Optimal Control with Heterogeneous Agents in Continuous Time

Galo Nuño — Banco de España
Benjamin Moll* — Princeton University

September 11, 2014

Abstract

This paper introduces the problem of a planner who wants to control a population of heterogeneous agents subject to idiosyncratic shocks. This is equivalent to a deterministic control problem in which the state variable is a distribution. We show how, in continuous time, the problem can be broken down into a dynamic programming equation plus the law of motion for the distribution, and provide a numerical algorithm to solve it. We particularize this method to analyze constrained efficiency. By comparing the planning solution with the competitive equilibrium we obtain a criterion to check whether a market allocation is constrained efficient.

Keywords: Kolmogorov forward equation, calculus of variations, social welfare function, mean field control.

JEL codes: C6, D3, D5, E2

*The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the European Central Bank or the Bank of Spain. This paper supersedes a previous version entitled "Constrained Social Optimum with Heterogeneous Agents in Continuous Time." The authors are very grateful to Fernando Álvarez, Luca Dedola, Alessio Moro, Giulio Nicoletti, Carlos Thomas, Oreste Tristani and seminar participants at the ECB and CEF 2014 for helpful comments and suggestions. All remaining errors are ours.

1 Introduction

Optimal control is an essential tool in economics and finance. In optimal control, a planner deals with the problem of finding a set of control variables for a given system such that a certain optimality criterion is achieved. The state of the system is typically characterized by a finite number of state variables.[1] Some systems of interest are composed of a very large number of agents; an economy, for example, is composed of millions of households and firms and a network may contain thousands of nodes. In these cases, assuming a continuous distribution of state variables seems to be a reasonable approximation to the real problem under consideration.

The aim of this paper is to introduce optimal control problems in which there is an infinite number of ex-ante identical agents. The state of each of these agents is characterized by a finite set of state variables. The evolution of the idiosyncratic state variables follows a controllable stochastic process, that is, there exists a set of controls that allows the planner to modify the individual state of each agent. The state variables are subject to some random disturbances. In addition there are some aggregate states. In this problem, the aim of the planner is to maximize an optimality criterion over the full distribution (across agents) of state variables.

We focus on the continuous time version of the problem. The key advantage of working in continuous time, compared to discrete time, is that the evolution of the state distribution across agents can be characterized by the Kolmogorov forward (KF) equation (also known as Fokker-Planck equation). This is a partial differential equation (PDE) that describes the evolution of the distribution given the controls. Despite the random evolution of each individual state, the dynamics of the distribution are deterministic due to the Law of Large Numbers. Thanks to this, the control of an infinite number of agents subject to idiosyncratic shocks can be expressed as the deterministic problem of controlling a distribution that evolves according to the KF equation, subject to the aggregate constraints. As we work with state distributions, we should employ calculus of variations in order to solve the problem.

The main contribution of the paper is to present the necessary conditions for a solution to this problem. These conditions are characterized by a system of two coupled PDEs: a Hamilton-Jacobi-Bellman (HJB) equation and the KF equation that describes the evolution of the distribution of agents. In addition, there is a number of equations that relate some nonlinear functions of the aggregated variables to the state distribution. The role of the value function in individual stochastic control is played here by the *marginal social value*, that is, the social value of an agent in a certain state.

A particular case of interest is the analysis of constrained efficiency in heterogeneous agents economies. The constrained efficient allocation is defined as the one of a social planner who

[1] See Bertsekas (2005, 2012) or Fleming and Soner (2006).

maximizes a utilitarian social welfare function (SWF) subject to the same equilibrium budget constraints and competitive price setting as the individual agents. The planner cannot complete markets or use any transfers between agents. We employ the techniques developed here to compare the solution of this problem with the case of a competitive equilibrium in which each agent maximizes its own discounted utility subject to its state dynamics taking the aggregate conditions and the dynamics of the other agents as given. Our results show that both cases yield a system of two coupled PDEs, with the difference that the variable in the decentralized HJB is the individual value function whereas in the social planning case it is the marginal social value. In addition, in the planner's HJB there is an extra term that accounts for the impact of the aggregate variables on the social value. When this extra term is zero, the competitive equilibrium yields the constrained-efficient allocation. Therefore, by checking whether this term is zero in the decentralized economy we have a criterion to evaluate the constrained optimality of a heterogeneous-agent model which does not require to solve the planner's problem. If the term is not zero we show how the constrained-inefficient stationary competitive allocation can be replicated by a planner with a non-utilitarian SWF.

We introduce a numerical algorithm to solve optimal control problems with heterogeneous agents. We build on the recent work of Achdou et al. (2014) on finite difference methods to solve the coupled HJB and KF equations and present a relaxation algorithm that finds the value function, distribution and optimal policies as well as the Lagrange multipliers of the aggregate conditions. We illustrate it with a example: the computation of the constrained efficient solution in a continuous time version of the neoclassical growth model with uninsurable idiosyncratic shocks as in Davila et al. (2012).[2]

Literature review. Our paper is related to the large literature studying general equilibrium models with heterogeneous agents. Early contributions are Bewley (1986), Imrohoroğlu (1989), Huggett (1993), and Aiyagari (1994). See Heathcote, Storesletten and Violante (2009) for a recent survey. Some existing papers analyze heterogeneous agent models in continuous time. Examples are Luttmer (2007), Alvarez and Shimer (2011) and Achdou et al. (2014). These models can be seen as particular cases of *mean field games* (MFG). Introduced by Lasry and Lions (2007), MFG equilibria are a generalization of Nash equilibria in stochastic differential games as the number of players tend to infinite and agents only care about the distribution of other players' states. The name is borrowed from the mean-field approximation in statistical physics, in which the effect on any given individual of all the other individuals is approximated by a single averaged effect. MFG are characterized by a HJB equation describing the value function of each player and a KF equation describing the evolution of the state distribution. See, for example, Guéant, Lasry and

[2] Davila et al. (2012) compute their numerical results using a 3-state Markov chain whereas we provide a method to compute it for any difussion. In particular we consider an Ornstein–Uhlenbeck process, which is the continuous-time counterpart of an AR(1).

Lions (2011) or Carmona, Delarue and Lachapelle (2013).

In particular, our paper is related to the case of *mean field control*. This is the optimal control counterpart to MFG in which a planner should control a population of infinite agents characterized by a distribution evolving according to the KF equation. A recent survey may be found in Bensoussan, Frehse and Yam (2013). In economics, according to our knowledge Lucas and Moll (2013) were the first ones to employ mean field control techniques to obtain the optimal allocation in a continuous time model with heterogeneous agents. However, their formulation does not allow the possibility of including aggregate constraints, such as price equations, which are prevalent in most economic problems. Our work builds on theirs by introducing both a general theory and a computational algorithm in order to solve general problems in economics.

The structure of the paper is as follows. In section 2 we introduce an example as a motivation. In section 3 we analyze the general case and present the main results. In section 4 we compare the optimal control with the competitive equilibrium. In section 5 we introduce the numerical algorithm. Finally, in section 6 we conclude.

2 An example: constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks

We begin with an example. It is a continuous-time counterpart to the Aiyagari-Bewley-Huggett economy described in Aiyagari (1994).

Workers. There is a continuum of mass unity of workers that are heterogeneous in their wealth a and labor productivity z. The state of the economy is the joint distribution $f(a, z)$. Workers have standard preferences over utility flows from future consumption c_t^j discounted at rate $\rho \geq 0$:

$$E \int_t^\infty e^{-\rho(s-t)} u(c_t^j) ds, \tag{1}$$

The function u is strictly increasing and strictly concave. A worker supplies z_t^j efficiency units of labor to the labor market and these get valued at wage w. A worker's wealth evolves according to

$$da_t^j = \left[w z_t^j + r a_t^j - c_t^j \right] dt, \tag{2}$$

where r is the interest rate. Workers also face a borrowing limit,

$$a_t^j \geq \bar{a} \tag{3}$$

where $\bar{a} \leq 0$. Finally, a worker's efficiency units evolve stochastically over time on a bounded interval $[\underline{z}, \bar{z}]$ with $\underline{z} \geq 0$, according to a stationary diffusion process that either stays in the interval by itself or is reflected at the boundaries:

$$dz_t^j = \eta(z_t^j)dt + \sigma_z(z_t^j)dB_t^j,$$

where B_t^j is a Brownian motion. We impose an additional restriction on the borrowing limit, $-\bar{a} \leq w\underline{z}/r$, that is the borrowing limit is at least as tight as the "natural borrowing limit."

