Existence, Uniqueness and Stability of Invariant Distributions in Continuous-Time Stochastic Models

Christian Bayer[a] and Klaus Wälde[b]

[a] University of Vienna and
[b] University of Mainz, CESifo and Université catholique de Louvain[1]

July 2011

We study a dynamic stochastic general equilibrium model in continuous time. Related work has proven that optimal consumption in this model is a smooth function of state variables. This allows us to describe the evolution of optimal state variables (wealth and labour market status) by stochastic differential equations. We derive conditions under which an invariant distribution for state variables exists and is unique. We also provide conditions such that initial distributions converge to the long-run distribution.

JEL Codes:	C62, D91, E24, J63
Keywords:	uncertainty in continuous time, Poisson process, existence, uniqueness, stability

1 Introduction

Dynamic stochastic general equilibrium models are widely used for macro economic analysis and also for many analyses in labour economics. When the development of these models started with the formulation of stochastic growth models, a lot of emphasis was put on understanding formal properties of these models. Does a unique solution exist, both for the control variables and general equilibrium itself? Is there a stationary long-run distribution (of state variables being driven by optimally chosen control variables) to which initial distributions of states converge? This literature is well developed for discrete time models (see below for a short overview). When it comes to continuous time models, however, only initially there were some articles looking at stability issues (Merton, 1975; Bismut, 1975; Magill, 1977; Brock and Magill, 1979; Chang and Malliaris, 1987). In recent decades, applications to economic questions have been the main focus. This does not mean, however, that all formal problems have been solved. In fact, we argue in this paper that formal work is badly

1

missing for continuous time uncertainty which has strong implications for applied work.

An area in economics where continuous-time stochastic models are especially popular is the labour-market search and matching literature. While all of these models are dynamic and stochastic by their basic structure (new job opportunities arise at random points in time only and/or new wage offers are random), there has been very little effort in understanding the stability properties of their distributional predictions.[2] This might be due to two reasons. Some papers work with a law of large numbers right from the beginning which allows to focus on means. The classic example is Pissarides (1985). Papers in the tradition of Burdett and Mortensen (1998) that do focus on distributions construct these distributions by focusing on "steady states", i.e. on distributions which are time-invariant.[3]

While this approach is extremely fruitful to understand a lot of important issues, one might also want to understand how distributions evolve over time.[4] Once this becomes the objective, the issue of stability and uniqueness of a long-run stationary distribution becomes central.

The goal of this paper is threefold: First, we introduce methods for analysing existence and stability of distributions described by stochastic differential equations from the mathematical literature. The approach to proving the existence and uniqueness of an invariant distribution and its ergodicity, i.e. of convergence to the said distribution, builds on the work of Meyn and Tweedie (1993 a,b,c) and Down et al. (1995). Their work is especially useful for understanding properties of systems driven by jump processes.[5] The methods we use here would be the ones that can be applied to the search and matching analyses cited above.

Second, we use these methods to analyse stability properties of a model of search and matching where individuals can smooth consumption by accumulating wealth. The model was originally developed and analyzed in a companion paper (Bayer and Wälde, 2011). Individuals have constant relative risk aversion and an infinite planning horizon.[6] Optimal behaviour implies that the two state variables of an individual, wealth and employment status, follow a process described by two stochastic differential equations. We analyse under which conditions a distribution for wealth and employment status exists, is unique and converges to an invariant long-run distribu-

---

[2]The literature on assortative matching carefully analyses existence and uniqueness issues (see e.g. Shimer and Smith, 2000 or Lentz, 2010). These analyses do not take the dynamics of distributions into account.

[3]Papers in this tradition include Postel-Vinay and Robin (2002), Cahuc et al. (2006), Moscarini (2005) and Burdett et al. (2011), to name just a few.

[4]There is a very recent strand of the literature which inquires into the evolution of distributions over the business cycle. Examples include Moscarini and Postel-Vinay (2008, 2010), Coles and Mortensen (2011) and Kaas and Kircher (2011). These papers all analyse the evolution of distributions over time.

[5]These methods are also used for understanding how to estimate models that contain jumps (e.g. Bandi and Nguyen, 2003) or for understanding long-term risk-return trade-offs (Hansen and Scheinkman, 2009).

[6]See related work by Alvarez and Shimer (2011) who allow for risk-sharing within large households, endogenous labour supply and a differentiated production structure. They do not include a savings decision.

tion. The corresponding theorem is proven.

Our third objective consists in illustrating the usefulness of the methods presented here for other stochastic continuous-time models. The examples we have in mind (see below) mainly come from macro, labour, finance and international finance. Illustrating how these methods can be applied in other contexts should make it easy to develop some intuition for understanding which models with evolving distributions predict stability.

One crucial component of our proofs is a smoothing condition. As we allow for Poisson processes, we have to use more advanced methods based on T-processes than in the case of a stochastic differential equation driven by a Brownian motion. In the latter case the strong smoothing properties of Brownian motion can be used to obtain the strong Feller property. In this sense, the corresponding analysis will often be more straightforward than the one presented here. For the wealth-employment process of Bayer and Wälde, we find that the wealth process is not smoothing and the strong Feller property does not hold. However, for the economically relevant parameter case (the low-interest rate regime), we can still show a strong version of recurrence (namely Harris recurrence) by using a weaker smoothing property, and thus obtain uniqueness of the invariant distribution. Ergodicity is then implied by properties of discrete skeleton chains.

Let us relate our approach to the more formal literature. In discrete time models, a classic analysis was undertaken by Brock and Mirman (1972). They analyse an infinite-horizon optimal stochastic growth model with discounting where uncertainty results from total factor productivity. They show inter alia that "the sequence of distribution functions of the optimal capital stocks converges to a unique limiting distribution." Methodologically, they use parts of the classical stability theory of Markov chains, but mainly rely on properties of their model. A nice presentation of stability theory for Markov processes with a general state space is by Futia (1982). He uses an operator-theoretic approach exploiting results from the theory of continuous linear operators on Banach spaces. Hopenhayn and Prescott (1992) analyse existence and stability of invariant distributions exploiting monotonicity of decision rules that result from optimal behaviour of individuals. Their approach mainly relies on fixed point theorems for increasing maps and increasing operators on measures (in the sense of stochastic dominance). Bhattacharya and Majumdar (2004) provide an overview of results concerning the stability of random dynamic systems with a brief application to stochastic growth. Nishimura and Stachurski (2005) present a stability analysis based on the Foster–Lyapunov theory of Markov chains. For a survey of stochastic optimal growth models, see Olson and Roy (2006). In the literature on precautionary savings, Huggett (1993) analyses an exchange economy with idiosyncratic risk and incomplete markets. Agents can smooth consumption by holding an asset and endowment in each period is either high or low, following a stationary Markov process. This structure is similar in spirit to our setup. Huggett provides existence and uniqueness results for the value function and the optimal consumption function and shows that there is a unique long-run distribution function to which initial distributions converge. Regarding stability, he relies on the results of Hopenhayn and Prescott (1992).

The theory we will employ below provides a useful contribution to the economic

literature as the latter, as just presented, focuses on related, but different methods. For one, we treat Markov processes in continuous time, while references in the macro-economic literature in the context of Markov-process stability are mostly related to discrete time.[7] But even in discrete time, the theory of Meyn and Tweedie and their coauthors differs from the approaches cited above. Indeed, Meyn and Tweedie's theory is a direct generalization of the classical stability theory for Markov chains in discrete time, with finite, discrete state space.

In the economic continuous–time literature, the starting point is Merton's (1975) analysis of the continuous-time stochastic growth model. For the case of a constant saving rate and a Cobb-Douglas production function, the "steady-state distributions for all economic variables can be solved for in closed form". No such closed form results are available of course for the general case of optimal consumption. Chang and Malliaris (1987) also allow for uncertainty that results from stochastic population growth as in Merton (1975) and they assume the same exogenous saving function where savings are a function of the capital stock. They follow a different route, however, by studying the class of strictly concave production functions (thus including CES production function and not restricting their attention to the Cobb-Douglas case). They prove "existence and uniqueness of the solution to the stochastic Solow equation". The build their proof on the so-called Reflection Principle. More work on growth was undertaken by Brock and Magill (1979) building on Bismut (1975). Magill (1977) undertakes a local stability analysis for a many-sector stochastic growth model with Brownian motions using methods going back to Rishel (1970). All of these models use Brownian motion as their source of uncertainty and do not allow for Poisson jumps. To the best of our knowledge, not much (no) work has been done on these issues since then.

The structure of our paper is as follows. The next section presents the model. Section 3 introduces the mathematical background for our existence and stability analysis. We remind the reader of familiar results from discrete-time stability analysis and then introduce corresponding concepts in continuous time. Section 4 applies these methods to our model and proves existence and uniqueness of an invariant measure for the process describing the state variables. Section 4 also proves that initial conditions converge to the long-run invariant distribution. Section 5 applies our methods to three stochastic processes originating from a long list of other papers in the literature. Section 6 concludes.

2   The model

Consider an individual that maximizes a standard intertemporal utility function, $U(t) = E_t \int_t^\infty e^{-\rho[\tau-t]} u(c(\tau)) d\tau$, where expectations need to be formed due to the uncertainty of labour income which in turn makes consumption $c(\tau)$ uncertain. The expectations operator is E_t and conditions on the current state in t. The planning horizon starts in t and is infinite. The time preference rate ρ is positive. We assume

---

[7]As already remarked earlier, continuous time models are treated thoroughly, but under different conditions, in the finance literature.

that the instantaneous utility functions has a CRRA structure $u\left(c\left(\tau\right)\right) = \frac{c(\tau)^{1-\sigma}-1}{1-\sigma}$ with $\sigma \neq 1$. All proofs for the logarithmic case $\sigma = 1$ should work accordingly.

Each individual can save in an asset a. Her budget constraint reads

$$da\left(t\right) = \left\{ra\left(t\right) + z\left(t\right) - c\left(t\right)\right\} dt. \tag{1}$$

Per unit of time dt wealth $a\left(t\right)$ increases (or decreases) if capital income $ra\left(t\right)$ plus labour income $z\left(t\right)$ is larger (or smaller) than consumption $c\left(t\right)$. Labour income $z\left(t\right)$ is given either by constants w or b[8] and is intuitively described by the second constraint of the household, a stochastic differential equation,

$$dz\left(t\right) = \Delta dq_\mu - \Delta dq_s, \qquad \Delta \equiv w - b. \tag{2}$$

The Poisson process q_s counts how often our individual moves from employment into unemployment. The arrival rate of this process is given by $s > 0$ when the individual is employed and by $s = 0$ when the individual is unemployed. The Poisson process related to job finding is denoted by q_μ with an arrival rate $\mu > 0$ when unemployed and $\mu = 0$ when employed (as there is no search on the job). It counts how often the individual finds a job. In effect, $z(t)$ is a continuous time Markov chain with state space $\{w, b\}$, where the transition $w \to b$ happens with rate s and the transition $b \to w$ with rate μ. This description of z will be used in the remainder of the paper. As usual, the wealth-employment process (a, z), is defined on a probability space $(\Omega, \mathcal{F}, P)$.

