

# Post-crisis macrofinancial modelling: continuous time approaches

Chapter for "Handbook of Post-Crisis Financial Modelling" \*

Jukka Isohätälä, Nataliya Klimenko and Alistair Milne

April, 2015

## 1 Introduction

Prior to the crisis the dominant paradigm in macroeconomic modelling was the micro-founded ‘New-Keynesian’ DSGE model (described in many textbooks including the influential exposition of [Woodford \[2003\]](#)). In its most basic form this combines price-stickiness with forward looking decision making by both households and firms. This provides a tractable framework for capturing the response of output and inflation to both demand and supply shocks and explaining intuitively the transmission of monetary policy (with monetary policy characterized as a choice over rules for current and future interest rates).

DSGE models have proved to be remarkably adaptable, being easily extended in many ways, most commonly by incorporating the so called ‘financial accelerator’, a premium on the cost of external investment finance decreasing in firm net worth ([Bernanke et al. \[1999\]](#)) and hence creating an extended dynamic response to shocks. DSGE models could also be fitted closely to macroeconomic data, successfully capturing macroeconomic fluctuations observed over several past decades (as demonstrated by [Smets and Wouters \[2005\]](#)).

Despite these successes the crisis revealed fundamental weaknesses in this DSGE paradigm. DSGE models proved incapable of explaining the protracted decline in output and investment in the industrial countries following the crisis of 2008 (or similarly persistent declines following other previous financial crises as documented by [Reinhart and Rogoff \[2009\]](#)). Contrary to widespread perception, DSGE models can be relatively easily extended to incorporate banks and bank balance sheets.<sup>1</sup> However, even with banking and other financial frictions, DSGE models, in their usual linearised form, fail to reproduce the sudden, substantial and long-lasting changes in asset prices, output or investment inherent in the periods of financial crises including that of 2008.

---

\*We are grateful for comments from Marcus Brunnermeier and from the handbook editors

<sup>1</sup>An example of such a DSGE extension is [Meh and Moran \[2010\]](#) who generalise the financial accelerator to include a bank-moral hazard based on [Holmstrom and Tirole \[1997\]](#).

The objective of this chapter is to introduce an emerging literature, pursued since the financial crisis, employing non-linear continuous time specifications of economic dynamics to capture the possibility of marked and sometimes long lasting changes in financial asset prices and asset price volatility or in real economy aggregates such as output or investment.<sup>2</sup> Prominent contributions to this new literature include [He and Krishnamurthy \[2012\]](#) and [Brunnermeier and Sannikov \[2014b\]](#). We aim to explain the methods used in this new literature and demonstrate how they can be applied to a range of different modelling problems. This is however not a complete review of the literature on macroeconomics with financial frictions ([Brunnermeier et al. \[2012\]](#) provides more extended review than we do, discussing a wider range of macroeconomic consequences of market incompleteness with extensive references to prior literature). Our aim is more limited, providing a fairly full discussion of what we perceive as some of the key contributions and describing both the economic intuition and technical solution methods that underpins their results.

This new approach to macroeconomic modelling is still very much in its infancy and the specifications employed in this generation of models are highly stylised. One way of describing this new literature is to say that it applies the tool of continuous-time modelling widely used for derivative and other asset pricing problem to a new class of macroeconomic general equilibrium problems. This though is a bit of an oversimplification – the standard financial applications of continuous time modelling beginning with [Merton \[1969, 1971\]](#) and [Black and Scholes \[1973\]](#) all assume complete markets. By contrast, the key underlying assumption of this new literature is market incompleteness – not all risks can be costlessly traded. The reasons for this market incompleteness are not typically however modelled. Instead, the focus is on the implications of market incompleteness for aggregate macrodynamics and in particular the macrodynamic role of balance sheet structure (the net worth and leverage of households, companies and financial intermediaries).

Market incompleteness can also be modelled in a discrete time setting, so why employ continuous time? The reason is that specifying the dynamics of the economy in continuous time, using diffusion processes governed by stochastic differential equations or sometimes jump processes, allows for a convenient description of the fully non-linear macrodynamics. The possible realisations of the economy are characterised by a set of differential equations<sup>3</sup> and the solution of these equations, subject to appropriate boundary conditions, yielding both the macroeconomic outcomes (as a function of state) and the probabilities of these outcomes occurring (that is, the ‘ergodic’ density or the probability density function of the state variable). Knowledge of the probability distribution of states then allows the analysis of the full macroeconomic dynamics. In the models reviewed in this chapter this approach is used to characterise both the

---

<sup>2</sup>Our paper complements [Brunnermeier and Sannikov](#) who provide a detailed discussion of the solution methods employed in these continuous-time models of this kind.

<sup>3</sup>Generally, these are partial differential equations, but when the model in question has just a single state variable, as is the case in the models we review here, the equations become ordinary differential equations.

impact and persistence of fundamental shocks and how this can reproduce some characteristic crisis features.

A key determining feature of the properties of this new generation of macroeconomic models is the magnitude of shocks relative to the balance sheet constraints that arise because of market incompleteness. If these shocks are relatively small, the model dynamics are dominated by the deterministic components of equations of state motion (e.g., the planned or expected saving and investment) and the diffusion of state towards these net worth or leverage constraints occurs only rarely so that model predictions are not so very different from those of conventional macroeconomic models. In this case linearised models of the kind employed in the DSGE tradition can adequately approximate the fully non-linear solution.

However, if shocks are sufficiently large so that stochastic disturbance can on occasion become much more important than the deterministic components of state motion and net worth or leverage are pushed towards constrained levels relatively frequently, then *qualitative* changes in model predictions are possible. Agents (households, firms, governments) substantially alter their behaviour, not just when the constraints are actually binding but when they are close to binding and sometimes even quite far away from these constraints. They do so in order to self-insure, offsetting the absence of markets that they would like to use to protect themselves against risk. This collective attempt to avoid risk can then in turn create feedbacks at the macroeconomic level following a large disturbance. The latter induce additional volatility of asset prices encouraging even greater self-insurance and inefficient employment of real economic resources (amplification) that potentially trigger long lasting declines of real macroeconomic aggregates (persistence) such as output, employment and investment.

In these circumstances DSGE-based linearisation can no longer provide an adequate description of aggregate dynamics, as this requires explicit modelling of induced volatility rather than the trend. Note though that there is no necessary and direct relationship between the magnitude of shocks and the frequency of such crisis episodes. In many of these specifications a relatively small exogenous noise may cause agents to operate with relatively small buffers of net worth, in which case even comparatively small disturbances can result in substantial departures from the predictions of linearised macroeconomic models (this a key finding of [Isohätälä et al. \[2014b\]](#) and seems to be what underlies the ‘paradox of volatility’ described, for example, by [Brunnermeier and Sannikov \[2014b\]](#)).

This chapter contains three main sections and provides detailed discussion of six contributions to the literature. Section 2 provides a general overview of this new literature, discussing how a combination of specific economic assumptions and modelling strategy generates results which differ sharply from more established traditions of macrodynamic modelling. Section 3 reviews a number of recent applications, some journal published, others work of our own still at working paper stage. This section is itself divided into a number of subsections: 3.1 focuses on the continuous time modelling of the dynamics of asset prices, following the approach taken by [He and Krishnamurthy \[2012\]](#) and also a related problem of optimal savings and consumption in general equilibrium addressed

by Isohätälä et al. [2014a]; 3.2 then discusses the dynamic modelling of the interaction of sectoral balance sheets with production and investment, focusing on the work of Brunnermeier and Sannikov [2014b] and the closely related partial equilibrium model of Isohätälä et al. [2014b]; 3.3 then discusses the further extension of these models to an explicit treatment of financial intermediation, describing current work by Klimenko et al. [2015] and Brunnermeier and Sannikov [2014e]. Section 4 offers an illustrative example of the required solution methods in the context of a simple model, a simplification of Brunnermeier and Sannikov [2014b]. This section is supported by a technical appendix providing a heuristic outline of solution methods. Section 5 then discusses the substantial agenda for future research opened up by this new ‘post-crisis’ approach to macro-financial modelling. Section 6 concludes.

## 2 Strengths and weaknesses of the new literature

This section provides a general overview of the new literature on continuous time macrofinancial dynamics. Neither the economics nor the solution methods employed in this literature are in themselves especially novel. The contribution comes from combining balance sheet restrictions, in appropriately chosen contexts, with the tools of continuous time stochastic dynamic optimisation. This section therefore proceeds by outlining the economics of this new literature comparing it with an earlier substantial body of research, dating back to the late 1980s, that addresses the aggregate implications of market incompleteness. It also offers a short discussion of the technical strengths and shortcomings of this new approach.

Most of this earlier work focused on the absence of markets for insuring idiosyncratic household labour income risks, a market incompleteness that can reduce the equilibrium real interest rate (Huggett [1993], Aiyagari [1994]) and provides one potential explanation of the incompatibility of the equity market risk premium with complete market models of household consumption-savings decisions (Mankiw [1986]).<sup>4</sup> The particular strand of this work closest to the new macrofinancial dynamics (initiated by Krusell and Smith [1997, 1998]) considers the dynamics of capital accumulation in economies combining uninsurable idiosyncratic shocks to employment with aggregate shocks to the productivity of capital. As with the new continuous time macrofinancial literature there are no analytical solutions, so numerical methods must be applied. A comparison of these two literatures offers useful insight into their respective strengths and weaknesses.

Macrodynamics analysis with incomplete markets is only ever tractable with strong simplifying assumptions. In the presence of market incompleteness, such as limits on individual household borrowing or frictions in access of firms to capital markets, standard aggregation results no longer hold.<sup>5</sup> This calls into

---

<sup>4</sup>See Guvenen [2011] for a detailed review of this literature.

<sup>5</sup>The standard results are those of Gorman [1959], who considers restrictions on utility under which consumption of goods can be expressed as a linear function of wealth allowing

question the appropriateness of widely employed ‘representative agent’ models. Full solution, based on the standard assumptions of complete information and model consistent expectations, requires every decision maker to track the current state and laws of motion of the entire distribution of assets and liabilities across all individual agents. There are therefore at least as many state variables as there are agents in the economy.

The new continuous-time macrofinancial literature sidesteps this challenge of aggregation, reintroducing the representative agent by assuming either that all agents of a particular type are exactly the same, with the same tastes or technology and affected simultaneously by the same shocks (within sector homogeneity); or by assuming that all agents of a particular type can costlessly trade all financial and real assets with each other (within sector market completeness) with often at least some assets also traded between sectors.<sup>6</sup> These strong assumptions have allowed these models to capture qualitative changes in aggregate behaviour that arise when there is a substantial probability of balance sheet constraints binding or coming close to binding, and the possibility of feedbacks that then amplify shocks and generate persistent fluctuations in economic aggregates and asset prices. They do though illustrate one of the main points we draw from our review: this new literature is still immature with much work yet to be done to examine how well its predictions hold in more realistic settings.

The older literature on aggregate productivity shocks and uninsurable labour income deals with this aggregation problem in a quite different way, restricting attention to particular model specifications in which solution can be reasonably accurately approximated by individual agent decision rules based on a small number of summary statistics for the entire distribution of household wealth.

The influential contribution of [Krusell and Smith \[1997, 1998\]](#) was to solve such a model, with two idiosyncratic employment states (employed, unemployed) and two aggregate productivity states (high in boom, low in recession), using a numerical schema which enforced model consistent capital dynamics and demonstrating that the resulting outcome exhibited ‘approximate aggregation’ in the sense that increasing the number of summary statistics for the wealth distribution used by households in their consumption/saving decisions beyond a small manageable number did not affect model outcomes.<sup>7</sup> An entire branch of literature has emerged focused on the numerical accuracy of this and other

---

the choices of a large number of households to be restated as that of a representative consumer; and of [Rubinstein \[1974\]](#) and [Constantinides \[1982\]](#) who examine aggregation in the context of portfolio allocation-consumption decisions. [Constantinides \[1982\]](#) shows that under relatively weak conditions with complete financial markets the decisions of individual consumers can be replaced by that of a composite representative agent. See [Güvener \[2011\]](#) for more discussion.

<sup>6</sup>Similarly strong representative agent assumptions are also imposed in earlier literature on the macroeconomics of financial frictions, including in the influential work of [Kiyotaki and Moore \[1997\]](#) and [Bernanke et al. \[1999\]](#).

<sup>7</sup>The Krusell-Smith algorithm for obtaining model consistent capital dynamics is based on updating a linear rule for the period by period investment in the stock of capital through a regression on the simulated model output from the previous iteration, iteration is continued until the investment rule is model consistent and the accuracy of the numerical solution is judged by the fit of the regression.

alternative algorithms for solving models of this kind (for further discussion see [Algan et al. \[2010\]](#), [Den Haan \[2010\]](#)).

A weakness of the Krusell-Smith algorithm is its model dependence.<sup>8</sup> While it appears to work reasonably well for particular calibrations of the specific model for which it was developed, it is far from clear that it can provide a reliable approximation to the dynamics of the kind that emerge in the new continuous time macrofinancial models we review. One limitation is that it makes no allowance for the resulting dynamic changes in interest rates or other financial asset prices consequent on changes to individual agent balance sheets. Another limitation is that there is no guarantee against the algorithm converging on a ‘wrong’ outcome in which the particular model simulations generated at convergence contain insufficient examples of the balance sheet constraints leading to qualitative shifts in the decisions of households or other agents that in turn substantially influence macroeconomic dynamics.<sup>9</sup>

Another obvious difference is that the earlier literature on macroeconomic dynamics in the presence of market incompleteness follows the dominant practice in macroeconomic modelling of assuming that time is discrete rather than continuous. The choice between discrete and continuous time is, however, less important than might at first appear. It can admittedly be a barrier to understanding.<sup>10</sup> But numerical solution using a computer always eventually requires discretization. Our view is that these two assumptions (discrete vs continuous) are complementary, each with their own strengths and weaknesses. It should be possible to state any of these models using either approach, and the choice then comes down to which is more convenient for solution and communication of results.

Continuous time diffusion has some advantages. Provided that the model can be specified with a small number of heterogeneous agents, tractable solution can be computed using ordinary or partial differential equations sidestepping concerns about the existence of a ‘Markovian’ equilibrium. Another convenience that all paths are continuous so there is no need to be concerned about the possibility of assets or liabilities jumping beyond constrained values.<sup>11</sup> Solution via ordinary or partial differential equations provides an efficient way of captur-

---

<sup>8</sup>For further discussion see [Den Haan \[2010\]](#).

<sup>9</sup>Another way of thinking about these challenges of numerical convergence is that an algorithm of this kind in effect substitutes moments of the distribution of networth, both across individual agents and across time, for the full distribution. If insufficient moments are included then the algorithm may yield a poor approximation to the correct solution.

<sup>10</sup>In Section 4 of this chapter we discuss the technicalities of solution of a simple illustrative example of continuous time macrofinancial modelling, hoping in this way to make this literature accessible to readers who are much more familiar with discrete time modelling. We also recommend as good practice further steps to assist readers become acquainted with these methods. One helpful presentational device, used for example by [Klimenko et al. \[2015\]](#), is to first state a model in discrete time with time steps of length  $\Delta t$  and then derive the limit as  $\Delta t \rightarrow 0$ . Another helpful step is to develop standalone numerical solvers which allow readers to use ‘sliders’ to vary parameters and observe the consequent changes in solutions. The website [www.leveragecycles.lboro.ac.uk](http://www.leveragecycles.lboro.ac.uk) contains examples of such standalone solution software for two of the papers reviewed here, [Isohätälä et al. \[2014b\]](#) and [Isohätälä et al. \[2014a\]](#).

