Centre de Recherche en Économie et Statistique

CREST

Série des Documents de Travail

n° 2012-29

Modelling Tails of Aggregated Economic Processes in a Stochastic Growth Model

S. AURAY[1] – **A. EYQUEM**[2]
F. JOUNEAU-SION[3]

October 2012

[1] CREST-ENSAI, EQUIPPE (EA 4018), Université du Littoral, France, and GREDI, CIRPEE, Canada.

[2] Université de Lyon, Ecole Normale Supérieure de Lyon, CNRS (GATE UMR 5824), France and GREDI, Canada.

[3] EQUIPPE (EA 4018), Université Lille Nord-de-France.

Modelling Tails of Aggregated Economic Processes in a Stochastic Growth Model*

Stéphane Auray† Aurélien Eyquem‡ Frédéric Jouneau-Sion§

October 2012

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.[1]

(Exodus, 16 21-22)

Abstract

We present an annual sequence of wages in England starting in 1245. We show that a standard AK-type growth model with capital externality and stochastic productivity shocks is unable to explain important features of the data. We then consider random returns to scale. Moderate episodes of increasing returns to scale and growth are shown to be compatible with stationarity. Further, random returns to scale generate heteroskedasticity, a feature common to macroeconomic time series. Third, stationary distributions display fat tails if returns to scale are episodically increasing. We provide several inference results to support randomness of returns to scale.

Keywords: Economic Growth, Unified Growth Theory, Heteroskedasticity, Fat Tails.

JEL Classification: C22, C46, N13, O41, O47.

---

*We would like to thank Philippe Barbe, Christian Francq, Xavier Gabaix, Paul Gomme, Markus Poschke and Jean–Michel Zakoian for insightful comments, as well as Guillaume Daudin for providing us with the dataset.

†CREST-Ensai, EQUIPPE (EA 4018), Université du Littoral, France and, GREDI, CIRPEE, Canada. Email: stephane.auray@ensai.fr.

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

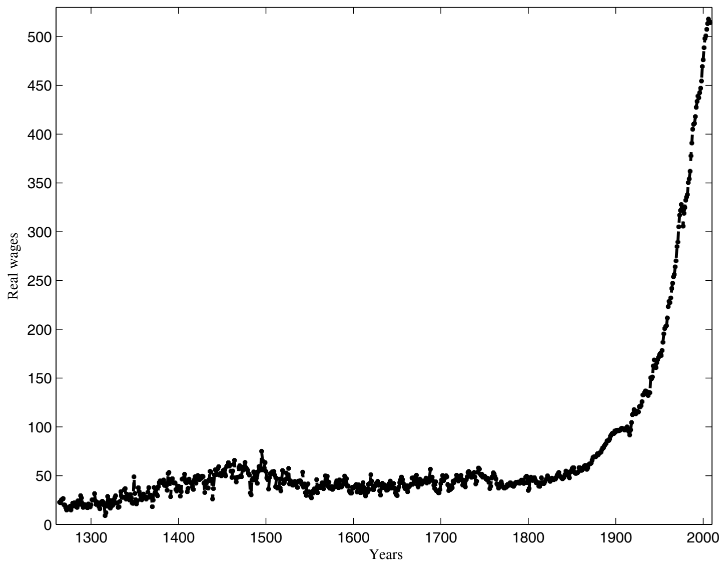
§EQUIPPE (EA 4018), Universités Lille Nord-de-France. Email: frederic.jouneau@univ-lille3.fr.

[1] Morning by morning they gathered it, each as much as he could eat; but when the sun grew hot, it melted. On the sixth day they gathered twice as much bread, two omers each. (translation from the English standard Version).

1 Introduction

In this paper, we make use of one of the longest available macroeconomic dataset: an annual sequence of real wages in England starting in 1245 reported in Figure 1.[2]

Figure 1: Time series of real wages (England 1245–2010)



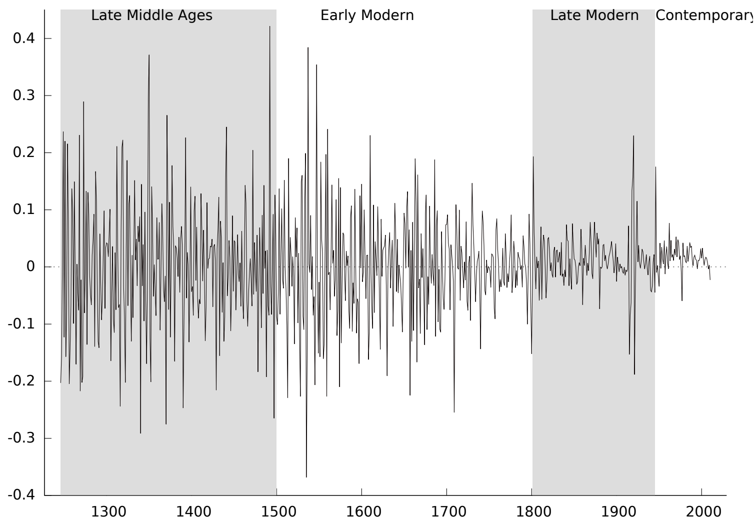
The most striking feature of this sequence is the dramatic increase in real wages during the XIX-th and XX-th centuries. Wage growth in the oldest part of the sample is undoubtedly smaller. Similar patterns may be obtained for other developed countries and other time series, such as GDP (see Maddison [2000]).

The second striking feature is the heteroskedasticity of this time series. In Figure 2, we plot the increment of the log of real wages, a measure of the annual real growth rate.

Obviously, the variability of the growth rate during Late Middle Ages (1245–1500) and Modern period (1500–1800) is much larger than in the most recent times. This high volatility of wage growth in the Middle Ages can mostly be explained by variations in nominal prices in the agricultural sector. During this period, climatic events and conflicts explain the erratic fluctuations in agricultural returns causing large fluctuations in wages. However, variations in the wage growth rate can hardly be considered as being caused by purely random short-term factors, as periods of large or low volatility seem to be quite persistent over time.

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

Figure 2: Annual Growth (first difference of log of real wages)



In the literature, it is usual to decompose such time series into a trend and a transitory component.[3] By and large, Growth Theory focuses on the trend component, whereas transitory components are the scope of business cycle models. Variations in the trend component most often rely on separate explanations of the so-called Malthusian, Post-Malthusian and Modern Growth Regimes. The analysis of business cycles is conducted once the trend component has been removed. The current Cycle Theory views business cycles as the consequence of the transmission of exogenous shocks across sectors, GDP components, and/or countries. Noticeably, the business cycle literature almost exclusively focuses on post-war data.

Recently however, the so-called 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) has proposed a single law of motion instead of using different ones to fit various historical episodes. In addition, an alternative and original approach to the understanding of business cycles has been proposed by Barro [2006], providing a middle-term perspective on economic crises. This proposal analyzes a large international panel of long time series that includes large crisis ("disasters") and allows to analyze their impact on the equity premium puzzle, among other things.

In this paper, we propose a structural model that allows for a joint analysis of variations in the level and the variability of growth in a long-term perspective. We eschew the decomposition into trend and transitory components, and confront the model predictions

[3]Many procedures and approaches have been proposed to obtain such decompositions, among which filtering and the co–integration are currently the most popular.

to the entire dynamics of wages.

The model is a modified version of the RBC model initially proposed by Hercowitz and Sampson [1991], that has the advantage to allow for a closed-form solution. As usual in the endogenous growth literature, an externality generated by capital accumulation generates growth. The originality of the model is to allow the strength of this externality to vary over time. We then derive the optimal response of agents to Total Factor Productivity and externality shocks. In the general case, we show that the dynamics of wages admits a Random–Coefficients ARMA representation, a particular case of Kesten's process (see Kesten [1973]), if the TFP process admits a strong ARMA representation. In the special case of a constant externality on capital accumulation, the series of real wages admits a strong ARMA representation with fixed coefficients.

The model has several interesting implications. First, we provide conditions under which the macroeconomic quantities admit stationary distributions in the long run. Second, we show that agent's optimal response to iid TFP and externality shocks generates dynamic heteroskedasticity. In the model, positive externality shocks can induce increasing returns to scale and result in recurrent transitory growth episodes. We derive analytical conditions under which these episodes are consistent with stationarity in the long run. Long-run stationarity emerges as long as increasing returns to scale are not too frequent. Third, the stationary sequences typically display fat tails, even if shocks are bounded. We show that fat tails are a direct consequence of episodically increasing returns to scale. Importantly, the model is plainly compatible with the vast empirical macroeconomic literature on random-walk hypothesis. More precisely we show that very small, and ultimately statistically insignificant, variations around the unit-root case are compatible with stationary and fat-tailed distributions of aggregate time series.

Further, the model also reconciles theory with the data. If we fit the data with a fixed coefficients ARMA model, the following empirical results emerge. First, ARMA estimates derived from the whole sample significantly differ from those obtained with sub-samples corresponding to the usual historical steps of economic development. Second, the sequence of residuals displays significant conditional heteroskedasticity. Both features reject stable fixed coefficient representations of economic processes in the long run. They are more consistent with the representation of the growth process we propose, in which intensity of the externality varies over time.

