Equilibrium Theory in Continuous Time

Tomas Björk
Stockholm School of Economics

Stockholm 2012

Contents

Equilibrium Theory in Continuous Time

Lecture 1

The connection between DynP and the Martingale Method

Tomas Björk

Main obejctives

In this lecture we will study a simple optimal investment problem using two standard approaches: Dynamic Programming (DynP) and the Martingale Method. The goal is to understand the deep connections that exist between these approaches. The results will be important when we move to equilibrium models later on in the course.

1.1

Model setup

A simple investment model

We consider a standard Black-Scholes model of the form

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt \end{aligned}$$

and the problem is that of maximizing expected utility of the form

$$E^P \left[\int_0^T U(t, c_t) dt + \Phi(X_T) \right]$$

with the usual portfolio dynamics

$$dX_t = X_t u_t (\alpha - r) dt + (r X_t - c_t) dt + X_t u_t \sigma dW_t$$

where we have used the notation

$$\begin{aligned} X_t &= \text{portfolio value,} \\ c_t &= \text{consumption rate,} \\ u_t &= \text{weight on the risky asset.} \end{aligned}$$

1.2

Dynamic programming

The HJB equation

The HJB equation for the optimal value function $V(t, x)$ is given by

$$\begin{aligned} V_t + \sup_{(c,u)} \left\{ U(t,c) + xu(\alpha - r)V_x + (rx - c)V_x + \frac{1}{2}x^2u^2\sigma^2V_{xx} \right\} &= 0, \\ V(T,x) &= \Phi(x) \\ V(t,0) &= 0. \end{aligned}$$

From the first order condition we obtain

$$\begin{aligned} U_c(t,\hat{c}) &= V_x(t,x), \\ \hat{u}(t,x) &= -\frac{(\alpha - r)}{\sigma^2} \cdot \frac{V_x(t,x)}{xV_{xx}(t,x)}. \end{aligned}$$

Plugging the expression for $\hat{u}$ into the HJB equation gives us the PDE

$$V_t + U(t, \hat{c}) + (rx - \hat{c})V_x - \frac{1}{2}\frac{(\alpha - r)^2}{\sigma^2} \cdot \frac{V_x^2}{V_{xx}} = 0,$$

with the same boundary conditions as above.

Problems with HJB:

- The HJB equation is highly nonlinear in V_x and V_{xx}.
- The optimal consumption $\widehat{c}$ is nonlinear in V_x.
- It is thus a hard task to solve the HJB.

1.3

The martingale approach

1.3.1

Basic arguments and results

The Martingale Method

Using standard arguments, the original problem is equivalent to that of maximizing the expected utility

$$E^P\left[\int_0^T U(t, c_t)dt + \Phi(X_T)\right]$$

over all consumption processes c and terminal wealth profiles X_T, under the budget constraint

$$E^P\left[\int_0^T e^{-rt}L_t c_t dt + e^{-rT}L_T X_T\right] = x_0,$$

where $L = dQ/dP$ has dynamics

$$\begin{cases} dL_t &= L_t \varphi_t dW_t, \\ L_0 &= 1 \end{cases}$$

and where the Girsanov kernel φ is given by

$$\varphi_t = \frac{r - \alpha}{\sigma}.$$

The Lagrangian for this problem is

$$E^P \left[\int_0^T \left\{ U(t, c_t) - \lambda e^{-rt} L_t c_t \right\} dt + \Phi(X_T) - e^{-rT} \lambda L_T X_T \right] + \lambda x_0$$

where λ is the Lagrange multiplier and x_0 the initial wealth. The first order conditions are

$$\begin{aligned} U_c(t, \hat{c}) &= \lambda M_t, \\ \Phi'(X_T) &= \lambda M_T. \end{aligned}$$

where M denotes the stochastic discount factor (SDF), defined by

$$M_t = B_t^{-1} L_t.$$

Recall

$$
\begin{aligned}
U_c(t, \hat{c}) &= \lambda M_t, \\
\Phi'(X_T) &= \lambda M_T.
\end{aligned}
$$

Introduce the following inverse (in x and c) functions

$$
\begin{aligned}
G(t, c) &= U_c^{-1}(t, c), \\
F(x) &= \left[\Phi'\right]^{-1}(x).
\end{aligned}
$$

We can then can write the optimality conditions on the form

$$
\begin{aligned}
\hat{c}_t &= G(t, \lambda M_t), \\
\hat{X}_T &= F(\lambda M_T).
\end{aligned}
$$

Now recall from DynP

$$
U_c(t, \hat{c}) = V_x(t, x).
$$

We have thus proved the following result.

Theorem

With notation as above we have

$$V_x(t, \hat{X}_t) = \lambda M_t,$$

In other words: Along the optimal trajectory, the indirect marginal utility of wealth is (up to a scaling factor) given by the stochastic discount factor process. Furthermore, the Lagrange multiplier λ is given by

$$\lambda = V_x(0, x_0).$$

Corollary: Let V be the solution of the HJB equation. We then have

$$E^P \left[\int_0^T V_x(t, \hat{X}_t) \cdot \hat{c}_t dt + V_x(T, \hat{X}_t) \cdot \hat{X}_T \right] = V_x(0, x_0) x_0.$$

1.3.2

The PDE of the martingale method

Some problems with the martingale method

The martingale approach is very nice, but there are, seemingly, some shortcomings.

- We have no explicit expression for the optimal portfolio weight $\hat{u}_t$.

- The formula $\hat{c}_t = G(t, \lambda M_t)$, for the optimal consumption is very nice, but it is expressed in the "dual" state variable $Z = \lambda M$, rather than as a feedback control in the "primal" state variable x.

- We would also like to have an explicit expression for the optimal wealth process $\hat{X}_t$.

Some comments

- We first note that the multiplier λ is determined by the budget constraint

$$E^Q\left[\int_0^T e^{-rt}G(t,\lambda M_t)dt + e^{-rT}F(\lambda M_T)\right] = x_0.$$

so we assume that we have computed λ.

- Define the process Z by

$$Z_t = \lambda M_t.$$

- We can then write

$$\begin{aligned}\hat{c}_t &= G(t, Z_t),\\ \hat{X}_T &= F(Z_T).\end{aligned}$$

General Strategy

1. From risk neutral valuation is easy to see that X_t is of the form
$$X_t = H(t, Z_t)$$
where H satisfies a Kolmogorov backward equation.

2. Using Ito on H we can compute dX.

3. We also know that the X dynamics are of the form
$$dX_t = (\ldots)\, dt + u_t X_t \sigma dW_t.$$

4. Comparing these two expressions for dX we can identify the optimal weight u from the diffusion part of dX.

5. We invert the formula $x = H(t, z)$ to obtain $z = K(t, x)$. This gives us u and c as functions of the primal state variable x.

6. Finally, we investigate what the Kolmogorov equation above looks like in the new variable x.

Computing X_t in terms of Z_t

Recall that

$$\begin{aligned}\hat{c}_t &= G(t, Z_t),\\ \hat{X}_T &= F(Z_T).\end{aligned}$$

From standard risk neutral valuation we thus have

$$X_t = E^Q\left[\int_t^T e^{-r(s-t)}G(s, Z_s)ds + e^{-r(T-t)}F(Z_T)\right|\left.\mathcal{F}_t\right].$$

Thus X_t can be expressed as

$$X_t = H(t, Z_t)$$

where H satisfies a Kolmogorov equation.

To find this equation we need the Q dynamics of Z.

The Q-dynamics of Z

Since $Z_t = B_t^{-1} L_t$ and the L dynamics are

$$dL_t = L_t \varphi_t dW_t,$$

with

$$\varphi = (r - \alpha)/\sigma$$

we see that the P dynamics of Z are

$$dZ_t = -rZ_t dt + Z_t \varphi dW_t$$

Thus, from Girsanov, the Q-dynamics of Z are

$$dZ_t = Z_t \left(\varphi^2 - r\right) dt + Z_t \varphi dW_t^Q.$$

where

$$dW_t = \varphi dt + dW_t^Q.$$

The PDE for $H(t, z)$

We recall that

$$X_t = H(t, Z_t) = E^Q \left[\int_t^T e^{-r(s-t)} G(s, Z_s) ds + e^{-r(T-t)} F(Z_T) \middle| \mathcal{F}_t \right].$$

and

$$dZ_t = Z_t \left(\varphi^2 - r \right) dt + Z_t \varphi dW_t^Q.$$

We thus obtain the Kolmogorov backward equation for H as

$$\begin{cases} H_t + z(\varphi^2 - r)H_z + \frac{1}{2}\varphi^2 z^2 H_{zz} + c(t, z) - rH & = & 0, \\ H(T, z) & = & F(z). \end{cases}$$

Determining $\hat{u}(t, z)$

Since

$$X_t = H(t, Z_t),$$

we can apply Ito to obtain

$$dX_t = (\ldots)\, dt + H_z(t, Z_t) Z_t \varphi dW_t.$$

Comparing this to

$$dX_t = (\ldots)\, dt + u_t X_t \sigma dW_t,$$

gives us the optimal weight on the risky asset as

$$u(t, z) = \frac{\varphi}{\sigma} \cdot \frac{z H_z(t, z)}{H(t, z)}.$$

We have thus proved...

Theorem

We have the following formulas for the optimal wealth, consumption, and portfolio weight.

$$\begin{aligned}
\widehat{X}_t &= H(t, Z_t), \\
\widehat{c}(t, z) &= G(t, z), \\
\widehat{u}(t, z) &= \frac{\varphi}{\sigma} \cdot \frac{z H_z(t, z)}{H(t, z)}.
\end{aligned}$$

Here

$$G = U_c^{-1}$$

and H is defined by

$$\begin{cases}
H_t + z(\varphi^2 - r)H_z + \frac{1}{2}\varphi^2 z^2 H_{zz} + G - rH = 0, \\
H(T, z) = F(z).
\end{cases}$$

1.4

The connection between HJB and Kolmogorov

HJB versus Kolmogorov

HJB:

$$V_t + U(t, \hat{c}) + (rx - \hat{c})V_x - \frac{1}{2}\frac{(\alpha - r)^2}{\sigma^2} \cdot \frac{V_x^2}{V_{xx}} = 0,$$

Kolmogorov:

$$H_t + z(\varphi^2 - r)H_z + \frac{1}{2}\varphi^2 z^2 H_{zz} + G - rH = 0$$

The Kolmogorov equation is **linear** in H, whereas the HJB equation is **non-linear** in H. The Kolmogorov eqn is thus **much** nicer that the HJB eqn.

There must be some connection between these equations. Which?

Drawbacks with Kolmogorov

- We have seen that The Kolmogorov eqn is much nicer that the HJB eqn.

- Thus the martingale approach seems to be preferable to DynP.

- Note, however, that with the MG approach the controls are determined as functions of the dual variable z.

- We would prefer to have the controls as feedback of the primal state variable x.

- This can in fact be achieved by a change of variables using the relation $x = H(t, z)$.

Changing variables

We have

$$x = H(t, z).$$

Assuming that H is invertible in the z-variable, we can write

$$z = K(t, x).$$

We can then substitute this into our formulas

$$\begin{aligned} \widehat{c}(t, z) &= G(t, z), \\ \widehat{u}(t, z) &= \frac{\varphi}{\sigma} \cdot \frac{z H_z(t, z)}{H(t, z)}. \end{aligned}$$

to obtain

$$\begin{aligned} \widehat{c}(t, x) &= G(t, K(t, x)), \\ \widehat{u}(t, x) &= \frac{\varphi}{\sigma} \cdot \frac{K(t, x) H_z(t, K(t, x))}{H(t, K(t, x))}. \end{aligned}$$

We now need a PDE for $K(t, x)$.

The PDE for $K(t, x)$.

By definition we have

$$H(t, K(t, x)) = x,$$

for all x. Differentiating this identity once in the t variable and twice in the x variable gives us,

$$H_t = -\frac{K_t}{K_x}, \quad H_z = \frac{1}{K_x}, \quad H_{zz} = -\frac{K_{xx}}{K_x^3}.$$

Substituting this into the Kolmogorov eqn for H gives us

$$\begin{cases} K_t + (rx - c)K_x + \frac{1}{2}\varphi^2 K^2 \frac{K_{xx}}{K_x^2} + (r - \varphi^2)K & = & 0, \\ K(T, x) & = & \Phi'(x), \end{cases}$$

which is a non-linear PDE for K.

What is going on?

To understand the nature of the PDE for K we recall that

$$V_x(t, \widehat{X}_t) = Z_t,$$

and since we also have

$$Z_t = K(t, \hat{X}_t)$$

this implies that we must have the interpretation

$$K(t, x) = V_x(t, x).$$

This can also be verified by differentiating the HJB eqn

$$V_t + U(t, \hat{c}) + (rx - \hat{c})V_x - \frac{1}{2}\frac{(\alpha - r)^2}{\sigma^2} \cdot \frac{V_x^2}{V_{xx}} = 0,$$

w.r.t x while using the optimality condition $U_c = V_x$.

Collecting results

- The process $Z_t = \lambda M_t$ has the representation

$$Z_t = V_x(t, \widehat{X}_t).$$

- The optimal wealth process is given by

$$\widehat{X}_t = H(t, Z_t),$$

  where the function H is defined by the Kolmogorov equation.

- The formulas for the optimal portfolio and consumption for the MG approach are mapped into the HJB formulas by the change of variable

$$\begin{aligned} x &= H(t, z), \\ z &= K(t, x), \end{aligned}$$

  where K is the functional inverse of H in the z variable.

- We have the identification

$$K(t,x) = V_x(t,x).$$

- After the variable change $z = K(t,x)$, the Kolmogorov equation for H transforms into the PDE for K.

- Since $K = V_x$ the PDE for K is identical to the PDE for V_x one obtains by differentiating the HJB equation w.r.t. the x variable.

Concluding remarks

- Using DynP we end up with the highly non linear HJB equation, which can be very difficult to solve.
- On the positive side for DynP, the controls are expressed directly in the natural state variable x.
- For the MG approach, the relevant PDE is much easier than the corresponding HJB equation for DynP. This is a big advantage.
- On the negative side for the MG approach, the optimal controls are expressed in the dual variable z instead of the wealth variable x, and in order to express the controls in the x variable, we need to invert the function H above.

Equilibrium Theory in Continuous Time

Lecture 2

A simple production equilibrium model

Tomas Björk

Where are we going?

- In the previous lecture the short rate r process was **exogenously** given.

- We now move to an **equilibrium model** where the the short rate process r_t will be determined **endogenously** within the model.

- In later lectures we will also discuss how other asset price processes (apart from r) are determined by equilibrium.

How do we do this?

Basic model structure

The simplest model has the following structure.