<sup>11</sup>This, however, does not apply to *jump*-diffusion processes.

ing the impact across the state space of constraints on behaviour at or close to boundaries. As discussed in the next section, specification in continuous time also allows the application of the convenient method of asymptotic expansion in order to capture the singularities that can emerge when financial constraints are hit. Finally, specification in continuous time with diffusion also means that decision rules can be expressed in relatively simple terms, namely, as functions of derivatives or partial derivatives of the value function (i.e., marginal values), thereby, providing useful economic intuition that is not so easily obtainable in discrete time.

Discrete time has the advantage that solution can be computed using the well developed and widely used tools of backward recursion. The literature offers a well developed discussion of both the existence and computation of equilibria in discrete time, including for macroeconomic models with incomplete markets.<sup>12</sup> There are larger amounts of available software developed for solution of discrete time models. Solution methods are now well understood both when disturbances are relatively small compared potential constraints so linearisation can be employed and for many non-linear models including several state variables (dynamic stochastic macroeconomic models can now be routinely solved with four or more states). Some forms of lagged response – e.g. the policy response lags resulting from delays in the release of statistical information are more naturally specified in discrete time.

So far our comparison of these two literatures has focused on the technical challenges of aggregation and numerical solution. Comparison of these two literatures also highlights some differences in economic assumptions. One is that in older literature, for example [Krusell and Smith \[1998\]](#), it is individual households who are financially constrained, whereas in [Brunnermeier and Sannikov \[2014b\]](#) it is the representative firm that is financially constrained (as we describe below in Section 3, they are unable to borrow more than the market value of their capital). [Krusell and Smith \[1998\]](#) find that these underlying financial constraints make relatively little difference to aggregate dynamics, whereas [Brunnermeier and Sannikov \[2014b\]](#) find that the constraints substantially reduce output and investment when firm net worth (as a proportion of the market value of the economy’s capital stock) falls close to zero.

But perhaps the most important advantage of the radical simplifying assumptions made in the continuous-time macrofinancial literature is the wide

---

<sup>12</sup>See [Krueger and Kubler \[2008\]](#) for a short overview, including discussion of the challenge of computing solution in a small number of state variables when it is no longer possible to obtain solution using contraction mapping theorems (theorems closely related to the aggregation results of [Constantinides \[1982\]](#) and the implied correspondence between market equilibrium and an equivalent central planning problem). The algorithm of [Krusell and Smith \[1998\]](#) is the most widely cited example of such methods applied in the context of incomplete markets. [Ljungqvist and Sargent \[2000\]](#) chapter 17 offers a number of other examples of solution for incomplete market economies and [FENG et al. \[2014\]](#) and [Guerrieri and Iacoviello \[2015\]](#) for two recent proposed methods for recursive numerical solution of incomplete market models in discrete time. Tractable solutions of these models are described as ‘Markovian’ because the stochastic dynamics can be expressed in terms of the equations of motion of a limited number of state variables.

range of issues that can then be addressed. This will become clearer from our review of individual models in the next section that explore the impact of constraints on households, firms and intermediaries for the dynamics of asset prices, output and investment. As we discuss in Section 5, there is scope for considerable further work of this kind on the dynamic consequences of market incompleteness for a range of other aggregate economic variables, including employment, price setting, government finances and macroeconomic policy. The price paid for these advances is not insubstantial, a clear data discrepancy at the microeconomic level since not all firms or all households are able to trade amongst each other to achieve common ratios of debt to assets (i.e. the assumption that each sector can be replaced by a single representative agent is not a realistic assumption in the context of incomplete markets). This though purchases valuable new understanding of a range of macroeconomic phenomena that are attracting attention in the wake of the global financial crisis.

### 3 A review of some recent continuous-time macro-financial models

In this section we review some recent continuous-time macrofinancial models. Our discussion covers three prominent papers that have attracted widespread attention together with three papers of our own. It is organized as follows. Section 3.1 discusses how continuous time models have been used to model the dynamics of assets prices, including the return on risk-free assets and the premium on risky assets. 3.2 reviews implications of the dynamic allocation of productive capital for financial stability. Section 3.3 discusses extensions of these models to the explicit treatment of the banking sector.

#### 3.1 Capital constraints and asset pricing

The series of papers developed by He and Krishnamurthy [2012, 2013] (hereafter, HK(2012) and HK(2013)) explores how market incompleteness affects the risk premium on risky assets in a Lucas Jr. [1978]-type endowment economy in which cash flow yields (dividends) on risky asset follow a random walk. The key distinguishing assumption of these models is that risky assets are held only by specialist financial intermediaries subject to agency frictions similar to those modelled in Holmstrom and Tirole [1997]. Incentive compatibility (i.e., avoiding the mismanagement of assets or ‘shirking’) requires that these intermediaries must finance their investments with a minimum proportion of their own equity. When intermediary capital is scarce, this equity capital constraint binds and works as a channel of amplification of fundamental shocks to intermediary assets and net worth, increasing the volatility of returns and the risk premium earned from investment in the risky asset.

Here we focus on the model developed in HK(2012) (the other model is similar). HK(2012) model the economy in which there is a single risky, non-tradable asset of a *fixed size* and the market price  $P_t$  that reflects the expected



discounted value of dividend streams.<sup>13</sup> The asset generates a stochastic flow of dividends  $D_t$  per unit of time, that evolves as a Geometric Brownian motion with a constant drift and volatility  $\sigma$ . There exists also a risk free asset (bonds) in zero net supply and interest rate  $r_t$ , i.e. there is the possibility of lending between the households and specialists. The risky asset's risk premium is then given by

$$\pi_{R,t} = \mathbb{E}\left[\frac{D_t dt + dP_t}{P_t}\right]/dt - r_t.$$

There are two classes of investors: *specialists* managing financial intermediaries that play the role of investment vehicles and *households* who delegate investment decisions to specialists, as they have no direct access to investment technologies (i.e., there is market segmentation). In this and all following models we review in this section, all agents belonging to a particular group are identical. Such a simplification is key for obtaining tractable solutions, as it allows working with a representative agent making the optimal decisions based on observations of her own level of wealth and (typically) a unique aggregate state.

Both specialists' and households' wealth is invested in intermediaries. The optimal contract between households and specialists determines  $\beta_t \in [0, 1]$  – the specialist's share of investment in the risky asset and hence, after allowing for a fee  $K_t dt$  that may be paid to specialists for managing entrusted funds, their claim on dividend income. Specialists choose the total volume of investment in the risky asset,  $\mathcal{E}_t$ , and make a working/shirking decision unobservable to households. As in [Holmstrom and Tirole \[1997\]](#), shirking reduces the cash-flow from risky assets by  $X_t dt$  but enables specialists to collect private benefits  $B_t dt$  which are assumed to be proportional to the reduction in the asset cash-flow, caused by shirking:

$$B_t dt = \frac{1}{1+m} X_t dt,$$

where the inverse of  $m$  captures the magnitude of agency frictions.<sup>14</sup>

The incentive contract preventing shirking places restrictions on outside equity financing. Namely, the households' equity stake must be limited to a fraction of the total risky investment that depends on the magnitude of agency frictions, which leads to the following equity capital constraint:

$$\mathcal{E}_t^h \leq m\mathcal{E}_t. \tag{1}$$

Put differently, to abstain from shirking, specialists must maintain some 'skin in the game', whose proportion is increasing with the magnitude of agency

<sup>13</sup>By contrast, the models by [Brunnermeier and Sannikov \[2014b\]](#) and [Brunnermeier and Sannikov \[2014e\]](#) that we review below enable the asset to be traded among two classes of agents, which allows capturing the impact of "fire sales" on asset prices and track their feedbacks into the dynamics of agents' wealth.

<sup>14</sup>A further assumption, introduced in order to avoid the challenges of solving for punishment and reward strategies as a dynamic game, is that the contract between households and specialists is lasting only from  $t$  to  $t + \Delta t$  after which the relationship between household and specialist is broken and each household is paired with a new specialist. This means that the equity constraint emerges as the solution to a static bargaining problem.

frictions. In terms of the sharing rule, the above constrain implies that

$$\beta_t^* \geq \frac{1}{1+m}, \quad \text{for } K_t \geq 0,$$

with equality when  $K_t > 0$ .

To obtain a closed form solution He and Krishnamurthy [2012] assume that both specialists and households have log-preferences over instantaneous consumption. With this assumption, the value function of any representative agent is additively-separable and can be written in the following form:

$$\frac{1}{\rho^i} \log(W_t^i) + Y_t^i,$$

where  $\rho^i$  is the discount rate of the agent  $i = \{s, h\}$  (specialist and household, respectively),  $W_t^i$  is the wealth of the agent  $i$  and  $Y_t^i$  is the function of the aggregate wealth and dividends, which are two state variables in this setting.

Due to the above property of the value function, the portfolio and consumption choices of agents are almost trivial. In particular, agents continuously consume an amount proportional to their net worth, where the consumption rates are given by their respective discount factors, i.e.,  $c_t^i = \rho^i W_t^i$  and the optimal exposure to risky asset is given by the mean-variance portfolio choice, yet, with a slight twist for households for whom the effective asset risk premium is reduced by the scaled intermediation fees  $k_t = K_t/\mathcal{E}_t^h$ .

Finding the unique equilibrium of this model requires solving for three processes – risky asset price  $P_t$ , riskless interest rate  $r_t$  and scaled intermediation fees  $k_t$  – compatible with the individual maximization and market clearing conditions.  $P_t$ ,  $r_t$  and  $k_t$  are the functions of the unique state variable – the aggregate specialists’ wealth scaled by aggregate dividends,  $w_t \equiv W_t^s/D_t$ .

Depending on the level of the scaled specialists’ net worth, at each moment of time the economy can find itself in one of two regimes: if scaled specialist net worth  $w_t$  exceeds a critical threshold  $w^c \equiv 1/(\rho^h m + \rho^s)$  then the solution is in an unconstrained regime in which the incentive constraint (1) is slack; otherwise the solution is in a constrained regime in which the incentive constraint (1) is binding.

In the unconstrained regime where the wealth of the specialist financial intermediaries is relatively high, the risk premium is constant and households pay zero fees for intermediation. There is no borrowing or lending (with the implicit ‘risk-free’ rate of interest  $r_t$ , a wealth dependent weighted average of the discount rates of households and specialists, that declines as  $w_t$  increases). Holdings in the risky asset ( $\beta_t$ ) are proportional to agent wealth. The price volatility of the risky asset is constant and is equal to the volatility of the dividend cash-flow, i.e.,  $\sigma_{R,t} = \sigma$ .<sup>15</sup>

---

<sup>15</sup>This property emerges essentially due to the absence of leverage in the unconstrained region. In the models we review next, the endogenous volatility is affected by the changes in leverage/feedbacks from asset prices and does not remain constant even when the capital/leverage constraints are far from binding.

In the constrained region, in which the wealth share of the specialist financial intermediary are relatively low (below  $w^c$ ), the equity constraint binds. Their relatively low level of wealth means that the specialists must borrow from households in order to maintain their required share of holdings of the risky asset. The volatility of the risky asset (endogenous volatility),

$$\sigma_{R,t} = \sigma \left[ \frac{(1+m)\rho^h}{(\rho^h m + \rho^s)(1 + (\rho^h - \rho^s)w_t)} \right] > \sigma,$$

then drives the level of both the risk premium and of intermediation fees in the constrained regime. These are both always higher than in the unconstrained regime, but decreasing with the scaled specialists wealth until the threshold between the two regimes is reached. The risk free interest rate (at least for the chosen parameterisations) also exhibit a different pattern than in the unconstrained region: namely, it becomes an increasing function of specialist wealth, i.e., in the constrained regime, the lower the specialist wealth, the higher the valuation placed on risk-free assets.

The HK(2012) model predicts that intermediaries only borrow in the constrained regime, otherwise intermediaries are unleveraged. In order to generate leverage in the unconstrained regime and so better match the data, HK(2013) amend their earlier model by introducing household labour income uncertainty and an exogenous demand by households for holding a minimum proportion of wealth in the form of risk-free lending to specialists.<sup>16</sup> Solution is now numerical, not closed form. Parameters are chosen so that, absent of any constraints, the risk-tolerant households hold all their wealth in the form of risky assets and as a result the equity constraint on specialists binds approximately fifty percent of the time. With this set up the model does a fairly good job of reproducing the dynamics of risk-premia during financial crises, with a ‘half-life’ (an expected decline of the risk-premia relative to unconstrained levels of 50%) of about 8 months.

Further insight into the impact of leverage constraints on the pricing of *risk-free* assets is provided by Isohätälä et al. [2014a](hereafter, IKMR(2014)). They consider the interaction of two household sectors receiving an endowment income subject to offsetting shocks: a positive shock to income and an equal and opposite negative shock in the other. There is a single consumption good.

<sup>16</sup>Note that households are no longer infinitely lived. Instead, HK(2013) consider the continuous time limit of an ‘overlapping generations’ setting in which households born and then die almost instantaneously. Specifically, households are born at  $t$  with a labour income proportional to the dividend on risky assets, and allocated in proportion  $\lambda : 1 - \lambda$  to one of two classes ‘risk-averse households’ whose wealth must all be held in the form of loans to specialists, and ‘risk-tolerant’ households who are free to choose the proportion of their wealth invested in risky assets managed by specialists and in loans to specialists. Households consume at  $t$  in order to maximise a utility function log linear in current consumption and an end-period bequest at  $t + \Delta t$  randomly allocated across the next generation (labour income is of infinitesimal size relative to inherited wealth and utility is logarithmic, implying that household consumption is a fixed proportion of their inherited wealth, the random allocation avoids the necessity of tracking the distributional impact of the allocation to risk-averse and risk-tolerant classes).

Cumulative income is a diffusion process with infinite local variation (the standard deviation of income over a period  $t$  to  $t + \Delta t$  is proportional to  $\sqrt{\Delta t}$  while expected income is proportional to  $\Delta t$ ). While there is no insurance contract that protects against this income uncertainty (the assumed market incompleteness), households can still smooth consumption by borrowing and lending from each other, subject to a constraint of some maximum level of borrowing. Both households seek to maximise a standard objective, the discounted expected utility with instantaneous ‘CRRA’ utility, i.e. constant relative risk aversion and intertemporal elasticity of substitution. One household is relatively impatient discounting consumption more than the other. The underlying microeconomics are not further developed although the constraint on borrowing might represent the possibility of repudiating debt and instead obtaining some alternative subsistence income.

These strong modelling assumptions yield a simple and intuitive outcome with buffer stock saving very similar to that predicted by standard microeconomic models of household precautionary saving. Both household consumption ( $c$ ) and expected saving, i.e. expected endowment and financial income net of consumption ( $a + r(w)w - c(w)$ ), are monotonic functions of wealth  $w$ , with consumption increasing and savings decreasing with  $w$ . Here wealth  $w$  is simply the net claims of impatient households on households in the other patient sector, so  $-w$  is a measure of impatient household leverage ( $w$  is almost always negative). Expected saving by the impatient household sector is positive whenever leverage is above a target level (buffer stock saving). The novel macrofinancial feature of the model is that the real interest rate  $r = r(w)$  adjusts to ensure goods market clearing i.e. total consumption by the two sectors equals their total endowment, with potentially large but relatively short lived declines of real interest rates whenever income shocks increase the leverage of impatient households close to their maximum levels of borrowing. This is thus a setting in which a financial problem (overleverage) is corrected in large part through adjustment of market prices (a temporary period of low real interest rates supporting deleveraging towards a long term desired level of borrowing) rather than through reduction of consumption.

