 2$, as the two branches of the inverse of f:

$$\begin{aligned} g_1(x) &= f^{-1}(x) \cap I_1 \\ g_2(x) &= f^{-1}(x) \cap I_2 \end{aligned}$$

Hence, g_1 and g_2 map $[0,1]$ bijectively onto I_1 and I_2 respectively[18]. Since $|f(x) - f(y)| \geq \alpha |x-y|$ for all $x, y \in I_j$, $j = 1, 2$, and for some $\alpha > 1$, it follows that the inverse functions g_1, g_2 are Lipschitz, both with Lipschitz constant $\alpha^{-1} < 1$: $|g_j(x) - g_j(y)| \leq \alpha^{-1}|x-y|$, $j = 1, 2$. Thus they are contractions on $[0,1]$ and for all probabilities $0 < p < 1$ the IFS $\{g_1, g_2, p\}$ has a unique compact set $A \subset [0,1]$ satisfying $A = g_1(A) \cup g_2(A)$ as attractor, which, thanks to (46), turns out to be the repeller of f.

The maps g_1, g_2 in case (b) of figure 7 are peculiar since they are linear: $g_1(x) = (1/3)x$ and $g_2(x) = 1 - (1/3)x$. As a matter of fact, their IFS is equivalent to system (5) for $\alpha = 1/3$ (the only difference is the sign of monotonicity of g_2), and it converges to an attractor which is the Cantor ternary set discussed in Section 2.2.

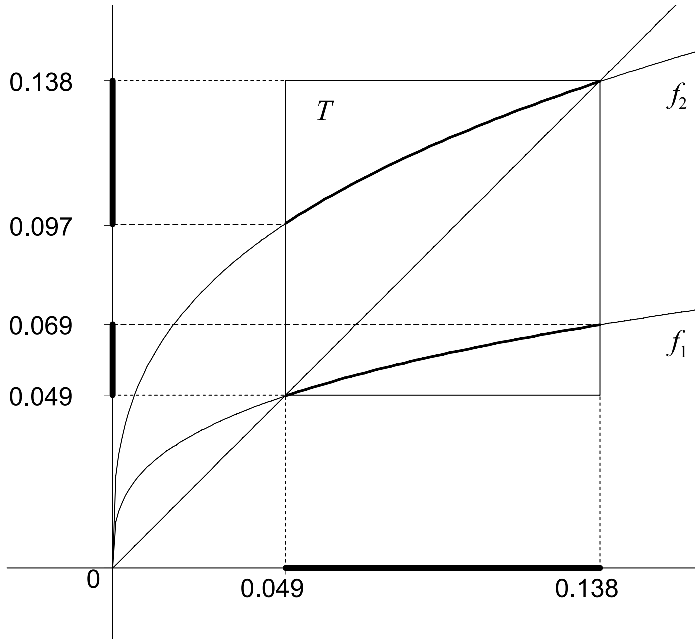


Figure 8: the relevant region of the IFS associated to $f_1 = (4/15)x^{1/3}$ and $f_2 = (2/15)x^{1/3}$.

It is also interesting to remark that the policy of the representative-agent optimal growth model tackled in [36] falls in the category of non-linear IFS, as it has the form

$$\begin{array}{l} f_1(x) = \alpha\beta r x^{\alpha} \\ f_2(x) = \alpha\beta x^{\alpha} \end{array} \tag{47}$$

where $0 < \alpha < 1$ is the parameter of the Cobb-Douglas production function $G(x) = x^{\alpha}$, $0 < \beta < 1$ is the discount factor and $0 < r < 1$ represents an exogenous technological perturbation that may occur with probability $1 - p$. In that paper, this non-linear IFS is being "linearized" into one of the type (5) thanks to a log-linear transformation, and it is normalized over the interval $[0,1]$ by applying formula (44). Hence, system (5) converging to some Cantor-like attractor is studied for values $\alpha < 1/2$. However, the true support of the invariant probability to which the economy converges in the long run is the Cookie-Cutter set obtained by an exponential transformation of the Cantor-like set, *i.e.*, it is a "distorted"

[18] In figure 7 (a), g_1 and g_2 are the inverse of the restrictions of the logistic map $f(x) = (17/4)x(1-x)$ to the intervals $[0, 0.38]$ and $[0.62, 1]$; while in figure 7 (b), g_1 and g_2 are the inverse of the restrictions of the tent map $f(x) = (3/2)(1 - |2x-1|)$ to the intervals $[0, 1/3]$ and $[2/3, 1]$.