When we allow the individual to optimally choose the consumption level $c\left(t\right)$, we make the following assumptions (see Bayer and Wälde, 2011 for details): (i) The interest rate is lower than the time-preference rate, $r < \rho$, (ii) optimal consumption $c\left(a, z\right)$ is continuously differentiable on $[-b/r, a_w^*] \times \{w, b\}$ in a, (iii) relative consumption $c(a, w)/c(a, b)$ is continuously differentiable in a and the derivative changes its sign only finitely often in every finite interval and (iv) the initial wealth $a(t)$ is chosen inside the interval $[-b/r, a_w^*]$. The discussion in Bayer and Wälde (2011, see the phase diagram there for an illustration) has revealed that wealth $a(\tau) \in]-b/r, a_w^*[$ is increasing when $z(\tau) = w$ and decreasing when $z(\tau) = b$. Moreover, wealth will never leave the interval $[-b/r, a_w^*]$.

3 Ergodicity results for continuous time Markov processes

The wealth-employment process $(a(\tau), z(\tau))$ is a continuous-time Markov process with a non-discrete state space $[-b/r, a_w^*] \times \{w, b\}$. Thus, we will rely on results from the general stability theory of Markov processes as presented in the works of Meyn and Tweedie and their coauthors cited above. In the present section, we will recapitulate

---

[8]In the full model of Bayer and Wälde (2011), w and b are endogenous functions of the equilibrium capital stock. This is of no importance for the distributional analysis here and we therefore immediately assume $w > b$ and b to be constant.

the most important elements of the stability for Markov processes in continuous time. Here, we will discuss the theory in full generality, i.e., we assume that we are given a Markov process $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ on a state space $\mathbf{X}$, which is assumed to be a locally compact separable metric space endowed with its Borel σ-algebra. All Markov processes are assumed to be time-homogeneous, i.e., the conditional distribution of X_{t+s} given $X_t = x$ only depends on s, not on t.

3.1 Preliminaries

Let $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ be a (homogeneous) Markov process with the state space $\mathbf{X}$, where $\mathbf{X}$ is assumed to be a locally compact and separable metric space, which is endowed with its Borel σ-algebra $\mathcal{B}(\mathbf{X})$. Let $P^t(x, A)$, $t \geq 0$, $x \in \mathbf{X}$, $A \in \mathcal{B}(\mathbf{X})$, denote the corresponding transition kernel, i.e.

$$P^t(x, A) \equiv P(X_t \in A | X_0 = x) \equiv P_x(X_t \in A), \tag{3}$$

where P_x is a shorthand-notation for the conditional probability $P(\cdot | X_0 = x)$. Note that $P^t(\cdot, \cdot)$ is a *Markov kernel*, i.e. for every $x \in \mathbf{X}$, the map $A \mapsto P^t(x, A)$ is a probability measure on $\mathcal{B}(\mathbf{X})$ and for every $A \in \mathcal{B}(\mathbf{X})$, the map $x \mapsto P^t(x, A)$ is a measurable function. Similarly, by a *kernel* we understand a function $K : (\mathbf{X}, \mathcal{B}(\mathbf{X})) \to \mathbb{R}_{\geq 0}$ such that $K(x, \cdot)$ is a measure, not necessarily normed by 1, for every x and $K(\cdot, A)$ is a measurable function for every measurable set A. Moreover, let us denote the corresponding semi-group by P_t, i.e.

$$P_t f(x) \equiv E(f(X_t) | X_0 = x) = \int_{\mathbf{X}} f(y) P^t(x, dy) \tag{4}$$

for $f : \mathbf{X} \to \mathbb{R}$ bounded measurable. For a measurable set A, we consider the stopping time τ_A and the number of visits of X in set A,

$$\tau_A \equiv \inf\{t \geq 0 | X_t \in A\}, \quad \eta_A \equiv \int_0^\infty \mathbf{1}_A(X_t) dt,$$

Definition 3.1 *Assume that there is a σ-finite, non-trivial measure φ on $\mathcal{B}(\mathbf{X})$ such that, for sets $B \in \mathcal{B}(\mathbf{X})$, $\varphi(B) > 0$ implies $E_x(\eta_B) > 0$, $\forall x \in \mathbf{X}$. Here, similar to P_x, E_x is a short-hand notation for the conditional expectation $E(\cdot | X_0 = x)$. Then X is called φ-irreducible.*

In the more familiar case of a finite state space and discrete time, we would simply require $\eta_{\{x\}}$ to have positive expectation for any state x. In the continuous case, such a requirement would obviously be far too strong, since singletons $\{x\}$ usually have probability zero. The above definition only requires positive expectation for sets B, which are "large enough", in the sense that they are non-null for some reference measure.

A simple sufficient condition for irreducibility is given in Meyn and Tweedie (1993b, prop. 2.1), which will be used to show irreducibility of the wealth-employment process.

Proposition 3.2 *Suppose that there exists a σ-finite measure μ such that $\mu(B) > 0$ implies that $P_x(\tau_B < \infty) > 0$. Then X is φ-irreducible, where*

$$\varphi(A) \equiv \int_{\mathbf{X}} R(x, A)\mu(dx), \quad R(x, A) \equiv \int_0^\infty P^t(x, A)e^{-t}dt.$$

Definition 3.3 *The process X is called* Harris recurrent *if there is a non-trivial σ-finite measure φ such that $\varphi(A) > 0$ implies that $P_x(\eta_A = \infty) = 1$, $\forall x \in \mathbf{X}$. Moreover, if a Harris recurrent process X has an invariant probability measure, then it is called positive Harris.*

Like in the discrete case, Harris recurrence may be equivalently defined by the existence of a σ-finite measure μ such that $\mu(A) > 0$ implies that $P_x(\tau_A < \infty) = 1$. As already remarked in the context of irreducibility, in the discrete framework one would consider sets $A = \{y\}$ with only one element.

Let μ be a measure on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$. We define a measure P_μ^t by

$$P_\mu^t(A) = \int_{\mathbf{X}} P^t(x, A)\mu(dx).$$

We say that μ is an *invariant measure*, iff $P_\mu^t = \mu$ for all t. Here, the measure μ might be infinite. If it is a finite measure, we may, without loss of generality, normalize it to have total mass $\mu(\mathbf{X}) = 1$. The resulting probability measure is obviously still invariant, and we call it an *invariant distribution*. (Note that any constant multiple of an invariant measure is again invariant.) In the case of an invariant distribution, we can interpret invariance as meaning that the Markov process has always the same marginal distribution in time, when starting with the distribution μ.

3.2   Existence of an invariant probability measure

The existence of finite invariant measures follows from a combination of two different types of conditions. The first property is a growth property. Several such properties have been used in the literature, a very useful one seems to be *boundedness in probability on average*.

Definition 3.4 *The process X is called* bounded in probability on average *if for every $x \in \mathbf{X}$ and every $\epsilon > 0$ there is a compact set $C \subset \mathbf{X}$ such that*

$$\liminf_{t \to \infty} \frac{1}{t} \int_0^t P_x(X_s \in C)ds \geq 1 - \epsilon. \tag{5}$$

The second property is a continuity condition.

Definition 3.5 *The Markov process X has the* weak Feller property *if for every continuous bounded function $f : \mathbf{X} \to \mathbb{R}$ the function $P_t f : \mathbf{X} \to \mathbb{R}$ from (4) is again continuous. Moreover, if $P_t f$ is continuous even for every bounded measurable function f, then X has the* strong Feller property.

Given these two conditions, Meyn and Tweedie (1993b, th. 3.1) establish the existence of an invariant probability measure in the following

Proposition 3.6 *If a Markov process X is bounded in probability on average and has the weak Feller property, then there is an invariant probability measure for X.*

3.3 Uniqueness

Turning to uniqueness, the following proposition is cited in Meyn and Tweedie (1993b, page 491). For a proof see Azéma, Duflo and Revuz (1969, Théorème 2.5).

Proposition 3.7 *If the Markov process X is Harris recurrent and irreducible for a non-trivial σ-finite measure φ, then there is a unique invariant measure (up to constant multiples).*

Proposition 3.7 gives existence and uniqueness of the invariant *measure*. A simple example shows that irreducibility and Harris recurrence do not guarantee existence of an invariant *probability* measure: Let $\mathbf{X} = \mathbb{R}$ and $X_t = B_t$ denote the one-dimensional Brownian motion. The Brownian motion is both irreducible and Harris recurrent – irreducibility is easily seen, while recurrence is classical in dimension one. Therefore, there is a unique invariant measure. By the Fokker-Planck equation, the density f of the invariant measure must satisfy $\Delta f = 0$. By non-negativity, this implies that f is constant, $f \equiv c$ for some $c > 0$. Thus, any invariant measure is a constant multiple of the Lebesgue measure, and there is no invariant probability measure for this example.

Given this example and as we are only interested in invariant *probability* measures, we need to combine this proposition with the previous section: Boundedness in probability on average together with the weak Feller property gives us the existence of an invariant probability measure as used in sect. 3.2, whereas irreducibility together with Harris recurrence imply uniqueness of invariant measures. Thus, for existence and uniqueness of the invariant probability measure, we will need all four conditions.

Whereas irreducibility, boundedness in probability on average and the weak Feller property are rather straightforward to check in practical situations, this seems to be harder for Harris recurrence. Thus, we next discuss some sufficient conditions for Harris recurrence. If the Markov process has the strong Feller property, then Harris recurrence will follow from a very weak growth property, namely that $P_x(X_t \to \infty) = 0$ for all $x \in \mathbf{X}$, see Meyn and Tweedie (1993b, th. 3.2). While the strong Feller property is often satisfied for models driven by Brownian motion (e.g., for hypo-elliptic diffusions), it may not be satisfied in models where randomness is driven by a pure-jump process. Thus, we will next formulate an intermediate notion between the weak and strong Feller properties, which still guarantees enough smoothing for stability.

Definition 3.8 *The Markov process X is called* T-process*, if there is a probability measure ν on $\mathbb{R}_{\geq 0}$ and a kernel T on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$ satisfying the following three conditions:*

1. *For every $A \in \mathcal{B}(\mathbf{X})$, the function $x \mapsto T(x, A)$ is continuous[9].*

2. *For every $x \in \mathbf{X}$ and every $A \in \mathcal{B}(\mathbf{X})$ we have $K_\nu(x, A) \equiv \int_0^\infty P^t(x, A)\nu(dt) \geq T(x, A)$.*

3. *$T(x, \mathbf{X}) > 0$ for every $x \in \mathbf{X}$.*

The kernel K_ν is the transition kernel of a discrete-time Markov process $(Y_n)_{n \in \mathbb{N}}$ obtained from $(X_t)_{t \geq 0}$ by random sampling according to the distribution ν: more precisely, let us draw a sequence σ_n of independent samples from the distribution ν and define a discrete time process $Y_n \equiv X_{\sigma_1 + \cdots + \sigma_n}$, $n \in \mathbb{N}$. Then the process Y_n is Markov and has transition probabilities given by K_ν. Using this definition and theorem 3.2 in Meyn and Tweedie (1993b), we can formulate

Proposition 3.9 *Suppose that X is a φ-irreducible T-process. Then it is Harris recurrent (with respect to φ) if and only if $P_x(X_t \to \infty) = 0$ for every $x \in \mathbf{X}$.*

Hence, in a practical sense and in order to prove existence of a unique invariant *probability* measure, one needs to establish that a process X has the weak Feller property and is an irreducible T-process which is bounded in probability on average (as the latter implies the growth condition $P_x(X_t \to \infty) = 0$ of prop. 3.9).

Let us shortly compare the continuous, but compact case – where boundedness in probability is always satisfied – with the discrete case. In the latter situation, existence of an invariant distribution always holds, while uniqueness is then given by irreducibility. In the compact, continuous case irreducibility and Harris recurrence only guarantee existence and uniqueness of an invariant measure, which might be infinite. On the other hand, existence of a finite invariant measure is given by the weak Feller property. Thus, for existence and uniqueness of an invariant probability measure, we will need the weak Feller property, irreducibility and Harris recurrence – which we will conclude from the T-property. Thus, the situation in the continuous (but compact) case is roughly the same as in the discrete case, except for some required continuity property, namely the weak Feller property.