In order to avoid the introduction of *state constraints* such as (3), we may modify the utility function in order to introduce a penalty function $\xi(a)$, such that

$$\xi(a) = \begin{cases} 0, \text{ if } a_t \geq \bar{a}, \\ -\infty, \text{ if } a_t < \bar{a}, \end{cases}.$$

Both problems are equivalent and the second one fits in the framework described below.[3]

The optimal value function results in

$$\hat{V}(a,z) = \sup_{c\geq 0} E \int_t^\infty e^{-\rho(s-t)} \left[u(c) + \xi(a)\right] ds,$$

subject to the price vector $\Gamma = [w, r]'$. The Hamilton-Jacobi-Bellman of this problem is

$$\rho\hat{V} = \sup_{c\in A} u(c) + \xi(a) + (wz + ra - c)\frac{\partial \hat{V}}{\partial a} + \eta(z)\frac{\partial \hat{V}}{\partial z} + \frac{\sigma_z^2(z)}{2}\frac{\partial^2 \hat{V}}{\partial z^2}. \tag{4}$$

Firms. There is representative firm with a constant returns to scale production function $Y = F(K, L)$. The total amount of capital supplied in the economy equals the total amount of wealth

$$K = \int af(a,z)dadz, \tag{5}$$

and we normalize the total amount of labor supplied in the economy to one. Capital depreciates at rate δ. Since factor markets are competitive, the wage and the interest rate are given by

$$r = \frac{\partial}{\partial K}F(K,1) - \delta, \qquad w = \frac{\partial}{\partial L}F(K,1). \tag{6}$$

The stationary distribution of agents is given by the KF equation

$$0 = -\frac{\partial}{\partial a}\left[(wz + ra - c)f\right] - \frac{\partial}{\partial z}\left[\eta(z)f\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma_z^2(z)f\right], \tag{7}$$

[3] See Capuzzo-Dolcetta and Lions (1990).

and the it should satisfy the normalization

$$\int f(a,z)dadz = 1.$$

We may define a stationary competitive equilibrium in this economy.

Definition 1 (Competitive equilibrium) *A stationary competitive equilibrium is a vector of prices $\Gamma = [w, r]'$, a value function $\hat{V}(a,z)$, a consumption policy $c(a,z)$ and a distribution $f(a,z)$ such that (i) given Γ and f, $\hat{V}$ solves the HJB equation (4) and the optimal control is $c(a,z)$; (ii) given c and Γ, f solves the KF equation (7); and given c and f, the price vectors satisfy the price equation (6).*

Constrained-efficient solution. We study the allocation of a benevolent social planner subject to the same constraints as the individual agents in the competitive equilibrium. Namely, market prices Γ still satisfy equations (6) and the aggregate distribution $f(a,z)$ the KF equation (7). The social planner chooses the consumption policy $c(a,z)$ of every agent $j \in [0,1]$. The planner also chooses the vector of prices Γ given the constraints. The social planner chooses the controls and the prices in order to maximize a discounted SWF $W(u)$ that aggregates individuals' utilities $u(c)$ into a social utility:

$$W[f] = \int \omega(a,z) [u(c) + \xi(a)] f(a,z)dadz, \tag{8}$$

where $\omega(a,z)$ are the Pareto weights. If $\omega(a,z) = 1$ then we have a purely utilitarian SWF. Notice that W is a functional as it maps a distribution function to a social welfare. The problem of the social planner is to maximize

$$\sup_{w,r,\ c\geq 0} \int_t^\infty \int e^{-\rho(s-t)} \omega(a,z) [u(c) + \xi(a)] f(a,z)dadz \tag{9}$$

subject to the KF equation (7) and to the price equations (6). Notice that this not a standard optimal control problem as the state variable $f(a,z)$ is a distribution and so it is the control $c(a,z)$. We devote the rest of the paper to develop a general theory for this kind of problems and we come back to this example below.

3 General approach

3.1 Statement of the problem

Agents' problem. We consider a continuous-time infinite-horizon economy. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ be a filtered probability space. There is a continuum of unit mass of agents indexed by $j \in [0,1]$. Let B_t^j is a n-dimensional $\mathcal{F}_t$-Brownian motion. Let X_t^j denote the state of an agent j at time $t \in [0, \infty)$. The state evolves according to a multidimensional Itô process of the form

$$dX_t^j = b\left(X_t^j, \mu(t, X_t^j), \Gamma_t\right) dt + \sigma\left(X_t^j\right) dB_t^j, \tag{10}$$

where $X_t^j \in \mathbb{R}^n$. Γ_t is a p-dimensional vector, which we denote as the price vector.[4] Here b (drift) and σ (diffusion) are measurable functions $b : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^n$, $b \in C^1(\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p)$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma \in C^2(\mathbb{R}^n)$.[5] All elements outside of the diagonal in $\sigma\left(X_t^j\right)$ are zero.[6] Notice that all the agents are identical in their drift and diffusion coefficients but potentially differ in their state and in the realization of the idiosyncratic Brownian motions. Notice too that only the drift coefficient can be controlled.[7]

The policy $\mu : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^m$, $m \in \mathbb{N}$, is a $m-$dimensional $\mathcal{F}_t$- adapted Markov control. The control $\mu(t, x) \in \mathbb{R}^m$ is admissible if for any initial point (t, x) such that $X_t^j = x$ the stochastic differential equation (10) has a unique solution.[8] We denote A as the space of all admissible controls contained in the set of all Markov controls. The control strategy is the same for every agent, but it depends on time and on the state of this particular agent.

Aggregate distribution and aggregate variables. Assume that the transition measure of X_t^j has a density $f(t, x; x_0^j) : [0, \infty) \times \mathbb{R}^n \to [0, 1]$, ($f \in C^2([0, \infty) \times \mathbb{R}^n)$), i.e., that

$$\forall \varphi \in C^2(\mathbb{R}^n), \qquad E\left[\varphi(X_t^j) | X_0^j = x_0\right] = \int_{\mathbb{R}^n} \varphi(x) f(t, x; x_0^j) dx.$$

---

[4]This is the price vector in our example, although it may represent other variables in different problems.

[5]$C^k(Q)$ is the set of all k-times continuously differentiable functions on Q.

[6]To ensure the existence of a solution of the stochastic differential equation (10), assume that

$$\begin{aligned} \|b(x,\mu)\| &\leq \kappa\left(1 + \|x\| + |\mu|\right), \quad \|\nabla_\Gamma b(x,\mu)\| &+ \|\nabla_x b(x,\mu)\| \leq \kappa, \\ \|\sigma(x)\| &\leq \kappa\left(1 + \|x\|\right), \quad \|\nabla_x \sigma(x)\| &\leq \kappa, \end{aligned}$$

for all $x \in \mathbb{R}^n$, some constant κ and $\mu \in \mathbb{R}^m$. Here $\|\sigma\|^2 = \sum_{i,j=1}^n |\sigma_{ij}|^2$. See Øksendal (2010) or Fleming and Soner (2006).

[7]This is done for simplification, the results in this paper can be extended to the case of controlled diffusion.

[8]This is guarantee if $\forall t < \infty$:

$$E \int_0^t \left\|\mu(s, X_s^i)\right\|^j ds < \infty \text{ for } j \in \mathbb{N}.$$

The initial distribution of X_t^j at time $t=0$ is $f(0,x)=f_0(x)$.

In this case, the dynamics of the distribution of agents $f(t,x)$ are given by the KF equation

$$\frac{\partial f}{\partial t} = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[b_i\left(x,\mu,\Gamma\right)f\right]+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}\left[\sigma_{ii}^2(x)f\right], \tag{11}$$

$$\int_{\mathbb{R}^n} f(t,x)dx = 1. \tag{12}$$

The vector of aggregate variables is determined by a system of p equations:

$$\Gamma_k(t)=G_k\left(\int_{\mathbb{R}^n} h(x,\mu)f(t,x)dx\right), \quad k=1,..,p. \tag{13}$$

where $G_k:\mathbb{R}\to\mathbb{R}$, $h:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, $G_k\in C^1\left(\mathbb{R}\right)$, $h\in C^1\left(\mathbb{R}^n\times\mathbb{R}^m\right)$.

Optimal control. We now study the allocation of a planner who chooses a vector of control variables $\mu(t,X_t^j)$ to be applied to every agent $j\in[0,1]$ with state dynamics (10). The planner also chooses the vector of prices Γ_t given the constraints (13). The planner chooses the controls and the prices in order to maximize a discounted functional

$$W\left[f,\mu\right]=\int_{\mathbb{R}^n} g(x,\mu)f(t,x)dx, \tag{14}$$

where $g\in C^1\left(\mathbb{R}^n\times\mathbb{R}^m\right)$ is an increasing concave function.[9] Notice that W is a functional that maps from the space of densities and the space of controls into the real numbers.

In this case, instead of an optimal value function, we have an optimal value functional $V[f]$, defined as

$$V\left[f(t,\cdot)\right]\equiv\sup_{\Gamma,\ \mu\in A,}\int_t^{\infty}\int_{\mathbb{R}^n} e^{-\rho(s-t)}g(x,\mu)f(t,x)dxds=\int_t^{\infty} e^{-\rho(s-t)}W\left[f,\mu^*\right]ds \tag{15}$$

subject to the KF equation (11, 12) and to the price equations (13). The supremum is taken over the family A of admissible controls and over the family of market prices. Notice that Γ appears now as a control. We assume the transversality condition

$$\lim_{t\uparrow\infty} e^{-rt}V\left[f(t,\cdot)\right]=0. \tag{16}$$

The constrained optimal control problem with heterogeneous agents is an extension of the classical deterministic optimal control problem to an infinite dimensional setting, in which the state is the whole distribution of individual states $f(t,x)$. The key point is that the law of motion

[9] $\|g(x,\mu)\|\leq\kappa\left(1+\|x)\|^k+\|\mu\|^k\right)$ for a suitable constant κ.

of this distribution is given by the KF equation.

