Model Risk in Financial Markets

Prof. Radu Tunaru, CeQuFin

November 2018

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Outline

In finance you need:

- a discount curve,
- a pricing/hedging model,
- a risk management model,
- a set of computational tools,

and lots of common sense...

- Model risk in interest rate modelling
- Evidence of model risk in interest rate literature
- Towards a Bayesian view on finance modelling
- Option pricing under uncertainty
- A new measure of parameter uncertainty model risk
- Pitfalls related to VaR and ES
- Model risk for credit models
- Application to rating transition matrix
- Why MLE can be wrong
- Problems with Implied Volatility
- Problems in Estimation of Jump-Diffusion models

Why Model Risk

- 1987 Merrill Lynch reported losses of 300 million USD on stripped mortgage-backed securities because of an incorrect pricing model
- 1992 J.P. Morgan lost about 200 million USD in the mortgage-backed securities market because of inadequate modelling of prepayments.
- 1997 Bank of Tokyo/Mitsubishi, its New York-subsidiary $83 million loss because of their internal pricing model overvalued a portfolio of swaps and options on USD interest rates.

Why Model Risk

- 1987 Merrill Lynch reported losses of 300 million USD on stripped mortgage-backed securities because of an incorrect pricing model
- 1992 J.P. Morgan lost about 200 million USD in the mortgage-backed securities market because of inadequate modelling of prepayments.
- 1997 Bank of Tokyo/Mitsubishi, its New York-subsidiary $83 million loss because of their internal pricing model overvalued a portfolio of swaps and options on USD interest rates.
  - Dowd (2002) pointed out that the loss was caused by wrongly using a one-factor Black-Derman-Toy (BDT) model to trade swaptions.

Why Model Risk

- 1987 Merrill Lynch reported losses of 300 million USD on stripped mortgage-backed securities because of an incorrect pricing model
- 1992 J.P. Morgan lost about 200 million USD in the mortgage-backed securities market because of inadequate modelling of prepayments.
- 1997 Bank of Tokyo/Mitsubishi, its New York-subsidiary $83 million loss because of their internal pricing model overvalued a portfolio of swaps and options on USD interest rates.
  - Dowd (2002) pointed out that the loss was caused by wrongly using a one-factor Black-Derman-Toy (BDT) model to trade swaptions.
  - The model was calibrated to ATM swaptions but used to trade out-of-the-money (OTM) Bermudan swaptions, which was not appropriate.

Why Model Risk

- 1987 Merrill Lynch reported losses of 300 million USD on stripped mortgage-backed securities because of an incorrect pricing model
- 1992 J.P. Morgan lost about 200 million USD in the mortgage-backed securities market because of inadequate modelling of prepayments.
- 1997 Bank of Tokyo/Mitsubishi, its New York-subsidiary \$83 million loss because of their internal pricing model overvalued a portfolio of swaps and options on USD interest rates.
    - Dowd (2002) pointed out that the loss was caused by wrongly using a one-factor Black-Derman-Toy (BDT) model to trade swaptions.
    - The model was calibrated to ATM swaptions but used to trade out-of-the-money (OTM) Bermudan swaptions, which was not appropriate.

Why Model Risk

- 1987 Merrill Lynch reported losses of 300 million USD on stripped mortgage-backed securities because of an incorrect pricing model
- 1992 J.P. Morgan lost about 200 million USD in the mortgage-backed securities market because of inadequate modelling of prepayments.
- 1997 Bank of Tokyo/Mitsubishi, its New York-subsidiary $83 million loss because of their internal pricing model overvalued a portfolio of swaps and options on USD interest rates.
  - Dowd (2002) pointed out that the loss was caused by wrongly using a one-factor Black-Derman-Toy (BDT) model to trade swaptions.
  - The model was calibrated to ATM swaptions but used to trade out-of-the-money (OTM) Bermudan swaptions, which was not appropriate.

- 1997, NatWest Capital Markets reported a $50 million loss because of a mispriced portfolio of German and U.K. interest rate derivatives on the book of a single derivatives trader in London who fed his own estimates of volatility into a model pricing OTC interest rate options with long maturities.
- Williams (1999) remarked that model risk was not included in standard risk management software and in 1999 about 5 billion USD losses were caused by model risk.
- A Deutsche Bank subsidiary in Japan used some smart models to trade electronically that went wild in June 2010, going into an infinite loop and taking out a $183 billion stock position.

Model risk has also been identified to some extent by the Basel Committee on Banking Supervision in the Basel II framework, see Basel (2006) and Basel (2011). Financial institutions ought to gauge their model risk. Furthermore, model validation is one component of the Pillar 1 Minimum Capital Requirements and Pillar 2 Supervisory Review Process.

- 1997, NatWest Capital Markets reported a $50 million loss because of a mispriced portfolio of German and U.K. interest rate derivatives on the book of a single derivatives trader in London who fed his own estimates of volatility into a model pricing OTC interest rate options with long maturities.
- Williams (1999) remarked that model risk was not included in standard risk management software and in 1999 about 5 billion USD losses were caused by model risk.
- A Deutsche Bank subsidiary in Japan used some smart models to trade electronically that went wild in June 2010, going into an infinite loop and taking out a $183 billion stock position.

Model risk has also been identified to some extent by the Basel Committee on Banking Supervision in the Basel II framework, see Basel (2006) and Basel (2011). Financial institutions ought to gauge their model risk. Furthermore, model validation is one component of the Pillar 1 Minimum Capital Requirements and Pillar 2 Supervisory Review Process.

- 1997, NatWest Capital Markets reported a $50 million loss because of a mispriced portfolio of German and U.K. interest rate derivatives on the book of a single derivatives trader in London who fed his own estimates of volatility into a model pricing OTC interest rate options with long maturities.
- Williams (1999) remarked that model risk was not included in standard risk management software and in 1999 about 5 billion USD losses were caused by model risk.
- A Deutsche Bank subsidiary in Japan used some smart models to trade electronically that went wild in June 2010, going into an infinite loop and taking out a $183 billion stock position.

Model risk has also been identified to some extent by the Basel Committee on Banking Supervision in the Basel II framework, see Basel (2006) and Basel (2011). Financial institutions ought to gauge their model risk. Furthermore, model validation is one component of the Pillar 1 Minimum Capital Requirements and Pillar 2 Supervisory Review Process.

- 1997, NatWest Capital Markets reported a $50 million loss because of a mispriced portfolio of German and U.K. interest rate derivatives on the book of a single derivatives trader in London who fed his own estimates of volatility into a model pricing OTC interest rate options with long maturities.
- Williams (1999) remarked that model risk was not included in standard risk management software and in 1999 about 5 billion USD losses were caused by model risk.
- A Deutsche Bank subsidiary in Japan used some smart models to trade electronically that went wild in June 2010, going into an infinite loop and taking out a $183 billion stock position.