3.4   Stability

By now we have established a framework for showing existence and uniqueness of an invariant distribution, i.e., probability measure. However, under *stability* we understand more, namely the convergence of the marginal distributions to the invariant distribution, i.e., that for any starting distribution μ, the law P_μ^τ of the Markov process at time τ converges to the unique invariant distribution for $\tau \to \infty$. In the context of T-processes, we are going to discuss two methods which allow to derive stability. But first, let us define the notion of stability in a more precise way.

---

[9]A more general definition requires lower semi-continuity only. As we can show continuity for our applications, we do not need this more general version here.

Definition 3.10 *For a signed measure μ consider the total variation norm*

$$\|\mu\| \equiv \sup_{|f| \leq 1} \left| \int_{\mathbf{X}} f(x)\mu(dx) \right|.$$

Then we call a Markov process $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ stable *or* ergodic *iff there is an invariant probability measure π such that*

$$\forall x \in \mathbf{X}: \quad \lim_{t \to \infty} \|P^t(x, \cdot) - \pi\| = 0.$$

Note that this implies in particular that the law P^t_μ of the Markov process converges to π, which is the unique invariant probability measure.

In the case of a finite state space in discrete time, ergodicity follows (inter alia) from aperiodicity. Down, Meyn and Tweedie (1995), also give one continuous result in this direction.

Definition 3.11 *A ψ-irreducible Markov process (X_t) is called* aperiodic *iff there is a measurable set C with $\psi(C) > 0$ satisfying the following properties:*

1. there is $\tau > 0$ and a non-trivial measure ν on $\mathcal{B}(\mathbf{X})$ such that

$$\forall x \in C, \ \forall A \in \mathcal{B}(\mathbf{X}): \quad P^\tau(x, A) \geq \nu(A);\,^{10}$$

2. there is $T > 0$ such that

$$\forall t \geq T, \ \forall x \in C: \quad P^t(x, C) > 0.$$

If we are given an irreducible, aperiodic Markov process, then stability is implied by conditions on the *infinitesimal generator*. In the following proposition we give a special case of Down, Meyn and Tweedie (1995, th. 5.2) suitable for the employment-wealth process in our model.

Proposition 3.12 *Given an irreducible, aperiodic T-process X_t with infinitesimal generator $\mathcal{A}$ on a compact state space. Assume we can find a measurable function $V \in \mathcal{D}(\mathcal{A})$ with $V \geq 1$ and constants $d, c > 0$ such that*

$$\mathcal{A}V \leq -cV + d.$$

Then the Markov-process is ergodic.

The problem with aperiodicity in the continuous time framework is that it seems hard to characterize the small sets appearing in def. 3.11. For this reason, we also give an alternative theorem, which avoids small sets (but is clearly related with the notion of aperiodicity). Given a fixed $\tau > 0$, the process $Y_n \equiv X_{\tau n}$, $n \in \mathbb{N}$, clearly defines a Markov process in discrete time, a so-called *skeleton* of X. These skeleton chains are a very useful construction for transferring results from Markov processes in discrete time to continuous time. In particular, Meyn and Tweedie (1993b, th. 6.1) gives a characterization of stability in terms of irreducibility of skeleton chains.

Proposition 3.13 *Given a Harris recurrent Markov process X with invariant probability measure π. Then X is stable iff there is some irreducible skeleton chain.*

---

[10]Such a set C is then called *small*.

4 Stability of the wealth-employment process

Turning to our main application, we would like to understand the stability properties of the model developed by Bayer and Wälde (2011). Understanding stability means understanding the dynamic properties of the state variables of the model. All other variables (like control variables or e.g. factor rewards) are known deterministic functions of the state variables. Hence, if we understand the process governing the state variables, we also understand the properties of all other variables in this model.

As the fundamental state variables in Bayer and Wälde (2011) are wealth and the employment status of an individual, the process we are interested in is the wealth-employment process $X_\tau \equiv (a(\tau), z(\tau))$. The state-space of this process is $\mathbf{X} \equiv [-b/r, \ a_w^*] \times \{w, b\}$ and has all the properties required for the state space in sect. 3. We now follow the structure of sect. 3.

4.1 Existence

In order to show existence for an invariant probability measure for X, prop. 3.6 tells us that we need to prove that X is (i) bounded in probability on average and (ii) has the weak Feller property. Showing that X is bounded in probability on average is straightforward: According to def. 3.4 we need to find a compact set for any initial condition x and any small number ϵ such that the average probability to be in this set is larger than $1 - \epsilon$. As our process $X_\tau \equiv (a(\tau), z(\tau))$ is bounded, we can choose the state-space $\mathbf{X} \equiv [-b/r, \ a_w^*] \times \{w, b\}$ as our set for any x and ϵ. Concerning the weak Feller property, we offer the following

Lemma 4.1 *The wealth-employment process has the weak Feller property.*

Proof. Let us first show that the wealth-employment process depends continuously on its initial values. To see this, fix some $\omega \in \Omega$, the probability space, on which the wealth-employment process is defined. Notice that $z_\tau(\omega)$ is certainly continuous in the starting values, because any function defined on $\{w, b\}$ is continuous by our choice of topology. Thus, we only need to consider the wealth process. For fixed ω, $a_\tau(\omega)$ is a composition of solutions to deterministic ODEs, each of which are continuous functions of the respective initial value. Therefore, $a_\tau(\omega)$ is a continuous function of the initial wealth.

Now assume, without loss of generality, that the wealth-employment process has a deterministic initial value (a_0, z_0) and fix some bounded, continuous function $f : [-b/r, \ a_w^*] \times \{w, b\} \to \mathbb{R}$. For the weak Feller property, we need to show that

$$P_\tau f(a_0, z_0) = E\left(f(a_\tau, z_\tau)\right)$$

is a continuous function in (a_0, z_0). Thus, take any sequence (a_0^n, z_0^n) converging to (a_0, z_0) and denote the wealth-employment process started at (a_0^n, z_0^n) by (a_τ^n, z_τ^n). Then, by continuous dependence on the initial value, $(a_\tau^n(\omega), z_\tau^n(\omega)) \to (a_\tau(\omega), z_\tau(\omega))$, for every $\omega \in \Omega$. By continuity of f, this implies convergence of $f(a_\tau^n(\omega), z_\tau^n(\omega))$. Since f is bounded, we may conclude convergence $P_\tau f(a_0^n, z_0^n) \to P_\tau f(a_0, z_0)$ by

the dominated convergence theorem. So, $P_\tau f$ is, indeed, bounded and continuous whenever f is bounded and continuous, and the weak Feller property holds. ∎

4.2 Uniqueness

Given the result of existence, we can use the results of sect. 3.3 to establish uniqueness of the probability measure if there is a unique invariant measure for X. Following prop. 3.7, the latter is established by proving Harris recurrence and irreducibility.

We establish irreducibility in the following

Lemma 4.2 *In the low-interest-regime with* $r < \rho$, $(a(\tau), z(\tau))$ *is an irreducible Markov process, with the non-trivial irreducibility measure* φ *introduced in prop. 3.2.*

Proof. Let $-b/r < a < a_w^*$, $z \in \{w, b\}$. Then, regardless of the initial point $a_t \in [-b/r, a_w^*]$ and regardless of z_t, it is possible to attain the state (a, z) in finite time with probability greater than zero. Thus, prop. 3.2 implies irreducibility, where we can take the Lebesgue measure on $[-b/r, a_w^*]$ times the counting measure on $\{w, b\}$ as measure μ. ∎

We now go for the more involved proof for Harris recurrence. In order to do this, we employ prop. 3.9 and need to prove that X is a T-process. Before we do so, consider the following useful auxiliary lemma.

Lemma 4.3 *The conditional density of the time of the first jump in employment given that there is precisely one such jump in* $[0, \tau]$ *and that* $z(0) = w$ *is given by*

$$g_\tau^{(1)}(u) = \begin{cases} \frac{\mu - s}{e^{(\mu-s)\tau} - 1} e^{(\mu-s)u}, \ 0 \le u \le \tau, & \mu \ne s, \\ 1/\tau, \ 0 \le u \le \tau, & \mu = s. \end{cases}$$

Proof. Since the formula is well-known for $\mu = s$, we only prove the result for $\mu \ne s$. The joint probability of the first jump $\tau_1 \le u \le \tau$ and $N_\tau = 1$, where N_τ denotes the number of jumps in $[0, \tau]$, is given by

$$P(\tau_1 \le u, \ N_\tau = 1) = P(\tau_1 \le u, \ \tau_2 \ge \tau - \tau_1) = \int_0^u P(\tau_2 \ge \tau - v) s e^{-sv} dv$$

$$= \int_0^u e^{-\mu(\tau - v)} s e^{-sv} dv = \frac{s}{\mu - s} e^{-\mu\tau} \left(e^{(\mu-s)u} - 1 \right).$$

Here, τ_2 denotes the time between the first and the second jump, and we have used independence of τ_1 and τ_2. Dividing through the probability of $N_\tau = 1$, we get

$$P(\tau_1 \le u | N_\tau = 1) = \frac{e^{(\mu-s)u} - 1}{e^{(\mu-s)t} - 1},$$

and we obtain the above density by differentiating with respect to u. ∎

Now we formulate the central technical lemma.

Lemma 4.4 *Assume that the optimal consumption function* $c(a, z)$ *is* C^1 *in wealth* a. *Then the employment-wealth-process* $(a(\tau), z(\tau))$ *is a T-process.*

12

Proof. Given $(a_t, z_t) \in \mathbf{X}$ and $A \in \mathcal{B}(\mathbf{X})$, the probability of $(a(\tau), z(\tau)) \in A$ for some $\tau > t$ can naturally be decomposed as

$$
\begin{aligned}
&P\left((a(\tau), z(\tau)) \in A | a(t) = a_t, \ z(t) = z_t\right) \\
&= \Sigma_{k=0}^{\infty} P((a(\tau), z(\tau)) \in A | a(t) = a_t, \ z(t) = z_t, \ N(\tau) = k) \cdot P(N(\tau) = k) \\
&\equiv \Sigma_{k=0}^{\infty} f_\tau^{(k)}(a_t, z_t; A) P(N(\tau) = k),
\end{aligned}
\tag{6}
$$

where obviously

$$
f_\tau^{(k)}(a_t, z_t; A) = P((a(\tau), z(\tau)) \in A | a(t) = a_t, \ z(t) = z_t, \ N(\tau) = k)
$$

is the probability of $(a(\tau), z(\tau))$ being in A conditional on initial condition and exactly k jumps and $N(\tau)$ denotes the number of jumps of the employment state between t and τ. The idea of the argument is that $f_\tau^{(0)}$ – corresponding to the event that no jump has occurred – has no smoothing property whatsoever, but all the other $f_\tau^{(k)}$, $k > 0$, do have smoothing effects, because of the random time of the jump. Therefore, we are going to prove that $f_\tau^{(1)}(a_t, z_t; A) P(N(\tau) = 1)$ satisfies the conditions of def. 3.8 on the kernel T using the Dirac measure $\nu = \delta_\tau$ which implies from def. 3.8.2. that $K_\nu((a, z), A) \equiv \int_0^\infty P^t((a, z), A) \delta_\tau(dt) = P^\tau((a, z), A)$.

