

Global Stochastic Properties of Dynamic Models and their Linear Approximations

Ana Babus
University of Cambridge

Casper G. de Vries*
Erasmus University Rotterdam

Abstract

The dynamic properties of micro based stochastic macro models are often analyzed through a linearization around the associated deterministic steady state. Recent literature has investigated the errors made by such a deterministic approximation. Complementary to this literature we investigate how the linearization affects the stochastic properties of the original model. We consider a simple real business cycle model with noisy learning by doing. The solution has a stationary distribution that exhibits moment failure and has an unbounded support. The linear approximation, however, yields a stationary distribution with possibly a bounded support and all moments finite.

Keywords: Linearization; ARCH process; Real business cycles model; Stochastic difference equation

JEL: C22, C62, E32

*Corresponding author. Address: Dept. of Economics H8-33, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; email: <cdevries@ese.eur.nl>.

1 Introduction

The dynamic properties of micro based stochastic macro models are often analyzed through a linearization around the associated deterministic steady state. In the seminal paper on real business cycles (RBC) Kydland and Prescott (1982) employed first order approximations to solve their dynamic, stochastic general equilibrium (DSGE) model. This method has become highly popular in analyzing DSGE's. Campbell (1994) and Uhlig (1997) provide overviews on how to perform the linearization of the dynamic micro based stochastic macro models. A number of papers has investigated the accuracy of the log linear approximation, by looking at the deterministic part of the approximate solution. Tesar (1995) and Kim (1997) prove that the loglinear approximation method may create welfare reversals, to the extent that the incomplete-markets economy produces a higher level of welfare than the complete-markets economy. Jin and Judd (2002) therefore recommend the use of second order perturbation methods. Sutherland (2002) and Kim and Kim (2003) have developed a bias selection method which can be as accurate as the perturbation method, but which requires less computational effort. The performance of the linear approximation in stochastic neoclassical growth models is studied by Dotsey and Mao (1992), and more recently in Arouba et al. (2006) and Fernandez-Villaverde and Rubio-Ramirez (2005).

We contribute to this literature by showing how the stochastic properties of the approximate solution differ from the equilibrium of the non-linear model. In particular, we investigate the simplest model in the business cycle literature with fixed labour supply, total depreciation of capital and a log-utility function. To this we add noisy learning by doing. The solution of the resulting stochastic difference equation has a stationary distribution which exhibits moment failure and has an unbounded support. The first order approximation, however, yields a stationary distribution with bounded support and all moments finite. Thus the linear approximation dramatically alters the stochastic properties of the model. We also consider briefly an application from asset pricing with stochastic volatility.

This note is organized as follows. In section 2 we analyze the RBC model and we show that while the exact solution of the model for the log of capital follows a stationary

distribution with unbounded support and exhibits moment failure, the approximation may nevertheless have bounded support and all moments finite. Section 3 further discusses the effects of linearization in the capital asset pricing model with changing conditional volatility of the ARCH variety. Section 4 concludes.

2 Application on the real business cycle model

Log-linearization is a well known method for solving business cycle models. It has its pros and cons, which are usually discussed in a deterministic setting. We join this literature by showing how linearization may change the stochastic equilibrium behavior of the solution of a dynamic RBC model.

The environment of the basic RBC model with fixed unitary labour supply and noisy learning by doing is as follows:

1. The production function is Cobb-Douglas $Y_t = I_t^\alpha K_t^{1-\alpha}$, where I is technology and K is capital.
2. With full depreciation, the next period capital equals the current period's savings:

$$K_{t+1} = I_t^\alpha K_t^{1-\alpha} - C_t.$$
3. The representative agent expected utility is: $U = E_t[\sum_{i=0}^{\infty} \beta^i \log(C_{t+i})]$.
4. Technological progress stems from learning by doing: $I_{t+1} = \phi_{t+1} Y_t^{\varepsilon_{t+1}}$, where $\phi_t > 0$ and ε_t are random variables independently distributed with mean ϕ and ε , respectively. The learning by doing effect stems from the aggregate production level. This externality is not taken into account by the individual consumer when planning his consumption pattern.
5. The gross rate of return on a one period investment in capital R_{t+1} equals the marginal product of capital: $R_{t+1} = (1 - \alpha) (I_{t+1}/K_{t+1})^\alpha$.

This special case of a stochastic dynamic general equilibrium model with full depreciation of capital and log utility function admits an exact solution. The first order condition

for utility maximization is: $1/C_t = \beta E_t[(1 - \alpha)I_{t+1}^\alpha K_{t+1}^{-\alpha}/C_{t+1}]$. In order to solve the system

$$\frac{1}{C_t} = \beta E_t\left[\frac{(1 - \alpha)I_{t+1}^\alpha K_{t+1}^{-\alpha}}{C_{t+1}}\right] \quad (1)$$

$$K_{t+1} = I_t^\alpha K_t^{1-\alpha} - C_t \quad (2)$$

$$Y_t = I_t^\alpha K_t^{1-\alpha} \quad (3)$$

$$I_{t+1} = \phi_{t+1} Y_t^{\varepsilon_{t+1}} \quad (4)$$

we guess the policy function

$$C_t = \mu I_t^\alpha K_t^{1-\alpha}. \quad (5)$$

Inserting (5) in (1) and using the equation for the capital accumulation process (2) determines the constant $\mu = 1 - \beta(1 - \alpha)$.