There are sharp contrasts between the investigations of IKMR(2014) and those of HK(2012) and HK(2013), but also striking similarities. Differences include: the specification of uncertainty (in HK(2012) this is a diffusion process for the productivity of assets, while in IKMR(2014) this is a diffusion process for cumulative endowment); the distinction between sectors (in HK(2012) this distinction is between specialist asset managers and outside investors, while in IKMR(2014) this distinction is between impatient borrowing households and patient lending households); the focus of the analysis (in HK(2012) this is the pricing of risky assets while in IKMR(2014) it is the pricing of risk-free instantaneous borrowing); and in the treatment of household optimisation (in HK(2013) the OLG setting abstracts from all issues intertemporal cash management while in IKMR(2014) both agents address a fully intertemporal optimisation).

The key similarity is that in both settings asset prices adjust so as to restore balance sheets fairly quickly towards long run expected values. Periods

of distress are relatively short lived. Following an initial disturbances, after a few months wealth shares gravitate back towards the steady state distribution (the ‘ergodic density’ across wealth). In particular, in all these settings risk-free interest rates decline dramatically during periods of extreme financial stress and this assists the process of deleveraging (see HK(2013) Figure 3 and IKMR(2014) Figure 6).

### 3.2 Models of output and investment *without* an explicit banking sector

In this section we describe the model of Brunnermeier and Sannikov [2014b] (hereafter, BS(2014a)) that focuses on the role of net worth in the allocation of productive capital in the economy and its implications for the dynamics of output and investment. In BS(2014a), capital is traded between more productive, risk-neutral, impatient *experts* and less productive, risk-averse, more patient *households*.<sup>17</sup> The productivity of capital follows a diffusion process, as in the complete market setting of Lucas Jr. [1978] and employed by HK(2012) and HK(2013). Also as in HK(2012) and HK(2013), the state of the economy is described by the single state variable, the ratio of expert net worth to household net worth.

As well as sharing in the risky investment opportunity, households may invest in risk-free debt issued by experts. Debt contracts are short-term, and experts continuously adjust their level of debt in order to balance a desire to consume early (impatience) against the potential costs of incomplete insurance against productivity shocks.<sup>18</sup> While BS(2014a) model features no explicit leverage or capital constraint, a constraint emerges implicitly because reductions in the market value of capital limit the ability of firms to borrow. The absence of a market for insuring against fluctuation against in the productivity of capital and hence net worth mean that, in effect, debt is subject to a collateral constraint, not unlike that featured in Kiyotaki and Moore [1997].

This implicit need for collateralisation is consistent with a standard paradigm of financial intermediation literature considering financial intermediaries (particularly, banks) as the providers of safe and liquid investment opportunity (demand deposits), given that some economic agents may have strong preferences for this kind of investment (see, e.g. Diamond and Dybvig [1983], DeAngelo

<sup>17</sup>The title of their paper ‘A macroeconomic model with a financial sector’ needs some explanation. Their productive experts who engage in investment and production could be real economy firms but on this interpretation their model does not have a financial sector at all; the title reflects their assumption that the assets held by these firms can be freely bought and sold between experts and households suggesting that they actually have in mind a very similar setting to that of HK(2012) and HK(2013) and that their experts are financial intermediaries who manage tradeable assets (see Brunnermeier and Sannikov [2014c] the online appendix to Brunnermeier and Sannikov [2014b] , where an equivalent version of their model distinguishing financial intermediaries and productive firms is discussed).

<sup>18</sup>In BS(2014a) setting experts do not need to maintain any liquid reserves, as arises in structural corporate finance models in which there are costs of adjusting liabilities (see e.g., Bolton et al. [2011]).

and Stulz [2013] for the arguments along this line). In the environment in which financial intermediaries act as the liquidity providers, while facing financial frictions, this feature creates a role for intermediaries' net worth as a loss-absorbing buffer that is needed to guarantee the safety of debt issued to households.

The productivity of capital in the BS(2014a) economy fluctuates over time according to a diffusion process with standard deviation  $\sigma$ . This in turn alters both expert net worth and the share of expert net worth (a positive shock to productivity of capital increases the net worth of both experts and households; as long as experts are leveraged then this also increases the share of expert net worth). They assume in addition that new physical capital can be built via an investment technology with adjustment costs.<sup>19</sup> The main friction in this economy refers to the fact that experts do not have 'deep pockets' and cannot raise outside equity (this, in fact, can be interpreted as the extreme form of the agency problem present in HK(2012)). As a result, a decline in net worth caused by negative productivity shocks increases the effective risk aversion of experts. This induces them to 'self-insure' by shrinking the scale of operation (simultaneously, reducing the volume of debt) and selling capital to less productive households, which ultimately leads to the reductions in output. Moreover, sales of capital by experts depress the asset price, which, in turn, feeds back into the dynamics of net worth, thereby amplifying the impact of the adverse productivity shock.<sup>20</sup> We illustrate the detailed modelling of this mechanism in Section 4 by using a simplified version of BS(2014a) model.

In equilibrium, the dynamics of capital prices, capital and experts' net worth, as well as the optimal consumption and investment decisions of agents (and their respective holdings of capital), can be characterized as the functions of a single state variable – the *experts' share in the total net worth*. Expression in terms of a single state variable is possible due to the linearity of the agents' value functions in individual agent's net worth (scale-invariance property). The optimal consumption decisions of experts (who face the non-negative consumption constraint) are determined by the marginal value of their net worth, which is a decreasing function of the state.

In the baseline model explored by BS(2014a) the optimal consumption pattern is similar to the optimal payout policies emerging in many (partial equilibrium) corporate finance models: as long as the value of the state is relatively low and thus an expert's net worth is highly valuable, it is optimal to retain earnings; however, once the marginal value of the state falls to one, experts consume all positive profits so as to maintain the state at the level associated to the unit marginal value.<sup>21</sup> Such a 'barrier-type' consumption strategy determines

<sup>19</sup>As shown in our illustrative example in Section 4, this feature is not crucial. Aside from the investment impact, the principal model results hold when this channel is switched off.

<sup>20</sup>By contrast, if experts could costlessly issue new equity, there would be no capital traded and all capital would instead be held by experts. The price of capital then would be constant and would reflect the expected discounted value of the perpetual output stream under the more productive technology.

<sup>21</sup>BS(2014a) also present an alternative version of their model in which both households and experts have logarithmic preferences (once again this choice of preferences simplifies solution because the value function is then additively separable and optimal consumption is a fixed

the upper bound of the state. The fluctuations of the state between zero and the consumption boundary drive the effective risk aversion of experts and thus the equilibrium allocation of capital in the economy: as long as the share of expert’s net worth is relatively high, all capital is concentrated in the experts’ hands; however, below a certain critical level, the fraction of capital held by experts is always lower than one and is an increasing function of the state.

An important effect captured by BS(2014a) model is extended persistence of the aggregate shocks, a consequence of the response of experts to the incomplete opportunities for insurance against productivity risks. As the share of expert net worth declines and an increasing proportion of capital is sold to and managed by households, it becomes relatively difficult for experts to rebuild net worth via retained earnings. This means that for at least some parameter combinations the economy may spend quite a lot of time in recession states with low asset prices and a large fraction of capital concentrated in the hands of less productive agents. This property manifests itself via the ergodic density of the state being spiked in the neighborhood of its lower boundary.<sup>22</sup>

A point that is not entirely clear in BS(2014a) is the respective importance of the ‘self-insurance’ effect and amplification effect generated by the endogenous volatility of the price of capital in generating these protracted dynamics. Certainly it is possible to get similarly protracted dynamics without endogenous price volatility. This point is illustrated by the closely related partial equilibrium model of Isohäätä et al. [2014b] (hereafter, IMR(2014)). In this paper identical impatient firms manage a risky asset and the diffusion process affects aggregate accumulated cash flow rather than the productivity of capital. Moreover, in order to reduce risk exposure, capital is rented by experts to patient households rather than sold. Preferences are the same as in the baseline model of BS(2014a), i.e. both experts and households have linear preferences but experts are subject to a ‘non-negativity’ constraint on consumption, i.e. in effect a prohibition on issue of new equity capital. Unlike BS(2014a), this model also parameterises the deadweight costs of equity issuance. The merit of this model specification is its relative simplicity and tractability, as there is no need to take any account of the complications of asset pricing or optimal portfolio allocation.

The optimal risk exposure chosen by a representative firm in IMR(2014) depends on its leverage and is implemented via the optimal rental decisions: at each moment of time, firms may unload some risk by leasing a fraction of capital to less productive households in return for a fee (assumed equal to the productivity of capital in the hands of households).

The IMR(2014) economy exhibits a very similar behaviour to the one that emerges in the BS(2014a) setting, albeit without price volatility: under the combination of relatively high uncertainty and large financing frictions (i.e., high recapitalization costs) the economy spends a lot of time in the recession

---

proportion of the market value of agent net worth). As long as experts are more impatient than households this generates very similar dynamics to the baseline model, but now with positive expert consumption (i.e. some payment of dividends) for all values of the state variable.

<sup>22</sup>See our Section 4 and Appendix 4.2.2 for discussion of the calculation of this ergodic density.

states characterized by low experts' net worth and a large fraction of capital concentrated in the hands of less productive households.

### 3.3 Models of output and investment *with* an explicit banking sector

In this section we consider two continuous time macrofinancial models with a more explicit treatment of the banking sector. The first is that of [Klimenko et al. \[2015\]](#) (hereafter, KPR(2015)) which distinguishes the banking sector from the productive sector in order to address the role of bank capital in the fluctuations of credit and output. This model captures a complementary channel for output distortions that works via the adjustments of credit volumes in the economy.<sup>23</sup> The second is the more ambitious modelling of [Brunnermeier and Sannikov \[2014e\]](#) (thereafter, BS(2014b)) who develop a monetary analysis in which net worth limits the ability of banks to create 'inside-money' and thus affects both the real economy and the nominal price level. While the dynamics of risk-premia and of output and investment generated by these models are similar to those reviewed earlier in this section, the explicit treatment of banking allows a much fuller discussion of policy instruments, including bank capital regulation, as well as monetary and fiscal policy. We should emphasise that work on both these models is ongoing – when eventually published in peer-review journals they could have evolved substantially from the versions we discuss here. Still we think these two models are worth highlighting as examples of where the continuous-time macrofinancial literature may be heading in the future.

KPR(2015) study the impact of bank capital on the cost of credit in the economy where the firms' projects are financed exclusively via bank loans. The model shares some similarities with [HK\(2012\)](#) and [BS\(2014a\)](#). Again there are two classes of agents, in this case relatively impatient banks and relatively patient households. Banks are risk-neutral and by implication (since they are maximising expected utility) have an infinite intertemporal elasticity of substitution. Households also have a infinite intertemporal elasticity of substitution, with a time discount rate of  $r$ , and are willing to provide unlimited deposits at an interest rate  $r$  but only as long as there is no risk of any loss on deposits.

The economy is subject to aggregate shocks, which affect the firms' default probability (and cannot be diversified) and ultimately the banks' profits. Cumulative profits (retained earnings) are described by a diffusion process with drift and diffusion proportional to the volume of bank lending. The firms' demand for credit is an exogenous decreasing function of the nominal loan rate  $R_t$ , where the latter is determined at equilibrium as a function of aggregate bank capitalisation  $E_t$ . Banks continuously adjust the volume of lending, as well as the volume of deposits they collect.<sup>24</sup> However, their capacity to adjust book equity (net worth) is limited, because banks face a proportional deadweight cost

---

<sup>23</sup>[Phelan \[2014\]](#) also introduces the banking sector in a continuous-time macrofinancial model, however, without explicitly modelling this lending channel.

<sup>24</sup>Both loans and deposits are assumed to be short term, and the full depreciation of productive capital is allowed.



$\gamma$ , when raising new capital. This parameterisation is similar to that employed by IMR(2014), albeit the additional assumption that  $\gamma$  is a decreasing function of bank aggregate capitalization which captures the idea that recapitalizations become costlier when the banking sector is in distress.

A convenient property of the model is a linearity of the value function of an individual bank in the level of its book equity. Banks in KPR(2015) economy behave competitively in both loan and deposit markets and make the same decisions. As a result, all banks' decisions (lending, recapitalization and dividend payouts) are driven by their individual market-to-book value of their equity, which in turn is a function of aggregate bank capitalization.<sup>25</sup> Aggregate lending, recapitalisation and dividends are then functions of aggregate variables and the level of aggregate bank capital that follows the Markov diffusion process reflected at two boundaries: banks are paying dividends at the upper boundary and recapitalize as soon as the book equity is depleted. In other words, to reduce the frequency of costly recapitalizations, banks maintain equity buffers, whose target size is optimally chosen so as to maximize shareholder value. As a consequence of the risk-neutrality of banks, dividend behaviour is of the same 'barrier control' form as in the baseline model of BS(2014a) and in IMR(2014) with payments only when bank equity climbs to an upper level  $E_{\max}$ . There is also recapitalisation at a lower barrier  $E_{\min}$ , which turns out to be zero in the competitive equilibrium.

The value function that emerges from optimal solution represent the expected value of the bank shareholders' claim and can be expressed as the product of a book equity times the market-to-book value. Such a structure of the value function helps understand the source of a positive lending premium (the margin between loan and deposit rates  $R_t - r > 0$  with equality at the upper dividend paying boundary) emerging from this model: any negative shock to bank earnings not only depletes book equity (directly reducing lending capacity) but is further amplified via a decline in the market-to-book value. The effect of a positive shock, in contrast, is dampen via the same channel. As bank equity declines, bank shareholders become effectively more and more risk-averse (even though their preferences are risk-neutral) and demand a strictly positive premium in order to lend to the real sector. This lending premium (as well as the loan rate itself) is a decreasing function of aggregate bank capitalization. Via this lending premium channel, the reductions in aggregate bank capitalization ultimately translates into a higher cost of credit and a reductions of the firms' demand for bank loans and thus the decline of output.

The explicit dynamics of the loan rate that emerges in KPR(2015) model (with the further assumption that the deposit rate  $r = 0$ , the drift and volatility of the loan rate can be obtained in closed form) allows for a tractable analysis of the long run behavior of the economy using the loan rate as the state variable. As in BS(2014a) or IMR(2014) models discussed above, the dynamics can be described by the ergodic density function of the state. The analysis of the

---

<sup>25</sup>The distribution of equity capital across individual banks then has no impact on economy wide outcomes.

ergodic density patterns shows that the economy spends a lot of time in the states with the lower endogenous volatility and, under strong financing frictions, can get trapped in the states with low bank capitalization, a high loan rate (low lending) and thus low output.

Due to its simple dynamics, the model also turns out to be suitable for a tractable welfare analysis. The latter shows that the competitive equilibrium is not constrained efficient. In particular, a social planner facing the same frictions as banks would choose lower volumes of lending than emerge in the competitive equilibrium. This effect is driven by a pecuniary externality: competitive banks fail to internalize the equilibrium impact of their individual lending decisions on the loan rate that feeds back into their expected profits. Moreover, competitive banks recapitalize too late and distribute too many dividends than would be optimal from the social perspective. Recognizing inefficiencies of the competitive equilibrium motivates the analysis of the effect of different regulatory tools, among which a combination of mild capital requirements and a tax on bank dividends turns out to be particularly efficient.

BS(2014b) also develop a model in which the experts are financial intermediaries or banks. The basic assumptions are that banks have a superior monitoring technology than households (in this respect their setup is similar to that of [Diamond \[1984\]](#) and [KPR\(2015\)](#)).

Their goal is however much more ambitious than that of [KPR\(2015\)](#). The bank share in aggregate net worth (the usual state variable) determines the extent to which they can issue short term liabilities (inside or ‘i’ money) and hence drives aggregate macroeconomic dynamics, both real economy output and investment and nominal pricing.