According to our model, if the returns to scale are episodically increasing, the limiting distributions of several aggregate macroeconomic time series (and, in particular GDP, real wages and capital stock) exhibit fat tails. While empirical evidence for fat tails are numerous in macroeconometrics (Engle [1982] originally tested ARCH models on macroeconomic series), structural explanations in dynamic models are much fewer. Granger [1980] structural explanations for long memory process generated by the aggregation of heterogeneous economic series is an important exception. Economic explanations for heavy tails in cross–

sectional data have been provided by Gabaix [1999]. Various occurrences of heavy tails in economics may be also be found in Gabaix [2009] whereas Gabaix, Gopikrishnan, Plerou and Stanley [2006] propose micro–structure explanations for heavy tails in finance. Recently Fagiolo, Napoletano and Roventini [2008] also investigate the presence of fat tails in post-war output data. The presence of fat tails and/or conditional heteroskedasticity has important consequences, as it precludes analysis relying on approximations in the neighborhood of the steady state. Because our model admits a closed–form representation, our ability to track the exact dynamics of the economy does not depend on the size of shocks. We consider this as an important advantage of our approach. Ultimately, the economic structure of the model may be used to derive inference on the fat index as it is directly linked to the occurrence of increasing returns to scale.

The paper is structured as follows. In Section 2, we describe a stochastic AK growth model augmented with neutral productivity shocks and a non-neutral source of shocks that affects the intensity of the externality on capital accumulation. We also derive the dynamics of endogenous macroeconomic series (that is capital, output, real wages and growth). In Section 3, we study the long term behavior of our series. In particular, we derive conditions for convergence to stationary sequences, we investigate the tails of the distributions, and stress the differences between our model and the fixed AR(I)MA representation implied by the standard endogenous growth model with constant externality. In section 4, we use our data to present evidence about parameter instability of ARMA representations and conditional heteroskedasticity. We also present various approaches to estimate the tail index of our data. They all provide similar results, with a fat tail index of 4 or 5. Section 5 concludes.

2 A simple growth model with random shocks

We describe a simple model that allows for endogenous growth through an externality on capital accumulation. One key feature of this model is to allow explicit solution without relying on the assumption that the data generating process remains in a small neighborhood around a deterministic steady state. As we explain in the introduction, we are mainly interested in extreme behavior of the endogenous aggregated macro series. Hence, our approach should allow for arbitrary large variations. This extreme behavior results from varying externalities. In the case of fixed externalities, the aggregated macro series admits a strong ARMA representation. If the strength of externalities varies, our model predicts that ARMA-based inference will produce instable estimates and heteroskedasticity. The empirical plausibility of these predictions are investigated in the Section 4.

2.1 The model

We consider a representative agent with infinite life maximizing[4]

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t) + \chi \log(1 - N_t)\right), \tag{1}$$

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

$$\begin{aligned} C_t + I_t &= W_t N_t + R_t K_{t-1} + \pi_t, && (2) \\ K_t &= A_K K_{t-1}^{\delta} I_t^{1-\delta}. && (3) \end{aligned}$$

The first equation is the budget constraint, where W_t and R_t stand respectively for real wages and real interest rates. In this equation, π_t is the representative firms' profit $Y_t - W_t N_t - R_t K_{t-1}$, where

$$Y_t = A_t K_{t-1}^{\alpha} N_t^{1-\alpha} \overline{K}_{t-1}^{\gamma_t}, \tag{4}$$

is the aggregate production function. The production function is a standard Cobb–Douglas function where A_t is the neutral Total Factor Productivity shock, except for the presence of the term $\overline{K}_{t-1}$. This term represents an externality generated by the average stock of capital available at the beginning of period t and is, as such, considered as fixed by individual agents in their optimization. The model is thus a variation of the well–known AK model. We do not rationalize explicitly the presence of the externality but several justifications have been proposed including learning–by–doing (see Lucas [1993] and Stokey [1988], among others) or human capital (see Becker, Murphy and Tamura [1990] and Stokey [1991], among others). In the literature, it is generally assumed that γ_t is a constant parameter. Our main message derives from the fact that $Var[\gamma_t] > 0$, *i.e.* from the introduction of time–varying externalities.

Equation (3) reflects the accumulation of capital. This multiplicative law of capital accumulation proposed by Lucas and Prescott [1971] followed by Hercowitz and Sampson [1991] and Collard [1999], is a slight variation of the usual linear case. The parameter $0 < \delta < 1$ can be interpreted as the quality of the installed stock of capital (see above references for details). This formulation can also account for the presence of adjustment costs, as the capital stock at time t is a concave function of investment I_t. We use this particular model because it admits a closed-form solution and allows an explicit derivation of the dynamic structure of economic aggregates.

[4]The parameter β is the usual discounting factor and we assume $1 > \beta > 0$. The parameter $\chi > 0$ controls the intensity of the demand for leisure.

2.2 Closed-form representations for the laws of motion

First-order conditions are

$$\begin{aligned}
\frac{1}{C_t} &= \frac{\lambda_t K_t(1-\delta)}{I_t}, &(5)\\
\frac{\chi N_t}{1-N_t} &= \frac{\lambda_t K_t(1-\delta)}{I_t}(1-\alpha)Y_t, &(6)\\
\lambda_t K_t &= \beta E_t\left[\lambda_{t+1}K_{t+1}\left(\delta+(1-\delta)\alpha\frac{Y_{t+1}}{I_{t+1}}\right)\right]. &(7)
\end{aligned}$$

A possible solution to the rational equation system is given by assuming $\lambda_t K_t$ is constant and $I_t = SY_t$ where the (constant) saving rate is

$$S = \frac{\alpha\beta(1-\delta)}{1-\delta\beta}. \tag{8}$$

The second first order condition (Equation (7)) implies $N_t = N$, since the RHS is also constant. Looking at symmetrical solutions in which each agent chooses the same level of investment, so that average and individual stocks of capital coincide, we derive the following system of three equations linking the capital stock (k_t), output (y_t), and the real wage (w_t) to the exogenous TFP process (a_t) and the externality process (γ_t):

$$\begin{aligned}
k_t &= a_k + (1-\delta)s + \delta k_{t-1} + (1-\delta)y_t, &(9)\\
y_t &= (\alpha+\gamma_t)k_{t-1} + a_t + (1-\alpha)n, &(10)\\
w_t &= \log(1-\alpha) + y_t - n. &(11)
\end{aligned}$$

Using the production function to substitute for y_t, we get

$$\begin{aligned}
k_t &= a_k + (1-\delta)s + \delta k_{t-1} + (1-\delta)((\alpha+\gamma_t)k_{t-1} + a_t + (1-\alpha)n) &(12)\\
&= a_k + (1-\delta)(s + (1-\alpha)n + a_t) + (\delta + (1-\delta)(\alpha+\gamma_t))k_{t-1}, &(13)\\
w_t &= \log(1-\alpha) + (\alpha+\gamma_t)k_{t-1} + a_t - \alpha n. &(14)\\
& &(15)
\end{aligned}$$

Assuming $\alpha + \gamma_t = 0$ is a zero probability event, both equations combine to yield

$$w_{t+1} = (\log(1-\alpha) - \alpha n)(1-\phi_{t+1}) + \phi_{t+1}w_t + a_{t+1} - \delta\theta_{t+1}a_t, \tag{16}$$

where $\theta_t = \frac{\alpha+\gamma_t}{\alpha+\gamma_{t-1}}$ and $\phi_t = (\delta + (1-\delta)(\alpha+\gamma_{t-1}))\theta_t$.

If γ_t is constant, this dynamics simplifies to

$$w_{t+1} = (\log(1-\alpha) - \alpha n)(1-\delta)(1-\alpha-\gamma) + (\delta + (1-\delta)(\alpha+\gamma))w_t + a_{t+1} - \delta a_t. \tag{17}$$

In this case, if a_t admits a strong ARMA(p,q) representation, the process w_t admits a strong ARMA$(p+1, q+1)$ representation. In the general case, w_t follows a Random Coefficients ARMA process.

3 Long term behavior

We now characterize the long term behavior of our aggregate macroeconomic times series. In a first subsection, we derive sufficient conditions for the convergence in distribution of the relevant quantities (*i.e.* output, wages, the capital stock, growth).[5] In case a stationary distribution can be derived, we study the question of the tails of these random variables in a second subsection. More precisely, we deal with the so-called extreme behavior, *i.e.* the dominant term of the cumulative density function in the neighborhood of $+\infty$. In a third subsection, we comment the economic plausibility of our assumptions. In particular, we stress the difference between variable *versus* fixed externality on capital accumulation. We also address the reliability of usual stationarity/unit root test in the context of Random Coefficient ARMA models.

3.1 Sufficient conditions for stationarity

In our model, the asymptotic behavior of all macroeconomic sequences is ultimately driven by the asymptotic behavior of the capital stock. We first establish a sufficient condition for the convergence of relevant aggregated variables towards a stationary distributions. We then shows that these distribution may display fat tail behavior. In particular, we show that arbitrarily small variations in the externality may lead to fat tails while being compatible with the existence of a stationary distribution in the long term.

As it is usual in the macro-literature, we assume that the TFP process a_t converges to a stationary sequence. More precisely we suppose that the TFP process a_t admits a strong stationary AR(2) representation.[6] Define $a'_t = a_k + (1-\delta)(a_t + (1-\alpha)n + s)$. The processes a'_t and a_t share the same strong AR(1) representation:

$$a'_t = \phi_1 a'_{t-1} + \phi_2 a'_{t-2} + \epsilon_t, \tag{18}$$

where ϵ_t is an iid process.

Proposition 1. *Assume the following conditions hold*

1. $\psi_t = \delta + (1-\delta)(\alpha + \gamma_t)$ *is an iid process such that* $P(\psi_1 > 0) = 1$ *and* $E[\log(\psi_1)] < 0$.
2. k_0 *is a given quantity.*
3. $E[\log^+(\epsilon_1)] < \infty$.
4. $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$, $|\phi_2| < 1$.

[5] In this model, hours worked are constant. Hence, the long term behavior is trivial.