Subsequently substitute (5) and (3) into (2). This shows that the log of capital k_{t+1} satisfies:¹

$$k_{t+1} = \log \beta(1 - \alpha) + y_t. \quad (6)$$

Transform (4) into logs

$$i_{t+1} = \log \phi_{t+1} + \varepsilon_{t+1} y_t. \quad (7)$$

Advancing (3) one period, taking logarithms as well and inserting (6) and (7), we obtain the first order stochastic difference equation for log income:

$$y_{t+1} = (1 - \alpha) \log \beta(1 - \alpha) + \alpha \log \phi_{t+1} + (\alpha \varepsilon_{t+1} + 1 - \alpha) y_t. \quad (8)$$

This difference equation can be conveniently summarized as

$$X_t = A_t + B_t X_{t-1}, \quad (9)$$

¹In this section capital letters stand for level values and small letters for log transformed variables.

where A_t and B_t are i.i.d. random variables, with

$$A_t = (1 - \alpha) \log \beta(1 - \alpha) + \alpha \log \phi_t$$

and

$$B_t = \alpha \varepsilon_t + 1 - \alpha.$$

Kesten (1973) provides a characterization of the stochastic steady state of the process (9), given certain conditions. One important property of the solution is that the tail of the stationary distribution may be asymptotic to the tail of a Pareto distribution, see the fourth claim in Theorem 1 below. Thus the stationary distribution is characterized by moment failure, even though the innovations may not exhibit this property; e.g. are bounded or exponential. The following theorem collects the full statement of Kesten's theorem (1973, th.5):

Theorem 1 (Kesten) *Consider the first order stochastic difference equation*

$$X_t = A_t + B_t X_{t-1},$$

where (A_t, B_t) , $t \geq 1$, are independent and identically distributed with absolutely continuous distribution functions.

Suppose there is a $\kappa > 0$ such that, $E[\log |B_1|] < 0$, $E[|B_1|^\kappa] = 1$, $E[|B_1|^\kappa \log^+ |B_1|] < \infty$, $0 < E[|A_1|^\kappa] < \infty$.² Then the following hold:

1. The equation $X_\infty \stackrel{d}{=} A_1 + B_1 X_\infty$, X_∞ and (A_1, B_1) are independent, has a solution unique in distribution given by $X_\infty \stackrel{d}{=} \sum_{j=1}^{\infty} A_j \prod_{i=1}^{j-1} B_i$.
2. If in (9) we take $X_0 \stackrel{d}{=} X_\infty$, then the process X_t is stationary.
3. No matter how the process X_t is initialized $X_t \xrightarrow{d} X_\infty$.
4. The limits $\lim_{x \rightarrow \infty} x^\kappa \Pr(X_\infty > x)$ and $\lim_{x \rightarrow \infty} x^\kappa \Pr(X_\infty < -x)$ exist and are finite; at least one of these limits is strictly positive.

²The following notation $\log^+ |B_1| = \max(\log |B_1|, 0)$ is used.

The basic intuition for the Kesten theorem to hold is as follows.³ Kesten's theorem basically applies if there exists a κ such that $E[|B_1|^\kappa] = 1$. The essential idea behind this requirement is that the multiplicative innovations should be sufficiently spread out to allow realizations larger than one, such that past realization have a more than proportionate impact on the current realization. But this should not occur too frequently, otherwise the process becomes non-stationary. This implies that at least one of the tails of the stationary distribution will decline according to a power-law.

Hence, when the stochastic coefficients of the dynamic RBC equation given by (8) meet the conditions above, Kesten's theorem predicts that log income over time becomes spread out over the positive half axis and has a distribution with a power like upper tail. Standard macro has it that log of national income is normally distributed. Recent evidence, however, shows that the growth rates have a heavy upper tail. See for instance, Cumperayot and de Vries (2006) who calculate the tail index κ for several macroeconomic variables, including log income, using time series data for various countries.⁴

In the remainder of this section we focus on calculating the approximate solution of the RBC equation (8) and study how it compares with the original process. When the system of equations (1-4) does not have a closed form solution, the standard approach is to consider a linearization of the Euler equations around the steady state of the model. For instance, when the capital does not depreciate fully, Campbell (1994) takes the log-approximation of the capital accumulation equation (2). We depart from this practice, since in our model the nonlinearity stems from the learning by doing specification for technological progress. In particular, the nonlinearity is generated by two random variables that enter multiplicatively in (7). The standard log-linearization would linearize equations that need not be linearized, without solving the stochastic nonlinearity. Since the nonlinearity from the stochastic learning by doing is transferred directly in the relation for log-income (8), we take the approximation at this point.

³For the interested reader, Embrechts, Kluepelberg and Mikosch (1997, ch.8.4) contains an accessible summary treatment.

