GRÉDI

Groupe de Recherche en Économie et Développement International

Cahier de recherche / Working Paper
09-17

Extremal behavior of aggregated economic processes in a structural growth model

Stéphane Auray

Aurélien Eyquem

Frédéric Jouneau-Sion

UNIVERSITÉ DE SHERBROOKE

Extremal Behavior of Aggregated Economic Processes in a Structural Growth Model*

Stéphane Auray† Aurélien Eyquem‡ Frédéric Jouneau-Sion§

January 2010

Colligebant autem mane singuli quantum sufficere poterat ad vescendum cumque incaluisset sol liquefiebat

In die vero sexta collegerunt cibos duplices id est duo gomor per singulos homines venerunt autem omnes principes multitudinis et narraverunt Mosi

(Exodus, 16 21-22)

Abstract

An AK-type model in which returns to scale may be strictly increasing or decreasing depending on random shocks is studied. We show that relevant observable series (including capital, growth, various relative prices) are all related to a Random Autoregressive Coefficient model. Recent works on extreme behavior on dependent process are presented and their link with our model is discussed. First, this model typically displays fat tail behavior even if the shocks do not. Second extremely (un)lucky events tend to appear in "cluster". We show that both fat tails and clustering of extreme values are consistent with arbitrarily small variations of returns to scale around the unit-root case. An annual sequence of real wages in England recorded since the XIII-th century onwards is used to provide empirical support for our model. The tests clearly reject the constant AR model and support the random coefficient hypothesis. We finally estimate the parameters related to extreme events and find empirical grounds for major booms/crises to appear once or twice per century.

Keywords: economic growth, extremal behavior, dependent processes.
JEL Classification: C22, C46, N13, O41, O47.

*We would like to thank Christian Francq and Philippe Barbe for insightful comments and Guillaume Daudin for providing us with the data set.

†EQUIPPE (EA 4018), CNRS–MESHS (USR 3185), Universités Lille Nord de France (ULCO), Université de Sherbrooke (GREDI) and CIRPÉE, Canada. Email: stephane.auray@gmail.com

‡GATE (UMR 5824), Université de Lyon, Ecole Normale Supérieure LSH, France and GREDI, Canada. Email: aurelien.eyquem@ens-lyon.fr.

§EQUIPPE (EA 4018), Universités Lille Nord de France. Email: frederic.jouneau@univ-lille3.fr.

1 Introduction

Most neo-classical growth models now belong to the so-called ”endogenous growth” literature. In many of these models, growth is obtained by increasing returns to scale.[1] In this case, the law of motion for the capital usually follows an AR(1) model in which the autoregressive coefficient is equal or larger than 1. It is then customary to decompose the series between a trend (which describes the long term behavior) and a transitory component (which captures the cyclical properties). Many procedures and approaches have been proposed to obtain such decompositions among which the Hodrick-Prescott filter and the co-integration are the most popular.

At least two pitfalls of this literature may be identified. Empirically, evidence for non stationarity is not always clear-cut especially when the sample size is moderate. Indeed, many unit root tests lack of power and size distortions may be large (see DeJong, Nankervis, Savin and Whiteman (1992), Jaeger (1990), Perron and Phillips (1987) and Shiller and Perron (1985)). Theoretically, the above framework implies that if returns to scale have been increasing in the past this will lasts forever. Of course this implication is at odds with economic evidence when the data set cover periods before and after the Industrial Revolution.

This last point has been particularly stressed by the proponents of the “unified growth theory” (see Galor (2005), Galor (2007), Galor and Moav (2004), Galor, Moav and Vollrath (2009), Galor and Weil (1999), and Galor and Weil (2000), among many others). To this end they propose models which encompass the main steps of development (namely Malthusian, post-Malthusian and permanent growth episodes). The main ingredients of these models (see for instance Galor (2000)) are twofold. During the first Malthusian epoch the increase in population size is related to the relative abundance of consumption goods. Once population is large enough, a “quality” effect creates increasing returns to scale and permanent growth. The main difficulty is to explain the transition path from one regime to the other.

[1] As it is well-known increasing returns to scale may hamper the mere existence of general competitive equilibrium. To overcome this difficulty, it is often assumed that returns are decreasing at the individual level and increasing in the aggregate, see below for details.

In this paper, we follow an alternative approach based on a stochastic growth model, which we confront to long period data (more precisely English wage data from 1264 to 2008). As mentioned above, a unit root-type model will face difficulty to cope with the earlier period (*i.e.* before 1800). On the other hand, a single stationary ARMA-type model cannot account for the dramatic growth episode of the XIX-th and XX-th centuries. However different ARMA-type model may successfully be fitted to parts of the observed sequence. We argue that a stochastic growth model in which returns to scale are subject to random shocks may perform satisfactorily over the whole period. Time series generated by the model admit Random-Coeffcients-ARMA (hereafter RC-ARMA) structure. We then investigate the statistical properties of these series.

Several interesting features arise. First, stationary behavior may coexist with episodes of increasing returns to scale. Second, time-varying returns to scale affect the extreme behavior of the series. More precisely, if we allow for periods with increasing returns to scale then the stationary distribution typically displays a Pareto-type behavior. Evidence for fat tail are numerous in macro-econometrics (recall indeed Engle (1982) originally tested ARCH models on macro-economic series). Structural explanations are much sparse (Granger's (1980) aggregation model being an exception). Finally, statistical literature on extreme behavior of dependent processes emphasizes the fact that very large (and small) values typically tend to arrive in "clusters" (see, among many others, Rootzén (1978), Leadbetter (1983), Leadbetter and Rootzén (1988)). This feature is also documented for quite a long time in economics. In particular, it has long been noticed that economic series tends to follow similar patterns during large "crises" or "booms" (see Barro (2006), Barro and Ursua (2008) and Barro (2009) for recent investigations).

The rest of the paper is as follows. In Section 2, we present the data set and investigate the fit of ARMA models for several sub-periods. In Section 3, we present our stochastic growth model and derive the dynamics for observable series. In Section 4, we recall some striking features of the literature on extremal behavior of dependent processes and apply its relevant results to our model. In particular, we show that an arbitrary small randomness of the returns to scale has dramatic consequences on the extremal behavior of the series. Finally in Section 5, we use our model to investigate the link between the parameters governing the extreme phenomena and the structural parameters of the model. We use this connection to provide estimates of extremal parameters. Section 6 concludes and contains some insights for future research.

2 Evidence from wages in England 1264-2008

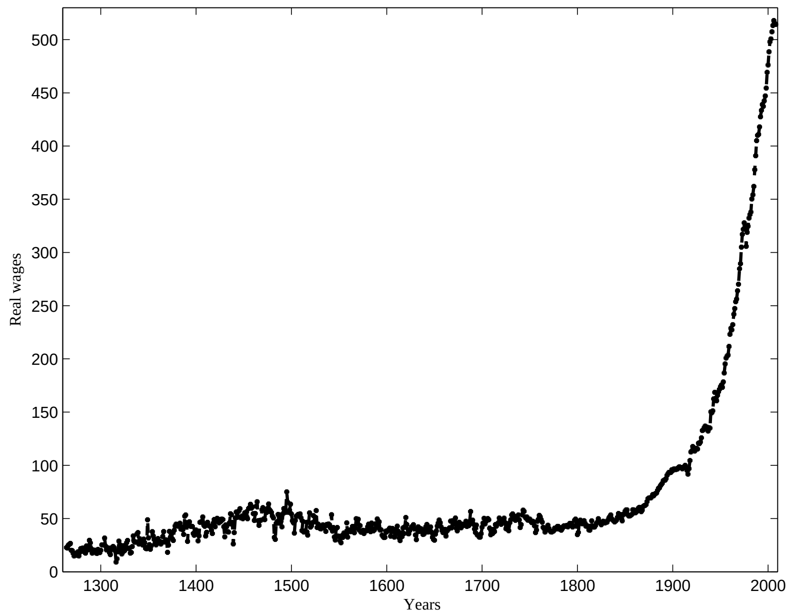
The usefulness of varying coefficient in macro-economic series has been a strongly debated question since the famous Keynes-Tinbergen controversy. More recently, several papers failed to reject the hypothesis of stable AR(1) structure both under non-stationary (for a recent investigation, see Distaso (2008)) and possibly stationary hypotheses. However, these investigations have been conducted on post-war US data. When considering very long term relationships such as those involved in growth theory, longer data sets should be considered. To this end, we look at one of the longest available macro-economic data set: an annual sequence of real wages in England starting in 1264.[2]