Model risk has also been identified to some extent by the Basel Committee on Banking Supervision in the Basel II framework, see Basel (2006) and Basel (2011). Financial institutions ought to gauge their model risk. Furthermore, model validation is one component of the Pillar 1 Minimum Capital Requirements and Pillar 2 Supervisory Review Process.

- 1997, NatWest Capital Markets reported a $50 million loss because of a mispriced portfolio of German and U.K. interest rate derivatives on the book of a single derivatives trader in London who fed his own estimates of volatility into a model pricing OTC interest rate options with long maturities.
- Williams (1999) remarked that model risk was not included in standard risk management software and in 1999 about 5 billion USD losses were caused by model risk.
- A Deutsche Bank subsidiary in Japan used some smart models to trade electronically that went wild in June 2010, going into an infinite loop and taking out a $183 billion stock position.

Model risk has also been identified to some extent by the Basel Committee on Banking Supervision in the Basel II framework, see Basel (2006) and Basel (2011). Financial institutions ought to gauge their model risk. Furthermore, model validation is one component of the Pillar 1 Minimum Capital Requirements and Pillar 2 Supervisory Review Process.

A Definition of Model Risk

Gibson et al. (1999) state

> *"Model risk results from the inappropriate specification of a theoretical model or the use of an appropriate model but in an inadequate framework or for the wrong purpose."*

while for McNeil et al. (2005) model risk can be defined as

> *"the risk that a financial institution incurs losses because its risk-management models are misspecified or because some of the assumptions underlying these models are not met in practice."*

For Barrieu and Scandolo (2013)

> *"The hazard of working with a potentially not well-suited model is referred to as model risk"*

and Boucher et al. (2014) define model risk as

> *"the uncertainty in risk forecasting arising from estimation error and the use of an incorrect model".*

What model risk is not

- It is not operational risk: Example: *The Vancouver stock exchange started a new index initialized at the level of 1000.000 in 1982. However, less than two years later it was observed that the index was constantly decreasing to about 520 despite the exchange setting records in value and volume as described in the Wall Street Journal in 1983. Upon further investigations it was revealed that the index, which was updated after every transaction, was recalculated by removing the decimals after the third decimal instead of rounding off. Hence, the correct value of 1098.892 became the published value of 520*.
- fiscal-legal updating. Sudden changes in law may expose a bank to great losses. Example: *In the UK a law on lower dividend tax credit was exploited by UBS in the 1990s. The law was changed in 1997 and caused many banks to suffer immediate losses with UBS incurring huge losses.* In general, see Gibson (2000), models used by banks simply ignore the impact of sudden fiscal change.

What model risk is not

- It is not operational risk: Example: *The Vancouver stock exchange started a new index initialized at the level of 1000.000 in 1982. However, less than two years later it was observed that the index was constantly decreasing to about 520 despite the exchange setting records in value and volume as described in the Wall Street Journal in 1983. Upon further investigations it was revealed that the index, which was updated after every transaction, was recalculated by removing the decimals after the third decimal instead of rounding off. Hence, the correct value of 1098.892 became the published value of 520*.
- fiscal-legal updating. Sudden changes in law may expose a bank to great losses. Example: *In the UK a law on lower dividend tax credit was exploited by UBS in the 1990s. The law was changed in 1997 and caused many banks to suffer immediate losses with UBS incurring huge losses*. In general, see Gibson (2000), models used by banks simply ignore the impact of sudden fiscal change.

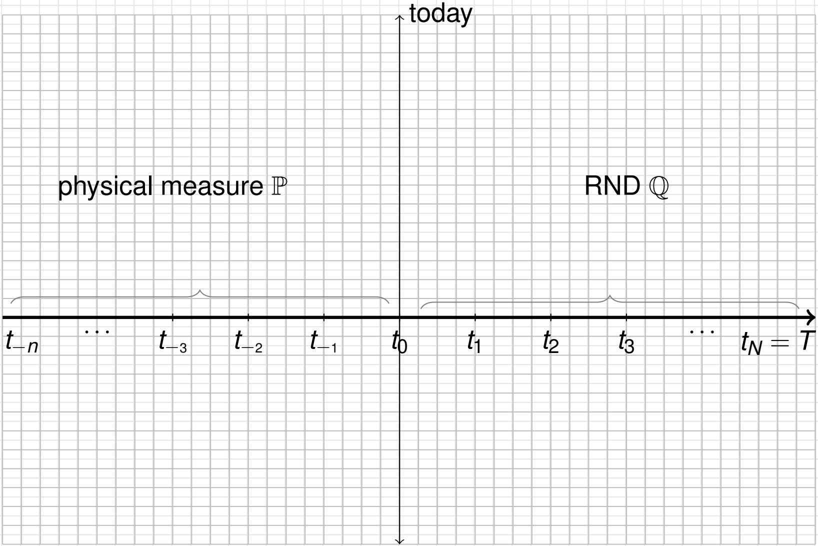


Figure: Two worlds of Finance through a mathematical eye

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:

Derman (1996) three main reasons for model risk:

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
   1. the true model belongs to a known class of models;

Derman (1996) three main reasons for model risk:

- it may be misspecified;

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
   1. the true model belongs to a known class of models;
   2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
   3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

1. the model parameters may not be estimated correctly;
2. it may be mis-specified;
3. it may be incorrectly implemented.

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

1. the model parameters may not be estimated correctly;
2. it may be mis-specified;
3. it may be incorrectly implemented.

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

1. the model parameters may not be estimated correctly;
2. it may be mis-specified;
3. it may be incorrectly implemented.

Bernardo and Smith (1994) and Draper (1995) enumerate the main sources of uncertainty related to the quantity under study:

1. uncertainty caused by the stochastic specification of the model; deterministic models as used in mechanics for example do not carry any degree of uncertainty.
2. uncertainty in the estimated values of the parameters underpinning the model; from a finite set of data we may not be sure about the true population value of the parameters.
3. uncertainty in the model used; it is difficult to know with certainty that a given model is the correct one. This category can be classified further:
    1. the true model belongs to a known class of models;
    2. the class of models used are known to be approximations to a more complex model that is cumbersome to work with;
    3. the class of models may provide a proxy for a more complex true model about which the modeller has no prior knowledge.

Derman (1996) three main reasons for model risk:

1. the model parameters may not be estimated correctly;
2. it may be mis-specified;
3. it may be incorrectly implemented.

1. *parameter estimation risk*
2. *model selection risk* within a given family of models
3. *model identification risk*. This category is more related to Knightian uncertainty than to model risk per se.
4. *computational implementation risk* which is generated by overlooking technical conditions under which particular computational mathematical techniques work.
5. *model protocol risk*.