Like BS(2014a), BS(2014b) model considers two classes of agent (households and experts) but now with the same rate of time preference. The experts are now financial intermediaries distinguished because the monitoring technology of intermediaries allows them to achieve superior performance from investment in a subset set of available technologies. Banks also benefit from diversification because they can invest in many technologies. Households in contrast can invest only in a single technology (at any point in time). The inability of households to diversify idiosyncratic risk again creates a demand for holding monetary deposits, which in this model are risky – because of the risk of changes in the nominal price level – but still carry a lower risk than any other technology in the economy.

As in BS(2014a), experts’ net worth serves as a loss absorbing buffer. Again this because markets are incomplete and experts cannot fully insure against fluctuations in the productivity of capital. Their net worth then affects the level of ‘inside money’ (i.e. bank deposits). This value is determined by a simple equilibrium mechanism: when negative shocks deplete the experts’ net worth, in order to reduce exposure to further shocks, they shrink their balance sheet by selling capital (loans to end-borrowers) to households. Due to the balance sheet adjustment, this automatically leads to the reduction of their deposit taking capacity, i.e., the supply of inside money shrinks. However, the households’ demand for deposits (money) remains almost unchanged, and hence the ‘price’

of money in terms of goods ( $p$ ) must rise at the same time. Thus a contraction of intermediary net worth both reduces the price of capital in terms of goods ( $q$ ) and increases the price of money. A rise in the nominal price of money is a fall in the price of goods, so this becomes a model of *disinflation* (assuming that monetary policy i.e. the supply of outside money, remains fixed).

The BS(2014b) model is a promising framework for a tractable analysis of macroprudential policies and both orthodox and unorthodox monetary policy. It is though difficult to relate their model to the widely accepted ‘new-Keynesian’ treatment of monetary policy widely employed in DSGE modelling. In the ‘new-Keynesian’ world money stocks, indeed all balance sheets, are essentially irrelevant, the main market friction is sluggishness of price adjustment usually determined in the optimisation setting of Calvo (Calvo [1998]) by assuming a fraction of price-setters in imperfectly competitive final goods markets can readjust prices at any point in time (without this feature DSGE models would exhibit price-neutrality, nominal pricing and monetary policy would then be entirely irrelevant to the real economy). In conventional DSGE stocks of money (as opposed to monetary policy) play no role at all.

There are of course many macroeconomic models in which the stock of money does play an essential role. These include many models in which money is required as a means of payment, either using the relatively ad-hoc mechanism of a ‘cash-in-advance’ constraint (Lucas and Stokey [1987]) and also search models in which money provides a solution to the problem of exchange between anonymous parties who have no mechanism to commit to contractual agreements (for example, the relatively tractable model of Lagos and Wright [2005]).

The role played by money in BS(2014b) is not a means of payment but a store of value. In this respect its role is comparable to that in the many overlapping generation models of money originating with Samuelson [1958]. The simplest example is the two period overlapping generations endowment economy with a single non-storable good. Without money younger generations are unable to lend to or borrow from the current older generation at period  $t$  in order to consume less or more than their period  $t$  endowment, the problem being that the older generation are no-longer around to receive or make repayment in the following period  $t + 1$ . The equilibrium is autarky with each generation consuming its own current endowments. With standard assumptions about preferences there is though an alternative welfare improving equilibrium (at least one) with ‘money’. For example young generations at  $t$  anticipating a large decline in their future endowment may save for old age by acquiring money. In the subsequent time period  $t + 1$  (when they themselves are old) they spend this money acquiring goods from the new younger generation. Money serves as a store of value and allows exchange to take place because of the belief that it will have an exchange value for goods in each following period. Such an equilibrium exists provided that there is no terminal time at which a new generation is no longer born and money has no value.

The demand for money as a store of value in BS(2014b) is different from that in these overlapping generation models, arising because of the risk diversification available to households from holding money. Still, as in the substantial

literature on overlapping generations models with money, this basic model in which money serves as a store of value can be extended to investigate several issues in monetary policy. Government can alter the equilibrium outcome by issuing an alternative ‘outside money’ as an alternative store of value, entirely equivalent from the perspective of households to inside money issued by intermediaries. Government can also offer interest on this outside money and issue long term bonds. Overlapping generations models of money have been used to explore many monetary issues, including the distributional and efficiency impact of different monetary policy rules (for example providing support to the Friedman rule that dynamic efficiency requires that the supply of outside money should contract, and its value increase, at a rate equal to the equilibrium rate of interest).

The major difference and the key contribution of BS(2014b) is that their setting incorporates business cycle fluctuations and so they are able to consider the role of these various monetary policies not just in steady state, but also as a tool for countering macroeconomic fluctuations through the redistributive effect of altering the distribution of net worth between creditors (in their case households) and debtors (in their case financial intermediaries). Policies which redistribute wealth from debtors to creditors following large shocks can help limit the occurrence and duration of extended downturns (deflations) in which output contracts, the price of money  $p$  is high and the price of capital  $q$  and hence investment is low.

It is clear that there is considerable scope for further research, investigating the robustness of these BS(2014b) findings in a range of other settings. It is possible that similar results could be obtained using other models of ‘inside money’.<sup>26</sup> The question of how to integrate market incompleteness and balance sheet constraints with conventional models of monetary policy remains a central issue for future research and continuous time macrofinancial models, building further on the work of BS(2014b), may yet provide considerable further insight.

## 4 An Illustrative Example: Output in general equilibrium

The purpose of this section is to present a simple and tractable example of a continuous-time macrofinancial model, in order to illustrate both methods of solution and some of the insight that can be obtained from this kind of model. The model we present here is essentially that of BS(2014a), but slightly simplified in that there is no investment. The solution method we apply to solve this model differs from the one employed in the original BS(2014a) model, but leads to the same results.

We develop this example with three objectives in mind: first, it shows how financing constraints mathematically appear in continuous time general equilib-

---

<sup>26</sup>It is noteworthy that many of the BS(2014b) results were originally obtained using a quite different underlying model of risks to bank asset returns, based on Poisson shocks, see Brunnermeier and Sannikov [2014d].

rium models; second, it gives a quick recipe for numerically solving such models; finally, it provides a concrete illustration of how such a model can, at least under some parametrization, explain persistence of fundamental shocks reflected by a protracted reduction of output. The Appendix to this chapter provides a short heuristic summary of the mathematical solution methods used in this literature, and further technical references containing a more rigorous presentation of these methods.

## 4.1 Model

In this illustrative example we consider a hypothetical economy that consists of two types of agents, experts and households (we will use an overbar to denote state variables and parameters corresponding to households). A representative expert (household) is characterized by two state variables: cash  $c$  ( $\bar{c}$ ) and capital  $k$  ( $\bar{k}$ ). Cash holdings earn interest at a constant exogenous rate  $r$ , while capital gives production yields at rates  $a$  and  $\bar{a}$ . Negative cash holdings are interpreted as debt. Agents consume their wealth at rates  $\kappa$  and  $\bar{\kappa}$  that are to be determined by maximising appropriate objective functions. Experts and households are identical, except for the following three differences: (i) households are less productive,  $\bar{a} < a$ , (ii) households are more patient than experts, which is captured by the difference in their respective discount rates  $\bar{\rho} \equiv r < \rho$ , and (iii) their consumption is not constrained, whereas an expert must have a non-negative consumption, *i.e.*,  $\kappa \geq 0$ .

Capital can be freely traded between experts and households at a stochastically varying price  $q_t$ . Capital does not depreciate, but is subject to productivity shocks with an amplitude  $\sigma$  per unit capital and square root unit time. At equilibrium, market for capital and debt must clear.

Under the above assumptions, the expert cash and capital follow the stochastic differential equations (here for experts only, analogous equations hold for households)

$$dc_t = (ak_t + rc_t - q_t\tau_t k_t - \kappa_t) dt, \quad (2a)$$

$$dk_t = \tau_t k_t dt + \sigma k_t dz_t, \quad (2b)$$

where  $\tau_t$  is the rate at which the agent trades capital (positive  $\tau$  buys, negative sells) and  $dz_t$  captures the aggregate productivity shocks. The capital price is supposed to be stochastic and follows the equation

$$dq_t = \mu_t^q q_t dt + \sigma_t^q q_t dz_t, \quad (3)$$

with initial data  $q_0$  and where the drift and diffusion functions  $\mu_t^q$  and  $\sigma_t^q$  are some functions of time to be determined in equilibrium.

Since the capital trade is unconstrained, the agents are free to allocate whatever proportion of their net worth,  $n_t = c_t + q_t k_t$ , between the risk-free asset and capital. Let  $\varphi_t = q_t k_t / n_t$  denote the proportion of an agent's net worth invested in capital (note that  $c_t = (1 - \varphi_t) k_t$ ). Applying the Itô's Lemma to  $n_t$

[Technical Appendix A.1, Eq. (37)], we get

$$dn_t = \left[ r + \left( \frac{a}{q_t} + \mu_t^q + \sigma \sigma_t^q - r \right) \varphi_t - \lambda_t \right] n_t dt + (\sigma + \sigma_t^q) \varphi_t n_t dz_t. \quad (4)$$

Note that, for convenience, we have also re-written consumption as  $\kappa_t = \lambda_t n_t$ . The structure of the above equation is essentially the same as in classical Merton's portfolio problem [Merton, 1969]: The agent makes the allocation choice  $\varphi$  between the risky (capital) and risk-free (cash) assets with the goal of maximising the value of pay-off from a (self-financing) portfolio. The main difference pertains to the fact that the price of capital,  $q$ , does not follow a geometric Brownian motion, as coefficients  $\mu^q$  and  $\sigma^q$  (that will be endogenously determined below) are not constant.

Following Brunnermeier and Sannikov [2014b], we hypothesize that the aggregate state of the economy is given by some one dimensional diffusion process which we here call  $x$ , and posit the equation of motion

$$dx_t = \mu_t^x x_t dt + \sigma_t^x x_t dz_t. \quad (5)$$

At this point, we do not say what  $x$  actually corresponds to. After formally writing down the agents' optimisation problems and aggregating, we will see that the system can indeed be described by a single variable  $x$  – the experts' share of the total net worth.<sup>27</sup>

Assuming then, that the aggregate state is fully specified by  $x$ , it follows that its drift and diffusion coefficients are functions of  $x$ ,  $\mu_t^x = \mu^x(x_t)$ ,  $\sigma_t^x = \sigma^x(x_t)$ , and importantly, so is the the price process  $q$ :

$$q_t = q(x_t), \quad \mu_t^q = \mu^q(x_t), \quad \sigma_t^q = \sigma^q(x_t).$$

The Itô's Lemma allows us now to create a mapping from the aggregate state  $x$  to price  $q$ . Applying it to  $q(x_t)$  yields

$$dq_t = \left[ \mu^x(x_t) x_t q'(x_t) + \frac{1}{2} \sigma^x(x_t)^2 x_t^2 q''(x_t) \right] dt + \sigma^x(x_t) x_t q'(x_t) dz_t. \quad (6)$$

Matching the drift and volatility terms in Eq. (6) with those from the original stochastic differential equation for the  $q$  process Eq. (3) yields the system of two equations:

$$\mu^q(x) = \mu^x(x) \frac{xq'(x)}{q(x)} + \frac{1}{2} \sigma^x(x)^2 \frac{x^2 q''(x)}{q(x)}, \quad (7a)$$

$$\sigma^q(x) = \sigma^x(x) \frac{xq'(x)}{q(x)}. \quad (7b)$$

Returning now to the agents' optimisation problem, the controls consumption  $\lambda$  and asset allocation  $\varphi$  are to be chosen so as to maximise the objective

<sup>27</sup>Of course, any invertible function of  $x$  could be considered the macrostate as well. In this particular example, one could alternatively use the capital price  $q$  as a state variable, since the mapping between  $q$  and  $x$  is invertible.

function that now depends on the present agent net worth  $n$  ( $\bar{n}$ ) and macro-state  $x$ . In our example, we assume that agents have linear consumption preferences, so that an expert's value function is

$$V(n, x) = \max_{\varphi, \lambda} \mathbb{E} \left[ \int_0^\infty e^{-\rho t} \lambda_t n_t dt \right]. \quad (8)$$

The value function must satisfy the Hamilton-Jacobi-Bellman (HJB) equation [Technical Appendix A.2, Eq. (41)] which here reads

$$\begin{aligned} \rho V(n, x) = \max_{\lambda, \varphi} & \left\{ \lambda(n, x) n \right. \\ & + \left[ r + \left( \frac{a}{q(x)} + \mu^q(x) + \sigma \sigma^q(x) - r \right) \varphi(n, x) - \lambda(n, x) \right] n \frac{\partial V(n, x)}{\partial n} \\ & + \mu^x(x) x \frac{\partial V(n, x)}{\partial x} + \frac{1}{2} \sigma^x(x)^2 x^2 \frac{\partial^2 V(n, x)}{\partial x^2} \\ & + \sigma^x(x) x [\sigma + \sigma^q(x)] \varphi(n, x) n \frac{\partial^2 V(n, x)}{\partial n \partial x} \\ & \left. + \frac{1}{2} [\sigma + \sigma^q(x)]^2 \varphi(n, x)^2 n^2 \frac{\partial^2 V(n, x)}{\partial n^2} \right\}. \end{aligned} \quad (9)$$

We cannot fix all boundary conditions for  $V$  at this stage, as we do not know what  $x$  is. Nonetheless, it is clear from Eq. (4) that if an agent has zero net worth, then  $n$  will always remain zero, as  $dn = 0$ . Consumption will then also be zero, and so  $V(0, x) = 0$  for all  $x$ . The objective function is linear in  $n$ , cf. Eq. (8), as are the  $n$  equations of motion, provided the controls are independent of  $n$ , and thus

$$V(n, x) = nW(x), \quad (10)$$

where  $W(x)$  can be interpreted as the marginal value of net worth.

Substituting the factored  $V$  into Eq. (9), we reduce it to an ordinary differential equation that depends only on a single state variable  $x$ :

$$\begin{aligned} (\rho - r)W(x) = \max_{\lambda, \varphi} & \left\{ \lambda(x)(1 - W(x)) \right. \\ & + \varphi(x) \left[ \frac{a}{q(x)} + \mu^q(x) + \sigma \sigma^q(x) - r + \sigma^x(x)(\sigma + \sigma^q(x)) \frac{xW'(x)}{W(x)} \right] W(x) \left. \right\} \\ & + \mu^x(x) x W'(x) + \frac{1}{2} [\sigma^x(x) x]^2 W''(x). \end{aligned} \quad (11)$$

It is easy to see that the right-hand side of (11) is linear in controls  $\varphi$  and  $\lambda$ . Thus, maximisation in consumption  $\lambda$  implies

$$\lambda(x) = \begin{cases} 0, & \text{if } W(x) - 1 < 0, \\ \text{unbounded}, & \text{if otherwise.} \end{cases} \quad (12)$$

For households, the consumption  $\bar{\lambda}$  choice is simpler: As they are not facing the non-negative consumption constraint, they choose their  $\bar{\lambda}$  so that  $\bar{W} = 1$ .

If the coefficient of  $\varphi$  in Eq. (11) were positive, all experts would allocate an unbounded amount of their net worth to  $k$  (using infinite leverage to do so). As total  $k$  is constrained, the capital allocations must all be finite, which is consistent with the agents' optimisation only if

$$\frac{a}{q(x)} + \mu^q(x) + \sigma\sigma^q(x) - r + \sigma^x(x)[\sigma + \sigma^q(x)]\frac{xW'(x)}{W(x)} = 0. \quad (13)$$

An equivalent formula holds for households and their capital allocation  $\bar{\varphi}$ , with the difference that they might prefer not to hold any capital at all:

$$\frac{\bar{a}}{q(x)} + \mu^q(x) + \sigma\sigma^q(x) - r \leq 0, \quad \text{with equality if } \bar{\varphi} > 0. \quad (14)$$

Under (12) and (13), the expert HJB equation reduces to

$$(\rho - r)W(x) = \mu^x(x)xW'(x) + \frac{1}{2}[\sigma^x(x)x]^2W''(x), \quad (15)$$

for any value of  $\varphi$  and for all values of  $x$  such that  $W(x) > 1$  holds. To fully close the model, one needs to pin down the equations of motion for the aggregate state – in other words, find and solve conditions determining diffusion coefficients  $\mu^x(x)$  and  $\sigma^x(x)$ .