The chart presented in Figure 1 is well known among historians and economist familiar with the evidences on currently developed countries. Real wages have dramatically increased during XIX-th and XX-th centuries and look virtually stationary previously. One obtains similar patterns for other developed countries and other series (such as GDP for instance). Evidence from large panel data (see, *e.g.* Maddison (2000)) also witnesses a dramatic increase in the difference of growth rates between countries. Explaining this phenomenon (coined as "Great Divergence") is a major goal in growth theory.

We argue in Section 3 below that the usual AK growth model with iid technological shocks implies that the logarithm of wages admits an ARMA(1,1) structure. Obviously a single ARMA(1,1) model cannot fit the previous sequence. Table 1 provides the estimates of

[2]This exceptionally long sequence is available at http://www.measuringworth.org/ukearncpi/

Figure 1: Time series of real wages (England 1264-2008)



ARMA(1,1) models computed from real wages subsequences of our entire data set (we follow here the conventional historical distinction between Middle Age, Modern, Contemporary, and Immediate -*e.g* post-WWII- periods).

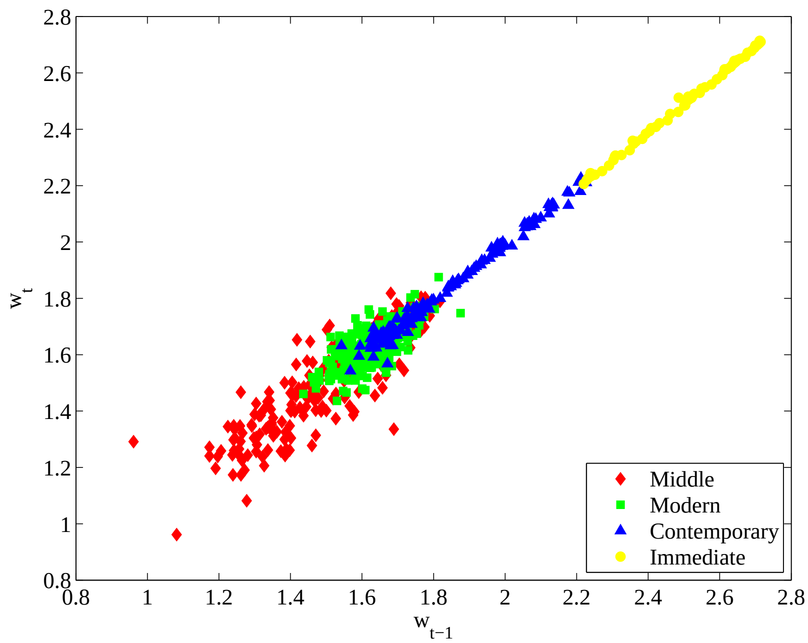
Table 1: Estimates of ARMA(1,1)

	order-1 autocorrelation		order-1 moving average	
period	estimate	std. err.	estimate	std. err.
1265-1492	0.906	0.033	-0.054	0.064
1493-1789	0.806	0.045	-0.212	0.069
1790-1946	0.998	0.013	0.382	0.045
1947-2008	0.991	0.009	0.174	0.107
1264-2008	0.999	0.008	-0.3189	0.029

The coefficients and in particular the order-1 autocorrelation clearly display significative variations. Although Contemporary and Immediate periods account for most of the growth, our estimation shows significant variations between the two first sub-periods. It also appears that the unit-root problem shows up precisely when the industrial revolution begins. Finally, notice that the estimate for the autoregressive coefficient over the whole period is also very close to unity.

As another informal approach we propose a scatter plot of (w_t, w_{t-1}). The dispersion around the 45-degree line appears larger during Middle Age than during Modern period, the dispersion being the lowest in the most recent times. Also, the variability of wages during Middle Age is much larger than during Contemporary period.

Figure 2: Scatter plot of actual wages versus lagged wages (data are in log).



As it is well known among historians, the major part of the high volatility of wages in Middle Age period is explained by variations of nominal prices in the agricultural sector. This largely due to erratic fluctuations in the agricultural returns caused by very different climatic episodes and conflicts. Finally contemporary and Immediate periods differ dramatically from previous times with very low dispersion around the 45-degree line together with large variation on the axis. The amplitudes during the Immediate and Middle Age periods are similar but the sample sizes are very different.

3 An economic growth model with random capital externality

We now propose a stochastic growth model which allows to capture the most relevant patterns of the previous data.

Consider the case in which the representative agent maximizes

$$E_0 \left[\sum_{t=0}^{+\infty} \beta^t \log \left(C_t - h N_t^{1+\mu} \right) \right]$$

with respect to consumption (C_t) and labor (N_t) paths, subject to the following constraints

$$\begin{aligned} C_t + I_t &= W_t N_t + R_t K_{t-1} + \pi_t \\ K_t &= A_K K_{t-1}^{\delta} I_t^{1-\delta} \end{aligned}$$

where π_t is the representative firms' profit $Y_t - W_t N_t - R_t K_{t-1}$ with $Y_t = A_t K_{t-1}^{\alpha} N_t^{1-\alpha} \overline{K}_{t-1}^{\gamma_t}$ the production level.

The first equation is the budget constraint (where W_t and R_t stand respectively for real wages and real interest rates) whereas the second equation reflects the accumulation of capital K_t. This equation is a slight variation of the usual linear case. Such a multiplicative structure has been proposed by Lucas and Prescott (1971) followed by Hercowitz and Sampson (1991) and Collard (1999). The parameter $0 < \delta < 1$ may be interpreted as a quality of installed capital (see above references for details). Also remark that this formulation may account for adjustment costs, the capital at date t being a concave function of the investment I_t. This multiplicative structure is useful to derive explicit dynamic structure of the relevant economic series (see below). The production function is of the standard Cobb-Douglas form with A_t the Solow residual, except for the presence of the term $\overline{K}_{t-1}$. It represents the average stock of capital available at the beginning of period t. It is considered as fixed by individual agents.

We assume $K_0 > 1$ is a given quantity.[3] We use the convention $x_t = \log(X_t)$ for every a.s. positive sequence X_t.

Equality between marginal disutility and marginal productivity of labor gives

$$h(1+\mu)N_t^{1+\mu} = (1-\alpha)Y_t.$$

Taking this relationship in the production function we derive

$$Y_t = \left(\frac{1-\alpha}{h(1+\mu)}\right)^{\frac{1-\alpha}{\alpha+\mu}} \left(A_t K_{t-1}^{\alpha} \overline{K}_{t-1}^{\gamma_t}\right)^{\frac{1+\mu}{\alpha+\mu}}.$$

We then recognize the well-known AK model. The term $\overline{K}$ plays a crucial role in the dynamic of growth. It creates a positive externality. Several justifications have been proposed including, learning-by-doing (see Lucas (1993) and Stokey (1988), among others), human capital (see Becker, Murphy and Tamura (1990) and Stokey (1991), among others).