1. *parameter estimation risk*
2. *model selection risk* within a given family of models
3. *model identification risk*. This category is more related to Knightian uncertainty than to model risk per se.
4. *computational implementation risk* which is generated by overlooking technical conditions under which particular computational mathematical techniques work.
5. *model protocol risk*.

1. *parameter estimation risk*
2. *model selection risk* within a given family of models
3. *model identification risk*. This category is more related to Knightian uncertainty than to model risk per se.
4. *computational implementation risk* which is generated by overlooking technical conditions under which particular computational mathematical techniques work.
5. *model protocol risk*.

1. *parameter estimation risk*
2. *model selection risk* within a given family of models
3. *model identification risk*. This category is more related to Knightian uncertainty than to model risk per se.
4. *computational implementation risk* which is generated by overlooking technical conditions under which particular computational mathematical techniques work.
5. *model protocol risk*.

1. *parameter estimation risk*
2. *model selection risk* within a given family of models
3. *model identification risk*. This category is more related to Knightian uncertainty than to model risk per se.
4. *computational implementation risk* which is generated by overlooking technical conditions under which particular computational mathematical techniques work.
5. *model protocol risk*.

Short rate Models I

- Vasicek (1977) developed a model for the risk-free rate of interest $\{r_t\})_{t\geq 0}$, given by the following continuous-time SDE

$$dr_t = k(b - r_t)dt + \sigma dW_t \quad (1)$$

where $k, b, \sigma > 0$.

- b is interpreted as the long-run mean rate; $\lim_{t\to\infty} \mathbb{E}[r_t] = b$, k represents the speed of mean reversion to b and σ is the *local* volatility parameter.
-

$$\mathbb{E}[r_{t+u}|r_t] = b + (r_t - b)e^{-ku}, \quad \mathrm{var}[r_{t+u}|r_t] = \sigma^2 \frac{1 - e^{-2ku}}{2k}$$

- the long-term standard deviation of r_t is $\frac{\sigma}{\sqrt{2k}}$.

Short rate Models II

- zero-coupon bond prices at time t for maturity T

$$p(t,T) = \exp[A(t,T) - B(t,T)r_t]$$

where $B(t,T) = \frac{1-e^{-k(T-t)}}{k}$ and

$$A(t,T) = \left[B(t,T) - (T-t)\left(b - \frac{\sigma^2}{2k^2}\right)\right] - \frac{\sigma^2}{4k}B(t,T)^2.$$

- under the Vasicek model the price of European call is

$$Call_t = p(t,T^*)\Phi(d_1) - Kp(t,T)\Phi(d_2) \tag{2}$$

where

$$d_1 = \frac{1}{\sigma^*}\ln\left(\frac{p(t,T^*)}{Kp(t,T)}\right) + \frac{\sigma^*}{2}, \quad d_2 = d_1 - \sigma^*$$

and where we denote by $\sigma^* = \frac{\sigma}{k}[1 - e^{-k(T^*-T)}]\sqrt{\frac{1-e^{-2k(T-t)}}{2k}}$

Short rate Models III

- One major "shortcoming" of the Vasicek model is that the rate r_t can become negative and depending on the parameters' values, with quite significant probabilities.
- The Cox, Ingersoll and Ross (CIR) model given by

$$dr_t = k(b - r_t)dt + \sigma\sqrt{r_t}dW_t \tag{3}$$

  where $k, b, \sigma > 0$.
- the zero-coupon bond prices

$$p(t,T) = \exp[a(T-t) - b(T-t)r_t] \tag{4}$$

  where $b(u) = \frac{2(e^{\gamma u}-1)}{(\gamma+k)(e^{\gamma u}-1)+2\gamma}$, $\gamma = \sqrt{k^2+2\sigma^2}$ and $a(u) = \frac{2kb}{\sigma^2}\ln\left[\frac{2\gamma e^{(\gamma+k)u/2}}{(\gamma+k)(e^{\gamma u}-1)+2\gamma}\right]$.
- If $q = \frac{\sigma^2(1-e^{-kT})}{4k}$ it can be proved that conditional on r_0 the distribution of $\frac{r_t}{q}$ is a non-central chi-squared distribution with $d = \frac{4kb}{\sigma^2}$ degrees of freedom and non-centrality parameter $\alpha = \frac{4kr_0}{\sigma^2(e^{kT}-1)}$.

Short rate Models IV

- the European call options on zero-coupon bonds with maturity T^*, exercise date T and strike price K

$$Call_t = p(0, T^*)\chi^2(d, \alpha_1; v_1) - Kp(0, T)\chi^2(d, \alpha_2; v_2) \tag{5}$$

where $\chi^2(d, \alpha; v)$ is the cumulative distribution function of the non-central chi-squared distribution with d degrees of freedom and non-centrality parameter α, and $d = \frac{4kb}{\sigma^2}$, $\gamma = \sqrt{k^2 + 2\sigma^2}$

Short rate Models V

- $$\begin{aligned}
\alpha_1 &= \frac{8\gamma^2 e^{\gamma T} r_t}{\sigma^2(e^{\gamma T}-1)(2\gamma+(\gamma+k+\sigma^2 b(T^*-T))(e^{\gamma T}-1))}\\
\alpha_2 &= \frac{8\gamma^2 e^{\gamma T} r_t}{\sigma^2(e^{\gamma T}-1)(2\gamma+(\gamma+k)(e^{\gamma T}-1))}\\
\delta &= \frac{a(T^*-t)-\ln K}{b(T^*-T)}\\
v_1 &= \frac{2\delta[2\gamma+(\gamma+k+\sigma^2 b(T^*-T))(e^{\gamma T}-1)]}{\sigma^2(e^{\gamma T}-1)}\\
v_2 &= \frac{2\delta[2\gamma+(\gamma+k)(e^{\gamma T}-1)]}{\sigma^2(e^{\gamma T}-1)}
\end{aligned}$$
- if $r_0 > 0$ and $2kb \geq \sigma^2$ then $r_t > 0$ almost surely.
- when $2kb < \sigma^2$ then there is a time t such that $r_t \leq 0$ almost surely.

Short rate Models VI

- looking at the zero-coupon bond prices

$$p(T, T^*) < \exp\{a(T^* - T)\}$$

  and if it also happens that $a(T^* - T) < 0$ at the maturity T of the option then, for exercise price K close enough to 1,

$$p(T, T^*) < \exp\{a(T^* - T)\} < K < 1$$

  which is impossible and will automatically give a call option price equal to zero!
- The Vasicek model does not behave well in the proximity of exercise price 1 either since it will provide a positive call option price for $K = 1$. This option price inflation is caused by the fact that r_t can be occasionally negative under the Vasicek model.