Definition 2 (optimal control) *An optimal control problem is an optimal value functional* $V[f(t,\cdot)]$*, a control function* $\mu^* = \mu^*(t,x) \in A$ *and a price vector* $\Gamma^*(t)$ *such that the control and the prices maximize (15) subject to the KF equation (11, 12) and the price equations (13).*

3.2 A verification theorem

Marginal social value. The price equations (13)

$$\Gamma_k^*(t) = G_k\left(\int_{\mathbb{R}^n} h(x,\mu^*)f(t,x)dx,\right), \quad k = 1,..,p,$$

can be expressed in functional form

$$G_k^{-1}\left(\Gamma_k^*(t)\right) = H[f(t,\cdot),\mu^*], \quad k = 1,..,p,$$

where

$$H[f(t,\cdot),\mu^*] = \int_{\mathbb{R}^n} h(x,\mu^*)f(t,x)dx,$$

and $G_k^{-1}(\Gamma)$ is the inverse function of G_k.

If an optimal Markov control $\mu^* = (\mu_1^*,...,\mu_m^*)$ and an optimal price vector $\Gamma^*(t)$ exist and satisfy the price constraints (13), they should be an extremal of the functional Lagragian

$$\mathcal{L}[f(t,\cdot),\mu;\Gamma] = \int_t^\infty e^{-\rho(s-t)}\left\{W[f(s,\cdot),\mu] + \sum_{k=1}^p \lambda_k(s)\left(H[f(s,\cdot),\mu] - G_k^{-1}(\Gamma_k(s))\right)\right\} ds, \quad (17)$$

where $\lambda_k(t) : \mathbb{R}^+ \to \mathbb{R}$ are the Lagrange multipliers of the problem. Notice that in the case of optimal controls and prices the Lagragian coincides with the optimal value function $\mathcal{L}[f(t,\cdot),\mu^*;\Gamma^*] = V[f(t,\cdot)]$.

Assume that $f(t,\cdot) \in L^2(\mathbb{R}^n)$, that is, f^2 is Lebesgue-integrable in $\mathbb{R}^n$. This is a mild assumption in most of the problems of interest. Notice that $\mathcal{L}[f(t,\cdot),\mu;\Gamma]$ is a linear functional in f and μ. Therefore we may apply the Riesz representation theorem and state that there exists a unique function $v(t,\cdot;\mu,\Gamma) \in L^2(\mathbb{R}^n)$ such that

$$\mathcal{L}[f(t,\cdot),\mu;\Gamma] = \langle v(t,\cdot;\mu,\Gamma), f(t,\cdot)\rangle_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} v(t,x;\mu,\Gamma)f(t,x)dx.^{10} \quad (18)$$

In the case of the optimal control and prices

$$V[f(t,\cdot)] = \mathcal{L}[f(t,\cdot),\mu^*;\Gamma^*] = \int_{\mathbb{R}^n} v(t,x;\mu^*,\Gamma^*)f(t,x)dx,$$

that is, the optimal value functional can be expressed as average across agents of $v(t,x) = v(t,x;\mu^*,\Gamma^*)$. Given (18), we can compute $v(t,x)$ as

$$v(t,x) = \frac{\delta V\left[f(t,\cdot)\right]}{\delta f(t,x)} = \sup_{\Gamma,\ \mu\in A,} \frac{\delta \mathcal{L}\left[f(t,\cdot),\mu;\Gamma\right]}{\delta f(t,x)}, \tag{19}$$

that is, v is the functional derivative of the Lagrangian with respect to the state distribution. Lucas and Moll (2013) provide an economic interpretation of $v(t,x)$ as the *marginal social value* at time t of an agent in state x. Appendix A briefly introduces the concept of functional derivative.

Necessary conditions. Given (19), we provide necessary conditions to the problem.

Proposition 3 (Necessary conditions) *Assume that a marginal social value v, an optimal admissible Markov control μ^* and an optimal price vector Γ^* exist. Then, they satisfy*

$$\rho v = \sup_{\mu\in A} g(x,\mu) + \sum_{k=1}^{p}\lambda_k(t)h(x,\mu) + \frac{\partial v}{\partial t} + \sum_{i=1}^{n} b_i\left(x,\mu,\Gamma^*\right)\frac{\partial v}{\partial x_i} + \sum_{i=1}^{n}\frac{\sigma_{ii}^2(x)}{2}\frac{\partial^2 v}{\partial x_i^2}, \tag{20}$$

with $\lambda_k(t) : [0,\infty) \to \mathbb{R},\ k = 1,..,p$ *:*

$$\lambda_k(t) = -G_k' \int_{\mathbb{R}^n} v(t,y)\left(\sum_{i=1}^{n}\left[\frac{\partial^2 b_i\left(y,\mu^*,\Gamma^*\right)}{\partial \Gamma_k \partial x_i} f(t,y) + \sum_{j=1}^{m}\frac{\partial^2 b_i}{\partial \Gamma_k \partial \mu_j}\frac{\partial \mu_j^*}{\partial x_i} f + \frac{\partial b_i}{\partial \Gamma_k}\frac{\partial f}{\partial x_i}\right]\right) dy. \tag{21}$$

and $G_k' = G_k'\left(\int_{\mathbb{R}^n} h(x,\mu) f(t,x)dx\right)$ *, together with the KF equation (11, 12), price equations (13) and the transversality condition*

$$\lim_{t\uparrow\infty} e^{-rt} v(t,x) = 0. \tag{22}$$

The proof can be found in the Appendix B. This proposition is the central result of the paper. It provides a system formed by a HJB, a KF and a number of price equations which link the dynamics of $v(t,x)$, $\mu^*(t,x)$, $\Gamma^*(t)$ and $f(t,x)$. The coupled system is forward-backward in the sense that the HJB equation for v is a function of the distribution f and it has boundary conditions at time $t \uparrow \infty$, whereas the KF is a function of the optimal controls μ and it has boundary conditions at time $t = 0$.[11]

4 Competitive equilibrium and constrained optimality

In this section we compare the constrained solution of the problem with the competitive equilibrium.

[11] Notice that we do not impose that v should be differentiable. This is due to the fact that the result is general enough to accommodate *viscosity solutions* as it will be described below.

4.1 Competitive equilibrium

Imagine that, instead of a social planner, individual agents' in the problem above maximize their discounted utility. The optimal value function $\hat{V}(t,x)$ is defined as

$$\hat{V}(t,x) = \sup_{\mu \in A} E \int_t^\infty e^{-\rho(s-t)} u(X_s^j, \mu) ds, \tag{23}$$

subject to (10) and $X_t^j = x$, where utility $u(x,\mu) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $\rho > 0$ is a constant. We assume that

$$\lim_{t \uparrow \infty} e^{-\rho t} \hat{V}(t,x) = 0. \tag{24}$$

The solution to this problem is given by the HJB equation

$$\rho \hat{V} = \frac{\partial \hat{V}}{\partial t} + \sup_{\mu \in A} u(x,\mu) + \sum_{i=1}^n b_i(x,\mu,\Gamma) \frac{\partial \hat{V}}{\partial x_i} + \sum_{i=1}^n \frac{\sigma_{ii}^2(x)}{2} \frac{\partial^2 \hat{V}}{\partial x_i^2}. \tag{25}$$

Its solution is the optimal value function $\hat{V}(t,x)$ and the control strategy $\mu(t,x)$.

We may define a competitive equilibrium in this economy.

Definition 4 (Competitive equilibrium) *A competitive equilibrium is a vector of prices Γ_t, a value function $\hat{V}(t,x)$, a control $\mu(t,x)$ and a distribution $f(t,x)$ such that (i) given Γ_t and $f(t,x)$, $\hat{V}(t,x)$ solves the HJB equation (25) and the optimal control is $\mu(t,x)$; (ii) given $\mu(t,x)$ and Γ_t, $f(t,x)$ solves the KF equation (11, 12); and given $\mu(t,x)$ and $f(t,x)$, the price vectors satisfy the price equation (13).*

This definition of competitive equilibrium is the continuous-time counterpart to the standard recursive competitive equilibrium in models with heterogeneous agents à la Aiyagari-Bewley-Huggett.

4.2 Constrained optimality of the competitive equilibrium

Imagine that a benevolent social planner chooses the controls and the prices in order to maximize a discounted SWF, W, that aggregates individuals' utilities $u(x,\mu)$ into a social utility. We consider the (symmetric) generalized utilitarian functional

$$W[f,\mu] = \int_{\mathbb{R}^n} \omega(x) u(x,\mu) f(t,x) dx, \tag{26}$$

where $\omega(x)$ are the Pareto weights. If $\omega(x) = 1$ then we have a purely utilitarian SWF. This is just a particular case of the general approach described in the previous section, and therefore the solution is given by Proposition 3 with $g(x,\mu) = \omega(x)u(x,\mu)$.

Notice that the necessary conditions for a constrained social optimum are the same as for a competitive equilibrium except that the individual HJB equation (25) and the individual optimal policy functions $\hat{V}(t,x)$ are now replaced by the planner's HJB equation (20) and the marginal social value $v(t,x)$. However, both equations are equal except that in the individual HJB we have the individual utility $u(x,\mu)$ and in the planner's HJB we have instead the sum $\omega(x)u(x,\mu) + \sum_{k=1}^{p} \lambda_k(t) G'_k h(x,\mu)$. Therefore, it is trivial to prove the following corollary.