An adaptation of Hercowitz and Sampson (1991) arguments shows that a possible solution is given by $Y_t = SI_t$ where $S = \alpha\beta(1-\delta)/(1-\delta/\beta)$. Moreover the law of motion for (the logarithm of) capital at the symmetric equilibrium is given by[4]

$$k_t = a_K + (1-\delta)s + \rho_t k_{t-1} + (1-\delta)\frac{1+\mu}{\alpha+\mu}a_t \tag{1}$$

where

$$\rho_t = \delta + \frac{(1-\delta)(1+\mu)}{\alpha+\mu}(\alpha+\gamma_t)$$

It should be stressed that although the above model looks very close to usual neoclassical

[3]The justification for this inequality will be given in Section 4. Notice it may be implied by a proper choice of unit of measurement as long as $K_0 > 0$. Now this choice may be tailored by appropriate choice of A_k.

[4]We assume $log(\overline{K}_t) = k_t$ which amounts to say that the externality arises from the un-weighted geometric average of individual stocks of capital. The only restriction here is that all individual stocks of capital must be strictly positive. It may not be very attractive an assumption if agents were heterogenous but it is harmless in the symmetric case.

growth models, the new dynamical feature introduced by time varying ρ_t coefficient makes a big change in both short and long run behavior of capital. Indeed, when ρ_t and a_t processes are independent, k_t provides an example of a structural Random Coefficient Autoregressive model (RCA hereafter). Granger and Swanson (1997) briefly mentioned some connections between Hall's (1978) permanent income model and the RCA process. They recall Hall's model leads to an AR(1) model in which the autoregressive coefficient is $(1-\beta)/R_t$ -using our notations-. In the stationary case where $R_t = R$ is constant the optimal decision of the consumer then implies $\beta = (1+R)^{-1}$. Granger and Swanson (1997) then claimed that allowing for randomness of the interest rate leads to an RCA model. However, the extra source of randomness modifies the optimal agent's decision, so their argument remains mainly heuristical from the economic viewpoint. In our model, agents observe the shocks and modify their behavior accordingly. More precisely, it may be shown that if $K_{t-1} > 1$ positive shocks on γ_t are alike positive technological shocks for individuals (*i.e.* employment, investment, consumption increase). Note however that positive shocks on γ_t and a_t have different dynamic consequences (see below).

Before we investigate the statistical properties of our model, we briefly derive some other relevant economic series. First consider growth. From $y_t = a_t + \alpha k_{t-1} + (1-\alpha)n_t$ we derive

$$y_t = \eta + \psi k_{t-1} + \frac{1+\mu}{\alpha+\mu} a_t$$

where

$$\eta = \tfrac{1-\alpha}{\alpha+\mu} \log\left(\tfrac{1-\alpha}{h(1+\mu)}\right)$$
$$\psi = \alpha \tfrac{1+\mu+\alpha}{\alpha+\mu}.$$

Using Equation (1) we get

$$y_{t+1} = \eta(1-\rho_t) + \rho_t y_t + \psi(a_K + (1-\delta)(s+a_t)) - \rho_t a_t + a_{t+1}. \tag{2}$$

In the usual framework (when ρ_t is fixed), the long term growth, computed for fixed $a_t =$

a, is constant and is a function of the parameters $\alpha, \delta, \mu, \gamma, a_k$ and s. The usual assumption is to fix a_t to its so-called "long term" value (which coincide with the marginal expectation of y_t in the stationary case). The growth rate then only depends on the autoregressive coefficient. If the external effects are strong enough *i.e.* if $\delta + (1-\delta)\alpha\frac{1+\mu+\alpha}{\alpha+\mu} + (1-\delta)\frac{1+\mu}{\alpha+\mu}\gamma$ is larger than 1), the long term growth rate is strictly positive. Notice k_t is then a non stationary sequence. Also remark that if γ_t is constant, the logarithm of GDP admits a (possibly non stationary) ARMA(1,1) structure whenever a_t is an iid process.

In our framework, the growth rate is affected by shocks even when a_t is fixed. Notice that varying ρ_t both affect the level and the dynamic of growth.

Now consider (real) prices. The maximization of private profits implies

$$
\begin{aligned}
r_t &= \ln(\alpha) + y_t - k_{t-1} = \ln(\alpha) + \eta + (\psi - 1)k_{t-1} + \tfrac{1+\mu}{\alpha+\mu}a_t \\
w_t &= \ln(1-\alpha) + y_t - n_t = \tfrac{\mu}{1+\mu}(y_t + \ln(1-\alpha)) + \tfrac{1}{1+\mu}\ln(h(1+\mu)).
\end{aligned}
$$

Hence real wages and real interest rates inherit from the dynamical properties of the capital. In particular, the previous computations explicitly show that when γ_t is constant the logarithm of real wages admit an ARMA(1,1) representation.

All the relevant economic series here follow similar types of processes. In the following, except otherwise mentioned, we then consider k_t only.

4 Statistical properties of economic processes governed by RCA models

Several authors (among which Granger and Swanson (1997), Abadir (2004), Nicolls and Quinn (1982)) already discussed some important statistical features of the RCA model. Most notably the time varying structure of *conditional* second order moments (both variance and auto-covariance functions) has received considerable attention. As our framework put the emphasis on growth, long term properties of the series are more relevant.

4.1 Existence of stationary solution

First consider the existence of a stationary solution to Equation (1). A stationary distribution clearly exists if $P(\rho_t < 1) = 1$ but then long term growth rate is zero (even if γ_t is constant). On the other hand, assume $P(\rho_1 > 1) = 1$ then clearly no positive stationary solution to Equation (1) exists. We shall then consider the case $1 > P(\rho_1 > 1) > 0$. From the economic viewpoint, this amounts to say that increasing returns to scale is a transitory phenomenon. Kesten (1973) is the first to derive conditions under which randomly varying dynamic models admit stationary solutions. For our purposes, we make extensive use of a paper by de Haan, Resnick, Rootzén and de Vries (1989). They consider the following model:

$$k_t = \rho_t k_{t-1} + b_t$$

where (ρ_t, b_t) for $t = 1, 2, \ldots$ is an iid drawing in some distribution taking values in R^2_{++}. The requirement $\rho_t > 0$ and $b_t > 0$ are easily justified. Indeed, if we have $k_{t-1} > 0$ and $\rho_t > 0$ the capital at date t is an increasing function of capital at date $t - 1$, a minimum requirement if we want to model growth. Now if $\rho_t > 0$ the conditions $b_t > 0, k_0 > 0$ are sufficient to guarantee $k_t > 0$ always.

More unusual is the iid assumption. Indeed, it is often assumed that the technological shocks are auto-correlated. Typically estimates of the autoregressive coefficient based on available short-time series are close to unit root which implies that persistence is always large (regardless of the positions in the cycle). Justifications for this structure invoke the time needed for diffusion of innovations, but this has no direct consequence on economic decisions. Moreover the argument is not very appealing in the medieval time where main source of shocks are related to natural conditions and the occurrence of conflicts. Some indirect consequences are also not very intuitive. For instance, the variance of stationary AR(1) model explodes as the autoregressive coefficient goes to one. Hence if agents are risk-averse welfare is a decreasing function of this autocorrelation coefficient. In other words, this dependence is not part of structural assumptions of the model. Assuming iid shocks implies that the dynamic features of the observable time series depends solely on agents decisions and explicit feasibility constraints.[5] The following results are established by de Haan, Resnick, Rootzén and de Vries (1989).