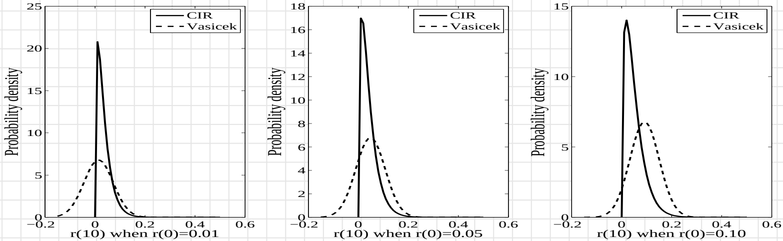


Figure: Comparison of probability density functions under the Vasicek and CIR models for the value rate r at $T = 10$.

the probability densities of r under each model at some horizon $T = 10$, the parameters of the two models calibrated over the same set of data. This point and an ad-hoc solution to get equivalent sets of parameters for the two short rate models has been discussed in Cairns (2004).

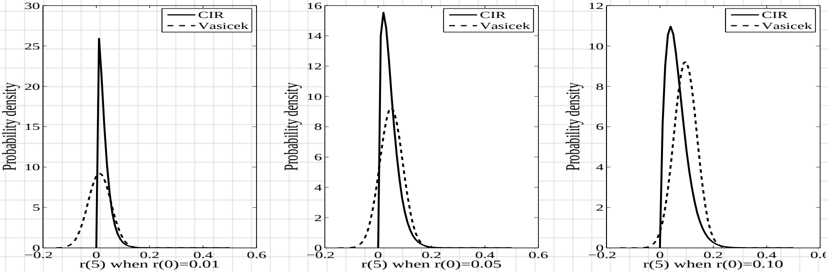


Figure: Comparison of probability density functions under the Vasicek and CIR models for the value rate r at $T = 5$.

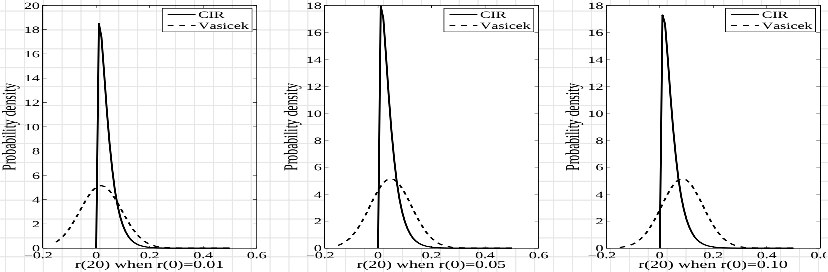


Figure: Comparison of probability density functions under the Vasicek and CIR models for the value rate r at $T = 20$.

Path Simulation

- for Vasicek is
$$\Delta r_t = k(b - r_t)\Delta t + \sigma \varepsilon_t \sqrt{\Delta t} \tag{6}$$
- for the CIR model the corresponding equation is
$$\Delta r_t = k(b - r_t)\Delta t + \sigma \sqrt{r_t} \varepsilon_t \sqrt{\Delta t} \tag{7}$$
where $\varepsilon_t \sim N(0,1)$ for all t.
- $r_0 = 3\%$, $T = 1$, $\Delta t = 0.004$, $k = 0.07$, $b = 2.50\%$ and $\sigma = 2.25\%$
- simulate paths from both data generating processes.

Standard Paths Non-Standard Paths

- Buraschi and Corielli (2005) describe the time inconsistencies that may appear when using models from the HJM family. This is a very important question for the risk manager. Is the model selected by a bank or financial institution complex enough to generate curves that cover the observed or realised term structure curves from the past? On the other hand, is the model too complex and is generating curves that have never been observed in practice?
- Filipovic (2009) discusses the nonexistence of HJM models with proportional volatility which apparently was one of the major reasons for the introduction of LIBOR market models in fixed income markets.
- Jarrow (2009), multi-factor models with more than three factors are actually required in practice, particularly when exotic interest rate derivatives are traded.

- Buraschi and Corielli (2005) describe the time inconsistencies that may appear when using models from the HJM family. This is a very important question for the risk manager. Is the model selected by a bank or financial institution complex enough to generate curves that cover the observed or realised term structure curves from the past? On the other hand, is the model too complex and is generating curves that have never been observed in practice?
- Filipovic (2009) discusses the nonexistence of HJM models with proportional volatility which apparently was one of the major reasons for the introduction of LIBOR market models in fixed income markets.
- Jarrow (2009), multi-factor models with more than three factors are actually required in practice, particularly when exotic interest rate derivatives are traded.

- Buraschi and Corielli (2005) describe the time inconsistencies that may appear when using models from the HJM family. This is a very important question for the risk manager. Is the model selected by a bank or financial institution complex enough to generate curves that cover the observed or realised term structure curves from the past? On the other hand, is the model too complex and is generating curves that have never been observed in practice?
- Filipovic (2009) discusses the nonexistence of HJM models with proportional volatility which apparently was one of the major reasons for the introduction of LIBOR market models in fixed income markets.
- Jarrow (2009), multi-factor models with more than three factors are actually required in practice, particularly when exotic interest rate derivatives are traded.

Proposition (Nelson-Siegel vs Hull-White)

The Hull-White extended Vasicek model is inconsistent with the Nelson-Siegel family of forward curves.

- Hence the NS manifold is not large enough for the HW model. If the initial forward rate curve is on the manifold, then the HW dynamics will force the term structure off the manifold within an arbitrarily short period of time!

Proposition (Nelson-Siegel vs Ho-Lee)

The full Nelson-Siegel family is inconsistent with the Ho-Lee model. The degenerate family $G(z; x) = z_1 + z_3 x$ is in fact consistent with Ho-Lee.

Filipovic (1998) proved the following important result.

Proposition

There is no non-trivial Wiener driven model that is consistent with the Nelson-Siegel family of forward curves.

Option Pricing with Uncertainty

For option pricing the most important quantity driving the prices is the volatility, various classes of models emerged for the estimation of the same quantity, volatility σ. Bunnin et al. (2002) classified them:

1. Implied volatility. A pointwise estimate of σ is derived as an inverse problem; the option price is given and the Black-Scholes formula is used to retrieve the $\widehat{\sigma}$ that makes the formula match the market option price.
2. Discrete time GARCH models. GARCH models were primarily developed for the evolution of variance but calculating the value of σ is straightforward.
3. Frequentist econometric pointwise estimation of σ. One can use the historical series and some error specification obtained after discretizing a continuous-time model, and then use OLS, or MLE or GMM to estimate $\widehat{\sigma}$.
4. Bayesian pointwise estimation. One could use solely the pointwise posterior estimators such as posterior mean or posterior median.