⁴Recall that Pareto's power law is derived from the observed distribution of individual income levels in a given period, while the RBC model studied here derives the distribution of the log of national income over time. Within the current model Pareto's law can be captured by assuming that the random variable ϕ is individualized and is Pareto distributed. Note that this in no way drives the result for the distribution of log of national income over time, as the (individualized) $\log\phi$ would be exponentially distributed.

To this end, rewrite (9) as follows:

$$y_{t+1} = a + m_{t+1} + (b + n_{t+1})y_t, \quad (10)$$

where $a = (1 - \alpha) \log \beta(1 - \alpha) + \alpha \log \phi$ and $m_t = \alpha \log \phi_t - \alpha \log \phi$, while $b = \alpha\varepsilon + 1 - \alpha$, $n_t = \alpha\varepsilon_t - \alpha\varepsilon$, such that m_t and n_t are mean zero random variables (assuming the mean exists).

To eliminate the nonlinearity from equation (10) due to the multiplicative random shock $(b + n_{t+1})$, we apply a first order Taylor expansion around the stochastic steady state:

$$y_{t+1} = a + m_{t+1} + by_t + \frac{a}{1-b}n_{t+1}. \quad (11)$$

If the deterministic part b of the multiplicative shock is less than one, the linearized equation for log income (11) is a stationary AR(1) process. However, $b < 1$ is not a necessary requirement for the Kesten result to hold. As $b > 1$ the AR(1) approximation becomes non-stationary; but (10) may still have a stationary limit distribution.

The starkest contrast between the linearized solution and the solution to the original model is obtained if we assume that the innovations ε_t and ϕ_t have a distribution with bounded support. For example, consider m_t and n_t uniformly distributed on the intervals $[-m, m]$ and $[-b, b]$ respectively, where it is assumed that $0 < m < a$ and $0 < b < 1$. We show that y_t generated by (11) has a bounded support, while (10) implies an equilibrium distribution of y_t with unbounded support. Taking $y_0 = 0$ and iterating (11) we obtain:

$$y_t = a\left(\sum_{i=1}^t b^{t-i}\right) + \left(\sum_{i=1}^t m_i b^{t-i}\right) + \frac{a}{1-b}\left(\sum_{i=1}^t n_i b^{t-i}\right). \quad (12)$$

Setting $m_t = m$ and $n_t = b$ we get

$$y_t \leq \max y_t \leq \frac{a}{1-b} + \frac{m}{1-b} + \frac{ab}{(1-b)^2}. \quad (13)$$

Thus y_t is bounded from above.

Turning to (10), if we can show that the Kesten theorem 1 applies, it would immediately follow that the support of the equilibrium distribution is unbounded due to the Pareto

type tail. From the assumptions it follows that B_1 is uniformly distributed on the interval $[0, 2b]$. In other words we need to show that the equation in κ

$$E[B_1^\kappa] = \int_0^{2b} t^\kappa \frac{1}{2b} dt = \frac{(2b)^\kappa}{(\kappa + 1)} = 1 \quad (14)$$

has a strictly positive solution. This holds for any $b \in (1/2, e/2)$.⁵ For example, it can be easily checked that the pair $(\kappa, b) = (2, \sqrt{3}/2)$ satisfies the equation.

The second case we consider gives less stark of a contrast if we assume that $(a + m_t)$ and $(b + n_t)$ are exponentially distributed random variables with respective means a and b . The backward iteration of the process in eq. (11) implies that y_t depends on two weighted sums of exponentially distributed random variables. Proposition 1 below describes the behavior at the limit of a weighted sum of random variables that follow a Gamma distribution. Since the exponential distribution is a special case of the Gamma distribution, we appeal to the this proposition to infer that the approximate solution y_t follows a distribution with exponential declining tail. The original model, however, has a solution for which the stationary distribution exhibits power decline.

We use the remainder of the section to show that the Kesten theorem applies when $A_t = a + m_t$ and $B_t = b + n_t$ are exponentially distributed. It is sufficient to show that the three conditions of the Kesten theorem hold. First we argue that $E[\log B_1] < 0$. Given that $E[\log B_1] = \int_0^\infty \ln(x) \frac{1}{b} e^{-x/b} dx$, by a transformation of variable

$$E[\log B_1] = \int_0^\infty (\ln y) e^{-y} dy + \ln b \int_0^\infty \frac{1}{b} e^{-x/b} dx = -\mathbb{C} + b \ln b$$

if $b \in (0, 1)$, and where \mathbb{C} denotes Euler's constant. Clearly $-\mathbb{C} + b \ln b < 0$, for $b < 1$. Second, we show that the main condition is fulfilled. Note that

$$E[B_1^\kappa] = \int_0^\infty x^\kappa \frac{1}{b} e^{-x/b} dx = b^\kappa \int_0^\infty y^\kappa e^{-y} dy = b^\kappa \Gamma(\kappa + 1)$$

⁵One needs $b < e/2$ for $E[\log |B|] = \ln 2b - 1$ to be negative. We are grateful to a referee to point out that the upper bound can be as high as $e/2$. Note furthermore that $E[|B|^\kappa \log^+ |B|]$ equals $b \ln 2b - b/2$ and hence is finite for $b \in (1/2, e/2)$; moreover, since $0 \leq |a + m_{t+1}| \leq a + m$ and given the uniform distribution of m_{t+1} , it follows that $0 < (a - m)^\kappa < E[|A|^\kappa] < (a + m)^\kappa < \infty$.