- We assume the existence of one or several economic **agents** with given utility functions for consumption.
- The agents are exogenously given a **production technology**.
- The agents make decisions about
  - Investment in the production technology.
  - Consumption
  - Investment in a risk free asset B.
- The agents act so as to maximize expected utlity.
- The short rate process r is then determined by the equilibrium condition that supply equals demand on the market for B.

2.1

Model, agents, and equilibrium

A simple production model

We consider an economy with one consumption good, referred to as "apples" or "dollars". All prices are in terms of this consumption good.

We now a give a formal assumption which is typical for this theory.

Assumption: *We assume that there exists a constant returns to scale physical production technology process* S *with dynamics*

$$dS_t = \alpha S_t dt + S_t \sigma dW_t.$$

The economic agents can invest unlimited positive amounts in this technology, but since it is a matter of physical investment, short positions are not allowed.

What exactly does this mean?

Interpretation

- At any time t you are allowed to invest dollars in the production process.
- If you, at time t_0, invest q dollars, and wait until time t_1 then you will receive the amount of

$$q \cdot \frac{S_{t_1}}{S_{t_0}}$$

  in dollars. In particular we see that the return on the investment is linear in q, hence the term "constant returns to scale".

- Since this is a matter of physical investment, shortselling is not allowed.

A moment of reflection shows that, from a purely formal point of view, investment in the technology S is in fact **equivalent to the possibility of investing in a risky asset** with price process S, but again with the constraint that shortselling is not allowed.

The risk free asset

Assumption: *We assume that there exists a* **risk free asset** *in zero net supply with dynamics*

$$dB_t = r_t B_t dt,$$

where r is the short rate process, which will be determined endogenously. The risk free rate r is assumed to be of the form

$$r_t = r(t, X_t)$$

where X denotes portfolio value.

Comment: The term **zero net supply** means that if someone buys a unit of B then someone else has to sell it. The aggregate demand, and supply, of B is thus equal to zero.

The wealth dynamics

Interpreting the production technology S as above, the wealth dynamics will be given, by the standard expression

$$dX_t = X_t u_t(\alpha - r)dt + (r_t X_t - c_t)dt + X_t u_t \sigma dW_t.$$

We note again that we have a shortselling – or rather short-investing – constraint on S.

Finally we need an economic agent.

The agent

Assumption: *We assume that there exists a* **representative agent** *who wishes to maximize the usual expected utility*

$$E^P\left[\int_0^T U(t, c_t)dt + \Phi(X_T)\right].$$

Comment: One would obviously like to have more than one agent, but we note the following.

- Assuming a representative agent facilitates the computations enormously.
- We will show later, that the general case with a finite number of different agents can be reduced to the case of a representative agent.
- We may thus WLOG assume the existence of a representative agent.

The control problem for the agent

Given the functional form $r(t,x)$, the agent wants to maximize

$$E^P\left[\int_0^T U(t,c_t)dt + \Phi(X_T)\right].$$

over c and u, subject to the X-dynamics

$$dX_t = X_t u_t(\alpha - r)dt + (r_t X_t - c_t)dt + X_t u_t \sigma dW_t.$$

and the constraints

$$u_t \geq 0, \quad c_t \geq 0.$$

Note: All results of the previous lecture are still valid if we replace expressions like $e^{-r(T-t)}$ by

$$e^{-\int_t^T r_s ds}$$

where r_t is shorthand for $r(t,X_t)$.

Equilibrium definition

An equilibrium of the model is a triple $\{\hat{c}(t,x), \hat{u}(t,x), r(t,x)\}$ of real valued functions such that the following hold.

1. Given the risk free short rate process $r(t, X_t)$, the optimal consumption and investment are given by $\hat{c}$ and $\hat{u}$ respectively.

2. The market for the risk free asset clears, i.e there is zero demand for B, so $1 - \hat{u}(t,x) = 0$

3. The market clears for the risk free asset, i.e.

$$\hat{u}(t,x) \equiv 1.$$

   (This is of course a consequence of market clearing for B).

In equilibrium, everything which is not consumed is thus invested in the production technology.

2.2

Dynamic programming

2.2.1

The HJB equation and market equilibrium

The HJB equation

We recall the agent's control problem as maximizing

$$E^P\left[\int_0^T U(t,c_t)dt + \Phi(X_T)\right].$$

over c and u, subject to the X-dynamics

$$dX_t = X_t u_t(\alpha - r)dt + (r_t X_t - c_t)dt + X_t u_t \sigma dW_t.$$

The HJB equation is thus given by

$$V_t + \sup_{(c,u)}\left\{U(t,c) + xu(\alpha - r)V_x + (rx - c)V_x + \frac{1}{2}x^2u^2\sigma^2 V_{xx}\right\} = 0,$$

Optimal consumtion and portfolio weight

The HJB equation was

$$V_t + \sup_{(c,u)} \left\{ U(t,c) + xu(\alpha - r)V_x + (rx - c)V_x + \frac{1}{2}x^2u^2\sigma^2V_{xx} \right\} = 0,$$

The optimal consumption and portfolio weight are given by

$$\begin{aligned} U_c(t,\hat{c}) &= V_x(t,x), \\ \hat{u}(t,x) &= -\frac{(\alpha - r)}{\sigma^2} \cdot \frac{V_x(t,x)}{xV_{xx}(t,x)}. \end{aligned}$$

Using the equilibrium condition $\widehat{u}_t \equiv 1$ we obtain the main result.

Equilibrium Theorem

- The equilibrium short rate is given by $r(t, \widehat{X}_t)$ where

$$r(t,x) = \alpha + \sigma^2 \, \frac{xV_{xx}(t,x)}{V_x(t,x)}.$$

- The dynamics of the equilibrium wealth process are

$$d\widehat{X}_t = \left(\alpha \widehat{X}_t - \hat{c}_t \right) dt + \widehat{X}_t \sigma dW_t.$$

- The Girsanov kernel has the form $\varphi(t, \widehat{X}_t)$ where

$$\varphi(t,x) = \frac{r(t,x) - \alpha}{\sigma}, \qquad (1)$$

or, alternatively,

$$\varphi(t,x) = \sigma \frac{xV_{xx}(t,x)}{V_x(t,x)}. \qquad (2)$$

- The optimal value function V is determined by the HJB equation

$$\left\{ \begin{array}{rcl} V_t + U(t, \hat{c}) + (\alpha x - \hat{c})V_x + \frac{1}{2}\sigma^2 x^2 V_{xx} & = & 0, \\ V(T, x) & = & \Phi(x). \end{array} \right.$$

Note: We see that although the (non-equilibrium) HJB equation

$$V_t + U(t, \hat{c}) + (rx - \hat{c})V_x - \frac{1}{2}\frac{(\alpha - r)^2}{\sigma^2} \cdot \frac{V_x^2}{V_{xx}} = 0,$$

is non-linear in V, the **equilibrium** HJB is (apart from the $\hat{c}$ terms) in fact **linear** in V.

2.2.2

A central planner

Introducing a central planner

- So far we have assumed that the economic setting is that of a **representative agent** investing in and consuming in a market.

- As an alternative to this setup, we now consider a **central planner** who does have access to the production technology, but who does **not** have access to the financial market, i.e. to B.

- The optimization problem for the central planner is simply that of maximizing expected utility when everything that is not consumed is invested in the production process.

- This looks very much like the problem of a representative agent who, in equilibrium, does not invest anything in the risk free asset.

- A natural conjecture is then that the equilibrium consumption of the representative agent coincides with the optimal consumption of the central planner.

The control problem

The formal problem of the central planner is to maximize

$$E^P\left[\int_0^T U(t,c_t)dt + \Phi(X_T)\right].$$

over the control c, given the wealth dynamics

$$dX_t = (\alpha X_t - c_t)dt + \sigma X_t dW_t.$$

The HJB equation for this problem is

$$\begin{cases} V_t + \sup_c \left\{ U(t,c) + (\alpha x - c)V_x + \frac{1}{2}\sigma^2 x^2 V_{xx} \right\} &= 0, \\ V(T,x) &= \Phi(x). \end{cases}$$

with the usual first order condition

$$U_c(t,c) = V_x(t,x).$$

Substituting the optimal c we thus obtain the PDE

$$\begin{cases} V_t + U(t,\hat{c}) + (\alpha x - \hat{c})V_x + \frac{1}{2}\sigma^2 x^2 V_{xx} & = & 0, \\ V(T,x) & = & \Phi(x). \end{cases}$$

and we see that this is identical to the HJB eqn for the representative agent. We have thus proved the following result.

Theorem: *Given assumptions as above, the following hold.*

- *The optimal consumption for the central planner coincides with the equilibrium consumption of the representative agent.*
- *The optimal wealth process for the central planner is identical with the equilibrium wealth process for the representative agent.*

Conclusion

- Solve the (fairly simple) problem for the central planner and, in particular, compute V.

- Define the "shadow interest rate" r by

$$r(t,x) = \alpha + \sigma^2 \, \frac{xV_{xx}(t,x)}{V_x(t,x)}.$$

- Now forget about the central planner and consider the optimal consumption/investment problem of a representative agent with access to the production technology S and a risk free asset B with dynamics

$$dB_t = r(t, X_t)B_t dt$$

  where r is defined as above.

- The economy will then be in equilibrium, so $\hat{u} = 1$, and we will recover the optimal consumption and wealth processes of the central planner.

2.3

The martingale approach

2.3.1

Model specification and equilibrium

Model specification

The model is almost exactly as before. The only difference is that, in order to have a Markovian model, we assume that the short rate process is of the form

$$r_t = r(t, Z_t).$$

Note the difference with the earlier assumption

$$r_t = r(t, X_t)$$

Optimality results

Using the results from Lecture 1.3 we have

$$
\begin{aligned}
\widehat{X}_t &= H(t, Z_t), \quad U_c(t, \hat{c}) = Z_t, \\
\widehat{c}(t, z) &= G(t, z), \quad \widehat{u}(t, z) = \frac{\varphi}{\sigma} \cdot \frac{z H_z(t, z)}{H(t, z)}.
\end{aligned}
$$

where G is the inverse of U_c, and H solves the PDE

$$
\begin{cases}
H_t + z(\varphi^2 - r)H_z + \frac{1}{2}\varphi^2 z^2 H_{zz} + G(t, z) - rH = 0, \\
H(T, z) = F(z),
\end{cases}
$$

where as usual

$$
\varphi = \frac{r - \alpha}{\sigma}
$$

Equilibrium

The equilibrium condition $\hat{u} = 1$ gives us the Girsanov kernel φ and the equilibrium rate r as

$$\varphi(t,z) = \sigma \frac{H(t,z)}{zH_z(t,z)}, \tag{3}$$

$$r(t,z) = \alpha + \sigma^2 \frac{H(t,z)}{zH_z(t,z)}. \tag{4}$$

In order to compute φ and r we must solve the PDE for H. On the surface, this PDE looks reasonable nice, but we must of course substitute the expressions for φ and r into the PDE for H. We then have the following result.

Theorem

The equilibrium interest rate is given by

$$r(t,z) = \alpha + \sigma^2 \frac{H(t,z)}{zH_z(t,z)}$$

where H solves the PDE

$$\begin{cases} H_t - \alpha z H_z + \frac{1}{2}\sigma^2 \frac{H^2}{H_z^2} H_{zz} + G - (\alpha + \sigma^2)H & = & 0, \\ H(T,z) & = & F(z). \end{cases}$$

Remark: We note that the equilibrium condition introduces a nonlinearity into the PDE for the MG approach.

Change of variable

We may again argue as in Lecture 1.4, and perform a change of variable by

$$x = H(t, z) \quad z = K(t, x).$$

Exactly as in Lecture 1.4, the PDE for H will then be transformed into the following PDE for K.

$$K_t + (\alpha + \sigma^2)xK_x - \hat{c}K_x + \frac{1}{2}\sigma^2 x^2 K_{xx} = 0.$$

As before we also have the indentification

$$K(t, x) = V_x(t, x),$$

and the PDE for K can also be derived by differentiating the equilibrium HJB equation in the x variable.

A remark on the shortselling constraint

We recall that since our process S has the interpretation of physical investment, then we have a shortselling constraint, the market becomes incomplete, and we are not formally allowed to use the MG approach. There seems to exist at lest two ways to handle this problem.

- We accept the reality of the shortselling constraint and interpret the results above as the equilibrium results in an extended model where shortselling **is** formally allowed. Since there is in fact no shortselling in equilibrium we then conclude that the extended equilibrium is indeed also an equilibrium for the original model. This, however, leaves open the question whether there can exist an equilibrium in the original model, which is not an equilibrium in the extended model.

- We gloss over the problem, abstain from even mentioning it, and hope that it will disappear. This seems to be a rather common strategy in the literature.

2.3.2

A central planner

Introducing a central planner

In the DynP approach we introduced, with considerable success, a central planner who maximized expected utility of wealth and consumption

$$E^P\left[\int_0^T U(t,c_t)dt + \Phi(X_T)\right].$$

given the wealth dynamics

$$dX_t = (\alpha X_t - c_t)dt + \sigma X_t dW_t.$$

The important assumption here is that the central planner does not have access to the risk free asset B. This implies that the market is incomplete so, as far as I understand, this implies that we cannot use the usual MG approach.

Concluding remarks

- In Lecture 1 we found that the Komogorov PDE in the MG approach had a much simpler structure than the HJB equation for the DynP approach.

- It seems, however, that this advantage of the MG approach over the DynP approach vanishes completely when we move from the pure optimization model to the equilibrium model.

- The equilibrium PDE for H

$$H_t - \alpha z H_z + \frac{1}{2}\sigma^2 \frac{H^2}{H_z^2} H_{zz} + G - (\alpha + \sigma^2) H = 0$$

does not look easier than the equilibrium HJB eqn for V

$$V_t + U(t, \hat{c}) + (\alpha x - \hat{c}) V_x + \frac{1}{2}\sigma^2 x^2 V_{xx} = 0$$

Equilibrium Theory in Continuous Time

Lecture 3

The CIR production factor model

Tomas Björk

Where are we going?

In this lecture we will study the famous Cox-Ingersoll-Ross factor model for a production equilibrium. The model is an extension of the model studied in the previous lecture, so the general strategy remains exactly the same.

3.1

The model

In the model some objects are assumed to be given exogenously whereas other objects are determined by equilibrium, and we also have economic agents.

Exogenous objects

Assumption: *The following objects are considered as exogenously given.*

1. *A 2-dimensional Wiener process W.*

2. *A scalar factor process Y of the form*

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dW_t$$

   where μ is a scalar real valued function and σ is a 2-dimensional row vector function.

3. *A constant returns to scale production technology process S with dynamics*

$$dS_t = \alpha(Y_t)S_t dt + S_t\gamma(Y_t)dW_t$$

The interpretation of this is that Y is a process which in some way influences the economy. It could for example describe the weather. The interpretation of the production technology is as in Lecture 2 and we have again a shortselling constraint.