Corollary 5 (Constrained optimality of the competitive equilibrium) *A competitive equilibrium is constrained efficient in the utilitarian sense* ($\omega(x) = 1$) *if*

$$\sum_{k=1}^{p} \tilde{\lambda}_k(t) h(x,\mu) = 0, \tag{27}$$

where $\tilde{\lambda}_k(t)$ *are given by*

$$\tilde{\lambda}_k(t) = -G'_k \int_{\mathbb{R}^n} \hat{V}(t,y) \left(\sum_{i=1}^{n} \left[\frac{\partial^2 b_i(y,\mu,\Gamma)}{\partial \Gamma_k \partial x_i} f + \sum_{j=1}^{m} \frac{\partial^2 b_i}{\partial \Gamma_k \partial \mu_j} \frac{\partial \mu_j}{\partial x_i} f + \frac{\partial b_i}{\partial \Gamma_k} \frac{\partial f}{\partial x_i} \right] \right) dy. \tag{28}$$

Notice that in this case $\hat{V} = v$, that is, the marginal social value equals the individual value. Therefore, it is enough to solve the competitive equilibrium and to compute (27) to check whether it is constrained efficient. It is not necessary to solve the social planner's problem.

In the case that the competitive equilibrium is not constrained efficient, we can provide, given a non-zero utility, some information regarding the aggregate preferences implicit in the market allocation.

Corollary 6 (Equivalent Pareto weights) *Assume that* $\forall x,\ u(x,\mu) \neq 0$. *Then any stationary competitive equilibrium allocation can be replicated by a constrained social planner with Pareto weights*

$$\omega(x) = 1 - \frac{\sum_{k=1}^{p} \tilde{\lambda}_k h(x,\mu)}{u(x,\mu)}. \tag{29}$$

This corollary states that, given a certain (stationary) market allocation, it is always possible to obtain a SWF that would produce the same allocation. This Pareto weights summarize all the social preferences and can be used to compare alternative market allocations. Of course if the competitive equilibrium is constrained efficient and (27) is satisfied then $\omega(x) = 1$.

5 The example revisited

Given the results above we can compute the constrained efficient solution to the Aiyagari-Bewley-Huggett economy. We do not discuss issues related to the existence or uniqueness of the solution, but we provide a numerical algorithm that can find a solution to the optimal control problem.

5.1 Overview

Given the notation introduced in section 3, $x = [a, z]'$, $\Gamma = [w, r]'$, $\mu(x) = c$, $b(x, \mu, \Gamma) = [(wz + ra - c), \eta(z)]'$, $\sigma(x) = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_z(z) \end{bmatrix}$, $h(x) = a$, $G_1(\cdot) = \frac{\partial}{\partial K} F(\cdot, 1) - \delta$ and $G_2(\cdot) = \frac{\partial}{\partial L} F(\cdot, 1)$. The optimal value functional $V[f(\cdot)]$ is given by

$$V[f(\cdot)] = \int v(a, z) f(a, z) dadz = \sup_{w,r;\ c\in A,} \int_t^{\infty} \int e^{-\rho(s-t)} [u(c) + \xi(a)] f(a, z) dadz,$$

subject to (7) and (6).

The HJB equation of the planner (20) in this problem is

$$\rho v = \sup_{c\in A} u(c) + \xi(a) + (\lambda_1 + \lambda_2) a + (wz + ra - c) \frac{\partial v}{\partial a} + \eta(z) \frac{\partial v}{\partial z} + \frac{\sigma_z^2(z)}{2} \frac{\partial^2 v}{\partial z^2}, \tag{30}$$

which is the same as the one of the individual problem (4) plus the term $(\lambda_1 + \lambda_2) a$.

In this case, the values of the modified Lagrange multipliers (28) are

$$\lambda_1 = -\frac{\partial^2}{\partial K^2} F(K, 1) \int v(a, z) \left(f(a, z) + a \frac{\partial f}{\partial a} \right) dadz. \tag{31}$$

$$\lambda_2 = -\frac{\partial^2}{\partial K \partial L} F(K, 1) \int v(a, z) z \frac{\partial f}{\partial a} dadz, \tag{32}$$

where K is the aggregate capital given by (5). Notice that λ_1 and λ_2 play a similar role as the net effect Δ in Davila et al. (2012). If we define an auxiliary value functional

$$\tilde{V}[f(\cdot); r, w] \equiv \sup_{c\in A} \int_t^{\infty} \int e^{-\rho(s-t)} [u(c) + \xi(a)] f(a, z) dadz,$$

subject to (7), this is the value functional optimal with respect to individual consumption but without considering the impact on prices. For the optimal prices r^* and w^* we have that

$$\tilde{V}[f(\cdot); r^*, w^*] = V[f(\cdot); r, w].$$

Then, using the results in Appendix B, we have that

$$\lambda_1 = \frac{\partial^2}{\partial K^2} F(K,1) \frac{\partial \tilde{V}[f(\cdot); r^*, w^*]}{\partial r}, \quad (33)$$

$$\lambda_2 = \frac{\partial^2}{\partial K \partial L} F(K,1) \frac{\partial \tilde{V}[f(\cdot); r^*, w^*]}{\partial w}, \quad (34)$$

that is, the Lagrange multipliers reflect the marginal impact on aggregate welfare of changing the prices. In the case of a state constraint such as (3), they may be different from zero.

5.2 Numerical solution

Here we provide a numerical algorithm to solve the model. The algorithm can easily be extended to tackle the general family of problems. In order to compute the numerical solution to the optimal control problem we employ a relaxation method. Given $\theta \in (0,1)$, begin with an initial guess of the aggregate capital K^0 and the Lagrange multipliers $\lambda_1^0 = \lambda_2^0 = 0$, set $n = m = 0$.[12] Then:

1. Compute $r^n = \frac{\partial}{\partial K} F(K^n, 1) - \delta$ and $w^n = \frac{\partial}{\partial L} F(K^n, 1)$.

2. Given r^n and w^n, solve the planner's HJB equation (30) to obtain an estimate of the value function v^n and of the consumption c^n.

3. Given c^n, solve the KF equation (4) and compute the aggregate distribution f^n.

4. Compute the aggregate capital stock $\hat{K}^n = \int f^n da dz$.

5. Compute $K^{n+1} = \theta K^n + (1-\theta)\hat{K}^n$. If K^{n+1} is close enough to K^n, stop. If not set $n := n+1$ and go to step 1.

6. Compute the Lagrange multipliers $\hat{\lambda}_1^m$ and $\hat{\lambda}_2^m$ using (31) and (32).

7. Compute $\lambda_i^{m+1} = \theta \lambda_i^m + (1-\theta)\hat{\lambda}_i^m$, $i = 1,2$. If λ_1^{m+1} and λ_2^{m+1} are close enough to λ_1^m and λ_2^m, stop. If not set $m := m+1$ and go to step 1.

In order to solve the HJB and the KF equations, we employ a finite difference method described in Appendix C. It approximates the value function $V(a,z)$ and the distribution $f(a,z)$ on a finite grid with steps Δa and Δz : $a \in \{a_1, ..., a_I\}$, $z \in \{z_1, ..., z_J\}$.[13] We use the notation $V_{i,j} \equiv V(a_i, z_j)$,

[12] Do not confuse the use of n and m here as indexes with the state and control dimension in section 3.

[13] Notice that subindexes i and j have a different meaning here than in the previous sections.

$f_{i,j} \equiv f(a_i, z_j)$, $i = 1, ..., I$; $j = 1, ..., J$. In this case $\hat{K} = \sum_{i=1}^{I} \sum_{j=1}^{J} f_{i,j} \Delta a \Delta z$ and

$$\lambda_1 \approx -\left(\frac{\partial^2}{\partial K^2} F(K,1)\right) \sum_{i=1}^{I} \sum_{j=1}^{J} V_{i,j} \left(f_{i,j} + a_i \frac{f_{i+1,j} - f_{i,j}}{\Delta a}\right) \Delta a \Delta z,$$

$$\lambda_2 \approx -\left(\frac{\partial^2}{\partial K \partial L} F(K,1)\right) \sum_{i=1}^{I} \sum_{j=1}^{J} V_{i,j} z_j \frac{f_{i+1,j} - f_{i,j}}{\Delta a} \Delta a \Delta z.$$

As discussed in Achdou et al. (2014), the appropriate solution concept of HJB equation with state constraints is that of a "viscosity solution" (Crandall and Lions, 1983; Crandall, Ishii and Lions, 1992). The proposed finite difference method converges to the unique viscosity solution of this problem (Barles and Souganidis, 1991).

5.3 Results

Calibration. We employ a similar calibration to the one in Aiyagari (1994). We consider the year as the unit of time. The utility function is CRRA $u(c) = \frac{c^{1-\phi}}{1-\phi}$, with $\phi = 3$. The discount rate ρ is set to 0.04. The production function is Cobb-Douglas $F(K, L) = ZK^{\alpha} L^{1-\alpha}$, with $Z = 1$, $\alpha = 0.36$. The depreciation rate is $\delta = 0.08$. The idiosyncratic productivity shock z_t follows a reflected Ornstein–Uhlenbeck process with unit mean:

$$dz_t = \eta(z_t)dt + \sigma_z(z_t)dB_t = \epsilon(\varpi - z)dt + \sigma_z dB_t,$$

where $\epsilon = 0.5$ and $\sigma_z = 0.2$. The bounded interval is $[\underline{z}, \bar{z}] = [0.5, 1.5]$. The parameter ϖ is set to 1 so that the mean of z_t is also 1. The borrowing constrain $\bar{a}$ is set to -1.

The simulation parameters are the following. The range for a is $[-1, 30]$ and the number of grid points is set to $I = 100$ and $J = 40$. The relaxation parameter θ is 0.99.