Cantor set. Figure 8 shows the (true) relevant portion, included in the square T determined by the two fixed points $b \simeq 0.138$ and $a \simeq 0.049$ of the maps f_1, f_2, of the optimal dynamics (47) for $\alpha = 1/3$, $\beta = 4/5$ and $r = 1/2$. The trapping region is the interval $[0.049, 0.138]$ (bold on the horizontal axis), and the two maps f_1, f_2, have disjoint image sets, the intervals $I_1 = [0.049, 0.069]$ and $I_2 = [0.097, 0.138]$ (bold on the vertical axis) having different length.

Thanks to the log-linearization and normalization over the unit interval, it is immediately seen that the occurrence of non-overlapping image sets in system (47) depends exclusively on parameter α (which must be less than 1/2 to have a Cantor-like attractor), and not on other parameters, like the discount factor β or the value of the exogenous shock r.

6.3 Polarization/Pulverization Generated by Non-Linear IFS

It is important to remark that, as it happened in our models of Section 3, the "normalization" procedure described in Section 6.1 points out what are the parameters truly affecting the appearance of a fractal (possibly distorted) Cantor-like attractor. In Section 3 we have shown that the only relevant parameters are the degree of intergenerational altruism β and the exogenous growth rate γ. Any change in other aspects of the mode (like direct redistribution from the luckies to the unluckies discussed in Section 3.3) does not modify the shape of the attractor. We now generalize this result by providing sufficient conditions for polarization/pulverization and by extending the ineffectiveness of direct redistribution schemes to the case of non-linear IFS.

Proposition 9 *Let $\{f_1, f_2, p\}$ be a non-linear contractive IFS with $f_1 < f_2$, and let $[a, b]$ be their relevant interval as defined in Lemma 3. Then, if*

$$\max_{x\in[a,b]} f_1(x) < \min_{x\in[a,b]} f_2(x) \tag{48}$$

the attractor of the IFS $\{f_1, f_2, p\}$ is a cookie-cutter set (a "distorted" Cantor-like set), and so is the attractor of the normalized IFS $\{g_1, g_2, p\}$ as in (43) obtained through formula (44).

Proof. The first part is obvious from the discussion in Section 6.2. Let us show that condition (48) implies

$$\max_{x\in[0,1]} g_1(x) < \min_{x\in[0,1]} g_2(x). \tag{49}$$

Let $y_1 = \max_{x\in[a,b]} f_1(x)$ and $y_2 = \min_{x\in[a,b]} f_2(x)$. Since by assumption $y_1 < y_2$, by formula (44),

$$\max_{x\in[0,1]} g_1(x) = k^{-1}y_1 - a < k^{-1}y_2 - a = \min_{x\in[0,1]} g_2(x),$$

which proves the claim. Moreover, by reversing the argument, also (49) implies (48), thus establishing that (48) holds if and only if (49) holds. ■

The following proposition relates the appearance of polarization/pulverization to the Lipschitz constants of the (contractive) maps f_1, f_2.

Proposition 10 *Let $\{f_1, f_2, p\}$ be a non-linear contractive IFS, with f_1, f_2 defined over some interval $I \subseteq \mathbb{R}$, such that $f_1 < f_2$. For $j = 1, 2$, let $0 < \alpha_j < 1$ be such that $|f_j(x) - f_j(y)| \leq \alpha_j |x - y|$ for all $x, y \in I$. Then, if $\alpha_1 + \alpha_2 < 1$, the attractor of the IFS $\{f_1, f_2, p\}$ is a cookie-cutter set (a "distorted" Cantor-like set).*

Proof. Consider the normalized IFS $\{g_1, g_2, p\}$ as in (43) obtained by applying formula (44) to $\{f_1, f_2, p\}$. It is immediately seen that the maps g_1, g_2 are Lipschitz with α_1, α_2 as Lipschitz constants, since, for $j = 1, 2$,

$$\begin{aligned} |g_j(x) - g_j(y)| &= k^{-1} |f_j(kx + a) - a - f_j(ky + a) + a| \\ &= k^{-1} |f_j(kx + a) - f_j(ky + a)| \\ &\leq k^{-1} \alpha_j |kx + a - ky - a| \\ &= \alpha_j |x - y|, \end{aligned}$$

where $k = b - a$ and b, a are the numbers provided by Lemma 3. Since the normalized IFS $\{g_1, g_2, p\}$ has the interval $[0, 1]$ as trapping region and $\min_{x\in[0,1]} g_1(x) = 0$ and $\max_{x\in[0,1]} g_2(x) = 1$, it follows that $\max_{x\in[0,1]} g_1(x) \leq \alpha_1$ and $\min_{x\in[0,1]} g_2(x) \geq 1 - \alpha_2$; thus, $\max_{x\in[0,1]} g_1(x) < \min_{x\in[0,1]} g_2(x)$ whenever $\alpha_1 + \alpha_2 < 1$. In this case, the image set $g_1([0, 1]) \cup g_2([0, 1])$ is some subset of $[0, \alpha_1] \cup [1 - \alpha_2, 1]$, and has a "gap" with amplitude larger or equals to $1 - \alpha_1 - \alpha_2 > 0$. This is enough to generate a cookie-cutter set as attractor. ■

Proposition 10 basically states that the nature of the attractor of a IFS depends on the slope of the maps f_j: if they are sufficiently flat, then polarization/pulverization occurs. This observation leads to the next corollary, which generalizes Proposition 3 to the non-linear case.