Noting that the drift and diffusion of expert(households) net worth is linear in  $n$  ( $\bar{n}$ ), cf. Eq. (4), the total expert net worth, denoted  $N$ , follows

$$\begin{aligned} dN_t = & \left[ r + \left( \frac{a}{q(x_t)} + \mu^q(x_t) + \sigma\sigma^q(x_t) - r \right) \varphi(x_t) + \lambda(x_t) \right] N_t dt \\ & + (\sigma + \sigma^q(x_t)) \varphi(x_t) N_t dz_t. \end{aligned} \quad (16)$$

Similar dynamics would emerge for total household net worth,  $\bar{N}$ . Now the aggregate state is determined by two state variables,  $N$  and  $\bar{N}$  ( $x$  is of course still there, but here we are trying to identify what it should be). This reduces to one when one notes that debt and capital market clearing imply

$$N_t + \bar{N}_t = q_t K_t^{\text{tot}}, \quad (17)$$

where  $K_t^{\text{tot}}$  is the total capital in the economy. Aggregating the  $k$  equations motion the same way as was done above for  $n$ , we have that  $dK_t^{\text{tot}} = \sigma K_t^{\text{tot}} dz_t$ . We can now define the aggregate state variable to be the experts' share of the total net worth,

$$x_t \equiv \frac{N_t}{q_t K_t^{\text{tot}}}. \quad (18)$$



Itô differentiating the definition of  $x$ , we then have

$$dx_t = \left\{ \frac{a}{q(x_t)} \psi(x_t) + [\mu^q(x_t) - \sigma^2 - \sigma\sigma^q(x_t) - \sigma^q(x_t)^2 - r] [\psi(x_t) - x_t] - \lambda(x_t)x_t \right\} dt + [\sigma + \sigma^q(x_t)][\psi(x_t) - x_t] dz_t, \quad (19)$$

where  $\psi$  is the fraction of total capital held by the experts,  $\psi(x) \equiv x\varphi(x)$ . Equating the drift and diffusion terms of  $x$  as given by Eq. (19) with those coming from our earlier definition, Eq. (5), gives us what we will refer to as the *closure conditions*:

$$x\mu^x(x) = \frac{a}{q(x)}\psi(x) + [\mu^q(x) - \sigma^2 - \sigma\sigma^q(x) - \sigma^q(x)^2 - r] [\psi(x) - x] \quad (20a)$$

$$x\sigma^x(x) = [\sigma + \sigma^q(x)][\psi(x) - x]. \quad (20b)$$

Finally, we can state the remaining boundary conditions for  $q$  and  $W$ . Experts will have unbounded consumption at the point where  $W$  reaches one, cf. Eq. (12). This introduces a reflecting upper boundary  $x^*$ , as whenever expert net worth share is over this point, they consume until  $x$  returns to the level  $x^*$ . By the properties of a reflecting boundary [Technical Appendix A.4], the derivatives at  $x^*$  must vanish, and we then have in total

$$q'(x^*) = 0, \quad W'(x^*) = 0, \quad W(x^*) = 1. \quad (21)$$

At the lower boundary, share of experts' net worth is stuck at zero, and so the price of capital there must be such that households are willing to hold it forever. As excess returns from holding capital for households are  $\bar{a}/q(0) - r$  when price remains at  $q(0)$ , the least possible  $q$  must be  $\bar{a}/r$ . Finally, the marginal value of wealth  $W(x)$  for experts must tend to infinity as  $x \rightarrow 0$ :<sup>28</sup> From Eq. (20b) we have that  $\lim_{x \rightarrow 0} \varphi(x) = 1 + \sigma^x(0)/\sigma$ . Assuming that experts are always leveraged,  $\varphi(x) > 1$ , we must have that  $\sigma^x(0) > 0$ . Subtracting Eq. (14) from Eq. (13) we get that

$$\lim_{x \rightarrow 0} \frac{xW'(x)}{W(x)} = -\frac{a - \bar{a}}{q(0)\sigma^x(0)} < 0. \quad (22)$$

This implies that  $W(x) \rightarrow \infty$  as  $x \rightarrow 0$ . Thus, at the lower boundary we have:

$$q(0) = \frac{\bar{a}}{r}, \quad \lim_{x \rightarrow 0} W(x) = \infty. \quad (23)$$

<sup>28</sup>Brunnermeier and Sannikov [2014b] obtain the same boundary condition as follows: At  $x = 0$  experts get excess returns of  $a/q(0) - r > 0$ . Choosing  $\varphi$  high enough, their rate of returns exceeds their discount rate  $\rho$ , and value function becomes infinite. However, since  $x$  can never escape from 0, and experts only consume at  $x = x^*$ , it is not totally clear that  $V$  can indeed grow unboundedly. The condition is therefore plausible but may require more careful analysis to be rigorously justified.

In total, from Eqs. (21) and (23), we have five conditions, which is the correct number for two second order ordinary differential equations, plus the yet unknown consumption boundary  $x^*$ .

We now have a sufficient number of equations to find the aggregate state drift and diffusion coefficients  $\mu^x$  and  $\sigma^x$ , and the expert capital share  $\psi$ . In addition we should also state the differential equations determining  $W$  and  $q$  – we will obtain these when we solve our equations for  $W''$  and  $q''$ . For these five unknowns, the five equations we need are the optimal capital allocation conditions, Eq. (13) and (14), the pair of closure conditions, Eqs. (20), and finally the expert HJB, Eq. (15). Solution is straight-forward, albeit the result is not particularly pretty:

$$\sigma^x(x) = \frac{a - \bar{a}}{q(x)} \left[ -\frac{W(x)}{xW'(x)} \right] \left[ \frac{\sigma}{2} + \sqrt{\left(\frac{\sigma}{2}\right)^2 - \frac{a - \bar{a}}{q(x)} \frac{xq'(x)}{q(x)} \frac{W(x)}{xW'(x)}} \right]^{-1}, \quad (24a)$$

$$\mu^x(x) = \frac{a}{q(x)} - \sigma^x(x) \left\{ \sigma + \sigma^x(x) \left[ \frac{xW'(x)}{W(x)} + \frac{xq'(x)}{q(x)} \right] \right\}, \quad (24b)$$

$$\psi(x) = x \left[ 1 + \frac{\sigma^x(x)}{\sigma + \sigma^x(x) \frac{xq'(x)}{q(x)}} \right], \quad (24c)$$

$$q''(x) = \frac{2q(x)}{x^2\sigma^x(x)^2} \left\{ r - \frac{a}{q(x)} \left( 1 + \frac{xq'(x)}{q(x)} \right) + \sigma^x(x) \left[ \sigma^x(x) \left( \frac{xq'(x)}{q(x)} \right)^2 - \sigma \frac{xW'(x)}{W(x)} \right] \right\}, \quad (24d)$$

$$W''(x) = \frac{2W(x)}{x^2\sigma^x(x)^2} \left[ \rho - r - \mu^x(x) \frac{xW'(x)}{W(x)} \right]. \quad (24e)$$

Sign in front of the square root in Eq. (24a) is here chosen so that  $\sigma^x(0) > 0$ .

In the region where households do not hold capital, we have only four equations, as Eq. (14) used above becomes an inequality. On the other hand, we have only four unknowns as  $\psi = 1$ . Solving the remaining Eqs. (13, 20, 15), one now finds

$$\sigma^x(x) = \sigma \frac{1 - x}{x - (1 - x) \frac{xq'(x)}{q(x)}}, \quad (25a)$$

$$\mu^x(x) = \frac{a}{q(x)} - \frac{1 - x}{x} \left[ \sigma + \sigma^x(x) \frac{xq'(x)}{q(x)} \right] \left[ \sigma + \sigma^x(x) \left( \frac{xq'(x)}{q(x)} + \frac{xW'(x)}{W(x)} \right) \right], \quad (25b)$$

$$q''(x) = \frac{2q(x)}{x^2\sigma^x(x)^2} \left\{ r - \frac{a}{q(x)} - (\sigma\sigma^x(x) + \mu^x(x)) \frac{xq'(x)}{q(x)} - \left[ \sigma + \sigma^x(x) \frac{xq'(x)}{q(x)} \right] \left[ \sigma^x(x) \frac{xW'(x)}{W(x)} \right] \right\}, \quad (25c)$$

$$W''(x) = \frac{2W(x)}{x^2\sigma^x(x)^2} \left[ \rho - r - \mu^x(x) \frac{xW'(x)}{W(x)} \right]. \quad (25d)$$

The model solution is complete once we numerically solve the  $q$  and  $W$  differential equations.

## 4.2 Numerical solution

### 4.2.1 Marginal value $W$ and aggregate state $x$

The numerical solution of Eqs. (24) and (25) poses some challenges. Standard, local iterative ordinary differential equation solvers such as Runge-Kutta or predictor-corrector methods (see e.g. Hairer et al. [1993]) work by propagating the solution from a given point  $x$  forward or backward by evaluating the derivatives at and around  $x$ . In the case of Eqs. (24) and (25), such methods run into the problem of division by zero: At the left-hand side boundary, the derivatives of  $q$  and  $W$  evaluate to plus or minus infinity.

One solution to this problem is ignoring it: The solution of the equations can be attempted with the derivatives apparently evaluating to  $1/0$  or  $0/0$ . Infinite initial values are replaced by very large, but finite numbers, and the derivatives are evaluated with the hope that numerical round-off error sends zeros to small but finite values, so that the undefined division by zero condition does not occur.

A more satisfactory approach is to remove the singularities altogether by some change of variables, or to use an approximate analytic solution near the critical point. In this example, we do the latter by constructing the asymptotic expansion of  $q$  and  $W$  near the boundary [Technical Appendix section A.5]. We begin by assuming a power law form for the solution near the lower boundary that is consistent with the boundary conditions of Eq. (23):

$$q(x) = \frac{\bar{a}}{r} + q_1 x^\alpha + o(x^\alpha), \quad W(x) = W_1 x^{-\beta} + o(x^{-\beta}), \quad \alpha, \beta > 0. \quad (26)$$

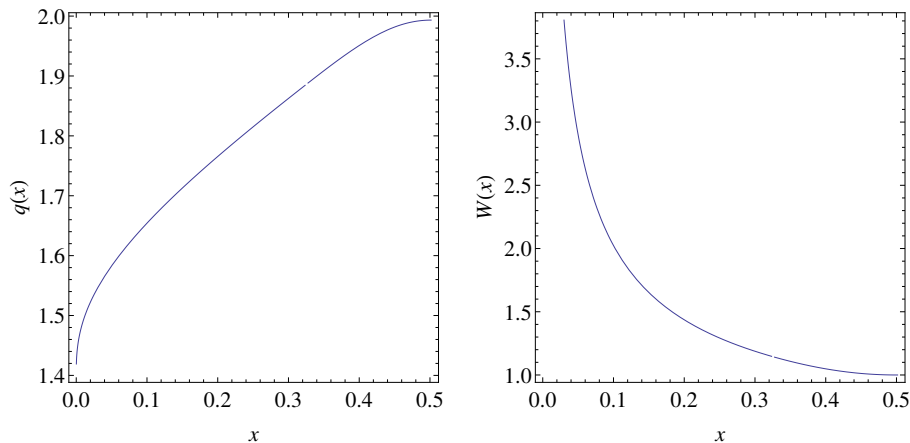
The exponents are determined by first substituting the trial functions into Eqs. (24d) and (24e), expanding the equations for small  $x$ , and then solving  $\alpha$  and  $\beta$  so that the leading order term vanishes. Albeit the algebra is tedious, a solution eventually emerges:

$$\alpha = \frac{1}{2} - \beta \left( 1 + \frac{a\bar{a}\beta\sigma^2}{(a-\bar{a})^2 r} \right) + \sqrt{\left[ \frac{1}{2} - \beta \left( 1 + \frac{a\bar{a}\beta\sigma^2}{(a-\bar{a})^2 r} \right) \right]^2 + \frac{2(\bar{a}\beta\sigma)^2}{(a-\bar{a})^2 r}}, \quad (27a)$$

$$\beta = \frac{1}{2} - \frac{\bar{a}\rho}{2ar} - \frac{(a-\bar{a})^2 r}{4a\bar{a}\sigma^2} + \sqrt{\left[ \frac{1}{2} - \frac{\bar{a}\rho}{2ar} - \frac{(a-\bar{a})^2 r}{4a\bar{a}\sigma^2} \right]^2 + \frac{(a-\bar{a})^2 r}{2a\bar{a}\sigma^2}}. \quad (27b)$$

Four different combinations for the signs in front of the square roots are possible. Clearly, however, only the above choice yields solutions that are both positive. The coefficients  $q_1$  and  $W_1$  will be determined by the boundary conditions.

For small  $x$ , we can now use the trial solutions of Eq. (26), truncated to the displayed terms, with  $\alpha$  and  $\beta$  from Eqs. (27). Then, for  $x$  greater than some small cross-over value  $\varepsilon$ , we use a standard iterative local ordinary differential equation solver, with initial conditions at  $\varepsilon$  coming from the asymptotic



(a) Capital price  $q$  as a function of the aggregate expert net worth share  $x$ . (b) Marginal value of expert net worth as a function of the aggregate expert net worth share  $x$ .

Figure 1: Numerical solution of Eqs. (24) and (25).

expansion. Value of  $\varepsilon$  should be large enough to ensure that evaluating the derivatives is not significantly affected by round-off error, but small enough so that the asymptotic expansion is accurate. As a first guess, the square root of the maximum relative error of the used floating point arithmetic,  $\varepsilon \sim 10^{-8}$  for double precision, can be used.

Using the above approach, we can now solve the  $q$  and  $W$  differential equations given  $q_1$  and  $W_1$ , the coefficients of the leading non-constant terms in the asymptotic expansion of  $q$  and  $W$  Eq. (26). We can arbitrarily fix  $W_1$  as the equations are invariant in linear scaling of  $W$ , and scale the solution ex-post in order to satisfy the condition  $W(x^*) = 1$  at the consumption boundary. We still have  $q_1$  and the position of the upper boundary  $x^*$  to be set so that the remaining boundary conditions  $W'(x^*) = q'(x^*) = 0$  are satisfied.

A simple numerical scheme that finds  $q_1$  and  $x^*$  can be setup as follows: Define  $\Theta(q_1)$  as  $q'$  evaluated at first  $x$  such that  $W'(x) = 0$  where  $q$  and  $W$  are solutions to the model equations for the given  $q_1$ . This point can be found by solving the differential equations forward from the initial point, until the  $W'(x)$  boundary is crossed; the exact crossing point is then polished using standard root finding methods. The correct  $q_1$  can then be determined by finding  $\Theta(q_1) = 0$ , where again, any standard root finding method can be employed.

In Fig. 1 we have plotted the numerical solution following the method described above. The parameter values used are  $a = 0.11$ ,  $\bar{a} = 0.07$ ,  $\sigma = 0.1$ ,  $r = 0.05$ , and  $\rho = 0.06$ .

