The Dynamics of Distributions in Continuous-Time Stochastic Models

Christian Bayer[(a)] and Klaus Wälde[(b)]
[(a)] Weierstraß Institute Berlin and [(b)]University of Mainz[1]

October 2013

We study an optimal precautionary-saving problem in continuous time. The evolution of optimally evolving state variables, wealth and labour market status, can be described 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. By deriving Fokker-Planck equations for these state variables, we can provide an intuitive interpretation of the evolution and determinants of the implied distribution of wealth.

JEL Codes: C62, D91, J63
Keywords: uncertainty in continuous time, Poisson process, existence, uniqueness, stability, Fokker-Planck equations

1 Introduction

[Motivation] Dynamic and stochastic models are widely used for macro economic analysis and also for many analysis 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 missing for continuous time uncertainty.

[Our application] 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 deterministic means. The classic example is Pissarides (1985). Papers in the tradition of Burdett and Mortensen (1998) that do focus on distributions construct these

[1]Christian Bayer: Weierstrass Institute, Mohrenstr. 39, 10117 Berlin, Germany. christian.bayer@wias-berlin.de. Klaus Wälde: University of Mainz, Gutenberg School of Management and Economics, Jakob-Welder-Weg 4, 55128 Mainz, Germany. klaus@waelde.com, www.waelde.com. Klaus Wälde is also a member of CESifo and an extramural fellow of the University of Bristol and the Université catholique de Louvain. We are grateful to William Brock, John Stachurski and Stephen Turnovsky for comments and suggestions.

[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.

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.

[Objectives] 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 are therefore particularly relevant for 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. This model would be an ingredient of any extension of textbook search and matching models in continuous time that allow for self-insurance. Individuals have constant relative risk aversion and an infinite planning horizon. 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 an invariant (stationary) distribution for wealth and employment status exists, is unique and when the model is stable in the sense that the distribution converges for any initial distribution to the unique invariant one. The corresponding theorem is proven.

Our third objective consists in providing some economic interpretations for the determinants of the distribution of wealth implied by the matching and saving process. To this end, we provide a tool that embeds the analysis of distributions into a standard mathematical tool - the so-called Fokker-Planck equations. These equations describe the distributional properties of stochastic processes in a fairly general but still intuitive way. The advantage of these equations consists in the fact that one is no longer restricted to specific distributions for which closed-form solutions can be found. The entire dynamics of distributions is described and not simply distributions in a "steady-state". They can also be applied to much more general processes than has been done so far in the literature. By their nature, all existing distributions must be special cases of these general equations.[6]

[Findings] 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 our model, 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

[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, 2013), Coles and Mortensen (2011) and Kaas and Kircher (2011). None of these papers uses the methods we propose here.

[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] As Fokker-Planck equations describe densities, this method would allow for structural maximum likelihood estimation of models that include additional features to those usually captured in labour models (see e.g. van den Berg, 1990; Postel-Vinay and Robin, 2002; Flinn, 2006 or Launov and Wälde, 2013). In work in progress, we use the Fokker-Planck equations derived in this paper to understand determinants of the wealth distribution of a cohort in the US based on the NLSY79.

thus obtain uniqueness of the invariant distribution. Ergodicity is then implied by properties of discrete skeleton chains.

Using the Dynkin formula, we compute the Fokker-Planck equations for the wealth-employment status system, a two-dimensional partial differential equation system. It describes the evolution of the density of wealth and employment status over time, given some initial condition. When we are interested in long-run properties only, we can set time derivatives equal to zero in the Fokker-Planck equations and obtain an ordinary two-dimensional non-autonomous differential equation system. Boundary conditions can be motivated from our phase diagram analysis.

The big advantage of our example for illustrating the usefulness of Fokker-Planck equations, consists in the generic nature of the resulting stochastic system. There will be one fundamental equation that describes the ins into and outs out of employment. Then, there will be one "dependent" equation that describes the accumulation of wealth. If wealth is replaced by firm-size, human capital, entitlement to benefits or duration in employment or unemployment, exactly the same structure occurs.

[Table of contents] The structure of our paper is as follows. The next section relates our analysis to the literature. Section 3 presents the consumption-saving problem and derives the stochastic differential equations describing the evolution of state variables. Section 4 proves existence and uniqueness of an invariant measure for the state variables together with convergence to the long-run invariant distribution. Section 5 provides a more applied approach to describing the dynamics of distributions by presenting the Fokker-Planck equations. The final section concludes. Appendix A provides some general background to stochastic processes in continuous time to which we refer in the main text. Appendix B derives the Fokker-Planck equations step by step.

2 Related literature

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). Partially inspired by Bhattacharya and Lee (1988), Kamihigashi and Stachurski (2012, 2013) considerably weaken the assumptions of Hopenhayn and Prescott, in particular allowing for unbounded shocks in models for exploitation of renewable resources or dynamic models for wealth distributions in discrete time. 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). A stochastic growth model with capital and human capital accumulation is studied by Krebs (2006). The effect of predicting shocks was analysed recently by Roy and Zilcha (2012).

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 T-processes of Meyn and Tweedie (a weaker version of strong Feller processes), seems new in the economics literature. While relying on other results from Meyn and Tweedie (1993a), Kamihigashi and Stachurski (2012, 2013), for instance, infer stability from order mixing properties instead.

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 principles behind and the derivation of the Fokker-Planck equation (FPE) for Brownian motion are treated e.g. in Friedman (1975, ch. 6.5) or Øksendal (1998, ch. 8.1). For our case of a stochastic differential equation driven by a Markov chain, we use the infinitesimal generator as presented e.g. in Protter (1995, ex. V.7). From general mathematical theory, we know that the density satisfies the corresponding FPE $\frac{\partial}{\partial t}p(t,x) = \mathcal{A}^* p(t,x)$, where p denotes the density of the process with state variable x at time t and $\mathcal{A}^*$ is the adjoint operator of the infinitesimal generator $\mathcal{A}$ of this process. We follow this approach in our framework and obtain the FPE for the law of the employment-wealth process.

In economics, versions of Fokker-Planck equations, equivalently called Kolmogorov forward equations, are rarely used or referred to so far. Lo (1988) derives a FPE for a one-dimensional process. Merton (1975) applies the method to analyse distributional properties of a stochastic Solow growth model. Bertola and Caballero (1994) study the distribution of capital when investment is irreversible. Klette and Kortum (2004) employ a method related to FPEs to derive firm-size distributions. Moscarini (2005) uses them to derive the distribution of the belief about the quality of a match. Koeniger and Prat (2007) obtain an employment distribution and Prat (2007) describes the distribution of detrended productivity. Impullitti et al. (2011) study the firm-size distribution in an international trade context. Stokey (2008) provides a text-book treatment of Kolmogorov forward and backward equations for Brownian motion.

[7] Continuous time models are treated thoroughly, but under different conditions, in the finance literature. As an example, Raimondo (2005) proves existence of equilibrium in a model with incomplete and with complete markets. Anderson and Raimondo (2008) prove dynamic completeness of the equilibrium price process.

The main difference in our application consists in its considerable generalization (as we allow for a system of stochastic differential equations with jumps), in the detailed derivation and in the explanations linking the derivation to standard methods taught in advanced graduate courses. The only new tool we require is the Dynkin formula. This approach focusing on the principles of FPEs in a tractable and accessible way should allow and encourage a much wider use of this tool for other applications. We would like to move Fokker-Planck equations much more into the mainstream. In fact, one could argue that Fokker-Planck equations should become a tool as common as Keynes-Ramsey rules.[8]

By transforming the FPEs from equations describing densities into equations describing distribution functions, we obtain a description of densities whose intuitive interpretation is very similar to derivations of less complex distributions as in Burdett and Mortensen (1998) or Burdett et al. (2011). In addition, however, our equations exhibit new "advection" terms that capture the shift of the distribution due to the evolution of the additional state variable, i.e. due to wealth.

3 The model

3.1 The setup and optimal consumption

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 denoted 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 that the instantaneous utility functions has a CRRA structure

$$u(c(\tau)) = \frac{c(\tau)^{1-\sigma} - 1}{1 - \sigma} \tag{1}$$

with $\sigma \neq 1$. All proofs for the logarithmic case $\sigma = 1$ should work accordingly.

Each individual can save in an asset a. Allowing optimal consumption to be a function of state variables, $c(a(t), z(t))$, the optimal evolution of individual wealth is given by

$$da(t) = \{ra(t) + z(t) - c(a(t), z(t))\}\, dt. \tag{2}$$

Wealth $a(t)$ increases (or decreases) per unit of time dt if capital income $ra(t)$ plus labour income $z(t)$ is larger (or smaller) than optimally chosen consumption $c(a(t), z(t))$. Labour income $z(t)$ is given by constants w and b[9] and is described by the second constraint of the household, a stochastic differential equation,

$$dz(t) = \Delta dq_\mu - \Delta dq_s, \qquad \Delta \equiv w - b. \tag{3}$$

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

[8] We would like to thank Philipp Kircher for having put this so nicely.

[9] In some broader equilibrium perspective, w and b would be endogenous objects. As long as there is only idiosyncratic risk and income is a deterministic function of time, all of our proofs below would work as well.

the paper. As usual, the wealth-employment process (a, z), is defined on a probability space $(\Omega, \mathcal{F}, P)$.[10]

We now let the individual maximize her objective function by choosing a consumption path subject to the budget constraint (2) and the equation for the employment status (3). Optimal consumption is described by the following generalized Keynes-Ramsey rules which extends the approach suggested by Wälde (1999) for the case of an uncertain interest rate to our case of uncertain labour income. We suppress the time argument for readability. Consumption $c(a_w, w)$ of an employed individual with current wealth a_w follows (see app. C.1)

$$\begin{aligned} -\frac{u''(c(a_w,w))}{u'(c(a_w,w))} dc(a_w,w) = &\left\{ r - \rho + s\left[\frac{u'(c(a_w,b))}{u'(c(a_w,w))} - 1\right]\right\} dt \\ &- \frac{u''(c(a_w,w))}{u'(c(a_w,w))}\left[c(a_w,b) - c(a_w,w)\right] dq_s \end{aligned} \tag{4a}$$

while her wealth evolves according to

$$da_w = \left[ra_w + w - c(a_w,w)\right] dt. \tag{4b}$$

Analogously, solving for the optimal consumption of an unemployed individual with current wealth a_b yields

$$\begin{aligned} -\frac{u''(c(a_b,b))}{u'(c(a_b,b))} dc(a_b,b) = &\left\{ r - \rho - \mu\left[1 - \frac{u'(c(a_b,w))}{u'(c(a_b,b))}\right]\right\} dt \\ &- \frac{u''(c(a_b,b))}{u'(c(a_b,b))}\left[c(a_b,w) - c(a_b,b)\right] dq_\mu \end{aligned} \tag{4c}$$

and her wealth follows

$$da_b = [ra_b + b - c(a_b,b)]dt. \tag{4d}$$

Without uncertainty about future labor income, i.e. $s = \mu = dq_s = dq_\mu = 0$, the above Keynes-Ramsey rules reduce to the classical deterministic consumption rule, $-\frac{u''(c)}{u'(c)}\dot{c} = r - \rho$. The additional $s[\ldots]$ term in (4a) shows that consumption growth is faster under the risk of a job loss. Note that the expression $[u'(c(a_w,b))/u'(c(a_w,w)) - 1]$ is positive as consumption $c(a_w,b)$ of an unemployed worker is smaller than consumption of an employed worker $c(a_w,w)$ (see lem. C.12 for a proof) and marginal utility is decreasing, $u'' < 0$. Similarly, the $\mu[\ldots]$ term in (4c) shows that consumption growth for unemployed workers is smaller.

As the additional term in (4a) contains the ratio of marginal utility from consumption when unemployed relative to marginal utility when employed, this suggests that it stands for precautionary savings (Leland, 1968, Aiyagari, 1994, Huggett and Ospina, 2001).[11] When marginal utility from consumption under unemployment is much higher than marginal utility from employment, individuals experience a high drop in consumption when becoming unemployed. If relative consumption shrinks as wealth rises, i.e. if $\frac{d}{da}\frac{c(a,w)}{c(a,b)} < 0$, reducing this gap and smoothing consumption is best achieved by fast capital accumulation. This fast capital accumulation would go hand in hand with fast consumption growth as visible in (4a).

In the case of unemployment, the $\mu[\ldots]$ term in (4c) suggests that the possibility to find a new job induces unemployed individuals to increase their current consumption level. Relative to

[10] Our model can be seen as a continuous-time model in the spirit of Aiyagari (1994). A similar model is used by Lise (2013) that allows for on-the-job search. He does not employ Fokker-Planck equations and abstracts from existence and stability analyses.

[11] If the individual knew the points in time where she moves to another state, the Keynes-Ramsey rule would not display this term. In fact, an explicit solution for the consumption *level* would be available for any wage path (see e.g. Wälde, 2012, eq. (5.6.10)).

a situation in which unemployment is an absorbing state (once unemployed, always unemployed, i.e. $\mu = 0$), the prospect of a higher labor income in the future reduces the willingness to give up today's consumption. With higher consumption levels, wealth accumulation is lower and consumption growth is reduced.

The stochastic dq-terms in (4a) and (4c) (tautologically) represent the discrete jumps in the level of consumption whenever the employment status changes. We will understand more about these jumps after the phase-diagram analysis below.

For our analysis to follow, we assume that the interest rate is lower than the time-preference rate, $r < \rho$. For convenience, we also assume that the initial wealth level $a(t)$ is chosen inside the interval $[-b/r, a_w^*]$. The lower bound $-b/r$ is a natural borrowing constraint as discussed below and the upper bound a_w^* is endogenously determined below as well.[12]

3.2 An illustration of consumption and wealth dynamics

The dynamics of consumption and wealth can be illustrated in the wealth-consumption space. The background for this illustration results from initially focusing on the evolution between jumps and by eliminating time as exogenous variable. Computing the derivatives of consumption with respect to wealth in both states and considering wealth as the exogenous variable, we obtain a two-dimensional system of non-autonomous ordinary differential equations (ODE). As wealth is now the argument for these two differential equations, there is no longer a need to distinguish between wealth of employed and unemployed workers (i.e. between a_w and a_b). The dynamics between jumps therefore follows

$$-\frac{u''(c(a,w))}{u'(c(a,w))}\frac{dc(a,w)}{da} = \frac{r-\rho+s\left[\frac{u'(c(a,b))}{u'(c(a,w))}-1\right]}{ra+w-c(a,w)}, \tag{5a}$$

$$-\frac{u''(c(a,b))}{u'(c(a,b))}\frac{dc(a,b)}{da} = \frac{r-\rho-\mu\left[1-\frac{u'(c(a,w))}{u'(c(a,b))}\right]}{ra+b-c(a,b)}. \tag{5b}$$

With two boundary conditions, this system provides a unique solution for $c(a,w)$ and $c(a,b)$. Once solved, the effect of a jump is then simply the effect of a jump of consumption from, say, $c(a,w)$ to $c(a,b)$.

Properties of this system can then be illustrated in the usual way by plotting zero-motion lines and by plotting the sign of the derivatives into a phase diagram. Following these steps, it turns out (see app. C.2) that there is an endogenous upper limit a_w^* of the wealth distribution determined by the zero-motion line for consumption. The ratio of consumption at this point is given by

$$\frac{u'(c(a_w^*,b))}{u'(c(a_w^*,w))} \equiv 1 - \frac{r-\rho}{s}. \tag{6}$$

Joint with an endogenous natural borrowing limit of $a \geq -b/r$ (see app. C.3), this allows us to plot a phase diagram as in fig. 1.[13] This figure displays wealth on the horizontal and consumption $c(a,z)$ on the vertical axis. It plots dashed zero-motion lines for a_w and $c(a,w)$ and a solid zero-motion line for a_b following from (4b), (55) and (4d), respectively. We assume for this figure that the threshold level a_w^* is positive.[14] The intersection point of the zero-motion

[12] Our discussion below suggests that wealth will lie within this intervall after a finite lenght of time with probability one even when initial wealth $a(t)$ lies outside the intervall.

[13] App. C proves various properties of our system used for plotting this phase diagram under a mild technical condition. A proof of the existence of an optimal consumption path is in app. D.

[14] This is of course a quantitative issue. In ongoing numerical work, the threshold is positive for reasonable parameter values. It approaches infinity for r approaching ρ.

lines for $c(a,w)$ and a_w is the temporary steady state (TSS),

$$\Theta \equiv (a_w^*, c(a_w^*, w)) . \tag{7}$$

We call this point *temporary* steady state for two reasons. On the one hand, employed workers experience no change in wealth, consumption or any other variable when at this point (as in a standard steady state of a deterministic system). On the other hand, the expected spell in employment is finite and a random transition into unemployment will eventually occur. Hence, the state in Θ is steady only temporarily.

As we know from the proposition in app. C.2 that consumption for the unemployed always falls, both consumption and wealth fall above the zero-motion line for a_b. The arrow-pairs for the employed workers are also added. They show that one can draw a saddle-path through the TSS. To the left of the TSS, wealth and consumption of employed workers rise, to the right, they fall.

Relative consumption when the employed worker is in the TSS is given by (6). A trajectory going through $(a_w^*, c(a_w^*, b))$ and hitting the zero-motion line of a_b at $-b/r$ is in accordance with laws of motions for the unemployed worker.

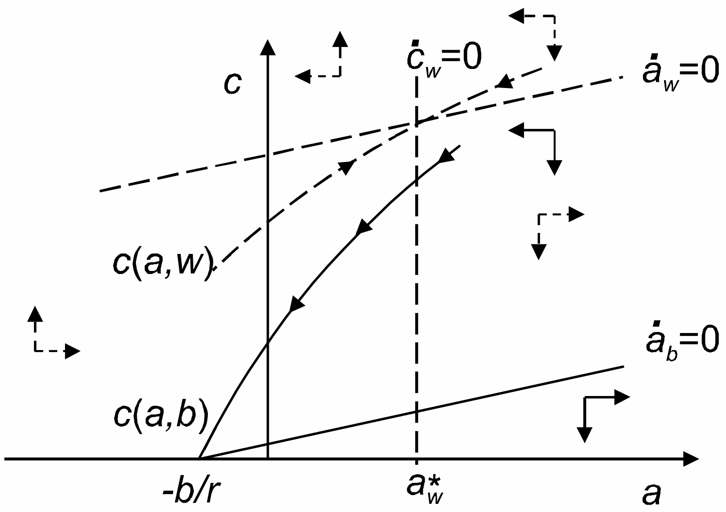

Figure 1 *Policy functions for employed and unemployed workers*

For our assumption of an interest rate being lower than the time preference rate, $r < \rho$, the range of wealth a worker can hold is bounded. Whatever the initial wealth level, there is a positive probability that the wealth level will be in the range $[-b/r, a_w^*]$ after some finite length of time. For an illustration, consider the policy functions in fig. 1: Wealth decreases both for employed and unemployed workers for $a > a_w^*$. The transition into the range $[-b/r, a_w^*]$ will take place only in the state of unemployment which, however, occurs with positive probability.