Endogenous objects

In this model we also have some processes which are to be determined endogenously in equilibrium. They are as follows, where we use the notation

$$X_t = \text{the portfolio value at time } t,$$

to be more precisely defined below.

1. A risk free asset B, in zero net supply, with dynamics

$$dB_t = r_t B_t dt$$

   where the risk free rate r is assumed to be of the form

$$r_t = r(t, X_t, Y_t).$$

2. A financial derivative process $F(t, X_t, Y_t)$, in zero net supply, defined in terms of X and Y, without dividends and with dynamics of the form

$$dF_t = \beta_t F_t dt + F_t h_t dW_t$$

The processes β and h are assumed to be of the form

$$\beta_t = \beta(t, X_t, Y_t), \quad h_t = h(t, X_t, Y_t),$$

and will be determined in equilibrium.

We also need an important assumption.

Assumption: We assume that the 2×2 diffusion matrix

$$\begin{pmatrix} -\gamma- \\ -h- \end{pmatrix}$$

is invertible P-a.s. for all t

Note: The implication of the invertibility assumption is that, apart from the shortselling constraint for S, the market consisting of S, F, and B is **complete**. This is very important.

Economic agents

The basic assumption in CIR-85a is that there are a finite number of agents with identical initial capital, identical beliefs about the world, and identical preferences. In the present complete market setting this implies that we may as well consider a single representative agent. The object of the agent is (loosely) to maximize expected utility of the form

$$E^P\left[\int_0^T U(t, c_t, Y_t)dt\right]$$

where c is the consumption rate (measured in dollars per time unit) and U is the utility function.

3.2

The Dynamic Programming Approach

This is the approach taken in the original CIR paper. We will follow CIR rather closely, but at some points we use modern arbitrage theory in order to have shorter and more clear arguments. In Lecture 3.3 we will present the same theory using the martingale approach.

3.2.1

The control problem and HJB

Portfolio dynamics

The agent can invest in S, F, and B and. We will use the following notation

$$\begin{aligned} X &= \text{portfolio market value}, \\ a &= \text{portfolio weight on } S, \\ b &= \text{portfolio weight on } F, \\ 1-a-b &= \text{portfolio weight on } B \end{aligned}$$

Using standard theory we see that the portfolio dynamics are given by

$$dX_t = a_t X_t \frac{dS_t}{S_t} + b_t X_t \frac{dF_t}{F_t} + (1 - a_t - b_t) X_t \frac{dB_t}{B_t} - c_t dt$$

This gives us the portfolio dynamics as

$$dX_t = X_t \left\{a(\alpha - r) + b(\beta - r)\right\} dt + (rX_t - c)\, dt + X_t \left\{a\gamma + bh\right\} dW_t,$$

and we write this more compactly as

$$dX_t = X_t m(t, X_t, Y_t, u_t)dt - c_t dt + X_t g(t, X_t, Y_t, u_t)dW_t,$$

where we use the shorthand notation

$$u = (a, b),$$

and where m and g are defined by

$$\begin{aligned} m &= a\left[\alpha - r\right] + b\left[\beta - r\right] + r, \\ g &= a\gamma + bh. \end{aligned}$$

The control problem

The control problem for the agent is to maximize

$$E^P\left[\int_0^\tau U(t, c_t, Y_t)dt\right]$$

where

$$\tau = \inf\{t \geq 0 : \; X_t = 0\} \wedge T$$

subject to the portfolio dynamics

$$dX_t = X_t m(t, X_t, Y_t, u_t)dt - c_t dt + X_t g(t, X_t, Y_t, u_t)dW_t,$$

and the control constraints

$$c \geq 0, \quad a \geq 0.$$

The HJB equation

The HJB equation for this is straightforward and reads as

$$
\left\{
\begin{array}{rcl}
V_t + \sup_{c,u} \{U + \mathbf{A}^u V\} & = & 0, \\
V(T,x) & = & 0, \\
V(t,0) & = & 0,
\end{array}
\right.
\qquad (5)
$$

The infinitesimal operator $\mathbf{A}^u$ is defined by

$$
\mathbf{A}^u V = (xm-c)V_x + \mu V_y + \frac{1}{2}x^2 g^2 V_{xx} + \frac{1}{2}\sigma^2 V_{yy} + xg\sigma V_{xy}.
$$

For the vectors σ and g in R^2, we have used the notation

$$
\begin{array}{rcl}
\sigma g & = & (\sigma, g), \\
g^2 & = & \|g\|^2, \\
\sigma^2 & = & \|\sigma\|^2
\end{array}
$$

where (σ, g) denotes inner product.

3.2.2

Equilibrium

Equilibrium definition

Since B and F are in zero net supply, we have the following definition of equilibrium.

Definition: *An equilibrium is a list of processes*

$$\{r, \beta, h, a, b, c, V\}$$

such that (V, a, b, c) *solves the HJB equation given* (r, β, h)*, and the market clearing conditions*

$$a_t = 1, \quad b_t = 0.$$

are satisfied.

We will now study the implications of the equilibrium conditions on the short rate r and the dynamics of F. We do this by studying the first order conditions for optimality in the HJB equations, with the equilibrium conditions in force.

First order conditions

The first order conditions, with the equilibrium conditions $a = 1$ and $b = 0$ inserted, are easily seen to be as follows.

$$
\begin{array}{lrcl}
(a) & x(\alpha - r)V_x + x^2\gamma^2 V_{xx} + x\gamma\sigma V_{xy} & = & 0, \\
(b) & x(\beta - r)V_x + x^2\gamma h V_{xx} + x\sigma h V_{xy} & = & 0, \\
(c) & U_c & = & V_x,
\end{array}
$$

where (a) indicates that it is the FOC for a etc.

The equilibrium HJB eqn

In equilibrium, the following hold.

- The HJB equations takes the form

$$\begin{aligned} V_t + \sup_c \left\{ U + (\alpha x - \hat{c})V_x + \mu V_y + \frac{1}{2}x^2\gamma^2 V_{xx} + \frac{1}{2}\sigma^2 V_{yy} + x\sigma\gamma V_{xy} \right\} &= 0, \\ V(T, x, y) &= 0, \\ V(t, 0, y) &= 0. \end{aligned}$$

- The equilibrium portfolio dynamics are given by

$$d\widehat{X}_t = (\alpha \widehat{X}_t - \hat{c}_t)dt + \widehat{X}_t \gamma dW_t$$

Remark

We will see below that "everything" in the model, like the risk free rate, the Girsanov kernel, risk premia etc, are determined by the equilibrium optimal value function V.

It is then important, and perhaps surprising, to note that the equilibrium HJB equation is completely determined by **exogenous data**, i.e. by the Y and S dynamics. In other words, the equilibrium short rate, risk premia etc, do **not** depend on the particular choice of derivative F (or on the F dynamics) that we use in order to complete the market.

3.2.3

The equilibrium short rate

The short rate

From the FOC for a

$$x(\alpha - r)V_x + x^2\gamma^2 V_{xx} + x\gamma\sigma V_{xy} = 0$$

we immediately obtain our first main result.

Proposition: *The equilibrium short rate* $r(t, x, y)$ *is given by*

$$r = \alpha + \gamma^2 \frac{xV_{xx}}{V_x} + \gamma\sigma \frac{V_{xy}}{V_x}$$

With obvious notation we can write this as

$$r = \alpha - \left(-\frac{xV_{xx}}{V_x}\right) Var\left(\frac{dX}{X}\right) - \left(-\frac{V_{xy}}{V_x}\right) Cov\left[\frac{dX}{X}, dY\right].$$

3.2.4

Risk premium, the SDF and the EMM

The risk premium

From the equilibrium optimality condition for b

$$x(\beta - r)V_x + x^2\gamma h V_{xx} + x\sigma h V_{xy} = 0$$

we obtain the risk premium for F in equilibrium as

$$\beta - r = -\left[\frac{xV_{xx}}{V_x}\gamma h + \frac{V_{xy}}{V_x}\sigma h\right]$$

The martingale measure

Since every equilibrium must be arbitrage free, we can in fact push the analysis further. We denote by φ the Girsanov kernel for the likelihood process $L = \frac{dQ}{dP}$, so L has dynamics

$$dL_t = L_t \varphi_t dW_t.$$

We know from arbitrage theory that the martingale conditions for S and F are

$$\begin{aligned} r &= \alpha + \gamma\varphi, \\ r &= \beta + h\varphi \end{aligned}$$

On the other hand we have, from the equations for the short rate, and the risk premium for F, respectively

$$\begin{aligned} r &= \alpha + \left\{ \frac{xV_{xx}}{V_x}\gamma + \frac{V_{xy}}{V_x}\sigma \right\} \gamma, \\ r &= \beta + \left\{ \frac{xV_{xx}}{V_x}\gamma + \frac{V_{xy}}{V_x}\sigma \right\} h \end{aligned}$$

Since, by assumption, the matrix

$$\begin{pmatrix} -\gamma- \\ -h- \end{pmatrix}$$

is invertible, we have the following result.

Proposition: *The Girsanov kernel φ is given by*

$$\varphi = \frac{xV_{xx}}{V_x}\gamma + \frac{V_{xy}}{V_x}\sigma.$$

The stochastic discount factor

We expect to have the relation

$$V_x(t, X_t, Y_t) = \lambda M_t,$$

along the equilibrium X-path, where M is the stochastic discount factor

$$M_t = B_t^{-1} L_t,$$

and λ is the Lagrange multiplier, which can be written as

$$\lambda = V_x(0, X_0, Y_0).$$

This result is clear from general martingale theory theory, but one can also derive it using a more bare hands approach by first recalling that the dynamics of $Z_t = \lambda M_t$ are given by

$$dZ_t = -rZ_t dt + Z_t \varphi dW_t,$$

with φ as above. We can then use the Ito formula on V_x and the envelope theorem on the HJB equation in equilibrium to compute dV_x. After lengthy calculations we obtain

$$dV_x = -rV_x dt + V_x \varphi dW_t.$$

Comparing this with the Z dynamics above gives us the following result.

Proposition: *The stochastic discount factor in equilibrium is given by*

$$M_t = \frac{V_x(t, X_t, Y_t)}{V_x(0, X_0, Y_0)}.$$

3.2.5

Risk neutral valuation

Risk neutral valuation

We now go on to derive the relevant theory of risk neutral valuation within the model. This can be done in (at least) two ways:

- We can follow the argument in the original CIR paper and use PDE techniques.
- We can use more general arbitrage theory using martingale measures.

To illustrate the difference we will in fact present both arguments, and we start with the martingale argument. The reader will notice that the modern martingale argument is considerably more streamlined the the traditional PDE argument.

The martingale argument

From general arbitrage theory we immediately have the standard risk neutral valuation formula

$$F(t,x,y) = E^Q_{t,x,y}\left[e^{-\int_t^T r_s ds} H(X_T, Y_T)\right]$$

where H is the contract function for F. The equilibrium Q-dynamics of X and Y are given by

$$\begin{aligned} d\hat{X}_t &= \hat{X}_t\left[\alpha + \gamma\varphi\right]dt - \hat{c}_t dt + \hat{X}_t \gamma dW_t^Q, \\ dY_t &= \left[\mu + \sigma\varphi\right]dt + \sigma dW_t^Q. \end{aligned}$$

We thus deduce that the pricing function F is the solution of the PDE

$$\begin{cases} F_t + F_x x(\alpha + \gamma\varphi) - cF_x + \frac{1}{2}x^2\gamma^2 F_{xx} & \\ +F_y(\mu + \sigma\varphi) + \frac{1}{2}F_{yy}\sigma^2 + xF_{xy}\sigma\gamma - rF = 0 & \\ F(T,x,y) = H(x,y) & \end{cases}$$

which is Kolmogorov backward equation for the expectation above.

The PDE argument of CIR

Using the Ito formula to compute dF and comparing with the dynamics

$$dF = F\beta dt + FhdW_t$$

allows us to identify β as

$$\beta = \frac{1}{F}\left\{F_t + (\alpha x - c)F_x + \mu F_y + \frac{1}{2}x^2\gamma^2 F_{xx} + \frac{1}{2}\sigma^2 F_{yy} + x\sigma\gamma F_{xy}\right\}$$

On the other hand we have

$$\beta - r = -\varphi h$$

with φ given above, and we also have

$$h = \frac{1}{F}\left\{xF_x\gamma + F_y\sigma\right\}$$

so we have

$$\beta = r - \frac{1}{F}\left\{xF_x\gamma\varphi + F_y\sigma\varphi\right\}$$

Comparing the two expressions for β gives us the basic pricing PDE

$$\left\{\begin{array}{rcl} F_t + F_x x(\alpha + \gamma\varphi) - cF_x + \frac{1}{2}x^2\gamma^2 F_{xx} & & \\ +F_y(\mu + \sigma\varphi) + \frac{1}{2}F_{yy}\sigma^2 + xF_{xy}\sigma\gamma - rF & = & 0 \\ F(T,x,y) & = & H(x,y) \end{array}\right.$$

which is (of course) identical to the Kolmogorov eqn above. Using Feynman-Kac we then obtain the standard risk neutral valuation formula as

$$F(t,x,y) = E^Q_{t,x,y}\left[e^{-\int_t^T r_s ds} H(X_T, Y_T)\right]$$

Another formula for φ

We recall the formula

$$\varphi = \frac{xV_{xx}}{V_x}\gamma + \frac{V_{xy}}{V_x}\sigma$$

for the Girsanov kernel. We also recall from the first order condition for consumption, that

$$U_c = V_x.$$

Let us now specialize to the case when the utility function has the form

$$U(t, c, y) = e^{-\delta t}U(c)$$

Along the equilibrium path we then have

$$V_x(t, X_t, Y_t) = e^{-\delta t}U'(\hat{c}(t, X_t, Y_t))$$

and differentiating this equation proves the following result.

Proposition: *Under the assumption*

$$U(t, c, y) = e^{-\delta t} U(c)$$

the Girsanov kernel is given by

$$\varphi = \frac{U''(\hat{c})}{U'(\hat{c})} \left\{ x \hat{c}_x \gamma + \hat{c}_y \sigma \right\}$$

along the equilibrium path.

3.2.6

A central planner

Introducing a central planner

As in Lecture 2.2 we now introduce a central planner into the economy. This means that there is no market for B and F, so the central planner only chooses the consumption rate, invests everything into S, and the problem is thus to maximize

$$E^P\left[\int_0^\tau U(t, c_t, Y_t)dt + \Phi(X_T)\right]$$

subject to the dynamics

$$\begin{aligned}
dX_t &= (\alpha X_t - c)dt + X_t \gamma dt, \\
dY_t &= \mu(Y_t)dt + \sigma(Y_t)dW_t
\end{aligned}$$

and the constraint $c \geq 0$.