### 4.2.2 Equilibrium probability distribution

The statistics of the possible realisations of the economy are given by the probability density function of  $x$ ,  $f(x)$  which itself is obtained from the Kolmogorov Forward equation [Technical Appendix section A.3, Eq. (45)]. In equilibrium, for the process  $x$  of Eq. (19), this reads

$$0 = \mu^x(x)xf(x) - \frac{\partial}{\partial x} \left[ \frac{1}{2}\sigma^x(x)^2x^2f(x) \right], \quad (28)$$

where the coefficients  $\mu^x$  and  $\sigma^x$  come from Eqs. (24) or (25) depending whether households are holding capital or not, and where the  $q$  and  $W$  functions are presumed to have already been solved for. Here, we have also used the fact that  $x^*$  is a reflecting boundary, and set the left-hand side of Eq. (28) to zero [Technical Appendix section A.4.2, Eq. (50)].

Rather than solving Eq. (28) directly, it is easier to define

$$f(x) = 2g(x)/(x\sigma^x(x))^2,$$

and solve for the function  $g$  instead. This change of variables avoids us having to differentiate  $\sigma^x$ , a straight-forward but laborious task. For  $g$ , the equation reads

$$0 = \frac{2\mu^x(x)}{x\sigma^x(x)^2}g(x) - g'(x). \quad (29)$$

The solution needs to be normalised to unit integral over the  $x$  range, and to do this, we can solve the differential equation  $F'(x) = f(x)$  in parallel to the one above.  $F$  will then be the cumulative probability distribution, if we further ask that  $F(0) = 0$ . Numerically, the equations can be solved with arbitrary initial conditions: If  $f$  and  $F$  are such un-normalised solutions, one simply replaces them according to  $f(x) \mapsto f(x)/[F(x^*) - F(0)]$  and  $F(x) \mapsto [F(x) - F(0)]/[F(x^*) - F(0)]$ . This is valid since the  $f$  and  $F$  differential equations are invariant in scaling, the  $F$  equation also in the addition of constants.

As was the case with  $q$  and  $W$  equations, we have a singularity at  $x = 0$ , since  $2\mu^x(x)/(x\sigma^x(x))$  tends to infinity at zero. An asymptotic expansion could be used here as well, and for completeness, we shall do it. But for the numerical solution, a simpler approach is possible: we can choose some interior point, and solve from there left towards the  $x = 0$  boundary, and right up to  $x = x^*$  edge. Due to the singularity, the left hand side solution is likely to fail before reaching  $x = 0$ , but this is fine as long as we got near enough to 0, and that the integral of density tends to a finite value (if not, in equilibrium all probability mass is at  $x = 0$ ).

The  $x \rightarrow 0$  asymptotic form of  $f$  can be found using the methods we have already used above. Alternatively, it would suffice to note that in the  $x \rightarrow 0$

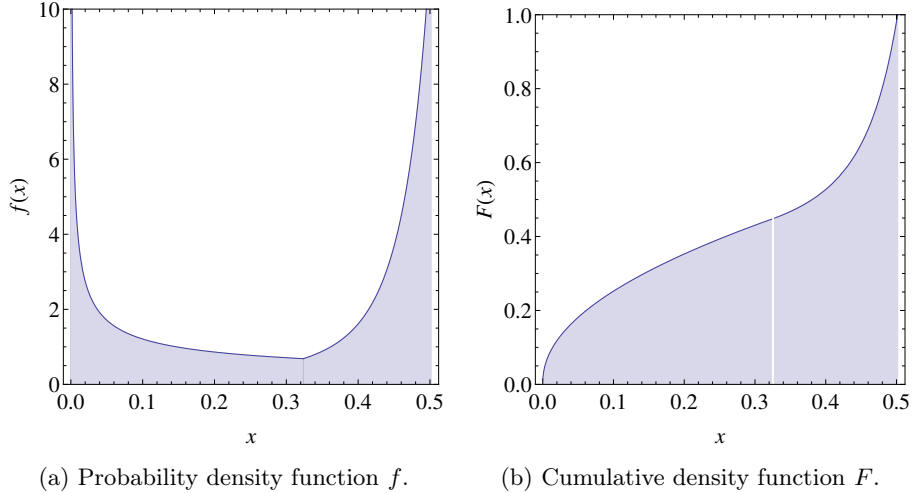


Figure 2: Equilibrium probability distribution functions.

limits of the relative drift and diffusion of  $x$  are

$$\lim_{x \rightarrow 0} \sigma^x(x) = \frac{r}{\bar{a}\beta} \frac{a - \bar{a}}{\sigma}, \quad (30a)$$

$$\lim_{x \rightarrow 0} \mu^x(x) = \frac{r}{\bar{a}\beta} \left[ \frac{r}{\bar{a}} \left( \frac{a - \bar{a}}{\sigma} \right)^2 + \bar{a} + (\beta - 1)a \right], \quad (30b)$$

and we can replace  $\mu^x$  and  $\sigma^x$  by these limits in Eq. (28) and solve  $f$  analytically. The result is

$$f(x) \propto x^\gamma + o(x^\gamma), \quad \text{where } \gamma = 2 \left[ \beta - 1 + \beta \frac{\bar{a}(\bar{a} + (\beta - 1)a)\sigma^2}{(a - \bar{a})^2 r} \right]. \quad (31)$$

One immediately sees that if  $\gamma < -1$ , the integral of  $f$  is infinite over arbitrarily small interval  $[0, \delta]$ ,  $\delta > 0$ . If this is the case, for those parameters, all probability condenses to  $x = 0$ .

The numerically solved probability distribution function  $f$  and the cumulative probability function  $F$  are plotted in Fig. 2. Comparing to BS(2014a), even in this simplified model, the two peaked structure of the density  $f$  is still visible. This is relatively unsurprising since the models differ mainly by the inclusion of investment dynamics. We refrain from analysing the economic implications of the result, as the main goal was to present an easy example to understand derivation of the model and the solution methods.

## 5 Some paths for future research

We complete this chapter with a short discussion of the range of further issues that models of this kind might usefully address emphasising once again

that this is still an immature literature, that many technical challenges remain and therefore we can make no firm predictions about where contributions and breakthroughs will come.

There are many paths of future investigation that could be followed. A number of contributions to the new continuous-time macrofinancial literature assume an ‘AK’ production function i.e. output is a linear function of the stock of capital. This assumption, which can be traced back to the contribution of Frankel [1962], was widely employed in the earlier contributions to the literature on endogenous growth (see Aghion and Howitt [2009] for review of this literature). Both Isohätälä et al. [2014b] and Brunnermeier and Sannikov [2014b] can be interpreted as models of endogenous growth yielding the prediction that capital accumulation and hence the rate of economic growth will fall when corporate net worth falls and that the economy may then remain for an extended period in a phase of low investment and low growth (the ‘net worth’ trap). These models may thus already provide some insight into the major puzzle of the slowdown of aggregate productivity growth in many countries since the global financial crisis.

This is though a rather simplistic account of growth. The theoretical growth literature has moved on to focus on other mechanisms such as investment in product variety (as in Romer [1987, 1990]) and in the discovery of new more efficient methods of production (innovation) that replace older inferior methods (e.g. the model of Schumpeterian ‘creative destruction’ of Aghion and Howitt [1992]). A natural further development will therefore be to employ similar continuous time models to examine the impact of balance sheet constraints and financial distress on investment in new products and processes and hence on productivity growth. Doing this though may be difficult within the current assumption of only a single state variable.

A related issue is that of structural adjustment following financial crises. Many countries need to adjust the structure of their economies, for example switching labour and capital resources from non-traded to traded output. Rising risk-premia in periods of financial distress can act as a barrier to such investments, providing another form of trap in a low-output low-income state. Other similar issues arise in understanding the low elasticities of traded sector output to changes in exchange rates following financial crises or in the response of small open economies to ‘sudden stops’ of capital flows (on this issue Brunnermeier and Sannikov [2014a] have made a promising start using continuous time methods).

Further understanding will surely also need to take account of the interaction of balance sheet constraints with household, corporate and bank expectations about future productivity and incomes. To date the continuous-time macrofinancial literature has imposed the conventional but rather strong assumption of model-consistent expectations. Every agent is assumed to know both the current state of the economy and stochastic processes that drives both the state of the economy and market prices. Different, possibly even more extended, dynamics can be expected when agents update their expectations about unobserved states and processes in response to their current observations. Thus, for example, a

fully adequate model of endogenously created bank inside-money would seem to need to take account of the possibility that optimistic expectations about future income result in a period of rapid and self-reinforcing expansion of both bank credit and bank money. Introducing learning of this kind into these models will be a further technical challenge.

A further potential line of inquiry is to make the treatment of financial markets more realistic, for example by allowing for alternative (financial markets) source of financing for the productive sector, on top of bank loans, since the empirical evidence suggests that firms tend to partly substitute bank financing by market financing when credit conditions tighten (see e.g., [Becker and Ivashina \[2014\]](#)). This would help to get a better understanding of how the substitutability of funding sources can add to/mitigate the propagation of fundamental shocks, with the implications for growth and financial stability. Another similar departure would be to allow for the internal source of financing for the productive sector (i.e., capital accumulation within firms). Introducing this feature, most likely, would require introducing an additional state variable in play – the net worth of the productive sector, which might be technically challenging.<sup>29</sup>

Similarly there is need for better understanding of the role of both commercial and central bank balance sheets in macroeconomic transmission and the supply of credit. The DSGE assumption that all that matters in monetary policy is interest rates is now accepted as an oversimplification, but we are not yet in the position of having tractable incomplete market models, in either continuous or discrete time, in which the role of commercial and central bank balance sheets is clearly articulated. While the work of [Brunnermeier and Sannikov \[2014e\]](#) that we have reviewed offers a particularly promising start, it is still not yet possible to say that the present ‘state of the art’ is sufficiently developed to provide a full understanding of the impact of unorthodox monetary policy (central bank balance sheet expansion) or macroprudential tools (such as cyclically varying capital requirements or limits on loan-to value ratios).

These paths for future research are far from exhausting the list of possible applications of new continuous-time macrofinancial modelling. There are also opportunities to apply this framework in a number of other settings more routinely explored using standard linearised DSGE models. Examples include modelling the labour market, real wages and employment and product markets and price setting. If balance sheet constraints affect investment and asset markets then they should also affect labour and goods markets. Writing down such models in continuous-time with balance sheet and net worth constraints does not seem so difficult. Solution though could be challenging because of the need to include additional state variables.

The challenges of numerical solution should give pause for thinking carefully about the choice of modelling strategy. There are well developed tools for numerical solution of macroeconomic models with several state variables in

---

<sup>29</sup>In a discrete time set-up an attempt to accommodate this feature is made by [Rampini and Viswanathan \[2012\]](#).



discrete time. Replicating all this technical work on solution with many state variables in continuous-time may not be an efficient way to proceed. It may instead be more useful to find ways to incorporate the insights of continuous-time macrofinancial modelling into more widely known and understood discrete time settings. This is one reason why, as we have already suggested, we believe that over time the ‘gap’ between continuous-time and discrete time specifications can and should be closed. There is no reason why the impact of balance sheet constraints and net worth cannot be incorporated into otherwise standard discrete-time specifications (although this may come at some cost, for example the need to introduce more explicit modelling of what happens when disturbances result in constraints binding, something that can often be conveniently put to one side when uncertainty is modelled as a continuous time diffusion).

Finally, of course, it will be essential to take these models closer to data. Some initial steps in this direction are already made by [He and Krishnamurthy \[2013, 2014\]](#) who seek to replicate the asset market behaviour during the 2007-2009 financial crisis. Improving predictive and simulation properties of these models, however, carries a risk of losing tractability and transparency. An illustration is the challenge of explaining the counterfactual prediction of [He and Krishnamurthy \[2012\]](#) and [He and Krishnamurthy \[2013\]](#) that the leverage of financial intermediaries increases substantially during financial crises (the underlying mechanism is that the emergence of crisis results from low specialist equity requiring them to leverage in order to maintain ‘skin in the game’).

As an example of the challenges of bring these models to the data, we can briefly describe how the model by [Adrian and Boyarchenko \[2013\]](#) employs continuous-time macro-financial modelling tools to develop an explanation of observed pro-cyclical intermediary leverage. There are several distinctive features of this work. Three of these seem to be particularly important. First that financial intermediaries do not maximise an objective function (such as present discounted future dividends); instead, they behave mechanically, first using earnings to pay floating rate coupons on long term bonds issued to households (the coupon rate is determined by equilibrium of the supply and demand for these bonds) and retaining all remaining earnings to build up equity and invest in productive capital. Second that if intermediary equity falls below a lower boundary, then financial intermediaries are restructured, with debt-holders wiped out and re-established under new equity holders. Since all intermediaries are identical and hit by the same shocks this is a systemic crisis. Third that financial intermediary leverage, and hence their investment in productive capital, is continuously determined by a regulatory driven value at risk type constraint that depends on the short term volatility of the price of productive capital.

Similar to the other macrofinancial models that we have described, all variable of interests (intermediary equity, the price of productive capital, the expected excess returns on holding intermediaries’ debt and productive capital) can be described as the functions of a single state variable. Just as in [He and Krishnamurthy \[2012\]](#) and [Brunnermeier and Sannikov \[2014b\]](#) this state variable is the share of financial intermediaries’ net worth in total wealth. The

novel contribution are two closely related empirical predictions not captured by other models. First, as a direct consequence of the assumed leverage constraint, it generates the empirically-observed procyclical pattern of intermediary leverage. Second it also explains what [Adrian and Boyarchenko \[2013\]](#) describe as the volatility paradox, i.e. the well known observation that systemic risks tend to increase during periods of perceived low volatility e.g. for example during the ‘great moderation’ that preceded the global financial crisis. In their model this appears as a negative relationship between the instantaneous endogenous volatility of the returns to holding capital with the probability of a systemic default. As endogenous volatility declines, leverage rises and thus also does the risk of a systemic default on a six-month ahead horizon (see their Figure 5.)

While [Adrian and Boyarchenko \[2013\]](#) make a valuable further contribution, their work can also be read as an illustration of the very substantial challenges of bringing models of this kind to the data. Departures from forward looking behaviour or inter-temporal optimisation are not necessarily wrong, but these departures open up such a large menu of possible modelling choices that it is difficult to know what is the best way forward. There is an almost unlimited range of possible underlying assumptions of this kind that can generate macro-financial interactions.

In order to impose the required intellectual discipline, it may prove necessary to focus on developing models of incomplete markets that are consistent not just with observed aggregate outcomes such as asset prices or national accounting measures of output and investment (there are simply too many potential modelling choices for doing this), but also with micro-level data at the level of individual firms, households and financial institutions. In this context it will be difficult to ignore the simplifying aggregation assumptions employed in the continuous-time macrofinancial literature. Instead it may eventually prove necessary to work with large scale agent-based models, in which distribution of net worth and leverage within sectors is tracked as well as the aggregate net worth (see [Haldane \[2015\]](#) for further discussion of why macrofinancial modelling should use agent-based modelling). Work of this kind will though need a different approach to research than has been conventionally used in macroeconomics, requiring relatively large teams of researchers in order to collect and match the underlying microlevel data.

## 6 Conclusion

This chapter has reviewed several contributions to a new and promising current literature, employing continuous time models to capture some of the macrofinancial interactions that have been highlighted by the global financial crisis. Though using highly stylized specifications, these models demonstrate how the interaction of market incompleteness with the balance sheet constraints of economic agents can generate dynamics of macroeconomic variables much more consistent with the empirically observed patterns at times of crisis than those generated by conventional DSGE models. These dynamics include substantial

variations in risk-premia and asset prices, as well as subsequent substantial and highly persistent declines of macroeconomic aggregates such as output and investment. In the models reviewed in this chapter these dynamics are largely driven by what are in effect changes in attitudes to risk, and by externalities stemming from the fact that individual agents do not internalize the impact of their individual risk taking or other decisions on the wider economy.