For ease of notation, let us assume that $z_t = w$ and that $t = 0$ – of course, it is easy to extend the argument to the general case. Let $g_\tau^{(1)}$ denote the density of the time of the first jump in employment conditional on the assumption that there is precisely one such jump in $[0, \tau]$. This density is given by

$$
g_\tau^{(1)}(u) = \frac{\mu - s}{e^{(\mu-s)\tau} - 1} e^{(\mu-s)u}, \quad 0 \leq u \leq \tau,
$$

if $\mu \neq s$ and $g_\tau^{(1)}(u) \equiv 1/\tau$ otherwise, see lemma 4.3 above.

Given that the only jump in employment between 0 and τ happens at time $0 \leq u \leq \tau$, let us denote the mapping between the starting value a_0 and the value of the wealth at time τ by $\phi_1(a_0, u; \tau)$, i.e. ϕ_1 is the solution map of the stochastic differential equation for the wealth, given that only one jump of z happens in the considered time-interval. By adding the time of that single jump as a variable, ϕ_1 is a deterministic function. We can understand it as the flow generated by the underlying ODE. Indeed, let ψ_z denote the solution map of the ODE

$$
\frac{da(\tau)}{d\tau} = ra(\tau) + z - c(a(\tau), z)
\tag{7}
$$

for $z \in \{w, b\}$ fixed. More precisely, $\psi_z(a, u)$ denotes the solution of the ODE (7) evaluated at time u given that the starting value (at time 0) is a. Then we may obviously write

$$
\phi_1(a_0, u; \tau) = \psi_b(\psi_w(a_0, u), \tau - u).
\tag{8}
$$

Using the function ϕ_1 and the fact that $z(\tau) = b$ (given that $z_t = w$ and that there

13

is precisely one jump in $[0, \tau]$), we may re-write $f_\tau^{(1)}$ as

$$f_\tau^{(1)}(a_0, w; A) = \int_0^\tau \mathbf{1}_A(\phi_1(a_0, u; \tau), b) g_\tau^{(1)}(u) du \tag{9}$$

$$= \int_{low}^{up} \mathbf{1}_A(y, b) g_\tau^{(1)}(\phi_1^{-1}(a_0, y; \tau)) \left| \frac{\partial}{\partial y} \phi_1^{-1}(a_0, y; \tau) \right| dy. \tag{10}$$

Before we proceed, note that in equation (10) we have made the substitution

$$y \equiv \phi_1(a_0, u; \tau),$$

understood as a change from the u-variable to the y-variable. Consequently, $y \mapsto \phi_1^{-1}(a_0, y; \tau)$ is to be understood as the inverse map of $u \mapsto \phi_1(a_0, u; \tau)$, with a_0 and τ being held fix. Therefore, we first need to check the validity of that substitution. Note that $u \mapsto \phi_1(a_0, u; \tau)$ is a strictly increasing map by our assumption, see Bayer and Wälde (2011, lem. 3 and 4). Thus, the substitution $y = \phi_1(a_0, u; \tau)$ is well-defined. Moreover, this implies that the lower and upper boundaries of the integration in (10) are given by

$$low \equiv \phi_1(a_0, 0; \tau), \quad up \equiv \phi_1(a_0, \tau; \tau),$$

respectively. In order to justify the application of the substitution rule, we need to establish continuous differentiability of ϕ_1^{-1} in y. Formally, we immediately have that

$$\frac{\partial}{\partial y} \phi_1^{-1}(a_0, y; \tau) = \frac{1}{\frac{\partial}{\partial u} \phi_1(a_0, u; \tau) \big|_{u = \phi_1^{-1}(a_0, y; \tau)}}.$$

By (8), we get

$$\frac{\partial}{\partial u} \phi_1(a_0, u; \tau) = \frac{\partial}{\partial u} \psi_b(\psi_w(a_0, u), \tau - u)$$

$$= -\frac{\partial \psi_b}{\partial u}(\psi_w(a_0, u), \tau - u) + \frac{\partial \psi_b}{\partial a}(\psi_w(a_0, u), \tau - u) \frac{\partial \psi_w}{\partial u}(a_0, u)$$

$$= -\underbrace{\left[r \psi_b(\psi_w(a_0, u), \tau - u) + b - c(b, \psi_b(\psi_w(a_0, u), \tau - u)) \right]}_{<0} +$$

$$+ \underbrace{\frac{\partial \psi_b}{\partial a_0}(\psi_w(a_0, u), \tau - u)}_{\geq 0} \underbrace{\left[r \psi_w(a_0, u) + w - c(w, \psi_w(a_0, u)) \right]}_{\geq 0} > 0.$$

The inequality follows, because wealth of the unemployed is strictly decreasing (see Bayer and Wälde 2011), the final wealth is an increasing function of the initial wealth, and the wealth of the employed is increasing. Moreover, we can easily see that all the terms are continuous in u. Thus, we can justify the substitution (10).

Having justified the substitution, we can return to (10) and establish continuity of $f_\tau^{(1)}(a_0, w; A)$ in a_0 since the integrand as well as the integration-boundaries in (10) are continuous in a_0. Therefore, $T(a_0, z_t; A) \equiv f_\tau^{(1)}(a_0, z_t; A) P(N(\tau) = 1)$ satisfies the

14

first condition in def. 3.8 – continuity in the z-variable is trivial. The second and the third conditions are obvious, and we have finished the proof of the lemma. ∎

We can now complete our proof of uniqueness by proving the remaining conditions of prop. 3.9.

Theorem 4.5 *Suppose that $r < \rho$ and that $c(a, z)$ is continuously differentiable in a. Then there is a unique invariant probability measure for the wealth-employment process $(a(\tau), z(\tau))$.*

Proof. By lemma 4.2 and lemma 4.4, the employment-wealth process $(a(\tau), z(\tau))$ is an irreducible T-process. Thus, prop. 3.9 implies that $(a(\tau), z(\tau))$ is Harris recurrent, given that $P_x(X_t \to \infty) = 0$ holds for our bounded state space. By prop. 3.7, there is a unique invariant measure (up to a constant multiplier), and prop. 3.6, finally, implies that we may choose the invariant measure to be a probability measure. ∎

4.3   Stability

Using the skeleton approach of prop. 3.13, we now show stability.

Corollary 4.6 *Under the assumptions of theorem 4.5, the employment-wealth process is stable in the sense of def. 3.10.*

Proof. Recall that the employment-wealth-process is a T-process, see lemma 4.4. Moreover, we have shown irreducibility in lemma 4.2. Proposition 3.13 will imply the desired conclusion, if we can show irreducibility of a skeleton chain. Take any $\tau > 0$ and consider the corresponding skeleton Y_n, $n \in \mathbb{N}$, with transition probabilities P^τ. By the proof of lemma 4.4, we see that (Y_n) is also a T-process, where the definition of T-processes is generalized to discrete-time processes in the obvious way. By Meyn and Tweedie (1993, prop. 6.2.1), the discrete-time T-process Y is irreducible if there is a point $x \in \mathbf{X}$ such that for any open neighborhood O of x, we have

$$\forall y \in \mathbf{X} : \quad \Sigma_{n=1}^\infty P^{n\tau}(y, O) > 0. \tag{11}$$

This property, however, can be easily shown for the wealth-employment process (a, z). Indeed, take $x = (-b/r, b)$. Then any open neighborhood O of x contains $[-b/r, -b/r + \epsilon[\times \{b\}$ for some $\epsilon > 0$. We start at some point $y = (a_0, z_0) \in \mathbf{X}$ and assume the following scenario: if necessary, at some time between 0 and τ, the employment status changes to b, then it stays constant until the random time $N\tau$ defined by $N \equiv \inf\{n \mid a(n\tau) < -b/r + \epsilon\}$. Note that the wealth is decreasing in a deterministic way while $z = b$. Thus, we can find a deterministic upper bound $N \leq K(a_0)$. The event that the employment attains the value b during the time interval $[0, \tau]$ and retains this value until time $K(a_0)\tau$ has positive probability. In this case, however, the trajectory of the wealth-employment process reaches O, implying that $\Sigma_{n=1}^\infty P^{n\tau}(y, O) > 0$. Thus, the τ-skeleton chain is irreducible and the wealth-employment process is stable. ∎

15

5 Application to other models

Let us now apply the conditions obtained above for the existence, uniqueness and stability of an invariant distribution to other models used in the literature. The models are chosen as they are either widely used or they have very useful properties. We start with a geometric Brownian motion process which is extremely popular in the economics literature. It is not, however, characterized by an invariant unique and stable distribution. It nevertheless does have, however, some other features which makes it very attractive. The subsequent section then presents a version of the Ornstein-Uhlenbeck process. This version also implies a lognormal distribution but it has all the properties one would wish for. While there are some recent applications of the process, it is not as widely used as it should. The final example studies properties of a model driven by Poisson uncertainty. The structure of this model is somewhat simpler than the one studied above. It is therefore especially useful for illustrating the ideas of the above proofs. The concluding subsection points out why proofs of stable invariant and unique distributions might fail for other modeling setups.

5.1 Geometric models

5.1.1 A geometric Brownian motion model

One of the most popular fundamental structures is the geometric Brownian motion model. Using capital $K(t)$ as an example, its structure is

$$dK(t) = K(t)\left[gdt + \sigma dB(t)\right], \tag{12}$$

where $B(t)$ is standardized Brownian motion, i.e. $B(t) \sim N(0,t)$, σ is the constant variance parameter and g is a constant which collects various parameters depending on the specific model under consideration. This reduced form can be derived from or appears in the models in Evans and Kenc (2003), Gabaix (1999), Gertler and Grinols (1982), Gokan (2002), Gong and Zou (2002, 2003), Grinols and Turnovsky (1998a, 1998b), Obstfeld (1994), Prat (2007), Steger (2005) or Turnovsky (1993) and many others. For more background and a critique, see Wälde (2011).

It is well-known that the capital stock as described by (12) for any $\tau > t$ given a fixed initial condition of $K(t) = K_0$ is log-normally distributed (this would also hold for a random (log-normal) initial condition): As the solution to (12) is $K(\tau) = K_0 e^{\left(g-\frac{1}{2}\sigma^2\right)(\tau-t)+\sigma z(\tau)}$ (see e.g. Wälde, 2010, ch. 10.4.1) and $B(t)$ is normally distributed, $K(t)$ is lognormally distributed,

$$K(\tau) \sim \log N\left(\mu_\tau, \sigma_\tau^2\right) \equiv f(K,\tau) = \frac{1}{K\sqrt{2\sigma_\tau^2\pi}}e^{-\frac{(\log K - \mu_\tau)^2}{2\sigma_\tau^2}} \tag{13}$$

where the mean and variance of $K(\tau)$ are (see app. A.1)

$$\mu_\tau \equiv \mu(\tau, K_0) \equiv E_t K(\tau) = K_0 e^{g[\tau-t]}, \tag{14}$$

$$\sigma_\tau^2 \equiv \sigma^2(\tau, K_0) \equiv \mathrm{var}_t K(\tau) = K_0^2 e^{2g[\tau-t]}\left(e^{\sigma^2[\tau-t]} - 1\right). \tag{15}$$

16

The variance σ_τ^2 increases for all $\tau \geq t$ if $g \geq -\sigma^2/2$. For $g < -\sigma^2/2$, the variance first increases and then decreases after some finite point τ^* in time defined by $(2g + \sigma^2) e^{\sigma^2[\tau^*-t]} = 2g$. As the variance approaches zero, the case of $g < -\sigma^2/2$ implies a degenerate long-run distribution.