Corollary 1 *Let $\{f_1, f_2, p\}$ be as in Proposition 10 with Lipschitz constants α_1, α_2. Any transformation of the maps f_1, f_2 that does not modify their slope, do not affect the nature of the attractor of the IFS. In particular, any additive transformation $\widehat{f_1} = f_1 + z$, $\widehat{f_2} = f_2 - z$, with $z < \min_{x\in I} (1/2) [f_2(x) - f_1(x)]$, does not affect relative distances between points of the (possibly fractal) attractor of the IFS.*

Proof. Since additive transformations $\widehat{f_j}$ have the same Lipschitz constants α_j of f_j, so have the "normalized" maps $g_j : [0, 1] \to [0, 1]$, for $j = 1, 2$, as shown in the proof of Proposition 10. Hence, Proposition 10 states that relative distances between points of the attractor of the IFS $\{f_1, f_2, p\}$ depend only on Lipschitz constants α_1, α_2 and not on additive constants. ■

The intuition behind this result is easily captured through the normalization argument of Section 6.1: "vertical" rigid translations of the maps f_1, f_2 are compensated by "horizontal" changes in the length of their trapping interval $[a, b]$. The simplest case is when f_1, f_2 are affine maps (similarities), as in Sections 2 and 3: as a matter of fact, rigid translations of affine maps leave the normalized IFS obtained through transformation (44) unaltered.

7 IFS with State-Dependent Probabilities

In the previous sections we have always assumed that the probability p of success, that is, the probability that the system (43) jumps to position $x_t = g_2(x_{t-1})$ at time t given its position x_{t-1} at time $t - 1$, is constant through time. Such assumption may appear too restrictive in the context of wealth polarization with social mobility. It is in fact questionable that the probability of success is the same for the rich and the poor; for example, the poor might find educational costs unbearable or access to credit market precluded, thus indirectly reducing their probability of success, while for the rich an easier access to education and credit markets improves their probability of being rich also in the future[19]. It seems reasonable to assume some degree of monotonicity of the probability of success with respect to the wealth: a probability of success $p = p(x)$ which is (at least) non-decreasing in x could well represent this scenario. Note that such assumption would introduce a "damping effect" on the dynamics of the models described in Sections 3 and 4 with the consequence of curbing social mobility.

While we do not attempt to modify our models of Sections 3 and 4 by assuming, for example, an individual utility effort e_t which depends on the wealth available at the beginning of period t, or by introducing an educational cost, in this section we briefly discuss the case in which the probability p of success at time t is allowed to depend on the state x_{t-1}, *i.e.*, in terms of the models of Sections 3 and 4, on the wealth inherited from the ancestor. In view of Section 6, a generic non-linear IFS $\{g_1, g_2, p(x)\}$ normalized over $[0,1]$ with Lipschitz maps $g_j : [0,1] \to [0,1]$ and probability function $p : [0,1] \to [0,1]$ is considered. Thus, system (43) can be rewritten as

$$x_t = \begin{cases} g_1(x_{t-1}) & \text{with probability } 1 - p(x_{t-1}) \\ g_2(x_{t-1}) & \text{with probability } p(x_{t-1}) \end{cases} \tag{50}$$

Such type of IFS is often called a *learning system* (see e.g. Karlin [32]), since the system "learns" from position x_{t-1} a new strategy $p(x_{t-1})$ for choosing the next stop.

It is immediately clear from the discussion in Section 2.2 that the geometric properties of the attractor of system (50) does not depend on probability $p(x)$, since the Barnsley operator (6) is defined only by means of the maps g_j, and is not affected by probabilities. In other words, the polarization results of Section 6.3 remain unchanged: polarization/pulverization appears as long as the image sets $g_1([0,1])$ and $g_2([0,1])$ do not overlap, and, if the Lipschitz constants α_j of maps g_j are such that $\alpha_1 + \alpha_2 < 1$, any rigid vertical shift leaves the relative shape of the attractor unaltered. This is important, since it implies that our main results remain valid also for more general models with state-dependent probabilities.

However there is the delicate question of convergence of the marginal probabilities to some (unique) invariant probability, independently of initial conditions, that must be tackled carefully.