[6] As carefull reading of the proofs reveals, the results of this section may also be established – at the cost of much heavier notation – when a_t follows strong strictly stationary AR(p) process for any given value of p. As the AR(2) assumption is supported by our data, we restrict our attention to this case. Also remark that the AR(2) case may cope for the possible presence of – exogenous – cycles in the TFP process.

Then k_t admits a unique limiting distribution as t goes to infinity. The same result applies to y_t and w_t.

Proof. See Appendix A. □

Notice the above result may be established under weaker conditions. For instance we could simply assume that a_t is a stationary process such that $\lim_{t \to +\infty} E[\log^+(a_t)] < \infty$.

3.2 Extreme behavior

From now on, we assume that conditions of Proposition 1 hold. Let k_∞ be the limiting random variable associated with the process k_t. We show that if $P(\psi_1 > 1) > 0$ then k_∞ displays fat tail behavior. More formally we have:

$$\lim_{x \to +\infty} x^\kappa P(k_\infty > x) = l > 0. \tag{19}$$

Of course, under the conditions of proposition 1 the capital stock would also admit fat tail if $P(\psi_1 > 1) = 0$ and a'_t display fat tail. However, recall the main message of the paper is that variable externalities can trigger drastic changes in the asymptotic behavior of the aggregated macroeconomic series. Our result will be strengthened if we show this holds even when shocks are small (see also Section 3.3 below for a discussion of this assumption). To this end, we consider the case in which ϵ_t and ψ_t are both bounded positive iid processes on $[\underline{\psi}, \overline{\psi}] \times [\underline{\epsilon}, \overline{\epsilon}]$. We also assume $k_0, \underline{\psi}, \underline{\epsilon}$ are all strictly positive quantities. Indeed, it makes sense from an economic perspective to consider that the capital stock increases when the externality γ_t gets larger, all else equal, and this is true iff $k_t > 0$ always.

Proposition 2. *Assume ϵ_t and ψ_t are both bounded positive iid processes on $[\underline{\psi}, \overline{\psi}] \times [\underline{\epsilon}, \overline{\epsilon}]$. We also assume $k_0, \underline{\psi}, \underline{\epsilon}$ are all strictly positive quantities.*

1. *The equation*

$$E[\psi_1^\kappa] = 1, \tag{20}$$

admits a unique solution in $]0, +\infty[$ denoted κ_0.

2.

$$\lim_{c \to +\infty} c^{\kappa_0} P(k_\infty > c) = l > 0. \tag{21}$$

Proof. See Appendix A. □

The conditions of Propositions 1 and 2 are sufficient for the derivation of the asymptotic behavior of other macroeconomic series.

Proposition 3. *Under the conditions of Proposition 2,*

1. *The stationary distributions of the processes y_t and w_t share the same extreme behavior as that of the capital stock.*
2. *The growth rates of k_t, y_t and w_t also admit fat tailed limiting distributions.*

3.3 Comments on fixed *vs* random returns to scale

The results of Propositions 1 to 3 show that the limiting distributions of aggregated series exist and may display fat tails. These results ultimately rest on the random nature of returns to scale. If returns to scale are sometimes (but not too often) increasing (*i.e.* $E[\log(\psi_1)] < 0$ and $P(\psi_1 > 1) > 0$), the model generates stationary series with Pareto-like tails, even if the shocks affecting the returns to scale are arbitrarily small.

As this result has strong economic implications, we shall now motivate our choice to model returns to scale as random. Let us start from the difficulties faced by the the usual model with constant returns to scale and consider the case where $\gamma_t = \gamma$. Stationarity holds iff the roots of the polynomial

$$(1 - (\alpha + \gamma)x)(1 - \phi_1 x - \phi_2 x^2), \tag{22}$$

have modulus larger than one. If this is the case, stationarity of the TFP process implies that returns to scale are decreasing, and that growth will stop asymptotically. Hence, sustained positive growth and stationarity are incompatible in a model with a constant externality on capital accumulation, decreasing returns to scale, and a stationary TFP process.

From an empirical point of view, stationarity has been studied extensively. In particular, the upward jump witnessed from 1800 onwards is difficult to explain based on stationary processes. Further, considering long term data, growth is not a stable phenomenon, as current growth figures in developed countries are very high from an historical perspective.

One way to reconcile these facts is to allow for fat tails in the distribution of shocks. Indeed, in this case, the contemporary period in the developed countries may be viewed as a very lucky episode of large positive shocks. This may be done “directly” by assuming that the distribution of the TFP displays fat tails behavior. For instance, Fagiolo et al. [2008] recently attempt to introduce fat tails in the distribution of shocks in the context of a growth model. The problem is then to justify the fact that several developed countries have experienced historically high growth rates for about two centuries. To overcome this implausible sequence of lucky events, the fat tail may be linked to dynamic heteroskedasticty. In this case, ARCH models are good candidates. Engle [1982] studied the presence of ARCH effects in several macroeconomic time series. As ARCH models typically entail leptokurticity, extreme behavior is more frequent in comparison to the usual homoskedastic case.

In the context of business cycle models, several problems with this approach arise. First, the justification of large TFP shocks remains purely empirical and does not rely on agents' behavior. Second, as agents take the distribution of TFP shocks into account in the decision process, frequent and large shocks may hamper the mere existence of their objective functions. Third, large shocks are not consistent with the usual practice of studying the dynamics of models around the deterministic steady state. Hence, if returns to scale are fixed, it is difficult to reconcile facts with both applied and theoretical modeling.

When returns to scale are random, the picture is totally different. First, from the previous section, we know that large swings are compatible with both stationarity and small shocks. If the TFP process admits a strong stationary AR(1) representation, the two sufficient conditions for this type of behavior are $E[\log(\psi_1)] < 0$ and $P(\psi_1 > 1) > 0$. The first condition (together with stationarity of the TFP process) implies stationarity. It basically says, that, *on average*, the roots of the random polynomial

$$(1-(\alpha+\gamma_t)x)(1-\phi_1-\phi_2x^2), \tag{23}$$

are inside the unit circle. The second condition states that returns to scales are increasing episodically. When γ_t is large enough, the above random polynomial admits one root outside the unit circle and k_t *temporarily* follows an explosive path. As γ_t is iid, this happens infinitely often, although not too often otherwise processes would become non-stationary. The stationary distribution then displays fat tails, even if the TFP process does not.

Processes derived from Random Coefficients ARMA equations also display time-dependent variance, even with homoskedastic shocks. Indeed

$$Var_t[k_{t+1}] = Var[\psi_{t+1}]k_t^2 + (1-\delta)^2Var[a_{t+1}]. \tag{24}$$

Hence, randomness of returns to scale may endogenously explain the presence of dynamic heteroskedasticity.

Finally, the conditions $E[\log(\psi_1)] < 0$ and $P(\psi_1 > 1) > 0$ are compatible with an arbitrarily close-to-unit-root process. In our view, this is of primary importance from an econometric perspective. Indeed, the huge empirical macroeconometric literature on unit root testing strongly backs the 'random-walk hypothesis'. Therefore, if the data generating process is close to the unit-root case, there would be no econometric tools to distinguish between a stationary process and a random walk.

For example, consider the case where ψ_1 is uniformly distributed on $[\underline{\psi}, \overline{\psi}]$. The condition $E[\log(\psi_1)] < 0$ is equivalent to

$$\overline{\psi}\log(\overline{\psi}) - \underline{\psi}\log(\underline{\psi}) < \overline{\psi} - \underline{\psi}. \tag{25}$$

Write $\overline{\psi} = 1+\eta$ and $\underline{\psi} = 1-\theta\eta$ where η and θ are strictly positive values. The above condition is met whenever

$$0 < \eta < \frac{2\theta}{1+\theta^2}. \tag{26}$$

If θ is fixed, this condition is always met for a sufficiently small value of η. If η is very small, it is empirically virtually impossible to distinguish the process k_t from a unit-root process.[7]

Hence, in our model, fat tails do not come from exogenous large shocks. They are driven by an explicit, small, source of randomness. Moreover, it is agent's optimal behavior that conveys this source of randomness towards observable variables. As agents' decisions explicitly take into account this special feature of the shocks, we obtain a fully coherent picture linking observable facts, time series econometrics and macro-economic modeling. Of course this ideal viewpoint must be confronted to real-life figures, and this is what the next section is about.

4 Evidence from wages in England 1245–2010

We now use our dataset to present evidence backing our assumption of time-varying externality on capital accumulation. We focus our attention on three consequences of this hypothesis: instability of fixed coefficients ARIMA estimates, dynamic heteroskedasticity and fat tails. Of course, all these effects may be inherited from the TFP process. Therefore consistent testing is difficult. However, if the strength of the externality is time-varying, these features emerge endogenously as optimal responses to shocks.

The presentation of empirical results follows a route from the most simplistic to more sophisticated tests. First, we consider the instability of ARIMA fitting as an indication that a model with a constant externality is rejected by the data. Second, we perform formal stability tests guided by the structural equations of our model. Third, we show that dynamic heteroskedasticity is present in the data, a feature that our model is able to capture. Finally, we perform various estimations of the tail index of our wage series, and propose a method to extract the path of the time-varying autoregressive coefficient (ψ_t), which, according to our model, is driven by the time-varying strength of the externality on capital accumulation.