Option Pricing with Uncertainty

For option pricing the most important quantity driving the prices is the volatility, various classes of models emerged for the estimation of the same quantity, volatility σ. Bunnin et al. (2002) classified them:

1. Implied volatility. A pointwise estimate of σ is derived as an inverse problem; the option price is given and the Black-Scholes formula is used to retrieve the $\widehat{\sigma}$ that makes the formula match the market option price.
2. Discrete time GARCH models. GARCH models were primarily developed for the evolution of variance but calculating the value of σ is straightforward.
3. Frequentist econometric pointwise estimation of σ. One can use the historical series and some error specification obtained after discretizing a continuous-time model, and then use OLS, or MLE or GMM to estimate $\widehat{\sigma}$.
4. Bayesian pointwise estimation. One could use solely the pointwise posterior estimators such as posterior mean or posterior median.

Option Pricing with Uncertainty

For option pricing the most important quantity driving the prices is the volatility, various classes of models emerged for the estimation of the same quantity, volatility σ. Bunnin et al. (2002) classified them:

1. Implied volatility. A pointwise estimate of σ is derived as an inverse problem; the option price is given and the Black-Scholes formula is used to retrieve the $\widehat{\sigma}$ that makes the formula match the market option price.
2. Discrete time GARCH models. GARCH models were primarily developed for the evolution of variance but calculating the value of σ is straightforward.
3. Frequentist econometric pointwise estimation of σ. One can use the historical series and some error specification obtained after discretizing a continuous-time model, and then use OLS, or MLE or GMM to estimate $\widehat{\sigma}$.
4. Bayesian pointwise estimation. One could use solely the pointwise posterior estimators such as posterior mean or posterior median.

Option Pricing with Uncertainty

For option pricing the most important quantity driving the prices is the volatility, various classes of models emerged for the estimation of the same quantity, volatility σ. Bunnin et al. (2002) classified them:

1. Implied volatility. A pointwise estimate of σ is derived as an inverse problem; the option price is given and the Black-Scholes formula is used to retrieve the $\widehat{\sigma}$ that makes the formula match the market option price.
2. Discrete time GARCH models. GARCH models were primarily developed for the evolution of variance but calculating the value of σ is straightforward.
3. Frequentist econometric pointwise estimation of σ. One can use the historical series and some error specification obtained after discretizing a continuous-time model, and then use OLS, or MLE or GMM to estimate $\widehat{\sigma}$.
4. Bayesian pointwise estimation. One could use solely the pointwise posterior estimators such as posterior mean or posterior median.

1. An important step in the evolution of modelling for financial markets was marked by the introduction of stochastic volatility models, in discrete time and continuous time.
2. Semi-parametric models. Not that many are available but they allow a very high degree of uncertainty since σ is constrained to a finite interval but no other specification of volatility is made. One important paper in this class is Avellaneda et al. (1995).
3. Volatility surfaces. There is great research in this area recognizing that at one time the option maturity spectrum is defined by a term structure. Combine that with assets requiring modelling of a term structure of prices, such as bonds, and the cross combination leads to a volatility surface that needs to be estimated.

the pricing equation should be rewritten as

$$u(S_t, t) = B(t, T) \int_0^\infty \psi(S_T) p(S_T, T | S_t, t, \theta) dS_T \quad (8)$$

making explicit the conditioning on θ.

1. An important step in the evolution of modelling for financial markets was marked by the introduction of stochastic volatility models, in discrete time and continuous time.
2. Semi-parametric models. Not that many are available but they allow a very high degree of uncertainty since σ is constrained to a finite interval but no other specification of volatility is made. One important paper in this class is Avellaneda et al. (1995).
3. Volatility surfaces. There is great research in this area recognizing that at one time the option maturity spectrum is defined by a term structure. Combine that with assets requiring modelling of a term structure of prices, such as bonds, and the cross combination leads to a volatility surface that needs to be estimated.

the pricing equation should be rewritten as

$$u(S_t, t) = B(t, T) \int_0^\infty \psi(S_T) p(S_T, T | S_t, t, \theta) dS_T \tag{8}$$

making explicit the conditioning on θ.

1. An important step in the evolution of modelling for financial markets was marked by the introduction of stochastic volatility models, in discrete time and continuous time.
2. Semi-parametric models. Not that many are available but they allow a very high degree of uncertainty since σ is constrained to a finite interval but no other specification of volatility is made. One important paper in this class is Avellaneda et al. (1995).
3. Volatility surfaces. There is great research in this area recognizing that at one time the option maturity spectrum is defined by a term structure. Combine that with assets requiring modelling of a term structure of prices, such as bonds, and the cross combination leads to a volatility surface that needs to be estimated.

the pricing equation should be rewritten as

$$u(S_t, t) = B(t, T) \int_0^\infty \psi(S_T) p(S_T, T | S_t, t, \theta) dS_T \tag{8}$$

making explicit the conditioning on θ.

1. An important step in the evolution of modelling for financial markets was marked by the introduction of stochastic volatility models, in discrete time and continuous time.
2. Semi-parametric models. Not that many are available but they allow a very high degree of uncertainty since σ is constrained to a finite interval but no other specification of volatility is made. One important paper in this class is Avellaneda et al. (1995).
3. Volatility surfaces. There is great research in this area recognizing that at one time the option maturity spectrum is defined by a term structure. Combine that with assets requiring modelling of a term structure of prices, such as bonds, and the cross combination leads to a volatility surface that needs to be estimated.

the pricing equation should be rewritten as

$$u(S_t, t) = B(t, T) \int_0^\infty \psi(S_T) p(S_T, T | S_t, t, \theta) dS_T \tag{8}$$

making explicit the conditioning on θ.

1. An important step in the evolution of modelling for financial markets was marked by the introduction of stochastic volatility models, in discrete time and continuous time.
2. Semi-parametric models. Not that many are available but they allow a very high degree of uncertainty since σ is constrained to a finite interval but no other specification of volatility is made. One important paper in this class is Avellaneda et al. (1995).
3. Volatility surfaces. There is great research in this area recognizing that at one time the option maturity spectrum is defined by a term structure. Combine that with assets requiring modelling of a term structure of prices, such as bonds, and the cross combination leads to a volatility surface that needs to be estimated.

the pricing equation should be rewritten as

$$u(S_t,t) = B(t,T)\int_0^{\infty} \psi(S_T)p(S_T,T|S_t,t,\theta)dS_T \tag{8}$$

making explicit the conditioning on θ.