7.1 Invariant Distributions for IFS with State-dependent Probabilities

Recall that Theorem 1 holds for a constant probability $0 < p < 1$. To obtain an equivalent Theorem for state-dependent probability $p(x)$ some more assumptions, specifically on the function $p : [0,1] \to [0,1]$, are required. They are summarized in the following Theorem. As usual, let $\mathcal{B}$ be the σ-algebra of Borel measurable subsets of $\mathbb{R}$ and $\mathcal{P}$ the space of probability measures on $(\mathbb{R}, \mathcal{B})$. Define the Markov operator $M : \mathcal{P} \to \mathcal{P}$ as

$$M\mu(B) = \int_{g_1^{-1}(B)} [1 - p(x)]\, \mu(dx) + \int_{g_2^{-1}(B)} p(x)\, \mu(dx), \qquad \text{for all } B \in \mathcal{B}.$$

Successive iterations of M starting from some initial probability μ_0, $\mu_t(B) = M[M^{t-1}\mu_0(B)]$, yield the evolution of marginal probabilities of the system as time elapses. As in Section 2.1, we are interested in the limit $\lim_{t\to\infty} M^t \mu_0$ for any starting probability μ_0.

[19] These observations suggest that models on wealth inequality from the traditional stream of research, like the ones in Galor and Zeira [27] or in Aghion and Bolton [1] (see also the whole literature cited in the introduction), which assume imperfect capital markets, may easily fit our framework with the necessary modifications.

Theorem 2 *Assume that the IFS* $\{g_1, g_2, p(x)\}$ *over* $[0,1]$ *satisfies the following:*

i) *for* $j = 1, 2$*, a constant* $\alpha_j > 0$ *(not necessarily less than* 1*) exists such that* $|g_j(x) - g_j(y)| \leq \alpha_j |x - y|$ *for all* $x, y \in [0,1]$*;*

ii) *a constant* $\delta < 0$ *exists so that* $[1 - p(x)] \ln \alpha_1 + p(x) \ln \alpha_2 \leq \delta$ *for all* $x \in [0,1]$*;*

iii) $0 < p(x) < 1$ *for all* $x \in [0,1]$*;*

iv) *a constant* $\lambda > 0$ *exist so that* $|p(x) - p(y)| \leq \lambda |x - y|$ *for all* $x, y \in [0,1]$*.*

Then, there is a unique probability distribution μ^* *on* $([0,1], \mathcal{B}([0,1]))$ *satisfying the functional equation*

$$\mu^*(B) = \int_{g_1^{-1}(B)} [1 - p(x)] \mu^*(dx) + \int_{g_2^{-1}(B)} p(x) \mu^*(dx) \qquad \text{for all } B \in \mathcal{B}([0,1])$$

which is supported on the attractor of the Barnsley operator (6), and, for any probability μ_0 *on* $([0,1], \mathcal{B}([0,1]))$*, the sequence* $\mu_t = M^t(\mu_0)$ *converges weakly to* μ^**.*

A general proof for the theorem above has been first provided by Barnsley et al. [9]. A slightly modified proof can be found in [34]. See also [20] and [47] for recent surveys on ergodic results for generalized IFS. It should be noted that Theorem 2 is a slightly simplified version of the general theorems available from the mathematical literature cited above; some assumptions can actually be weakened. For example, in [34] a similar results is given for Polish state spaces which may be not (even locally) compact, and in [45] it is further extended to IFS with a continuum of maps. Moreover, Lipschitz condition (*iv*) on probabilities it is usually replaced by a weaker notion of Dini continuity[20].

In Theorem 2 conditions (*i*) and (*ii*) generalize strict contractivity of maps g_j that has been assumed everywhere in the previous sections. Here the maps are still required to be Lipschitz but the Lipschitz constants α_j are allowed to be larger than 1. In fact, it has been shown that what really matters is some "average contractivity" property of the type[21] expressed by condition (*ii*), which allows one of the g_j's to have slope larger than 1 on some subset of $[0,1]$ (only one of the two on the same sub-interval, however). Note that, since Theorem 2 holds in particular for a constant probability $p(x) \equiv p$, the assumption of strict contractivity of both maps in Theorem 1 may be weakened and replaced by conditions (*i*) and (*ii*) of Theorem 2,

[20] We say that a probability vector $(p_1(x), \ldots, p_n(x))$ satisfies the Dini condition if there is a continuous, nondecreasing, concave function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\sum_{j=1}^{n} |p_j(x) - p_j(y)| \leq \varphi(|x - y|) \qquad \text{for all } x, y$$

and

$$\int_0^1 \frac{\varphi(t)}{t} dt < \infty.$$

Such condition is more restrictive than continuity of functions p_j, but less restrictive than Lipschitz (like our condition (*iv*) in Theorem 2) or Hölder continuity. See, *e.g.*, [34] for more details.