Since $\Gamma(2) = 1$ and $\Gamma(\kappa + 1) = \kappa!$ eventually overtakes $(1/b)^\kappa$ as κ increases, the equation $E[B_1^\kappa] = 1$ has a solution. Furthermore, since

$$E[B^\kappa \log^+ B] < E[B^{\kappa+1}] = b^{\kappa+1} \Gamma(\kappa + 2) < \infty$$

$E[|B|^\kappa \log^+ |B|] < \infty$ is satisfied for $b \in (0, 1)$. Lastly, the condition $0 < E[A^\kappa] < \infty$ follows from a similar argument.

Thus if one analyzes the original dynamic RBC equation (8), one finds that log income becomes spread out over the positive half axis and, moreover, has a distribution with a heavy Pareto upper tail. Under the same conditions however, when one starts from the approximation (11), income either remains bounded, or exhibits exponentially thin tails.

3 Conditional volatility in the CAPM model

In this section we illustrate with another example how the linearization affects the stochastic properties of the original model. We consider an intertemporal version of the Capital Asset Pricing Model (CAPM) as presented in Campbell, Lo and MacKinlay (1997, p.323 and p.494). The CAPM relates the expected return of an asset to the covariance of its return with the market portfolio return. In a dynamic setting, when applied to the market portfolio itself, the intertemporal CAPM model predicts that the expected market portfolio excess returns depend linearly on the variance of market portfolio.

It has been widely observed that periods of turbulence in the stock market are generally followed by further periods of turbulence, while periods of tranquility are followed by periods of tranquility. The intertemporal CAPM model can account for this empirical observation by allowing the variance of the market portfolio to be time varying. In fact, there is ample evidence that the volatility is an autoregressive process for which the ARCH process specification is a natural choice. Moreover, the ARCH process also captures the martingale property of the returns. The properties of the stationary solution of the ARCH process have direct implications for the equilibrium distribution of the returns of the market portfolio. We investigate how this equilibrium distribution is affected when both the ARCH specification and its linear approximation are considered.

Consider again the first order stochastic difference equation (9)

$$X_t = A_t + B_t X_{t-1}.$$

The ARCH process is related to (9) in the following way. Consider $q_t = \sigma_t s_t$, where $s_t \sim IID N(0,1)$ and where the variance σ_t^2 of q_t is driven by its own past squared $\sigma_t^2 = v + wq_{t-1}^2$; see Engle (1982). Squaring the mean equation $q_t^2 = \sigma_t^2 s_t^2$ and substituting the variance equation in this expression, gives

$$q_t^2 = v s_t^2 + w s_t^2 q_{t-1}^2. \quad (15)$$

Equating q_t^2 with X_t , $v s_t^2$ with A_t and $w s_t^2 q_{t-1}^2$ with B_t , yields the difference equation (9). The parameters of this difference equation are random variables that follow chi square distributions. The ARCH process is the most popular specification for volatility clusters in finance.

We first investigate the properties of the linearized version of (9). Let A_t and B_t in (9) have mean a and b respectively (where $0 < b < 1$, by assumption), so that we may write $A_t = a + m_t$ and $B_t = b + n_t$. Equation (9) only contains a single non-linear term. To approximate the non-linear term $B_t X_{t-1}$, we take again a first order Taylor expansion around the stochastic steady state with respect to X_{t-1} and n_t . This transforms (9) into:

$$X_t = a + m_t + b X_{t-1} + \frac{a}{1-b} n_t. \quad (16)$$

In the case of ARCH the analogous first order Taylor approximation will give the following AR(1) process for q_t^2 :

$$q_t^2 = v + w q_{t-1}^2 + \frac{v}{1-w} (s_t^2 - 1) \quad (17)$$

and where $a = v$, $m_t = v(s_t^2 - 1)$, $b = w$ and $n_t = w(s_t^2 - 1)$.

We show that the conditions of Theorem 1 apply to the ARCH process (15). In the case of ARCH, $B_1 = w s_1^2$, and we need to ascertain that there exists a κ such that $E[(w s_1^2)^\kappa] = 1$. Given the normality assumption regarding s_1 , we can rewrite this condition

as follows

$$\Gamma(\kappa + \frac{1}{2}) = \pi^{1/2}(2w)^{-\kappa}.$$

Note that for $w = 1$ the equation has the solution $\kappa = 1$. Since for given $(2w)^{-\kappa}$ is increasing as w decreases from 1 and the gamma function is increasing in κ , there exists a solution for any $0 < w < 1$ such that $\kappa > 1$. For example $(w, \kappa) = (1/\sqrt{3}, 2)$ is a solution. This solution implies that for $w > 1/\sqrt{3}$ the stationary solution of (15) has a finite variance (as $\kappa > 2$). By the fact that $\Pr(X_\infty > x) \sim cx^{-\kappa}$, the r -th moment exists as long as $\int_1^\infty x^r x^{-\kappa-1} dx$ is bounded, which requires $r < \kappa$. The other conditions can be verified to hold as well.⁶