4.1 ARIMA fitting

As shown in Section 2.2, w_t admits a strong ARMA(p+1,q+1) representation if the TFP process admits a strong ARMA(p,q) representation and if γ_t is constant. If the externality is large enough, then $\psi = \delta + (1-\delta)(\alpha+\gamma)$ may be equal to or larger than 1. If $\psi = 1$,

[7]Distaso [2008] proposes a test for random walk against the alternative of random coefficient AR(1) process. The test procedure is designed so that the probability of rejection of the random walk is asymptotically controlled. But it cannot lead to a *consistent* test for *stationarity* since the alternative hypothesis contains both stationary and non-stationary processes. In particular, the alternative hypothesis contains very heavy-tailed stationary distributions without any moment. Under such an alternative, the asymptotic behavior of the test statistic proposed by Distaso [2008] is difficult to assess.

the wage process admits a unit root.[8]

Following the standard procedure of ARMA fitting, we first consider the issue of unit roots. ADF and KPSS tests may be highly unreliable to test stationarity, especially in the neighborhood of the unit-root case. However, they may be used to provide information about the number of roots that are close to unity. Indeed, if more than one root lies inside the unit circle, then the process of TFP shocks cannot be considered stationary.

Table 1: Stationarity/unit root tests (asymp. p-values)

	ADF			GLS-ADF		KPSS	
	no trend	linear t.	quadratic t.	no t.	linear t.	no t.	linear t.
Whole Sample	0.950	0.890	0.490	0.780	> 0.1	< 0.01	< 0.01
Late Medieval	0.190	0.001	0.004	0.020	0.05	< 0.01	< 0.01
Early modern	0.020	0.080	0.003	0.040	> 0.1	< 0.01	< 0.01
Late Modern	0.860	0.450	0.810	0.950	> 0.1	< 0.01	0.04
Contemporary	0.370	0.940	0.960	0.990	> 0.1	< 0.01	0.01

From the results reported in Table 1, stationarity is rejected for the whole sample as well as for the three most recent historical periods.[9] However, the decisions between various models are not always clearcut, especially for the oldest parts of the sample. Comparing ADF results (ADF and GLS-ADF), we see that the best decision regarding the presence of trend may change if we correct for possible heteroskedasticity. On the contrary, KPSS and ADF tests seem to point in the same directions.

A more informative approach is given by ARIMA fitting. Given the mixed evidence about unit roots, we choose to fit ARIMA specifications. We select the model from ARIMA(p,1,q) specifications for p and q ranging from 0 to 3 according to the Akaike Information Criterion among the models for which the p-values of the Ljung-Box and/or Box-Pierce test is above 5%.[10] The results are reported in Table 2.

Best ARIMA fits are changing quite drastically from a period to another. Moreover, the result obtained for the whole sample does not coincide with any of the fits obtained in the sub-samples. Further, residuals for the whole sample estimation barely succeed the

[8]Of course, a unit root may also be present in the process of TFP shocks. We ruled out this case in the theoretical section, as it corresponds to “exogenous growth” in the sense that growth results from a purely exogenous factor. See also footnote 14 below.

[9]The two first sub–samples have been chosen according to the usual convention *i.e.* Late Medieval up to 1500, Early Modern 1500-1800. The limit between Late Modern and Contemporary has been placed in 1913 in order to balanced the sample sizes.

[10]The number of lags for the Portmanteau tests have been adapted to match the square root of the (sub)-sample size.

Table 2: ARIMA fits

AR(1) (Est. / Std. err. / St. p-val.)	AR(2) (Est. / Std. err. / St. p-val.)	MA(1) (Est. / Std. err. / St. p-val.)	Inter. (Est. / Std. err. / St. p-val.)	Checks (AIC / B-P p-val. / L-B p-val.)	Roots (1st AR Mod. / 2nd AR Mod. / MA Mod.)
		Whole Sample : ARIMA(2,1,1)			
0.499	-0.194	-0.710	0.004	-1590	2.26
0.057	0.040	0.049	0.001	0.054	2.26
1.46 E-17	1.57 E-06	2.39 E-42	0.006	0.040	1.41
		Late Medieval : ARIMA(2,1,0)			
-0.132	-0.269		0.002	-399	1.93
0.061	0.061		0.005	0.065	1.93
0.030	1.13 E-05		0.347	0.055	
		Early Modern : ARIMA(1,1,1)			
0.530		-0.904	-0.001	-614	1.88
0.062		0.028	0.001	0.133	
8.70 E-17		2.09 E-148	0.272	0.119	1.11
		Late Modern : ARIMA(0,1,0)			
			0.012	-426	
			0.003	0.412	
			0.002	0.368	
		Contemporary : ARIMA(2,1,1)			
0.743	-0.311	-0.703	0.016	-300	1.79
0.171	0.108	0.145	0.003	0.165	1.79
1.61 E-05	0.004	1.50 E-06	3.04 E-09	0.132	1.42

Portmanteau test.[11] We thus consider Table 2 as the first piece of evidence of instability of the ARIMA coefficients.

Of course, it could be claimed that this instability reflects the different economic periods that England went through during the last 8 centuries. However, current Growth Theory tries to avoid using several *post hoc* explanations to get a broader unifying picture. Moreover, a closer inspection of the dynamics captured by sub-sample ARIMA fitting does not quite match what we known about economics in each historically relevant periods. Indeed, we should expect very low growth in the first period, with strong, possibly long lasting, MA effects coming from storage smoothing as reactions to bad/good harvest times. For the last period, we should obtain large AR effects echoing the slow diffusion of R&D improvements in heterogenous sectors. But the results are at odds with this simplistic view. Autoregressive (resp. Moving Average) effects are strong even in the first (resp. last) periods.

We may nevertheless use this first step estimation to gather some useful information about

[11] Slightly larger p-values for the Portmanteau tests may be obtained by increasing autoregressive and moving average orders, but the estimates are barely identified.

a more convincing model.

First, it does not seem justified to introduce moving average coefficients in the TFP process since the largest order of the MA coefficient estimated on wages is 1. According to our model, the dynamics of w_t contains current and first lag values of a_t. Bridging these results and our model with constant γ_t, an approximate estimation of δ is given by the MA(1) term. The results point to an estimated value of δ of 0.71.[12]

Second, the estimations reported in Table 2 point to an AR(2) specification for the process of TFP shocks. In addition, this representation is likely to be stationary with complex roots and with a dominant frequency around 0.15, which is consistent with (exogenous) cycles lasting 6 to 7 years. From now on, we thus assume that a_t admits a strong, stationary AR(2) representation.

4.2 Formal stability tests

The above investigation provides useful information that can be used to perform a more structural analysis. It also serves as an indirect indication that a model with a constant externality would probably be rejected by the data. However, it cannot be considered as a formal test of (in)stability. We now propose a more direct way to assess the interest of our model. To this end, we start from Equations (6)-(7) to get

$$k_t = a_k + (1-\delta)s + \delta k_{t-1} + (1-\delta)(w_t - \log(1-\alpha) + n), \tag{27}$$
$$w_t - \log(1-\alpha) + n = \gamma'_t k_{t-1} + a_t + (1-\alpha)n, \tag{28}$$
$$a_t = w_t + \alpha n - \log(1-\alpha) - \gamma'_t k_{t-1}, \tag{29}$$

where $\gamma'_t = \alpha + \gamma_t$. As $0 < \delta < 1$ we may write

$$k_t = s' + (1-\delta) \sum_{i=-\infty}^{t} \delta^{t-i} w_i, \tag{30}$$
$$a_{t+1} - \alpha n + \log(1-\alpha) = w_{t+1} - \gamma'_{t+1} k_t, \tag{31}$$

where $s' = s + \frac{a_k}{1-\delta} + n - \log(1-\alpha)$. Finally, we derive the dynamics of wages

$$w_{t+1} - \gamma'_{t+1}(1-\delta) \sum_{i=-\infty}^{t} \delta^{t-i} w_i = a_{t+1} - \alpha n + \log(1-\alpha) + \gamma'_{t+1} s'. \tag{32}$$

Provided δ is known, the series $x_t(\delta) = (1-\delta)\sum_{i=-\infty}^{t} \delta^{t-i} w_i$ may be computed directly from the data. In addition, the RHS of the above equation admits a strong, stationary, ARMA representation under the assumptions of Section 3.

[12] This is quite lower than what we could expect, but remember that δ can be interpreted as capturing the quality of the installed capital stock. As we deal with very old data, a small value of δ, *i.e.* a low "average" quality of the capital stock over the whole period, seems justified.

On one hand, with a constant γ_t, the series w_t and $x_t(\delta)$ are stationary iff $\gamma' < 1$. On the other hand, if $\gamma' > 1$, w_t and $x_t(\delta)$ are cointegrated if a_t is a stationary process. Therefore Equation (32) is particularly relevant to test various hypotheses about our process without relying on strong assumptions. We made clear in the discussion of the results of Table 2 that a rough estimate of δ can be derived from the case where γ_t is constant, and that $\hat{\delta} = 0.71$. After using this estimated value to compute $x_t(\delta)$, we perform cointegration tests between w_t and $\hat{x}_t = x_t(\hat{\delta})$ and report the results in Table 3 below.

Table 3: P-values of ADF tests

Series	P-values
w_t	0.9135
$\hat{x}_t$	0.9994
Residuals	6.38 E-58

In Table 3, 'Residuals' correspond to the linear regression of w_t on $\hat{x}_t$. The estimated value of γ' is 1.04 with a standard error of 0.007. At first glance, the cointegration relationship seems well established.[13] The estimated value of γ is larger than 1, which is consistent with the fact that w_t and $\hat{x}_t$ should be non stationary if γ_t is constant and large enough.[14] Further, a formal test of stability can be conducted along the lines proposed by Kuan and Hornik [1995].[15] Figure 3 provides the CUM-SUM and Moving Estimates derived from the cointegration relationship identified above.[16]

Both tests give the same clear message: the relationship is unstable.[17] Importantly, this feature does not start in the last part of the sample (*i.e.* from the Industrial Revolution onwards), it is a much older phenomenon. Hence, the stability problems identified with a model with a constant γ_t cannot be solved using, say, two different values of γ (before and after 1800, for instance).