For the empirically more relevant case of $g \geq -\sigma^2/2$, we see from the expressions for μ_τ and σ_τ^2 that one can not find parameter values such that the mean and the variance are both constant over time, not even asymptotically. It immediately follows that conditions for tightness or boundedness in probability on average are violated. Hence, an *invariant* probability distribution does not exist. Obviously, trying to show uniqueness of an *invariant* probability distribution is then also futile. By contrast, (13) tells us that a unique (but not invariant) probability distribution exists. Stability could be analyzed for a random initial condition which is not log-normal. We would find that the distribution converges to the unique (not invariant) lognormal distribution.

5.1.2   Other geometric models

There are various extensions of the geometric Brownian motion process. One popular extensions consists in allowing for jumps. Unfortunately, this does not affect the empirically implausible property of a variance approaching infinity in the long-run. When we consider the version of Lo (1988), Wachter (2009), Posch (2009, 2010) or many others,

$$dK(t) = K(t) [gdt + \sigma dB(t) + \gamma dq(t)], \tag{16}$$

we obtain (cf. app. A.2) a mean of $E_t K(\tau) = K(t) e^{(g+\gamma\lambda)(\tau-t)}$ and a variance of $\text{var}_t K(\tau) = K(t)^2 \left[e^{\omega[\tau-t]} - e^{2(g+\gamma\lambda)(\tau-t)} \right]$, where $\omega \equiv 2[g + \gamma\lambda] + \sigma^2 + \gamma^2\lambda$. Obviously, by assuming $g + \gamma\lambda = 0$, i.e. a negative jump size γ with a positive trend or a negative trend g with positive jump size, we can obtain a constant mean. But the variance would even in this case increase over time. Clearly, an increasing variance per se does not preclude stability. Indeed, the limiting distributions for models with jumps characterized by e.g. Cauchy processes or, more generally, so-called stable processes are characterized by infinite variance. Economically speaking, however, an infinite variance appears unsatisfactory if not from an empirical then at least from a modelling perspective.

Merton's (1975) analysis of the Solow growth model with constant saving rate and Cobb-Douglas production function also starts from a geometric process with a structure identical to (12), specifying population growth. He obtains an invariant gamma-distribution for the capital to labour ratio (and for other endogenous quantities) on an infinite support. Starting with a specification of a non-stationary distribution therefore sometimes still leads to stationary distributions for endogenous variables.

5.2 The Ornstein Uhlenbeck process

5.2.1 The basic structure and some properties

The version of the Ornstein-Uhlenbeck (OU) process we employ here is described by

$$dx(t) = \beta[\mu - x(t)]dt + \sigma dB(t), \tag{17}$$

where μ and $\sigma > 0$ are constants, $\beta > 0$ can be called the speed-of-convergence parameter and $B(t)$ is standardized Brownian motion as before. One of the most convenient properties is that $x(t) \sim N\left(\mu + (x_0 - \mu)e^{-\beta[\tau-t]}, \frac{\sigma^2}{2\beta}\left[1 - e^{-2\beta[\tau-t]}\right]\right)$ for a non-random or normal initial value x_0 (Karlin and Taylor, 1981, 1998). It has been used in economics recently e.g. by Shimer (2005).[11] For an application in finance, see e.g. Menzly et al. (2004). Its solution for an initial condition $x(t) = x_0$ is

$$x(\tau) = \mu + (x_0 - \mu)e^{-\beta[\tau-t]} + \int_t^\tau \sigma e^{-\beta[\tau-s]}dB(s).$$

This process is most useful for modelling positive variables like the capital stock or TFP if an exponential transformation $e^{x(\tau)}$ is considered. This transformation also allows us to see most clearly the advantages of this process as compared to a geometric Brownian motion as in (12). Let us imagine we have a model where prices, TFP or capital follow a process as described by

$$K(\tau) \equiv e^{x(\tau)}. \tag{18}$$

As $x(t)$ fluctuates around μ, $e^{x(t)}$ fluctuates around $K^* \equiv e^\mu$ with lower bound at zero and upper bound infinity. Now define $K_0 \equiv e^{x_0}$ and write

$$K(\tau) = e^\mu e^{(x_0-\mu)e^{-\beta[\tau-t]}} e^{\int_t^\tau \sigma e^{-\beta[\tau-t]}dB(\tau)} = K^* \left[\frac{K_0}{K^*}\right]^{e^{-\beta[\tau-t]}} e^{\int_t^\tau \sigma e^{-\beta[\tau-s]}dB(s)}. \tag{19}$$

The mean and variance of this expression are, assuming a deterministic initial condition K_0 (cf. app. A.4),

$$E_t K(\tau) = K^* \left[\frac{K_0}{K^*}\right]^{e^{-\beta[\tau-t]}} \frac{1}{2}\frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta[\tau-t]}\right), \tag{20}$$

$$\mathrm{var}_t K(\tau) = \left(K^* \left[\frac{K_0}{K^*}\right]^{e^{-\beta[\tau-t]}}\right)^2 e^{\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta[\tau-t]}\right)} \left(e^{\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta[\tau-t]}\right)} - 1\right). \tag{21}$$

An immediate implication of the construction of $K(\tau)$ in (18) is that $K(\tau)$ is log-normally distributed with mean and variance given by (20) and (21). The transition density of $K(\tau)$ given $K_t = K_0$ is therefore given by $f(K, \tau; K_0, t)$ as in (13) with moments given by (20) and (21). In contrast to the moments in (14) and (15) for the geometric process, however, (20) and (21) both converge to constants. There is therefore an invariant lognormal distribution.

---

[11]A process inspired by the OU-process was used in Shimer (2007). A geometric-reverting process was used by Posch (2009).

5.2.2 Conditions for existence, uniqueness and stability

After these preliminary steps, we will now illustrate the proofs of section 3.2 and apply them to the exponential OU process. While sections 3.2 and 4 have already shown the line of thought, we briefly recapitulate what needs to be proven. The following can also be seen as a "check-list" of conditions to be checked by authors of other models.

Consider a Markov process X. We need to investigate whether it has the following properties.

1) X is bounded in probability on average
2) X has the weak Feller property
3) X is a φ-irreducible T-process
4) There is a skeleton chain of process X which is irreducible.

Property 1 and 2 (by prop. 3.6) imply the existence of an invariant *probability* measure. Properties 3 and the growth condition $P_x(X_t \rightarrow \infty) = 0$ for every $x \in \mathbb{R}$ imply the existence of a unique invariant measure (by prop. 3.7 via prop. 3.9). As this growth condition is implied by the stronger property 1, we will establish the existence of a unique invariant measure by checking properties 1 and 3. Properties 1 *to* 3 imply the existence of a unique invariant *probability* measure. Property 4 implies stability of an invariant distribution by prop. 3.13.

Existence - The Markov process K of this application is described by (18) with (17). It is tight as discussed after def. 3.4: Fix an initial condition K_0 and an $\epsilon > 0$. A set C for which $P_{K_0}(K(\tau_1) \in C) = 1 - \epsilon$ for some τ_1 can trivially be found. If the variance and mean, and thereby the entire distribution, stay constant, the same set is characterized by $P_{K_0}(K(t) \in C) \geq 1 - \epsilon$ for all t. As X is tight, it is also 'bounded in probability on average'.

The process K also has the weak Feller property: $K(t)$ is an Ito diffusion (Øksendal, 1998, 5th edition, ch. 7.1) and "any Ito diffusion is Feller-continuous" (Øksendal, 1998, 5th edition, lem. 8.1.4). As Feller-continuous is a different name for weak Feller property, we conclude that a long-run invariant probability measure exists.

Uniqueness - We now argue that K is a φ-irreducible T-process. We can follow the arguments in the proof of lemma 4.2 to first show φ-irreducibility: For any initial condition K_0, there is a positive probability to reach any subset of the state space in finite time, essentially by irreducibility of the underlying Brownian motion. Using as measure μ the Lebesgue measure, for instance, we have established irreducibility of K.[12]

To show that K is a T-process, def. 3.8 tells us that we need to find a function $T(K, A)$ which satisfies the three conditions of this definition. The function should be positive, continuous and smaller than $K_\nu(K_0, A) \equiv \int_0^\infty P^\tau(K_0, A)\nu(d\tau)$ for any initial condition K_0, any set A and a suitable probability measure ν. The latter

---

[12]In contrast to our original setup of section 2, the OU-process x here is uniformly elliptic. This immediately implies irreducibility for x and also K. Indeed, by the results of Kusuoka and Stroock (1987), we see that the transition density of x is bounded below (and above) by a Gaussian transition density. Thus, it is positive, implying irreducibility. Of course, in the case of an OU-process, we can directly write down the transition density and verify its positivity.

measure is completely unrelated to the probability measure implicit in $P^{\tau}(K_0, A)$.

Let us first compute $P^{\tau}(K_0, A)$ for (18). Using the lognormal density of (13), we get

$$P^{\tau}(K_0, A) = \int_A f(K, \tau) \, dK. \tag{22}$$

Integrating yields $\int_0^{\infty} P^{\tau}(K_0, A) \nu(d\tau) = \int_0^{\infty} \int_A f(K, \tau) \, dK \nu(d\tau)$. Now imagine that v is a smooth probability measure with support $[0, \infty[$. We can then write this expression as $\int_0^{\infty} \int_A f(K, \tau) \, dK \nu(\tau) \, d\tau$. As the density $f(K, \tau)$ defined in (13), using the mean and variance from (20) and (21), is continuous in K_0 and integrating in (22) preserves this continuity and so does integration with a smooth probability measure $v(\tau)$, the function $K_{\nu}(K_0, A)$ itself is continuous in K_0. Taking it as our function $T\left(\hat{K}, A\right)$, we found again a T-process. This establishes the existence of a unique invariant probability measure.

It is worth noting at this point that the application of T-processes for processes driven only by Brownian motion (i.e. there are no jumps) is somewhat "unusual". Our process K actually has the strong Feller property, which is much stronger than the T-property. Indeed, this is a common feature of all so-called hypo-elliptic diffusion processes, see, for instance, Meyn and Tweedie (1993b). Hypo-ellipticity is a generalization of ellipticity, which is satisfied here.

Stability - The skeleton chain is constructed by fixing a $\tau > 0$ and considering the process $Y_n \equiv K(\tau n)$ with $n \in \mathbb{N}$. As we have noted above, the transition densities of K are strictly positive. Thus, the transition density of Y is positive, which immediately implies irreducibility.

Consider the following example. We start in $y = K_0$ and consider a neighborhood $O = \left[\bar{K}, \bar{K} + \varepsilon\right]$ with some $\varepsilon > 0$. Then $P^{n\tau}\left(K_0, \left[\bar{K}, \bar{K} + \varepsilon\right]\right) = \int_{\bar{K}}^{\bar{K}+\varepsilon} f(K, n\tau)$. Given the density defined in (13) with (20) and (21), this probability is strictly positive. As this works for any $\bar{K} > 0$, we have shown stability.

5.3 A model of natural volatility

The following model is characterized by jumps as well. The strong Feller property therefore does not hold and the application of the tools of section 3 and 4 is needed.