Constrained efficient solution. We solve the constrained optimum of the model. Figure 1 displays the savings functions

$$s(a, z) = wz + ra - c(a, z),$$

as well as the state distribution $f(a, z)$. The values of the main aggregate variables are shown in Table 1. Aggregate capital is 5.27. The values of the Lagrange multipliers are $\lambda_1 = 1.7374$ and $\lambda_2 = 0.3007$. Given equations (33) and (34), the fact that λ_1 and λ_2 are positive indicates that an increase in wages or interest rates would be welfare improving, that is, both prices would be higher in the unconstrained first best. The constrained inefficiency of the competitive equilibrium and the higher prices in the first best allocation are in line with Davila et al. (2012).

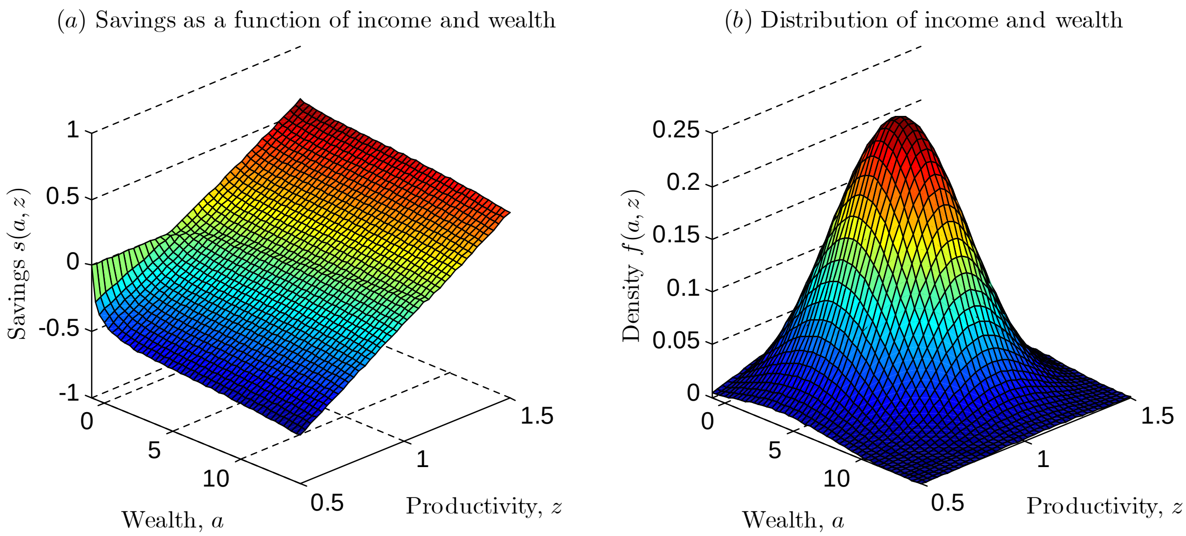

Figure 1: Savings policy and distribution of income and wealth: constrained optimum.

Table 1. Comparison between the constrained optimum and the competitive equilibrium

	Constrained optimum	Competitive equilibrium
Aggregate capital	5.2740	5.7534
Output	1.8196	1.8775
Interest rate (%)	4.42	3.75
Capital-output ratio	2.8985	3.0644
Consumption	1.4002	1.4198

Competitive equilibrium. We also solve the competitive equilibrium of this economy. The numerical procedure can be seen as a particular case of the one described above in which λ_1 and λ_2 are set to zero. Figure 2 displays the results. Aggregate capital is 5.75, larger than in the constrained optimum.

6 Conclusions

This paper introduces the problem of a planner who tries to control a population of heterogeneous agents subject to idiosyncratic shocks in order to maximize an optimality criterion related to the distribution of states across agents. If the problem is analyzed in continuous time, the KF equation provides a deterministic law of motion of the entire distribution of state variables across agents. The problem can thus be analyzed as a deterministic optimal control in which both the control and the state are distributions. We provide necessary conditions by combining dynamic programming with calculus of variations. If a solution to the problem exists and satisfies some differentiability

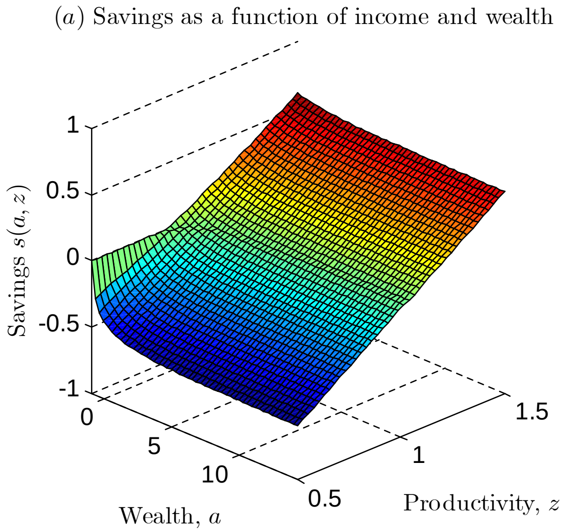

Figure 2: Savings policy and distribution of income and wealth: competitive equilibrium.

conditions, we show how it should satisfy a system of PDEs including a generalization of the HJB equation and a KF equation.

As an example, we employ this technique to analyze the welfare properties of heterogeneous-agent models with idiosyncratic shocks. In particular, we analyze the constrained social optimum in which a social planner maximizes a SWF subject to the same equilibrium budget constraints and competitive price setting as the individual agents. We introduce two main results. First, we provide a simple criterion to check whether a competitive equilibrium is constrained efficient. The criterion does not require the computation of the planner's problem, it is just enough to have the competitive equilibrium solution. Second, we consider the case in which an economy is constrained inefficient and we show that, under some extra assumptions, any stationary competitive equilibrium can be replicated by a constrained social planner.

Finally, we provide a numerical algorithm in order to find the solution to these kind of problems. The algorithm is based on finite difference techniques in order to solve the HJB and KF equations plus a relaxation algorithm. This methodology can be applied to a variety of problems in both micro and macroeconomics.

References

[1] Achdou, Y., J.-M. Lasry, P.-L. Lions and B. Moll (2014), "Heterogeneous Agent Models in Continuous Time," mimeo.

[2] Alvarez, F., and R. Shimer (2011), "Search and Rest Unemployment." *Econometrica*, 79 (1), pp. 75–122.

[3] Aiyagari, R., (1994), "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109 (3), pp. 659-84.

[4] Barles, G. and P. E. Souganidis (1991), "Convergence of Approximation Schemes for Fully Nonlinear Second Order Equations," *J. Asymptotic Analysis*, 4, pp. 271-283.

[5] Bensoussan, A., J. Frehse and P. Yam (2013), *Mean Field Games and Mean Field Typer Control Theory*, Springer, Berlín.

[6] Bewley,T. (1986) "Stationary Monetary Equilibrium With a Continuum of Independently Fluctuating Consumers," in *Contributions to Mathematical Economics in Honor of Gérard Debreu*, ed. By W. Hildenbrand, and A. Mas-Colell. North Holland, Amsterdam.

[7] Brezis, H. (2011), *Functional Analysis, Sobolev Spaces and Partial Differential Equations,* Springer, Berlín.

[8] Capuzzo-Dolcetta, I., and P. L. Lions (1990), "Hamilton-Jacobi Equations with State Constraints," *Transactions of the American Mathematical Society*, 318(2), pp. 643+.

[9] Carmona, R., F. Delarue and A. Lachapelle (2013), "Control of McKean–Vlasov Dynamics versus Mean Field Games," *Mathematics and Financial Economics*, 7 (2), pp. 131-166.

[10] Crandall, M. G., H. Ishii, and P.-L. Lions (1992), "User's guide to viscosity solutions of second order partial differential equations." *ArXiv Mathematics e-prints.*

[11] Crandall, Michael G., and Pierre L. Lions (1983), "Viscosity Solutions of Hamilton-Jacobi Equations." *Transactions of the American Mathematical Society*, 277(1), pp. 1–42.

[12] Dávila, J., J. H. Hong, P. Krusell and J. V. Ríos-Rull (2012) "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks," *Econometrica*, 80(6), pp. 2431-2467.

[13] Fleming, W. H. and H. M. Soner (2006), *Controlled Markov Processes and Viscosity Solutions*, Springer, Berlín.

[14] Gelfand, I. M. and S. V. Fomin (1991), *Calculus of Variations*, Dover Publications, Mineola, NY.

[15] Guéant, O., J.-M. Lasry and P.-L. Lions (2011), "Mean Field Games and Applications," *Paris-Princeton Lectures on Mathematical Finance*, 2011.

[16] Heathcote, J., K. Storesletten, and G. L. Violante (2009), "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*,1(1), pp. 319–354.

[17] Imrohoruglu, A. (1989). "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," J*ournal of Political Economy*, 97(6), pp. 1364-83.

[18] Lasry, J.-M. and P.-L. Lions (2007), "Mean Field Games," *Japanese Journal of Mathematics*, 2 (1), pp. 229-260.

[19] Lucas, R. and B. Moll (2013), "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, forthcoming.

[20] Luttmer, E. G. J. (2007), "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 122 (3), pp. 1103–1144.

[21] Øksendal, B. (2010), *Stochastic Differential Equations: An Introduction with Applications*, Springer, Berlín.

[22] Sagan, H. (1992), *Introduction to the Calculus of Variations*, Dover Publications, Mineola, NY.