Together with the indirect stability tests from ARIMA fitting, the formal stability tests point to a rejection of a model with a constant γ_t. We interpret this empirical rejection as indirect evidence that our model with time-varying γ_t is more in accordance with the data over the long run.

[13] We performed robustness checks with respect to the value of δ used to compute $x_t(\delta)$ and the results were not altered.

[14] The very low P-value associated with the ADF test performed on the series of residuals is also consistent with the stationarity assumption made in point 4 of Proposition 1.

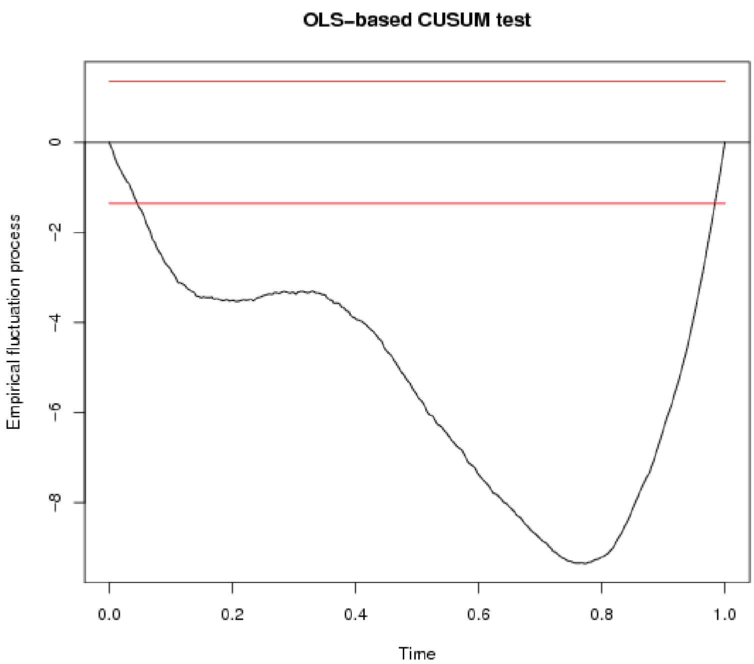
[15] We choose this type of test because the more usual F-tests procedures proposed by Andrews [1993] are not suitable for our alternative (namely iid γ_t). This may induce lower power, but this is harmful in our setting since generalized tests clearly *reject* the assumption of stability.

[16] Both charts have been obtained using the 'strucchange' package of **R**.

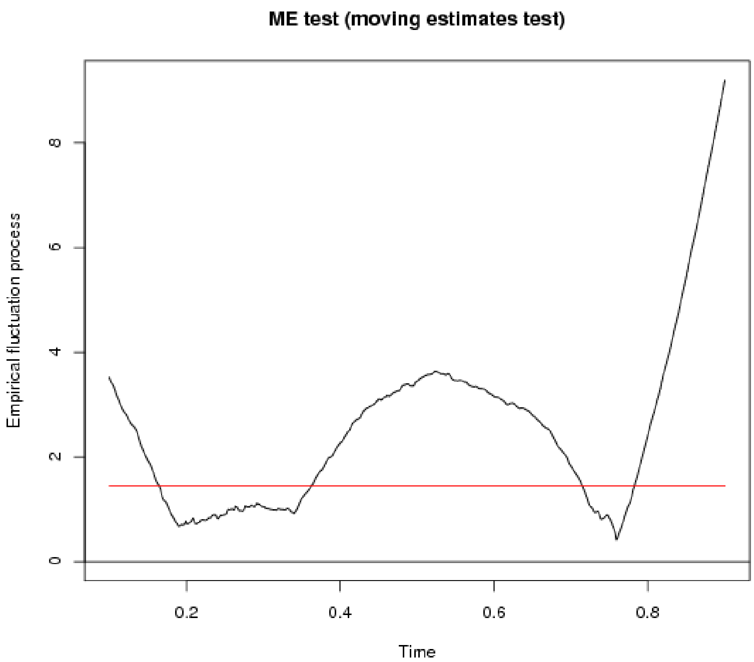
[17] Again this conclusion is robust to small changes in the value δ.

Figure 3: Generalized fluctuation tests (with 95% confidence boundaries)

(a) CUM-SUM



(b) ME

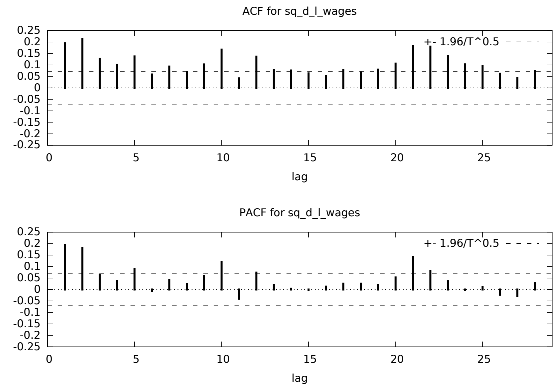


4.3 Dynamic heteroskedasticity

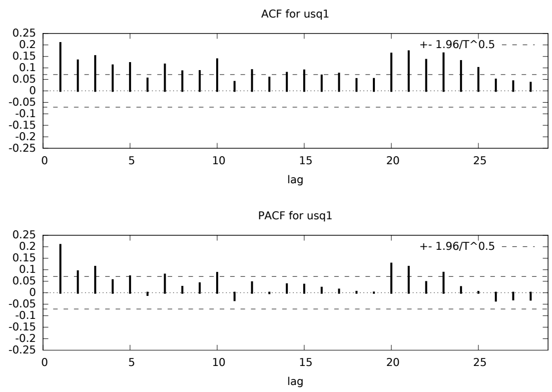
In the model with a time-varying externality on capital accumulation, heteroskedasticity can be considered from two perspectives. First, w_t and $w_t - w_{t-1}$ should display significant heteroskedasticity. Second, since ARIMA models with fixed coefficients are obviously misspecified, the residuals of the ARIMA fits should also display heteroskedasticity. Both features are captured in Figure 4.

Figure 4: Heteroskedasticity: evidence from correlograms

(a) Correlogram of squared increments of w_t



(b) Correlogram of squared residuals from ARIMA(2,1,1) fitting on w_t



The presence of dynamic heteroskedasticity is confirmed by the KPSS tests performed on these series, that also reject the hypothesis of stationarity. For completeness, we also provide the ARIMA fitting performed on the logarithm of squared residuals from the ARIMA(2,1,1) fitting in Table 4.

Table 4: ARIMA (1,1,1) estimates for log of squared residuals

AR(1) $\begin{pmatrix}\text{Est.}\\ \text{Std. err.}\\ \text{St. p-val.}\end{pmatrix}$	MA(1) $\begin{pmatrix}\text{Est.}\\ \text{Std. err.}\\ \text{St. p-val.}\end{pmatrix}$	Inter. $\begin{pmatrix}\text{Est.}\\ \text{Std. err.}\\ \text{St. p-val.}\end{pmatrix}$	Checks $\begin{pmatrix}\text{AIC}\\ \text{B-P p-val.}\\ \text{L-B p-val.}\end{pmatrix}$	Roots $\begin{pmatrix}\text{1st AR Mod.}\\ \text{MA Mod.}\end{pmatrix}$
0.108	-0.979	-0.027	3362	9.250
0.037	0.009	0.002	0.365	1.021
0.002	0.000	0.088	0.333	

Table 4 shows that the heteroskedasticity detected displays a significant dynamics. In particular, it is highly implausible that heteroskedasticity purely results from differences in the nature of shocks along the period and/or from a lack of precision in the oldest part of the sample. Together with evidence for instability presented in the two last subsections, the presence of dynamic heteroskedasticity strongly backs our assumption of time-varying γ_t.

4.4 Inference for the tail index

We now consider the second, more specific, consequence of time-varying returns to scale, namely the fact that aggregate time series should display fat-tail behavior.

Several techniques can be used to estimate the tail index of time series. As none of these methods has been precisely designed for RC-ARMA models, we present the estimations resulting from the various methods. First, we propose three techniques to derive inference about the tail index for stationary time series. Second, we propose an alternative approach based on the specific features of our model.

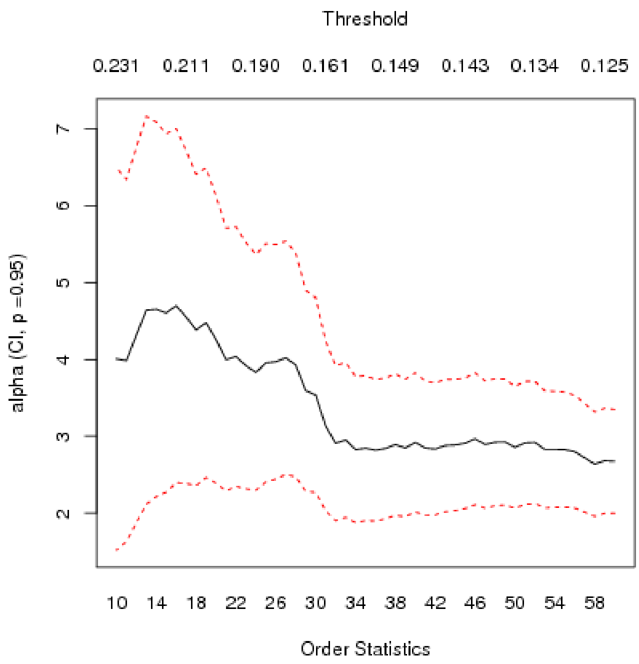
The most common estimator derives from the Hill plot. The consistency of the Hill estimator is established in various cases, but the convergence may be really slow and results in large imprecision, especially if dependence is very strong and/or if the distribution departs from the Pareto case. Additionally, the Hill plot, as well as other tail estimators, derives from high-order statistics. Applying this technique to our series of wages, will select only the last part of the sample (see Figure 1), leading to large dependence. We thus suggest using real wage growth (*i.e.* increments in w_t) instead of real wages to derive inference about the presence of a fat tail.[18]

[18] Indeed, we established in Proposition 3 that

$$\lim_{c,t\to+\infty} P(w_t - w_{t-1} > c) = \lim_{c,t\to+\infty} P(w_t > c)P(\psi_t > 1). \tag{33}$$

so that real wage growth should also display Pareto-like tails.