HJB for the central planner

The Bellman equation for this problem is

$$\begin{cases} V_t + \sup_c \left\{ U + (\alpha x - c)V_x + \mu V_y \frac{1}{2}\gamma^2 V_{xx} + \frac{1}{2}\sigma^2 V_{yy} + V_{xy}\sigma\gamma \right\} & = & 0 \\ V(T,x) & = & \Phi(x) \\ V(t,0) & = & 0 \end{cases}$$

We now see that this is exactly the equilibrium Bellman equation in the CIR model. We thus have the following result.

Central planner theorem

Given assumptions as above, the following hold.

- The optimal consumption for the central planner coincides with the equilibrium consumption of the representative agent.

- The optimal wealth process for the central planner is identical with the equilibrium wealth process for the representative agent.

Central planner vs equilibrium

- Solve the problem for the central planner, thus computing V.

- Define the "shadow interest rate" r and the Girsanov kernel φ by

$$\begin{aligned} r &= \alpha + \left\{ \frac{xV_{xx}}{V_x}\gamma + \frac{V_{xy}}{V_x}\sigma \right\}\gamma, \\ \varphi &= \frac{xV_{xx}}{V_x}\gamma + \frac{V_{xy}}{V_x}\sigma. \end{aligned}$$

- For a derivative with contract function H, define F by

$$F(t,x,y) = E^Q_{t,x,y}\left[e^{-\int_t^T r_s ds} H(X_T, Y_T) \right]$$

- Define and h and β by

$$\begin{aligned} h &= \frac{1}{F}\left\{ xF_x\gamma + F_y\sigma \right\} \\ \beta &= r - \frac{1}{F}\left\{ xF_x\gamma\varphi + F_y\sigma\varphi \right\} \end{aligned}$$

- The F dynamics will now be

$$dF = \beta F dt + F h dW_t.$$

- Now forget about the central planner and consider the optimal consumption/investment problem of a representative agent with access to the production technology S, the derivative F and the risk free asset B with dynamics

$$dB_t = r(t, X_t) B_t dt$$

where r is defined as above.

- The economy will then be in equilibrium, so $a = 1$, $b = 0$ and we will recover the optimal consumption and wealth processes of the central planner.

3.3

The Martingale Approach

In this section we study the CIR model from a a martingale point of view. This was not done in the original paper (the relevant martingale theory was not well known at the time of the CIR paper), and we will see that the martingale method greatly simplifies the analysis.

3.3.1

Generalities

The problem

Applying the usual arguments we then want to maximize expected utility

$$E^P\left[\int_0^\tau U(t, c_t, Y_t)dt + \Phi(X_T)\right]$$

over (c, X) given the budget constraint

$$E^P\left[\int_0^\tau c_t M_t dt + \Phi(X_T)M_T\right] = x_0$$

where, as usual, M is the stochastic discount factor and L is the likelihood process $L = dQ/dP$. We note that M will be determined endogenously in equilibrium. The Lagrangian for this problem is

$$E^P\left[\int_0^T \{U - Z_t c_t\} dt + \Phi(X_T) - Z_t X_T\right] + \lambda x_0$$

where

$$Z_t = \lambda M_t.$$

The first order conditions are

$$\begin{aligned} U_c(t, \hat{c}_t, Y_t) &= Z_t, \\ \Phi'(\hat{X}_T) &= Z_T, \end{aligned}$$

and, comparing the FOC for c with the FOC in the HJB eqn gives us the following expected result.

Proposition: *In equilibrium we have the identification*

$$V_x(t, \hat{X}_t, Y_t) = \lambda M_t,$$

where

$$\lambda = V_x(0, x_0, y_0)$$

Denoting the inverse of $U_c(t, c, y)$ in the c variable by $G(t, z, y)$ and the inverse of Φ' by F we have

$$\begin{aligned} \hat{c}(t, z, y) &= G(t, z, y), \\ \hat{X}_T &= F(Z_T). \end{aligned}$$

3.3.2

The short rate and the EMM

A Markovian assumption

We need a slight modification of an earlier assumption. **Assumption:** *We assume that the equilibrium short rate r and the equilibrium Girsanov kernel φ have the form*

$$
\begin{aligned}
r &= r(t, Z_t, Y_y), \\
\varphi &= \varphi(t, Z_t, Y_t).
\end{aligned}
$$

From risk neutral valuation we obtain the optimal wealth process X

$$
X_t = E^Q \left[\int_t^T e^{-\int_t^s r_u du} G(s, Z_s, Y_s) ds + e^{-\int_t^T r_u du} F(Z_T) \middle| \mathcal{F}_t \right]
$$

The Kolmogorov equation

The Markovian structure allows us to express X as

$$X_t = H(t, Z_t, Y_t)$$

where H solves a Kolmogorov equation. In order to find this equation we need the Q dynamics of Z, and these are easily obtained as

$$dZ_t = (\varphi^2 - r)Z_t dt + Z_t \varphi dW_t^Q.$$

The Kolmogorov equation is now

$$\begin{cases} H_t + \mathcal{A}H + G - rH & = & 0, \\ H(T, x, y) & = & F(z) \end{cases}$$

where the infinitesimal operator $\mathcal{A}$ is defined by

$$\mathcal{A}H = (\varphi^2 - r)zH_z + \mu H_y + \frac{1}{2}\varphi^2 z^2 H_{zz} + \frac{1}{2}\sigma^2 H_{yy} + \varphi\sigma H_{zy}$$

We can now use Ito to express the X dynamics as

$$dX_t = (\ldots)dt + \left\{ Z_t H_z \varphi + H_y \sigma \right\} dW_t$$

On the other hand, we know from general theory that the X dynamics in equilibrium are given by

$$dX_t = (\ldots)dt + X_t \gamma dW_t,$$

so, using $X_t = H(t, Z_t, Y_t)$ we obtain

$$zH_z \varphi + H_y \sigma = H\gamma,$$

giving us

$$\varphi = \frac{H}{zH_z}\gamma - \frac{H_y}{zH_z}\sigma.$$

The martingale condition for S is obviously

$$r = \alpha + \varphi\gamma,$$

which is our formula for the equilibrium interest rate. We may now summarize.

Proposition

The equilibrium interest rate $r(t,z,y)$ and Girsanov kernel $\varphi(t,z,y)$ are given by

$$\begin{aligned} r &= \alpha + \frac{H}{zH_z}\gamma^2 - \frac{H_y}{zH_z}\sigma\gamma, \\ \varphi &= \frac{H}{zH_z}\gamma - \frac{H_y}{zH_z}\sigma, . \end{aligned}$$

Here the function $H(t,z,y)$ is determined by the PDE

$$\begin{cases} H_t + \mathcal{A}H + G - rH &= 0, \\ H(T,x,y) &= F(z) \end{cases}$$

with $\mathcal{A}$ is defined by

$$\mathcal{A}H = (\varphi^2 - r)zH_z + \mu H_y + \frac{1}{2}\varphi^2 z^2 H_{zz} + \frac{1}{2}\sigma^2 H_{yy} + \varphi\sigma H_{zy}$$

and r and φ replaced by the formulas above.

Comment

Inserting the expressions for r and φ into the PDE above will result in a really horrible PDE, and I am rather at a loss to see how to proceed with that object. One alternative is to derive H under P instead of under Q, but I did not have time to do this yet.

Equilibrium Theory in Continuous Time

Lecture 4

The CIR interest rate model

Tomas Björk

Main objective

The goal of this lecture is to derive the CIR interest rate model. This model is in fact a simple special case of the general CIR factor model studied in the previous lecture.

4.1

The model

Log utility

The CIR interest rate model is a special case of the production model. We start by specializing to log utility.

Assumption: *We assume that the local utility function is of the form*

$$U(t, c, y) = e^{-\delta t} \ln(c)$$

where δ is the subjective discount factor of the agent.

We will now study (special cases of) this model using both DynP and the martingale approach.

4.2

Dynamic programming

HJB

It is easy to see that the HJB equation has a solution of the form

$$V(t,x,y) = e^{-\delta t} f(t,y) \ln(x) + e^{-\delta t} g(t,y).$$

and we obtain the following PDE for f.

$$\begin{cases} f_t + \mu f_y + \frac{1}{2}\sigma^2 f_{yy} - \delta f + 1 & = & 0, \\ F(T,y) & = & 0. \end{cases}$$

Using Feynman-Kac it is easily seen that f is in fact given by the simple formula

$$f(t,y) = \frac{1}{\delta}\left[1 - e^{-\delta(T-t)}\right].$$

so we have

$$\frac{xV_{xx}}{V_x} = -1, \quad \frac{V_{xy}}{V_x} = 0,$$

The short rate

We recall the formula for the short rate:

$$r = \alpha + \gamma^2 \frac{xV_{xx}}{V_x} + \gamma\sigma \frac{V_{xy}}{V_x}$$

so using

$$\frac{xV_{xx}}{V_x} = -1, \quad \frac{V_{xy}}{V_x} = 0,$$

gives us the short rate as follows.

Proposition: *The short rate is given by*

$$r(t, y) = \alpha(y) - \gamma^2(y).$$

We now specialize further.

The CIR special case

Recall

$$r(t, y) = \alpha(y) - \gamma^2(y).$$

Given this formula it is now natural to specialize further by assuming that

$$\alpha(y) = \alpha \cdot y, \quad \gamma(y) = \gamma \cdot \sqrt{y}.$$

which means that the S dynamics are of the form

$$dS_t = \alpha S_t Y_t dy + \gamma S_t \sqrt{Y_t} dW_t$$

In order to have a positive Y we introduce the assumption that the Y dynamics are of the form

$$dY_t = \{AY_t + B\} dt + \sigma \sqrt{Y_t} dW_t$$

where A, B and σ are positive constants so in the earlier notation we have

$$\begin{aligned} \mu(y) &= Ay + B, \\ \sigma(y) &= \sigma \sqrt{y}. \end{aligned}$$

We can now use our old formula

$$\varphi = \frac{xV_{xx}}{V_x}\gamma + \frac{V_{xy}}{V_x}\sigma$$

and the relations

$$\frac{xV_{xx}}{V_x} = -1, \quad \frac{V_{xy}}{V_x} = 0,$$

$$\alpha(y) = \alpha \cdot y, \quad \gamma(y) = \gamma \cdot \sqrt{y}.$$

to obtain our final result.

CIR Main Theorem

Assuming log utility and Y-dynamics of the form

$$dY_t = \{AY_t + B\}\,dt + \sigma\sqrt{Y_t}dW_t$$

the following hold.

- *The short rate is given by*

$$r(t, Y_t) = (\alpha - \gamma^2)Y_t.$$

- *The short rate dynamics under P are*

$$dr_t = [A + B_0]\,dt + \sigma_0\sqrt{r_t}dW_t,$$

  where

$$B_0 = (\alpha - \gamma^2)B, \quad \sigma_0 = \sigma\sqrt{\alpha - \gamma^2}.$$

- *The Girsanov kernel is given by*

$$\varphi(t, y) = -\gamma\sqrt{y}.$$

- *The Q dynamics of r are*

$$dr_t = [A_0 r_t + B_0]\, dt + \sigma_0 \sqrt{r_t} dW_t^Q$$

where

$$A_0 = A - \gamma\sigma\sqrt{\alpha - \gamma^2}.$$

Remark: The condition guaranteeing that the Y equation has a positive solution is

$$2A \geq \sigma^2.$$

This will also guarantee that the SDE for the short rate has a positive solution. In order to have a positive short rate we obviously also need to assume that

$$\alpha \geq \gamma^2.$$

4.3

Martingale analysis

The control problem

The problem is to maximize expected utility

$$E^P \left[\int_0^T e^{-\delta t} \ln(c_t) dt \right]$$

subject to the budget constraint

$$E^P \left[\int_0^T M_t c_t dt \right] = x_0$$

Performing the usual calculations, we obtain the optimal consumption as

$$\hat{c}_t = \lambda^{-1} M_t^{-1} e^{-\delta t}.$$

The short rate

From Lecture 3.3.2 we recall the formula

$$r = \alpha + \frac{H}{zH_z}\gamma^2 - \frac{H_y}{zH_z}\sigma\gamma,$$

where

$$H(t,z,y) = E^Q_{t,z,y}\left[\int_t^T \frac{B_t}{B_s}\hat{c}_s ds\right].$$

We can also write H as

$$H(t,z,y) = \frac{1}{M_t}E^P_{t,z,y}\left[\int_t^T M_s\hat{c}_s ds\right].$$

Inserting $\hat{c}_t = \lambda^{-1}M_t^{-1}e^{-\delta t}$ and recalling that $Z_t = \lambda M_t$ gives us the formula

$$H(t,z,y) = \frac{1}{z}g(t)$$

where

$$g(t) = \frac{1}{\delta}\left\{e^{-\delta t} - e^{-\delta T}\right\}$$

so we obtain

$$r(t, y) = \alpha(y) - \gamma^2(y),$$

and we can proceed as in the DynP analysis.

Equilibrium Theory in Continuous Time

Lecture 5

Endowment equilibrium models

Tomas Björk

Where are we going?

In the previous chapters we have studied equilibrium models in economies with a production technology. An alternative to that setup is to model an economy where each agent is exogenously endowed with a stream of income/consumption.

Endowment models are either **unit net supply** models or **zero net supply** models. Both model classes are treated in the lecture notes, but here we will only cover unit net supply models. Zero net supply models are technically more messy and provide identical results to those obtained by unit net supply models.

5.1

The model

Exogenous objects

Assumption: *The following objects are considered as given* a priori*.*

1. *A 1-dimensional Wiener process W.*

2. *A scalar and strictly positive process e of the form*

$$de_t = a(e_t)dt + b(e_t)dW_t \tag{6}$$

   where a and b is a scalar real valued functions.

The interpretation of this is that e is a an **endowment process** which provides the owner with a consumption stream at the rate e_t units of the consumption good per unit time, so during the time interval $[t, t + dt]$ the owner will obtain $e_t dt$ units of the consumption good.

Endogenous objects

The endogenous object in the model are as follows.

1. A risk free asset B, in **zero net supply**, with dynamics
$$dB_t = r_t B_t dt$$
where the risk free rate r is determined in equilibrium.

2. A price dividend pair (S, D) in **unit net supply**, where by assumption
$$dD_t = e_t dt.$$
In other words: Holding the asset S provides the owner with the dividend process e over the time interval $[0, T]$. Since S is defined in terms of e we can write the dynamics of S as
$$dS_t = \alpha_t S_t dt + \gamma_t S_t dW_t$$
where α and γ will be determined in equilibrium.

Comment

We stress the fact that, apart for providing the owner with the dividend process e over $[0, T]$, the asset S gives no further benefits to the owner. In equilibrium we will thus have

$$S_t = \frac{1}{M_t} E^P \left[\int_t^T M_s e_s ds \middle| \mathcal{F}_t \right],$$

where M is the equilibrium stochastic discount factor. In particular we will have

$$S_T = 0.$$

Economic agents

We consider a single representative agent who wants to maximize expected utility of the form

$$E^P\left[\int_0^T U(t, c_t)dt\right]$$

where c is the consumption rate (measured in dollars per time unit) and U is the utility function.