[21] Some authors, *e.g.* Lasota [34] and Stenflo [45], assume a slightly more restrictive "average contractivity" property of the form

$$[1 - p(x)] \alpha_1 + p(x) \alpha_2 < 1.$$

Clearly, by Jensen inequality, this is implied by condition (*ii*) in Theorem 2. See also Chapter 12 in [33].

at least for the non-linear case discussed in the Section 6. That is, even if the slope of one of the f_j's is larger than 1 somewhere in the relevant interval $[a, b]$, which means $\alpha_j > 1$ for some $j \in \{1, 2\}$, there is still weak convergence to a unique invariant probability independent of initial conditions as long as $(1 - p) \ln \alpha_1 + p \ln \alpha_2 \leq \delta < 0$. However we do not develop the details here.

We already had a flavor of how difficult is to determine the complete nature of the invariant probability μ^* for i.i.d. IFS (*i.e.*, with state-independent probabilities) in Section 2.3; of course, letting probability p depend on the state of the process at each time cannot provide any improvement. However, Barnsley et al. [9] proved that the "pure type" result still holds: μ^* is either singular or absolutely continuous on its whole support.

It is interesting to see what happens if some of the assumptions in Theorem 2 are relaxed. In particular, removing condition (*iii*) (but this is not the only way, see, *e.g.*, [46]), will in general lead to the appearance of more than one distinct invariant probability dependent on the initial condition. That is, conditions (*iii*) and (*iv*) of Theorem 2 guarantee uniqueness of the invariant probability μ^*. We conclude this section with an example where condition (*iii*) is relaxed, and which yields an interesting economic interpretation when applied to the model of Section 3.

7.2 An Example of a (*a.s.*) Totally Polarized Economy

Consider system (50) with $g_1(x) = \alpha x$, $g_2(x) = \alpha x + (1 - \alpha)$ for $0 < \alpha < 1$, and $p(x) = x$. Clearly, the probability function does not satisfy (*iii*) in Theorem . Hence, it is easily seen that the IFS has two invariant probabilities: the Dirac measure $\delta_0(x)$ concentrated fully on $x = 0$ and the Dirac measure $\delta_1(x)$ concentrated fully on $x = 1$ (see Karlin [32]). If the initial position is $x_0 = 0$ (that is, if the initial marginal probability is $\mu_0 = \delta_0$ itself), then the system remains at the same position forever, while if the initial position is $x_0 = 1$ (or $\mu_0 = \delta_1$), then the system is trapped at $x_t = 1$ for all t. For all other starting points $0 < x_0 < 1$, the system eventually is being attracted (with probability 1) into one of the absorbing states $x = 0$ or $x = 1$, with an initial probability to hit $x = 1$ which is an increasing function of x_0.

We now apply this example to our model developed in Section 3. Probability $p(x) = x$, the probability of success, is an increasing function of the (productivity adjusted and normalized) wealth, hence it does make sense from the economic point of view. As time goes to infinity, whatever the initial distribution the whole population is being eventually trapped into these two absorbing states. The society ends up into two wealth clusters, "totally polarized" at the extrema of interval $[0, 1]$. It is important to observe that $x = 0$ and $x = 1$ are absorbing states for all $0 < \alpha < 1$; hence, the society in the long run becomes divided into two groups, the "completely derelict" and the "super rich" also for $\alpha > 1/2$, namely, also when the attractor of the Barnsley operator (6) is the full interval $[0, 1]$, and thus, with a probability function satisfying (*iii*) of Theorem 2, one would expect an invariant probability spread over all $[0, 1]$. Therefore, the polarization obtained through state-dependent probabilities in this extreme example is independent of the results related to the geometrical properties of the attractor discussed in the previous sections.

The weight of the two groups polarized at $x = 0$ and $x = 1$ depend, of course, on the initial distribution of wealth. Recall, however, that, in order to let every individual put the required effort into learning the new technology, some assumption equivalent to A.1 must hold. By assuming $e_0 > 0$, from A.1 we get

$$p(x) = x > \frac{e_0}{\rho A_0} = m$$

where $\rho = (1-\beta)^{1-\beta}\beta^{\beta}$, which implies that all individuals starting with wealth $x_0 \in [0, m]$ will not undertake any effort and thus they will eventually go to swell the group of the (completely) poor with certainty; while the model in Section 3 should be be re-normalized over the individuals with initial (productivity adjusted and normalized) wealth $x_0 \in [m, 1]$. In other words, the group of the poor will always have strictly positive mass, whenever there is a strictly positive mass of individuals with initial wealth $x_0 \leq m$.

Finally, observe that whenever we replace $p(x) = x$ with any Lipschitz increasing function $p(x)$ such that $p(0) = 0$ and $p(1) < 1$, $x = 0$ is the unique absorbing state, and, with probability one, eventually the whole society extinguishes with all individuals totally poor, regardless of the value $0 < \alpha < 1$.