Figure 5: Hill plot of real wages growth



The results seem to favor a fat tail index between 2 and 5.[19] Taking this feature into account, we perform the estimation method proposed by Gabaix and Ibragimov [2011], which is a version of the Hill estimator that corrects for small sample bias. Using the method of Gabaix and Ibragimov [2011], we obtain an estimated value of the tail index of 3.94 (with a standard error for the slope equal to 0.15).

Francq and Zakoian [2011] recently questioned the previous approaches in the context of time series on the ground that the dynamic dimension is neglected. We use Francq and Zakoian [2011]'s method to estimate a confidence interval for our tail index. More precisely, we fit a Generalized Extreme Value distribution to our data and compute the asymptotic confidence interval as explained in Francq and Zakoian [2011].[20]

Figure 6 assesses the goodness of fit between the distribution of increments of w_t and the Generalized Extreme Value distribution and Table 5 presents the estimation results.[21]

[19] Using the 'evir' package of R, we obtain the Hill plot in Figure 5.

[20] We thank Christian Francq for kindly providing his **R code.**

[21] The parametrization in Table 5 below refers to the notation in Extreme Value Theory used by Kotz and Nadarajah [2002]. In particular, the parameter ξ is denoted γ in Francq and Zakoian [2011]'s paper. We rely on the the parametrization by Kotz and Nadarajah [2002] to avoid confusion.

Figure 6: densities comparison GEV/actual

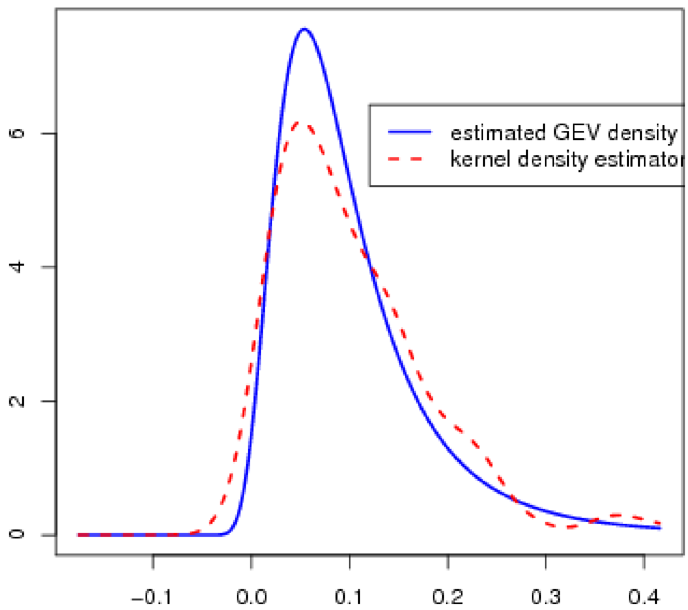


Table 5: Estimation results of GEV parameters

Parameter	Est.	Std	(Asy.) 95% C.I.
μ	0.063	0.006	[0.051;0.075]
σ	0.050	0.005	[0.04; 0.06]
ξ	0.208	0.100	[0.008; 0.408]
Tail index	4.70	12.37	[2.17; 26.09]

The estimated value of the tail index is somewhat larger than the estimated value obtained with Gabaix and Ibragimov [2011]'s method, and the confidence interval is wide. This is not surprising since Francq and Zakoian [2011]'s method corrects the confidence interval for serial correlation. The point estimate for the tail index is larger than the results usually obtained in the financial literature. This is consistent with the fact that real markets exhibit less volatility than financial ones.

We now discuss a final approach using the specific features of our model and dataset. We simply sketch the relevant part of this analysis, and refer to Appendix B for details. From our model, k_t is observable up to two constants, namely δ and s'. Appendix B explains our approach to derive reasonable estimates of these two quantities, so that an estimated sequence $\hat{k}_t$ can be computed from the dataset.[22]

We may write

$$k_t/k_{t-1} = \psi_t + a'_t/k_{t-1}, \tag{34}$$

so that ψ_t is observed up to the "error term" a'_t/k_{t-1}. One strategy is to recover an estimate for the density of ψ_t by deconvolution. This can not be done directly since a'_t/k_{t-1} is in fact a dependent and heteroskedastic process. To deal with dependence, we propose to sub–sample our data and deconvolution with (observed) heteroskedasticity is achieved using Delaigle and Meister [2008]'s method.[23] Figure 7 plots the density estimation obtained using Delaigle and Meister [2008]'s method, assuming the innovations affecting a'_t are Gaussian.

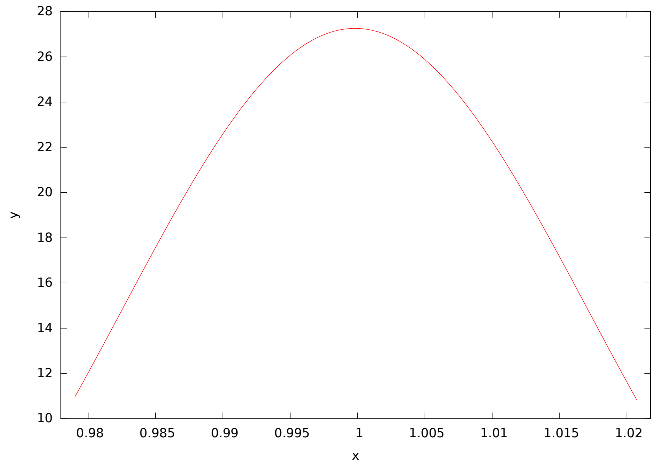
As expected, the density is concentrated around one, with a clear mode at the unit-root case. This density estimate $\hat{f}_\psi$ may then be used to estimate κ by the moment condition

$$\int x^{\hat{\kappa}} \hat{f}_\psi(x)dx = 1. \tag{35}$$

[22]In the following, for simplicity, we remove the ˆ.

[23]We choose a 7 years lag in the estimation. The estimated autocorrelation coefficient at this lag is close to 3×10^{-4}. Implementation is provided by the **R**-package 'Decon' (see Wang and Wang [2011] for explanations).

Figure 7: Density estimation for ψ (heteroskedastic deconvolution)



We get $\hat{\kappa} = 3.89$ with a 95%-confidence interval equal to $[1.98, 18.54]$.[24] The estimate is close to – albeit somewhat smaller than – the results obtained using the the methods proposed by Francq and Zakoian [2011] and Gabaix and Ibragimov [2011]. The confidence interval derived with the method relying on deconvolution is similar to the confidence interval obtained using Francq and Zakoian [2011]'s method.

4.5 Extraction of the path of returns to scale

Finally, the estimation of the density of ψ_1 may be used to derive an extracted path for ψ_t. Consider

$$\begin{aligned} k_t &= \psi_t k_{t-1} + a'_t \\ &= \psi_t k_{t-1} + \sigma \epsilon_t + \phi_1 (k_{t-1} - \psi_{t-1} k_{t-2}) + \phi_2 (k_{t-2} - \psi_{t-2} k_{t-3}) + (1 - \phi_1 - \phi_2) a_\infty \end{aligned}$$

where $\epsilon_t | \psi_t, \psi_{t-1}, \psi_{t-2}, k_{t-1}, k_{t-2}, k_{t-3}$ is distributed as $\mathcal{N}(0,1)$. Hence, the distribution of k_t given $(\psi_t, \psi_{t-1}, \psi_{t-2}, k_{t-1}, k_{t-2}, k_{t-3})$ is also Gaussian, with mean $(\psi_t + \phi_1) k_{t-1} + (\phi_2 - \phi_1 \psi_{t-2}) k_{t-2} - \phi_2 \psi_{t-2} k_{t-3} + (1 - \phi_1 - \phi_2) a_\infty$ and variance σ^2. Applying Bayes' rule and taking into account the fact that ψ_t is independent from the past, we get

$$\begin{aligned} &\hat{f}(\psi_t | \psi_{t-1}, \psi_{t-2}, k_t, k_{t-1}, k_{t-2}, k_{t-3}) \propto \hat{f}_\psi(\psi_t) \times \\ &exp - \frac{1}{2\sigma^2} \left((\psi_t + \phi_1) k_{t-1} + (\phi_2 - \phi_1 \psi_{t-2}) k_{t-2} - \phi_2 \psi_{t-2} k_{t-3} + (1 - \phi_1 - \phi_2) a_\infty - k_t \right)^2 \end{aligned}$$

[24] The confidence interval is constructed using 1000 bootstrap replications in the sub-sampled sequence of k_t / k_{t-1}.

If the first two values of ψ_1, ψ_2 are fixed, the distribution of ψ_3 given the data is provided by the previous formula. Then ψ_4 may be simulated for any given valid simulation of ψ_3 and so on.[25] Figure 8 Panel (a) plots the average and interquartile bands for the path extracted from the data. The mean extracted path is very concentrated around the case of constant returns to scale, *i.e.* $\psi_t = 1$, but variations are quite important whatever the period. A smoothed path may be inferred using a polynomial trend approximation. We report it in Figure 8 Panel (b). It must be taken with care, as the dataset is not very informative on the path's history. Despite these reservations, the coincidence with currently admitted historical trends is remarkable.