5.3.1 Basic structure and some properties

The natural volatility literature stresses that growth and volatility are two sides of the same coin. Determinants of long-run growth are also determinants of short-run volatility. Both short-run volatility and long-run growth are endogenous. This point has been made in the papers of Bental and Peled (1996), Matsuyama (1999), Wälde (2002, 2005), Francois and Lloyd-Ellis (2003,2008), Posch and Wälde (2009) and others. The (slightly simplified) reduced form of the models by Wälde (2005) and Posch and Wälde (2009) reads

$$d\hat{K}(\tau) = \left(\hat{K}(\tau)^{\alpha} - \beta_1 \hat{K}(\tau)\right) dt - \beta_2 \hat{K}(\tau) \, dq, \tag{23}$$

where $\hat{K}$ is the (stochastically detrended) capital stock, q is a Poisson process with some constant (but endogenous) arrival rate and $0 < \alpha, \beta_i < 1$ are constant parameters.

This equation immediately reveals that in times without jumps ($dq = 0$) and an initial capital stock of $\hat{K}_0 < \hat{K}^* \equiv \beta_1^{-1/(1-\alpha)}$, the capital stock monotonically increases over time to asymptotically approach $\hat{K}^*$. When there is a jump, the (stochastically detrended) capital stock reduces by β_2 percent. We also see that the range of the capital stock with an initial condition given by $0 < \hat{K}_0 < \hat{K}^*$ is given by $\left]0, \hat{K}^*\right[$.

5.3.2 Conditions for existence, uniqueness and stability

The Markov process $\hat{K}$ of this application is described by (23). We follow our "checklist" of ch. 5.2.2.

Existence - The process $\hat{K}$ is bounded in probability on average as its state space $\left]0, \hat{K}^*\right[$ is finite. The process $\hat{K}$ also has the weak Feller property as it has the same properties as shown in 4.1: With fixed ω, i.e. fixed jump-times, $\hat{K}(\tau)$ is a continuous function of the initial condition $\hat{K}_0$. Second, for any bounded and continuous function $f : \left]0, \hat{K}^*\right[\to \mathbb{R}$, the mean $E\left(f\left(\hat{K}(\tau)\right)\right)$ is a continuous function in $\hat{K}_0$. The proof follows exactly, mutatis mutandis, the same steps as the arguments in the proof of lemma 4.1. This implies existence of an invariant probability measure.

Uniqueness - Let us establish that $\hat{K}$ is a φ-irreducible T-process. We can follow the arguments in the proof of lemma 4.2 to first show φ-irreducibility: Let the state space be $\mathbb{R}_{\geq 0}$. Regardless of the initial point $K(t) \in \mathbb{R}_{\geq 0}$, it is possible to attain the state $K(\tau)$ in finite time with probability greater than zero. Thus, prop. 3.2 implies irreducibility, where we can take the Lebesgue measure on $\mathbb{R}_{\geq 0}$ as measure μ.

To show that $\hat{K}$ is a T-process, def. 3.8 tells us that we need to find a function $T\left(\hat{K}, A\right)$ which satisfies the three conditions of this definition. Using the same Dirac measure as before in the proof of lem. 4.4, $v = \delta_\tau$, the three conditions require that $T\left(\hat{K}, A\right)$ be continuous with $K_a\left(\hat{K}; A\right) \equiv \int_0^\infty P^t(x, A)\delta_\tau(dt) = P^\tau\left(\hat{K}, A\right) \geq T\left(\hat{K}, A\right) > 0$.

The idea for finding such a function is also borrowed from lem. 4.4: We consider the probability that $\hat{K}(\tau) \in A$ for some $\tau > t$. We decompose it as in (6) into

$$P\left(\hat{K}(\tau) \in A | \hat{K}(t) = \hat{K}_0\right)$$
$$= \Sigma_{k=0}^\infty P\left(\hat{K}(\tau) \in A | \hat{K}(t) = \hat{K}_0, N(\tau) = k\right) P\left(N(\tau) = k\right)$$

and argue that $T\left(\hat{K}; A\right) \equiv P\left(\hat{K}(\tau) \in A | \hat{K}(t) = \hat{K}_0, N(\tau) = 1\right) P\left(N(\tau) = 1\right)$ is continuous in $\hat{K}_0$. Given that $T\left(\hat{K}; A\right)$ is the product of two probabilities, it is positive. As $T\left(\hat{K}; A\right) \leq P\left(\hat{K}(\tau) \in A | \hat{K}(t) = \hat{K}_0\right) = K_a\left(\hat{K}; A\right)$, we are done and

have established that $\hat{K}$ is a T-process. All of this implies that $\hat{K}$ has a unique invariant measure. Combined with the existence of an invariant probability measure, we can conclude that there is a unique invariant probability measure.

Stability - This works as in the OU-case above. We construct a skeleton chain and see that it is a T-process. Then we argue that we can reach any neighborhood with a positive probability. This establishes stability.

6 Conclusion

This paper has introduced methods that allow to prove existence, uniqueness and stability of distributions described by stochastic differential equations which result from the solution of an optimal control problem. These methods are first applied to a matching and saving model. Existence, uniqueness and stability of the optimal process for the state variables, wealth and labour market status, were proven. We proceed by applying these methods to three model classes from the literature. We show that proving existence, uniqueness and stability is relatively straightforward, following the steps in the detailed proofs for the matching model.

We finally provide some caveats about sensible issues for more complex models than the ones used here. Properties 3 and 4 of our "check list" imply Harris recurrence which is crucial for uniqueness. It is well-known that Brownian motion of dimension three or higher is not recurrent, see Øksendal (1998, Example 7.4.2). As a consequence, one can easily find two functions (and any linear combination with positive coefficients) that are densities of invariant measures of the three-dimensional Brownian motion.

The T-property is especially useful for models where randomness is introduced by finite-activity jump processes, i.e., by compound Poisson processes. In diffusion models, usually even the strong Feller property holds, which makes it easy to conclude the T-property. On the other hand, in models driven by infinite-activity jump processes, e.g., Lévy processes with infinite activity, it does not seem clear whether the T-property can lead to useful results. Indeed, in these models, the strong Feller property may and may not hold, see, for instance, Picard (1995/97). On the other hand, the weak Feller property is satisfied for all Lévy processes, implying existence of invariant distributions, see Applebaum (2004, theorem 3.1.9). Looking at these issues in economic applications offers many fascinating research projects for years to come.

A Appendix

- see after references -

References

Alvarez, F., and R. Shimer (2011): "Search and Rest Unemployment," Econometrica, 79, 75–122.

Applebaum, D. (2004): Lévy processes and stochastic calculus, vol. 93 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge.

Azema, J., M. Duflo, and D. Revuz (1969): "Mesure invariante des processus de Markov recurrents.," Sem. Probab. III, Univ. Strasbourg 1967/68, Lect. Notes Math. 88, 24-33 (1969).

Bandi, F. M., and T. H. Nguyen (2003): "On the functional estimation of jump-diffusion models," Journal of Econometrics, 116(1-2), 293–328.

Bayer, C., and K. Wälde (2011): "Describing the Dynamics of Distributions in Search and Matching Models by Fokker-Planck Equations," mimeo - available at www.waelde.com/pub.

Bental, B., and D. Peled (1996): "The Accumulation of Wealth and the Cyclical Generation of New Technologies: A Search Theoretic Approach," International Economic Review, 37, 687–718.

Bhattacharya, R., and M. Majumdar (2004): "Random dynamical systems: a review," Economic Theory, 23(1), 13–38.

Bismut, J.-M. (1975): "Growth and Optimal Intertemporal Allocation of Risks," Journal of Economic Theory, 10(2), 239–257.

Brock, W., and M. Magill (1979): "Dynamics under Uncertainty," Econometrica, 47(4), 843–868.

Brock, W., and L. Mirman (1972): "Optimal Economic Growth and Uncertainty: The Discounted Case," Journal of Economic Theory, 4(3), 479–513.

Burdett, K., C. Carrillo-Tudela, and M. G. Coles (forthcoming): "Human Capital Accumulation and Labour Market Equilibrium," International Economic Review.

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," International Economic Review, 39, 257–273.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-job Search: Theory and Evidence," Econometrica, 74, 323–365.

Chang, F.-R., and A. Malliaris (1987): "Asymptotic Growth under Uncertainty: Existence and Uniqueness," Review of Economic Studies, 54(1), 169–174.

Coles, M. G., and D. T. Mortensen (2011): "Dynamic Monopsonistic Competition and Labor Market Equilibrium," mimeo.

Down, D., S. P. Meyn, and R. L. Tweedie (1995a): "Exponential and Uniform Ergodicity of Markov Processes," Annals of Probability, 23, 1671 – 1691.

———— (1995b): "Exponential and uniform ergodicity of Markov processes," Ann. Probab., 23(4), 1671–1691.

Evans, L., and T. Kenc (2003): "Welfare cost of monetary and fiscal policy shocks," Macroeconomic Dynamics, 7(2), 212–238.

Francois, P., and H. Lloyd-Ellis (2003): "Animal Spirits Trough Creative Destruction," American Economic Review, 93, 530–550.

———— (2008): "Implementation Cycles, Investment, And Growth," International Economic Review, 49(3), 901–942.

Futia, C. (1982): "Invariant Distributions and the Limiting Behavior of Markovian Economic Models," Econometrica, 50(2), 377–408.

Gabaix, X. (1999): "Zipf's Law For Cities: An Explanation," Quarterly Journal of Economics, 114(3), 739–767.

García, M. A., and R. J. Griego (1994): "An Elementary Theory of Stochastic Differential Equations Driven by A Poisson Process," Communications in Statistics - Stochastic Models, 10(2), 335–363.

Gertler, M., and E. Grinols (1982): "Monetary randomness and investment," Journal of Monetary Economics, 10(2), 239–258.

Gokan, Y. (2002): "Alternative government financing and stochastic endogenous growth," Journal of Economic Dynamics and Control, 26(4), 681–706.

Gong, L., and H.-F. Zou (2002): "Direct preferences for wealth, the risk premium puzzle, growth, and policy effectiveness," Journal of Economic Dynamics and Control, 26(2), 247–270.

———— (2003): "Military spending and stochastic growth," Journal of Economic Dynamics and Control, 28(1), 153–170.

Grinols, E. L., and S. J. Turnovsky (1998): "Consequences of debt policy in a stochastically growing monetary economy," International Economic Review, 39(2), 495–521.

———— (1998b): "Risk, Optimal Government Finance and Monetary Policies in a Growing Economy," Economica, 65(259), 401–427.

Hansen, L. P., and J. A. Scheinkman (2009): "Long-Term Risk: An Operator Approach," Econometrica, 77(1), 177–234.

Hopenhayn, H., and E. Prescott (1992): "Stochastic Monotonicity and Stationary Distributions for Dynamic Economies," Econometrica, 60(6), 1387–1406.

Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," Journal of Economic Dynamics and Control, 17, 953–969.

Kaas, L., and P. Kircher (2011): "Efficient Firm Dynamics in a Frictional Labor Market," mimeo Konstanz and LSE.

Karlin, S., and H. Taylor (1981): A second course in stochastic modeling. Academic Press.

——— (1998): An introduction to stochastic modeling. Academic Press, 3rd ed.

Kusuoka, S., and D. Stroock (1987): "Applications of the Malliavin calculus. III," J. Fac. Sci. Univ. Tokyo Sect. IA Math., 34(2), 391–442.

Lentz, R. (2010): "Sorting by search intensity," Journal of Economic Theory, 145(4), 1436–1452.

Lo, A. W. (1988): "Maximum likelihood estimation of generalized Ito processes with discretely sampled data," Econometric Theory, 4, 231–247.

Magill, M. (1977): "A Local Analysis of N-Sector Capital Accumulation under Uncertainty," Journal of Economic Theory, 15(1), 211–219.

Matsuyama, K. (1999): "Growing Through Cycles," Econometrica, 67, 335–347.