8 Conclusions

In this paper we have pointed out how wealth polarization/ pulverization is not to be contrasted with socioeconomic mobility, but that instead it can be the effect of a very strong mobility of the individuals.

We have also shown that the existence of a middle class and, more generally, of a non polarized wealth distribution and a non-disconnected society can be the result of a slowly growing economy with little fiscal intervention. What really matters for polarization/ pulverization is the reward from being successful, which is increasing in the size of the technological jump.

Private investment in the human capital necessary to adopt an exogenous innovation stream can be a cause of a "fractal society". The private investment in research aimed at improving everybody's productivity can be another cause. Both the risky adoption of an exogenous innovation stream and the risky development of innovations can have the same consequences on the support of the limit distribution of productivity adjusted individual wealths. Despite the differences between these two engines of growth, they generate fractal distribution support and polarization if and only if the implied growth rate of the economy is higher than a common threshold.

We have shown that in our successive generations framework polarization and pulverization cannot be eliminated by fiscal measures such as wealth redistribution through taxation of the successful people with equally rebated tax revenues, while taxation can even create the fractalization of society. Also a random taxation scheme is substantially incapable of reducing polarization, at least if the incentive compatibility constraint is tight enough. As a general result we can say that in the examples of this paper redistribution is unable to eliminate socioeconomic disconnection and that it can even increase it.

Finally, we have briefly illustrated how the study of dynamics in stochastic economies by means of IFS can be generalized to models evolving through non-linear maps, also with associated probabilities which are state-dependent. Sufficient conditions for polarization/pulverization are available also for economies exhibiting these generalized dynamics. We plan, in a future paper, to extend our main results to models with infinitely lived agents.

References

[1] Aghion, P. and P. Bolton, A Theory of Trickle-Down Growth and Development, *Review of Economics Studies* **64**: 151-172, 1997.

[2] Aghion, P. and P. Howitt, *Endogenous Growth Theory*, Cambridge, MA: MIT Press, 1998.

[3] Aghion, P., E. Caroli and C. Garcia-Penalosa, Inequality and Economic Growth: The Perspective of the New Growth Theories, *Journal of Economic Literature* **37**: 1615-1660, 1999.

[4] Alesina, A. and D. Rodrik, Distributive Politics and Economic Growth, *Quarterly Journal of Economics* **109**: 465-490, 1994.

[5] Andreoni, J., Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence, *Journal of Political Economy* **6**: 1447-1458, 1989.

[6] Arnold L., *Random Dynamical Systems*, Berlin: Springer-Verlag, 1998.

[7] Banerjee A. and A. Newmann, Occupational Choice and the Process of Development, *Journal of Political Economy* **98**: 274-298, 1993.

[8] Barnsley, M. F. and S. Demko, Iterated Function Systems and the Global Construction of Fractals, *Proceeding Royal Society London Series* **A 399**: 243-275, 1985.

[9] Barnsley, M. F., S. G. Demko, J. H. Elton and J. S. Geronimo, Invariant Measures for Markov Processes Arising from Iterated Function Systems with Place-Dependent Probabilities, *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques* **24**: 367-394, 1988 (Errata Corrige, *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques* **25**: 589-590, 1989).

[10] Barro, R., Inequality, Growth, and Investment, *NBER Working Paper* **7038**, 1999.

[11] Barro, R., Inequality and Growth in a Panel of Countries, *Journal of Economic Growth* **5**: 5-30, 2000.

[12] Benabou, R., Inequality and Growth, *NBER Macroeconomics Annual*, Cambridge, Mass.: MIT Press, 1996.

[13] Benabou, R., Unequal Societies: Income Distribution and the Social Contract, *American Economic Review* **90**: 96-129, 2000.

[14] Benhabib, J. and M. Spiegel, Cross-Country Growth Regressions, *C.V. Starr Center Research Report* **9720**, 1997.

[15] Benhabib, J. and A. Rustichini, Social Conflict and Growth, *Journal of Economic Growth* **1**: 125-142, 1996.

[16] Bhattacharya, R. and M. Majumdar, Convergence to Equilibrium of Random Dynamical Systems Generated by IID Monotone Maps, With Applications to Economics, in Ghosh, S. (Ed.), *Asymptotics, Nonparametrics, and Time Series*, New York: Marcel Dekker Inc., 713-741, 1999.

[17] Bhattacharya, R. and M. Majumdar, On a Theorem of Dubins and Freedman, *Journal of Theoretical Probability* **12**: 1067-1087, 1999.

[18] Bhattacharya, R. and M. Majumdar, On a Class of Stable Random Dynamical Systems: Theory and Applications, *Journal of Economic Theory* **96**: 208-229, 2001.