Finally, we checked whether this extracted sequence displays any particular frequency pattern. Indeed, one could argue that very long cycles of variations in the externality may be related to middle–term pattern of growth's path. This quest for evidence of very long cycles in the economic development has a long–standing history (see for instance Korotayev and Tsirel [2010] for a recent inquiry). To this end, we compute a spectral density using Bartlett's window choice and report the results in Figure 8 Panel (c). The evidence is inconclusive, as the spectrum barely displays any significant peak.

5 Conclusion

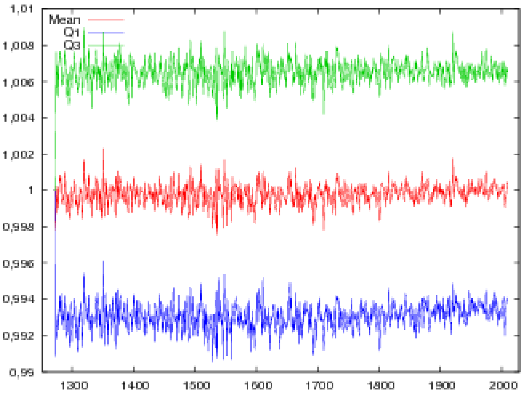
In this paper, we study a stochastic growth model in which returns to scale may vary randomly over time. The dynamics of main aggregate economic series including capital, growth, various relative prices are related to a Random Coefficients ARMA model. A nice feature of our model is that transitory growth episodes could be combined with the existence of a limiting stationary solution. We give the sufficient conditions under which the stationary distributions of macroeconomic sequences display fat tails. The index of fat tail is closely related to the occurrence of temporary increasing returns to scale. These conditions are consistent with available empirical evidence on very long macroeconomic time series.

We show that it is difficult to distinguish our model from a more usual growth model in which externalities are constant with available macroeconomic data, as this task typically requires very long datasets. To this end, we propose evidence based on an annual sequence of real wages in England recorded since 1245. Dynamic heteroskedasticity is present in the data and the instability of ARIMA estimations over sub–samples of the data is consistent with random returns to scale. Finally, the data is used to estimate the fat tail index using various methods, including a method that help us extracting a path of returns to scale. The analysis concludes that the wage sequence is mildly fat tailed. All these results support our assumption of small variations in returns to scale around the case of constant

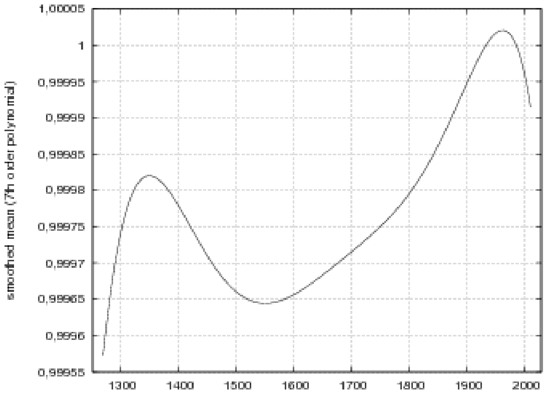
[25] From a computational perspective, signal extraction is much less involving than smoothing (that is extraction of the values of ψ_t from the whole dataset). Whether smoothing is feasible in such a context is beyond the scope of this paper.

Figure 8: Characteristics of the extracted path of ψ_t

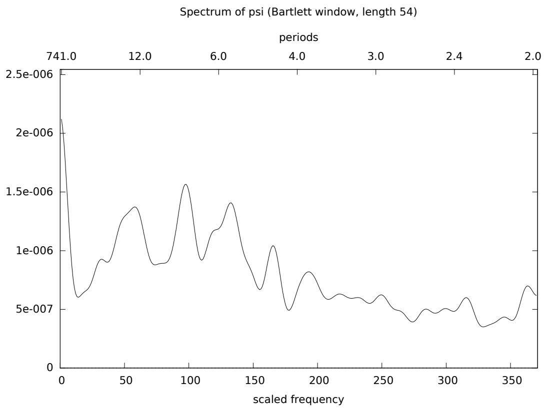
(a) Extracted path of ψ_t (5000 simulated paths)



(b) Smooth extracted trend of ψ_t (7th-order polynomial)



(c) Spectral density decomposition of ψ_t



returns to scale. These results are confirmed by the extracted path of returns to scale, as their variations are very small but consistent with major historical trends.

References

Andrews, D.K.W., Tests for parameter instability and structural change with unknown change point, *Econometrica*, 1993, *61*, 821–856.

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

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

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

de Haan, L., S. Resnick, H. Rootzén, and C. de Vries, Extremal Behavior of Solutions to a Stochastic Difference Equation, with Applications to ARCH–processes, *Stochastic Processes and their Applications*, 1989, *32*, 213–224.

Delaigle, A. and A. Meister, Deinsty estimation with heteroscedastic errors, *Bernoulli*, 2008, *14*, 562–579.

Distaso, W., Testing for Unit Root Processes in Random Coefficient Autoregressive Models, *Journal of Econometrics*, 2008, *142*, 581–609.

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

Fagiolo, G., M. Napoletano, and A. Roventini, Are Output Growth-Rate Distributions Fat-Tailed ? Some Evidence from OECD Countries, *Journal of Applied Econometrics*, 2008, *23(5)*, 639–669.

Francq, C. and J-M. Zakoian, Estimating the Marginal Law of a Time Series with Applications to Heavy Tailed Distributions, *w.p. CREST 2011-30*, 2011.

Gabaix, X., Zipf's Law for Cities: An Explanation, *Quarterly Journal of Economics*, 1999, *114*, 739–767.

——, Power Laws in Economics and Finance, *Annual Review of Economics*, 2009, *1*, 255–293.

—— and R. Ibragimov, Rank–1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents, *Journal of Business Economics and Statistics*, 2011, *29*, 24–39.

——, P. Gopikrishnan, V. Plerou, and H. E. Stanley, Institutional Investors and Stock Market Volatility, *Quarterly Journal of Economics*, 2006, *121*, 461–504.

Galor, O., From Stagnation to Growth: Unified Growth Theory, in Handbook of Economic Growth, North–Holland, 2005.

——, Multiple Growth Regimes – Insights from Unified Growth Theory, *Journal of Macroeconomics*, 2007, *29*, 470–475.

—— and D. Weil, From Malthusian Stagnation to Modern Growth, *American Economic Review*, 1999, *89*, 150–154.

—— and ——, Population, Technology and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond, *American Economic Review*, 2000, *90*, 806–828.

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

——, ——, and D. Vollrath, Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions and the Great Divergence, *Review of Economic Studies*, 2009, *76*, 143–179.

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

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

Kesten, H., Random Difference Equations and Renewal Theory for Products of Random Matrices, *Acta Mathematica*, 1973, *131*, 207–248.

Korotayev, A. V. and S.V. Tsirel, A Spectral Analysis of World GDP Dynamics : Kondratieff Waves, Kuznets Swings, Juglar and Kitchin cycles in Global Economic Development, and the 2008-2009 Economic Crisis, *Structure and Dynamics*, 2010, *4*.

Kotz, S. and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*, World Scientific: Imperial College Press, 2002.

Kuan, C.M. and K. Hornik, The generalized fluctuation test: A unifying overview, *Econometric Reviews*, 1995, *14*, 135–161.

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

—— and E. Prescott, Investment under Uncertainty, *Econometrica*, 1971, *39*, 659–681.

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

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

——, Human Capital, Product Quality, and Growth, *Quarterly Journal of Economics*, 1991, *106*, 587–616.

Wang, X.F. and B. Wang, Deconvolution in Measurement Error Models : the R package decon, *Journal of statistical software*, 2011, *39.*

A Proofs

A.1 Proposition 1

Proof. Clearly, if k_t admits a unique stationary distribution, then so do y_t and w_t. Indeed, recall we have

$$y_t = (\alpha + \gamma_t)k_{t-1} + a_t + (1-\alpha)n. \tag{36}$$

Under the conditions of Proposition 1, the processes a_t and $\alpha + \gamma_t$ are both stationary. Also w_t is an affine function of y_t.

Now the result for k_t may be derived as follows

$$k_t = \psi_t k_{t-1} + a'_t = k_0 \prod_{i=1}^{t} \psi_i + \sum_{i=1}^{t} a'_i \prod_{j=t}^{i+1} \psi_j, \tag{37}$$

with the convention $\prod_{j\in\emptyset} = 1$. A sufficient condition for stationarity is the absolute convergence of the sequences $k_0 \prod_{i=1}^{t} \psi_i$ and $\sum_{i=1}^{t} a'_i \prod_{j=t}^{i+1} \psi_j$ as t goes to infinity. Let us consider the second one. Using the Cauchy principle, a sufficient condition is

$$\overline{lim}_{i\to+\infty} |a'_i|^{1/i} \prod_{j=t}^{i+1} |\psi_j|^{1/i} < 1. \tag{38}$$

Taking logarithms, a sufficient condition is

$$\overline{lim}_{i\to+\infty} \frac{1}{i} \log|a'_i| + \frac{1}{i} \sum_{j=t}^{i+1} \log|\psi_j| < 0. \tag{39}$$

Now the condition 4 in Proposition 1 entails a'_i converges to a stationary distribution. Moreover as $E[\log|\epsilon_t|] < +\infty$ we have

$$\overline{lim}_{i\to+\infty} \frac{1}{i} \log|a'_i| = 0. \tag{40}$$

Now, as for the second term, we have

$$\overline{lim}_{i\to+\infty} \frac{1}{i} \sum_{j=t}^{i+1} \log|\psi_j| = E[\log(\psi_1)](a.s.) \tag{41}$$

and the result follows from the assumption $E[\log(\psi_1)] < 0$. Now this condition together with the fact that k_0 is a given quantity insures $k_0 \prod_{i=1}^{t} \psi_i$ is an absolute convergent random sequence.