Assumption: *We assume that the agent has initial wealth $X_0 = S_0$. In other words: The agent has enough money to buy the right to the dividend process Y.*

We will use the notation

$$\begin{aligned} u_t &= \text{portfolio weight on the risky asset,} \\ 1 - u_t &= \text{portfolio weight on the risk free asset,} \\ c_t &= \text{rate of consumption.} \end{aligned}$$

Equilibrium

The natural equilibrium conditions are that the agent will, at all times, hold the risky asset (unit net supply), invest nothing in the risk free asset (zero net supply), and consume all dividends. Formally this reads as follows.

$$
\begin{aligned}
u_t &= 1, \quad ((S, D) \text{ in unit net supply}), \\
1 - u_t &= 0, \quad (B \text{ in zero net supply}), \\
c_t &= e_t, \quad (\text{market clearing for consumption}).
\end{aligned}
$$

5.2

Martingale analysis

DynP vs Martingale approach

It turns out that the DynP analysis of this model is quite tricky and a bit messy. Since the martingale approach is so much easier we confine ourselves to this method. See the lecture notes for details concerning the DynP approach.

5.2.1

The control problem and the equilibrium

The control problem

We assume again that the initial wealth of the agent is given by $X_0 = S_0$. The agent's control problem is then to maximize

$$E^P \left[\int_0^T U(t, c_t) dt \right]$$

subject to the following constraints.

$$\begin{aligned} c_t &\geq 0, \\ E^P \left[\int_0^T M_t c_t dt \right] &\leq S_0. \end{aligned}$$

Where M denotes the stochastic discount factor. The first constraint is obvious and the second one is the budget constraint.

Since the asset S provides the owner with the income stream defined by e and nothing else we can apply arbitrage theory to deduce that

$$S_0 = E^P \left[\int_0^T M_t e_t dt \right].$$

We can thus rewrite the budget constraint as

$$E^P \left[\int_0^T M_t c_t dt \right] \leq E^P \left[\int_0^T M_t e_t dt \right].$$

The Lagrangian is thus given by

$$E^P \left[\int_0^T \{U(t, c_t) - \lambda M_t c_t dt\} \right] + \lambda E^P \left[\int_0^T M_t e_t dt \right],$$

where λ is the Lagrange multiplier. and the optimality condition for c is thus

$$U_c(t, c_t) = Z_t,$$

where

$$Z_t = \lambda M_t.$$

Equilibrium conditions

As before we make the natural assumption that the processes α, γ and r are of the form

$$\begin{aligned}
\alpha_t &= \alpha(t, Z_t, e_t), \\
\gamma_t &= \gamma(t, Z_t, e_t), \\
r_t &= r(t, Z_t, e_t).
\end{aligned}$$

The equilibrium conditions are

$$\begin{aligned}
u_t &\equiv 1, \quad (S \text{ in unit net supply}), \\
1 - u_t &\equiv 0, \quad (B \text{ in zero net supply}), \\
c_t &\equiv e_t, \quad (\text{market clearing for consumption}).
\end{aligned}$$

The equilibrium short rate and the Girsanov kernel

The clearing condition $c = e$ and the optimality condition $U_c(t, c_t) = Z_t$ gives us

$$Z_t = U_c(t, e_t),$$

so we have

$$dZ_t = \left\{ U_{ct}(t, e_t) + a(e_t) U_{cc}(t, e_t) + \frac{1}{2} b^2(e_t) U_{ccc}(t, e_t) \right\} dt + b(e_t) U_{cc}(t, e_t) dW_t.$$

Using the formula

$$dZ_t = -r_t Z_t dt + Z_t \varphi_t dW_t.$$

we can thus identify the equilibrium rate and the Girsanov kernel as follows.

Theorem

The equilibrium short rate is given by

$$r(t,e) = -\frac{U_{ct}(t,e) + a(e)U_{cc}(t,e) + \frac{1}{2}b^2(e)U_{ccc}(t,e)}{U_c(t,e)}$$

and we see that the short rate r does in fact not depend explicitly on z. Furthermore, the Girsanov kernel is given by

$$\varphi(t,e) = \frac{U_{cc}(t,e)}{U_c(t,e)} \cdot b(e).$$

5.2.2

A special case

Log utility

To exemplify we now specialize to the log utility case when the local utility function is of the form

$$U(t,c) = e^{-\delta t}\ln(c).$$

In this case we have

$$U_c = \frac{1}{c}e^{-\delta t}, \quad U_{tc} = -\frac{\delta}{c}e^{-\delta t}, \quad U_{cc} = -\frac{1}{c^2}e^{-\delta t}, \quad U_{ccc} = \frac{2}{c^3}e^{-\delta t}$$

Plugging this into the formula for r and φ gives us

$$\begin{aligned} r(t,e) &= \delta + \frac{a(e)}{e} - \frac{b^2(e)}{e^2}, \\ \varphi(t,e) &= -\frac{b(e)}{e}. \end{aligned}$$

Specializing further

Given the expressions

$$
\begin{aligned}
r(t,e) &= \delta + \frac{a(e)}{e} - \frac{b^2(e)}{e^2}, \\
\varphi(t,e) &= -\frac{b(e)}{e},
\end{aligned}
$$

it is natural to specialize further to the case when the e dynamics are if the form

$$de_t = ae_t dt + be_t dW_t,$$

so that

$$a(e) = a \cdot e, \quad b(e) = b \cdot e.$$

We then obtain constant r and φ of the form

$$
\begin{aligned}
r &= \delta + a - b^2, \\
\varphi &= -b.
\end{aligned}
$$

5.3

Extending the model

In this section we extend the model to allow for a more general endowment process. As a special case we consider a factor model.

5.3.1

The general case

Basic assumptions

- We assume that the endowment process has the structure
$$de_t = a_t dt + b_t dW_t,$$
where W is a k-dimensional Wiener process, and where a and b are adapted to some given filtration $\mathbf{F}$. We assume that we have $N+1$ random sources in the model.

- We assume the asset-dividend pair (S, D) where $dD_t = e_t dt$, and we assume, as before, that S is in unit net supply.

- We assume the existence of a risk free asset B in zero net supply.

- We assume the existence of a number of assets $F_1, \ldots, F_N$, in zero net supply, which are defined in terms of the random sources, so that the market consisting of $S, B, F_1, \ldots, F_N$ is complete.

Main Result

We can now apply the usual martingale approach, and a moment of reflection will convince you that the argument in Lecture 5.2 goes through without any essential change. We thus conclude that for this extended model we have the following result.

Theorem: *With assumptions as above, the following hold.*

- *The equilibrium short rate process is given by*

$$r_t = -\frac{U_{ct}(t,e) + a_t U_{cc}(t,e_t) + \frac{1}{2}\|b_t\|^2 U_{ccc}(t,e_t)}{U_c(t,e_t)}.$$

- *The Girsanov kernel is given by*

$$\varphi_t = \frac{U_{cc}(t,e_t)}{U_c(t,e_t)} \cdot b_t.$$

5.3.2

A factor model

The model

We exemplify the theory of the previous section by considering a factor model of the form

$$
\begin{aligned}
de_t &= a(e_t, Y_t)dt + b(e_t, Y_t)dW_t, \\
dY_t &= \mu(Y_t)dt + \sigma(Y_t)dW_t.
\end{aligned}
$$

where W is a standard two dimensional Wiener process. For simplicity we assume log utility, so

$$
U(t, c) = e^{-\delta t} \ln(c).
$$

In this case the equilibrium rate and the Girsanov kernel will be of the form $r_t = r(e_t, Y_t)$, $\varphi_t = \varphi(e_t, Y_t)$ and after some easy calculations we obtain

$$
\begin{aligned}
r(e, y) &= \delta + \frac{a(e, y)}{e} - \frac{\|b(e, y)\|^2}{e^2}, \\
\varphi(e, y) &= -\frac{b(e, y)}{e}.
\end{aligned}
$$

Specializing further

Given these expressions it is natural to make the further assumption that a and b are of the form

$$a(e,y) = e \cdot a(y), \quad b(e,y) = e \cdot b(y),$$

which implies

$$\begin{aligned} r(y) &= \delta + a(y) - \|b(y)\|^2 \\ \varphi(y) &= -b(y). \end{aligned}$$

We now specialize further to the case when

$$a(y) = a \cdot y, \quad b(y) = \sqrt{y} \cdot b,$$

and in order to guarantee positivity of Y we assume

$$\begin{aligned} \mu(y) &= \beta + \mu \cdot y, \\ \sigma(y) &= \sqrt{y} \cdot \sigma \end{aligned}$$

where $2\beta \geq \|\sigma\|^2$. We then have the following result.

Theorem

Assume that the model has the structure

$$
\begin{aligned}
de_t &= ae_tY_tdt + e_t\sqrt{Y_t}bdW_t, \\
dY_t &= \{\beta + \mu Y_t\} dt + \sqrt{Y_t}\sigma dW_t.
\end{aligned}
$$

Then the equilibrium short rate and the Girsanov kernel are given by

$$
\begin{aligned}
r_t &= \delta + \left(a - \|b\|^2\right) Y_t, \\
\varphi_t &= \sqrt{Y_t} \cdot b.
\end{aligned}
$$

We have thus essentially re-derived the Cox-Ingersoll-Ross short rate model, but now within an endowment framework.

A comment

We finish this section with a remark on the structure of the Girsanov transformation. Let us assume that, for a general utility function $U(t, c)$, the processes e and Y are driven by independent Wiener processes, so the model has the form

$$
\begin{aligned}
de_t &= a(e_t, Y_t)dt + b(e_t, Y_t)dW_t^e, \\
dY_t &= \mu(Y_t)dt + \sigma(Y_t)dW_t^Y.
\end{aligned}
$$

where W^e and W^Y are independent and where b and σ are scalar. Then the Girsanov kernel has the vector form

$$\varphi_t = \frac{U_{cc}(t, e_t)}{U_c(t, e_t)} \cdot [b(e_t, Y_t), 0]$$

so the likelihood dynamics are

$$dL_t = L_t \frac{U_{cc}(t, e_t)}{U_c(t, e_t)} \cdot b(e_t, Y_t)dW_t^e,$$

implying that the Girsanov transformation will only affect W^e and not W^Y.

Equilibrium Theory in Continuous Time

Lecture 6

The existence of a representative agent

Tomas Björk

Where are we going?

In this lecture we will show that, under rather general condition, a model with several agents can be replaced by an equivalent model with one representative agent. The model is a multi-agent version of the endowment model above.

Exogenous object

The following objects are considered as given *a priori*.

1. An n-dimensional Wiener process W.

2. An n-dimensional strictly positive column vector process $e = (e_1, \ldots, e_n)'$ with dynamics of the form

$$de_t = a_t dt + b_t dW_t$$

where a is an (adapted) R^n valued process and b is an adapted process taking values in the space of $n \times n$ matrices. With obvious notation we will write the dynamics of e_i as

$$de_{it} = a_{it} dt + b_{it} dW_t$$

The interpretation of this is that, for each i, e_i is a an endowment process which provides the owner with a consumption stream at the rate e_{it} units of the consumption good per unit time.

Endogenous object

The endogenous object in the model are as follows.

1. A risk free asset B, in zero net supply, with dynamics

$$dB_t = r_t B_t dt$$

   where the risk free rate r is determined in equilibrium.

2. A sequence of price dividend pairs $\left\{(S^i, D^i)\right\}_{i=1}^n$, all in in unit net supply, where by assumption

$$dD_t^i = e_{it} dt.$$

   In other words: Holding the asset S^i provides the owner with the dividend rate e_i. We write the dynamics of S^i as

$$dS_t^i = \alpha_{it} S_t^i dt + \gamma_{it} S_t^i dW_t, \quad i = 1, \dots n.$$

   where α and γ will be determined in equilibrium.

Economic agents

We consider d economic agent who wants to maximize expected utility of the form

$$E^P\left[\int_0^T U_i(t, c_{it})dt\right], \quad i = 1, \ldots, d,$$

where c_i is the consumption rate and U_i is the utility function for agent i. We assume that U_i is strictly concave in the c variable, and we also need an assumption on initial wealth.

Denoting the wealth process of agent i by X_i we assume that

$$\sum_{i=1}^{d} X_{i0} = \sum_{j=1}^{n} S_0^j$$

In other words: As a group, the agents have enough money to buy the dividend paying assets $S^1, \ldots S^n$.

Notation

We will use the notation

$$
\begin{aligned}
u_{ijt} &= \text{portfolio weight for agent } i \text{ on the risky asset } S^j, \\
u_{it} &= (u_{i1t}, \ldots, u_{int}), \quad \text{portfolio weights process for the risky assets} \\
1 - \sum_{j=1}^{n} u_{ijt} &= \text{portfolio weight for agent } i \text{ on the risk free asset}, \\
c_{it} &= \text{consumption rate for agent } i.
\end{aligned}
$$

Equilibrium conditions

The natural equilibrium conditions are

- The aggregate net demand will, at all times, be exactly one unit of each asset $S^1, \dots, S^n$.
- There is zero net demand of the risk free asset B.
- The consumption market will clear.

Formally this reads as follows.

Equilibrium definition

An equilibrium is a family of processes $\{u_{it}^{\star}\}_{i=1}^{d}$, $\{c_{it}^{\star}\}_{i=1}^{d}$, and $(S_t^1, \ldots, S_t^n)$ such that

1. Given the asset prices $(S_t^1, \ldots, S_t^n)$, the processes $u_{it}^{\star}$ and $c_{it}^{\star}$ are optimal for agent i.

2. The markets for risky assets clear:

$$\sum_{i=1}^{d} u_{ijt} X_{it} = S_t^j, \quad j = 1, \ldots, n.$$

3. There is zero net demand for the risk free asset:

$$\sum_{i=1}^{d} X_{it} \left(1 - \sum_{j=1}^{n} u_{ijt} \right) = 0.$$

4. The consumption market clears:

$$\sum_{i=1}^{d} c_{it} = \sum_{j=1}^{n} e_{jt}.$$

Optimality for the individual agent

We assume the existence of an equilibrium, with a corresponding stochastic discount factor process $M^\star$.

Using the martingale approach, the problem of the agent i is that of maximizing

$$E^P\left[\int_0^T U_i(t, c_{it})dt\right],$$

subject to the budget constraint

$$E\left[\int_0^T M_t^\star c_{it}dt\right] \le x_{i0}.$$

The Lagrange function for this is

$$\int_0^T \left\{U_i(t, c_{it}) - \lambda_i^\star M_t^\star c_{it}\right\} dt + \lambda_i^\star x_{i0}.$$

where $\lambda_i^\star$ is the Lagrange multiplier for agent i. Assuming an interior optimum, this gives us the following first order conditions.