Proposition 1 *Assume there exist a constant* $\kappa > 0$ *such that* $E[\rho_1^\kappa] = 1$, $E[\rho_1^\kappa \log(\rho_1)^+]$ *exists and* $E[b_1^\kappa]$ *exists and is strictly positive. Also assume* $b_1/(1-\rho_1)$ *is non degenerate and the distribution of* $\log(\rho_1)$ *given* $\rho_1 > 0$ *is non lattice.*

i) The equation

$$k_\infty \stackrel{d}{=} \rho_1 k_\infty + b_1$$

admits a single solution.

ii) For all value of $k_0 > 0$ *any random sequence such that*

$$k_t = \rho_t k_{t-1} + b_t \, t = 1, 2, \ldots$$

converges in distribution to k_∞. *Moreover if* $k_0 = k_\infty$ *then* $k_0, k_1, \ldots$ *is stationary random sequence.*

iii) $\lim_{t\to+\infty} t^\kappa P(k_\infty > t) = c > 0$.

Let us now comment the economic implications of the above result.

Consider the assumptions. Notice that $E[\rho_1^\kappa] = 1$, together with $\kappa > 0$ implies $1 > P(\rho_1 > 1) > 0$. Indeed, assume $P(\rho_1 > 1) < 0$. Then there exist $\overline{\rho} < 1$ such that $P(\rho_1 < \overline{\rho}) = 1$. Now $E[\rho_1^\kappa] < \overline{\rho}^\kappa < 1$ for all $\kappa > 0$. On the other hand assume $P(\rho_1 > 1) = 1$. Then $E[\rho_1^\kappa] = 1$ implies $\kappa < 0$. Hence the assumptions of Haan, Resnick, Rootzén and de Vries (1989) are slightly stronger than what we already assumed.[6] Finally remark Jensen's inequality implies

[5] The assumption that the random sequence b_t is iid could be relaxed as long it remains stationary. For instance, the results presented in this section may easily be extended to the MA(1) case -as in the Equation (2).

[6] Notice the existence of the individual welfare implies $E\left[\sum_{t\geq 0} \beta^t k_t\right] < +\infty$ and $E\left[\sum_{t\geq 0} \beta^t b_t\right] < +\infty$. Also

$E[\log(\rho_1)] < 0$.

Results i) and ii) implies that our model display no (stochastic) indeterminacy.

Finally result iii) has two consequences. From an econometric viewpoint, one should be cautious when assuming the existence of high order moments of the series. Indeed, if $\kappa \leq 3$ then the random variable k_∞ has no second-order moment. From an economic viewpoint, this result implies that extreme events are likely to occur even if the shocks driving the economy are of moderate size. Consider for instance the case where both b_1 and ρ_1 are bounded random variables. Short run growth is bounded from above. However, somewhere in the future there will be a time where k_t will be as large as desired. Compare this to the case where γ_t is constant. If b_1 is bounded from above and $\rho_1 < 1$ then so is k_∞. Now if $\rho \geq 1$ then k_t is a non stationary sequence. Our model then allow for the capital sequence to reach arbitrarily large values while remaining stationary.

As arbitrarily large values of the capital are directly linked to the occurrence of extreme events, specific investigation of these is needed. We now give a very brief introduction to the main results on extreme behavior of processes. Readers familiar with Leadbetter and Rootzén (1988) or Leadbetter, Lindgren and Rootzén (1983) may skip the following two subsections whereas those interested in more detailed presentations are referred to these authors.

4.2 Extremal behavior of an iid random sequence

Consider an arbitrary real sequence $x_1, \ldots, x_n$ and define $M_n = \max_{1 \leq i \leq n} x_i$. The following property may be found for instance in Leadbetter, Lindgren and Rootzén (1983).

Proposition 2 *Let $x_1, \ldots, x_n$ be an iid sequence of real-valued random variables. If there exist two real deterministic sequences $\alpha_n > 0$, β_n and a non-degenerate distribution function*

it is easy to see that if we assume the existence of all the moments of b_t and that ρ_1 is a bounded random variable, the existence of a positive κ fulfilling the conditions of the previous proposition is equivalent to $1 > P(\rho_1 > 1) > 0$.

$G(.)$ *such that for all* x

$$\lim_{n\to+\infty} P\left(\alpha_n(M_n-\beta_n)\le x\right)=G(x)$$

then there exist some constants $\alpha>0,\mu>0$ *and* β *such that*

$$G(\mu x+\beta)=\begin{cases}\exp(-e^{-x}) & (type\, I)\\ or\ \exp(-x^{-\alpha})\mathbf{1}_{x>0} & (type\, II)\\ or\ \mathbf{1}_{x>0}+\exp(-(-x)^{-\alpha})\mathbf{1}_{x\le 0} & (type\, III)\end{cases}$$

This result resembles the central limit theorem except that the limit distribution is not Gaussian. Instances in which α_n and β_n sequences may be found are many. For instance, in the standard Gaussian distribution case if we set $\alpha_n=\sqrt{2\log(n)}$, $\beta_n=\alpha_n-\frac{\log\log(n)+\log(4\pi)}{2\alpha_n}$ then $\lim_{n\to+\infty}P\left(\alpha_n(M_n-\beta_n)\le x\right)=\exp(-e^{-x})$. The Cauchy and Pareto distributions are related to type II which is then associated to fat tail behavior. At the other extreme, bounded distributions (uniform, beta...) lead to the type III. Of course, the sequences α_n and β_n are different for all cases (see Leadbetter, Lindgren and Rootzén (1983), pp 19–27 for examples). There exist cases in which no sequence α_n and β_n exist such that proposition 2 applies. Classical examples are the Poisson and geometric distributions.

It should be stressed that the rate of convergence is also different from the central limit theorem. For instance, in the iid Gaussian case we have

$$P\left(\alpha_n(M_n-\beta_n)<x\right)-exp(-e^{-x})=\frac{exp(-e^{-x})e^{-x}}{16}\frac{(\log\log n)^2}{\log n}(1+o(1)).$$

Although $\lim_{n\to+\infty}\frac{(\log\log n)^2}{\log n}=0$ this term decreases only from 0.54 to 0.43 when n ranges from 10^2 to 10^{10}. Hence, for many practical issues, it may considered as constant. Note however that $\frac{exp(-e^{-x})e^{-x}}{16}<0.023$ so that the maximum difference in the Gaussian case is less than 5% for $n=10^3$.[7]

[7] Recall we will be mainly interested in the non iid case. As we shall see the limiting distributions are typically the same. Note however that the rate of convergence is often smaller in the depend case.

Similar results may be obtained for other order statistics. Trivially, by considering the sequence $-x_1, \ldots, -x_n$ the same result applies to the $\min_{1\leq i\leq n} x_i$. But less trivially, one may also consider the joint distribution of the maximum and the k-th largest values for $k = 2, 3...$ For instance let $M_n^{(1)} = M_n$ and $M_n^{(2)}$ be the second largest value of the sequence $x_1, \ldots, x_n$. If the conditions of Proposition 2 apply for the sequences α_n, β_n and for the distribution G we also have for any $x_1 > x_2$

$$\lim_{n\to+\infty} P(\alpha_n(M_n^{(1)} - \beta_n) < x_1, \alpha_n(M_n^{(2)} - \beta_n) < x_2) = G(x_2)(\log(G(x_1) - \log(G(x_2)) + 1).$$

In particular, this result implies that the second largest value of an iid sequence has the same limiting distribution and the same normalizing sequence α_n, β_n than the maximum.

This result may be further generalized. Consider a sequence u_n such that $\lim_{n\to+\infty} n(1 - F(u_n)) = \tau$. [8] Let $x_1, \ldots, x_n$ be an iid sequence such that $P(x_1 < x) = F(x)$. Define for any $E \subset]0, 1]$ the random quantity $N_n(E) = \#\{i : x_i > u_n, i/n \in E\}$. Then the random mapping $N_n(.)$ converges in distribution (as n goes to infinity) to a Poisson process on $[0, 1]$ with intensity τ. The extension of this result to the depend case will play an important role (see below).