When wealth of an individual is within the range $[-b/r, a_w^*]$, consumption and wealth will rise while employed and fall while unemployed. While employed, precautionary saving motives drive the worker to accumulate wealth. While unemployed, the worker runs down current wealth as higher income for the future is anticipated – "postcautionary dissaving" takes place. When a worker loses a job at a wealth level of, say, $a_w^*/2$, his consumption level will drop from $c(a_w^*/2, w)$ to $c(a_w^*/2, b)$. Conversely, if an unemployed worker finds a job at, say, $a = 0$, her consumption increases from $c(0, b)$ to $c(0, w)$. A worker will therefore be in a permanent consumption and wealth cycle. Given these dynamics, wealth will never leave the interval $[-b/r, a_w^*]$ and one can easily imagine a distribution of wealth over the range $[-b/r, a_w^*]$.

4 Stability of the wealth-employment process

We would now like to formally understand the stability properties of the model just presented. As the fundamental state variables are wealth (2) and the employment status (3) of an individual, the process we are interested in is the wealth-employment process $X_\tau \equiv (a(\tau), z(\tau))$. All other variables (like control variables or e.g. factor rewards in a general equilibrium version) 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. The state-space of this process X_τ is $\mathbf{X} \equiv [-b/r,\ a_w^*] \times \{w, b\}$ and has all the properties required for the state space in appendix A. Moreover, for the sake of simplicity, we now set the initial time $t = 0$ – following the usual practice in the mathematical literature. For the necessary mathematical background on the ergodicity theory of Markov processes we refer to appendix A.

4.1 Existence

In order to show existence for an invariant probability measure for X, we need (i) some compactness result for X like boundedness in probability on average recalled in def. A.4 and (ii) a continuity property like the weak Feller property, see prop. A.6. Showing that X is bounded in probability on average is straightforward: According to def. A.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. Thus, $P_\tau f$ is, indeed, bounded and continuous whenever f is bounded and continuous, and the weak Feller property holds. ■

4.2 Uniqueness

Given existence of an invariant distribution, uniqueness will follow from (Harris) recurrence together with irreducibility of the process X. The details are spelled out in appendix A.3, in particular in prop. A.7.

We prove 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. A.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. A.2 implies irreducibility with irreducibility measure

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

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. A.9 and need to prove that X is a T-process, see def. A.8. 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

$$\begin{aligned} P(\tau_1 \le u, \ N_\tau = 1) &= P(\tau_1 \le u, \ \tau_2 \ge \tau - \tau_1) = \int_0^u P(\tau_2 \ge \tau - v)se^{-sv}dv \\ &= \int_0^u e^{-\mu(\tau - v)}se^{-sv}dv = \frac{s}{\mu - s}e^{-\mu\tau}\left(e^{(\mu-s)u} - 1\right). \end{aligned}$$

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. ■

As discussed in app. A, uniqueness of the invariant distribution of a Markov process is implied by smoothing properties of the process. Obviously, the wealth-employment process (a, z) does not satisfy the strong Feller property (see def. A.5). Indeed, assume that $f : [a_b^*, a_w^*] \times \{w, b\} \to \mathbb{R}$ is bounded measurable, but not continuous. For the sake of concreteness, let us assume that f has a jump at some point $a_b^* < a_0 < a_w^*$. If there is no jump in the employment status until time τ (an event with positive probability), then the trajectory of the wealth process a is deterministic until time τ and z is even constant. Hence, on this event the jump cannot be smeared out. On the other hand, the distribution of the jump times has a smooth density. If there is at least one jump until time τ, we, therefore, expect the discontinuity of f to be smeared out due to the density of the jump times. If both these heuristics are true, then

- the wealth-employment process is not strong Feller, as

$$P_\tau f(a_0, z_0) = E\left[f(a_\tau, z_\tau)\right] = \underbrace{E\left[f(a_\tau, z_\tau)\mathbf{1}_{N_\tau=0}\right]}_{\text{discontinuous in }(a_0,z_0)} + \underbrace{E\left[f(a_\tau, z_\tau)\mathbf{1}_{N_\tau>0}\right]}_{\text{continuous in }(a_0,z_0)}$$

  is discontinuous in (a_0, z_0) – where N denotes the number of jumps in the employment status;

- the wealth-employment status conditioned on the number of jumps being greater then zero should satisfy the strong Feller condition. Hence, the kernel $T((a_0, z_0), A) = P^\tau((a_0, z_0), A \cap \{N_\tau > 0\})$ should be a continuous component of P^τ in the sense of def. A.8. In other words, the wealth-employment process is a T-process.

Indeed, it turns out that these heuristic considerations lead to a correct conclusion.

Theorem 4.4 *The wealth-employment process* $(a(\tau), z(\tau))$ *is a* T*-process.*

Given that there are some technical difficulties concerning the proof of th. 4.4, we first illustrate our approach heuristically. A formal proof is provided afterwards. The main step in establishing that a kernel T is a continuous component of P^τ in the sense of def. A.8 is to show continuity. To this end, let us consider a measurable set $A \subset [a_b^*, a_w^*] \times \{w, b\}$ and define

$$\begin{aligned} T_{>0}((a_0, z_0), A) \equiv P^\tau((a_0, z_0), A \cap \{N_\tau > 0\}) = \\ \int \mathbf{1}_A(a, w) p^\tau_{>0}((a_0, z_0), (a, w)) da P\left(N_\tau > 0\right) + \\ \int \mathbf{1}_A(a, b) p^\tau_{>0}((a_0, z_0), (a, b)) da P\left(N_\tau > 0\right), \end{aligned}$$

where $p^\tau_{>0}((a_0, z_0), (a, z))$ denotes the transition density of the wealth-employment process *conditioned on* $\{N_\tau > 0\}$. Obviously, continuity of $T_{>0}$ is equivalent to continuity of $a_0 \mapsto p^\tau_{>0}((a_0, z_0), (a, w))$ and $a_0 \mapsto p^\tau_{>0}((a_0, z_0), (a, b))$. Moreover, if the heuristic argument is correct, we may actually restrict ourselves to the case when there is exactly one jump in the employment process until time τ. This means, we consider the kernel

$$\begin{aligned} T_1((a_0, z_0), A) \equiv P^\tau((a_0, z_0), A \cap \{N_\tau = 1\}) = \\ \int \mathbf{1}_A(a, z_0'^{\tau}{}_1((a_0, z_0), (a, z_0')) da P\left(N_\tau = 1\right), \end{aligned}$$

where $z_0' \in \{w, b\}$, $z_0' \neq z_0$ and p_1^τ denotes the transition density conditioned on the event that there is exactly one jump until time τ. Now the picture becomes much clearer. Indeed, let us assume that the jump in employment status happens at some time $u < \tau$. Up to time u, the wealth process moves deterministically according to the ODE (2), after time u it again moves in a deterministic way according to (2). Hence, there is a deterministic function ϕ_{z_0} (see (10) for the precise definition) such that

$$a_\tau = \phi_{z_0}(a_0, u; \tau)$$

provided that there is precisely one jump of the employment status at time u (and no other jump before τ). Hence, we may express T_1 by

$$T_1((a_0, z_0), A) = \int_0^\tau \mathbf{1}_A(\phi_{z_0}(a_0, u; \tau), z_0'^{(1)}{}_\tau(u) du P\left(N_\tau = 1\right).$$

If $u \mapsto \phi_{z_0}(a_0, u; \tau)$ where smooth and invertible with smooth inverse $y \mapsto \phi_{z_0}^{-1}(a_0, y; \tau)$, then we could re-write the equation as

$$T_1((a_0, z_0), A) = \int_{low(a_0)}^{up(a_0)} \mathbf{1}_A(y, {z_0'}^{(1)}{}_\tau(\phi_{z_0}^{-1}(a_0, y; \tau))\left|\frac{\partial}{\partial y}\phi_{z_0}^{-1}(a_0, y; \tau)\right| dy, \tag{8}$$

which is continuous in a_0 provided that $a_0 \mapsto \left|\frac{\partial}{\partial y}\phi_{z_0}^{-1}(a_0, y; \tau)\right|$ and $a_0 \mapsto low(a_0)$, $a_0 \mapsto up(a_0)$ are continuous (plus some boundedness assumption). Assuming that we can make all these steps rigorous, we thus have proved the theorem.

In order to verify the various assumptions made in the above sketch , we need to understand the solution of the ODE

$$\frac{da_z(\tau)}{d\tau} = ra_z(\tau) + z - c(a_z(\tau), z) \tag{9}$$

better. Indeed, the properties would be essentially trivial, if it were not for the (possible) singularity of the consumption function $c(a, z)$ at $a = a_b^*$ and $a = a_w^*$ induced by the explosion of the right hand side in (5). Nevertheless, by careful analysis we can establish the assumptions made above, at least when we further restrain the domain.

We denote the solution of (9) started at $a_0 \in [a_b^*, a_w^*]$ at time 0 evaluated at time $\tau = u$ by $\psi_z(a_0, u)$, i.e., $\psi_z(a_0, 0) = a_0$. Let $\mathfrak{T}(a, z) \in [0, \infty]$ be the time it takes for the deterministic function $\psi_z(a, \cdot)$ to reach the boundary $\{a_b^*, a_w^*\}$ of the domain. Note that $\mathfrak{T}$ may be infinite, which is actually the good situation, as the consumption function $c(a, z)$ is actually C^1 in that case – and, hence, stability holds. While it seems not clear how to obtain C^1 on the whole interval $[a_b^*, a_w^*]$, it is clear how to get it on the interior of the domain. Of course, if $\mathfrak{T}(a, z) = \infty$ for some $a \in]a_b^*, a_w^*[$, then it is infinite for any such a.

Lemma 4.5 *For $z = w, b$, the map $a \mapsto c(a, z)$ is C^1 in the interior $]a_b^*, a_w^*[$ of the domain.*

Proof. $x(a) \equiv (c(a, w), c(a, b))$ solves an ODE in a (the reduced form ODE system), with a right hand side which is locally Lipschitz in the interior of the domain. Fix some interior value a_0 and consider the initial value problem for x started at $x(a_0)$ on the a-domain $[a_0, a_w^*[$. As the right hand side is locally Lipschitz, we can apply the usual existence and uniqueness theorem, which gives, in particular, that the solution is C^1 up to (but not necessarily including) $a = a_w^*$. On the other hand, for $a \in]a_b^*, a_0]$, we just revert the direction, which gives another locally Lipschitz right hand side, and, hence, C^1 follows in the same way. ■

This directly implies that $\psi_z(a, u)$ is C^1 in both a and u for $u < \mathfrak{T}(a, z)$, and continuous in both variables even for $u \leq \mathfrak{T}(a, z)$.

Lemma 4.6 *The map $a \mapsto \mathfrak{T}(a, z)$ is continuous on $[a_b^*, a_w^*] \setminus \{a_z^*\}$. Moreover, if $\mathfrak{T}(a, z) < \infty$ for any $a_b^* < a < a_w^*$, then $\mathfrak{T}(\cdot, z)$ is continuous on the whole domain.*[15]

Proof. Let $\psi_z(a, u)$ denote the solution map of the ODE driving a_z evaluated at time u for initial value $\psi_z(a, 0) = a$. Obviously, $\psi_w(a, \cdot)$ is strictly increasing (until the time that a_w^* is hit), while $\psi_b(a, \cdot)$ is strictly decreasing. Hence, they have continuous inverse functions (in t, for fixed a).

Fix any point $a^0 \in]a_b^*, a_w^*[$ and the corresponding value $\mathfrak{T}^0(z) \equiv \mathfrak{T}(a^0, z)$. For any positive t we obviously have

$$\mathfrak{T}(\psi_z(a, t), z) = \mathfrak{T}(a, z) - t.$$

Denoting $\psi_z^0(t) \equiv \psi_z(a^0, t)$, we get for any $a < a^0$ for $z = b$ and any $a > a^0$ for $z = w$ that

$$\mathfrak{T}(a, z) = \mathfrak{T}(\psi_z^0((\psi_z^0)^{-1}(a)), z) = T^0(z) + (\psi_z^0)^{-1}(a),$$

[15] Otherwise, we have a jump from $+\infty$ to 0 at $a = a_z^*$.

which is continuous in a. As a^0 was arbitrary in the interior of the interval, the claim follows. ■

Let us introduce a little bit of notation: for $z \in \{w, b\}$ we denote by z' the *other* element of $\{w, b\}$. Moreover, we define

$$\phi_z(a, u; \tau) \equiv \psi_{z'}\left(\psi_z(a, u), \tau - u\right), \quad 0 \leq u \leq \tau,\ z \in \{w, b\}. \tag{10}$$

In words, ϕ_z denotes the value of the wealth process at time τ given that the wealth process at time 0 has the value a and there is precisely one change of the employment status (from z to z') in $[0, \tau]$, which takes place at time u. We are going to identify a sufficiently large set of us on which $u \mapsto \phi_z(a, u; \tau)$ is differentiable and invertible with differentiable inverse.

Lemma 4.7 *Define the set*

$$\mathfrak{S}(a, z; \tau) \equiv \{u \in [0, \tau] \mid u > \tau - \mathfrak{T}\left(\psi_z(a, u), z'\right)\}.$$

If $\mathfrak{T}(a, z') = \infty$ *for some* $a_b^* < a < a_w^*$*, then*

$$\mathfrak{S}(a, z; \tau) = \begin{cases} [0, \tau], & a \neq a_{z'}^*, \\ ]0, \tau], & a = a_{z'}^*. \end{cases}$$

Otherwise, the following three properties hold:

1. *There are numbers* $\mathfrak{s}(a, z; \tau)$ *such that* $\mathfrak{S}(a, z; \tau) =]\mathfrak{s}(a, z; \tau), \tau]$.
2. $a \mapsto \mathfrak{s}(a, z; \tau)$ *is continuous on* $]a_b^*, a_w^*[$.
3. *For every* $(a, z) \in [a_b^*, a_w^*] \times \{w, b\}$ *we have (uniformly)* $\tau - \mathfrak{s}(a, z; \tau) > 0$.

Proof. The description for $\mathfrak{T}(a, z') = \infty$ is obvious, so we assume that $\forall a \in [a_b^*, a_w^*] \setminus \{a_{z'}^*\}$: $\mathfrak{T}(a, z') < \infty$.

First note that $\tau \in \mathfrak{S}(a, z; \tau)$. Moreover, for $u < v < \tau$ we have that $u \in \mathfrak{S}(a, z; \tau)$ implies $v \in \mathfrak{S}(a, z; \tau)$, since

$$\tau - \mathfrak{T}\left(\psi_z(a, v), z'\right) \leq \tau - \mathfrak{T}\left(\psi_z(a, u), z'\right) < u < v,$$

which shows that $\mathfrak{S}(a, z; \tau)$ is an interval. However, for its lower endpoint the inequality is no longer strict, implying that the interval is closed to the right, but open to the left.

For the continuity of $\mathfrak{s}$, let us consider any (monotone) converging sequence $a_n \to a \in [a_b^*, a_w^*]$. First, assume that $u \in \mathfrak{S}(a_n, z; \tau)$ for all $n \geq N$. Then $u > \tau - \mathfrak{T}(\psi_z(a_n, u), z')$. Thus, continuity of $\psi_z(\cdot, u)$ and $\mathfrak{T}(\cdot, z')$ (cf. Lemma 4.6) imply that

$$u \geq \tau - \mathfrak{T}\left(\psi_z(a, u), z'\right).$$

The right hand side of the inequality is decreasing in u, so that we can infer that every $u' > u$ is contained in $\mathfrak{S}(a, z; \tau)$, hence $u \in \overline{\mathfrak{S}(a, z; \tau)}$. In a similar way, we can show that $u \in [0, \tau] \setminus \overline{\mathfrak{S}(a_n, z; \tau)}$ for every $n \geq N$ implies that $u \in [0, \tau] \setminus \mathfrak{S}(a, z; \tau)$. However, this is only possible if $\mathfrak{s}(a_n, z; \tau) \to \mathfrak{s}(a, z; \tau)$, proving continuity in the interior of the domain.

It is obvious that $\tau > \mathfrak{s}(a, z; \tau)$ as $\tau \in \mathfrak{S}(a, z; \tau)$ and $\mathfrak{S}(a, z; \tau)$ is half-open. The uniformity is also clear. ■

Lemma 4.8 *The map* $u \mapsto \phi_z(a, u; \tau)$ *is differentiable on* $\mathfrak{S}(a, z; \tau)$ *and we have*

$$\left|\frac{\partial}{\partial u}\phi_z(a, u; \tau)\right| > 0.$$

Proof. By (10), ϕ_z is differentiable in u provided that $a' \mapsto \psi_{z'}(a', \tau - u)$ is differentiable at $a' = \psi_z(a, u)$. It is a well-known fact that the solution map of an ODE is differentiable in its initial value provided that the right hand side is C^1. By Lemma 4.5, the right hand side of (9) (for $z = z'$) is C^1 (in a) as long as we do not hit $a^*_{z'}$, which is precisely guaranteed by $u \in \mathfrak{S}(a, z; \tau)$. Hence, we can apply the chain rule and obtain

$$\begin{aligned}\frac{\partial}{\partial u}\phi_z(a, u; \tau) &= -\frac{\partial \psi_{z'}}{\partial u}\left(\psi_z(a, u), \tau - u\right) + \frac{\partial \psi_{z'}}{\partial a}\left(\psi_z(a, u), \tau - u\right)\frac{\partial \psi_z}{\partial u}(a, u) \\ &= -\underbrace{[r\phi_z(a, u; \tau) + z' - c(\phi_z(a, u, \tau), z')]}_{I} + \\ &\quad + \underbrace{\frac{\partial \psi_{z'}}{\partial a}\left(\psi_z(a, u), \tau - u\right)}_{II}\underbrace{[r\psi_z(a, u) + z - c(\psi_z(a, u), z)]}_{III}.\end{aligned}$$

For $z = w$, we have $I < 0$ (with strict inequality as $u \in \mathfrak{S}(a, z; \tau)$), and $II \geq 0$, $III \geq 0$, implying that

$$\frac{\partial}{\partial u}\phi_w(a, u; \tau) > 0.$$

On the other hand, for $z = b$, we have $I > 0$ (again, with strict inequality), $II \geq 0$ and $III \leq 0$, implying that

$$\frac{\partial}{\partial u}\phi_b(a, u; \tau) < 0.$$

■

By Lemma 4.8 together with Lemma 4.7 we now understand rigorously on which domains of integration we can do the change of variables in (8), which is crucial for establishing continuity. Therefore, we are now prepared to finish the proof of the theorem.