[19] Crownover, R. M., *Introduction to Fractals and Chaos*, Boston: Jones & Bartlett Publishers, 1995.

[20] Diaconis P. and D. Freedman, Iterated Random Functions, *SIAM Review* **41**: 45-76, 1999.

[21] Edgar, G. A., *Measure, Topology and Fractal Geometry*, New York: Springer-Verlag, 1990.

[22] Erdös, P., On a Family of Symmetric Bernoulli Convolutions, *American Journal of Mathematics* **61**: 974-975, 1939.

[23] Esteban, J: M. and D. Ray, On the Measurement of Polarization, *Econometrica* **62**: 819-851, 1994.

[24] Falconer, K. J., *Fractal Geometry*, Chichester: John Wiley & Sons, 1990.

[25] Falconer, K. J., *Techniques in Fractal Geometry*, New York: John Wiley & Sons, 1997.

[26] Forbes K. J., A Reassessment of the Relationship Between Inequality and Growth, *American Economic Review* **90**: 869-887, 2000.

[27] Galor, O. and J. Zeira, Income Distribution and Macroeconomics, *Review of Economic Studies* **60**: 35-52, 1993.

[28] Helpman, E. (Ed.), *General Purpose Technologies and Economic Growth*, Cambridge, MA: MIT-Press, 1998.

[29] Howitt, P., Steady Endogenous Growth with Population and R&D Inputs Growing, *Journal of Political Economy* **107**: 715-730, 1999.

[30] Hutchinson, J., Fractals and Self-similarity, *Indiana University Mathematics Journal* **30**: 713-747, 1981.

[31] Jones, C., R&D-Based Models of Economic Growth, *Journal of Political Economy* **103**: 759-784, 1995.

[32] Karlin, S., Some Random Walks Arising in Learning Models, *Pacific Journal of Mathematics* **3**: 725-756, 1953.

[33] Lasota, A. and M. C. Mackey, *Chaos, Fractals, and Noise*, *Stochastic Aspects of Dynamics* New York: Springer-Verlag, 1994.

[34] Lasota, A., From Fractals to Stochastic Differential Equations, in Garbaczewski P., M. Wolf and A. Weron (Eds.), *Chaos - The Interplay Between Stochastic and Deterministic Behaviour, Proceedings of the XXXIst Winter School of Theoretical Physics*, Berlin Heidelberg: Springer-Verlag, 235-255, 1995.

[35] Loury, G., Intergenerational Transfers and the Distribution of Earnings, *Econometrica* **49**: 843-867, 1981.

[36] Mitra, T., L. Montrucchio and F. Privileggi, The Nature of the Steady State in Models of Optimal Growth Under Uncertainty, *CAE Working Paper* **01-04**, Cornell University, 2001 (dwnloadable http://www.arts.cornell.edu/econ/cae/01wp.htm). Forthcoming in *Economic Theory*.

[37] Montrucchio, L. and F. Privileggi, Fractal Steady States in Stochastic Optimal Control Models, *Annals of Operations Research* **88**: 183-197, 1999.

[38] Perotti, R., Growth, Income Distribution, and Democracy: What the Data Say, *Journal of Economic Growth* **1**: 149-187, 1996.

[39] Persson, T. and G. Tabellini, Is Inequality Harmful for Growth? Theory and Evidence, *American Economic Review* **84**: 600-621, 1994.

[40] Piketty, T., The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing, *Review of Economic Studies* **64**: 173-189, 1997.

[41] Sargent, T. J., *Dynamic Macroeconomic Theory*, Cambridge, Mass.: Harvard University Press, 1987.

[42] Schumpeter, J. A., *The Theory of Economic Development*, Cambridge, MA: Harvard University Press, 1934.

[43] Schumpeter, J. A., *Business Cycles. A Theoretical, Historical and Statistical Analysis of the Capitalist Process*, New York: Mc Graw Hill, 1939.

[44] Stark, J. and P. Bressloff, Iterated Function Systems and their Applications, in M. Farge, J. C. R. Hunt and J. C. Vassilicos, eds., *Wavelets, Fractals and Fourier Transforms,* Oxford: Clarendon Press, 65-90, 1993.

[45] Stenflo, Ö., Ergodic Theorems for Markov Chains Represented by Iterated Function Systems, *Bulletin of the Polish Academy of Sciences: Mathematics* **49**: 27-43, 2001 (downloadable http://abel.math.umu.se/~stenflo/research.html).

[46] Stenflo, Ö., A Note on a Theorem of Karlin, *Statistics and Probability Letters* **54**: 183-187, 2001 (downloadable http://abel.math.umu.se/~stenflo/research.html).