Thus the remarkable feature of the ARCH process is that while the driving random variables have distributions with exponential type upper tails (chi-square distribution), the linear process has a solution which is distributed with a hyperbolic type upper tail implying moment failure. Per contrast, the approximation (17) has a distribution with all the moments finite. To show this, we first obtain a more convenient expression for q_t^2 by iterating (17) backwards

$$q_t^2 = v \frac{1 - w^t}{1 - w} + \frac{v}{1 - w} \sum_{i=1}^t w^{t-i} (s_i^2 - 1) + w^t q_0^2 \quad (18)$$

This expression indicates that the asymptotic behavior of q_t^2 is driven by the properties of the limit distribution of $\sum_{i=1}^t w^{t-i} s_i^2$. The next proposition summarizes a result that enables us to prove how the first-order approximation to the ARCH process changes the stochastic properties of the solution.

Proposition 1 *Let $S_t = \sum_{i=1}^t \lambda^{i-1} Q_i$, where Q_i is a random variable with a Gamma distribution with parameters A and B strictly positive and λ is a constant in the interval $(0, 1)$. Then the limit distribution of S_t as $t \rightarrow \infty$ exhibits thin tails.*

Proof. A proof is available in the Appendix of the working paper version. ■

The chi-square distributed random variables are a particular case of gamma distrib-

⁶De Haan et al. (1989) first linked the ARCH process to the Kesten equation and discussed the extremal behavior of the process. Engle (1982) already described the finite number of bounded moments.

uted random variables, so that the q_t^2 from the approximation (17) has all its moments finite. Theorem 1, however, implies that the original q_t^2 from (15) has heavy tails, so that the higher moments are unbounded. Thus while the first order approximation has in expectation a solution which is identical to the solution (in expectation) of the original ARCH process, the stochastic properties differ regarding the tail area of the distributions. Nevertheless, both the approximation and the ARCH process pick up the volatility clustering.

It follows that the equilibrium solution for the return of the market portfolio is affected by the linear approximation of the ARCH specification for the volatility. The original ARCH specification predicts fat-tailed returns, an empirical finding now widely accepted in the field. However, the linearization implies that the market returns would exhibit thin-tails as in the original CAPM with constant volatility and normally distributed returns. Even though (17) does pick up the clustering of volatilities, it fails to induce heavy tails.

Connecting the ARCH and SV model

We discuss the properties of the stochastic steady state of the ARCH model in relation with a stochastic volatility model (SV), as introduced by Harvey et al. (1994). A comparison of the two models builds on the approximation discussed above. In essence, we show that the limiting behavior of the SV model and the loglinearized version of the ARCH model can be identical.

Stochastic volatility models are a popular alternative for the ARCH models. The SV model has the advantage that it can allow for asymmetric leverage effects. In the SV model, the volatility process is rewritten as

$$q_t^2 = s_t^2 e^{h_t}, \tag{19}$$

where, using the previous notation, $h_t = \log \sigma_t^2$. The difference with the standard ARCH model is the way in which the variance σ_t^2 of q_t is driven by its own past and an exogenous innovation z_t :

$$h_t = a + bh_{t-1} + z_t. \tag{20}$$

Here z_t is assumed to be IID $N(u, \phi)$. To this point we do not make any assumption on the distribution of s_t .

By taking logarithms in (19) we can rewrite the process for the variance (20) as

$$\log q_t^2 - \log s_t^2 = a + b \log q_{t-1}^2 - b \log s_{t-1}^2 + z_t.$$

It is then straightforward to show through backward iteration that

$$\log q_t^2 \rightarrow \log s_t^2 + \frac{a}{1-b} + \sum_{i=0}^{\infty} b^i z_{t-i}, \quad (21)$$

where the convergence is in probability.

The standard ARCH process $q_t = \sigma_t s_t$, with variance $\sigma_t^2 = v + wq_{t-1}^2$, was given in (15). While so far we have concentrated on the effects of straightforward linearization of the process given in (15), for a sharper comparison with the stochastic volatility model we analyze the solution of the log-linearized ARCH model. As before, the non-stochastic steady state is $q^2 = v/(1-w)$ and it is assumed that $E[s_t^2] = 1$. Log-linearization around the steady state gives

$$q^2 \log \frac{q_t^2}{q^2} = [v + wq^2] s^2 \log \frac{s_t^2}{s^2} + w s^2 q^2 \log \frac{q_{t-1}^2}{q^2}.$$

Iterating backwards gives

$$\log \frac{q_t^2}{q^2} = \sum_{i=0}^{\infty} w^i \log s_{t-i}^2. \quad (22)$$

We show that for certain choices of parameters and innovations the stationary solution (22) can be made identical to the limit distribution of (21). To this end, suppose that $s_t = Q\sqrt{x_t}$, where the x_t are lognormally distributed with mean 1 and variance $e - 1$; and Q is a Bernoulli random variable that equals 1 with probability 1/2 and -1 with probability 1/2. This implies that $\log s_t^2 \sim N(-1/2, 1)$. Under this assumption we show that the standard ARCH process implies a limit distribution for q_t^2 that is heavy tailed, while both the log-linearization of the ARCH process and the SV process yield solutions that are distributed with an exponential type tail.