Menzly, L., T. Santos, and P. Veronesi (2004): "Understanding Predictability," Journal of Political Economy, 112(1), 1–47.

Merton, R. C. (1975): "An Asymptotic Theory of Growth under Uncertainty," The Review of Economic Studies, 42(3), 375–393.

Meyn, S. P., and R. L. Tweedie (1993a): Markov chains and stochastic stability, Communications and Control Engineering Series. Springer-Verlag London Ltd., London.

Meyn, S. P., and R. L. Tweedie (1993b): "Stability of Markovian processes. II. Continuous-time processes and sampled chains," Adv. in Appl. Probab., 25(3), 487–517.

——— (1993c): "Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes," Adv. in Appl. Probab., 25(3), 518–548.

Moscarini, G. (2005): "Job Matching and the Wage Distribution," Econometrica, 73(2), 481–516.

Moscarini, G., and F. Postel-Vinay (2008): "The Timing of Labor Market Expansions: New Facts and a New Hypothesis," NBER Macroeconomic Annual 23.

——— (2010): "Stochastic Search Equilibrium," mimeo Bristol and Yale.

Nishimura, K., and J. Stachurski (2005): "Stability of stochastic optimal growth models: a new approach," Journal of Economic Theory, 122(1), 100–118.

Obstfeld, M. (1994): "Risk-Taking, Global Diversification, and Growth," American Economic Review, 84, 1310–29.

Øksendal, B. (1998): Stochastic Differential Equations. Springer, Fifth Edition, Berlin.

Olson, L., and S. Roy (2006): "Theory of Stochastic Optimal Economic Growth," in Handbook of Optimal Growth, Vol. I: Discrete Time, ed. by R. A. Dana, C. L. Van, T. Mitra, and K. Nishimura. Springer, Berlin.

Picard, J. (1995/97): "Density in small time for Levy processes," ESAIM Probab. Statist., 1, 357–389 (electronic).

Pissarides, C. A. (1985): "Short-run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages," American Economic Review, 75, 676–90.

Posch, O. (2009): "Structural estimation of jump-diffusion processes in macroeconomics," Journal of Econometrics, 153, 196–210.

Posch, O., and K. Wälde (2009): "On the Non-Causal Link between Volatility and Growth," University of Aarhus and University of Louvain la Neuve, Discussion Paper.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," Econometrica, 70, 2295–2350.

Prat, J. (2007): "The impact of disembodied technological progress on unemployment," Review of Economic Dynamics, 10, 106–125.

Rishel, R. (1970): "Necessary and Sufficient Dynamic Programming Conditions for Continuous Time Stochastic Optimal Control," SIAM Journal on Control and Optimization, 8(4), 559–571.

Scott, L. O. (1987): "Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application," Journal of Financial and Quantitative Analysis, 22(4), 419–438.

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," American Economic Review, 95, 25–49.

——— (2007): "Mismatch," American Economic Review, 97(4), 1074–1101.

Shimer, R., and L. Smith (2000): "Assortative Matching and Search," Econometrica, 68(2), 343–369.

Steger, T. M. (2005): "Stochastic growth under Wiener and Poisson uncertainty," Economics Letters, 86(3), 311–316.

Turnovsky, S. J. (1993): "Macroeconomic policies, growth, and welfare in a stochastic economy," International Economic Review, 34(4), 953–981.

Wachter, J. A. (2009): "Can time-varying risk of rare disasters explain aggregate stock market volatility?," mimeo Wharton School.

Wälde, K. (2002): "The economic determinants of technology shocks in a real business cycle model," Journal of Economic Dynamics and Control, 27(1), 1–28.

——— (2005): "Endogenous Growth Cycles," International Economic Review, 46, 867–894.

——— (2010): Applied Intertemporal Optimization. Gutenberg University Press, available at www.waelde.com/aio.

——— (2011): "Production technologies in stochastic continuous time models," Journal of Economic Dynamics and Control, 35, 616–622.

A   Appendix

This appendix contains various derivations omitted in the main part.

A.1   Mean and variance in (14) and (15)

Consider the geometric Brownian process (12). Given the solution $K\left(\tau\right) = K_0 e^{\left(g-\frac{1}{2}\sigma^2\right)\left(\tau-t\right)+\sigma z(\tau)}$, we would now like to know what the expected level $K\left(\tau\right)$ in τ from the perspective of t is. To this end, apply the expectations operator E_t and find

$$E_t K\left(\tau\right) = K_0 E_0 e^{\left(g-\frac{1}{2}\sigma^2\right)\left(\tau-t\right)+\sigma z(\tau)} = K_0 e^{\left(g-\frac{1}{2}\sigma^2\right)\left(\tau-t\right)} E_0 e^{\sigma z(\tau)}, \qquad (24)$$

where the second equality exploited the fact that $e^{\left(g-\frac{1}{2}\sigma^2\right)\left(\tau-t\right)}$ is non-random. As $z\left(\tau\right)$ is standard normally distributed, $z\left(\tau\right) \tilde{} N\left(0, \tau-t\right)$, $\sigma z\left(\tau\right)$ is normally distributed with $N\left(0, \sigma^2\left[\tau-t\right]\right)$. As a consequence, $e^{\sigma z(\tau)}$ is log-normally distributed with mean $e^{\frac{1}{2}\sigma^2[\tau-t]}$ (see e.g. Wälde, 2010, ch. 7.3.3). Hence,

$$E_t K\left(\tau\right) = K_0 e^{\left(g-\frac{1}{2}\sigma^2\right)\left(\tau-t\right)} e^{\frac{1}{2}\sigma^2[\tau-t]} = K_0 e^{g[\tau-t]}. \qquad (25)$$

Note that we can also determine the variance of $K\left(\tau\right)$ by simply applying the variance operator to the solution for $K\left(\tau\right)$,

$$\operatorname{var}_t K\left(\tau\right) = \operatorname{var}_t\left(K_0 e^{\left(g-\frac{1}{2}\sigma^2\right)\left(\tau-t\right)} e^{\sigma z(\tau)}\right) = K_0^2 e^{2\left(g-\frac{1}{2}\sigma^2\right)\left(\tau-t\right)} \operatorname{var}_t\left(e^{\sigma z(\tau)}\right).$$

We now make similar arguments to before when we derived (25). As $\sigma B\left(t\right) \tilde{} N\left(0, \sigma^2\left[\tau-t\right]\right)$, the term $e^{\sigma z(\tau)}$ is log-normally distributed with variance $e^{\sigma^2[\tau-t]}\left(e^{\sigma^2[\tau-t]}-1\right)$. Hence, we find

$$\operatorname{var}_t K\left(\tau\right) = K_0^2 e^{2\left(g-\frac{1}{2}\sigma^2\right)[\tau-t]} e^{\sigma^2[\tau-t]}\left(e^{\sigma^2[\tau-t]}-1\right) = K_0^2 e^{\left(2\left(g-\frac{1}{2}\sigma^2\right)+\sigma^2\right)[\tau-t]}\left(e^{\sigma^2[\tau-t]}-1\right)$$

$$= K_0^2 e^{2g[\tau-t]}\left(e^{\sigma^2[\tau-t]}-1\right).$$

We see that even with $g=0$ where $K\left(\tau\right)$ is a martingale, i.e. for $E_t K\left(\tau\right) = K_0$, the variance increases over time. The variance also obviously increases for $g>0$. To understand the case of a negative g, we write the variance as

$$\sigma_\tau^2 = K_0^2 \left[e^{\left(2g+\sigma^2\right)[\tau-t]} - e^{2g[\tau-t]}\right].$$

This shows that as long as $2g + \sigma^2 \geq 0 \Leftrightarrow (0 >) g \geq -\sigma^2/2$, the variance also increases as the first term, $e^{(2g+\sigma^2)[\tau-t]}$, increases and the second term, $e^{2g[\tau-t]}$, falls and the minus sign implies that σ_τ^2 unambiguously rises. For $g < -\sigma^2/2$, we have ambiguous effects. Computing the time τ derivative gives

$$\frac{d}{d\tau}\mathrm{var}_t K\left(\tau\right) = K_0^2\left(\left(2g + \sigma^2\right)e^{\left(2g+\sigma^2\right)[\tau-t]} - 2ge^{2g[\tau-t]}\right) \geq 0 \Leftrightarrow$$
$$\left(2g + \sigma^2\right)e^{\sigma^2[\tau-t]} \geq 2g.$$

This inequality is satisfied for τ "close to" t but is violated for τ sufficiently large. In other words, the variance first increases and then falls.

As a remark - life would be easier when we consider the log solution. It reads

$$\ln K\left(t\right) = \ln K_0 + \left(g - \frac{1}{2}\sigma^2\right)t + \sigma B\left(t\right).$$

The mean and variance of the log is simply

$$E_t \ln K\left(\tau\right) = \left(g - \frac{1}{2}\sigma^2\right)\left(\tau - t\right),$$
$$\mathrm{var}\left(\ln K\left(\tau\right)\right) = \sigma^2\left[\tau - t\right].$$

Here, the variance unambiguously increases.

A.2   The mean and variance of (16)

- The mean

Consider (16), reproduce here, $dK\left(t\right) = K\left(t\right)\left[gdt + \sigma dB\left(t\right) + \gamma dq\left(t\right)\right].$ The integral version reads

$$K\left(\tau\right) - K\left(t\right) = \int_t^\tau K\left(s\right)gds + \int_t^\tau K\left(s\right)\sigma dB\left(s\right) + \int_t^\tau K\left(s\right)\gamma dq\left(s\right).$$

Forming expectations yields with $\mu\left(u\right) \equiv E_t K\left(u\right)$ and $E_t \int_t^\tau K\left(s\right)\sigma dB\left(s\right) = 0$ from the martingale property of $\int_t^\tau K\left(s\right)\sigma dB\left(s\right)$ and $E_t \int_t^\tau K\left(s\right)\gamma dq\left(s\right) = \int_t^\tau E_t K\left(s\right)\gamma\lambda ds$ from the martingale properties of Poisson processes (e.g. Garcia and Griego, 1994 or Wälde, 2010, ch. 10.5.3) with λ being the arrival rate of the Poisson process

$$\mu\left(\tau\right) - \mu\left(t\right) = \int_t^\tau \mu\left(s\right)gds + \int_t^\tau \mu\left(s\right)\gamma\lambda ds.$$

Differentiating with respect to time τ gives $\dot{\mu}\left(\tau\right) = \mu\left(\tau\right)g + \mu\left(\tau\right)\gamma\lambda$ and solving yields

$$E_t K\left(\tau\right) = K\left(t\right)e^{\left(g+\gamma\lambda\right)\left(\tau-t\right)}. \tag{26}$$

- The variance

Computing the variance is based on

$$\text{var}_t K(\tau) = E_t\left[K(\tau)^2\right] - (E_t K(\tau))^2 = E_t\left[K(\tau)^2\right] - K(t)^2 e^{2(g+\gamma\lambda)(\tau-t)}. \quad (27)$$