First order conditions

The equilibrium is characterized by the following conditions.

$$
\begin{aligned}
U'_{ic}(t, c^{\star}_{it}) &= \lambda^{\star}_i M^{\star}_t, \\
E\left[\int_0^T M^{\star}_t c^{\star}_{it} dt\right] &= x_{i0}, \\
\sum_{i=1}^{d} c^{\star}_{it} &= \eta_t,
\end{aligned}
$$

where the aggregate endowment η is given by

$$
\eta_t = \sum_{j=1}^{n} e_{jt}.
$$

Constructing the representative agent

We consider the equilibrium SDF $M^\star$, consumption rates $c_1^\star, \ldots, c_d^\star$, and Lagrange multipliers $\lambda_1^\star, \ldots, \lambda_d^\star$. These objects will, in particular, satisfy the first order conditions above.

Definition: *The utility function* $U : R_+ \times R_+ \to R$ *is defined by*

$$U(t, c) = \sup_{c_1, \ldots, c_d} \sum_{i=1}^{d} \frac{1}{\lambda_i^\star} U_i(t, c_i)$$

subject to the constraints

$$\begin{aligned} \sum_{i=1}^{d} c_i &= c, \\ c_i &\geq 0, \quad i = 1, \ldots, d. \end{aligned}$$

For a given c we denote the optimal $c_1, \ldots, c_d$ by $\hat{c}_1(c), \ldots, \hat{c}_d(c)$.

Using elementary optimization theory, we know that (for a given $c \in R_+$) there exists a nonnegative Lagrange multiplier $q(c)$ such that the Lagrange function

$$\sum_{i=1}^{d} \frac{1}{\lambda_i^\star} U_i(t, c_i) - q(c) \left\{ \sum_{i=1}^{d} c_i - c \right\},$$

is maximized by $\hat{c}_1(c), \ldots, \hat{c}_d(c)$.

Assuming an interior optimum, we thus see that $\hat{c}_1(c), \ldots, \hat{c}_d(c)$ are characterized by the first order conditions

$$U'_{ic}(t, \hat{c}_i(c)) = \lambda_i^\star q(c), \quad i = 1, \ldots, d.$$

From the Envelope Theorem we also know that

$$U'_c(t, c) = q(c).$$

The existence result

We recall the multi-agent market model given above, with the corresponding equilibrium, characterized by the price system $(S^\star, B^\star)$, the stochastic discount factor $M^\star$, investment policies $u_1^\star, \dots, u_d^\star$, consumption policies $c_1^\star, \dots, c_d^\star$, and the corresponding Lagrange multipliers $\lambda_1^\star, \dots, \lambda_1^\star$.

Now let us consider the same market but with a single agent, namely the representative agent of the previous section, with the utility function specified above, and initial wealth $x_0 = \sum_{i=1}^d x_{i0}$. Using the martingale approach, the problem of the representative agent is that of maximizing

$$E^P \left[\int_0^T U(t, c_t) dt \right],$$

subject to the budget constraint

$$E \left[\int_0^T M_t^\star c_t dt \right] \le x_0.$$

The Lagrange function for this is

$$\int_0^T \{U(t, c_t) - \lambda M_t^\star c_t\} \, dt + \lambda x_0.$$

where λ is the Lagrange multiplier for the representative agent. The FOC:s are

$$U_c'(t, \hat{c}_t) = \lambda M_t^\star,$$

where λ is determined by

$$E\left[\int_0^T M_t^\star \hat{c}_t dt\right] = x_0.$$

Note: Note that, because of convexity, these conditions are necessary and sufficient for the determination of $\hat{c}$ and λ.

Existence Theorem

(i) *The equilibrium price system* $(S^\star, B^\star)$, *and stochastic discount factor* $M^\star$ *is also an equilibrium for the single agent, so*

$$\hat{c}_t = \eta_t,$$

where $\eta_t = \sum_i e_{it}$.

(ii) *In equilibrium, the multiplier* λ *for the representative agent is given by*

$$\lambda = 1.$$

(iii) *The multi agent equilibrium consumption processes* $c^\star_{1t}, \ldots, c^\star_{dt}$ *are given by*

$$c^\star_{it} = \hat{c}_i(\eta_t),$$

where $\hat{c}_i(c)$ *is defined earlier in connection with the utility function for the representative agent.*

Proof

It is enough (see the note above) to show that

$$\begin{aligned} U_c'(t, \eta_t) &= M_t^\star, \\ E\left[\int_0^T M_t^\star \eta_t dt\right] &= x_0, \\ c_{it}^\star &= \hat{c}_i(\eta_t). \end{aligned}$$

We now go on to prove these items.

Proof of $c_{it}^{\star} = \hat{c}_i(\eta_t)$

From the properties of $U(t, c)$ we have

$$U'_{ic}(t, \hat{c}_i(\eta_t)) = \lambda_i^{\star} q(\eta_t), \quad i = 1, \ldots, d.$$

and from the multi agent equilibrium condition we have

$$U'_{ic}(t, c_{it}^{\star}) = \lambda_i^{\star} M_t^{\star}, \quad i = 1, \ldots, d.$$

Since $U'_{ic}(t, c)$ is strictly decreasing in c, and since

$$\sum_{i=1}^{d} c_{it}^{\star} = \sum_{i=1}^{d} \hat{c}_i(\eta_t) = \eta_t,$$

it is easy (how?) to deduce that we have

$$\begin{aligned} q(\eta_t) &= M_t^{\star}, \\ \hat{c}_i(\eta_t) &= c_{it}^{\star}, \quad i = 1, \ldots, d, \end{aligned}$$

and we have thus proved $c_{it}^{\star} = \hat{c}_i(\eta_t)$.

Proof of $U_c'(t, \eta_t) = M_t^\star$

From the properties of $U(t, c)$ we recall

$$U_c'(t, \eta_t) = q(\eta_t),$$

and since $q(\eta_t) = M_t^\star$ we obtain

$$U_c'(t, \eta_t) = M_t^\star.$$

Proof of $E\left[\int_0^T M_t^{\star}\eta_t dt\right] = x_0$

This is just the aggregate budget constraint for the multi agent equilibrium. ∎

Equilibrium Theory in Continuous Time

Lecture 7

Non linear filtering theory

Tomas Björk

Where are we going?

In this lecture we give a brief overview of nonlinear filtering theory. The OH slides actually contain much more than we will really need for the purpose of equilibrium theory, but the extra material will (it is hoped) enhance the understanding of the theory.

The material that will be directly used in equilibrium theory later on in the course is contained in sections 7.2-7.5.

The single most important result for our equilibrium applications is the innovations theorem in section 7.2.2.

Contents

7.1

A motivating problem

A problem with stochastic rate of return

Model:

$$dS_t = S_t\alpha(Y_t)dt + S_t\sigma dW_t$$

W is scalar and Y is some factor process. We assume that (S, Y) is Markov and adapted to the filtration $\mathbf{F}$.

Wealth dynamics

$$dX_t = X_t u_t \left(\alpha - r\right) dt + rX_t dt + u_t X_t \sigma dW_t$$

Objective:

$$\max_u \quad E^P\left[\Phi(X_T)\right]$$

- If we can observe S **and** Y, so $u \in \mathbf{F}$, then this is a standard problem which can be treated with DynP.
- In this lecture we will, however, study the case with **partial observations**.

Partial observations

Recall S dynamics

$$dS_t = S_t\alpha(Y_t)dt + S_t\sigma dW_t$$

- We assume that we can only observe S, so $u \in \mathbf{F}^S$.
- Although we cannot observe Y and $\alpha(Y_t)$ directly, we can estimate Y_t on the basis of past observations of S.
- It thus seems natural to compute the conditional mean $E\left[Y_t \middle| \mathcal{F}_t^S\right]$, or even the entire conditional distribution $\mathcal{L}(Y_t|\mathcal{F}_t^S)$
- We need **filtering theory**.

7.2

Non linear filtering theory

7.2.1

Setup and problem formulation

Setup

Given some filtration $\mathbf{F}$:

$$\begin{aligned} dX_t &= a_t dt + dM_t \\ dZ_t &= b_t dt + dW_t \end{aligned}$$

Here all processes are $\mathbf{F}$ adapted and

$$\begin{aligned} X &= \text{state process,} \\ Z &= \text{observation process,} \\ M &= \text{martingale w.r.t. } \mathbf{F} \\ W &= \text{Wiener w.r.t. } \mathbf{F} \end{aligned}$$

We assume (for the moment) that M and W are **independent**.

Problem:
Compute (recursively) the filter estimate

$$\widehat{X}_t = \Pi_t [X] = E\left[X_t | \mathcal{F}_t^Z \right]$$

Typical example

A very commen example is given by

$$
\begin{aligned}
dX_t &= \mu(t, X_t)dt + \sigma(t, X_t)dV_t, \\
dZ_t &= b(t, X_t)dt + dW_t
\end{aligned}
$$

where W and V are Wiener.

7.2.2

The innovations process

The innovations process

Recall:

$$dZ_t = b_t dt + dW_t$$

Our best guess of b_t is $\hat{b}_t$, so the genuinely new information should be

$$dZ_t - \hat{b}_t dt$$

Definition:
The **innovations process** ν is defined by

$$\nu_t = dZ_t - \hat{b}_t dt$$

Theorem: The process ν is $\mathbf{F}^Z$-Wiener.

Proof: By Levy it is enough to show that

- ν is an $\mathbf{F}^Z$ martingale.
- $\nu_t^2 - t$ is an $\mathbf{F}^Z$ martingale.

I. ν is an $\mathbf{F}^Z$ martingale:

From definition we have

$$d\nu_t = \left(b_t - \widehat{b}_t\right) dt + dW_t \qquad (7)$$

so

$$E_s^Z\left[\nu_t - \nu_s\right] = \int_s^t E_s^Z\left[b_u - \widehat{b}_u\right] du + E_s^Z\left[W_t - W_s\right]$$

$$= \int_s^t E_s^Z\left[E_u^Z\left[b_u - \hat{b}_u\right]\right] du + E_s^Z\left[E_s\left[W_t - W_s\right]\right] = 0$$

I. $\nu_t^2 - t$ is an $\mathbf{F}^Z$ martingale:

From Itô we have

$$d\nu_t^2 = 2\nu_t d\nu_t + \left(d\nu_t\right)^2$$

Here $d\nu$ is a martingale increment and from (7) it follows that $\left(d\nu_t\right)^2 = dt$.

Important fact

Note that we now have two expressions for the Z dynamics. We have the original dynamics

$$dZ_t = b_t dt + dW_t,$$

and we have

$$dZ_t = \hat{b}_t dt + d\nu_t,$$

- It is extremely important to realize that the Z process in the left hand of these equations is, trajectory by trajectory, **exactly the same process**.
- The first equation gives us the Z- dynamics relative to the filtration $\mathbf{F}$, whereas the second equation gives us the Z-dynamics w.r.t. the $\mathbf{F}^Z$-filtration.
- We express this by saying that we have projected the Z dynamics onto the $\mathbf{F}^Z$ filtration.
- This projection technique is the basic tool for the equilibrium theory in the next lecture.

7.2.3

The filter dynamics

Filter dynamics

From the X dynamics we guess that

$$d\widehat{X}_t = \hat{a}_t dt + \text{martingale}$$

Definition: $dm_t = d\widehat{X}_t - \hat{a}_t dt.$

Proposition: m is an $\mathcal{F}_t^Z$ martingale.

Proof:

$$E_s^Z\left[m_t - m_s\right] = E_s^Z\left[\hat{X}_t - \hat{X}_s\right] - E_s^Z\left[\int_s^t \widehat{a}_u du\right]$$

$$= \quad E_s^Z\left[X_t - X_s\right] - E_s^Z\left[\int_s^t a_u du\right]$$

$$= \quad E_s^Z\left[M_t - M_s\right] = E_s^Z\left[E_s\left[M_t - M_s\right]\right] = 0$$

Filter dynamics

We now have the filter dynamics

$$d\widehat{X}_t = \hat{a}_t dt + dm_t$$

where m is an $\mathcal{F}_t^Z$ martingale.

If the **innovations hypothesis**

$$\mathcal{F}_t^Z = \mathcal{F}_t^\nu$$

is true, then the martingale representation theorem would give us an $\mathcal{F}_t^Z$ adapted process h such that

$$dm_t = h_t d\nu_t \tag{8}$$

The innovations hypothesis is not generally correct but FKK have proved that in fact (8) is always true.

Filter dynamics

We thus have the filter dynamics

$$d\widehat{X}_t = \widehat{a}_t dt + h_t d\nu_t$$

and it remains to determine the gain process h.

Proposition: The process h is given by

$$h_t = \widehat{X_t b_t} - \widehat{X}_t \widehat{b}_t$$

We give a slighty heuristic proof.

Proof sketch

From Itô we have

$$d\left(X_t Z_t\right) = X_t b_t dt + X_t dW_t + Z_t a_t dt + Z_t dM_t$$

using

$$d\widehat{X}_t = \widehat{a}_t dt + h_t d\nu_t$$

and

$$dZ_t = \hat{b}_t dt + d\nu_t$$

we have

$$d\left(\widehat{X}_t Z_t\right) = \widehat{X}_t \widehat{b}_t dt + \widehat{X}_t d\nu_t + Z_t \widehat{a}_t dt + Z_t h_t d\nu_t + h_t dt$$

Formally we also have (why?)

$$E\left[d\left(X_t Z_t\right) - d\left(\widehat{X}_t Z_t\right) \middle| \mathcal{F}_t^Z\right] = 0$$

which gives us

$$\left(\widehat{X_t b_t} - \widehat{X}_t \widehat{b}_t - h_t\right) dt = 0.$$

The FKK filter equations

For the model

$$
\begin{aligned}
dX_t &= a_t dt + dM_t \\
dZ_t &= b_t dt + dW_t
\end{aligned}
$$

where M and W are independent, we have the Fujisaki-Kallianpur-Kunita (FKK) non-linear filter equations

$$
\begin{aligned}
d\widehat{X}_t &= \widehat{a}_t dt + \left\{ \widehat{X_t b_t} - \widehat{X}_t \widehat{b}_t \right\} d\nu_t \\
d\nu_t &= dZ_t - \widehat{b}_t dt
\end{aligned}
$$

Remark: It is easy to see that

$$
h_t = E\left[\left(X_t - \widehat{X}_t \right) \left(b_t - \widehat{b}_t \right) \middle| \mathcal{F}_t^Z \right]
$$

The general filter equations

For the model

$$\begin{aligned} dX_t &= a_t dt + dM_t \\ dZ_t &= b_t dt + \sigma_t dW_t \end{aligned}$$

where

- The process σ is $\mathcal{F}_t^Z$ adapted and positive.
- There is no assumption of independence between M and W.

we have the filter

$$\begin{aligned} d\widehat{X}_t &= \widehat{a}_t dt + \left[\widehat{D}_t + \frac{1}{\sigma_t} \left\{ \widehat{X_t b_t} - \widehat{X}_t \widehat{b}_t \right\} \right] d\nu_t \\ d\nu_t &= \frac{1}{\sigma_t} \left\{ dZ_t - \widehat{b}_t dt \right\} \\ D_t &= \frac{d\langle M, W \rangle_t}{dt} \end{aligned}$$

Comment on $\langle M, W \rangle$

This requires semimartingale theory but there are two simple cases

- If M is Wiener then

$$d\langle M, W \rangle_t = dM_t dW_t$$

  with usual multiplication rules.