Proof of th. 4.4. We choose the measure $\nu(dt) = \delta_\tau(dt)$ for some fixed $\tau > 0$ and define a candidate $\widetilde{T}$ for a continuous component of P^τ by

$$\widetilde{T}((a, z), A) \equiv \int_0^\tau \mathbf{1}_A(\phi_z(a, u; \tau), z')\mathbf{1}_{\mathfrak{S}(a,z;\tau)}(\phi_z(a, u; \tau))g_\tau^{(1)}(u)du P(N(\tau) = 1), \qquad (11)$$

for $a \in [a^*_b, a^*_w]$, $z \in \{w, b\}$, $A \subset [a^*_b, a^*_w] \times \{w, b\}$ measurable, i.e.,

$$\widetilde{T}((a, z), A) = P^\tau((a, z), A \cap \{N_\tau = 1\} \cap \{T_1 \in \mathfrak{S}(a, z; \tau)\}),$$

where T_1 denotes the first jump time of the Poisson process N. Hence, it is clear that $\widetilde{T} \leq P^\tau$. Now, introduce a change or variables $u \to y \equiv \phi_z(a, y; \tau)$ as in (8). By Lemma 4.8, we get

$$\begin{aligned}\widetilde{T}((a, z), A) = \int_{L(a,z;\tau)}^{U(a,z;\tau)} &\mathbf{1}_A(y, z')\mathbf{1}_{\mathfrak{S}(a,z;\tau)}\left(\phi_z^{-1}(a, y; \tau)\right) \times \cdots \\ &\cdots \times g_\tau^{(1)}\left(\phi_z^{-1}(a, y; \tau)\right)\left|\frac{\partial}{\partial y}\phi_z^{-1}(a, y; \tau)\right| dy, \quad (12)\end{aligned}$$

where the lower and upper limits of the integration are given by

$$L(a, z; \tau) \equiv \begin{cases}\phi_z(a, 0; \tau), & z = w, \\ \phi_z(a, \tau; \tau), & z = b,\end{cases} \qquad U(a, z; \tau) \equiv \begin{cases}\phi_z(a, \tau; \tau), & z = w, \\ \phi_z(a, 0; \tau), & z = b,\end{cases}$$

respectively. Here, $y \mapsto \phi_z^{-1}(a, y; \tau)$ denotes the inverse function of $u \mapsto \phi_z(a, u; \tau)$. Comparing (12) with (11), we note two important differences: the integrand (including the limits of the integration) in (12) is continuous in a almost everywhere but, on the other hand, generally unbounded.

By a slight abuse of notation, let us denote $\mathfrak{S}(a,z;\tau) \equiv]\mathfrak{s}(a,z;\tau),\tau]$.[16] Lemma 4.7 implies that we may choose $0 < \epsilon < \inf_{(a,z)}(\tau - \mathfrak{s}(a,z;\tau))$. Now define $\mathfrak{S}_\epsilon(a,z;\tau) \equiv]\mathfrak{s}(a,z;\tau)+\epsilon,\tau]$ and

$$T((a,z),A) \equiv \int_0^\tau \mathbf{1}_A(\phi_z(a,u;\tau),z')\mathbf{1}_{\mathfrak{S}_\epsilon(a,z;\tau)}(\phi_z(a,u;\tau))g_\tau^{(1)}(u)duP(N(\tau)=1). \quad (13)$$

By the same change of variables as above, we arrive at

$$T((a,z),A) = \int_{L(a,z;\tau)}^{U(a,z;\tau)} \mathbf{1}_A(y,z')\mathbf{1}_{\mathfrak{S}_\epsilon(a,z;\tau)}\left(\phi_z^{-1}(a,y;\tau)\right) \times \cdots$$
$$\cdots \times g_\tau^{(1)}\left(\phi_z^{-1}(a,y;\tau)\right)\left|\frac{\partial}{\partial y}\phi_z^{-1}(a,y;\tau)\right| dy. \quad (14)$$

Since the term I in the proof of Lemma 4.8 only gets close to 0 when u is close to $\mathfrak{s}(a,z;\tau)$, now

$$\mathbf{1}_{\mathfrak{S}_\epsilon(a,z;\tau)}\left(\phi_z^{-1}(a,y;\tau)\right)\left|\frac{\partial}{\partial y}\phi_z^{-1}(a,y;\tau)\right|$$

is uniformly bounded, implying that $(a,z) \mapsto T((a,z),A)$ is continuous for any measurable set A.

As, by construction, $\tau - (\mathfrak{s}(a_0,z_0;\tau)+\epsilon) > 0$ we have $T((a,z),[a_b^*,a_w^*]\times\{w,b\}) > 0$. Finally, it is obvious that $T((a,z),A) \le \widetilde{T}((a,z),A) \le P^\tau((a,z),A)$ for any (a,z) and any measurable function A. ■

We can now complete our proof of uniqueness by proving the remaining conditions of prop. A.9.

Theorem 4.9 *Suppose that $r < \rho$. Then there is a unique invariant probability measure for the wealth-employment process $(a(\tau), z(\tau))$.*

Proof. By lemma 4.2 and theorem 4.4, the employment-wealth process $(a(\tau), z(\tau))$ is an irreducible T-process. Thus, prop. A.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. A.7, there is a unique invariant measure (up to a constant multiplier), and prop. A.6, finally, implies that we may choose the invariant measure to be a probability measure. ■

4.3 Stability

Stability, i.e., convergence of the distribution of $(a(\tau), z(\tau))$ to the unique invariant distribution for any given initial distribution is implied by the existence of an irreducible skeleton chain, see prop. A.13.

Corollary 4.10 *Under the assumptions of theorem 4.9, the employment-wealth process is stable in the sense of def. A.10.*

Proof. Recall that the employment-wealth-process is a T-process, see theorem 4.4. Moreover, we have shown irreducibility in lemma 4.2. Proposition A.13 will imply the desired conclusion, if we can show irreducibility of a skeleton chain. Take any $\tau > 0$ and consider

[16] This means that $\mathfrak{s}(a,z;\tau) \equiv 0$ in the case $\mathfrak{T}(a,z') = \infty$ and $\mathfrak{S}(a,z;\tau) = [0,\tau]$ is replaced by $]0,\tau]$ in that case.

the corresponding skeleton Y_n, $n \in \mathbb{N}$, with transition probabilities P^τ. By the proof of theorem 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

$$\forall y \in \mathbf{X}: \quad \sum_{n=1}^{\infty} P^{n\tau}(y, O) > 0. \tag{15}$$

This property, however, can be easily shown for the wealth-employment process (a, z) as illustrated in fig. 1 and formally analysed in app. C and D. 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 $\sum_{n=1}^{\infty} P^{n\tau}(y, O) > 0$. Thus, the τ-skeleton chain is irreducible and the wealth-employment process is stable. ∎

5 Describing the distribution of labour income and wealth

We now come to the applied part of this paper where we describe distributional properties of $z(\tau)$ and $a(\tau)$ by Fokker-Planck equations. This is of importance per se for our setup and serves as an example that can be adapted for many other applications.

5.1 Labour market probabilities

Consider first the distribution of the labour market state. Given that the transition rates between w and b are constant, the conditional probabilities of being in state $z(\tau)$ follow e.g. from solving Kolmogorov's backward equations as presented e.g. in Ross (1993, ch. 6). As an example, the probability of being employed in $\tau \geq t$ conditional on being in state $z \in \{w, b\}$ in t are

$$P(z(\tau) = w | z(t) = w) \equiv p_{ww}(\tau) = \frac{\mu}{\mu+s} + \frac{s}{\mu+s} e^{-(\mu+s)(\tau-t)}, \tag{16}$$

$$P(z(\tau) = w | z(t) = b) \equiv p_{bw}(\tau) = \frac{\mu}{\mu+s} - \frac{\mu}{\mu+s} e^{-(\mu+s)(\tau-t)}. \tag{17}$$

The complementary probabilities are $p_{wb}(\tau) = 1 - p_{ww}(\tau)$ and $p_{bb}(\tau) = 1 - p_{bw}(\tau)$. Letting $p_w(t)$ denote the probability of $z(t) = w$, i.e. letting it describe the initial distribution of $z(t)$, the unconditional probability of being in state z in τ is

$$p_z(\tau) = p_w(t) p_{wz}(\tau) + (1 - p_w(t)) p_{bz}(\tau). \tag{18}$$

Equations (16) and (17) nicely show the influence of the initial condition on the probability of having a job. Consider a point in time τ which is just an instant after t. Let this instant be so small that τ is basically identical to t. Then, the probability of being employed in τ (where $\tau = t$) is given by $\frac{\mu}{\mu+s} + \frac{s}{\mu+s} = 1$. Similarly, the probability of being unemployed in τ where τ is very close to t is given by (set $\tau = t$ in (17)) $\frac{\mu}{\mu+s} - \frac{\mu}{\mu+s} = 0$. The longer the point τ lies into the future, the less important the initial state becomes and the closer both probabilities approach the unconditional probability of being employed, which is $\frac{\mu}{\mu+s}$.

5.2 Fokker-Planck equations for wealth

5.2.1 The question and how to answer it

Our individual faces an uncertain future labour income stream $z(\tau)$. We would like to understand the joint distribution of $a(\tau)$ and $z(\tau)$ for $\tau \geq t$. To this end, we consider the stochastic processes of $a(\tau)$ in (2) and $z(\tau)$ in (3). After defining the (joint) density of $(a(\tau), z(\tau))$, we apply the "Fokker-Planck machinery" to obtain a description of the densities.

We denote the joint density by $p(a, z, \tau)$. For each point in time τ, there is obviously a discrete and a continuous random variable. We can therefore split the density into two "subdensities" $p(a, w, \tau)$ and $p(a, b, \tau)$, both drawn in fig. 2 for some $\tau \geq t$. The subdensities can be understood as the product of a conditional density $p(a, \tau|z)$ times the probability of being in employment state z,

$$p(a, z, \tau) \equiv p(a, \tau|z)\, p_z(\tau). \tag{19}$$

The probability $p_z(\tau)$ of an individual to be in a state z in τ is given by (18). As is clear from (19), $p(a, z, \tau)$ are not conditional densities – they rather integrate to the probability of $z(\tau) = z$. Looking at an individual who is in state z in τ, we get

$$\int p(a, z, \tau)\, da = \int p(a, \tau|z)\, p_z(\tau)\, da = p_z(\tau) \int p(a, \tau|z)\, da = p_z(\tau). \tag{20}$$

The density of a at some point in time τ is then simply

$$p(a, \tau) = p(a, w, \tau) + p(a, b, \tau). \tag{21}$$

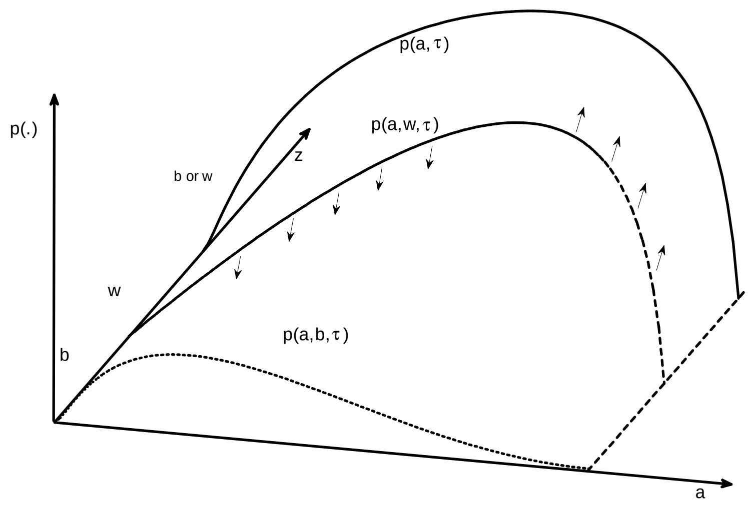

Figure 2 *The subdensities $p(a, b, \tau)$ and $p(a, w, \tau)$ and the density $p(a, \tau)$*

Note that the distribution of $(a(\tau), z(\tau))$ certainly depends on the initial condition $(a(t), z(t))$, which needs to be specified in order to calculate $p(a, z, \tau)$. In the notation we do not distinguish between the following two possibilities. Firstly, $(a(t), z(t))$ can be deterministic numbers, in which case $p(a, z, t)$ is a Dirac-distribution centered in $(a(t), z(t))$ (more precisely, the mapping $a \to p(z, a, t)$ is a Dirac-distribution). Secondly, $(a(t), z(t))$ can itself be random, either because we regard them as outcomes of the employment-wealth-process started at an even earlier time, or because there is some intrinsic uncertainty in measuring $a(t)$ (as e.g. the exact value of some asset, think e.g. of a house, is not known).

Let us now step back and ask how this approach can be applied to other setups. If one would like to understand the process of accumulation and depreciation of skills and experience during different employment states, one would have to specify a differential equation for skill similar to the budget constraint (2). Joint with the fundamental process (3) one could then derive Fokker-Planck equations for densities. If one would like to model the endogenous distribution of entitlement to unemployment benefits, one would have to "translate" regulations concerning entitlement into a differential equation, add again (3) and proceed to derive Fokker-Planck equations. Similar procedures are possible for analysing distributions over the business cycle where some aggregate shock process would be added to (2), (3) or both. Note that this approach works for processes driven e.g. by Brownian motion just as well.

5.2.2 The equations and their economic interpretation

The derivation of the Fokker-Planck equations is in app. B. The result is a system of two non-autonomous quasi-linear partial differential equations in $p(a,w,\tau)$ and $p(a,b,\tau)$,

$$\begin{aligned}&\frac{\partial}{\partial\tau}p(a,w,\tau)+\{ra+w-c(a,w)\}\frac{\partial}{\partial a}p(a,w,\tau)=\\&-\left\{r-\frac{\partial}{\partial a}c(a,w)+s\right\}p(a,w,\tau)+\mu p(a,b,\tau),\end{aligned}\tag{22a}$$

$$\begin{aligned}&\frac{\partial}{\partial\tau}p(a,b,\tau)+\{ra+b-c(a,b)\}\frac{\partial}{\partial a}p(a,b,\tau)=\\&sp(a,w,\tau)-\left\{r-\frac{\partial}{\partial a}c(a,b)+\mu\right\}p(a,b,\tau).\end{aligned}\tag{22b}$$

The system is a *partial* differential equation system as there are two derivatives, one with respect to time τ and one with respect to wealth a – which is not surprising: As the FPEs describe the evolution of the density for wealth over time, two derivatives are needed. The derivative with respect to a describes the "cross-sectional" property of the density for a given τ. The time derivative describes how a density changes over time.[17] The differential equations are called *quasi*-linear as the factors in front of the wealth-derivatives are functions of a. The PDEs are *non-autonomous* as some of the terms (other than the densities) also depend explicitly on one of the exogenous variables (exogenous in a differential equation sense), i.e. on wealth a.

As we can see, the density depends on properties of optimizing behaviour through the consumption levels $c(a,w)$ and $c(a,b)$ and through the marginal propensities to consume out of wealth, $\partial c(a,w)/\partial a$. These FPEs therefore describe the evolution of wealth for any specification of the utility function (e.g. CRRA, CARA, log, etc.). Modifying the utility function (e.g. allowing for labour supply or separating the intertemporal elasticity of substitution from risk aversion) affects the density of wealth through the effect on the optimal consumption plan $c(a,z)$.

Before we give an economic interpretation to these equations, we transform them such that they do not describe densities but distribution functions. To this end, define "subdistribution" functions as

$$P(a,z,\tau)\equiv\int_{-b/r}^{a}p(a,z,\tau)\,da.\tag{23}$$

[17] Compare this to the Pearson system of distributions that describes densities by ordinary non-autonomous differential equations (see e.g. Johnson, Kotz and Balakrishnan, 1994, ch. 12). These ordinary differential equations describe the density of one random variable. Here, we analyse a stochastic process, i.e. a sequence of random variables, and therefore need two derivatives.

The term $P(a,w,\tau)$ gives the probability that an individual will be employed in τ and own wealth equal or lower to a. Given our definition of subdensities and their property in (20), we know that $\lim_{a\rightarrow\infty}P(a,w,\tau)=p_{zw}(\tau)$ where the term $p_{zw}(\tau)$ is given in either (16) or (17), depending on the initial state in t.

The transformation of our FPEs is subject to the condition that $p\left(-\frac{b}{r},z,\tau\right)=0$ for all τ. This means that there is no worker with wealth equal to $-b/r$. As a wealth of $-b/r$ for unemployed workers would imply zero consumption, $c(-b/r,b)=0$, this can be ruled out indeed as marginal utility from consumption would then be infinity. This would violate optimality. As employed workers with wealth of $-b/r$ can only originate from unemployed workers with this wealth level (as wealth of employed workers increases) and as $p\left(-\frac{b}{r},b,\tau\right)=0$ for all τ, we know that $p\left(-\frac{b}{r},w,\tau\right)=0$ for all τ as well.