As we highlight in Section 5 of this chapter, the modelling approach introduced by this new literature has the potential to address a large spectrum of macroeconomic problems. The pursuit of these avenues of future research is an exciting challenge, from both technical and economic perspectives. Where then will future research ultimately take us? We have no crystal ball but our judgment is that the eventual destination of this literature will be a relatively small number of comparatively simple but influential continuous-time models of the kind we review here, providing widely accepted economic intuition and policy insight into macrofinancial interactions rather than predicting accurately macroeconomic and financial market developments.

An important impact of these modelling efforts may be persuading researchers working in more conventional discrete-time frameworks of the necessity of taking incomplete markets and balance sheet and net worth constraints seriously. Matching with data is then in turn likely to require more ‘agent-based’ approaches in order to meet the fundamental challenge of aggregation.

Such research will need a very careful process of matching against both microlevel and aggregate data in order to develop useful models. The final outcome could be a shift in modelling paradigms, from the typical small team work found in much current macro-economics to more resource intensive investigations involving many investigators and massive efforts at data-collection and calibration, with the overall direction of research guided to an important degree by the insights of the new continuous-time approach to macrofinancial modelling.

This does not avoid the need for considerable efforts in order to find a reasonable balance between on the one hand transparency and clear economic intuition, and on the other the realism and accuracy of underlying economic assumptions. We believe that the new macrofinancial modelling, despite the use of techniques of continuous-time stochastic modelling with which most economists are not very familiar, are especially useful because of the relatively clear and simple intuitions they provide about the macroeconomic consequences of incomplete markets and hence the resulting impact of balance sheet and net worth on macroeconomic outcomes. We therefore hope this chapter can be helpful in acquainting our readers with the conceptual and technical features of this new generation of models and stimulating interest in this field of research.

## A Appendix: Basics of continuous time models

In this appendix, we review some technical issues that arise when trying to understand, construct, or solve models that incorporate binding financing con-

straints in continuous time models. We begin with a brief summary of key concepts, eschewing formal proofs and favouring heuristic derivations that nonetheless can in principle be used as a basis for more rigorous approach.

## A.1 Stochastic differential equations

The primary modeling tool for describing the state of an agent is the *stochastic differential equation* (SDE), a generalisation of the ordinary differential equation that incorporates some form of external stochastic forcing. Gardiner [2009] gives an excellent and approachable review of many of the topics covered here and subsequent sections; other solid reference texts on stochastic differential equations include Øksendal [2003] and Feller [1971].

Let  $x_t$  stand for an agent's state, e.g. wealth or net worth, at time  $t$ , with the initial,  $t = 0$  state  $x_0$  given. We say  $x$  is a *diffusion process* and formally write its equation of motion as the SDE

$$dx_t = \mu(x_t) dt + \sigma(x_t) dz_t, \quad (32)$$

where  $z_t$  is a *Wiener process*, a continuous stochastic process having independent and normally distributed increments,  $z_{t+h} - z_t = \sqrt{h}\eta$ , where  $\eta$  is a unit normal distributed random variable. The differentials,  $dx_t, dz_t, \dots$ , can be viewed as the zero time step limits of corresponding finite differences, so that Eq. (32) becomes the limit of the discrete time model

$$x_{(k+1)T} = x_{kT} + T\mu(x_{kT}) + \sqrt{T}\sigma(x_{kT})\eta_{kT}, \quad (33)$$

where  $T, T \rightarrow 0$ , is the length of the period and  $\eta_{kT}, k = 0, 1, \dots$ , are independent unit normal distributed random variables<sup>30</sup>. More rigorously, the stochastic differential equation, Eq. (32) should be understood as a short-hand way of writing the *stochastic integral*

$$x_t = x_0 + \int_0^t \mu(x_s) ds + \int_0^t \sigma(x_s) dz_s, \quad (34)$$

where the integral against the Wiener process, or any diffusion process in general, is defined analogously to the Riemann-Stieltjes integral. *Stochastic calculus* is the theory of stochastic differential equations; for a more focused reference on the topic, we refer the reader to e.g. Klebaner [2005].

An important analytical tool is the *Itô's Lemma* which allows one to differentiate functions of diffusion processes such as  $x$  as given by Eq. (32). If  $f$  is

---

<sup>30</sup>The above uses the Itô interpretation of the SDE (32) which amounts to assuming that the noise amplitude  $\sigma$  is evaluated at the start of each period. In general one can set  $x_{(k+1)T} = x_{kT} + T\mu(x_{kT}) + \sqrt{T}\sigma(x_{kT+\alpha})\eta_{kT}$  where  $0 \leq \alpha \leq 1$ . Choice of  $\alpha$  does influence the form of later formulas. The Itô interpretation,  $\alpha = 0$  is the default choice in economics applications, as equations of motion are supposed not to pre-empt the shocks. In natural sciences, the Stratonovich convention  $\alpha = 1/2$  is commonly used.

any twice differentiable function defined in the domain of  $x$ , then  $f$  evaluated at  $x_t$ , denoted  $f_t = f(x_t)$ , is also a diffusion process with the increments

$$df_t = \left[ \frac{\partial f(x_t)}{\partial t} + \mu(x_t) \frac{\partial f(x_t)}{\partial x} + \frac{1}{2} \sigma(x_t)^2 \frac{\partial^2 f(x_t)}{\partial x^2} \right] dt + \frac{\partial f(x_t)}{\partial x} \sigma(x_t) dz_t. \quad (35)$$

Although we focus on single state variable problems, equivalent equations for multivariate case are still useful. For instance, in the construction of our example model, we needed multivariable formulas as we reduced an initially two variable problem to a single variable. Suppose  $X_t$  is an  $N$ -dimensional diffusion process taking values on some subset of  $\mathbb{R}^N$ ,  $X_t = (x_t^1, x_t^2, \dots, x_t^N)^\top$  (here  $^\top$  stands for matrix transposition). We write the  $X$  equation of motion as

$$dX_t = \mu(X_t) dt + \sigma(X_t) dZ_t, \quad (36)$$

where now  $Z$  is a vector of Wiener processes,  $Z_t = (z_t^1, \dots, z_t^K)^\top$ ,  $K$  is the number of independent shock sources,  $\mu(X_t) = (\mu^1(X_t), \dots, \mu^N(X_t))^\top$  is the drift vector and  $\sigma(X_t) = [\sigma^{ij}(X_t)]_{ij}$  is the  $N \times K$  covariance matrix. The elements of  $Z_t$  can be without loss of generality taken to be independent: any and all instantaneous correlations between shocks are encoded in the matrix  $\sigma$ .

The multivariate version of Itô's Lemma, Eq. (35), for a function  $f = f(X)$ ,  $f: \mathbb{R}^N \mapsto \mathbb{R}$ , reads

$$df_t = \left\{ \mu(X_t)^\top \nabla_X f(X_t) + \frac{1}{2} \text{tr} [\sigma(X_t)^\top \mathbf{H}_X f(X_t) \sigma(X_t)] \right\} dt + \nabla_X f(X_t)^\top \sigma(X_t) dZ_t, \quad (37)$$

where  $\text{tr}$  stands for matrix trace, and  $\nabla_X f$  and  $\mathbf{H}_X f$  give the gradient vector and the  $N \times N$  Hessian matrix of the function  $f$ ,  $[\nabla_X f(X)]_i = \partial f(X) / \partial x^i$ ,  $[\mathbf{H}_X f(X)]_{ij} = \partial^2 f(X) / \partial x^i \partial x^j$ .

These equations assume that the SDEs do not depend explicitly on time  $t$ , that is, they are time-homogeneous. An easy way of extending all of these definitions to account for explicit time dependence is to consider  $t$  additional state variable in  $X$ , with the trivial SDE  $dt = dt$ .

## A.2 Hamilton-Jacobi-Bellman equation

The standard tool in stochastic dynamical programming is the *Hamilton-Jacobi-Bellman equation* [Fleming and Soner, 2006]. Suppose that equations of motion, Eqs. (32) and (36) in the multivariate case depend on some controls  $y$ , which we now wish to choose so that our discounted future utility is maximal. For an infinite time-horizon problem, with standard exponential discounting and time-preference rate  $\rho$ , the objective function to maximise is

$$\Omega(x_0; \{y_t\}_{t=0}^\infty) = \mathbb{E} \int_0^\infty e^{-\rho t} u(y_t) dt, \quad (38)$$

where  $u$  is the utility function. Let  $V$  then be the maximal  $\Omega$ , and henceforth assume that  $y$  refers to the maximiser:

$$V(x) = \max_{\{y_t\}_{t=0}^{\infty}} \Omega(x_0; \{y_t\}_{t=0}^{\infty}). \quad (39)$$

Assuming that  $V$  is twice differentiable, it can be found as a solution to the Hamilton-Jacobi-Bellman (HJB) equation:

$$\rho V(x) = \max_y \left\{ u(y) + \mu(x, y)V'(x) + \frac{1}{2}\sigma(x, y)^2V''(x) \right\}. \quad (40)$$

An easy heuristic derivation goes as follows: (i) Start with the definition of  $V$  and divide the integral into two parts: from 0 to some small  $h$ , and from  $h$  to infinity; (ii) Use Bellman's principle on the second integral to make it  $V$  a time  $h$  later, appropriately discounted,  $e^{-\rho h}V(x_h)$ ; (iii) Use Itô's Lemma to approximate  $V(x_h)$ , Taylor expand in  $h$ , and then let  $h \rightarrow 0$ . This basic derivation of the Hamilton-Jacobi-Bellman equation requires that  $V$  be twice differentiable. It is well known that solutions do not always have this property. The theory of viscosity solutions addresses this problem, however, this topic is beyond the scope of this introduction (See e.g. Fleming and Soner [2006], or Crandall et al. [1992] for a rigorous but self-contained guide to the subject).

Generalising to the multivariate case, the Hamilton-Jacobi-Bellman equation for the process of Eq. (36) with controls  $y$  reads:

$$\rho V(X) = \max_y \left\{ u(y) + \mu(X)^\top \nabla_X V(X) + \frac{1}{2} \text{tr} [\sigma(X)^\top H_X V(X) \sigma(X)] \right\}. \quad (41)$$

### A.3 Equilibrium characterisation

In the conventional view, persistent e.g. in much of DSGE modeling today, what constitutes an equilibrium is a point in state space, plus random fluctuations induced by shocks. Such equilibria are a feature of, say, models linearised around a deterministic steady state, and which treat shocks as relatively small perturbations. When the modeling paradigm allows for large deviations, as is the case in the models we are highlighting here, this point-plus-perturbations picture of the equilibrium breaks down. Shocks, possibly amplified by feedback effects, can now drive the system far from what would have traditionally been seen as a relatively tranquil equilibrium point. Rather then, the equilibrium is characterised by a *probability distribution over the whole state space*.

In continuous time, the probability distribution of a diffusion process  $x$  is given by the *Kolmogorov forward equation*, also known as the *Fokker-Planck equation* (See e.g. Gardiner [2009]; Risken [1996] is solely dedicated this equation): If  $f(t, x|x_0)$  is the probability density function of process  $x$  following Eq. (32) with initial data  $x_0$ , then  $f$  satisfies the partial differential equation (omitting the explicit conditioning on the initial  $x$ ):

$$\frac{\partial f}{\partial t}(t, x) = -\frac{\partial}{\partial x} \left\{ \mu(x)f(t, x) - \frac{\partial}{\partial x} \left[ \frac{1}{2}\sigma(x)^2 f(t, x) \right] \right\}. \quad (42)$$

Probability densities are integrated to get actual probabilities: Given a large number of independent realisations of  $x$ , the probability of finding  $x$  in the interval  $[x_0, x_1]$  at time  $t$  is

$$\mathbb{P}_t(x_t \in [x, x + \Delta x]) = \int_{x_0}^{x_1} f(t, x') dx'. \quad (43)$$

The “large number of independent realisations” can be understood either as many simultaneously running independent processes (and so with independent shocks), or as a large number of samples of a single process, taken over an infinitely long time period. In the former view,  $f(x)$  represents the cross-sectional density of the state variables following the same dynamic stochastic equations of motion. Which view is correct depends of course on what one aims to model.

For additional intuition, the forward equation can be written in the form of a continuity equation relating the temporal change of  $f$  to spatial variation of a probability flux:

$$\frac{\partial f}{\partial t}(t, x) = -\frac{\partial j}{\partial x}(t, x), \quad (44a)$$

$$j(t, x) = \mu(x)f(t, x) - \frac{\partial}{\partial x} \left[ \frac{1}{2} \sigma(x)^2 f(t, x) \right], \quad (44b)$$

where  $j(t, x)$  is the probability current, that is, the rate of flow of probability through the point  $x$  to the positive  $x$  direction. In equilibrium  $\partial f(t, x)/\partial t = 0$ , and therefore the Fokker-Planck equation reduces to a first order ordinary differential equation

$$\mu(x)f(t, x) - \frac{\partial}{\partial x} \left[ \frac{1}{2} \sigma(x)^2 f(t, x) \right] = j_0, \quad (45)$$

where  $j_0$  is a constant to be determined by the boundary conditions.

The multivariate Fokker-Planck equation corresponding to the process  $X$ , again taking values on  $\mathbb{R}^N$  and following the SDE (36), is in turn

$$\frac{\partial f}{\partial t}(t, X) = -\nabla_X \cdot J(t, X) \quad (46a)$$

$$J(t, X) = \mu(X)f(t, X) - \nabla_X \cdot \left[ \frac{1}{2} \sigma(x)^\top \sigma(x) f(t, X) \right], \quad (46b)$$

where  $J = (j^1, \dots, j^N)^\top$  is now an  $N$ -dimensional probability current, and  $\nabla_X \cdot$  is the divergence operator,  $\nabla_X \cdot J(X) = \sum_{i=1}^N \partial j^i(X)/\partial x^i$ ,  $[\nabla_X \cdot \sigma(x)^\top \sigma(x)]_i = \sum_{j=1}^N \partial [\sigma(x)^\top \sigma(x)]_{ij} / \partial x^j$ .

#### A.4 Boundary conditions

Solutions to the HJB and the Fokker-Planck equations, Eqs. (40) and (44), or Eqs. (41) and (46) in the multivariate case, are not uniquely fixed until appropriate boundary conditions are given. Nonlinearity, capital constraints, and the

need for proper treatment boundary conditions go hand in hand: If one is to construct a model that can account for large fluctuations, one must account for the possibility of a state variable hitting a hard bound, e.g. a capital or leverage constraint. Here, we consider the two most common boundary conditions: the absorbing and (instantaneously) reflecting boundary.

#### A.4.1 Absorbing boundary

An absorbing boundary, say placed at position  $x^\dagger$ , is such that upon reaching it, the process is stopped and removed from the distribution. A stopped process (an agent who goes out of business, or a firm that has been liquidated) can no longer generate utility, and so it is natural to require that at an absorbing boundary value is zero,

$$V(x^\dagger) = 0, \tag{47}$$

whenever of course the utility function is non-negative for all controls. For the probability density, an absorbing boundary at  $x^\dagger$  means that

$$f(x^\dagger) = 0, \tag{48}$$

which has the natural interpretation of asking that point  $x^\dagger$  is always completely free of the process  $x$ .

Multivariate generalisations are obvious: The absorbing boundary is not a point anymore, but some surface in the embedding space, and the same zero value or zero density requirement holds.

#### A.4.2 Reflecting boundary

Although the word reflection invokes a picture of a very certain type of motion, such as elastic bouncing of a ball off of a rigid wall, or specular reflection of a beam of light, a reflecting boundary is here understood somewhat more generally: We will call a boundary reflecting whenever it conserves probabilities in the sense that it does not leak probability in or out, or allow the process accumulate or stop there for a finite time period.