[47] Stenflo, Ö., Uniqueness of Invariant Measures for Place-Dependent Random Iterations of Functions, forthcoming in *The IMA Volumes in Mathematics and its Applications*, New York: Springer-Verlag (downloadable http://abel.math.umu.se/~stenflo/research.html)

[48] Stokey, N. L. and R. E. Lucas, *Recursive Methods in Economic Dynamics*, Cambridge, Mass.: Harvard University Press, 1989.

[49] Strogatz, S. H., *Nonlinear Dynamics and Chaos*, New York: Addison-Wesley, 1994.

[50] Vrscay, E. R., Iterated Function Systems: Theory, Applications and the Inverse Problem, in J. Bélair and S. Dubuc, eds., *Fractal Geometry and Analysis*, Dordrecht: Kluwer Academic Publishers, 405-468, 1991.

INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH
APPLIED MATHEMATICS WORKING PAPER SERIES

1. Luigi Montrucchio and Fabio Privileggi, “On Fragility of Bubbles in Equilibrium Asset Pricing Models of Lucas-Type,” *Journal of Economic Theory* 101, 158-188, 2001 (ICER WP 2001/5).
2. Massimo Marinacci, “Probabilistic Sophistication and Multiple Priors,” *Econometrica* 70, 755-764, 2002 (ICER WP 2001/8).
3. Massimo Marinacci and Luigi Montrucchio, “Subcalculus for Set Functions and Cores of TU Games,” April 2001 (ICER WP 2001/9).
4. Juan Dubra, Fabio Maccheroni, and Efe Ok, “Expected Utility Theory without the Completeness Axiom,” *Journal of Economic Theory*, forthcoming (ICER WP 2001/11).
5. Adriana Castaldo and Massimo Marinacci, “Random Correspondences as Bundles of Random Variables,” April 2001 (ICER WP 2001/12).
6. Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci, and Marciano Siniscalchi, “A Subjective Spin on Roulette Wheels,” July 2001 (ICER WP 2001/17).
7. Domenico Menicucci, “Optimal Two-Object Auctions with Synergies,” July 2001 (ICER WP 2001/18).
8. Paolo Ghirardato and Massimo Marinacci, “Risk, Ambiguity, and the Separation of Tastes and Beliefs,” *Mathematics of Operations Research* 26, 864-890, 2001 (ICER WP 2001/21).
9. Andrea Roncoroni, “Change of Numeraire for Affine Arbitrage Pricing Models Driven By Multifactor Market Point Processes,” September 2001 (ICER WP 2001/22).
10. Maitreesh Ghatak, Massimo Morelli, and Tomas Sjoström, “Credit Rationing, Wealth Inequality, and Allocation of Talent”, September 2001 (ICER WP 2001/23).
11. Fabio Maccheroni and William H. Ruckle, “BV as a Dual Space,” *Rendiconti del Seminario Matematico dell'Università di Padova*, forthcoming (ICER WP 2001/29).
12. Fabio Maccheroni, “Yaari Dual Theory without the Completeness Axiom,” October 2001 (ICER WP 2001/30).
13. Umberto Cherubini and Elisa Luciano, “Multivariate Option Pricing with Copulas,” January 2002 (ICER WP 2002/5).
14. Umberto Cherubini and Elisa Luciano, “Pricing Vulnerable Options with Copulas,” January 2002 (ICER WP 2002/6).
15. Steven Haberman and Elena Vigna, “Optimal Investment Strategies and Risk Measures in Defined Contribution Pension Schemes,” *Insurance: Mathematics and Economics* 31, 35-69, 2002 (ICER WP 2002/10).
16. Enrico Diecidue and Fabio Maccheroni, “Coherence without Additivity,” *Journal of Mathematical Psychology*, forthcoming (ICER WP 2002/11).
17. Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci, “Ambiguity from the Differential Viewpoint,” April 2002 (ICER WP 2002/17).

18. Massimo Marinacci and Luigi Montrucchio, “A Characterization of the Core of Convex Games through Gateaux Derivatives,” April 2002 (ICER WP 2002/18).
19. Fabio Maccheroni and Massimo Marinacci, “How to cut a pizza fairly: fair division with decreasing marginal evaluations,” *Social Choice and Welfare*, forthcoming (ICER WP 2002/23).
20. Erio Castagnoli, Fabio Maccheroni and Massimo Marinacci, “Insurance premia consistent with the market,” *Insurance: Mathematics and Economics* 31, 267-284, 2002 (ICER WP 2002/24).
21. Fabio Privileggi and Guido Cozzi, “Wealth polarization and pulverization in fractal societies,” September 2002 (ICER WP 2002/39).
22. Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci, “Certainty Independence and the Separation of Utility and Beliefs,” December 2002 (ICER WP 2002/40).