The transition probability density for the case when parameters “are known” should be replaced with the *predictive density*

$$p(S_T, T|S_t, t) = \int p(S_T, T|S_t, t, \theta)p(\theta|Y_t)d\theta \tag{9}$$

where $p(\theta|Y_t)$ is the posterior density of θ given the occurrence of Y_t, which is all observed data such as returns or changes or level prices of S, up to time t. Then (8) can be rewritten as

$$u(S_t, t) = B(t, T)\int_0^\infty \psi(S_T)\left[\int_\Theta p(S_T, T|S_t, t, \theta)p(\theta|Y_t)d\theta\right]dS_T \tag{10}$$

The transition probability density for the case when parameters “are known” should be replaced with the *predictive density*

$$p(S_T, T|S_t, t) = \int p(S_T, T|S_t, t, \theta)p(\theta|Y_t)d\theta \tag{9}$$

where $p(\theta|Y_t)$ is the posterior density of θ given the occurrence of Y_t, which is all observed data such as returns or changes or level prices of S, up to time t.

Then (8) can be rewritten as

$$u(S_t, t) = B(t, T)\int_0^{\infty} \psi(S_T)\left[\int_{\Theta} p(S_T, T|S_t, t, \theta)p(\theta|Y_t)d\theta\right]dS_T \tag{10}$$

The transition probability density for the case when parameters “are known” should be replaced with the *predictive density*

$$p(S_T, T|S_t, t) = \int p(S_T, T|S_t, t, \theta)p(\theta|Y_t)d\theta \tag{9}$$

where $p(\theta|Y_t)$ is the posterior density of θ given the occurrence of Y_t, which is all observed data such as returns or changes or level prices of S, up to time t. Then (8) can be rewritten as

$$u(S_t, t) = B(t, T)\int_0^\infty \psi(S_T)\left[\int_\Theta p(S_T, T|S_t, t, \theta)p(\theta|Y_t)d\theta\right]dS_T \tag{10}$$

- For some options payoffs it may be possible to derive closed-form solutions, for a given value of parameter θ.
- denoting by $u^*(S_t, t, \theta) = B(t, T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)|\theta])$

$$
\begin{aligned}
u(S_t, t) &= B(t, T)\int_{\Theta}\left[\int_0^{\infty}\psi(S_T)p(S_T, T|St, t\theta)dS_T\right]p(\theta|Y_t)d\theta \\
&= \int_{\Theta} B(t, T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)|\theta]p(\theta|Y_t)d\theta \\
&= \int_{\Theta} u^*(S_t, t, \theta)p(\theta|Y_t)d\theta \\
&\approx \frac{1}{M}\sum_{i=1}^{M} u^*(S_t, t, \theta_i)
\end{aligned}
$$

where the last approximation formula is calculated by drawing parameter values for their posterior distribution

$$
\theta_i \sim p(\theta|Y_t). \tag{11}
$$

- For some options payoffs it may be possible to derive closed-form solutions, for a given value of parameter θ.
- denoting by $u^*(S_t, t, \theta) = B(t, T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)|\theta])$

$$
\begin{aligned}
u(S_t, t) &= B(t, T) \int_{\Theta} \left[\int_0^{\infty} \psi(S_T) p(S_T, T | St, t\theta) dS_T \right] p(\theta | Y_t) d\theta \\
&= \int_{\Theta} B(t, T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)|\theta] p(\theta | Y_t) d\theta \\
&= \int_{\Theta} u^*(S_t, t, \theta) p(\theta | Y_t) d\theta \\
&\approx \frac{1}{M} \sum_{i=1}^{M} u^*(S_t, t, \theta_i)
\end{aligned}
$$

where the last approximation formula is calculated by drawing parameter values for their posterior distribution

$$\theta_i \sim p(\theta | Y_t). \tag{11}$$

- generate samples of possible parameter values from the posterior distribution of the parameter θ given the observed data, $p(\theta|Y_t)$.
- assume that $Y_t = \{S_t, S_{t-1}, \ldots, S_0\}, \forall t \geq 0$ and denote by $Y_{[s,t]} = Y_t \setminus Y_s$ for any $s < t$.
- Since Itô diffusions are Markov processes, applying Bayes' formula gives

$$p(\theta|Y_t) = \frac{p(Y_{[s,t]}|\theta, S_s)p(\theta|Y_s)}{\int_\Theta p(Y_{[s,t]}|\theta, S_s)p(\theta|Y_s)} \tag{12}$$

Bunnin et al. (2002) present two distinct algorithms for sampling from $p(\theta|Y_t)$, which is needed in order to compute the sample option prices.

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t)$, $i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight

  $$w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \quad i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight

  $$w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

  1. [illegible]
  2. [illegible]
  3. [illegible]

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \;\; i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight

  $$w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \;\; i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight

  $$w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

  1. Split the interval $(0, 1]$ into n subintervals $(a_i, b_i]$, where the end of values are $a_i = \sum_{j=1}^{j=i-1} w_j$ and $b_i = \sum_{j=1}^{i} w_j$.
  2. Draw M i.i.d. $Uniform(0,1)$ random numbers
  3. If $U_k \in (a_i, b_i]$ then θ_i becomes the k-th sample value

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \quad i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight
  $$w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

  1. Split the interval $(0, 1]$ into n subintervals $(a_i, b_i]$, where the end of values are $a_i = \sum_{j=1}^{j=i-1} w_j$ and $b_i = \sum_{j=1}^{i} w_j$.
  2. Draw M i.i.d. *Uniform*$(0, 1)$ random numbers $\{U_k\}_{k\in\{1,\ldots,M\}}$
  3. If $U_k \in (a_i, b_i]$ then θ_i becomes the k-th sample value.

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \quad i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight $w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

  1. Split the interval $(0, 1]$ into n subintervals $(a_i, b_i]$, where the end of values are $a_i = \sum_{j=1}^{j=i-1} w_j$ and $b_i = \sum_{j=1}^{i} w_j$.
  2. Draw M i.i.d. $Uniform(0, 1)$ random numbers $\{U_k\}_{k\in\{1,\ldots,M\}}$
  3. If $U_k \in (a_i, b_i]$ then θ_i becomes the k-th sample value.

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \quad i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight $w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

    1. Split the interval $(0, 1]$ into n subintervals $(a_i, b_i]$, where the end of values are $a_i = \sum_{j=1}^{j=i-1} w_j$ and $b_i = \sum_{j=1}^{i} w_j$.
    2. Draw M i.i.d. $Uniform(0, 1)$ random numbers $\{U_k\}_{k \in \{1,\ldots,M\}}$
    3. If $U_k \in (a_i, b_i]$ then θ_i becomes the k-th sample value.

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \quad i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight $w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

  1. Split the interval $(0, 1]$ into n subintervals $(a_i, b_i]$, where the end of values are $a_i = \sum_{j=1}^{j=i-1} w_j$ and $b_i = \sum_{j=1}^{i} w_j$.
  2. Draw M i.i.d. $Uniform(0, 1)$ random numbers $\{U_k\}_{k\in\{1,\ldots,M\}}$
  3. If $U_k \in (a_i, b_i]$ then θ_i becomes the k-th sample value.