4.3 Extreme values of depend processes

Extensions of the above results to the depend case have been obtained at the end of 70's and 80's. It is still a very active research domain. We now very briefly summarize most relevant available results. Obviously, in the depend case, some assumptions are needed otherwise if all random variables may be chosen equal, the maximum, the minimum and any other order statistic would have the same pre-specified distribution. In the following, we shall assume that $x_1, \ldots, x_n$ is a stationary strong mixing real-valued random sequence.[9] We then have

[8] Whenever the conditions of Proposition 1 apply, we have $\lim_{n\to+\infty} n(1 - F(u_n)) = -\log(G(x))$ if $u_n = \beta_n + x/\alpha_n$. Also, if $F(.)$ is continuous, we may compute $u_n(\tau)$ such that $n(1 - F(u_n(\tau))) = \tau$. Hence the existence of such a sequence u_n is not very demanding.

[9] Strong mixing is not needed, see the above quoted references for details.

the following

Proposition 3 *If $x_1, \ldots, x_n$ is a stationary strong mixing real-valued random sequence, if there exist two real deterministic sequences $\alpha_n > 0$, β_n and a non-degenerate distribution function $G(.)$ such that for all x*

$$\lim_{n \to +\infty} P\left(\alpha_n(M_n - \beta_n) \leq x\right) = G(x)$$

then there exist some constants $\alpha > 0, \mu > 0$ and β such that

$$G(\mu x + \beta) = \begin{cases} \exp(-e^{-x}) & (\textit{type}\, I) \\ \textit{or}\ \exp(-x^{-\alpha})\mathbf{1}_{x>0} & (\textit{type}\, II) \\ \textit{or}\ \mathbf{1}_{x>0} + \exp(-(-x)^{-\alpha})\mathbf{1}_{x\leq 0} & (\textit{type}\, III) \end{cases}$$

The first result then shows that moderate dependence does not modify the possible asymptotic distributions for the maximum of stationary random sequences. Though, it does not implies that dependence makes no difference.

Define the associated iid sequence $\hat{x}_1, \ldots \hat{x}_n$ such that $P(\hat{x}_1 \leq x) = P(x_1 \leq x) = F(x)$. Let $\hat{M}_n = \max_{1 \leq i \leq n} \hat{x}_i$. Further assume for each positive value τ we have a deterministic sequence $u_n(\tau)$ such that $\lim_{n \to +\infty} n(1 - F(u_n(\tau))) = \tau$. There exist $0 \leq \theta \leq 1$ such that for all τ we have

$$\lim_{n \to +\infty} P(M_n \leq u_n(\tau)) = exp(-\theta\tau).$$

Recall we have $\lim_{n \to +\infty} P(\hat{M}_n \leq u_n(\tau)) = exp(-\tau)$.

In such a case θ is the extremal index of the sequence $x_1, \ldots, x_n$. The underlying intuition to this result is that $\hat{M}_n$ is stochastically as large as M_n. Notice the extremal index equals 1 when the sequence is iid. Rather pathological cases in which $\theta = 0$ or no extremal index exists are available.

The notion of extremal index plays a key role in this literature. For instance, we have the following result.

Proposition 4 *If $x_1, \ldots, x_n$ is a stationary strong mixing real-valued random sequence with extremal index θ, and if there exist two real deterministic sequences $\alpha_n > 0$, β_n and a non-degenerate distribution function $G(.)$ such that for all x*

$$\lim_{n\to+\infty} P\left(\alpha_n(\hat{M}_n - \beta_n) \leq x\right) = G(x)$$

then

$$\lim_{n\to+\infty} P\left(\alpha_n(M_n - \beta_n) \leq x\right) = G^{\theta}(x)$$

and conversely.

Very roughly, this result shows that the maximum of a dependent stationary sequence of size n with extremal index θ behave as the maximum of an iid sequence with the same marginal distribution of size $n\theta$.

The results obtained in the iid case extend to the dependent case if the extremal index is 1. However, when $\theta < 1$ then large values of the sequence tend to appear at dates close to each others (the so-called clustering phenomenon). Hence the exceedances of a deterministic sequence u_n by the random sequence tends to appear in groups. When $\theta < 1$ the spatial process $N_n(.)$ does not converge to a Poisson process as in the iid case. Indeed, the paths of a Poisson process has jumps of size one (a.s.).

Let us be slightly more formal.[10] Consider a deterministic sequence r_n such that $\lim_{n\to+\infty} r_n = +\infty$ and $\lim_{n\to+\infty} r_n/n = 0$. Define for $i = 1, 2, \ldots$

$$\pi_n(i) = P\left(i = \sum_{j=1}^{r_n} \mathbf{1}_{x_j > u_n} \middle| 0 < \sum_{j=1}^{r_n} \mathbf{1}_{x_j > u_n}\right).$$

The quantity $\pi_n(i)$ is the probability that there is exactly i exceedance(s) of the sequence u_n by the random sequence $x_1, \ldots, x_{r_n}$ provided there is at least one. Typically if $u_n(\tau)$ is such that $\lim_{n\to+\infty} n(1 - F(u_n(\tau))) = \tau$ then in the iid case we would have $\lim_{n\to+\infty} \pi_n(1) = \tau$

[10] An illustrative example is provided below.

and $\lim_{n\to+\infty} \pi_n(i) = 0$ for all $i > 1$. The same result applies if the extremal index is one, but not when $\theta < 1$. More precisely, if $u_n(\tau)$ is such that $\lim_{n\to+\infty} n(1 - F(u_n(\tau))) = \tau$ and r_n is chosen such that the sequence $\pi_n(i)$ converges for all i then $N_n(](j-1)r_n/n, jr_n/n])$ converges to a *compound* Poisson process (more precisely in the limit the probability of jump in any interval of size $\epsilon \to 0$ is $\theta\tau\epsilon$ and, conditionally on the event that in such an interval a jump occurs its height is i with probability $\pi_i = \lim_{n\to+\infty} \pi_n(i)$).

Leadbetter (1983) mentions the following link between the extremal parameter θ and the compounding probabilities $(\pi_n(1), \pi_n(2), \ldots)$:

$$\lim_{n\to+\infty} \sum_{i=1}^{+\infty} i\pi_n(i) = \frac{1}{\theta}.$$

Hence, in the limit, the smaller θ, the larger the number of exceedances of the sequence $u_n(\tau)$ by the random sequence $x_1, \ldots, x_{r_n}$ *provided there exists at least one.* When θ is small Proposition 4 entails that the *average* waiting time to the next extreme event is larger than in the iid case. However, if an extreme event has been observed in the recent past, the average time for the next one is relatively smaller.

4.4 An illustrative example by Chernik

As the two previous subsections are rather involving, it may be interesting to consider a nice illustrative example due to Chernik (1961). Although it does not correspond exactly to the type of process we derived from our model, it is simple enough to grasp the main message.