- If M is a pure jump process then

$$d\langle M, W \rangle_t = 0.$$

7.3

Filtering a Markov process and dimensionality problems

Filtering a Markov process

Assume that X is Markov with generator $\mathcal{G}$. We want to compute $\Pi_t[f(X_t)]$, for some nice function f. Dynkin's formula gives us

$$df(X_t) = (\mathcal{G}f)(X_t)dt + dM_t$$

Assume that the observations are

$$dZ_t = b(X_t)dt + dW_t$$

where W is independent of X.

The filter equations are now

$$\begin{aligned} d\Pi_t[f] &= \Pi_t[\mathcal{G}f]\,dt + \{\Pi_t[fb] - \Pi_t[f]\,\Pi_t[b]\}\,d\nu_t \\ d\nu_t &= dZ_t - \Pi_t[b]\,dt \end{aligned}$$

Remark: To obtain $d\Pi_t[f]$ we need $\Pi_t[\mathcal{G}f]$, $\Pi_t[fb]$, and $\Pi_t[b]$. This leads generically to an infinite dimensional system of filter equations.

On the filter dimension

$$d\Pi_t[f] = \Pi_t[\mathcal{G}f]\,dt + \{\Pi_t[fb] - \Pi_t[f]\,\Pi_t[b]\}\,d\nu_t$$

- To obtain $d\Pi_t[f]$ we need $\Pi_t[\mathcal{G}f]$, $\Pi_t[fb]$, $\Pi_t[b]$.
- We apply the FKK equations to $\mathcal{G}f$, fb, and b.
- This leads to new filter estimates to determine and generically to an **infinite dimensional** system of filter equations.
- The filter equations are really equations for the **entire conditional distribution** of X.
- You can only expect the filter to be finite when the conditional distribution of X is finitely parameterized.
- There are only very few examples of finite dimensional filters.
- The most well known finite filters are the Wonham and the Kalman filters.

7.4

The Wonham filter

The Wonham setting

Assume that X is a continuous time Markov chain on the state space $\{1, \ldots, n\}$ with (constant) generator matrix H, i.e.

$$P\left(X_{t+h} = j \,|X_t = i\right) = H_{ij}h + o(h),$$

for $i \neq j$ and

$$H_{ii} = -\sum_{j \neq i} H_{ij}$$

Define the indicator processes by

$$\delta_i(t) = I\left\{X_t = i\right\}, \quad i = 1, \ldots, n.$$

Dynkin's Theorem gives us

$$d\delta_i = \sum_j H_{ji}\delta_j dt + dM_t^i, \quad i = 1, \ldots, n.$$

The Wonham filter

Recall

$$d\delta_i = \sum_j H_{ji}\delta_j dt + dM_t^i, \quad i = 1, \ldots, n.$$

Observations are

$$dZ_t = b(X_t)dt + dW_t.$$

The filter equations are

$$d\Pi_t[\delta_i] = \sum_j H_{ji}\Pi_t[\delta_j]\,dt + \{\Pi_t[\delta_i b] - \Pi_t[\delta_i]\,\Pi_t[b]\}\,d\nu_t$$

$$d\nu_t = dZ_t - \Pi_t[b]\,dt$$

We note that

$$b(X_t) = \sum_i b_i \delta_i(t)$$

so

$$\begin{aligned} \Pi_t[\delta_i b] &= b_i \Pi_t[\delta_i], \\ \Pi_t[b] &= \sum_j b_j \Pi_t[\delta_j] \end{aligned}$$

We finally have the Wonham filter

$$\begin{aligned} d\widehat{\delta}_i &= \sum_j H_{ji} \widehat{\delta}_j dt + \left\{ b_i \widehat{\delta}_i - \widehat{\delta}_i \sum_j b_j \widehat{\delta}_j \right\} d\nu_t, \\ d\nu_t &= dZ_t - \sum_j b_j \widehat{\delta}_j dt \end{aligned}$$

7.5

The Kalman filter

The Kalman model

The Kalman model is a linear Gaussian system

$$
\begin{aligned}
dX_t &= aX_t dt + c dV_t, \\
dZ_t &= X_t dt + dW_t
\end{aligned}
$$

where W and V are independent Wiener.

Remark. We can have correlation between V and W as long as the correlation is not perfect.

The filter equations

$$
\begin{aligned}
dX_t &= aX_t dt + cdV_t, \\
dZ_t &= X_t dt + dW_t
\end{aligned}
$$

FKK gives us

$$
\begin{aligned}
d\Pi_t[X] &= a\Pi_t[X]\,dt + \left\{\Pi_t\left[X^2\right] - \left(\Pi_t[X]\right)^2\right\} d\nu_t \\
d\nu_t &= dZ_t - \Pi_t[X]\,dt
\end{aligned}
$$

We need $\Pi_t\left[X^2\right]$, so use Itô to get write

$$
dX_t^2 = \left\{2aX_t^2 + c^2\right\} dt + 2cX_t dV_t
$$

From FKK:

$$
\begin{aligned}
d\Pi_t\left[X^2\right] &= \left\{2a\Pi_t\left[X^2\right] + c^2\right\} dt \\
&+ \left\{\Pi_t\left[X^3\right] - \Pi_t\left[X^2\right]\Pi_t[X]\right\} d\nu_t
\end{aligned}
$$

Now we need $\Pi_t\left[X^3\right]$! Etc!

Define the conditional error variance by

$$H_t = \Pi_t\left[(X_t - \Pi_t[X])^2\right] = \Pi_t\left[X^2\right] - (\Pi_t[X])^2$$

Itô gives us

$$d(\Pi_t[X])^2 = \left[2a(\Pi_t[X])^2 + H^2\right]dt + 2\Pi_t[X]\,H d\nu_t$$

and Itô again

$$\begin{aligned} dH_t &= \left\{2aH_t + c^2 - H_t^2\right\}dt \\ &+ \left\{\Pi_t\left[X^3\right] - 3\Pi_t\left[X^2\right]\Pi_t[X] + 2(\Pi_t[X])^3\right\}d\nu_t \end{aligned}$$

In **this particular case** we know (why?) that the distribution of X conditional on Z is Gaussian!

Thus we have

$$\Pi_t\left[X^3\right] = 3\Pi_t\left[X^2\right]\Pi_t[X] - 2(\Pi_t[X])^3$$

so H is deterministic (as expected).

The Kalman filter

Model:

$$
\begin{aligned}
dX_t &= aX_t dt + c dV_t, \\
dZ_t &= X_t dt + dW_t
\end{aligned}
$$

Filter:

$$
\begin{aligned}
d\widehat{X} &= a\widehat{X} dt + H_t d\nu_t \\
\dot{H}_t &= 2aH_t + c^2 - H_t^2 \\
d\nu_t &= dZ_t - \widehat{X} dt
\end{aligned}
$$

$$
H_t = E\left[\left(X_t - \widehat{X}_t \right)^2 \middle| \mathcal{F}_t^Z \right]
$$

Remark: Because of the Gaussian structure, the conditional distribution evolves on a two dimensional submanifold. Hence a two dimensional filter.

7.6

The SPDE for the conditional density

The SPDE for the conditional density

Recall the FKK equation

$$d\Pi_t[f] = \Pi_t[\mathcal{G}f]\,dt + \{\Pi_t[fb] - \Pi_t[f]\,\Pi_t[b]\}\,d\nu_t$$

Now **assume** that X has a conditional density process $p_t(x)$, with interpretation

$$p_t(x)dx = E\left[X_t \in dx \middle| \mathcal{F}_t^Z\right]$$

so

$$\Pi_t[f] = E\left[f(X_t)\middle| \mathcal{F}_t^Z\right] = \int_{R^n} f(x)p_t(x)dx$$

Using the pairing $\langle f, g\rangle = \int f(x)g(x)dx$ we can write FKK as

$$d\langle f, p_t\rangle = \langle \mathcal{G}f, p_t\rangle dt + \{\langle fb, p_t\rangle - \langle f, p_t\rangle\langle b, p_t\rangle\}\,d\nu_t$$

Recall

$$d\langle f, p_t\rangle = \langle \mathcal{G}f, p_t\rangle dt + \left\{\langle fb, p_t\rangle - \langle f, p_t\rangle\langle b, p_t\rangle\right\} d\nu_t$$

We can now dualize this to obtain

$$d\langle f, p_t\rangle = \langle f, \mathcal{G}^\star p_t\rangle dt + \left\{\langle f, bp_t\rangle - \langle f, p_t\rangle\langle b, p_t\rangle\right\} d\nu_t$$

Since this holds for all test functions f we have the following result.

Theorem: The density function $p_t(x)$ satisfies the following stochastic partial differential equation (SPDE)

$$dp_t(x) = \mathcal{G}^\star p_t(x)dt + p_t(x)\left\{b(x) - \int_{R^n} b(y)p_t(y)dy\right\} d\nu_t$$

This SPDE is know as the **Kushner-Stratonovic equation**.

7.7

Unnormalized filter estimates

The main idea

We consider the following model under a measure P.

$$\begin{aligned} dX_t &= a(X_t)dt + b(X_t)dV_t, \\ dZ_t &= h(X_t)dt + dW_t \end{aligned}$$

where V and W are independent Wiener processes.

The SPDE for $p_t(x)$ is quite messy. We now present an alternative along the following lines.

- Perform a Girsanov transformation from P to Q so that X and Z are independent under Q.
- Compute filtering estimates under Q. This should be very easy, due to the independence.
- Transform the filter estimetes back from Q to P, using the abstract Bayes Formula.

The Basic Construction

Consider a probability space $(Q, \mathcal{F}, V, Z)$ where V and Z are independent Wiener processes under Q. Define X by

$$dX_t = a(X_t)dt + b(X_t)dV_t$$

and define $\mathbf{F}$ by

$$\mathcal{F}_t = \mathcal{F}_t^Z \vee \mathcal{F}_\infty^V$$

Define the likelihood process L by

$$dL_t = h(X_t)L_t dZ_t, \quad L_0 = 1$$

and define P by $dP = L_t dQ$ on $\mathcal{F}_t$. From Girsanov we deduce that W, defined by

$$dZ_t = h(X_t)dt + dW_t$$

is $(P, \mathbf{F})$-Wiener. In particular it is independent of $\mathcal{F}_0 = \mathcal{F}_\infty^V$, so W and V are P-independent. It is also easy to see (how?) that (X, V) has the same distribution under P as under Q. Under P we now have our standard model.

The unnormalized estimate

Define $\Pi_t[f]$ as usual by

$$\Pi_t[f] = E^P\left[f(X_t) \middle| \mathcal{F}_t^Z \right].$$

We then have, from Bayes,

$$\Pi_t[f] = \frac{E^Q\left[L_t f(X_t) \middle| \mathcal{F}_t^Z \right]}{E^Q\left[L_t \middle| \mathcal{F}_t^Z \right]}$$

Now define $\sigma_t[f]$ by

$$\sigma_t[f] = E^Q\left[L_t f(X_t) \middle| \mathcal{F}_t^Z \right]$$

which gives us the **Kallianpur-Striebel formula**

$$\Pi_t[f] = \frac{\sigma_t[f]}{\sigma_t[1]}$$

We can view $\sigma_t[f]$ as an **unnormalized** filter estimate of $f(X_t)$, and we now define the SDE for $\sigma_t[f]$.

The Zakai Equation

We have

$$\sigma_t[f] = \Pi_t[f] \cdot \sigma_t[1]$$

By FKK we already have an expression for $d\Pi_t[f]$ and one can show that

$$d\sigma_t[1] = \Pi_t[h]\,\sigma_t[1]\,dZ_t$$

From Ito, and after lots of calculations, we have the following result.

Theorem: The unnormalized filter estimate satisfies the **Zakai Equation**

$$d\sigma_t[f] = \sigma_t[\mathcal{G}f]\,dt + \sigma_t[hf]\,dZ_t$$

The SPDE for the unnormalized density

Let us now assume that there exists an unnormalized density process $q_t(x)$ with interpretation

$$\sigma_t [f] = \int_{R^n} f(x) q_t(x) dx$$

Arguing as before we then obtain the following result.

Theorem: The unnormalized density Q satisfies the SPDE

$$dq_t(x) = \mathcal{G}^{\star} q_t(x) dt + h(x) q_t(x) dZ_t$$

This is a **much** nicer equation than the corresponding equation for $p_t(x)$, since

- It is linear in q_t whereas the SPDE for p_t is quadratic in p_t.
- The equation for q is driven directly by the observations process Z, rather than by the innovations process ν.

7.8

Appendix:

Dynkin's Theorem and the Kolmogorov Equation

The generator

We consider a real valued Markov process X, and a real valued function $f(x)$.

Definition: The **infinitesimal generator** $\mathcal{A}$ is defined, for all f in the **domain** $\mathcal{D}$, by

$$[\mathcal{A}f](t,x) = \lim_{h \downarrow 0} \frac{1}{h} E_{t,x}\left[f(X_{t+h}) - f(x)\right],$$

where $\mathcal{D}$ is the subspace of bounded functions for which the limit exists for all (t,x).

Note: The operator $\mathcal{A}$ maps functions into functions. More precisely: if f is a function of x only, then $\mathcal{A}f$ is a function of (t,x).