The subdistribution functions in (23) obey the following system (cf. app. E.2)

$$\frac{\partial}{\partial\tau}P(a,w,\tau)=-\{ra+w-c(a,w)\}\frac{\partial}{\partial a}P(a,w,\tau)-sP(a,w,\tau)+\mu P(a,b,\tau), \qquad (24a)$$

$$\frac{\partial}{\partial\tau}P(a,b,\tau)=-\{ra+b-c(a,b)\}\frac{\partial}{\partial a}P(a,b,\tau)+sP(a,w,\tau)-\mu P(a,b,\tau). \qquad (24b)$$

This system is now extremely easy to understand: Starting with the first equation, the evolution of the distribution function over time, i.e. the time derivative $\partial P(a,w,\tau)/\partial\tau$ on the left hand side depends on three terms. Starting at the end, there is an increase in the probability $P(a,w,\tau)$ if there is a high flow from the state of being unemployed. This flow can be high if the matching rate μ, the probability of being unemployed $P(a,b,\tau)$ or if a combination of the two is high. Similarly, the probability $P(a,w,\tau)$ decreases (ceteris paribus) exponentially at the rate s, and the faster so, the higher the separation rate. The interpretation of the last two terms in the second equation (24b) is identical (subject to reversed signs). These two terms are very familiar from derivations of wage distributions in the Burdett-Mortensen (1998) tradition.

We can think of these equations as describing how wealth of a worker flows up and down depending on her current state. The labor income levels of workers are stochastically moving back and forth between the different states w and b. The effect of these stochastic jumps on the distribution of wealth are captured by the two terms at the end of (24a,b). Wealth is moving non-stochastically within the states, either upwards (when employed) or downwards (when unemployed). The direction of the movement is on the wealth line, i.e. the partial derivative $\partial P(a,w,\tau)/\partial a$ gives the direction of a. The speed of this movement is determined by savings $ra+z-c(a,z)$. The speed is positive when employed and negative when unemployed. The overall effect of positive savings for the probability $P(a,w,\tau)$ of employed workers is then to *decrease* this probability. As wealth increases, the probability of having a wealth level equal to or lower than a certain level a obviously falls as there is a permanent flow towards higher wealth levels. This flow is then reversed in the state of unemployment where the speed (i.e. savings $ra-b-c(a,b)$) is negative. As a consequence, the probability $P(a,b,\tau)$ ceteris paribus increases over time as unemployed workers "gather" towards the lower end of the wealth distribution.

5.2.3 Initial conditions

Obtaining a unique solution for ODEs generally requires certain differentiability conditions and as many initial conditions as differential equations. Conditions for obtaining a unique solution for PDEs differ in various respects, of which the most important one from an intuitive perspective is the fact that instead of initial conditions (i.e. an initial value or vector), initial functions are required. This can easily be understood for our case: Let us assume two initial functions for a, one for each labour market state $z\in\{w,b\}$. The obvious interpretation for these initial functions are densities, just as illustrated in fig. 2. Initial functions would therefore be given by $p(a,b,t)=p^{ini}(a,b)$ and $p(a,w,t)=p^{ini}(a,w)$. Clearly, they take positive values on the range $[-b/r, a_w^*]$ only and need to jointly integrate to unity. Given these initial functions, one can then compute the partial derivatives with respect to a in (22). This gives an ODE system which allows us to compute the density for the "next" τ. Repeating this gives us the densities for all z, a and τ we are interested in.

5.2.4 A density gives a density

The Fokker-Planck equations have a very convenient property that easily allows to show that they indeed describe densities (in the sense that their solutions integrate to one). The only condition is that the initial functions integrate to one. We summarize this in the following

Proposition 5.1 *Define* $I(\tau) \equiv \int_{-\infty}^{\infty} p(a, w, \tau) + p(a, b, \tau)\, da$. *Given the laws of motion for* $p(a, z, \tau)$ *from (22) and the fact of a bounded support* $[-b/r, a_w^*]$, *this integral is mass-preserving, i.e.* $dI(\tau)/d\tau = 0$ *for all* τ. *Assuming initial densities, i.e. initial functions* $p(a, z, t) \geq 0$ *such that* $I(t) = 1$, *the PDEs in (22) indeed describe the dynamics of distributions over time.*

Proof. see app. E.1 ■

This is an extremely useful property as this implies that with an initial density we know that all other functions $p(a, w, \tau) + p(a, b, \tau)$ integrate to one and therefore represent densities.

5.2.5 The long-run distribution of individual wealth

When we are interested in the long-run distribution of wealth and income only, the time derivatives of the densities would be zero and the long-run densities would be described by two linear ordinary differential equations. This is true both for the system in densities (22) and for the system for distributions (24).

Initial conditions for this ordinary differential equations are given by

$$p(a_w^*, w) = 0, \quad p(a_w^*, b) = 0. \tag{25}$$

The intuition for $p(a_w^*, w) = 0$ comes from the saddle-path nature of the TSS Θ in (7): There is one path going into Θ from the left and one going into Θ from the right and two (not drawn) starting from Θ and going North and South. In saddle-points of ODE systems, one can prove by linearization around the fix point that local solutions of the ODE approach the saddle point asymptotically. Linearization here is more involved given the special structure of our system (see fn. 22). Assuming that the qualitative properties of local behaviour are not affected by this structure, we would observe asymptotic behaviour here as well and the TSS Θ would actually never be reached: $p(a_w^*, w) = 0$ would follow. The second boundary condition is then an immediate consequence. As the state (a_w^*, b) can occur only through a transition from (a_w^*, w) but the density at (a_w^*, w) is zero, $p(a_w^*, b) = 0$ as well.

6 Conclusion

This paper has introduced methods that allow to prove existence, uniqueness and stability of distributions described by stochastic differential equations driven by a jump process. These methods were applied to a model of precautionary saving. Existence, uniqueness and stability of the optimal process for the state variables, wealth and labour market status, were proven. The results hold for an interest rate being lower than the time-preference rate.

The T-property turned out to be 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.

From a more applied perspective, we derived Fokker-Planck equations for wealth and labour market status. We saw inter alia how matching and separation rates and savings shape the evolution of the wealth distribution over time. Our approach and our derivation provides a considerable generalization to existing applications in economics. This will facilitate the use these equations in many other applications in future work.

A Appendix on ergodicity results for continuous time Markov processes

The wealth-employment process $(a(\tau), z(\tau))$ described by (2) and (3) 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 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.

A.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.

$$P^t(x, A) \equiv P(X_t \in A | X_0 = x) \equiv P_x(X_t \in A), \tag{26}$$

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.

$$P_t f(x) \equiv E(f(X_t) | X_0 = x) = \int_{\mathbf{X}} f(y) P^t(x, dy) \tag{27}$$

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 A.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 A.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 A.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^t_μ by

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

We say that μ is an *invariant measure*, iff $P^t_\mu = \mu$ 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.

A.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 A.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*

$$\liminf_{t\to\infty} \frac{1}{t} \int_0^t P_x(X_s \in C)ds \geq 1 - \epsilon. \tag{28}$$

The second property is a continuity condition.

Definition A.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 (27) 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 A.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.*

A.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 A.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 A.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. A.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 A.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*[18].
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 $x \in \mathbf{X}$.*

[18] 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.

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 A.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. A.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.

A.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^τ_μ 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.

Definition A.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 A.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*

$$\forall x \in C,\ \forall A \in \mathcal{B}(\mathbf{X}) : \quad P^{\tau}(x, A) \geq \nu(A);$$[19]

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 A.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. A.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 A.13 *Given a Harris recurrent Markov process X with invariant probability measure π. Then X is stable iff there is some irreducible skeleton chain.*

B Appendix on deriving the Fokker-Planck equations

This appendix derives the Fokker-Planck equations (22) of the wealth-employment process $(a(t), z(t))$. We proceed step by step as this facilitates applications for other purposes. Step 1: We start with some function f having as arguments the variables whose density we would like to understand. We compute the differential of this function in the usual way and also compute its expected change. Step 2: The starting point here is Dynkin's formula. This formula, intuitively speaking, gives the expected value of some function f, whose arguments are the random variables we are interested in, as the sum of the current value of f plus the integral over expected future changes of f. The expected change of f is expressed by using the density of our random variables. The Dynkin formula is differentiated with respect to time. Step 3: By using integration by parts or the adjoint operator, we get an expression for the change of the expected value of f. Step 4: A different expression for this change of the expected value can be obtained by starting from the expected value and differentiating it. Step 5: Equating the two gives the differential equations for the density.

It should be kept in mind that this approach can be applied to systems beyond (2) and (3). As long as there are one to several stochastic processes described by stochastic differential equations, this approach can be used to obtain a description of the corresponding densities. Uncertainty can stem from Brownian motion, Poisson processes, a combination of the two or Levy processes.

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

B.1 The expected change of some function f

Assume there is a function f having as arguments the state variables a and z. This function has a bounded support S, i.e. $f(a,z)=0$ outside this support.[20] Heuristically, the differential of this function, using a change of variable formula,[21] gives

$$\begin{aligned} df\left(a\left(\tau\right),z\left(\tau\right)\right) &= f_a\left(.\right)\left\{ra\left(\tau\right)+z\left(\tau\right)-c\left(a\left(\tau\right),z\left(\tau\right)\right)\right\}d\tau \\ &+\left\{f\left(a\left(\tau\right),z\left(\tau\right)+\Delta\right)-f\left(a\left(\tau\right),z\left(\tau\right)\right)\right\}dq_\mu \\ &+\left\{f\left(a\left(\tau\right),z\left(\tau\right)-\Delta\right)-f\left(a\left(\tau\right),z\left(\tau\right)\right)\right\}dq_s. \end{aligned}$$

Due to the state-dependent arrival rates, see after (3), only one Poisson process is active at a time.

When we are interested in the expected change, we need to form expectations. Applying the conditional expectations operator E_τ and dividing by $d\tau$ yields the heuristic equation

$$\begin{aligned} \frac{E_\tau df\left(.\right)}{d\tau} &= f_a\left(.\right)\left\{ra\left(\tau\right)+z\left(\tau\right)-c\left(a\left(\tau\right),z\left(\tau\right)\right)\right\} \\ &+\mu\left(z\left(\tau\right)\right)\left[f\left(a\left(\tau\right),z\left(\tau\right)+\Delta\right)-f\left(a\left(\tau\right),z\left(\tau\right)\right)\right] \\ &+s\left(z\left(\tau\right)\right)\left[f\left(a\left(\tau\right),z\left(\tau\right)-\Delta\right)-f\left(a\left(\tau\right),z\left(\tau\right)\right)\right] \end{aligned} \tag{29}$$

In what follows, we denote this expression by

$$\mathcal{A}f\left(a\left(\tau\right),z\left(\tau\right)\right)\equiv\frac{E_\tau df\left(a\left(\tau\right),z\left(\tau\right)\right)}{d\tau} \tag{30}$$

which is, more precisely, the infinitesimal generator $\mathcal{A}$ defined by

$$\mathcal{A}f(a,z)=\lim_{\epsilon\searrow 0}\frac{E\left(f(z(\tau+\epsilon),a(\tau+\epsilon))|z(\tau)=z,\ a(\tau)=a\right)-f(a,z)}{\epsilon}.$$

Notice that $\mathcal{A}f(a,z)$ does not depend on τ, because the Markov-process $(a(\tau),z(\tau))$ is time-homogeneous. We understand $\mathcal{A}$ as an operator mapping functions (in a and z) to other such functions. Moreover, note that all test-functions, i.e. C^∞ functions of bounded support, are in the domain of the operator $\mathcal{A}$, i.e. the domain of all functions f such that the above limit exists (for all a and z).

B.2 Dynkin's formula and its manipulation

To abbreviate notation, we now define $x\left(\tau\right)\equiv\left(a\left(\tau\right),z\left(\tau\right)\right)$. The expected value of our function $f\left(x\left(\tau\right)\right)$ is by Dynkin's formula (e.g. Yuan and Mao, 2003) given by

$$Ef\left(x\left(\tau\right)\right)=Ef\left(x\left(t\right)\right)+\int_t^\tau E\left(\mathcal{A}f\left(x\left(s\right)\right)\right)ds. \tag{31}$$

To understand this equation, use the definition in (30) and formally write it as

$$Ef\left(x\left(\tau\right)\right)=Ef\left(x\left(t\right)\right)+\int_t^\tau\frac{Edf\left(x\left(s\right)\right)}{ds}ds=Ef\left(x\left(t\right)\right)+\int_t^\tau Edf\left(x\left(s\right)\right).$$

[20] We can make this assumption without any restriction. As we will see below, this function will not play any role in the determination of the actual density.

[21] There are formal derivations of this equation in mathematical textbooks like Protter (1995). For a more elementary presentation, see Wälde (2012, part IV).

Intuitively speaking, Dynkin's formula says that the expected value of $f(x(\tau))$ is the expectation for the current value, $Ef(x(t))$ (given that we allow for a random initial condition $x(t)$), plus the "sum of" expected future changes, $\int_t^\tau Edf(x(s))$.

Let us now differentiate (31) with respect to time τ and find

$$\frac{\partial}{\partial \tau} Ef(x(\tau)) = \frac{\partial}{\partial \tau} \int_t^\tau E(\mathcal{A}f(x(s)))\, ds = E(\mathcal{A}f(x(\tau))), \tag{32}$$

where the first equality used that $Ef(x(t))$ is a constant and pulled the expectations operator into the integral. This equation says the following: We form expectations in t about $f(x(\tau))$. We now ask how this expectation changes when τ moves further into the future, i.e. we look at $\frac{\partial}{\partial \tau} E[f(x(\tau))]$. We see that this change is given by the expected change of $f(x(\tau))$, where the change is $\mathcal{A}f(x(\tau))$.

We now introduce the densities we defined in sect. 5.2.1. The expectation operator E in (32) integrates over all possible states of $x(\tau)$. When we express this joint density as $p(a, z, \tau) \equiv p(a, \tau|z)\, p_z(\tau)$, we can write (32) as

$$\begin{aligned} \frac{\partial}{\partial \tau} Ef(x(\tau)) &= E(\mathcal{A}f(x(\tau))) \\ &= p_w(\tau) \int_{-\infty}^{\infty} \mathcal{A}f(a, w)\, p(a, \tau|w)\, da + p_b(\tau) \int_{-\infty}^{\infty} \mathcal{A}f(a, b)\, p(a, \tau|b)\, da. \end{aligned}$$

Now pull $p_w(\tau)$ and $p_b(\tau)$ back into the integral and use $p(a, z, \tau) \equiv p(a, \tau|z)\, p_z(\tau)$ again for $z = w$ and $z = b$. Then

$$\begin{aligned} \frac{\partial}{\partial \tau} Ef(x(\tau)) &= \int_{-\infty}^{\infty} \mathcal{A}f(a, w)\, p(a, w, \tau)\, da + \int_{-\infty}^{\infty} \mathcal{A}f(a, b)\, p(a, b, \tau)\, da \\ &\equiv \phi_w + \phi_b. \end{aligned} \tag{33}$$

B.3 The adjoint operator and integration by parts

This is now the crucial step in obtaining a differential equation for the density. It consists in applying an integration by parts formula which allows to move the derivatives in $\mathcal{A}f(x(\tau))$ into the density $p(x, \tau)$. Let us briefly review this method, without getting into technical details. Given two functions $f, g : \mathbb{R} \to \mathbb{R}$ and two fixed real numbers $c < d$, the factor rule of differentiation

$$d(f(x) \cdot g(x)) = df(x) \cdot g(x) + f(x) \cdot dg(x) \tag{34}$$

implies that $f(d)g(d) - f(c)g(c) = \int_c^d f'(x)g(x)dx + \int_c^d f(x)g'(x)dx$, a formula referred to as partial integration rule. In particular, it also holds for $c = -\infty$ and $d = +\infty$, if the function evaluations are understood as limits for $c \to -\infty$ and $d \to +\infty$, respectively. If f has bounded support, i.e. is equal to zero outside a fixed bounded set, then the function evaluations at $\pm\infty$ vanish and we get

$$\int_{-\infty}^{+\infty} f'(x)g(x)dx = -\int_{-\infty}^{+\infty} f(x)g'(x)dx. \tag{35}$$

We now apply (35) to equation (33). We can do this as the expressions in (33) "lost" all stochastic features. To this end, insert the definition of $\mathcal{A}$ given in (30) together with (29) into (33). To avoid getting lost in long expressions, we look at the both integrals in (33) in turn. For the second, observe that

$$\mathcal{A}f(a, b) = f_a(.)\{ra + b - c(a, b)\} + \mu[f(a, w) - f(a, b)],$$

i.e. the term with s in (29) is missing given that we are in state b. Hence,

$$\begin{aligned}\phi_b &= \int_{-\infty}^{\infty} \left[f_a(a,b)\{ra+b-c(a,b)\} + \mu\left[f(a,w)-f(a,b)\right]\right] p(a,b,\tau)\,da \\ &= \int_{-\infty}^{\infty} f_a(a,b)\{ra+b-c(a,b)\}\,p(a,b,\tau)\,da \\ &+ \int_{-\infty}^{\infty} \mu\left[f(a,w)-f(a,b)\right] p(a,b,\tau)\,da.\end{aligned}$$

Now integrate by parts. As this integral shows, we only need to integrate by parts for the f_a term. The rest remains untouched. This gives with (35), where $g(x)$ stands for $\{ra+b-c(a,b)\}\,p(a,b,\tau)$ and x for a,

$$\begin{aligned}\phi_b = &-\int_{-\infty}^{\infty} f(a,b)\left[\left\{r-\frac{\partial}{\partial a}c(a,b)\right\}p(a,b,\tau) + \{ra+b-c(a,b)\}\frac{\partial}{\partial a}p(a,b,\tau)\right]da \\ &+ \int_{-\infty}^{\infty} \mu\left[f(a,w)-f(a,b)\right]p(a,b,\tau)\,da. \end{aligned} \tag{36}$$

Now look at the first integral of (33). After similar steps (as the principle is the same, we replace b by w and the arrival rate μ by s in the last equation), this reads

$$\begin{aligned}\phi_w = &-\int_{-\infty}^{\infty} f(a,w)\left[\left\{r-\frac{\partial}{\partial a}c(a,w)\right\}p(a,w,\tau) + \{ra+w-c(a,w)\}\frac{\partial}{\partial a}p(a,w,\tau)\right]da \\ &+ \int_{-\infty}^{\infty} s\left[f(a,b)-f(a,w)\right]p(a,w,\tau)\,da. \end{aligned} \tag{37}$$

Summarizing, we find

$$\frac{\partial}{\partial \tau}Ef(x(\tau)) = \phi_w + \phi_b$$

$$\begin{aligned}&= \int_{-\infty}^{\infty} f(a,w)\left[-\left\{r-\frac{\partial}{\partial a}c(a,w)\right\}p(a,w,\tau) - \{ra+w-c(a,w)\}\frac{\partial}{\partial a}p(a,w,\tau)\right]da \\ &+ \int_{-\infty}^{\infty} s\left[f(a,b)-f(a,w)\right]p(a,w,\tau)\,da \\ &+ \int_{-\infty}^{\infty} f(a,b)\left[-\left\{r-\frac{\partial}{\partial a}c(a,b)\right\}p(a,b,\tau) - \{ra+b-c(a,b)\}\frac{\partial}{\partial a}p(a,b,\tau)\right]da \\ &+ \int_{-\infty}^{\infty} \mu\left[f(a,w)-f(a,b)\right]p(a,b,\tau)\,da. \end{aligned} \tag{38}$$

B.4 The expected value again

Let us now derive the second expression for the change in the expected value. By definition, and as an alternative to the Dynkin formula (31), we have

$$Ef(x(\tau)) = \int_{-\infty}^{\infty} f(a,b)\,p(a,b,\tau)\,da + \int_{-\infty}^{\infty} f(a,w)\,p(a,w,\tau)\,da. \tag{39}$$

When we differentiate this expression with respect to time, we get

$$\begin{aligned}\frac{\partial}{\partial \tau}Ef(x(\tau)) = &\int_{-\infty}^{\infty} f(a,b)\frac{\partial}{\partial \tau}p(a,b,\tau)\,da \\ &+ \int_{-\infty}^{\infty} f(a,w)\frac{\partial}{\partial \tau}p(a,w,\tau)\,da. \end{aligned} \tag{40}$$

Note that we can use

$$\frac{\partial}{\partial \tau} \int_{-\infty}^{\infty} f(a, z)\, p(a, z, \tau)\, da = \int_{-\infty}^{\infty} f(a, z)\, \frac{\partial}{\partial \tau} p(a, z, \tau)\, da$$

as z and a inside this integral are no longer functions of time.