The standard ARCH model yields a martingale for the log stock prices that can be positive and negative. Moreover one shows that Kesten's theorem applies as $E[(ws^2)^\kappa] = 1$ requires that

$$w^\kappa e^{-\kappa+\kappa^2/2} = 1$$

has a solution. By taking logs, one sees that this equation has a non-trivial root $\kappa = 2(1 - \log w) > 0$ (recall that $w \in (0, 1)$). The other conditions are easily verified: $E[\log(ws^2)] = \log w - 1/2 < 0$; furthermore, $E[(ws^2)^\kappa \log(ws^2)] < E[(ws^2)^{\kappa+1}] < \infty$ and $E[(ws^2)^\kappa] < \infty$. Thus by Kesten's theorem it follows that the stationary solution of this specific ARCH model is heavy tailed.

The log-linearized version of the ARCH model, however, has a stationary solution that follows a normal distribution

$$\log \frac{q_t^2}{q^2} \sim N\left(-\frac{1/2}{1-w}, \frac{1}{1-w^2}\right). \quad (23)$$

The stochastic volatility model also has the a normal distribution as its stationary solution

$$\log q_t^2 \sim N\left(\frac{u+a}{1-b} - \frac{1}{2}, 1 + \frac{\phi}{1-b^2}\right).$$

This implies that⁷

$$\log \frac{q_t^2}{q^2} \sim N\left(\frac{u+a}{1-b} - \frac{1}{2} - e^{\frac{u+a}{1-b} + \frac{1}{2} \frac{\phi}{1-b^2}}, 1 + \frac{\phi}{1-b^2}\right). \quad (24)$$

Compare the expressions for the stationary solution to the SV model in (24) to the stationary solution of the log-linearized ARCH model (23). There are a sufficient degrees of freedom in the choices of parameters such that the two limit distributions can be made identical. To conclude, the SV model can be seen as a linearized version of the ARCH model.

As we argued in the beginning, both the ARCH and SV models capture the time varying nature of the volatility. Since the intertemporal CAPM predicts that the expected excess return is linear in the variance of the market portfolio, it then follows that the

⁷Since q_t^2 is lognormally distributed we have $E[q_t^2] = \exp\left(\frac{u+a}{1-b} + \frac{1}{2} \frac{\phi}{1-b^2}\right)$.

unconditional distribution of the returns is thin or heavy tailed depending on whether the SV (the approximate ARCH), or the original ARCH process is used to model the volatility process.

4 Discussion

In this paper we have proposed two examples to analyze the failures of the first order stochastic approximation. Our focus has been to contrast the stochastic properties of the approximate solution with the stochastic properties of the original model. We study two simple frameworks that allow us to compare the approximation and the original model without taking recourse to simulations.

The first application considers a basic RBC model with full depreciation of capital and log utility function. Admittedly, these assumptions are stylized. For instance, it is more common in applied analysis to assume that the capital depreciates at a rate lower than unity. However, imposing simple conditions is necessary to derive a closed form solution for the RBC model (cf. Campbell, 1994). The transfer of capital between periods introduces nonlinearity in the system (1-4), which, in this case, needs to be solved by taking a log-linearization. The closed form solution gives us a useful benchmark to study the properties of the approximate solution in a meaningful way.

In our basic RBC model, the nonlinearity has a stochastic dimension. This becomes clear when looking at the difference equation that characterizes the process for log-income (8), where both coefficients are stochastic. In essence, the nonlinearity is generated by two random variables that enter multiplicatively in (8). The first order approximation will transform the equation such that all random variables enter additively. In our example, the source of nonlinearity stems from the learning by doing specification for technological progress. More generally, any RBC model that can be characterized through linear stochastic difference equations adhere to Kesten's result.

The second application shares similar features with the RBC model. Namely, the nonlinearity in the ARCH model has the same multiplicative-stochastic nature. The approximation works again to separate the two random variables q_{t-1}^2 and s_t in (15). This

explains why the solution obtained from log-linearization of ARCH has similar stochastic properties as the solution for the stochastic volatility model. Both the approximated model and the stochastic volatility model are based on a linear first order difference equation with deterministic coefficients.

5 Conclusions

The solution of a stochastic macro model is usually determined through a linearization around the associated deterministic steady state. Recently, a significant number of papers has thoroughly examined the errors that could potentially be made by such an approximation. This literature, however, is mainly preoccupied with the analysis of the deterministic part of the approximate solution.

Parallel to this literature, we have studied what are the effects of the linearization on the stochastic properties of the original model. To this end we have solved the simplest model in the business cycle literature with fixed labour supply, total depreciation of capital, a log-utility function and noisy learning by doing. We showed that the solution of the resulting stochastic difference equation yields a distribution for the log of income over time which is stationary, exhibits moment failure and has an unbounded upper support. The approximation, however, has a stationary distribution with bounded support and all moments finite. As a second example we considered the CAPM asset pricing equilibrium model from finance with stochastic volatility.