In general, a reflecting boundary is setup by some forcing term that is strong enough to overcome the drift and diffusion terms in Eq. (32), preventing the process from ever crossing the boundary. Such forcing can be due to e.g. a singular control term (a control that has unbounded magnitude and which optimally is always either fully on or off). A rigorous mathematical treatment of SDEs with reflection does not use infinitely strong drift terms<sup>31</sup>, but as a model to guide intuition, the idea that the boundary is enforced by infinitely strong and infinitely short kicks is reasonable enough.

---

<sup>31</sup>A standard approach is recasting the SDE with reflection into a so called Skorokhod problem, whereby the process  $x$  is seen as driven by an additional process  $k$ ,  $dx_t = \mu(x_t) dt + \sigma(x_t) dz_t + dk_t$ , and where  $dk_t$  is non-zero only on the boundary. See e.g. Lions and Sznitman [1984].



For functions of process  $x$ , a reflecting boundary implies a Neumann condition: it imposes a specific value on the derivative of the function. Say there is a reflecting boundary at  $x^*$ . For the value function, we then must have that

$$V'(x^*) = 0. \quad (49)$$

This can be justified as follows: imagine the  $x$  range inside the boundary reflected into the range outside the boundary. One can now view a process hitting the boundary as instead passing into the “mirror” space. In order for  $V$  to be smooth across the boundary, demanded by the smoothness of optimal  $V$ , then the derivative  $V'$  is zero. Similarly, for any continuously differentiable function of the process, the derivative should vanish at  $x^*$ .

For the density  $f$ , a reflecting boundary naturally corresponds to a point where the probability current  $j$ , Eq. (44), vanishes

$$j(x^*) = \mu(x)f(x) - \frac{\partial}{\partial x} \frac{1}{2} \sigma(x)^2 f(x) \Big|_{x=x^*} = j_0 = 0. \quad (50)$$

Note that in the one-dimensional case, in steady state, this condition fixes the probability flow to zero over the whole of the  $x$  range. If there is also a reachable absorbing state, the probability density must be over the  $x$  range, for then both the value and derivative of  $f$  vanish at the same time. The word reachable is key: An absorbing boundary may be such that it cannot be arrived at in finite time. In this case, both an absorbing and reflecting boundary can co-exist, with the probability density not collapsing to zero.

Extensions to  $N$ -dimensional processes  $X$  are somewhat more complicated than for the absorbing boundary. Suppose that the reflecting boundary is a surface in  $\mathbb{R}^N$ ,  $x^*$  is a point on that surface, and that  $\Gamma(x^*)$  is the direction of the boundary forcing term (assumed never perpendicular to the normal of the boundary). Then the boundary condition for  $V$  reads

$$\Gamma(x^*) \cdot \nabla_X V(x^*) = 0, \quad (51)$$

that is, the  $\Gamma$ -directed derivative of  $V$  is zero when on an reflecting boundary.

For the Fokker-Planck equation, the  $N$ -variable extension Eq. (50) is

$$\nu(x^*) \cdot J(x^*) = 0, \quad (52)$$

where  $\nu(x^*)$  is the inwards unit normal vector of the boundary surface at  $x^*$ , and  $J$  is the probability current as given by Eq. (46b). The natural interpretation is that the probability flow perpendicular to the surface is zero (no outflow of probability, or accumulation on the surface).

## A.5 Asymptotic analysis

A useful tool in studying the behaviour of continuous time models is the asymptotic expansion. These are simply approximate analytical solutions of the model equations that are valid only near the boundaries. Their utility lies in the fact

that they can yield analytic insight into the qualitative and quantitative behaviour of the model near the boundaries, which in turn can aid the model analysis or help with the numerical solution of the equations.

The exact way of constructing the expansion varies from problem to problem, but the general idea is to use the smallness of the distance to the boundary, or the greatness of the variable if very far from it, as a simplifying assumption. A fairly generally applicable recipe goes as follows:

1. Guess the limiting form of the solution, oftentimes a power law of the independent variable.
2. Substitute this trial function into the equation to be solved.
3. Expand the equation to leading order by neglecting terms that are guaranteed to be smaller than other terms in the equation.
4. Choose the parameters of the trial function so that a solution matching the boundary constraints are satisfied.

## B References

- Tobias Adrian and Nina Boyarchenko. Interdiary Leverage Cycles and Financial Stability. 2013. URL [http://www.newyorkfed.org/research/staff\\_reports/sr567.pdf](http://www.newyorkfed.org/research/staff_reports/sr567.pdf).
- Philippe Aghion and Peter Howitt. A Model of Growth Through Creative Destruction. *Econometrica*, 60(2):323–351, March 1992. ISSN 00129682. doi: 10.2307/2951599. URL <http://www.jstor.org/stable/2951599>.
- Philippe Aghion and Peter Howitt. *THE ECONOMICS OF GROWTH*. MIT Press, Boston, 2009.
- S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684, 1994. ISSN 0033-5533.
- Yann Algan, Olivier Allais, Wouter J Den Haan, and Pontus Rendahl. Solving and simulating models with heterogeneous agents and aggregate uncertainty. *Handbook of Computational Economics*, 2010.
- Bo Becker and Victoria Ivashina. Cyclicity of credit supply: Firm level evidence. *Journal of Monetary Economics*, 62:76–93, March 2014. ISSN 03043932. doi: 10.1016/j.jmoneco.2013.10.002.

- Ben Bernanke, Mark Gertler, and Simon Gilchrist. The Financial Accelerator in a Quantitative Business Cycle Framework. In John B Taylor and Michael Woodford, editors, *Handbook of Macroeconomics, Volume 1C*, pages 1341–1393. Elsevier Science, North-Holland, 1999.
- Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973. ISSN 0022-3808.
- Patrick Bolton, Hui Chen, and Neng Wang. A Unified Theory of Tobin’s  $q$ , Corporate Investment, Financing, and Risk Management. *The Journal of Finance*, 66(5):1545–1578, October 2011. ISSN 00221082. doi: 10.1111/j.1540-6261.2011.01681.x. URL <http://doi.wiley.com/10.1111/j.1540-6261.2011.01681.x>.
- Markus K Brunnermeier and Yuliy Sannikov. Macro, Money and Finance: A Continuous-Time Approach. In John Taylor and Harald Uhlig, editors, *Handbook of Macroeconomics Vol 2*.
- Markus K. Brunnermeier and Yuliy Sannikov. International Credit Flows, Pecuniary Externalities, and Capital Controls. 2014a.
- Markus K. Brunnermeier and Yuliy Sannikov. A Macroeconomic Model with a Financial Sector. *American Economic Review*, 104(2):379–421, February 2014b. ISSN 0002-8282. doi: 10.1257/aer.104.2.379. URL <http://www.ingentaconnect.com/content/aea/aer/2014/00000104/00000002/art00002>.
- Markus K Brunnermeier and Yuliy Sannikov. A Macroeconomic Model with a Financial Sector: online appendix. *American Economic Review*, (2), February 2014c. ISSN 0002-8282. doi: 10.1257/aer.104.2.379. URL <http://www.ingentaconnect.com/content/aea/aer/2014/00000104/00000002/art00002>.
- Markus K Brunnermeier and Yuliy Sannikov. The I Theory of Money: version of April 2014. 2014d.
- Markus K Brunnermeier and Yuliy Sannikov. The I Theory of Money: version of November 2014. 2014e.
- Markus K Brunnermeier, Thomas M Eisenbach, and Yuliy Sannikov. Macroeconomics with financial frictions: A survey. Technical report, 2012.
- Guillermo A. Calvo. Capital Market Crises. *Journal of Applied Economics*, 1(1):35–54, 1998.
- George M Constantinides. Intertemporal asset pricing with heterogeneous consumers and without demand aggregation. *Journal of Business*, pages 253–267, 1982. ISSN 0021-9398.

- Michael G Crandall, Ishii Hitoshi, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27:1–67, 1992.
- Harry DeAngelo and René M Stulz. Why high leverage is optimal for banks. Technical report, 2013.
- Wouter J. Den Haan. Comparison of solutions to the incomplete markets model with aggregate uncertainty. *Journal of Economic Dynamics and Control*, 34(1):4–27, January 2010. ISSN 01651889. doi: 10.1016/j.jedc.2008.12.010. URL <http://www.sciencedirect.com/science/article/pii/S0165188909001298>.
- Douglas Diamond and Phillip Dybvig. Bank runs, deposit insurance and liquidity. *Journal of Political Economy*, 91:401–419, 1983.
- Douglas W. Diamond. Financial Intermediation and Delegated Monitoring. *The Review of Economic Studies*, 51(3):393, July 1984. ISSN 00346527. doi: 10.2307/2297430. URL <http://restud.oxfordjournals.org/content/51/3/393.short>.
- William Feller. *An Introduction to Probability Theory and its Applications*, volume II. Wiley, New York, 2nd edition, 1971.
- ZHIGANG FENG, JIANJUN MIAO, ADRIAN PERALTA-ALVA, and MANUEL S SANTOS. NUMERICAL SIMULATION OF NONOPTIMAL DYNAMIC EQUILIBRIUM MODELS. *International Economic Review*, 55(1):83–110, February 2014. ISSN 1468-2354. doi: 10.1111/iere.12042. URL <http://dx.doi.org/10.1111/iere.12042>.
- Wendell Fleming and Halil Mete Soner. *Controlled {Markov} Processes and Viscosity Solutions*. Stochastic Modelling and Applied Probability. Springer, New York, 2nd edition, 2006.
- Marvin Frankel. The production function in allocation and growth: a synthesis. *The American Economic Review*, pages 996–1022, 1962. ISSN 0002-8282.
- Crispin Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics. Springer, Berlin, 4th edition, 2009.
- William M Gorman. Separable utility and aggregation. *Econometrica: Journal of the Econometric Society*, pages 469–481, 1959. ISSN 0012-9682.
- Luca Guerrieri and Matteo Iacoviello. OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics*, 70:22–38, March 2015. ISSN 03043932. doi: 10.1016/j.jmoneco.2014.08.005. URL <http://www.sciencedirect.com/science/article/pii/S0304393214001238>.

- Fatih Guvenen. Macroeconomics with heterogeneity: a practical guide. *Economic Quarterly*, (3Q):255–326, 2011.
- Ernst Hairer, Gerhard Wanner, and Syvert P Nørsett. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Series in Computational Mathematics. Springer, Berlin, 1993.
- Andrew G Haldane. On microscopes and telescopes, 2015. URL <http://www.bankofengland.co.uk/publications/Documents/speeches/2015/speech812.pdf>.
- Zhiguo He and Arvind Krishnamurthy. A Model of Capital and Crises. *The Review of Economic Studies*, 79(2):735–777, 2012. ISSN 0034-6527.
- Zhiguo He and Arvind Krishnamurthy. Intermediary Asset Pricing. *American Economic Review*, 103(2):732–770, 2013. ISSN 0002-8282.
- Zhiguo He and Arvind Krishnamurthy. A macroeconomic framework for quantifying systemic risk. 2014.
- B Holmstrom and J Tirole. Financial Intermediation, Loanable Funds, and the Real Sector. *Quarterly Journal of Economics*, 112(3):663–691, 1997.
- Mark Huggett. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of economic Dynamics and Control*, 17(5):953–969, 1993. ISSN 0165-1889.
- Jukka Isohätälä, Feodor Kusmartsev, Alistair Milne, and Donald Robertson. Leverage constraints and real interest rates. 2014a. URL [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2526790](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2526790).
- Jukka Isohätälä, Alistair Milne, and Donald Robertson. The Net Worth Trap: Investment and Output Dynamics in the Presence of Financing Constraints. 2014b. URL [http://www.suomenpankki.fi/en/julkaisut/tutkimukset/keskustelualoitteet/Pages/dp2014\\_26.aspx](http://www.suomenpankki.fi/en/julkaisut/tutkimukset/keskustelualoitteet/Pages/dp2014_26.aspx).
- Nobuhiro Kiyotaki and John Moore. Credit Cycles. *Journal of Political Economy*, 105(2):211–248, 1997.
- Fima C Klebaner. *Introduction to Stochastic Calculus with Applications*. Imperial College Press, 2nd edition, 2005.
- Nataliya Klimenko, Sebastian Pfeil, and Jean-Charles Rochet. Bank Capital and Aggregate Credit. 2015.
- Dirk Krueger and Felix Kubler. Markov equilibria in macroeconomics. In *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, Basingstoke. 2008.

- Per Krusell and Anthony A Smith. Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns. *Macroeconomic dynamics*, 1(02):387–422, 1997. ISSN 1469-8056.
- Per Krusell and Anthony A Jr Smith. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, 1998. ISSN 0022-3808.
- Ricardo Lagos and Randall Wright. A Unified Framework for Monetary Theory and Policy Analysis. *Journal of Political Economy*, 113(3):463–484, 2005.
- P L Lions and A S Sznitman. Stochastic differential equations with reflecting boundary conditions. *Communications on Pure and Applied Mathematics*, 37(4):511–537, 1984. ISSN 1097-0312. doi: 10.1002/cpa.3160370408.
- Lars Ljungqvist and Thomas J Sargent. *Recursive Macroeconomic Theory*. MIT Press, 2nd edition, 2000.
- Robert E Lucas and Nancy L Stokey. Money and Interest in a Cash-in-Advance Economy. *Econometrica*, 55(3):491–513, 1987.
- Robert E Lucas Jr. Asset Prices in an Exchange Economy. *Econometrica*, 46(6):1429–1445, November 1978. ISSN 00129682. doi: 10.2307/1913837. URL <http://www.jstor.org/stable/1913837>.
- N Gregory Mankiw. The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics*, 17(1):211–219, 1986. ISSN 0304-405X.
- Césaire A Meh and Kevin Moran. The Role of Bank Capital in the Propagation of Shocks. *Journal of Economic Dynamics and Control*, 34(3):555–576, 2010.
- Robert C Merton. Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *Rev. Econ. Stat.*, 51(3):247–257, 1969. doi: 10.2307/1926560.
- Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373 – 413, 1971. URL <http://ideas.repec.org/a/eee/jetheo/v3y1971i4p373-413.html>.
- Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer, Berlin, 6th edition, 2003.
- Gregory Phelan. Financial intermediation, leverage, and macroeconomic instability. Technical report, 2014.
- Adriano A Rampini and S Viswanathan. Financial intermediary capital. *Available at SSRN 1785877*, 2012.

- Carmen Reinhart and Kenneth Rogoff. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press, 2009. URL [http://www.amazon.co.uk/This-Time-Different-Centuries-Financial-ebook/dp/B004EYT932/ref=sr\\_1\\_1?ie=UTF8&qid=1404847123&sr=8-1&keywords=this+time+is+different+rogoff](http://www.amazon.co.uk/This-Time-Different-Centuries-Financial-ebook/dp/B004EYT932/ref=sr_1_1?ie=UTF8&qid=1404847123&sr=8-1&keywords=this+time+is+different+rogoff).
- Hannes Risken. *The {Fokker-Planck} Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer, Berlin, 1996.
- Paul M Romer. Growth Based on Increasing Returns Due to Specialization. *American Economic Review*, 77(2):56–62, 1987. ISSN 00028282.
- Paul M Romer. Endogenous Technological Change. *Journal of Political Economy*, 98(5):S71–102, 1990. ISSN 00223808.
- Mark Rubinstein. An aggregation theorem for securities markets. *Journal of Financial Economics*, 1(3):225–244, 1974. ISSN 0304-405X.
- Paul A Samuelson. An exact consumption-loan model of interest with or without the social contrivance of money. *The journal of political economy*, pages 467–482, 1958. ISSN 0022-3808.
- F Smets and R Wouters. Comparing Shocks and Frictions in {US} and Euro Area Business Cycles: A {B}ayesian {DSGE} Approach. *Journal of Applied Econometrics*, 20:161–183, 2005.
- M Woodford. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, 2003.