B.5 Equating the two expressions

We now equate (38) with (40). Collecting terms belonging to $f(a, w)$ and $f(a, b)$ gives

$$\int_{-\infty}^{\infty} f(a, w)\, \varphi_w da + \int_{-\infty}^{\infty} f(a, b)\, \varphi_b da = 0, \tag{41}$$

where

$$\varphi_w \equiv -\left\{ r - \frac{\partial}{\partial a} c(a, w) + s \right\} p(a, w, \tau) - \{ra + w - c(a, w)\} \frac{\partial}{\partial a} p(a, w, \tau)$$
$$+ \mu p(a, b, \tau) - \frac{\partial}{\partial \tau} p(a, w, \tau)$$

and

$$\varphi_b \equiv -\left\{ r - \frac{\partial}{\partial a} c(a, b) + \mu \right\} p(a, b, \tau) - \{ra + b - c(a, b)\} \frac{\partial}{\partial a} p(a, b, \tau)$$
$$+ sp(a, w, \tau) - \frac{\partial}{\partial \tau} p(a, b, \tau).$$

Obviously, the above equation is satisfied if

$$\varphi_b = \varphi_w = 0. \tag{42}$$

These are the Fokker-Planck equations used in (22).

It is easy to see that the integral equation can only be satisfied for all functions f if these Fokker-Planck equations are satisfied. Indeed, assume that $\varphi_b > 0$ on an interval $I = [d-\epsilon, d+\epsilon]$. One can find a non-negative function f smooth in a such that $f(a, w) = 0$ for all a and

$$f(a, b) = \begin{cases} 1, & a \in [d - \epsilon/2, d + \epsilon/2], \\ 0, & a \in]-\infty, d - \epsilon] \cup [d + \epsilon, \infty[. \end{cases}$$

Inserting this test function into the integral equation gives

$$\int_{-\infty}^{\infty} f(a, w)\, \varphi_w da + \int_{-\infty}^{\infty} f(a, b)\, \varphi_b da = 0 + \int_{d-\epsilon}^{d+\epsilon} f(a, b)\, \varphi_b da > 0$$

by construction. Therefore, $\varphi_b = 0$ has to hold for all $a \in \mathbb{R}$, and similarly for φ_w.

C Referee appendix

For all further appendices, please see the Referees' appendix

References

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," Quarterly Journal of Economics, 109, 659–84.

Anderson, R., and R. Raimondo (2008): "Equilibrium in Continous-Time Financial Markets: Endogenously Dynamically Complete Markets," Econometrica, 76(4), 841–907.

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.

Bertola, G., and R. Caballero (1994): "Irreversibility and Aggregate Investment," Review of Economic Studies, 61(207), 223–246.

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.

Flinn, C. (2006): "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates," Econometrica, 74, 1013–1062.

Friedman, A. (1975): Stochastic differential equations and applications. Vol. 1. Academic Press [Harcourt Brace Jovanovich Publishers], New York, Probability and Mathematical Statistics, Vol. 28.

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

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.

Impullitti, G., A. A. Irarrazabal, and L. D. Opromolla (2011): "A Theory of Entry into and Exit from Export Markets," mimeo Cambridge University.

Johnson, N., S. Kotz, and N. Balakrishnan (1994): "Continuous Distributions (General)," in Continuous Univariate Distributions Vol. 1, ed. by N. Johnson, S. Kotz, and N. Balakrishnan, pp. 1–79. Wiley Publications.

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

Kamihigashi, T., and J. Stachurski (2012): "An order-theoretic mixing condition for monotone Markov chains," Statistics & Probability Letters, 82(2), 262–267.

——— (2013): "Stochastic Stabilty in Monotone Economies," Theoretical Economics, forthcoming.

Klette, T. J., and S. Kortum (2004): "Innovating Firms and Aggregate Innovation," Journal of Political Economy, 112(5), 986–1018.

Koeniger, W., and J. Prat (2007): "Employment protection, product market regulation and firm selection," Economic Journal, 117(521), F302–F332.

Krebs, T. (2006): "Recursive equilibrium in endogenous growth models with incomplete markets," Economic Theory, 29(3), 505–523.

Launov, A., and K. Wälde (2013): "Estimating Incentive and Welfare Effects of Non-Stationary Unemployment Benefits," International Economic Review, forthcoming.

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

Lise, J. (2013): "On-the-Job Search and Precautionary Savings," Review of Economic Studies, forthcoming.

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.

Mattheij, R., and J. Molenaar (2002): Ordinary differential equations in theory and practice, vol. 43 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, Reprint of the 1996 original.

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.

——— (2013): "Stochastic Search Equilibrium," Review of Economic Studies, forthcoming.

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

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.

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.

Protter, P. (1995): Stochastic Integration and Differential Equations. Springer-Verlag, Berlin.

Raimondo, R. C. (2005): "Market clearing, utility functions, and securities prices," Economic Theory, 25(2), 265–285.

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.

Ross, S. M. (1993): Introduction to Probability Models, 5th edition. Academic Press, San Diego.

Roy, S., and I. Zilcha (2012): "Stochastic growth with short-run prediction of shocks," Economic Theory, 51(3), 539–580.

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

Stokey, N. L. (2008): The Economics of Inaction: Stochastic Control Models with Fixed Costs. Princeton University Press.

van den Berg, G. J. (1990): "Nonstationarity in Job Search Theory," Review of Economic Studies, 57, 255–277.

Wälde, K. (1999): "Optimal Saving under Poisson Uncertainty," Journal of Economic Theory, 87, 194–217.

——— (2012): Applied Intertemporal Optimization. Know Thyself - Academic Publishers, available at www.waelde.com/KTAP.

Yuan, C., and X. Mao (2003): "Asymptotic stability in distribution of stochastic differential equations with Markovian switching," Stochastic Processes and Their Applications, 103, 277–291.

Referees' appendix to

The Dynamics of Distributions
in Continuous-Time Stochastic Models

by

Christian Bayer and Klaus Wälde

October 2013

C Consumption and wealth dynamics

This appendix provides preliminary results for deriving the generalized Keynes-Ramsey rules and which allow us to perform the phase diagram illustration and the subsequent existence proof for an optimal consumption path $c(a,z)$.

C.1 Derivation of the Keynes-Ramsey rules (4a) and (4c)

- Bellman equations and first-order conditions

We now let the individual maximize her objective function by choosing a path $\{c(\tau)\}$ of consumption subject to the budget constraint (2) and the equation for its employment status (3). Given that the state variables describing an individual are not only current labour income $z(t)$ but also current wealth $a(t)$, we define the value function as $V(a(t),z(t)) = \max_{\{c(\tau)\}} U(t)$ subject to (3) and (2). The Bellman equation for this problem reads (see Wälde, 2012, part IV)

$$\rho V(a(t),z(t)) = \max_{c(t)} \left\{ u(c(t)) + \frac{1}{dt} E_t dV(a(t),z(t)) \right\}. \tag{43}$$

Computing the differential $dV(a(t),z(t))$, taking the constraints (2) and (3) into account and forming expectations yields, suppressing the time argument t for simplicity,

$$\rho V(a,z) = \max_c \left\{ \begin{array}{c} u(c) + [ra + z - c] V_a(a,z) \\ +s(z)[V(a,b) - V(a,w)] + \mu(z)[V(a,w) - V(a,b)] \end{array} \right\}, \tag{44}$$

where V_a stands for the partial derivative of V with respect to a. Note that this Bellman equation holds for both the employment state and the unemployment state as the arrival rates are state dependent.

Given that the individual only needs to choose consumption, the only first-order condition equates marginal utility from consumption with the shadow price of wealth,

$$u'(c(a,z)) = V_a(a,z). \tag{45}$$

We know by the budget constraint (2) that one unit of consumption costs one unit of wealth. Hence, in the optimum, the instantaneous increase in utility due to marginally consuming more is identical to the present value increase in overall utility due to an additional unit of wealth.

- Evolution of the shadow price

Using the budget constraint (2) and the evolution of labour income (3), the differential of the shadow price of wealth reads

$$dV_a(a,z) = V_{aa}(a,z)\{ra+z-c\}\,dt \\ +[V_a(a,w)-V_a(a,b)]\,dq_\mu + [V_a(a,b)-V_a(a,w)]\,dq_s. \tag{46}$$

The maximized version of the Bellman equation (44) simply replaces the control variable c by its optimal value $c(a,z)$,

$$\rho V(a,z) = \left\{ \begin{array}{c} u(c(a,z)) + [ra+z-c(a,z)]\,V_a(a,z) \\ +s(z)[V(a,b)-V(a,w)] + \mu(z)[V(a,w)-V(a,b)] \end{array} \right\}. \tag{47}$$

Differentiating with respect to wealth yields, using the envelope theorem,

$$\rho V_a(a,z) = \left\{ \begin{array}{c} rV_a(a,z) + [ra+z-c(a,z)]\,V_{aa}(a,z) \\ +s(z)[V_a(a,b)-V_a(a,w)] + \mu(z)[V_a(a,w)-V_a(a,b)] \end{array} \right\}. \tag{48}$$

Rearranging yields

$$(\rho-r)V_a(a,z) - s(z)[V_a(a,b)-V_a(a,w)] - \mu(z)[V_a(a,w)-V_a(a,b)] \\ = [ra+z-c(a,z)]\,V_{aa}(a,z).$$

Inserting into (46) gives

$$dV_a(a,z) = \{(\rho-r)V_a(a,z) - s(z)[V_a(a,b)-V_a(a,w)] - \mu(z)[V_a(a,w)-V_a(a,b)]\}\,dt \\ +[V_a(a,w)-V_a(a,b)]\,dq_\mu + [V_a(a,b)-V_a(a,w)]\,dq_s.$$

- Inserting first-order condition

When we now replace the shadow price by marginal utility from the first-order condition (45), we get the Keynes-Ramsey rule for marginal utility,

$$du'(c(a,z)) = \left\{ \begin{array}{c} (\rho-r)u'(c(a,z)) - s(z)[u'(c(a,b)) - u'(c(a,w))] \\ -\mu(z)[u'(c(a,w)) - u'(c(a,b))] \end{array} \right\} dt \\ +[u'(c(a,w)) - u'(c(a,b))]\,dq_\mu + [u'(c(a,b)) - u'(c(a,w))]\,dq_s. \tag{49}$$

For an employed individual where $\mu(z)=0$ and $a=a_w$, this reads

$$du'(c(a_w,w)) = \{(\rho-r)u'(c(a_w,w)) - s[u'(c(a_w,b)) - u'(c(a_w,w))]\}\,dt \\ +[u'(c(a_w,b)) - u'(c(a_w,w))]\,dq_s.$$

Let $f(.)$ be the inverse function for u', i.e. $f(u')=c$ and apply the CVF to $f(u'(c(a_w,w)))$ This gives

$$df(u'(c(a_w,w))) = f'(u'(c(a_w,w)))\{(\rho-r)u'(c(a_w,w)) - s[u'(c(a_w,b)) - u'(c(a_w,w))]\}\,dt \\ +[f(u'(c(a_w,b))) - f(u'(c(a_w,w)))]\,dq_s.$$

As $f(u')=c$ and therefore $f'(u'(c(a_w,w))) = \frac{df(u'(c(a_w,w)))}{du'(c(a_w,w))} = \frac{dc(a_w,w)}{du'(c(a_w,w))} = \frac{1}{u''(c(a_w,w))}$, we get

$$dc(a_w,w) = \frac{1}{u''(c(a_w,w))}\{(\rho-r)u'(c(a_w,w)) - s[u'(c(a_w,b)) - u'(c(a_w,w))]\}\,dt \\ +[c(a_w,b) - c(a_w,w)]\,dq_s \Leftrightarrow$$

$$\frac{u''(c(a_w,w))}{u'(c(a_w,w))}dc(a_w,w) = \left\{\rho - r - s\left[\frac{u'(c(a_w,b))}{u'(c(a_w,w))} - 1\right]\right\}dt \\ +\frac{u''(c(a_w,w))}{u'(c(a_w,w))}[c(a_w,b) - c(a_w,w)]\,dq_s.$$

Using the instantaneous CRRA utility function (1), we get $\frac{u''(c(a_w,w))}{u'(c(a_w,w))} = \frac{-\sigma c(a_w,w)^{-\sigma-1}}{c(a_w,w)^{-\sigma}} = \frac{-\sigma}{c(a_w,w)}$ and therefore

$$-\sigma\frac{dc\left(a_w,w\right)}{c\left(a_w,w\right)} = \left\{\rho - r - s\left[\left(\frac{c\left(a_w,w\right)}{c\left(a_w,b\right)}\right)^{\sigma} - 1\right]\right\}dt - \sigma\left[\frac{c\left(a_w,b\right)}{c\left(a_w,w\right)} - 1\right]dq_s.$$

After dividing by $-\sigma$, we get (4a) in the main text. The derivation of (4c) also starts from (49) and steps are in perfect analogy.

C.2 Consumption growth and the interest rate

Our analysis focuses on paths $c(a,z)$ as depicted in fig. 1. In this figure, we implicitly considered solutions of our system in the set $Q = \{a \geq -b/r\} \cap \{c(a,w) \leq ra + w\} \cap \{c(a,b) \geq ra + b\} \cap \{c(a,b) \geq 0\} \cap \{c(a,w) \geq c(a,b)\}$. In words, wealth is at least as large as the maximum debt level b/r, consumption of the employed worker is below the zero-motion line for her wealth, consumption of the unemployed worker is above her zero-motion line for wealth, consumption of the unemployed worker is non-negative and consumption of employed workers always exceeds consumption of unemployed workers (see lem. C.12).

For the proofs we restrict this set in two ways. First, we consider the domain

$$Q_v = \{(a, c(a,w), c(a,b)) \in \mathbb{R}^3 | \, (a, c(a,w), c(a,b)) \in Q, c(a,w) \leq ra + w - v\}, \tag{50}$$

where v is the small positive constant, as an approximation to our "full" set Q. As $Q_0 = Q$, Q_v simply excludes the zero-motion line for wealth of the employed workers. We need to do this as the fraction on the right-hand side of our differential equation (5a) is not defined for the TSS.[22] As v is small, however, we can get arbitrarily close to this zero-motion line and Q_v approximates Q arbitrarily well.

Second, we consider

$$R_{v,\Psi} = \big\{(a, c(a,w), c(a,b)) \in \mathbb{R}^3 | \, (a, c(a,w), c(a,b)) \in Q_v, \\ c(a,w) \leq \Psi < \infty\big\}, \tag{51}$$

where Ψ is a finite large constant.[23] This additional restriction makes the set $R_{v,\Psi}$ bounded. This is a purely technical necessity.

We first focus on individuals in periods between jumps. The evolution of consumption is then given by the deterministic part, i.e. the dt-part, in (4a) and (4c). We then easily understand

Lemma C.1 *Individual consumption rises if and only if current consumption relative to consumption in the other state is sufficiently high.*

For the employed worker, consumption rises if and only if $c(a_w,w)$ relative to $c(a_w,b)$ is sufficiently high,

$$\frac{dc\left(a_w,w\right)}{dt} \geq 0 \Leftrightarrow \frac{u'\left(c\left(a_w,b\right)\right)}{u'\left(c\left(a_w,w\right)\right)} \geq 1 - \frac{r-\rho}{s} \Leftrightarrow \frac{c\left(a_w,w\right)}{c\left(a_w,b\right)} \geq 1/\psi, \tag{52}$$

[22] While this is a standard property of many steady states, the standard solutions (e.g. linearization around the steady state) do not work in our case. This is in part due to the fact that the original stochastic differential equation system (4a) to (4d) - even when stripped of its stochastic part - is not an ordinary differential equation system.