Appendix A: Functional Derivative

We introduce here the concept of functional derivative from the calculus of variations, see Gelfand and Fomin (1991) or Sagan (1992).

Definition 7 *Given a differentiable functional $\Phi[\mu_1, ..., \mu_m]$, the* functional derivative *of Φ with respect to $\mu_j : \mathbb{R}^n \to \mathbb{R}$ is*

$$\frac{\delta\Phi}{\delta\mu_j(x)} \equiv \lim_{\varepsilon\to 0} \frac{\Phi[\mu_1, ..., \mu_j(z) + \varepsilon\delta(z-x), ..., \mu_m] - \Phi[\mu_1, ..., \mu_j, ..., \mu_m]}{\varepsilon} = \frac{d}{d\varepsilon}\Phi[\mu_1, ..., \mu_j(z)+\varepsilon\delta(z-x), ..., \mu_m],$$

where $\delta(x)$ is the Dirac delta.

Given the definition of a partial functional derivative, in the case of functionals of the form

$$\Phi[\mu_1, ..., \mu_m] = \int_{\mathbb{R}^n} \varphi(x, \mu_1, ..., \mu_m, \frac{\partial\mu_1}{\partial x_1}, ..., \frac{\partial\mu_1}{\partial x_n}, ..., \frac{\partial\mu_m}{\partial x_1}, ..., \frac{\partial\mu_m}{\partial x_n})dx,$$

with φ an arbitrary function twicely differentiable, the functional derivative results in

$$\frac{\delta\Phi}{\delta\mu_j} = \frac{\partial\varphi}{\partial\mu_j} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \frac{\partial\varphi}{\partial\left(\frac{\partial\mu_j}{\partial x_i}\right)}. \tag{35}$$

Furthermore, the functional derivative satisfies the chain rule. If $G:\mathbb{R}\to\mathbb{R}$ is a real function, $G\in C^1(\mathbb{R})$, then

$$\frac{\delta G(\Phi)}{\delta\mu_j}=G'(\Phi)\frac{\delta\Phi}{\delta\mu_j}, \tag{36}$$

where $G'(\Phi)$ is the standard derivative of G.

Appendix B: Proof of Proposition 3

Proof. For any initial condition $f(t_0,x)$ such that $t_0\in[0,\infty)$, suppose that the admissible control $\mu^*\in A$ is a solution to the problem (15) for $t_0\le t<\infty$, then

$$V\left[f(t_0,\cdot)\right]=\int_{t_0}^{t}e^{-\rho(s-t_0)}W\left[f(s,\cdot),\mu^*\right]ds+e^{-\rho(t-t_0)}V\left[f(t,\cdot)\right], \tag{37}$$

This is a particular case of lemma 7.1 in Flemming and Soner (2006) replacing the value function by $V\left[f(t_0,\cdot)\right]$.

Taking derivatives with respect to time in equation (37):

$$\begin{aligned}\rho V\left[f(t,\cdot)\right] &= W\left[f(t,\cdot),\mu^*\right]+\frac{\partial}{\partial t}V\left[f(t,\cdot)\right]=W\left[f(t,\cdot),\mu^*\right]+\int_{\mathbb{R}^n}\frac{\delta V\left[f\right]}{\delta f(t,y)}\frac{\partial f(t,y)}{\partial t}dy \\ &= W\left[f(t,\cdot),\mu^*\right]+\int_{\mathbb{R}^n}\frac{\delta V\left[f\right]}{\delta f(t,y)}\left(-\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[b_i\left(y,\mu^*,\Gamma^*\right)f\right]+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}\left[\sigma_{ii}^2(y)f\right]\right)dy.\end{aligned} \tag{38}$$

Therefore, the optimal Markov control $\mu^*(t,\cdot)$ and the optimal prices $\Gamma_k^*(t)$ that maximize (38) subject to the constraints (13), should be extremals of

$$\begin{aligned}&W\left[f(t,\cdot),\mu\right]+\sum_{k=1}^{p}\lambda_k(t)\left\{H[f(t,\cdot),\mu]-G_k^{-1}\left(\Gamma_k(t)\right)\right\} \\ &+\int_{\mathbb{R}^n}v(t,y)\left(-\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[b_i\left(y,\mu,\Gamma\right)f\right]+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}\left[\sigma_{ii}^2(y)f\right]\right)dy,\end{aligned} \tag{39}$$

where $v(t,y)=\frac{\delta V[f]}{\delta f(t,y)}$.

If μ^* is an extremal of (39), all the functional derivatives of (39) with respect to μ_j, $j=1,...,m$, should be zero:

$$\frac{\delta W}{\delta\mu_j(t,x)}+\sum_{k=1}^{p}\lambda_k(t)\frac{\delta H}{\delta\mu_j(t,x)}+\frac{\delta}{\delta\mu_j(t,x)}\int_{\mathbb{R}^n}v(t,y)\left(-\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left(b_if\right)+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}\left(\sigma_{ii}^2f\right)\right)dy=0.$$

Notice that

$$\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[b_i\left(y,\mu^*,\Gamma^*\right)f\right]=\sum_{i=1}^{n}\left[\frac{\partial b_i}{\partial x_i}f+\sum_{j=1}^{m}\frac{\partial b_i}{\partial \mu_j}\frac{\partial \mu_j^*}{\partial x_i}f+b_i\frac{\partial f}{\partial x_i}\right],$$

therefore,

$$\begin{aligned}
&\frac{\delta}{\delta\mu_j(t,x)}\int_{\mathbb{R}^n}v(t,y)\left(-\sum_{i=1}^{n}\left[\frac{\partial b_i}{\partial x_i}f+\sum_{j=1}^{m}\frac{\partial b_i}{\partial \mu_j}\frac{\partial \mu_j^*}{\partial x_i}f+b_i\frac{\partial f}{\partial x_i}\right]+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}\left[\sigma_{ii}^2(y)f\right]\right)dy\\
=\;&-\sum_{i=1}^{n}\left[\frac{\partial^2 b_i}{\partial\mu_j\partial x_i}vf+\sum_{k=1}^{m}\frac{\partial^2 b_i}{\partial\mu_j\partial\mu_k}\frac{\partial\mu_k^*}{\partial x_i}vf+\frac{\partial b_i}{\partial\mu_j}\frac{\partial f}{\partial x_i}v\right]+\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left(vf\frac{\partial b_i}{\partial\mu_j}\right)\\
=\;&\sum_{i=1}^{n}\frac{\partial b_i}{\partial\mu_j}\frac{\partial v}{\partial x_i}f,
\end{aligned}$$

where in the second line we have computed the functional derivative using (35). We obtain the set of equations

$$\frac{\partial g(x,u(x,\mu^*))}{\partial\mu_j}f(t,x)+\sum_{k=1}^{p}\lambda_k(t)G_k'\frac{\partial h(x,\mu^*)}{\partial\mu_j}f(t,x)+\sum_{i=1}^{n}\frac{\partial b_i\left(x,\mu^*,\Gamma^*\right)}{\partial\mu_j}\frac{\partial v}{\partial x_i}f(t,x)=0,\quad j=1,...,m. \tag{40}$$

If $\Gamma^*(t)$ is a maximum of (39), all the partial derivatives of (39) with respect to Γ_k, $k=1,...,m$, should be zero:

$$\begin{aligned}
&-\frac{\lambda_k(t)}{G_k'}+\frac{\partial}{\partial\Gamma_k}\int_{\mathbb{R}^n}v(t,y)\left(-\sum_{i=1}^{n}\left[\frac{\partial b_i}{\partial x_i}f+\sum_{j=1}^{m}\frac{\partial b_i}{\partial \mu_j}\frac{\partial \mu_j^*}{\partial x_i}f+b_i\frac{\partial f}{\partial x_i}\right]\right)dy\\
=\;&-\frac{\lambda_k(t)}{G_k'}+\int_{\mathbb{R}^n}v(t,y)\left(-\sum_{i=1}^{n}\left[\frac{\partial^2 b_i}{\partial\Gamma_k\partial x_i}f+\sum_{j=1}^{m}\frac{\partial^2 b_i}{\partial\Gamma_k\partial\mu_j}\frac{\partial\mu_j^*}{\partial x_i}f+\frac{\partial b_i}{\partial\Gamma_k}\frac{\partial f}{\partial x_i}\right]\right)dy=0,
\end{aligned}$$

where we have appplied the inverse function theorem and $G_k'=G_k'\left(\int_{\mathbb{R}^n}h(x,\mu)f(t,x)dx\right)$. Therefore the value of the Lagrange multipliers is

$$\lambda_k(t)=-G_k'\int_{\mathbb{R}^n}v(t,y)\left(\sum_{i=1}^{n}\left[\frac{\partial^2 b_i\left(y,\mu^*,\Gamma^*\right)}{\partial\Gamma_k\partial x_i}f+\sum_{j=1}^{m}\frac{\partial^2 b_i\left(y,\mu^*,\Gamma^*\right)}{\partial\Gamma_k\partial\mu_j}\frac{\partial\mu_j}{\partial x_i}f+\frac{\partial b_i\left(y,\mu^*,\Gamma^*\right)}{\partial\Gamma_k}\frac{\partial f}{\partial x_i}\right]\right)dy.$$

We also compute the functional derivative with respect to f in equation (39) noticing that $\mu^*=\mu[f]$ and $\Gamma^*=\Gamma[f]$, that is, the optimal controls and the price vector depend on the state distribution:

$$\frac{\delta W\left[f,\mu^{*}\right]}{\delta f(t,x)}+\sum_{k=1}^{p}\lambda_{k}(s)\frac{\delta H\left[f,\mu^{*}\right]}{\delta f(t,x)}$$
$$+\frac{\delta}{\delta f(t,x)}\int_{\mathbb{R}^{n}}\frac{\delta V\left[f\right]}{\delta f(t,y)}\left(-\sum_{i=1}^{n}\frac{\partial}{\partial x_{i}}\left[b_{i}\left(y,\mu^{*},\Gamma^{*}\right)f\right]+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^{2}}{\partial x_{i}^{2}}\left[\sigma_{ii}^{2}(y)f\right]\right)dy=0,$$

where we have applied the envelope condition.