Let $r \geq 2$ be an integer such that

$$x_{n+1}^{(r)} = x_n^{(r)}/r + \epsilon_n^{(r)}$$

where $\epsilon_1^{(r)}, \ldots,$ is an iid sequence such that $P(\epsilon_1^{(r)} = k/r) = 1/r$ for $k = 0, 1, 2, \ldots, r-1,$. Also assume that $x_0^{(r)}$ is uniform in $[0, 1]$ and $\epsilon_1^{(r)}, \ldots,$ and $x_0^{(r)}$ are independent. Then Chernik (1961) shows that this AR(1) process has an extremal index equal to $(r-1)/r$. Chernik's example provides a nice illustration of the clustering phenomenon. We have $x_{n+1}^{(r)} > x_n^{(r)}$ if and only if $\epsilon_n^{(r)} = (r-1)/r$. Now this event has probability $1/r$ no matter how large $x_n^{(r)}$ is. For instance if $r = 3$ the autocorrelation coefficient is only $1/3$, but the probability for four successive records is than $1/27$, *no matter how large the first one.* Now the probability for four successive *independent* values to form an increasing sequence is $1/(4!)$ which is larger than -though close to- $1/27$. But if the first value is large, the probability of this event drops drastically in the iid case (whereas it remains constant in Chernik's case).

As Chernik (1961) shows that $x_n^{(r)}$ is (marginally) uniformly distributed over $[0, 1]$ a direct computation is easy. If $x_1 = 0.9$ we have in the iid case

$$\begin{aligned} P(x_4 > x_3 > x_2 > x_1 | x_1 = 0.9) &= P(x_4 > 0.9) \times P(x_3 > 0.9) \times P(x_2 > 0.9) \times P(x_4 > x_3 > x_2) \\ &= 10^{-3}/(3!) \end{aligned}$$

which approximatively 222 times smaller than $1/27$ and when $x_1 = 0.99$ the probability of the same event is $10^{-6}/6$.

The previous computation illustrates the main ingredient behind the clustering phenomenon. A sequence of consecutive extreme events is always rare, but for some particular dependence structures (such as in this example) they are *relatively* much more frequent than in the iid case. This clustering of extreme event appears only if the extremal parameter is far enough from 1.

Moreover this process may be used to study the link between the clustering phenomenon and growth. Consider indeed the case where $x_1 = 0.01$. If we known that $x_4 > x_3 > x_2 > x_1$ then $x_4 \simeq 0.963$. Assume that some quantity is governed by a process similar to that of Chernik's. Then a historic minimum followed by consecutive "lucky" events creates a dramatic increase in the growth of this quantity.[11] Now if the sequence is particularly "unlucky" (that is to say $\epsilon_1^{(r)} = \epsilon_2^{(r)} = \epsilon_3^{(r)} = 0$) then $x_4 = x_1/27 = 0.00037$. The extremal

[11] This is not the case if the same quantity is governed by an iid uniform process. Indeed x_4 is then distributed as the maximum of three iid uniform drawings. Hence $P(x_4 < 0.963 | x_1 = 0.01, x_4 > x_3 > x_2 > x_1) \simeq 0.892$.

behavior of this depend process may then create a phenomenon closely resembling the "Great Divergence" evidence.

5 Inference for extreme behavior of economic processes

Inference of extreme behaviors is notoriously difficult (see Robert (2009) for a recent contribution). Of course as we deal with rare events we need very long series, but as the previous section makes clear the statistical theory is also involving (rate of convergence and limiting distributions are rather exotic and the existence of higher-order moments is not guaranteed).

However in our model, it is possible to derive explicitly the link between our "deep" parameters and fat tail parameter κ, the extremal parameter θ and the compounding probabilities $\pi_1, \ldots$. As the "deep" paramers are easier to estimate, we shall use them as an indirect way to infer the extreme parameters.

According to de Haan, Resnick, Rootzén and de Vries (1989) there exists $c > 0$ such that for all $x > 0$

$$\lim_{t \to +\infty} P(t^{-1/\kappa} \max_{0 \leq s \leq t} \{k_t\} \leq x) = exp(-c\theta x^{-\kappa})$$

where θ is the extremal index of the sequence $k_0, k_1, \ldots$. More precisely we also have

$$\theta = \int_0^1 P\left(\max_{0 \leq t}\{\prod_{s=1}^{t} \rho_s\} \leq w^{-1}\right) \kappa w^{-\kappa-1} dw$$

Finally, consider for $x > 0$ the number of exceedances of the sequence $xt^{1/\kappa}$ by k_t

$$N_t(I) = \#\{s/t \in I : k_s > xt^{1/\kappa}\}.$$

Then $N_t(.)$ converges to a compound Poisson process with intensity $c\theta x^{-\kappa}$ and compounding probabilities $\pi_k = (\theta_k - \theta_{k+1})/\theta$ with

$$\theta_k = \int_0^1 P\left(\#\{t : \prod_{s=1}^{t} \rho_s > w^{-1}\} = k - 1\right) \kappa w^{-\kappa-1} dw$$

for $k = 1, 2, \dots$

Observe that the extreme behavior of our sequence depends on the distribution of ρ_1 only (except perhaps for the constant c). As ρ_1 is explicit in our model, we are now in a position to compute $\kappa, theta, \pi_1, \pi_2, \dots$ as functions of the deep parameters.

5.1 Parameters of extreme behavior in the vicinity of the unit-root case

We first investigate this link in an analytic way. To this end, we consider a sequence of economies, indexed $n \in N$. Let $(\eta_n)_{n \in N}$ be a strictly decreasing sequence of real variables such that $\lim_{n \to +\infty} \eta_n = 0$ and fix $\eta_0 = 1$. The sequence of economies will be generated as follows. Starting from a given (inter-temporal) economy $\mathcal{E}_0$ with a given sequence $\rho_{1,0}, \rho_{2,0}$.. such that the conditions of Proposition 1 apply, we generate $\mathcal{E}_n$ as the (inter-temporal) economy with the sequence $\rho_{1,n}, \rho_{2,n}, \dots$, such that $\rho_{t,n} = \rho_{t,0}^{\eta_n}$ for all $t > 1$. By this device we generate a sequence of economies which may be as close as desired from the unit-root case and such that the conditions of Proposition 1 are always fulfilled. Focusing on the unit root case is important on two grounds. Empirically, the estimation of the autocorrelation coefficient is close to unit root (see Section 2 above and subsection 5.2 below). Theoretically, long-lasting growth in AK-models is related to efficient accumulation of capital. We are now in a position to study the sequence of extreme behaviors associated to a given sequence of economies.

For all $n = 0, 1, \dots,$ define $\kappa_n, \theta_n, \theta_{1,n}, \theta_{2,n}, \dots$ such that

$$E[\rho_{1,n}^{\kappa_n}] = 1,\ E[\rho_{1,n}^{\kappa_n} \log(\rho_{1,n})^+] < +\infty,\ E[b_1^{\kappa_n}] < +\infty$$
$$\theta_n = \int_0^1 P\left(\max_{0\le t}\{\prod_{s=1}^t \rho_{s,n}\} \le w^{-1}\right) \kappa_n w^{-\kappa_n - 1} dw$$
$$\theta_{k,n} = \int_0^1 P\left(\#\{t : \prod_{s=1}^t \rho_{s,n} > w^{-1}\} = k-1\right) \kappa_n w^{-\kappa_n - 1} dw.$$

We shall also assume that $E[b_1^k]$ exists for all $k > \kappa_0$.

We then have the following proposition

Proposition 5 *For all $n = 0, 1, \dots,$*

i) $\kappa_n = \kappa_0/\eta_n$

ii) $\theta_n = \theta_0$

iii) $\theta_{k,n} = \theta_{k,0}$.