[23] The constant Ψ only serves to make $R_{v,\Psi} \subset \mathbb{R}^3$ a compact set, which we need to obtain global, uniform Lipschitz constants. We shall see below that Ψ has to be chosen larger than $\Psi_0 = \frac{\psi w - b}{(1-\psi)r}$. In this case, however, Ψ does not interfere with the construction.

where

$$\psi \equiv \left(1 - \frac{r-\rho}{s}\right)^{-1/\sigma}. \tag{53}$$

For the unemployed worker, consumption rises if and only if $c(a_b, b)$ *relative to* $c(a_b, w)$ *is sufficiently high,*

$$\frac{dc(a_b, b)}{dt} \geq 0 \Leftrightarrow \frac{u'(c(a_b, w))}{u'(c(a_b, b))} \geq 1 - \frac{r-\rho}{\mu} \Leftrightarrow \frac{c(a_b, b)}{c(a_b, w)} \geq \left(1 - \frac{r-\rho}{\mu}\right)^{1/\sigma}. \tag{54}$$

Proof. Rearranging (4a) and (4c) for $dq_s = dq_\mu = 0$ and taking (1) into account gives the results (see app. C.2). Note that in what follows ψ will be used only for r sufficiently small making sure that ψ is a real number. ■

We rely on the following lemma for our proposition. It reads

Lemma C.2 *Relative consumption* $c(a, w)/c(a, b)$ *is continuously differentiable in wealth* a.

Proof. Consumption levels $c(a, w)$ and $c(a, b)$ are understood as solutions to our ODE system (5). As the latter is well-behaved within the set Q_v from (50), consumption levels are differentiable and thereby continuous in Q_v. This implies the differentiability of $c(a, w)/c(a, b)$. ■

For technical reasons, we also need to make

Assumption 1 *The number of sign changes of the derivative of relative consumption with respect to wealth, i.e.* $d(c(a, w)/c(a, b))/da$, *in any interval of finite length is finite.*

This assumption is required to rule out "pathological cases". One can construct continuously differentiable functions that change sign infinitely often in a finite neighborhood (think of $x^3 \sin(1/x)$ in a neighborhood of zero). Any economic intuition suggests that such pathological cases are not relevant for our model. We employ this assumption only in this sect. C (and implicitly in sect. D where we refer to sect. C).

Proposition C.3 *Consider a low interest rate, i.e.* $0 < r \leq \rho$. *Define a threshold level* a_w^* *by*

$$\frac{u'(c(a_w^*, b))}{u'(c(a_w^*, w))} \equiv 1 - \frac{r-\rho}{s}. \tag{55}$$

For our instantaneous utility function (1), this definition reads

$$c(a_w^*, b) = \psi c(a_w^*, w) \tag{56}$$

where ψ *is from (53).*

(i) Consumption of employed workers increases if the worker owns a sufficiently low wealth level, $a < a_w^*$. *Employed workers with* $a > a_w^*$ *choose falling consumption paths.*

(ii) Consumption of unemployed workers always decreases.

(iii) Consumption of employed workers exceeds consumption of unemployed workers at the threshold a_w^*, *i.e.* $\psi \leq 1$ *in (56) for* $r \leq \rho$.

Proof. see app. C.4 ■

C.3 Natural borrowing limit

The subsequent analysis will be facilitated by noting that there is an endogenous "natural" borrowing limit. The idea is similar to Aiyagari's (1994) borrowing limit resulting from non-negative consumption. This limit is derived in the following

Proposition C.4 *Any individual with initial wealth $a \geq -b/r$ will never be able to or willing to borrow more than $-b/r$. Consumption of an unemployed worker at $a = -b/r$ is zero, $c(-b/r, b) = 0$.*

Proof. "willing to": An employed individual with $a \geq -b/r$ will increase wealth for any wealth levels below a_w^* from (55). If a_w^* is larger than $-b/r$ – which we can safely assume – employed workers with wealth below a_w^* increase wealth and are not willing to borrow more than $-b/r$.

"able to": Imagine an unemployed worker had wealth lower than $-b/r$. Even if consumption is equal to zero, wealth would further fall, given that $\dot{a} = ra + b < 0 \Leftrightarrow a < -b/r$. If an individual could commit to zero consumption when employed and if the separation rate was zero, the maximum debt an individual could pay back is $-w/r$. Imagine an unemployed worker succeeded in convincing someone to lend her "money" even though current wealth is below $-b/r$. Then, with a strictly positive probability, wealth will fall below $-w/r$ within a finite period of time. Hence, anyone lending to an unemployed worker with wealth below $-b/r$ knows that not all of this loan will be paid back with positive probability. This cannot be the case in our setup with one riskless asset. Hence, the maximum debt level is b/r and consumption is zero at $a = -b/r$ for an unemployed worker. ■

C.4 Proof of prop. C.3 concerning Keynes-Ramsey rule

C.4.1 Proof of part (i)

- A local result

We first show that consumption $c(a_w, w)$ rises in time for wealth smaller than but close to a_w^*.

Consider relative consumption $\chi(a) \equiv x(a)/y(a)$. By ass. 1, the number of sign changes of $\chi'(a)$ in any interval for a of finite length is finite. We can therefore for any a_0 find an $\varepsilon > 0$ such that $\chi(a)$ is monotonic in $[a_0 - \varepsilon, a_0]$. Exploiting this for a_w^*, whatever the properties of relative consumption, we can always find an ε such that one of the following three cases must hold for $\Omega_\varepsilon \equiv [a_w^* - \varepsilon, a_w^*[$

$$\left.\begin{matrix} \text{(i)} \\ \text{(ii)} \\ \text{(iii)} \end{matrix}\right\} \chi'(a)|_{a\in\Omega_\varepsilon} \left\{\begin{matrix} < \\ > \\ = \end{matrix}\right\} 0.$$

Note that we do not make any statement about the derivative in a_w^*. In fact, in case (i) $\chi'(a)|_{a\in a_w^*}$ can be negative or zero, in case (ii), it can be positive or zero.

Lemma C.5 *(a) Consumption of employed workers rises over time for a wealth level $a \in \Omega_\varepsilon$ if and only if case (i) holds,*

$$\frac{dc(a_w(\tau), w)}{d\tau} > 0 \text{ for } a_w(\tau) \in \Omega_\varepsilon \Leftrightarrow \text{case (i) holds.}$$

(b) Consumption $c(a_w(\tau), w)$ falls over time for $a_w(\tau) \in \Omega_\varepsilon$ if and only if (ii) holds.

Proof. (a) By (52), $\frac{dc(a_w(\tau),w)}{d\tau} > 0 \Leftrightarrow c(a_w(\tau), w) / c(a_w(\tau), b) > 1/\psi$. As $c(a_w^*, w) / c(a_w^*, b) = 1/\psi$ at a_w^*, as w and b are parameters and using ass. 1, this is a condition on the derivative of relative consumption with respect to wealth a in Ω_ε: $dc(a_w(\tau), w) / d\tau$ is positive for $a_w(\tau) \in \Omega_\varepsilon$ if and only if case (i) holds.

(b) By (52), consumption falls over time if relative consumption lies below $1/\psi$. This can be the case in Ω_ε only if case (ii) holds. ■

Lemma C.6 *Relative consumption falls in wealth for $a \in \Omega_\varepsilon$, $\chi'(a)|_{a\in\Omega_\varepsilon} < 0$, i.e. case (i) holds.*

Proof. a) Assume that case (ii) holds, i.e. $\chi'(a)|_{a\in\Omega_\varepsilon} > 0$. Then, by lem. C.5, $\frac{dc(a_w(\tau),w)}{d\tau} < 0$ for $a_w(\tau) < a_w^*$. Consumption of unemployed workers would still decrease in time for all wealth levels. In our set Q_v from (50), $\frac{da_w(\tau)}{d\tau} > 0$ and therefore $\frac{dx(a)}{da} < 0$. As $\frac{dc(a_b(\tau),b)}{d\tau} < 0$ and $\frac{da_b(\tau)}{d\tau} < 0$ in Q_v, we know that $\frac{dy(a)}{da} > 0$. As a consequence, $\chi'(a) < 0$. This contradicts the assumption that case (ii) holds and case (ii) can be excluded.

b) Now assume that case (iii) holds, i.e. relative consumption is flat, $\chi'(a)|_{a\in\Omega_\varepsilon \cup a_w^*} = 0$. As $c(a_w^*, w) / c(a_w^*, b) = 1/\psi$, $dc(a_w(\tau), w) / d\tau = 0$ for $a_w(\tau) \in \Omega_\varepsilon$. As $dc(a_b(\tau), b) / d\tau < 0$, relative consumption is not constant – which contradicts the assumption that relative consumption is flat in wealth. As case (iii) is thereby excluded as well, the proof is complete. ■

- A global result

We now complete the proof by a global result on consumption growth.

Lemma C.7 *Consumption $c(a_w, w)$ (a) rises in time for all $a < a_w^*$ and (b) decreases in time for all $a > a_w^*$.*

Proof. (a) Imagine to the contrary of "$c(a_w, w)$ rises in time for all $a < a_w^*$" that there is an interval $]\Gamma_1, \Gamma_2[$ with $\Gamma_2 < a_w^*$ such that this is is the last interval before a_w^* where $c(a_w, w)$ falls in time,

$$dc(a_w(\tau), w) / d\tau < 0, \quad \forall\, \Gamma_1 < a_w(\tau) < \Gamma_2 < a_w^*. \tag{57}$$

We now proceed as in the proof of lem. C.6. As $\frac{da_w(\tau)}{d\tau} > 0$ in Q_v, this would imply that $\frac{dx(a)}{da} < 0$ for $\Gamma_1 < a < \Gamma_2$. We know that $\frac{dy(a)}{da} > 0$ in Q_v. Hence, we would conclude that

$$\chi'(a) < 0, \quad \forall\, \Gamma_1 < a < \Gamma_2. \tag{58}$$

By (52), the assumption in (57) would hold if and only if relative consumption $\frac{c(a_w,w)}{c(a_w,b)}$ is below $1/\psi$ for $\Gamma_1 < a < \Gamma_2$: $\frac{dc(a_w(\tau),w)}{d\tau} < 0 \Leftrightarrow \frac{c(a_w(\tau),w)}{c(a_w(\tau),b)} < 1/\psi$. As $\frac{x(a)}{y(a)}$ is continuous in wealth by lem. C.2 and as case (i) holds by lem. C.6, $\frac{x(a)}{y(a)}$ can be smaller than $1/\psi$ only if there is some range $]\Gamma_3, \Gamma_2[$ in which $\chi'(a) > 0$. (An example of such a path is shown in fig. 3.) This is a contradiction to the conclusion in (58). Hence, consumption must rise in time for all $a < a_w^*$.

(b) This proof is in analogy to the proof of (a). ■

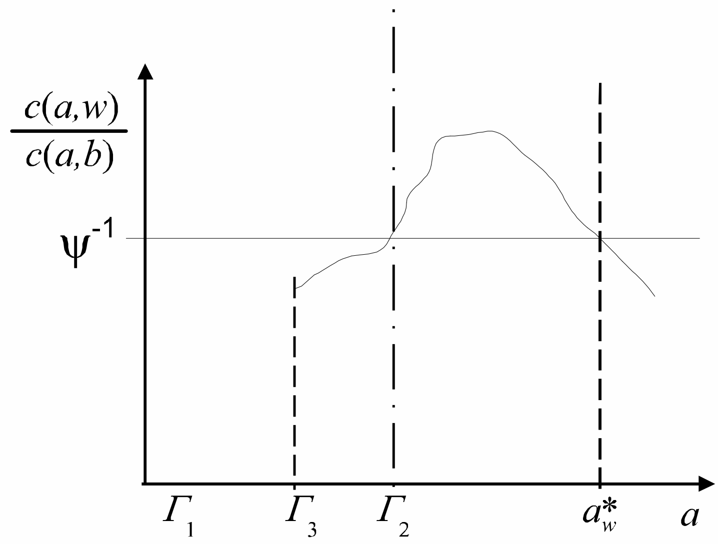

Figure 3 *An example for relative consumption* $\chi(a) \equiv \frac{x(a)}{y(a)}$

C.4.2 Intermediary steps

Before we prove the rest of prop. C.3, we need some further intermediary results – which, however, are of some interest in their own right. Given that marginal utility from (1) is positive and decreasing, $u'(c) > 0$ and $u''(c) < 0$, we can establish that $x(a) > y(a)$, i.e. consumption in the state of employment is larger than in the state of unemployment, keeping wealth constant. We prove in passing that the value functions $V(a, z)$ are strictly concave in wealth a.

Lemma C.8 *Consumption rises in wealth,* $c_a(a, z) > 0$.

Proof. Prop. C.3 (i) shows that $dc(a_w(\tau), w)/d\tau > 0$ in Q_v. As $da_w(\tau)/d\tau > 0$ as well, the derivative $dx(a)/da$ in (5) is positive in Q_v. ■

Lemma C.9 *As marginal utility from consumption is positive, the value function* $V(a, z)$ *rises in wealth,* $V_a(a, z) > 0$.

Proof. The first-order condition for optimal consumption is given by (45) in the Referees' appendix and reads

$$u'(c(a, z)) = V_a(a, z). \tag{59}$$

As marginal utility is positive by (1), the value function rises in wealth. ■

Lemma C.10 *As* $u''(c) < 0$ *and as consumption rises in* a *by lemma C.8, the value function is strictly concave in* a.

Proof. The partial derivative of the first-order condition with respect to wealth implies

$$u''(c(a, z))\, c_a(a, z) = V_{aa}(a, z). \tag{60}$$

As $u''(c(a, z)) < 0$ from the concavity of (1) and $c_a(a, z)$ is positive by lem. C.8, $V_{aa}(a, z)$ must be negative. With lem. C.9, the value function is strictly concave. ■

Lemma C.11 *The shadow price for wealth is higher in the state of unemployment,* $V_a(a, b) > V_a(a, w)$.

Proof. The derivation of the Keynes-Ramsey rule gives us (see app. C.1)

$$\begin{aligned}&(\rho - r) V_a(a,z) - s(z)[V_a(a,b) - V_a(a,w)] - \mu(z)[V_a(a,w) - V_a(a,b)]\\&= [ra + z - c(a,z)] V_{aa}(a,z).\end{aligned}$$

In state $z = w$, this means

$$(\rho - r) V_a(a,w) - s(z)[V_a(a,b) - V_a(a,w)] = [ra + w - x(a)] V_{aa}(a,w). \tag{61}$$

Given the region we are interested in (where $ra+w-x(a) > 0$) and given lemma C.10, the right-hand side is negative. Hence, the left-hand side must be negative as well. As $(\rho - r) V_a(a,w)$ is positive due to $r < \rho$, the second term must be negative. This is the case only for $V_a(a,b) > V_a(a,w)$. ■

Lemma C.12 *Consumption of the employed worker is higher than consumption of the unemployed worker,* $x(a) > y(a)$.

Proof. As $V_a(a,b) > V_a(a,w)$, the first-order condition implies $u'(y(a)) > u'(x(a))$. As the marginal utility is decreasing, $x(a) > y(a)$. ■

C.4.3 Proof of parts (ii) and (iii)

(ii) By (54), $dc(a_b(\tau), b)/d\tau < 0 \Leftrightarrow u'(c(a_b(\tau), w)) < \varkappa u'(c(a_b(\tau), b))$ where $\varkappa \equiv 1 - \frac{r-\rho}{\mu} \geq 1$ as $r \leq \rho$. As $u'(c(a_b(\tau), w)) < u'(c(a_b(\tau), b))$ with $c(a_b(\tau), w) > c(a_b(\tau), b)$ from lem. C.12, this condition always holds.

(iii) This follows from solving (55) for relative consumption.

D Existence of an optimal consumption path

This appendix provides a proof for the existence of a path $c(a,z)$ as depicted in fig. 1. We now introduce an auxiliary TSS (aTSS) in order to capture v. In analogy to the TSS Θ from (7), this point is defined by

$$\Theta_v \equiv (a_w^*, c_v(a_w^*, w)),$$

i.e. the wealth level a_w^* is unchanged but the consumption level is "a bit lower" than in the TSS. In the TSS, the consumption level is on the zero-motion line, i.e. $c(a_w^*, w) = ra_w^* + w$. In the aTSS, the consumption level is on the line $ra + w - v$ and therefore given by $c_v(a_w^*, w) = ra_w^* + w - v$. Let us now consider the following

Definition D.1 *(Optimal consumption path) A* consumption path *is a solution* $(a, c(a,w), c(a,b))$ *of the ODE-system (5) for the range* $-b/r \leq a \leq a_w^*$ *in* $R_{v,\Psi}$ *from (51) with terminal condition* $(a_w^*, c_v(a_w^*, w), c_v(a_w^*, b))$. *In analogy to the aTSS and to (56), the terminal condition satisfies* $c_v(a_w^*, w) = ra_w^* + w - v$ *and* $c_v(a_w^*, b) = \psi c_v(a_w^*, w)$ *for an arbitrary* $a_w^* > -b/r$. *An* optimal consumption path *is a consumption path which in addition satisfies* $c(-b/r, b) = 0$.

App. D.1 then proves

Theorem D.2 *There is an optimal consumption path.*

This establishes that we can continue in our analysis by taking the existence of a path $c(a,z)$ as given. Intuitively speaking, i.e. looking at v as very small constant close to zero, we know that there are paths $c(a,w)$ and $c(a,b)$ as drawn in fig. 1. The approximation implied by the auxiliary TSS is very small compared to any measurement error in the data. Values of $v = 10^{-3}$ worked perfectly in numerical solutions.