The generator

Recall

$$\left[\mathcal{A}f\right](t,x) = \lim_{h \downarrow 0} \frac{1}{h} E_{t,x}\left[f(X_{t+h}) - f(x)\right],$$

Interpretation: Intuitively speaking we have

$$E_{t,x}\left[df(X_t)\right] = \left[\mathcal{A}f\right](t,x)dt$$

Note: It is easy (how?) to see that for a function of the form $f(t,x)$ we have

$$E_{t,x}\left[df(t,X_t)\right] = \left\{\frac{\partial f}{\partial t}(t,X_t)dt + \mathcal{A}f(t,x)\right\}dt$$

The SDE Case

Suppose that X solves an SDE of the form

$$dX_t = \mu(t, X_t)dt + \sigma(T, X_t)dW_t$$

it is then an easy exercise to see that

$$\mathcal{A}f(t,x) = \mu(t,x)\frac{\partial f}{\partial x}(t,x) + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 f}{\partial x^2}(t,x)$$

Intuition

Since we have

$$\left[\frac{\partial f}{\partial t} + \mathcal{A}f\right](t,x)dt = E_{t,x}\left[df(t,X_t)\right]$$

we expect the "detrended increment"

$$df(t,X_t) - \left[\frac{\partial f}{\partial t} + \mathcal{A}f\right](t,x)dt$$

to be martingale increment. This is basically the Dynkin Theorem

Dynkin's Theorem

Theorem: (Dynkin) Assume that X is a Markov process with infinitesimal generator $\mathcal{A}$. Then, for every f in the domain of $\mathcal{A}$, the process

$$M_t = f(t, X_t) - \int_0^t \left\{ \frac{\partial}{\partial t} + \mathcal{A} \right\} f(s, X_s) ds$$

is a martingale. We often write this as

$$df(t, X_t) = \left\{ \frac{\partial}{\partial t} + \mathcal{A} \right\} f(t, X_t) dt + dM_t$$

Furthermore, the process $f(t, X_t)$ is a martingale if and only if

$$\left\{ \frac{\partial}{\partial t} + \mathcal{A} \right\} f(t, x) = 0, \quad (t, x) \in R_+ \times R$$

The Kolmogorov Backward Equation

Let X be Markov (relative to some filtration $\mathbf{F}$) with generator $\mathcal{A}$, and let Φ be a real valued function. We now define the function f by

$$f(t,x) = E_{t,x}\left[\Phi(X_T)\right].$$

Since X is Markov we have

$$f(t,X_t) = E\left[\Phi(X_T)\middle|\, X_t\right] = E\left[\Phi(X_T)\middle|\, \mathcal{F}_t\right]$$

so the process $f(t,X_t)$ is obviously (why?) a martingale. We can then apply Dynkin to get the following result.

Theorem: (Kolmogorov) The function f solves the boundary value problem

$$\begin{cases} \dfrac{\partial f}{\partial t}(t,x) + \mathcal{A}f(t,x) &= 0, \\ F(T,x) &= \Phi(x) \end{cases}$$

This equation is known as the Kolmogorov backward equation.

Equilibrium Theory in Continuous Time

Lecture 8

Models with partial observations

Tomas Björk

Where are we going?

In this lecture we will study equilibrium models where the agents cannot observe all relevant stochastic processes. In the typical case the agent can observe the asset price processes, and/or the endowment process, but not the underlying factor processes.

8.1

A production model

8.1.1

The model

Basic model assumptions

- We assume the existence of a scalar production technology (with the usual interpretation) with dynamics given by

$$dS_t = Y_t S_t dt + S_t \sigma dW_t^s,$$

where W^s is Wiener.

- The scalar factor process Y, determining the rate of return on physical investment, is assumed to have dynamics given by

$$dY_t = (AY_t + B)dt + CdW_t^y,$$

where W^y is a Wiener process. For notational simplicity we assume that W^s and W^y are independent.

- The filtration generated by W^s and W^y is denoted by $\mathbf{F}$, so $\mathcal{F}_t = \sigma\left\{W_u^s, W_u^y;\ 0 \leq u \leq t\right\}$.

The agent

- We consider a representative agent with utility function
$$E^P\left[\int_0^T U(t,c_t)dt\right].$$

- The agent can observe the S process,but **not** the Y process, so all his actions must be adapted to the filtration $\mathbf{F}^S$, where $\mathcal{F}_t^S=\sigma\left\{S_u;\ 0\leq u\leq t\right\}$.

- The agent can invest in the following assets.
  - The physical production process S.
  - A risk free asset B in zero net supply with dynamics
  $$dB_t=r_tB_tdt,$$
  where the $\mathbf{F}^S$-adapted risk free rate of return r will be determined in equilibrium.

8.1.2

Projecting the S dynamics, and filtering

Projecting the S dynamics

We define the process Z by

$$dZ_t = \frac{dS_t}{\sigma S_t} \tag{9}$$

and we note that $\mathbf{F}^Z = \mathbf{F}^S$. We can write the observation dynamics as

$$dZ_t = \frac{Y_t}{\sigma}dt + dW_t^s.$$

and the innovations process ν is defined as usual by

$$d\nu_t = dZ_t - \frac{\hat{y}_t}{\sigma}dt,$$

or

$$dZ_t = \frac{\hat{y}_t}{\sigma}dt + d\nu_t \tag{10}$$

Equations (9)-(10) gives us the S-dynamics projected onto the observable filtration $\mathbf{F}^S$ as

$$dS_t = \hat{y}_t S_t dt + \sigma S_t d\nu_t.$$

Filtering equations

Recalling the dynamics of the pair (Y, Z) we have

$$
\begin{aligned}
dY_t &= (AY_t + B)dt + CdW_t^y, \\
dZ_t &= \frac{Y_t}{\sigma}dt + dW_t^s,
\end{aligned}
$$

and we recognize this as a standard Kalman model. We thus have the Kalman filter equations

$$
d\hat{y}_t = (A\hat{y}_t + B) + H_t d\nu_t,
$$

where H is deterministic.

8.1.3

The control problem

Portfolio dynamics

Assumption: *We assume that the risk free rate process* r *is of the form*

$$r_t = r(t, X_t, \hat{y}_t)$$

where X *denotes portfolio value and* $r(t, x, y)$ *is a deterministic function.*

From the projected S-dynamics

$$dS_t = \hat{y}_t S_t dt + \sigma S_t d\nu_t.$$

and from standard theory we see that the portfolio value dynamics are given by

$$dX_t = u_t X_t(\hat{y}_t - r_t)dt + (r_t X_t - c_t)dt + u_t X_t \sigma d\nu_t.$$

where u is the weight on the risky asset.

The control problem

The object is to maximize the expected utility

$$E^P\left[\int_0^T U(t, c_t)dt\right].$$

over $\mathbf{F}^S$-adapted controls (c, u), given the system

$$\begin{aligned} dX_t &= u_t X_t(\hat{y}_t - r_t)dt + (r_t X_t - c_t)dt + u_t X_t \sigma d\nu_t, \\ d\hat{y}_t &= (A\hat{y}_t + B) + H_t d\nu_t. \end{aligned}$$

and the constraint $c_t \geq 0$.

Main point: This is a **standard problem with full information** so we can apply DynP in a standard manner.

The HJB equation

Denoting the optimal value function by $V(t, x, y)$ we have the following HJB equation.

$$
\begin{cases}
V_t(t, x, y) + \sup_{c,u} \{U(t, c) + \mathbf{A}^{c,u} V(t, x, y)\} & = 0, \\
V(T, x) & = 0,
\end{cases}
$$

where the operator $\mathbf{A}^{c,u}$ is defined as

$$
\begin{aligned}
\mathbf{A}^{c,u} V &= u(y - r)xV_x + (rx - c)V_x + \frac{1}{2}u^2 x^2 \sigma^2 V_{xx} \\
&+ (Ay + B)V_y + \frac{1}{2}H^2 V_{yy} + ux\sigma H V_{xy}.
\end{aligned}
$$

Assuming an interior optimum, we have the first order conditions

$$
\begin{aligned}
U_c' &= V_x, \\
\hat{u} &= \frac{r - y}{\sigma^2} \left(\frac{V_x}{xV_{xx}} \right) - \frac{H}{\sigma} \left(\frac{V_{xy}}{xV_{xx}} \right).
\end{aligned}
$$

8.1.4

Equilibrium

Main result

Since the risk free asset is in zero net supply, the equilibrium condition is $\hat{u} = 1$. Inserting this into the first order condition above we obtain the main result.

Proposition: The risk free rate and the Girsanov kernel φ are given by

$$
\begin{aligned}
r(t,x,y) &= y + \frac{xV_{xx}}{V_x}\sigma^2 + \frac{V_{xy}}{V_x}H\sigma, \\
\varphi(t,x,y) &= \frac{xV_{xx}}{V_x}\sigma + \frac{V_{xy}}{V_x}H.
\end{aligned}
$$

Some comments

It is instructive to compare this result to the result we would have obtained if the factor process Y had been observable. There are similarities as well as differences. At first sight it may seem that the only difference is that Y is replaced by $\hat{y}$, but the situation is in fact a little bit more complicated than that.

Dynamics

- For the fully observable model the (S, Y) dynamics are of the form

$$
\begin{aligned}
dS_t &= Y_t S_t dt + S_t \sigma dW_t^s, \\
dY_t &= (AY_t + B)dt + CdW_t^y,
\end{aligned}
$$

  where W^s and W^y are independent.

- For the partially observable model, the process Y is replaced by the filter estimate $\hat{y}$, and the $(S, \hat{y})$ dynamics are of the form

$$
\begin{aligned}
dS_t &= \hat{y}_t S_t dt + \sigma S_t d\nu_t. \\
d\hat{y}_t &= (A\hat{y}_t + B) + H_t d\nu_t.
\end{aligned}
$$

  Firstly we note that whereas S and Y are driven by independent Wiener processes, S and $\hat{y}$ are driven by **the same** Wiener process, namely the innovation ν. Secondly we note that the diffusion term C in the Y dynamics is replaced by H in the $\hat{y}$ dynamics.

The short rate

The formulas for the short rate in the observable and the partially observable case are given as follows.

$$r(t,x,y) = y + \frac{xV_{xx}}{V_x}\sigma^2,$$
$$r(t,x,y) = y + \frac{xV_{xx}}{V_x}\sigma^2 + \frac{V_{xy}}{V_x}H\sigma.$$

Apart from the fact that y refers to Y in the first formula and to $\hat{y}$ in the second one, there are two differences between these formulas. Firstly, there is no mixed term in the completely observable model. We would perhaps have expected a term of the form

$$\frac{V_{xy}}{V_x}C\sigma$$

but this term vanishes because of the assumed independence between W^s and W^y. Secondly, the function V is not the same in the two formulas. We recall that V is the solution to the HJB equation, and this equation differs slightly between the two models.

8.2

An endowment model

In this section we study a partially observable version of the endowment model from Lecture 5.

8.2.1

The model

The endowment

- We assume the existence of an endowment process e of the form

$$de_t = a_t dt + b_t dW_t.$$

- The observable filtration is given by $\mathbf{F}^e$, i.e. all observations are generated by the endowment process e.

- The process a is **not** assumed to be observable, so it is not adapted to $\mathbf{F}^e$.

- WLOG, the process b is adapted to $\mathbf{F}^e$. (Why?)

- The process b is assumed to satisfy the non-degeneracy condition

$$b_t > 0, \quad P\text{-a.s. for all } t. \tag{11}$$

Traded assets and agents

- We assume that there exists a risky asset S in unit net supply, giving the holder the right to the endowment e.
- There exists a risk free asset in zero net supply.
- The initial wealth of the representative agent is assumed to equal S_0 so the agent can afford to buy the right to the endowment e.
- The representative agent is as usual assumed to maximize utility of the form

$$E\left[\int_0^T U(t, c_t) dt\right].$$

8.2.2

Projecting the e-dynamics

Standard projection argument

We define the process Z by

$$dZ_t = \frac{de_t}{b_t}$$

so that

$$dZ_t = \frac{a_t}{b_t}dt + dW_t,$$

and define the innovation process ν as usual by

$$d\nu_t = dZ_t - \frac{\hat{a}_t}{b_t}dt,$$

where

$$\hat{a}_t = E\left[a_t \middle| \mathcal{F}_t^e\right]$$

This gives us the Z dynamics on the $\mathbf{F}^e$ filtration as

$$dZ_t = \frac{\hat{a}_t}{b_t}dt + d\nu_t,$$

Plugging this into $dZ_t = de_t/b_t$ gives us the e dynamics projected onto the $\mathbf{F}^e$ filtration as

$$de_t = \hat{a}_t dt + b_t d\nu_t.$$

8.2.3

Equilibrium

General result

Given e-dynamics of the form

$$de_t = \hat{a}_t dt + b_t d\nu_t.$$

we are now back in a completely observable model, so we can copy the main result from Lecture 5 to obtain the following

Proposition: *For the partially observed model above, the following hold.*

- *The equilibrium short rate process is given by*

$$r_t = -\frac{U_{ct}(t,e) + \hat{a}_t U_{cc}(t,e_t) + \frac{1}{2}\|b_t\|^2 U_{ccc}(t,e_t)}{U_c(t,e_t)}.$$

- *The Girsanov kernel is given by*

$$\varphi_t = \frac{U_{cc}(t,e_t)}{U_c(t,e_t)} \cdot b_t.$$

8.2.4

A factor model

An abstract factor model with log utility

In this section we specialize the model above to a factor model of the form

$$
\begin{aligned}
de_t &= a(e_t, Y_t)dt + b(e_t)dW_t^e, \\
dY_t &= \mu(Y_t)dt + \sigma(Y_t)dW_t^y,
\end{aligned}
$$

where, for simplicity, we assume that W^e and W^y are independent. Note that we cannot allow b to depend on the factor Y. We also assume log utility, so that

$$U(t, c) = e^{-\delta t} \ln(c).$$

We easily obtain

$$
\begin{aligned}
r_t &= \delta + \frac{\hat{a}_t}{e_t} - \frac{b^2(e_t)}{e_t^2}, \\
\varphi_t &= -\frac{b(e_t)}{e_t}.
\end{aligned}
$$

where

$$\hat{a}_t = E\left[a(e_t, Y_t) \middle| \mathcal{F}_t^e \right].$$

Specializing further

Given the expressions

$$\begin{aligned} r_t &= \delta + \frac{\hat{a}_t}{e_t} - \frac{b^2(e_t)}{e_t^2}, \\ \varphi_t &= -\frac{b(e_t)}{e_t}. \end{aligned}$$

it is natural to specialize to the case when

$$\begin{aligned} a(e, y) &= e \cdot a(y), \\ b(e) &= b \cdot e, \end{aligned}$$

where b is a constant. This gives us

$$\begin{aligned} r_t &= \delta + \hat{a}_t - b^2, \\ \varphi_t &= -b. \end{aligned}$$

In order to obtain a finite filter for $\hat{a} = E\left[a(Y_t) \middle| \mathcal{F}_t^e \right]$ it is now natural to look for a Kalman model and our main result is as follows.

Linear Y-dynamics

Proposition: *Assume log utility and a model of the form*

$$
\begin{aligned}
de_t &= a e_t Y_t dt + b e_t dW_t^e, \\
dY_t &= B Y_t dt + C dW_t^y,
\end{aligned}
$$

The risk free rate and the Girsanov kernel are then given by

$$
\begin{aligned}
r_t &= \delta - b^2 + a\hat{y}_t, \\
\varphi_t &= -b.
\end{aligned}
$$

where $\hat{y}$ is given by the Kalman filter

$$
d\hat{y} = B\hat{y}_t + H_t d\nu_t.
$$