To conclude, an approximation of the stochastic part of an equilibrium model needs to be considered with the same care as an approximation of the deterministic part. Although often disregarded, the approximation can alter the global data features in equilibrium dramatically.

Acknowledgments

The authors are grateful for the helpful suggestions by two anonymous referees and Antonio Di Cesare.

References

Arouba, S. B., Fernandez-Villaverde, J., Rubio-Ramirez, J. F., 2006. Comparing Solution Methods for Dynamic Equilibrium Economies. *Journal of Economic Dynamics and Control* 30, 2477–2508.

Campbell, J., 1994. Inspecting the Mechanism: An Analytical Approach to the Stochastic Growth Model. *Journal of Monetary Economics* 33, 463–506.

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

Cumperayot, P., de Vries, C. G., 2006. Large Swings in Currencies Driven by Fundamentals. Tinbergen Institute Discussion Papers 06-086/2.

Dotsey, M., Mao, C. S., 1992. How Well Do Linear Approximation Methods Work? *Journal of Monetary Economics* 29, 25–58.

Embrechts, P., Kluepelberg, C., Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*. Springer Verlag, Berlin.

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

Fernandez-Villaverde, J., Rubio-Ramirez, J. F., 2005. Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood. *Journal of Applied Econometrics* 20, 891–910.

Harvey, A., Ruiz, E., Shephard, N., 1994. Multivariate Stochastic Variance Models. *Review of Economic Studies* 61, 247–64.

Jin, H., Judd, K., 2002. Perturbation Methods for General Dynamic Stochastic Models. Discussion paper, Stanford University.

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

Kim, J., Kim, S., 2003. Spurious Welfare Reversals in International Business Cycle Models. *Journal of International Economics* 60, 471–500.

Kim, S., 1997. International Business Cycles in a Partial-Risk-Sharing Market with

Capital Income Taxation. Ph.D. thesis, Yale University.

Kydland, F., Prescott, E., 1982. Time to Build and Aggregate Fluctuations. *Econometrica* 50, 1345–1370.

Pratt, J., 1960. On Interchanging Limits and Integrals. *The Annals of Mathematical Statistics* 31, 74–77.

Sutherland, A., 2002. A Simple Second-Order Solution Method for Dynamic General Equilibrium Models. CEPR Discussion Paper No 3554.

Tesar, L., 1995. Evaluating the Gains from International Risksharing. *Carnegie-Rochester Conference Series on Public Policy* 42, 95–143.

Uhlig, H., 1997. A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily. Working paper, Tilburg University.

A Appendix

In what follows we prove the result described by Proposition 1: An infinite weighted sum of Gamma random variables (rvs), which include the chi-square and exponential distributed rvs, has a limit distribution with all moments bounded. Let $S_t = \sum_{i=1}^t \lambda^{i-1} Q_i$, where Q_i belongs to a Gamma family with parameters A and B strictly positive. We need to study the tail behavior of $S = \lim_{t \rightarrow \infty} S_t$. We first prove that all central moments of S are finite, by using the relationship that exists between the cumulants of a distribution and its moments. Let $M_{S_t}(x)$ be the moment generating function (mgf) of S_t , while $\ln M_{S_t}(x)$ is the cumulant generating function (cgf). From the properties of the mgf of a gamma distribution it follows that

$$M_S(x) = \lim_{t \rightarrow \infty} \left[\prod_{i=1}^t (1 - \lambda^{i-1} Bx) \right]^{-A}. \quad (25)$$

Lemma 1 *The moment generating function $M_S(x)$ and the cumulant generating function $\ln M_S(x)$ exist and are finite.*

Proof. Indeed, the sequence $(f_t)_{t \geq 0}$, with $f_t(x) = \prod_{i=1}^t (1 - \lambda^{i-1} Bx)$, is decreasing (that is, for every x , $f_t(x) > f_{t+1}(x)$ for any t) and bounded (for every x , $f_t(x) \in (0, 1)$ for any t)⁸. Given these two conditions, the *monotone convergence theorem* implies the convergence of $(f_t)_{t \geq 0}$ to a finite limit. Moreover, $f_t(x)$ converges *pointwise* to an $f(x) < \infty$. The moment generating function can be derived easily as $M_S(x) = f(x)^{-A}$. It also follows that the cumulant generating function is given by $\ln M_S(x) = -A \ln f(x)$. ■

To determine the cumulant κ_m , which is the coefficient of $x^m/m!$ in the Taylor expansion of the logarithm of the mgf of S , we need the derivative of $\ln M_S(x) = -A \lim_{t \rightarrow \infty} \sum_{i=1}^t \ln(1 - \lambda^{i-1} Bx)$. Pratt's theorem (1960)⁹ allows us to interchange limits and integrals.