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \quad i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight $w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

  1. Split the interval $(0,1]$ into n subintervals $(a_i, b_i]$, where the end of values are $a_i = \sum_{j=1}^{j=i-1} w_j$ and $b_i = \sum_{j=1}^{i} w_j$.
  2. Draw M i.i.d. $Uniform(0,1)$ random numbers $\{U_k\}_{k\in\{1,\ldots,M\}}$
  3. If $U_k \in (a_i, b_i]$ then θ_i becomes the k-th sample value.

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0,T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- knowing how to simulate from $p(Y_t|\theta)$, we need only to be able to draw samples from $p(\theta|Y_t)$ in order to calculate the predictive density.
- In order to avoid the calculation of the prior density, Bunnin et al. (2002) suggest applying for this important step the sampling importance resampling (SIR) algorithm that will go through the following procedure

  Step 1 Sample $\theta_i \sim p(\theta|Y_t), \quad i = 1, \ldots, n$

  Step 2 Given new data $Y_{[s,t]}$, calculate $\sum_i p(Y_{[s,t]}|\theta_i, S_s)$.

  Step 3 For each θ_i calculate the importance weight
  $w_i = \frac{p(Y_{[s,t]}|\theta_i, S_s)}{\sum_i p(Y_{[s,t]}|\theta_i, S_s)}$

  Step 4 Resample from the θ_i obtained, using the importance weighting given by w_i. This will result in $M < n$ samples from $p(\theta|Y_t)$. The resampling is done through the following subroutine:

  1. Split the interval $(0, 1]$ into n subintervals $(a_i, b_i]$, where the end of values are $a_i = \sum_{j=1}^{j=i-1} w_j$ and $b_i = \sum_{j=1}^{i} w_j$.
  2. Draw M i.i.d. $Uniform(0, 1)$ random numbers $\{U_k\}_{k\in\{1,\ldots,M\}}$
  3. If $U_k \in (a_i, b_i]$ then θ_i becomes the k-th sample value.

- option price sampling from the predictive density of S_T.
- when the SDE of the Itô diffusion has a closed form solution $W_T \sim N(0, T)$ and $S_T^{(i)} = g(W_T, S_0, \theta_i)$

- When the SDE does not have a closed form solution, the Euler-Maruyama discretization seems to be the only feasible route. The procedure is the following.
    - First discretize the SDE.

$$S_{t_{j+1}} - S_{t_j} = a(S_{t_j}, t_j)\Delta t + b(S_{t_j}, t_j)(W_{t_{j+1}} - W_{t_j})$$

where $t_1 = t, t_m = T$, so the simulation will be pathwise of length m between current valuation time t and maturity T.

- When the SDE does not have a closed form solution, the Euler-Maruyama discretization seems to be the only feasible route. The procedure is the following.
  - First discretize the SDE.

$$S_{t_{j+1}} - S_{t_j} = a(S_{t_j}, t_j)\Delta t + b(S_{t_j}, t_j)(W_{t_{j+1}} - W_{t_j})$$

    where $t_1 = t, t_m = T$, so the simulation will be pathwise of length m between current valuation time t and maturity T.
  - Then
    1. Take a sample of size M with $\theta_i \sim p(\theta|Y_t)$.
    2. Generate m standard Gaussian random draws that will help create the path to maturity;
    3. For each θ_i, generate an entire path of S values, leading to the final one, $S_T^{(i)}$ which will be drawn from the predictive density $p(S_T, T|S_t, t)$.
  - Then, the value of the option under parameter estimation uncertainty is given by

$$\begin{aligned} u(S_t, t) &= B(t,T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)] \\ &= B(t,T)\int_0^{\infty} \psi(S_T)p(S_T, T|S_t, t)dS_T \\ &\approx B(t,T)\frac{1}{M}\sum_{i=1}^{M}\psi(S_T^{(i)}) \end{aligned}$$

    where the last relationship reflects the Monte Carlo integration.

- When the SDE does not have a closed form solution, the Euler-Maruyama discretization seems to be the only feasible route. The procedure is the following.
  - First discretize the SDE.

$$S_{t_{j+1}} - S_{t_j} = a(S_{t_j}, t_j)\Delta t + b(S_{t_j}, t_j)(W_{t_{j+1}} - W_{t_j})$$

    where $t_1 = t, t_m = T$, so the simulation will be pathwise of length m between current valuation time t and maturity T.
  - Then
    1. Take a sample of size M with $\theta_i \sim p(\theta | Y_t)$,
    2. Generate m standard Gaussian random draws that will help create the path to maturity,
    3. For each θ_i, generate an entire path of S values, leading to the final one, $S_T^{(i)}$ which will be drawn from the predictive density $p(S_T, T | S_t, t)$.
  - Then, the value of the option under parameter estimation uncertainty, is given by

$$\begin{aligned} u(S_t, t) &= B(t,T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)] \\ &= B(t,T)\int_0^{\infty} \psi(S_T) p(S_T, T | S_t, t) dS_T \\ &\approx B(t,T)\frac{1}{M}\sum_{i=1}^{M} \psi(S_T^{(i)}) \end{aligned}$$

    where the last relationship reflects the Monte Carlo integration.

- When the SDE does not have a closed form solution, the Euler-Maruyama discretization seems to be the only feasible route. The procedure is the following.
  - First discretize the SDE.

$$S_{t_{j+1}} - S_{t_j} = a(S_{t_j}, t_j)\Delta t + b(S_{t_j}, t_j)(W_{t_{j+1}} - W_{t_j})$$

    where $t_1 = t, t_m = T$, so the simulation will be pathwise of length m between current valuation time t and maturity T.
  - Then
    1. Take a sample of size M with $\theta_i \sim p(\theta | Y_t)$,
    2. Generate m standard Gaussian random draws that will help create the path to maturity,
    3. For each θ_i, generate an entire path of S values, leading to the final one, $S_T^{(i)}$ which will be drawn from the predictive density $p(S_T, T | S_t, t)$.
  - Then, the value of the option under parameter estimation uncertainty, is given by

$$\begin{aligned} u(S_t, t) &= B(t, T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)] \\ &= B(t, T)\int_0^{\infty} \psi(S_T)p(S_T, T | S_t, t)dS_T \\ &\approx B(t, T)\frac{1}{M}\sum_{i=1}^{M} \psi(S_T^{(i)}) \end{aligned}$$

    where the last relationship reflects the Monte Carlo integration.