Notice that

$$\begin{aligned}\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^{2}}{\partial x_{i}^{2}}\left[\sigma_{ii}^{2}(y)f\right] &= \sum_{i=1}^{n}\frac{\partial}{\partial x_{i}}\left[\sigma_{ii}\frac{\partial\sigma_{ii}}{\partial x_{i}}f+\frac{\sigma_{ii}^{2}}{2}\frac{\partial f}{\partial x_{i}}\right]\\ &= \sum_{i=1}^{n}\left[\left(\frac{\partial\sigma_{ii}}{\partial x_{i}}\right)^{2}f+\sigma_{ii}\frac{\partial^{2}\sigma_{ii}}{\partial x_{i}^{2}}f+2\sigma_{ii}(x)\frac{\partial\sigma_{ii}}{\partial x_{i}}\frac{\partial f}{\partial x_{i}}+\frac{\sigma_{ii}^{2}}{2}\frac{\partial^{2}f}{\partial x_{i}^{2}}\right],\end{aligned}$$

then,

$$\begin{aligned}&\frac{\delta}{\delta f(t,x)}\int_{\mathbb{R}^{n}}\frac{\delta V\left[f\right]}{\delta f(t,y)}\left(-\sum_{i=1}^{n}\frac{\partial}{\partial x_{i}}\left[b_{i}\left(y,\mu^{*},\Gamma^{*}\right)f\right]+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^{2}}{\partial x_{i}^{2}}\left[\sigma_{ii}^{2}(y)f\right]\right)dy\\ =\;&\int_{\mathbb{R}^{n}}\frac{\delta^{2}V\left[f\right]}{\delta f(t,x)\delta f(t,y)}\frac{\partial f}{\partial t}dy+\sum_{i=1}^{n}\left[-\frac{\partial b_{i}}{\partial x_{i}}-\sum_{j=1}^{m}\frac{\partial b_{i}}{\partial\mu_{j}}\frac{\partial\mu_{j}^{*}}{\partial x_{i}}+\left(\frac{\partial\sigma_{ii}}{\partial x_{i}}\right)^{2}+\sigma_{ii}(x)\frac{\partial^{2}\sigma_{ii}}{\partial x_{i}^{2}}\right]v\\ &-\sum_{i=1}^{n}\frac{\partial}{\partial x_{i}}\left(-b_{i}v+2\sigma_{ii}\frac{\partial\sigma_{ii}}{\partial x_{i}}v\right)+\sum_{i=1}^{n}\frac{\partial^{2}}{\partial x_{i}^{2}}\left(2\sigma_{ii}\frac{\partial\sigma_{ii}}{\partial x_{i}}\frac{\sigma_{ii}^{2}}{2}v\right)\\ =\;&\int_{\mathbb{R}^{n}}\frac{\delta^{2}V\left[f\right]}{\delta f(t,x)\delta f(t,y)}\frac{\partial f}{\partial t}dy+\sum_{i=1}^{n}b_{i}\frac{\partial v}{\partial x_{i}}+\sum_{i=1}^{n}\frac{\sigma_{ii}^{2}}{2}\frac{\partial^{2}v}{\partial x_{i}^{2}}.\end{aligned}$$

where, from the first to the second and third lines we have computed the functional derivative using (35).[14]

Finally, taking into account that

$$\begin{aligned}\frac{\partial v(t,x)}{\partial t} &= \frac{\partial}{\partial t}\frac{\delta V\left[f\right]}{\delta f(t,x)}=\int_{\mathbb{R}^{n}}\frac{\delta^{2}V\left[f\right]}{\delta f(t,y)\delta f(t,x)}\frac{\partial f(t,y)}{\partial t}dy\\ &= \int_{\mathbb{R}^{n}}\frac{\delta^{2}V\left[f\right]}{\delta f(t,x)\delta f(t,y)}\frac{\partial f(t,y)}{\partial t}dy,\end{aligned}$$

[14] In this case, as there are also second order derivatives of the distribution, we obtain an extra term in (35) of the form $+\sum_{i=1}^{n}\frac{\partial^{2}}{\partial x_{i}^{2}}\frac{\partial\varphi}{\partial\left(\frac{\partial^{2}\mu_{j}}{\partial x_{i}^{2}}\right)}$.

we obtain the equation

$$g(x,\mu^*)+\sum_{k=1}^{p}\lambda_k(t)h(x,\mu^*)+\frac{\partial v(t,x)}{\partial t}+\sum_{i=1}^{n}b_i\left(x,\mu^*,\Gamma^*\right)\frac{\partial v(t,x)}{\partial x_i}+\sum_{i=1}^{n}\frac{\sigma_{ii}^2(x)}{2}\frac{\partial^2 v(t,x)}{\partial x_i^2}=\rho v(t,x). \tag{41}$$

Combining (41) and (40) we obtain (20).

Regarding the transversality condition, suppose that $t\uparrow\infty$. In this case, using (16) and taking functional derivatives with respect to f we obtain (22). ■

Appendix C: Description of the numerical algorithm

Step 1: Solution to the Hamilton-Jacobi-Bellman equation

The HJB equation is solved by a finite difference scheme following Achdou et al. (2014). It approximates the value function $V(a,z)$ on a finite grid with steps Δa and Δz : $a\in\{a_1,...,a_I\}$, $z\in\{z_1,...,z_J\}$.[15] We use the notation $V_{i,j}\equiv V(a_i,z_j)$, $i=1,...,I$; $j=1,...,J$. The derivative of V with respect to a can be approximated with either a forward or a backward approximation:

$$\frac{\partial V(a_i,z_j)}{\partial a} \approx \partial_{a,F}V_{i,j}\equiv\frac{V_{i+1,j}-V_{i,j}}{\Delta a}, \tag{42}$$

$$\frac{\partial V(a_i,z_j)}{\partial a} \approx \partial_{a,B}V_{i,j}\equiv\frac{V_{i,j}-V_{i-1,j}}{\Delta a}, \tag{43}$$

where the decision between one approximation or the other depends on the sign of the savings function $s_{i,j}=wz_j+ra_i-c_{i,j}$ through an "upwind scheme" described below. The derivatives of V with respect to z are approximated using a forward approximation

$$\frac{\partial V(a_i,z_j)}{\partial z} \approx \partial_z V_{i,j}\equiv\frac{V_{i,j+1}-V_{i,j}}{\Delta z}, \tag{44}$$

$$\frac{\partial^2 V(a_i,z_j)}{\partial z^2} \approx \partial_{zz}V_{i,j}\equiv\frac{V_{i,j+1}+V_{i,j-1}-2V_{i,j}}{(\Delta z)^2}. \tag{45}$$

The HJB equation (4)

$$\rho V=u(c)+(wz+ra-c)\frac{\partial V}{\partial a}+\eta(z)\frac{\partial V}{\partial z}+\frac{\sigma_z^2(z)}{2}\frac{\partial^2 V}{\partial z^2},$$

where

$$c=(u')^{-1}(\frac{\partial V}{\partial a}),$$

[15] Notice that subindexes i and j have a different meaning here than in the main text.

is approximated by an upwind scheme

$$\begin{aligned}\frac{V_{i,j}^{n+1}-V_{i,j}^{n}}{\Delta}+\rho V_{i,j}^{n+1} &= u(c_{i,j}^{n})+\partial_{a,F}V_{i,j}^{n+1}s_{i,j,F}^{n}\mathbf{1}_{s_{i,j,F}^{n}>0}+\partial_{a,B}V_{i,j}^{n+1}s_{i,j,B}^{n}\mathbf{1}_{s_{i,j,B}^{n}<0}\\ &\quad +\eta(z_j)\partial_z V_{i,j}^{n+1}+\frac{\sigma_z^2(z_j)}{2}\partial_{zz}V_{i,j}^{n+1},\end{aligned}$$

where

$$\begin{aligned}s_{i,j,F}^{n} &= wz_j+ra_i-(u')^{-1}\left(\partial_{a,F}V_{i,j}^{n}\right),\\ s_{i,j,B}^{n} &= wz_j+ra_i-(u')^{-1}\left(\partial_{a,B}V_{i,j}^{n}\right).\end{aligned}$$