For simple reference in what follows and to simplify notation, define

$$x(a) \equiv c(a,w), \quad y(a) \equiv c(a,b), \tag{62}$$

and express the reduced form (5) as

$$\dot{x}(a) = \frac{r - \rho + s\left[\left(\frac{x(a)}{y(a)}\right)^{\sigma} - 1\right]}{ra + w - x(a)} \frac{x(a)}{\sigma}, \tag{63a}$$

$$\dot{y}(a) = \frac{r - \rho - \mu\left[1 - \left(\frac{y(a)}{x(a)}\right)^{\sigma}\right]}{ra + b - y(a)} \frac{y(a)}{\sigma}. \tag{63b}$$

D.1 Proof of theo. D.2 - existence of an optimal consumption path

D.1.1 Preliminaries

The natural borrowing limit implies that any solution to (63) must satisfy

$$y(-b/r) = 0. \tag{64}$$

In what follows, we will use classical theorems for initial value problems for ODEs. Currently, we have formulated our system (63) as a terminal value problem, since the definition of the optimal consumption path in def. D.1 uses a terminal condition at the end of the interval $[-b/r, a_w^*]$ under consideration. Using the notation from (62), this terminal condition can be written in compact form as

$$\Phi \equiv \Phi_v(\hat{a}) = (\hat{a}, x_v(\hat{a}), y_v(\hat{a})). \tag{65}$$

Note that Φ depends on v

For ease of notation and to help intuition, we shall now recast the problem into a classical initial value problem, i.e. we will require the value Φ to be attained at the fixed beginning $\tau = 0$ of an interval $[0, \tau^*]$, on which we study the problem. To this end, it is more useful to work with an autonomous system. Hence, we rewrite (63) by including $m(a) = a$ as third variable which "replaces" wealth a, which now purely serves as a parameter, i.e. as the independent variable. By using (62), this gives the system

$$\dot{m}(a) = 1,$$

$$\dot{x}(a) = \frac{r - \rho + s\left[\left(\frac{x(a)}{y(a)}\right)^{\sigma} - 1\right]}{rm(a) + w - x(a)} \frac{x(a)}{\sigma},$$

$$\dot{y}(a) = \frac{r - \rho - \mu\left[1 - \left(\frac{y(a)}{x(a)}\right)^{\sigma}\right]}{rm(a) + b - y(a)} \frac{y(a)}{\sigma}.$$

Now define $\tau \equiv \hat{a} - a$, $x_1(\tau) \equiv m(\hat{a} - \tau)$, $x_2(\tau) \equiv x(\hat{a} - \tau)$, $x_3(\tau) \equiv y(\hat{a} - \tau)$. Then, $\frac{d}{d\tau} x_1(\tau) \equiv \dot{x}_1(\tau) = \frac{d}{d\tau} m(\hat{a} - \tau) = \frac{d}{d[\hat{a} - a]} m(a) = -\frac{d}{da} m(a) = -\dot{m}(a)$. Doing the same for x

and y, the "inverted" autonomous system therefore reads

$$\dot{x}_1(\tau) = -1, \tag{66a}$$

$$\dot{x}_2(\tau) = -\frac{r-\rho+s\left[\left(\frac{x_2(\tau)}{x_3(\tau)}\right)^{\sigma}-1\right]}{rx_1(\tau)+w-x_2(\tau)}\frac{x_2(\tau)}{\sigma}, \tag{66b}$$

$$\dot{x}_3(\tau) = \frac{r-\rho-\mu\left[1-\left(\frac{x_2(\tau)}{x_3(\tau)}\right)^{\sigma}\right]}{rx_2(\tau)+b-x_3(\tau)}\frac{x_3(a)}{\sigma}, \tag{66c}$$

where now $\dot{x}_i$ denotes the derivative of $x_i(\tau)$ with respect to τ, $i = 1, 2, 3$.

Definition D.3 *Given (66) and for $\tau \geq 0$, let $X(\tau;\Phi) = (x_1(\tau), x_2(\tau), x_3(\tau))$ denote the solution of (66) started at $X(0;\Phi) = \Phi \in R_{v,\Psi}$ from (65) where $-b/r \leq \hat{a} \leq \frac{\Psi+v-w}{r}$. For later use, we also introduce the notation $x_i(\tau) = x_i(\tau;\Phi)$, $i = 1, 2, 3$.*

By passing from (63) to (66) we have reverted the time-direction – more precisely, in our setting, the wealth-direction – and turned a non-autonomous system into an autonomous one by including the independent variable as an additional component of the solution. Thus, the curve $a \mapsto (a, x(a), y(a))$ with terminal value $x(\hat{a}) = x_v(\hat{a})$, $y(\hat{a}) = y_v(\hat{a})$ is equal to the curve $\tau \mapsto X(\tau;\Phi)$ with $\Phi = \Phi(\hat{a})$, which is the solution of an initial value problem in the classical sense. However, the parametrization is reverted in the sense that in the former case we start at the left endpoint ("left" in the sense of the smallest value of the a-component) and end in the right endpoint, whereas in the latter case we start at the right endpoint and end in the left one. In particular, the absolute value of the speed along the curve is equal, but the direction is reversed.

D.1.2 Continuity of the solution in initial values

In order to be able to apply classical theorems, we need finite derivatives on the right-hand side of an ODE system. The right-hand side of the ODE (63), however, exhibits singularities at the boundary $y = ra + b$ of Q_v. This is of particular importance as the definition of the optimal consumption path in Definition D.1 uses $y(-b/r) = 0$ – which lies on this boundary. We obtain finite derivatives by (i) a coordinate transformation and by (ii) (temporarily) reducing the set on which we are interested in a solution by demanding that $y \geq \varepsilon$. We will later show how this reduction can then be removed again by passing $\varepsilon \to 0$.

Lemma D.4 *(Coordinate transformation) Let $x(a)$ and $y(a)$ be solutions of (63). The mapping $a \mapsto y(a)$ is bijective. Change variables $a = a(y)$ and consider x and a as functions of y. Then*

$$x'(y) \equiv \frac{dx(y)}{dy} = \frac{r-\rho+s\left[\left(\frac{x(y)}{y}\right)^{\sigma}-1\right]}{r-\rho-\mu\left[1-\left(\frac{y}{x(y)}\right)^{\sigma}\right]}\frac{x(y)}{y}\frac{ra(y)+b-y}{ra(y)+w-x(y)}, \tag{67a}$$

$$a'(y) \equiv \frac{da(y)}{dy} = \frac{ra(y)+b-y}{r-\rho-\mu\left[1-\left(\frac{y}{x(y)}\right)^{\sigma}\right]}\frac{\sigma}{y}. \tag{67b}$$

Proof. Since $\dot{y}(a) > 0$, y is a bijective function of a. As $a'(y) = \frac{1}{\dot{y}(a)}$, we obtain the second equation by inserting (63b). The first equation follows from "dividing (63a) by (63b)". ■

We are going to avoid the singularity at $y(-b/r) = 0$ by temporarily requiring these properties only to hold "up to an arbitrarily small number ε". We do this by considering the domain $R_{\varepsilon,v,\Psi}$ as given in the following

Definition D.5 *Fix a numbers $\varepsilon > 0$ and define*

$$R_{\varepsilon,v,\Psi} = R_{v,\Psi} \cap \left\{(a, x, y) \in \mathbb{R}^3 \,|\, y \geq \varepsilon\right\}. \tag{68}$$

This definition implies that we temporarily replace the requirement that $y(-b/r) = 0$ by $y(a) = \varepsilon$ for some $-b/r \leq a \leq -b/r + \varepsilon/r$.

Lemma D.6 *The right-hand side given in (67) is uniformly Lipschitz on $R_{\varepsilon,v,\Psi}$.*

Proof. Consider the right-hand side of (67a). The only possible points, where the Lipschitz constant can explode, are when the denominators in the right-hand side become 0 or when a term under a fractional power (i.e. with exponent σ) becomes 0. In $R = R_{\varepsilon,v,\Psi}$, y is uniformly bounded away from 0 and x is uniformly bounded away from $ra + w$. Moreover, note that $r - \rho - \mu\left[1 - \left(\frac{y}{x}\right)^\sigma\right] = 0$ if and only if $\left(\frac{y}{x}\right)^\sigma = 1 - \frac{r-\rho}{\mu}$. Now $1 - \frac{r-\rho}{\mu} > 1$ by the assumption that $r < \rho$. On the other hand, $y < x$, implying that $\left(\frac{y}{x}\right)^\sigma < 1$. Consequently, all the denominators are uniformly bounded away from 0.

For the fractional powers, note that $x/y > 1$ is trivially uniformly bounded away from 0. As $x \leq \Psi$,

$$\frac{y}{x} > \frac{\epsilon}{\Psi}$$

is uniformly bounded away from 0 on $R_{\varepsilon,v,\Psi}$. This shows that (67a) is uniformly Lipschitz.

The same arguments show that the right-hand side of (67b) is uniformly Lipschitz, too. ■

Since the right hand side of (67) is uniformly Lipschitz, we can now apply the classical theory of ODEs. For instance, we have existence and uniqueness of the solution by the Picard-Lindelöf theorem, see Mattheij and Molenaar (2002, th. II.2.3, th. II.3.1). Moreover, the solution will be continuous as a function of the initial value, see, again, Mattheij and Molenaar (2002, th. II.4.7). In the lemma below, we will see how this even implies the corresponding properties for the non-transformed system (66).

Lemma D.7 *(Continuity in initial values) Consider the set $R = R_{\varepsilon,v,\Psi}$ from (68) and the solution $X(\tau; \Phi)$ from Definition D.3 with initial condition Φ given in (65). The solution $X(\tau; \Phi)$ depends continuously on its initial values Φ. More precisely, there is a constant $L > 0$ and an increasing map $\kappa : [0, \infty[\to [0, \infty[$ (a* modulus of continuity*) with $\lim_{t \searrow 0} \kappa(t) = \kappa(0) = 0$ such that*

$$\|X(\tau_1; \Phi_1) - X(\tau_2; \Phi_2)\| \leq L\|\Phi_1 - \Phi_2\| + \kappa(|\tau_1 - \tau_2|),$$

provided that $\Phi_1, \Phi_2 \in R$ and $X(\tau; \Phi_i) \in R$ for all $0 \leq \tau \leq \max(\tau_1, \tau_2)$, $i = 1, 2$. Here, $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^3$.

Proof. By classical results from the theory of ordinary differential equations, see for instance Mattheij and Molenaar (2002, th. II.4.7), the solution of an ODE-system depends continuously on the initial data as long as the right-hand side is uniformly Lipschitz. More precisely, let $Y(\tau; \Phi)$ denote the solution of an ODE with uniformly Lipschitz right-hand side (with Lipschitz constant C), started at $Y(\tau_0; \Phi) = \Phi$, then

$$\|Y(\tau; \Phi_1) - Y(\tau; \Phi_2)\| \leq \exp\left(C(\tau - \tau_0)\right) \|\Phi_1 - \Phi_2\|.$$

Now consider the transformed system $(a(y), x(y))$ from (67). By Lemma D.6, the right-hand side is uniformly Lipschitz. The solution of (67) therefore depends continuously on its initial data (a_0, x_0). It is then obvious that the trajectory $(a(y), x(y), y)$ depends continuously on (a_0, x_0, y_0). As system (67) is a reparameterized version of (63), the solution $(a, x(a), y(a))$ to (63) from def. D.1 is also continuous in its boundary conditions – even though the right hand side of (63) is not uniformly Lipschitz. Similarly, as (66) is just a reparameterization of (63), the solution $X(\tau;\Phi)$ to (66) from def. D.3 is also continuous in its initial condition Φ.

In order to get the estimate, we now consider the ODE (66) and note that we only consider it on the compact set $R_{\varepsilon,v,\Psi}$. In the parametrization by y given in (67), y is the independent variable, i.e. plays the role of τ in the above estimate. By compactness of $R_{\varepsilon,v,\Psi}$, y only runs through a bounded set, therefore we can rewrite the constant in the above inequality as $\exp(C(y-y_0)) \leq L$ for some suitable $L > 0$.

Given $\Phi \in R_{\varepsilon,v,\Psi}$. Then $a_w^* \leq \frac{\Psi-w+v}{r}$, which implies that the solution $X(\tau;w)$ can only stay inside $R_{\varepsilon,v,\Psi}$ until time $\tau = \frac{\Psi-w+v+b}{r}$, at most. Consider

$$D = \{(\tau,\Phi) \in [0,\infty[\times R_{\varepsilon,v,\Psi} \,|\, X(\tau;\Phi) \in R_{\varepsilon,v,\Psi}\}.$$

Then D is a closed subset of $\left[0, \frac{\Psi-w+v+b}{r}\right] \times R_{\varepsilon,v,\Psi}$, implying that D is compact. Consequently, $X : D \to R_{\varepsilon,v,\Psi}$ is uniformly continuous, which implies the existence of a modulus of continuity κ with

$$\|X(\tau_1;\Phi_1) - X(\tau_2;\Phi_2)\| \leq \kappa(|\tau_1-\tau_2| + \|\Phi_1-\Phi_2\|).$$

The inequality in the lemma then follows by the triangle inequality. ■

D.1.3 Continuity of the first hitting-wealth in initial values

While we have shown in the previous section that the solutions to all systems (63), (66) and (67) are continuous in initial values, this does not automatically imply that the solutions will be continuous on the boundary of the domain we are interested in, in the sense that the place where the solution leaves the domain R might not depend continuously on the initial data. This will now be proved in this section.

In the proofs and also in a later step, we will use the following

Definition D.8 *(First hitting-wealth) Consider the set $R_{\varepsilon,v,\Psi}$ from (68) and the solution $X(\tau;\Phi)$ to the system (66). Consider the path $y(a)$ that corresponds to $x_2(\tau)$ of this solution. Then we define $\hat{a}_{1st} = f(\hat{a})$ as the "first hitting-wealth" (in analogy to first hitting-time), i.e. the wealth level where the path $y(a)$ hits any boundary of $R_{\varepsilon,v,\Psi}$ for the first time. Similarly denote $\tau(\Phi) \equiv \inf\{\tau \geq 0 \,|\, X(\tau;\Phi) \in \partial R_{\varepsilon,v,\Psi}\}$ and $F(\Phi) \equiv X(\tau(\Phi);\Phi)$.*

We know that $\hat{a}_{1st}$ exists because in the set $R_{\varepsilon,v,\Psi}$ the derivatives in (66) are well-defined and a solution therefore exists. Notice that $\hat{a}_{1st}$ equals the first component of $F(\Phi(\hat{a}))$.

We also need

Definition D.9 *Let $N \subset R_{\varepsilon,v,\Psi}$ with*

$$N = \left\{\Phi(\hat{a}) \,\middle|\, \hat{a} \in \left[-\frac{b}{r}, \frac{\psi[w-v]-b}{r[1-\psi]}\right]\right\}$$

be the set of all potential initial conditions from (65) for a solution in the sense of def. D.1. Here we implicitly assume that Ψ is large enough that indeed $N \subset R_{\varepsilon,v,\Psi}$.[24] *Define M as*

$$M = M_1 \cup M_2 \cup M_3 \subset R_{\varepsilon,v,\Psi} \tag{69}$$

with

$$\begin{aligned} M_1 &= \{(a,x,y) \in R_{\varepsilon,v,\Psi} \,|\, y = ra+b\}, \\ M_2 &= \{(a,x,y) \in R_{\varepsilon,v,\Psi} \,|\, a = -b/r\}, \\ M_3 &= \{(a,x,y) \in R_{\varepsilon,v,\Psi} \,|\, y = \varepsilon\}. \end{aligned}$$

This set will turn out to be the set of all potential first hitting-wealths.

[24] This is the only necessary condition on Ψ for the construction to work. In the sequel, we shall assume this condition without further notice.

Since we know that $x > y$, the trajectory will not hit the boundary of R at the part $\{x = y\}$. Therefore, we have the

Corollary D.10 *$F : N \to M$ is a well-defined map, i.e. for every $\Phi \in N$, the corresponding solution path $X(\tau; \Phi)$ exists and stays in $R_{\varepsilon,v,\Psi}$ until it finally hits M (and no other boundary of $R_{\varepsilon,v,\Psi}$).*

Before formulating the main lemma of this section, let us first derive a simple bound on the derivative $\dot{y}(a)$ of the consumption of the unemployed.

Lemma D.11 *For (a, x, y) in the interior of Q_v from (50), we have*

$$\dot{y}(a) \geq \frac{r - \rho}{ra + b - y(a)} \frac{y(a)}{\sigma}.$$

Proof. By (63b) we have

$$\begin{aligned} \dot{y}(a) &= \frac{r - \rho - \mu \left[1 - \left(\frac{y(a)}{x(a)}\right)^{\sigma}\right]}{ra + b - y(a)} \frac{y(a)}{\sigma} \\ &= \left(\frac{r - \rho}{ra + b - y(a)} - \frac{\mu \left[1 - \left(\frac{y(a)}{x(a)}\right)^{\sigma}\right]}{ra + b - y(a)} \right) \frac{y(a)}{\sigma} > \frac{r - \rho}{ra + b - y(a)} \frac{y(a)}{\sigma}. \end{aligned}$$

The last inequality follows from the fact that $\frac{\mu\left[1-\left(\frac{y(a)}{x(a)}\right)^{\sigma}\right]}{ra+b-y(a)}$ is negative (and therefore $-\frac{\mu\left[1-\left(\frac{y(a)}{x(a)}\right)^{\sigma}\right]}{ra+b-y(a)}$ is positive) as $ra + b - y(a)$ is negative in the interior of Q_v. ■

The key result in this section is presented in

Lemma D.12 *The map $F : N \to M$ is continuous.*

Proof. We need to prove that for every $\Phi \in N$ and every $\delta > 0$ there is an $\eta > 0$ such that

$$\|\Phi_0 - \Phi\| < \eta \Longrightarrow \|F(\Phi_0) - F(\Phi)\| < \delta. \tag{70}$$

We start the proof by fixing Φ_0, $\Phi \in N$ such that $\|\Phi_0 - \Phi\| < \eta$ for some $\eta > 0$. Let us first *assume* that $\tau(\Phi_0) \leq \tau(\Phi)$. By the triangle inequality and Lemma D.7, we have

$$\begin{aligned} \|X(\tau(\Phi_0); \Phi_0) - X(\tau(\Phi); \Phi)\| &\leq \|X(\tau(\Phi_0); \Phi_0) - X(\tau(\Phi_0); \Phi)\| + \\ &\quad + \|X(\tau(\Phi_0); \Phi) - X(\tau(\Phi); \Phi)\| \\ &\leq L_1 \|\Phi_0 - \Phi\| + \kappa(|\tau(\Phi_0) - \tau(\Phi)|), \end{aligned} \tag{71}$$

for a constant $L_1 > 0$ and the modulus of continuity κ. In order to get an estimate for $|\tau(\Phi_0) - \tau(\Phi)|$, we have to distinguish between three different cases.

Case (i): $F(\Phi_0) \in M_1$.