□

A.2 Proposition 2

Point 1 is easily derived the function of the real positive quantity κ defined as $E[\psi_1^\kappa]$ is convex and twice differentiable, and its derivative at zero is $E[\log(\psi_1)] < 0$ whereas $P(\psi_1 > 0) > 1$ implies $E[\psi_1^0] = 1$. Now $P(\psi_1 > 1) > 0$ implies this function is larger than 1 for some large enough value of κ and the result follows.
As for Point 2, first notice that according to Equation (18)

$$k_t - \psi_t k_{t-1} = a'_t \Rightarrow k_t - \psi_t k_{t-1} = \epsilon_t + \phi(k_t - \psi_t k_{t-1}) \tag{42}$$

Let us define

$$A_1 = \begin{pmatrix} \psi_1 + \phi_1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$A_2 = \begin{pmatrix} \psi_2 + \phi_1 & \phi_2 - \phi_1\psi_1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$
$$A_t = \begin{pmatrix} \psi_t + \phi_1 & \phi_2 - \phi_1\psi_{t-1} & -\phi_2\psi_{t-2} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \forall t \geq 2$$
$$\mathbf{k}_t = \begin{pmatrix} k_t \\ k_{t-1} \\ k_{t-2} \end{pmatrix} \quad \forall t \geq 2$$
$$\mathbf{e}_t = \begin{pmatrix} \epsilon_t \\ 0 \\ 0 \end{pmatrix} \quad \forall t \geq 2$$

We have $\mathbf{k}_t = A_t\mathbf{k}_{t-1} + \mathbf{e}_t$ for all $t \geq 2$.

Under the conditions of Proposition 1 we shall study the existence of a solution to the following moment equation

$$\lambda(\kappa) = \lim_{t\to+\infty} E[|A_T A_{T-1} \dots A_1|^\kappa]^{1/T} = 1. \tag{43}$$

Indeed, after Kesten [1973], we known that if a solution to the above equation exists, then $\lim_{x\to+\infty} P(k_\infty > x)x^\kappa = l > 0$. Set $\phi' = (\phi_1, \phi_2)$ and define for all $t > 2$ $N_t(\phi) = A_T A_{T-1} \dots A_1$. Direct computations shows that for all $t > 2$ the only elements non null element of the matrix $N_t(\phi)$ are those of the first column. So that $|A_T A_{T-1} \dots A_1|$ is equal to the top left element of $N_t(\phi)$ which we denote $n_t(\phi)$. Moreover we have for all $t \geq 2$:

$$n_t(\phi) - \psi_t n_{t-1}(\phi) = \phi_1(n_{t-1}(\phi) - \psi_{t-1}n_{t-2}(\phi)) + \phi_2(n_{t-2}(\phi) - \psi_{t-2}n_{t-3}(\phi)), \tag{44}$$

with the initial conditions

$$n_0(\phi) = 1, \tag{45}$$
$$n_1(\phi) = \psi_1 + \phi_1, \tag{46}$$
$$n_2(\phi) = \psi_2(\psi_1 + \phi_1)\phi_2 + \phi_1^2. \tag{47}$$

By recurrence, we then see that $n_t(\phi) - \psi_t n_{t-1}(\phi)$ is a deterministic process (it does not depend on ψ_t). Moreover this quantity admit the same AR(2) representation as a_t, except that the innovation is zero. We the deduce

$$\lim_{t\to+\infty} n_t(\phi) - \psi_t n_{t-1}(\phi) = 0. \tag{48}$$

Now, as ψ_t and $n_{t-1}(\phi) - n_{t-1}(\phi)$ are independent and $P(\psi_t > 0) = 1$ we get:

$$\lim_{t\to+\infty} n_t(\phi) - n_t(0) = 0. \tag{49}$$

Using the boundary conditions, $n_t(\phi)$ and $n_t(0)$ are strictly positive and bounded sequences, which implies

$$\lim_{t\to+\infty} \frac{1}{t} E[\log(n_t(\phi))] = \lim_{t\to+\infty} \frac{1}{t} E[\log(n_t(0))]. \tag{50}$$

Hence, whenever it exists $\lambda(\kappa)$ does not depend on ϕ. We may now rely on de Haan, Resnick, Rootzén and de Vries [1989] to establish that the distribution of k_∞ display fat tail behavior in the case $\phi = 0$.

A.3 Proposition 3

As w_t is an affine strictly increasing function of y_t, both variables have the same extremal behavior. To derive the result for y_t, we use the relation between output y_t and the capital stock

$$y_t = (\alpha + \gamma_t)k_{t-1} + a_t + (1-\alpha)n. \tag{51}$$

Now under the conditions of Proposition 2, the processes a_t and $\alpha + \gamma_t$ are both bounded. Let $\underline{\gamma}$ and $\underline{a}$ be the lower bounds of these processes. Assuming $\alpha + \underline{\gamma} > 0$ we have for all x

$$P(y_t > x) \geq P((\alpha + \underline{\gamma})k_{t-1} + \underline{a} > x), \tag{52}$$

from which we deduce

$$\lim_{x\to+\infty} x^{\kappa_0} P(y_\infty > x) \geq l > 0, \tag{53}$$

where y_∞ is the stationary limit of y_t. Similarly, let $\overline{\gamma}$ and $\overline{a}$ be the upper bounds of a_t and $\alpha + \gamma_t$, we have

$$P(y_t > x) \leq P((\alpha + \overline{\gamma})k_{t-1} + \overline{a} > x), \tag{54}$$

from which we deduce

$$\lim_{x\to+\infty} x^{\kappa_0} P(y_\infty > x) \leq l. \tag{55}$$

Notice the same result may be derived for the (real) rate r_t as we have

$$r_t = y_t + \log(\alpha + \gamma_t) - k_{t-1} = a_t + (1-\alpha)n + \log(\alpha + \gamma_t) + (\alpha + \gamma_t - 1)k_{t-1}. \tag{56}$$

Consider the growth rate of the capital stock over one period

$$k_t - k_{t-1} = (-1 + \psi_t)k_{t-1} + a'_t. \tag{57}$$

Since a'_t is bounded

$$\lim_{c\to+\infty} P((-1+\psi_t)k_{t-1} + a_t > c) = \lim_{c\to+\infty} P((-1+\psi_t)k_{t-1} > c). \tag{58}$$

As $\psi_t = 1$ is zero probability event $P((-1+\psi_t)k_{t-1} > c)$ equals

$$P(k_{t-1} > c/(-1+\psi_t)|\psi_t > 1)P(\psi_t > 1) + P(k_{t-1} < c/(-1+\psi_t)|\psi_t < 1)P(\psi_t < 1). \tag{59}$$

As k_{t-1} and ψ_t are independent, and considering ψ_t is bounded

$$\lim_{c\to+\infty} P(k_{t-1} > c/(-1+\psi_t)|\psi_t > 1) = \lim_{c\to+\infty} P(k_{t-1} > c), \tag{60}$$

$$\lim_{c\to+\infty} P(k_{t-1} < c/(-1+\psi_t)|\psi_t < 1) = \lim_{c\to-\infty} P(k_{t-1} < c) = 0. \tag{61}$$

where the last result follows from $k_t > 0$. The desired result follows. Results for the growth rates of y_t and w_t follow accordingly.

B Details of computations used in Section 4.4

We describe how the first step estimates needed for the estimation performed in Section 4.4 are recovered. First, δ, ϕ_1, ϕ_2 are calibrated from the case with a constant γ_t, as explained in Section 4.2. We also use this approximation to derive an estimated value of the variance of TFP shocks. The only remaining parameter is s'. We propose the following solution. Consider $\gamma'_t = \gamma' + \zeta\eta_t$ where η_t is iid, $E[\eta_t] = 0$ and $Var[\eta_t] = 1$. We have

$$z_{t+1} = w_{t+1} - \gamma'(1-\delta)x_t = a_{t+1} - \alpha n + \log(1-\alpha) + \gamma' s' + \zeta\eta_t((1-\delta)x_t + s'), \tag{62}$$

and

$$\begin{aligned} z_{t+1} - \phi_1 z_t - \phi_2 z_{t-1} &= \epsilon_{t+1} + \zeta(\eta_t - \phi_1\eta_{t-1} - \phi_2\eta_{t-2})s' && (63)\\ &+ \zeta(1-\delta)(\eta_t x_t - \phi_1\eta_{t-1}x_{t-1} - \phi_2\eta_{t-2}x_{t-2}). && (64) \end{aligned}$$

We then compute

$$\begin{aligned} &E[(z_{t+3} - \phi_1 z_{t+2} - \phi_2 z_{t+1})(z_{t+1} - \phi_1 z_t - \phi_2 z_{t-1})|x_{t+2}, x_{t+1}, x_t, x_{t-1}, x_{t-2}] = && (65)\\ &E[\epsilon_t]^2 - \phi_2(s'\zeta)^2 - 2\phi_2\zeta^2(1-\delta)s'x_t - \phi_2\zeta^2(1-\delta)^2x_t^2. && (66) \end{aligned}$$

Hence, the quantity s' can be obtained by regressing $(z_{t+3} - \phi_1 z_{t+2} - \phi_2 z_{t+1})(z_{t+1} - \phi_1 z_t - \phi_2 z_{t-1})$ on x_t, x_t^2 and a constant. s' is then a given by the ratio of the slopes associated to x_t and x_t^2. Once s' is known, we may then compute $k_t = s' + (1-\delta)x_t$.