Moving all variables with $n+1$ superscripts to the left hand side and those with n superscripts to the right hand side:

$$\frac{V_{i,j}^{n+1}-V_{i,j}^{n}}{\Delta}+\rho V_{i,j}^{n+1}=u(c_{i,j}^{n})+V_{i-1,j}^{n+1}\varrho_{i,j}+V_{i,j}^{n+1}\beta_{i,j}+V_{i+1,j}^{n+1}\gamma_{i,j}+V_{i,j-1}^{n+1}\chi_j+V_{i,j+1}^{n+1}\varsigma_j, \tag{46}$$

where

$$\begin{aligned}c_{i,j}^{n} &= (u')^{-1}\left(\partial_{a,F}V_{i,j}^{n}\mathbf{1}_{s_{i,j,F}^{n}>0}+\partial_{a,B}V_{i,j}^{n}\mathbf{1}_{s_{i,j,B}^{n}<0}+u'(wz_j+ra_i)\mathbf{1}_{s_{i,j,F}^{n}<0,s_{i,j,B}^{n}>0}\right), \qquad (47)\\ \varrho_{i,j} &= -\frac{s_{i,j,B}^{n}\mathbf{1}_{s_{i,j,B}^{n}<0}}{\Delta a},\\ \beta_{i,j} &= -\frac{s_{i,j,F}^{n}\mathbf{1}_{s_{i,j,F}^{n}>0}}{\Delta a}+\frac{s_{i,j,B}^{n}\mathbf{1}_{s_{i,j,B}^{n}<0}}{\Delta a}-\frac{\eta(z_j)}{\Delta z}-\frac{\sigma_z^2(z_j)}{(\Delta z)^2},\\ \gamma_{i,j} &= \frac{s_{i,j,F}^{n}\mathbf{1}_{s_{i,j,F}^{n}>0}}{\Delta a},\\ \chi_j &= \frac{\sigma_z^2(z_j)}{2(\Delta z)^2},\\ \varsigma_j &= \frac{\sigma_z^2(z_j)}{2(\Delta z)^2}+\frac{\eta(z_j)}{\Delta z}.\end{aligned}$$

The state constraint (3) $a\geq\bar{a}$ is enforced by setting $s_{i,j,B}^{n}=0$.[16] Similarly, $s_{I,j,F}^{n}=0$. Therefore, the values $V_{0,j}^{n+1}$ and $V_{I+1,j}^{n+1}$ are never used. At the boundaries in the j dimension, equation (46)

[16] This is equivalent to the penalty function approach discussed in the main text.

becomes

$$\begin{aligned}
\frac{V_{i,j}^{n+1}-V_{i,j}^{n}}{\Delta}+\rho V_{i,j}^{n+1} &= u(c_{i,1}^{n})+V_{i-1,j}^{n+1}\varrho_{i,1}+V_{i,1}^{n+1}\left(\beta_{i,1}+\chi_1\right)+V_{i+1,1}^{n+1}\gamma_{i,1}+V_{i,2}^{n+1}\varsigma_1,\\
\frac{V_{i,j}^{n+1}-V_{i,j}^{n}}{\Delta}+\rho V_{i,j}^{n+1} &= u(c_{i,J}^{n})+V_{i-1,J}^{n+1}\varrho_{i,J}+V_{i,J}^{n+1}\left(\beta_{i,J}+\varsigma_J\right)+V_{i+1,J}^{n+1}\gamma_{i,J}+V_{i,J-1}^{n+1}\chi_J.
\end{aligned}$$

Equation (46) is a system of $I\times J$ linear equations which can be written in matrix notation as:

$$\frac{\mathbf{V}^{n+1}-\mathbf{V}^{n}}{\Delta}+\rho\mathbf{V}^{n+1}=\mathbf{u}^{n}+\mathbf{A}^{n}\mathbf{V}^{n+1},$$

where the matrix $\mathbf{A}^n$ and the vectors $\mathbf{V}^{n+1}$ and $\mathbf{u}^n$ are defined by:

$$\mathbf{A}^n=\begin{bmatrix}
\beta_{1,1}+\chi_1 & \gamma_{1,1} & 0 & \cdots & 0 & \varsigma_1 & 0 & 0 & \cdots & 0\\
\varrho_{2,1} & \beta_{2,1}+\chi_1 & \gamma_{2,1} & 0 & \cdots & 0 & \varsigma_1 & 0 & \cdots & 0\\
0 & \varrho_{3,1} & \beta_{3,1}+\chi_1 & \gamma_{3,1} & 0 & \cdots & 0 & \varsigma_1 & \cdots & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & 0 & \cdots & 0\\
0 & 0 & 0 & \varrho_{I,1} & \beta_{I,1}+\chi_1 & \gamma_{I,1} & 0 & 0 & \cdots & 0\\
0 & 0 & \cdots & 0 & \varrho_{1,2} & \beta_{1,2} & \gamma_{1,2} & 0 & \cdots & 0\\
0 & 0 & \cdots & 0 & 0 & \varrho_{2,2} & \beta_{2,2} & \gamma_{2,2} & \cdots & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0\\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & \varrho_{I-1,J} & \beta_{I-1,J}+\varsigma_J & \gamma_{I-1,J}\\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \varrho_{I,J} & \beta_{I,I}+\varsigma_J
\end{bmatrix},$$

$$\mathbf{V}^{n+1}=\begin{bmatrix} V_{1,1}^{n+1}\\ V_{2,1}^{n+1}\\ \vdots\\ V_{1,2}^{n+1}\\ V_{2,2}^{n+1}\\ \vdots\\ V_{I-1,J}^{n+1}\\ V_{I,J}^{n+1}\end{bmatrix},\qquad \mathbf{u}^{n}=\begin{bmatrix} u(c_{1,1}^{n})\\ u(c_{2,1}^{n})\\ \vdots\\ u(c_{1,2}^{n})\\ u(c_{2,2}^{n})\\ \vdots\\ u(c_{I-1,J}^{n})\\ u(c_{I,J}^{n})\end{bmatrix}.$$

The system can in turn be written as

$$\mathbf{B}^n\mathbf{V}^{n+1}=\mathbf{d}^n, \tag{48}$$

where $\mathbf{B}^n$= $\left(\frac{1}{\Delta}+\rho\right)\mathbf{I}-\mathbf{A}^n$ and $\mathbf{d}^n=\mathbf{u}^n+\frac{\mathbf{V}^n}{\Delta}$. $\mathbf{I}$ is the identity matrix.

The algorithm to solve the HJB equation runs as follows. Begin with an initial guess $V_{i,j}^{0} = u(ra_i + wz_j)/\rho$, set $n = 0$. Then:

1. Compute $\partial_{a,F}V_{i,j}^{n}$, $\partial_{a,B}V_{i,j}^{n}$, $\partial_{z}V_{i,j}^{n}$ and $\partial_{zz}V_{i,j}^{n}$ using (42)-(45).

2. Compute $c_{i,j}^{n}$ using (47).

3. Find $V_{i,j}^{n+1}$ solving the linear sustem of equations (48).

4. If $V_{i,j}^{n+1}$ is close enought to $V_{i,j}^{n}$, stop. If not set $n := n + 1$ and go to step 1.

Step 2: Solution to the Kolmogorov Forward equation

The KF equation is also solved using an upwind finite difference scheme. The equation (7) in this case is

$$\begin{aligned} 0 &= -\frac{\partial}{\partial a}\left[\left(wz + ra - c\right)f\right] - \frac{\partial}{\partial z}\left[\eta(z)f\right] + \frac{1}{2}\frac{\partial^2}{\partial z^2}\left[\sigma_z^2(z)f\right], && (49)\\ \int f(a,z)dadz &= 1. && (50) \end{aligned}$$

This case is simpler than the previous one, as the problem is linear in f, so no iterative procedure is needed.

We use the notation $f_{i,j} \equiv f(a_i, z_j)$. The system can be now expressed as

$$\begin{aligned} 0 = & -\frac{f_{i,j}s_{i,j,F}^{n} - f_{i-1,j}s_{i-1,j,F}^{n}}{\Delta a}\mathbf{1}_{s_{i,j,F}^{n}>0} - \frac{f_{i+1,j}s_{i+,j,B}^{n} - f_{i,j}s_{i,j,B}^{n}}{\Delta a}\mathbf{1}_{s_{i,j,B}^{n}<0} \\ & -\frac{f_{i,j}\eta(z_j) - f_{i,j-1}\eta(z_{j-1})}{\Delta z} + \frac{f_{i,j+1}\sigma_z^2(z_{j+1}) + f_{i,j-1}\sigma_z^2(z_{j-1}) - 2f_{i,j}\sigma_z^2(z_j)}{2\left(\Delta z\right)^2}, \end{aligned}$$

or equivalently

$$f_{i-1,j}\gamma_{i,j} + f_{i+1,j}\varrho_{i,j} + f_{i,j}\beta_{i,j} + f_{i,j+1}\chi_j + f_{i,j-1}\varsigma_j = 0, \quad (51)$$

then (51) is also a system of $I \times J$ linear equations which can be written in matrix notation as:

$$\mathbf{A^T f} = \mathbf{0}, \quad (52)$$

where $\mathbf{A^T}$ is the transpose of $\mathbf{A} = \lim_{n\to\infty}\mathbf{A}^n$. In order to impose the normalization constraint (50) we fix one value of $f_{i,j}$ equal to 0.1 by replacing the corresponding entry of the zero vector in (52) by 0.1 and the corresponding row of $\mathbf{B}$ by a row of zeros everywhere except for a one in the diagonal. We solve the system (52) and obtain a solution $\hat{\mathbf{f}}$. Then we renormalize as

$$f_{i,j} = \frac{\hat{f}_{i,j}}{\sum_{i=1}^{I} \sum_{j=1}^{J} \hat{f}_{i,j} \Delta a \Delta z}.$$