Proof

i) By the definition of κ_n and $\rho_{1,n}$ we have $1 = E[\rho_{1,n}^{\kappa_n}] = E[\rho_{1,0}^{\eta_n\kappa_n}]$. As the Proposition 1 implies $E[\log(\rho_{1,0})] < 0$, the quantity κ_0 such that $1 = E[\rho_{1,0}^{\eta_n\kappa_n}]$ is unique. Hence $\eta_n\kappa_n = \kappa_0$.

ii) Following de Haan, Resnick, Rootzén and de Vries (1989) we have

$$1 - \theta_n = E\left[\mathbf{1}_{\max_{1\le t}\sum_{i=1}^t \log(\rho_{n,i}) < -e_{\kappa_n}}\right]$$

where for all integer n we have $P(e_{\kappa_n} < x|\rho_{n,1}, \rho_{n,2}, \dots) = 1 - \exp(-\kappa_n x)\mathbf{1}_{x\ge 0}$. We then may write

$$1 - \theta_n = E\left[\mathbf{1}_{\max_{1\le t}\sum_{i=1}^t \log(\rho_{n,i}) < \log(u)/\kappa_n}\right]$$

where $P(u < x | \rho_{n,1}, \rho_{n,2}, \ldots) = x\mathbf{1}_{1 \geq x \geq 0}$. Now as $\kappa_n = \kappa_0/\eta_n$ we have

$$\begin{aligned}
1 - \theta_n &= E\left[\mathbf{1}_{\kappa_0 \max_{1\leq t} \sum_{i=1}^{t} \log(\rho_{n,i})/\eta_n < \log(u)}\right] \\
&= E\left[\mathbf{1}_{\kappa_0 \max_{1\leq t} \sum_{i=1}^{t} \log(\rho_{0,i}) < \log(u)}\right] \\
&= E\left[\mathbf{1}_{\max_{1\leq t} \sum_{i=1}^{t} \log(\rho_{0,i}) < -e_{\kappa_0}}\right] \\
&= 1 - \theta_0.
\end{aligned}$$

The result *iii*) above follows along similar lines as *ii*). □

Results *ii*) and *iii*) of proposition 5 imply that clustering phenomena remain unchanged as we approach the unit root case. This show that "exotic" extreme behavior are compatible with arbitrarily small externality effects and do not require episodes of implausible high returns to scale. Result *i*) implies that fat tails behavior are less important when we approach the unit root case since κ_n appears as an increasing sequence.[12] This does not imply however that fat tail behavior are incompatible with "close to unit-root" processes. Indeed, consider the case where $\log(\rho_1) = \nu + \eta u$ where $P(u < x) = (x+1)\mathbf{1}_{x \in [-1,1]}$. The condition $E[\log(\rho_1)] < 0$ implies $\nu < 0$ whereas the condition $P(\rho_1 > 1) > 0$ implies $\nu + \eta > 0$. It may then be shown that κ is the unique strictly positive solution of the equation

$$1 = \frac{\exp(\iota\zeta)\sinh(\iota)}{\iota}$$

where $\iota = \eta\kappa$ and $-1 < \zeta = \nu/\eta < 0$. For instance if we consider the case $\nu = -0.000025$ and $\eta = 0.01$ (which implies $\rho_1 \in [0.99002, 1.01002]$) we get $\kappa \simeq 1.4996$. In this case, the stationary distribution of the logarithm of the capital admits no first-order moment.[13]

[12] Ultimately the sequence κ_n diverges if η_n goes to zero.

[13] Proposition 5 and the above case consider in fact different neighborhoods of the unit root case. More precisely the sequence studied in Proposition 5 would amount to consider a sequence ζ_n decreasing to -1 if η_n goes to zero, whereas in the above example $\zeta = -0.025$.

5.2 Inference on extreme parameters: real wages in England since 1264

We now use the data set presented in Section 2 to infer extreme parameters. Recall there exists a triplet (a_w, b_w, c_w) such that

$$w_t = \alpha_w + \beta_w k_{t-1} + \sigma_w a_t.$$

Now as

$$k_t = a_k + (1-\delta)s + \rho_t k_{t-1} + (1-\delta)\frac{1+\mu}{\alpha+\mu}a_t.$$

We get

$$\begin{aligned} w_t &= \alpha_w + \beta_w(a_k + (1-\delta)s + \rho_{t-1}k_{t-2} + (1-\delta)\tfrac{1+\mu}{\alpha+\mu}a_{t-1}) + \sigma_w a_t \\ &= \alpha'_w + \rho_{t-1}w_{t-1} - \alpha_w\rho_{t-1} - \sigma_w a_{t-1}\rho_{t-1} + \theta_w a_{t-1} + \sigma_w a_t. \end{aligned}$$

Hence w_t follows an ARMA(1,1) model with random coefficient both in the autoregressive and the moving average parts. A direct estimation of this model is difficult, but an approximate approach will be sufficient for our purposes. Denote $E[\rho_t] = \rho$, and $Var[\rho_t] = \sigma_\rho^2$ we may write $\rho_t = \rho + \sigma_\rho \epsilon_{\rho,t}$ where $\epsilon_{\rho,t}$ is an iid white noise with unit variance. Similarly we write $a_t = a + \epsilon_{a,t}$ where $E[a_t] = a$. We then have

$$w_t = \alpha''_w + (\rho + \sigma_\rho \epsilon_{\rho,t-1})w_{t-1} + \theta_{w,\rho}\epsilon_{\rho,t-1} + \theta_{w,\rho,a}\epsilon_{\rho,t-1}\epsilon_{a,t-1} + \theta_{w,a}\epsilon_{a,t-1} + \sigma_w \epsilon_{a,t}.$$

Remark that the second order term $\epsilon_{\rho,t-1}\epsilon_{a,t-1}$ induces an endogeneity problem. To avoid bias as much as possible, we then performed a Double Least Square estimation using w_{t-2} as an instrument for w_{t-1} (see Table 2).

Notice the Hausman test backs the suspicion for endogenous regressor (the p-value for the absence of bias for the MCO estimates being 0.002). Also remark that ρ is close to 1. Recall from the previous subsection that this is compatible with arbitrarily small values of κ and θ.

Table 2: Model 1

	Coefficient	Std. err.	z-stat.	p-value
intercept	0.033	0.028	1.161	0.246
w_{t-1}	0.993	0.007	139.235	0.000

Note: DLS Estimation with 743 observations
Dependent var.: w_t, Explanatory var.: w_{t-1}
Instruments: intercept, w_{t-2}, w_{t-1}.

Now we also have

$$(w_t - \rho w_{t-1})^2 = \sigma_\rho^2 \epsilon_{\rho,t-1}^2 w_{t-1}^2 + 2\sigma_\rho \epsilon_{\rho,t-1} \eta_{w,t} w_{t-1} + \eta_{w,t}^2$$

with $\eta_{w,t} = \alpha_w'' + \theta_{w,\rho,a}\epsilon_{\rho,t-1}\epsilon_{a,t-1} + \theta_{w,a}\epsilon_{a,t-1} + \sigma_w \epsilon_{a,t}$.

Again running a DLS regression with w_{t-3}, w_{t-3}^2 as instruments [14] we get

Table 3: Model 2

	Coefficient	Std. err.	z-stat	p-value
intercept.	0.275	0.050	5.462	0.000
w_{t-1}	-0.108	0.023	-4.702	0.000
w_{t-1}^2	0.010	0.003	4.124	0.000

Note: DLS Estimation with 742 observations
Dependent var.: $(w_t - \rho w_{t-1})^2$, Explanatory vars. : w_{t-1}, w_{t-1}^2
Instruments: const, w_{t-3}, w_{t-3}^2.

It is interesting to notice that the p-value associated to the F-test for this second model is $1.89e-11$ hence the first order residuals if we assume ρ_t constant clearly displays time dependent heteroscedasticity. Notice also that the absence of randomness of ρ_t is clearly rejected.

All together we then obtain $\hat{\rho} = 0.992$ and $\hat{\sigma}_\rho = 0.0322$.

First, these estimates may be used to evaluate the probability of increasing returns to scale. Recall in our setting ρ_t is an iid sequence, so we are typically interested in $P(\rho_1 > 1)$.