Theorem 2 *If:*

1. $F_t \xrightarrow[t \rightarrow \infty]{} F$, $G_t \xrightarrow[t \rightarrow \infty]{} G$ and $H_t \xrightarrow[t \rightarrow \infty]{} H$ pointwise;

⁸By assumption, we have $\lambda \in (0, 1)$. In addition, for the moment generating function to exist, x has to be smaller than $1/B$. Moreover, we restrict the domain of $M_S(x)$ only to positive numbers. As a result we have the constraint: $0 < 1 - \lambda^{i-1} Bx < 1$

2. $F_t \leq G_t \leq H_t$ for all t ;

3. $\int H_t \xrightarrow{t \rightarrow \infty} \int H$ and $\int F_t \xrightarrow{t \rightarrow \infty} \int F$ with $\int H$ and $\int F$ finite;

then $\int G_t \xrightarrow{t \rightarrow \infty} \int G$ and $\int G$ is finite.

Informally, this criterion states that a convergent sequence permits exchange of limits and integrals if it is bracketed by two sequences which permit this exchange. From here, we move forward to showing how to interchange limits and derivatives. This result is summarized in Lemma 2.

Lemma 2 *Let $g_t(x) = -A \ln f_t(x)$, where $f_t(x)$ was defined above. It is then possible to interchange limits and derivatives. In other words, the following equation holds for all*

$$m \geq 1: \frac{\partial^m \left[\lim_{t \rightarrow \infty} g_t(x) \right]}{\partial x^m} = \lim_{t \rightarrow \infty} \frac{\partial^m g_t(x)}{\partial x^m}.$$

Proof. We will prove that the first derivative of $g_t(x)$ with respect to x , satisfies the conditions of Theorem 2.

The m th derivative of $g_t(x)$ with respect to x is given by $\frac{\partial^m g_t(x)}{\partial x^m} = AB^m (m-1)! \sum_{i=1}^t \frac{(\lambda^{i-1})^m}{(1-\lambda^{i-1}Bx)^m}$. One can easily check that $\frac{\partial^m g_t(x)}{\partial x^m}$ is increasing in t for every x and for every m . Note also that $\frac{\partial^m g_t(x)}{\partial x^m}$ is positive, as $0 < 1 - \lambda^{i-1}Bx < 1$. Knowing that $\lambda^i < 1$, it follows that $\left(\frac{\lambda^i}{1-\lambda^i Bx}\right)^m < \left(\frac{\lambda^i}{1-Bx}\right)^m$, which gives further $\sum_{i=0}^{t-1} \left(\frac{\lambda^i}{1-\lambda^i Bx}\right)^m < \frac{1}{(1-Bx)^m} \sum_{i=0}^{\infty} (\lambda^m)^i$. Thus $\frac{\partial^m g_t(x)}{\partial x^m}$ is bounded from above as well.

$$0 < \frac{\partial^m g_t(x)}{\partial x^m} < AB^m (m-1)! \frac{1}{(1-Bx)^m} \frac{1}{1-\lambda^m}$$

Again, the *monotone convergence theorem* implies that the sequence $\frac{\partial^m g_t(x)}{\partial x^m}$ is convergent when t goes to infinity. Furthermore the convergence is *pointwise*.

Taking $m = 1$, one can choose $G_t = \frac{\partial g_t(x)}{\partial x}$, $F_t = 0$, $H_t = AB \frac{1}{1-Bx} \frac{1}{1-\lambda}$, so that Theorem 2 applies, and hence $\int \frac{\partial g_t(x)}{\partial x} dx \xrightarrow{t \rightarrow \infty} \int G dx$. As $\int \frac{\partial g_t(x)}{\partial x} dx = g_t(x)$, and since Lemma 1 shows that $g_t(x)$ converges and its unique limit is the cumulant generating function $-A \ln f(x)$, it follows that $\int G dx = -A \ln f(x)$. This implies further that $G = \frac{\partial(-A \ln f(x))}{\partial x}$, or, in other words, $\frac{\partial \left[\lim_{t \rightarrow \infty} g_t(x) \right]}{\partial x} = \lim_{t \rightarrow \infty} \frac{\partial g_t(x)}{\partial x}$, which is the relationship we wanted to obtain,

for the case $m = 1$. Similar arguments are used for $m = 2, 3$, etc. so we can finally have

$$\frac{\partial^n \left[\lim_{t \rightarrow \infty} g_t(x) \right]}{\partial x^m} = \lim_{t \rightarrow \infty} \frac{\partial^m g_t(x)}{\partial x^m}. \quad \blacksquare$$

By Lemma 2 it becomes apparent that $\frac{\partial^m \ln M_S(x)}{\partial x^m} = \lim_{t \rightarrow \infty} \frac{\partial^m g_t(x)}{\partial x^m}$, and thus the m th cumulant is given by:

$$\kappa_m = \left. \frac{\partial^m \ln M_S(x)}{\partial x^m} \right|_{x=0} = AB^m (m-1)! \lim_{t \rightarrow \infty} \sum_{i=1}^t (\lambda^{i-1})^m = AB^m (m-1)! \frac{1}{1-\lambda^m}. \quad (26)$$

The m -th moment of a distribution is a polynomial in the $\kappa_1, \kappa_2, \dots, \kappa_m$. It follows that all moments of the limit distribution are bounded. Q.E.D.