The process for $K(t)^2$ with $F(K(t)) = K(t)^2$ and (16) is given by (this is an application of a change-of-variable formula (CVF) for a combined jump-diffusion process. See e.g. Wälde, 2010, ch. 10.2.4)

$$dK(t)^2 = dF(K(t))$$
$$= \left\{ F'(K(t)) gK(t) + \frac{1}{2} F''(K(t)) \sigma^2 K(t)^2 \right\} dt$$
$$+ F'(K(t)) \sigma K(t) dB(t) + [F((1+\gamma)K(t)) - F(K(t))] dq(t)$$
$$= \left\{ 2gK(t)^2 + \sigma^2 K(t)^2 \right\} dt + 2\sigma K(t)^2 dB(t) + \left[(1+\gamma)^2 K(t)^2 - K(t)^2\right] dq(t) \Leftrightarrow$$
$$d\kappa(t) = \left(2g + \sigma^2\right) \kappa(t) dt + 2\sigma\kappa(t) dB(t) + (2+\gamma)\gamma\kappa(t) dq(t)$$

where we defined $\kappa(t) \equiv K(t)^2$ for simplicity. The integral version reads

$$\kappa(\tau) - \kappa(t) = \left(2g + \sigma^2\right) \int_t^\tau \kappa(s)\, ds + 2\sigma \int_t^\tau \kappa(s)\, dB(s) + (2+\gamma)\gamma \int_t^\tau \kappa(s)\, dq(s).$$

With expectations, we get

$$E_t\kappa(\tau) - \kappa(t) = \left(2g + \sigma^2\right) \int_t^\tau E_t\kappa(s)\, ds + (2+\gamma)\gamma \int_t^\tau E_t\kappa(s)\,\lambda ds.$$

The time τ derivative gives

$$\frac{d}{d\tau} E_t\kappa(\tau) = \left(2g + \sigma^2\right) E_t\kappa(\tau) + (2+\gamma)\gamma\lambda E_t\kappa(\tau).$$

Solving this yields with $\omega \equiv 2g + \sigma^2 + (2+\gamma)\gamma\lambda = 2[g+\gamma\lambda] + \sigma^2 + \gamma^2\lambda$

$$E_t\kappa(\tau) = E_t\kappa(t)\, e^{\omega[\tau-t]} \Leftrightarrow E_t K(\tau)^2 = K(t)^2 e^{\omega[\tau-t]}.$$

The expectations operator in front of $K(t)^2$ was dropped as the latter is known in t.

Returning to (27), the variance therefore reads

$$\text{var}_t K(\tau) = K(t)^2 \left[e^{\omega[\tau-t]} - e^{2(g+\gamma\lambda)(\tau-t)}\right]. \quad (28)$$

Can we be sure that $\text{var}_t K(\tau) \geq 0$? (We know that this must hold by definition, but our calculations might contain an error.) Let's see,

$$\text{var}_t K(\tau) \geq 0 \Leftrightarrow e^{\omega[\tau-t]} \geq e^{2(g+\gamma\lambda)(\tau-t)} \Leftrightarrow \omega = 2g + \sigma^2 + (2+\gamma)\gamma\lambda \geq 2(g+\gamma\lambda)$$
$$\Leftrightarrow \sigma^2 + 2\gamma\lambda + \gamma^2\lambda \geq 2\gamma\lambda \Leftrightarrow \sigma^2 + \gamma^2\lambda \geq 0$$

which holds as the arrival rate λ is positive.

- Is there an invariant distribution?

There are two necessary conditions for the distribution to be invariant in the limit. The mean (26) is constant over time if $g + \gamma\lambda = 0$ which is possible for negative jump size or for a negative trend g with positive jumps. The variance (28) is constant, assuming a constant mean with $g + \gamma\lambda = 0$ and the implication that $\omega = 2g + \sigma^2 - (2 + \gamma)g = \sigma^2 - \gamma g$, if

$$\mathrm{var}_t K(\tau) = K(t)^2 \left[e^{(\sigma^2 - \gamma g)[\tau - t]} - 1 \right].$$

Note that $g + \gamma\lambda = 0$ implies that (remember that $\lambda \geq 0$) either g or γ are negative (i.e. either there is a negative trend and positive jumps or a positive trend and negative jumps). This implies that γg is negative and therefore $\sigma^2 - \gamma g > 0$. In other words, adding jumps allows to obtain a constant mean but does not allow to make the variance converge to a constant. Adding jumps to a geometric Brownian motion therefore does not imply an invariant distribution.

A.3   The differential of (18)

To compare the process $K(\tau)$ to the geometric Brownian motion in (12), we compute its differential. We need to bring (18) into a form which allows us to apply the appropriate version of Ito's Lemma (see Wälde, 2010, eq. 10.2.3). Use the expression $dx(t) = \beta[\mu - x(t)]dt + \sigma dB(t)$ from (17) and $K(t) \equiv F(x(t)) = e^{x(t)}$. Ito's lemma then gives

$$dK(t) = dF(x(t)) = \left\{ F_x(x(t))\beta[\mu - x(t)] + \frac{1}{2}F_{xx}(x(t))\sigma^2 \right\} dt + F_x(x(t))\sigma dB(t).$$

Inserting the first and second derivatives of $F(x(t))$ yields

$$dK(t) = \left\{ e^{x(t)}\beta[\mu - x(t)] + \frac{1}{2}e^{x(t)}\sigma^2 \right\} dt + e^{x(t)}\sigma dB(t) \Leftrightarrow$$

$$dK(t) = K(t)\left[\left\{ \beta[\mu - \ln K(t)] + \frac{1}{2}\sigma^2 \right\} dt + \sigma dB(t) \right],$$

where the "iff" reinserted $K(t) = e^{x(t)}$ and divided by $K(t)$.

Obviously, the essential difference to (12) consists in the term "$- \ln K(t)$". This differential representation also appears in the option pricing analysis of Scott (1987).

A.4   The mean and variance of (19)

Consider the solution of the exponentially transformed OU process in (19), i.e.

$$K(\tau) = K^* \left[\frac{K_0}{K^*} \right]^{e^{-\beta[\tau - t]}} e^{\int_t^\tau \sigma e^{-\beta[\tau - s]} dB(s)}.$$

A.4.1 The mean of $K(\tau)$

- The mean and variance of $m(\tau)$

Consider the Ito integral

$$m(\tau) \equiv \int_t^\tau g(\tau, s)\, dB(s) \;\; \text{and} \;\; g(\tau, s) \equiv \sigma e^{-\beta[\tau - s]}. \tag{29}$$

As the time τ argument can be pulled in front of the integral, $m(\tau) = \sigma e^{-\beta\tau} \int_t^\tau e^{\beta s} dB(s)$, and the remaining expression $\int_t^\tau e^{\beta s} dB(s)$ is a martingale (Øksendal, 1998, 5th ed., Corollary 3.2.6), we find that the mean of $m(\tau)$ is given by

$$E_t m(\tau) = m(t) = 0. \tag{30}$$

The variance of $m(\tau)$ follows from

$$\mathrm{var}_t m(\tau) = E_t\left[m(\tau)^2\right] - (E_t m(\tau))^2 = E_t\left[m(\tau)^2\right], \tag{31}$$

where we used the marginal property.

To compute $E_t\left[m(\tau)^2\right]$, we differentiate $m(\tau)$ and find

$$dm(\tau) = g(\tau, \tau)\, dB(\tau) + \int_t^\tau g_\tau(\tau, s)\, dB(s) = \sigma dB(\tau) - \beta \int_t^\tau \sigma e^{-\beta[\tau - s]} dB(s)$$

$$= \sigma dB(\tau) - \beta m(\tau)\, d\tau.$$

We replace τ by t for simplifying subsequent steps,

$$dm(t) = -\beta m(t)\, d\tau + \sigma dB(t).$$

Computing the differential for $m(t)^2$ is then straightforward by applying Ito's Lemma to $F(m(t)) = m(t)^2$,

$$dm(t)^2 = dF(m(t)) = \left\{ F'(m(t))\left[-\beta m(t)\right] + \frac{1}{2}F''(m(t))\sigma^2 \right\} dt + F'(m(t))\sigma dB(t)$$

$$= \left\{ 2m(t)\left[-\beta m(t)\right] + \sigma^2 \right\} dt + 2m(t)\sigma dB(t)$$

$$= \left\{ \sigma^2 - 2\beta m^2(t) \right\} dt + 2\sigma m(t)\, dB(t).$$

The integral version reads

$$m(\tau)^2 - m(t)^2 = \int_t^\tau \left\{ \sigma^2 - 2\beta m^2(s) \right\} ds + \int_t^\tau 2\sigma m(s)\, dB(s).$$

Forming expectations, taking the martingale property $E_t \int_t^\tau 2\sigma m(s)\, dB(s) = 0$ into account and moving the expectations operator into the integral, yields

$$E_t m(\tau)^2 - m(t)^2 = \int_t^\tau \left\{ \sigma^2 - 2\beta E_t m^2(s) \right\} ds.$$

Computing the time τ derivative gives $\frac{d}{d\tau} E_t m(\tau)^2 = \sigma^2 - 2\beta E_t m^2(\tau)$, an ordinary differential equation in the second moment $E_t m(\tau)^2$. It can be solved to give

$$E_t m(\tau)^2 = E_t m(t)^2 e^{-2\beta[\tau-t]} + \int_t^\tau e^{-2\beta[\tau-s]}\sigma^2 ds$$

$$= \frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta[\tau-t]}\right).$$

The final equality used $m(t)^2 = 0$ from (29).

Returning to our original inquiry from (31) into determining the variance of $m(\tau)$, we conclude

$$\text{var}_t m(\tau) = \frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta[\tau-t]}\right). \tag{32}$$

The variance is zero for $\tau = t$ and increases the larger $\tau > t$. It approaches $\sigma^2/2\beta$ in the limit.

- The mean of $E_t e^{m(\tau)}$

Let us continue and compute $E_t e^{m(\tau)}$ with $m(\tau)$ from (29). It is well-known that the Ito Integral $m(\tau)$ is normally distributed. This means that $e^{m(\tau)}$ is lognormally distributed. We can then conclude that the mean of $e^{m(\tau)}$ is given by

$$E_t e^{m(\tau)} = \frac{1}{2}\frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta[\tau-t]}\right) \tag{33}$$

- Finally, the mean of $K(\tau)$

Let us now compute the mean,

$$E_t K(\tau) = K^* \left[\frac{K_0}{K^*}\right]^{e^{-\beta[\tau-t]}} E_t e^{\int_t^\tau \sigma e^{-\beta[\tau-s]}dB(s)}. \tag{34}$$

Given (33), we find

$$E_t K(\tau) = K^* \left[\frac{K_0}{K^*}\right]^{e^{-\beta[\tau-t]}} \frac{1}{2}\frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta[\tau-t]}\right).$$

A.4.2 The variance

Let us now compute the variance.

$$\text{var}_t K(\tau) = \left(K^* \left[\frac{K_0}{K^*}\right]^{e^{-\beta[\tau-t]}}\right)^2 \text{var}_t\left[e^{m(\tau)}\right].$$

As the Ito integral $m(\tau)$ in (29) is normally distributed and mean and variance are given by (30) and (32), we have $m(\tau) \sim N\left(0, \frac{\sigma^2}{2\beta}\left(1 - e^{-2\beta[\tau-t]}\right)\right)$. Then, $e^{m(\tau)}$ is

log-normally distributed with the mean being given by (33) and the variance of $e^{m(\tau)}$ is

$$\text{var}_t \left[e^{m(\tau)} \right] = e^{\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta[\tau-t]}\right)} \left(e^{\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta[\tau-t]}\right)} - 1 \right).$$

The variance of the capital stock is then

$$\text{var}_t K \left(\tau \right) = \left(K^* \left[\frac{K_0}{K^*} \right]^{e^{-\beta[\tau-t]}} \right)^2 e^{\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta[\tau-t]}\right)} \left(e^{\frac{\sigma^2}{2\beta}\left(1-e^{-2\beta[\tau-t]}\right)} - 1 \right).$$