By Lemma D.11, there are constants $L_2, \ell_2 > 0$ such that $\dot{y} \geq L_2$ for $|y - (ra + b)| \leq \ell_2$. More precisely, we can choose $\ell_2 > 0$ freely and obtain the bound for $L_2 = \frac{1}{\ell_2} \frac{(\rho - r)\varepsilon}{\sigma}$. If $L_1 \eta \leq \ell_2$, we can bound the absolute value of the derivative of $x_3(\tau; \Phi)$ from below by L_2 (for $t \geq \tau(\Phi_0)$). This implies that the path $X(\tau; \Phi)$ hits M_1 before time $\tau(\Phi_0) + \tau$ for

$$\tau(L_2 - r) = \ell_2 \iff \tau = \frac{\ell_2}{L_2 - r},$$

unless it hits another boundary of $R_{\varepsilon,v,\Psi}$ before that. Inserting into (71), this gives the estimate

$$\|F(\Phi_0) - F(\Phi)\| \leq L_1\eta + \kappa\left(\frac{\ell_2}{L_2 - r}\right).$$

Choosing $\ell_2 = L_1\eta$, the bound is smaller than δ provided that

$$\kappa\left(\frac{L_1}{\frac{C}{L_1\eta} - r}\eta\right) + L_1\eta < \delta, \tag{72}$$

where $C \equiv \frac{(\rho-r)\varepsilon}{\sigma}$. Note that the left hand side in (72) converges to zero for $\eta \to 0$, therefore we can find an $\eta_0(\delta) > 0$ (only depending on the constants C, L_1 and r and the modulus of continuity κ, but not on Φ_0 or Φ) such that the desired inequality (70) holds for $\eta < \eta_0$. We have tacitly assumed that $L_2 = C/\ell_2 = \frac{C}{L_1\eta} > r$, which can be realized by choosing η small enough.

Case (ii): $F(\Phi_0) \in M_2$.
Let $\hat{a}$ denote the first component of Φ, and $\hat{a}_0$ the first component of Φ_0. Note that $x_1(\tau;\Phi) = \hat{a} - \tau$, for every $\tau \geq 0$. Since $X(\tau(\Phi_0);\Phi_0) \in M_2$, we have $-b/r = x_1(\tau(\Phi_0);\Phi_0) = \hat{a}_0 - \tau(\Phi_0)$, implying that $\tau(\Phi_0) = \hat{a}_0 + b/r$. On the other hand, $x_1(\tau(\Phi);\Phi) \geq -b/r$, implying that $\tau(\Phi) \leq \hat{a} + b/r$. Combining these two results, we obtain

$$|\tau(\Phi_0) - \tau(\Phi)| = \tau(\Phi) - \tau(\Phi_0) \leq \hat{a} - \hat{a}_0 \leq \|\Phi_0 - \Phi\|.$$

Consequently, the inequality (71) implies

$$\|F(\Phi_0) - F(\Phi)\| \leq L_1\|\Phi_0 - \Phi\| + \kappa(\|\Phi_0 - \Phi\|) \leq L_1\eta + \kappa(\eta),$$

and (70) holds for η small enough such that

$$L_1\eta + \kappa(\eta) < \delta. \tag{73}$$

Case (iii): $F(\Phi_0) \in M_3$.
Since $x_3(\tau(\Phi_0);\Phi_0) = \varepsilon$, we have $0 \leq x_3(\tau(\Phi_0);\Phi) - \varepsilon \leq L_1\eta$. By Lemma D.11, we can find a constant $L_3 > 0$ such that $\dot{y} \geq L_3$ on $R_{\varepsilon,v,\Psi}$ – note that L_3 depends on ε. Thus, $X(s;\Phi)$ will hit the boundary M_3 before time $\tau(\Phi_0) + \tau$ with $\tau = L_1\eta/L_3$, unless it hits another boundary of $R_{\varepsilon,v,\Psi}$ before. In any case, $|\tau(\Phi_0) - \tau(\Phi)| \leq L_1\eta/L_3$, and we obtain

$$\|F(\Phi_0) - F(\Phi)\| \leq L_1\eta + \kappa\left(\frac{L_1}{L_3}\eta\right),$$

and (70) is satisfied for

$$L_1\eta + \kappa\left(\frac{L_1}{L_3}\eta\right) < \delta. \tag{74}$$

Choosing η small enough that both (72) and (73) and (74) are satisfied, settles the proof for $\tau(\Phi_0) \leq \tau(\Phi)$. Notice that none of the conditions (72), (73) and (74) depends on Φ_0. Therefore, in the other case $\tau(\Phi_0) \geq \tau(\Phi)$, we can just revert the rôles of Φ and Φ_0 and obtain the same results in cases (i), (ii) and (iii). ■

D.1.4 Existence of a solution

This section proves our main result formulated in Theorem D.2.

Proof. Fix some $\varepsilon > 0$ and consider $R_{\varepsilon,v,\Psi}$. By an intermediate value theorem applied to $F : N \to M$, we will obtain a point or points $\Phi \in N$ such that $F(\Phi) \in M_3$ as used in (69), i.e. $x_3(\tau(\Phi);\Phi)=\varepsilon$ provided that we can show the existence of points (that could be called upper and lower bounds) $\Phi_v^{\min},\Phi_v^{\max}\in N$ with $F(\Phi_v^{\min})\in M_2$ and $F(\Phi_v^{\max})\in M_1$. (Note that $F=F_\varepsilon$ and all the $M_i=M_i(\varepsilon)$, $i=1,2,3$, depend on ε and v, but not on Ψ, provided that Ψ is large enough.)

Choose

$$\Phi_v^{\min}=\Phi(-b/r)=(-b/r,w-b-v,\psi\left[w-b-v\right]),\quad \Phi_v^{\max}=\Phi\left(\frac{\psi(w-v)-b}{(1-\psi)r}\right).$$

By construction, both $\Phi_v^{\min}$ and $\Phi_v^{\max}$ are contained in N. Moreover, we trivially have $F_\varepsilon(\Phi_v^{\min})\in M_2(\varepsilon)$, $F_\varepsilon(\Phi_v^{\max})\in M_1(\varepsilon)$ for every $\varepsilon>0$ small enough. Note, in particular, that Lemma D.12 also implies continuity of F in the boundary points $\Phi_v^{\min}$ and $\Phi_v^{\max}$ of N. Therefore, the image set $F_\varepsilon(N)$ is a connected set, with non-empty intersection with both M_1 and M_2. Since the distance

$$\operatorname{dist}(M_1,M_2)=\inf\left\{\|\Phi_1-\Phi_2\|\mid \Phi_1\in M_1,\ \Phi_2\in M_2\right\}=\frac{\varepsilon}{r}>0,$$

we may conclude that $F_\varepsilon(N)\cap M_3(\varepsilon)\neq\emptyset$. This establishes that there must be a Φ such that $F_\varepsilon(\Phi)\in M_3$. In words, there is an initial condition $\Phi(\hat a)$ such that the path $(a,x(a),y(a))$ hits the boundary at $y=\varepsilon$.

Now define

$$N_3(\varepsilon)\equiv F_\varepsilon^{-1}(M_3(\varepsilon))=\left\{\Phi\in N\,|\,F_\varepsilon(\Phi)\in M_3(\varepsilon)\right\}.$$

By continuity of $F_\varepsilon:N\to M(\varepsilon)$, the bounded set $N_3(\varepsilon)$ is closed and thus compact. Moreover, the family $(N_3(\varepsilon))_{\varepsilon>0}$ is directed in the sense that

$$0<\varepsilon_2<\varepsilon_1\Longrightarrow N_3(\varepsilon_2)\subset N_3(\varepsilon_1).$$

By standard results from topology, the intersection of a directed family of non-empty, compact sets is non-empty, i.e.

$$N_3(0)\equiv\bigcap_{\varepsilon>0}N_3(\varepsilon)\neq\emptyset.$$

Indeed, take a decreasing sequence $(\varepsilon_n)_{n\geq1}$ of positive numbers converging to zero. For every n choose some $\Phi_n\in N_3(\varepsilon_n)$. By compactness of the largest set $N_3(\varepsilon_1)$, we can find a subsequence n_k such that $(\Phi_{n_k})_{k\geq1}$ converges to some Φ. Note that $\Phi\in N_3(\varepsilon_{n_k})$ for every k, since $\Phi=\lim_{l\to\infty,\,l\geq k}\Phi_{n_l}$ and each such Φ_{n_l} lies in the closed set $N_3(\varepsilon_{n_k})$. Now choose any $\varepsilon>0$ and pick a k such that $\varepsilon_{n_k}<\varepsilon$. Then $\Phi\in N_3(\varepsilon_{n_k})\subset N_3(\varepsilon)$, implying that $\Phi\in\bigcap_{\varepsilon>0}N_3(\varepsilon)$.

We claim that every element $\Phi\in N_3(0)$ corresponds to an aTSS. Indeed, the path $(a,x(a),y(a))$ with terminal value $(\hat a,\hat x,\hat y)=\Phi$ (corresponding to the path $X(\tau;\Phi)$) satisfies the ODE (63) on $]-b/r,\hat a]$. Moreover, it starts at N by construction, and for every $\varepsilon>0$, it takes on the value ε somewhere on the interval $]-b/r,-b/r+\varepsilon[$. Thus, using monotonicity of y, we may conclude that

$$\lim_{a\searrow-b/r}y(a)=0.$$

This establishes that there is an initial condition $\Phi(\hat a)$ such that the path $y(a)$ hits the boundary at $y=0$ in the sense that $y(-b/r)=0$. ■

Note that it is essential for the proof of Theorem D.2 that the trajectory $X(\tau;\Phi)$ – or, equivalently, $(a,x(a),y(a))$ – does not depend on ε, which only determines "how long" we observe the trajectory. This means that we observe the trajectory $X(\tau;\Phi)$ for $0\leq\tau\leq\tau(\Phi)$, with the hitting time $\tau(\Phi)$ obviously depending on ε. Therefore, we can, for fixed $\Phi\in N_3(0)$, easily take the limit $\varepsilon\to0$, which means that we take the limit in $\tau(\Phi)$, but do not change the trajectory itself. As a consequence, the ODE is automatically satisfied for the limit, at least for $0\leq\tau<\lim_{\varepsilon\to0}\tau(\Phi)$.

Let us illustrate why we had to use the specific properties of the dynamic system (66) in the proof of lem. D.12. Continuity in initial conditions does not imply continuity of "first hitting values" in general. Indeed, the first hitting times are inherently non-continuous functionals, even if both the paths and the set, which determines the hitting times, are smooth.

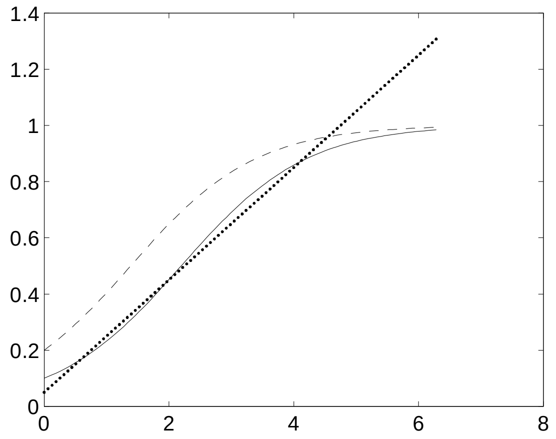

Figure 4 *Non-continuity of the first hitting time*

To see this most clearly, consider the differential equation $\dot{z}(t) = (1 - z(t)) z(t)$ whose solution is $z(t) = \left(1 + \left(z_0^{-1} - 1\right) e^{-t}\right)^{-1}$. This solution is continuous in the initial level z_0 (for $z_0 > 0$ which we assume) and the solution is plotted for $z_0 \in \{0.1, 0.2\}$ in fig. 4. Now consider the first-hitting time on the straight line $0.05 + t/5$ as drawn. Obviously, this time is not continuous in the initial values z_0.

E Properties of the wealth distribution

E.1 A density gives a density

Given the definition of $I(t)$ in prop. 5.1, write the time derivative as

$$\frac{d}{d\tau} I(\tau) = \frac{d}{d\tau} \int_{-\infty}^{\infty} p(a, w, \tau) + p(a, b, \tau)\, da = \int_{-\infty}^{\infty} \left[\frac{d}{d\tau} p(a, w, \tau) + \frac{d}{d\tau} p(a, b, \tau) \right] da.$$

Using the partial differential equations (22), we get

$$\begin{aligned}
&\frac{d}{d\tau} I(\tau) \\
&= \int_{-\infty}^{\infty} \left[-\{ra + w - c(a, w)\} \frac{\partial}{\partial a} p(a, w, \tau) - \left\{ r - \frac{\partial}{\partial a} c(a, w) + s \right\} p(a, w, \tau) + \mu p(a, b, \tau) \right] da \\
&+ \int_{-\infty}^{\infty} \left[-\{ra + b - c(a, b)\} \frac{\partial}{\partial a} p(a, b, \tau) - \left\{ r - \frac{\partial}{\partial a} c(a, b) + \mu \right\} p(a, b, \tau) + s p(a, w, \tau) \right] da \\
&= - \int_{-\infty}^{\infty} \left[\{ra + w - c(a, w)\} \frac{\partial}{\partial a} p(a, w, \tau) + \left\{ r - \frac{\partial}{\partial a} c(a, w) \right\} p(a, w, \tau) \right] da \\
&- \int_{-\infty}^{\infty} \left[\{ra + b - c(a, b)\} \frac{\partial}{\partial a} p(a, b, \tau) + \left\{ r - \frac{\partial}{\partial a} c(a, b) \right\} p(a, b, \tau) \right] da, \qquad (75)
\end{aligned}$$

where the last equality dropped the terms $\mu p(a, b\tau)$ and $sp(a, w, \tau)$. Now integrate by parts, i.e. use $\int_{-\infty}^{\infty} u(a) v'(a) \, da = [u(a) v(a)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} u'(a) v(a) \, da$ to find

$$\begin{aligned}
&\int_{-\infty}^{\infty} \{rz + z - c(a, z)\} \frac{\partial}{\partial a} p(a, z, \tau) \, da \\
&= [\{ra + z - c(a, z)\} p(a, z, \tau)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \left\{r - \frac{\partial}{\partial a} c(a, z)\right\} p(a, z, \tau) \, da.
\end{aligned}$$

As the density is zero for all $a < -b/r$ and $a > a_w^*$, the first term is zero. We therefore find for the derivative of our integral

$$\begin{aligned}
\frac{d}{d\tau} I(\tau) &= \int_{-\infty}^{\infty} \left\{r - \frac{\partial}{\partial a} c(a, w)\right\} p(a, w, \tau) \, da - \int_{-\infty}^{\infty} \left\{r - \frac{\partial}{\partial a} c(a, w)\right\} p(a, w, \tau) \, da \\
&+ \int_{-\infty}^{\infty} \left\{r - \frac{\partial}{\partial a} c(a, b)\right\} p(a, b, \tau) \, da - \int_{-\infty}^{\infty} \left\{r - \frac{\partial}{\partial a} c(a, b)\right\} p(a, b, \tau) \, da \\
&= 0.
\end{aligned}$$

E.2 The subdistribution functions (24)

Look again at the FPEs (22). Observe that

$$\frac{d}{da} [\{ra + z - c(a, z)\} p(a, z, \tau)] = \{ra + z - c(a, z)\} \frac{\partial}{\partial a} p(a, z, \tau) + \left\{r - \frac{\partial}{\partial a} c(a, z)\right\} p(a, z, \tau)$$

and write the FPEs as

$$\begin{aligned}
\frac{\partial}{\partial \tau} p(a, w, \tau) &= -\frac{d}{da} [\{ra + w - c(a, w)\} p(a, w, \tau)] \\
&\quad -sp(a, w, \tau) + \mu p(a, b, \tau), \\
\frac{\partial}{\partial \tau} p(a, b, \tau) &= -\frac{d}{da} [\{ra + b - c(a, b)\} p(a, b, \tau)] \\
&\quad +sp(a, w, \tau) - \mu p(a, b, \tau).
\end{aligned} \tag{76}$$

Now consider subdistributions (as we describe subdensities), i.e. consider

$$P(a, z, \tau) \equiv \int_{-b/r}^{a} p(a, z, \tau) \, da. \tag{77}$$

As a preliminary step, compute

$$\begin{aligned}
&\int_{-b/r}^{a} \frac{d}{da} [\{ra + z - c(a, z)\} p(a, z, \tau)] \, da \\
&= \{ra + z - c(a, z)\} p(a, z, \tau) - \left[\{-b + z - c(-b/r, z)\} p\left(-\frac{b}{r}, z, \tau\right)\right] \\
&= \{ra + z - c(a, z)\} p(a, z, \tau)
\end{aligned} \tag{78}$$

where we used the fact (see main text) that $p\left(-\frac{b}{r}, z, \tau\right) = 0$ for all τ. Using this, we can compute the time derivative of the subdistribution (77) for $z = w$ as

$$\begin{aligned}
\frac{d}{d\tau} \int_{-b/r}^{a} p(a, w, \tau) \, da &= \int_{-b/r}^{a} \frac{d}{d\tau} p(a, w, \tau) \, da \\
&= -\{ra + w - c(a, w)\} p(a, w, \tau) + \int_{-b/r}^{a} -sp(a, w, \tau) + \mu p(a, b, \tau) \, da,
\end{aligned}$$

where the second equality used the PDE from (76) and the result from (78) also for $z = w$.

Using the definition from (77) and doing the same steps for $z = w$, we therefore found a differential equation system for subdistributions that reads

$$\frac{d}{d\tau}P(a,w,\tau) = -\{ra + w - c(a,w)\}\,p(a,w,\tau) - sP(a,w,\tau) + \mu P(a,b,\tau),$$
$$\frac{d}{d\tau}P(a,b,\tau) = -\{ra + b - c(a,b)\}\,p(a,b,\tau) + sP(a,w,\tau) - \mu P(a,b,\tau).$$

Using the derivative of the definition of subdistributions in (77), we finally write this as

$$\frac{\partial}{\partial\tau}P(a,w,\tau) = -\{ra + w - c(a,w)\}\,\frac{\partial}{\partial a}P(a,w,\tau) - sP(a,w,\tau) + \mu P(a,b,\tau),$$
$$\frac{\partial}{\partial\tau}P(a,b,\tau) = -\{ra + b - c(a,b)\}\,\frac{\partial}{\partial a}P(a,b,\tau) - \mu P(a,b,\tau) + sP(a,w,\tau).$$