[14] We prefer to use one more lag for the instrument because of the presence of w_{t-1} on the left hand side.

The Chebychev inequality may be used to compute a lower bound for this probability. More precisely if $P(\overline{\rho} > \rho_1 > \underline{\rho}) = 1$ we have

$$\frac{\rho^2 + \sigma_\rho^2 + 2\underline{\rho}(1-\rho) - 1}{\overline{\rho} - \underline{\rho}} \leq P(\rho_1 > 1).$$

A data-based interval such as $\underline{\rho} = 0.4$ and $\overline{\rho} = 1.2$ seems very reasonable. We then obtain $P(\rho_1 > 1) > 0.001$. This non parametric lower bound may look small, but the actual value of $P(\rho_1 > 1)$ may be much larger depending on the distribution of ρ_1. For instance If the $log(\rho_1)$ is uniformly distributed over $[\nu, \nu + \eta]$ then this probability is 16.5%. Using the same parametric specification we compute $\hat{\kappa} = 2.53$ (a figure which cast doubts on the existence of moments larger then 3).

Finally, the extremal coefficient θ and the compounding probabilities $\pi_1, \ldots,$ may be derived from the same specification. Using the simulation techniques proposed in de Haan, Resnick, Rootzén and de Vries (1989) we obtain

$\hat{\theta}$	$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\pi}_3$	$\hat{\pi}_4$	$\hat{\pi}_5$	$\hat{\pi}_6$	$\hat{\pi}_7$	$\hat{\pi}_8$
0.611	0.418	0.035	0.058	0.022	0.015	0.016	0.027	0.009

If we identify major crises and/or boom as a consecutive sequence of exceptionally (un)lucky times lasting three years or more, the above figures are somewhat consistent with actual observation that these appear once or twice in a century.[15]

6 Conclusion

In this paper, we study a growth model in which returns to scale vary randomly. The dynamics of main aggregate economic series including capital, growth, various relative prices are all related to a RCA model. A nice feature of this model is that growth may be combined with the existence of a stationary solution. Recent works on extreme behavior on dependent process emphasize some properties of this model that are worth noting from the economic

[15] Also remark that the extreme parameter appear close to Chernik's example (in which it equals 2/3).

viewpoint. First, aggregate economic series typically display fat tail behavior even if the shocks do not. Second, historically high records are rare but tend to appear in cluster. We show that both fat tails and clustering of extreme values are consistent with arbitrarily small variations of the autoregressive coefficient around the usual unit-root case. Distinguishing RCA from a more usual constant returns to scale model is difficult with available macro data and typically require very long data set. To this end, we propose a direct test based on the annual sequence of real wages in England recorded since the XIII-th century onwards. The test clearly rejects the absence of randomness in the returns to scale and supports the RCA model. Finally this data is used to estimate the parameters associated to the extreme behavior of the series and to evaluate the importance of the clustering phenomenon.

References

Barro, R. (2006), "Rare Disasters and Asset Markets in the Twentieth Century", *Quarterly Journal of Economics*, **121**, 823–866.

Barro, R. (2009), "Rare Disasters, Asset Prices, and Welfare Costs", *American Economic Review*, **99**, 243–264.

Barro, R. and J. Ursua (2008), "Macroeconomic Crises since 1870", *Brookings Papers on Economic Activity*, **99**, 255–350.

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

Chernik, M. (1981), "A Limit Theorem for the Maximum of Autoregressive Processes with Uniform Marginal Distributions", *Annals of Probability*, **9**, 145–149.

Collard, F. (1999), "Spectral and Persistence Properties of Cyclical Growth", *Journal of Economic Dynamic and Control*, **23**, 463–488.

DeJong, D., J. Nankervis, N. Savin and C. Whiteman (1992), "Integration Versus Trend

Stationary in Time Series", *Econometrica*, **60**, 423–433.

Distaso, W. (2008), "Testing for unit root processes in random coefficient autoregressive models", *Journal of Econometrics*, **142**, 581–609.

Engle, R. (1982), "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation", *Econometrica*, **50**, 987–1008.

Galor, O. (2005), "From Stagnation to Growth: Unified Growth Theory, *Handbook of Economic Growth*, 171–293.

Galor, O. (2007), "Multiple Growth Regimes – Insights from Unified Growth Theory", *Journal of Macroeconomics*, **29**, 470–475.

Galor, O. and O. Moav (2004), "From Physical to Human Capital Accumulation: Inequality in the Process of Development", *Review of Economic Studies*, **71**, 1001–1026.

Galor, O., O. Moav and D. Vollrath (2009), "Inequality in Land ownership, the Emergence of Human Capital Promoting Institutions and the Great Divergence", *Review of Economic Studies*, **76**, 143–179.

Galor, O. and D. Weil (1999), "From Malthusian Stagnation to Modern Growth", *American Economic Review*, **89**, 150–154.

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

Granger, C. (1980), "Long Memory Relationships and the Aggregation of dynamic Models", *Journal of Econometrics*, **14**, 227–238.

Granger, C. and N. Swanson (1997), "An Introduction to Stochastic Unit Root Processes", *Journal of Econometrics*, **80**, 35–62.

de Haan, L., S. Resnick, H. Rootzén, and C. de Vries (1989), "Extremal behaviour of solutions to a stochastic difference equation, with applications to ARCH-processes" *Stochastic*

Processes and their Applications, **32**, 213–224.

Hall, R. (1978), "Stochastic implications of the life cycle-permanent income hypothesis", *Journal of Political Economy*, **86**, 971–87

Hercowitz, Z. and M. Sampson (1991), "Output Growth, the Real Wage, and Employment Fluctuations", *American Economic Review*, **81**, 1215–1237.

Jaeger, A. (1990), "Shock Persistence and the Measurement of Prewar Output Series", *Economics Letters*, **34**, 333–337.

Kesten, H. (1973), "Random difference equations and Renewal theory for products of random matrices", *Acta Mathematica*, **131**, 207–248.

Leadbetter, M. (1983), "Extremes and local dependence in stationary sequences", *Zeit. Wahrsch. und Ver. Geb.*, **65**, 291–306.

Leadbetter, M. and H. Rootzén (1988), "Extremal Theory for Stochastic Processes", *The Annals of Probability*, **16**, 431–478.

Leadbetter, M., G. Lindgren and H. Rootzén (1983), *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, N.Y.

Lucas, R. (1993), "Making a Miracle", *Econometrica*, **61**, 251–272.

Lucas, R. and E. Prescott (1971), "Investment under uncertainty", *Econometrica*, **39**, 659–681.

Maddison, A. (2000), *The World Economy: Historical Statistics*, Paris: Organization for Economic Cooperation and Development.

Nicholls, D. and B. Quinn (1982), *Random Coefficient Autoregressive Models: An Introduction*, Lecture Notes in Statistics vol. 11, Springer, N.Y.

Perron, P. and P. Phillips (1987), "Does GNP have a Unit Root? A Reevaluation", *Economics Letters*, **23**, 139–145.

Robert, C. (2009), "Asymptotic distributions for the intervals estimators of the extremal index and the cluster-size probabilities", *Journal of Statistical Planning and Inference*, **139**, 3288–3309.

Rootzén, H. (1978), "Extremes of moving averages of stable processes", *Annals of Probability*, **6**, 847–869.

Shiller, R. and P. Perron (1985), "Testing the Random Walk Hypothesis: Power Versus Frequency of Observation", *Economics Letters*, **18**, 381–386.

Stokey, N. (1988), "Learning by Doing and the Introduction of New Goods", *The Journal of Political Economy*, **96**, 701–717.

Stokey, N. (1991), "Human Capital, Product Quality, and Growth", *The Quarterly Journal of Economics*, **106**, 587–616.