- When the SDE does not have a closed form solution, the Euler-Maruyama discretization seems to be the only feasible route. The procedure is the following.
  - First discretize the SDE.

$$S_{t_{j+1}} - S_{t_j} = a(S_{t_j}, t_j)\Delta t + b(S_{t_j}, t_j)(W_{t_{j+1}} - W_{t_j})$$

    where $t_1 = t, t_m = T$, so the simulation will be pathwise of length m between current valuation time t and maturity T.
  - Then
    1. Take a sample of size M with $\theta_i \sim p(\theta | Y_t)$,
    2. Generate m standard Gaussian random draws that will help create the path to maturity,
    3. For each θ_i, generate an entire path of S values, leading to the final one, $S_T^{(i)}$ which will be drawn from the predictive density $p(S_T, T | S_t, t)$.
  - Then, the value of the option under parameter estimation uncertainty, is given by

$$\begin{aligned} u(S_t, t) &= B(t,T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)] \\ &= B(t,T)\int_0^{\infty} \psi(S_T)p(S_T, T | S_t, t)dS_T \\ &\approx B(t,T)\frac{1}{M}\sum_{i=1}^{M} \psi(S_T^{(i)}) \end{aligned}$$

    where the last relationship reflects the Monte Carlo integration.

- When the SDE does not have a closed form solution, the Euler-Maruyama discretization seems to be the only feasible route. The procedure is the following.
  - First discretize the SDE.

$$S_{t_{j+1}} - S_{t_j} = a(S_{t_j}, t_j)\Delta t + b(S_{t_j}, t_j)(W_{t_{j+1}} - W_{t_j})$$

    where $t_1 = t, t_m = T$, so the simulation will be pathwise of length m between current valuation time t and maturity T.
  - Then
    1. Take a sample of size M with $\theta_i \sim p(\theta|Y_t)$,
    2. Generate m standard Gaussian random draws that will help create the path to maturity,
    3. For each θ_i, generate an entire path of S values, leading to the final one, $S_T^{(i)}$ which will be drawn from the predictive density $p(S_T, T|S_t, t)$.
  - Then, the value of the option under parameter estimation uncertainty, is given by

$$\begin{aligned} u(S_t, t) &= B(t, T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)] \\ &= B(t, T)\int_0^{\infty} \psi(S_T)p(S_T, T|S_t, t)dS_T \\ &\approx B(t, T)\frac{1}{M}\sum_{i=1}^{M}\psi(S_T^{(i)}) \end{aligned}$$

    where the last relationship reflects the Monte Carlo integration.

- Bayesian model averaging is a technique that accounts for lacking the precise knowledge of which model is best.
- Consider now that the market agent has a finite suite of models $\{M_i\}_{i=1,\dots,k}$ at her disposal *A priori* the trader does not know which model will perform best so, *ceteris paribus*, the only reasonable thing she could do is to derive an option price that is averaging across the uncertainty regarding model selection.
- $$\begin{aligned}
u(S_t, t|\{M_i\}_{i=1,\dots,k}) &= B(t,T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)|\{M_i\}_{i=1,\dots,k}] \\
&= B(t,T)\int_0^{\infty}\psi(S_T)\sum_{i=1}^{k} p(S_T,T|S_t,t,M_i)p(M_i|Y_t)dS_T \\
&= B(t,T)\sum_{i=1}^{k}\int_0^{\infty}\psi(S_T)p(S_T,T|S_t,t,M_i)dS_T p(M_i|D_t) \\
&= B(t,T)\sum_{i=1}^{k}\mathbb{E}_t^{Q}[\psi(S_T)|M_i]p(M_i|Y_t) \qquad (13)
\end{aligned}$$

- Bayesian model averaging is a technique that accounts for lacking the precise knowledge of which model is best.
- Consider now that the market agent has a finite suite of models $\{M_i\}_{i=1,\dots,k}$ at her disposal *A priori* the trader does not know which model will perform best so, *ceteris paribus*, the only reasonable thing she could do is to derive an option price that is averaging across the uncertainty regarding model selection.
-

$$
\begin{aligned}
u(S_t, t|\{M_i\}_{i=1,\dots,k}) &= B(t,T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)|\{M_i\}_{i=1,\dots,k}] \\
&= B(t,T)\int_0^{\infty}\psi(S_T)\sum_{i=1}^{k} p(S_T,T|S_t,t,M_i)p(M_i|Y_t)dS_T \\
&= B(t,T)\sum_{i=1}^{k}\int_0^{\infty}\psi(S_T)p(S_T,T|S_t,t,M_i)dS_T p(M_i|D_t) \\
&= B(t,T)\sum_{i=1}^{k}\mathbb{E}_t^{Q}[\psi(S_T)|M_i]p(M_i|Y_t) \qquad (13)
\end{aligned}
$$

- Bayesian model averaging is a technique that accounts for lacking the precise knowledge of which model is best.
- Consider now that the market agent has a finite suite of models $\{M_i\}_{i=1,\dots,k}$ at her disposal *A priori* the trader does not know which model will perform best so, *ceteris paribus*, the only reasonable thing she could do is to derive an option price that is averaging across the uncertainty regarding model selection.
-

$$
\begin{aligned}
u(S_t, t|\{M_i\}_{i=1,\dots,k}) &= B(t,T)\mathbb{E}_t^{\mathbb{Q}}[\psi(S_T)|\{M_i\}_{i=1,\dots,k}] \\
&= B(t,T)\int_0^{\infty}\psi(S_T)\sum_{i=1}^{k} p(S_T, T|S_t, t, M_i)p(M_i|Y_t)dS_T \\
&= B(t,T)\sum_{i=1}^{k}\int_0^{\infty}\psi(S_T)p(S_T, T|S_t, t, M_i)dS_T p(M_i|D_t) \\
&= B(t,T)\sum_{i=1}^{k}\mathbb{E}_t^{Q}[\psi(S_T)|M_i]p(M_i|Y_t) \qquad (13)
\end{aligned}
$$

- where $p(M_i|Y_t)$ is the posterior probability associated with model M_i, so when this probability is high then the option price received from this model receives a larger weight in the final valuation.
- Bayes' formula gives the recursive calculation of the model posterior probabilities in the light of new data.

$$p(M_i|Y_t) = \frac{p(Y_{[s,t]}|M_i, S_s)p(M_i|Y_s)}{\sum_{i=1}^{k} p(Y_{[s,t]}|M_i, S_s)p(M_i|Y_s)} \quad (14)$$

- each model is given some informative or non-informative prior probabilities and then the calculation of posterior model probabilities

$$\theta_i \sim p(\theta|Y_s), \quad p(Y_{[s,t]}|\{M_i\}_{i=1,\dots,k}, S_s) \approx \frac{1}{n}\sum_{i=1}^{k} p(Y_{[s,t]}|\theta_i, S_s)$$

- where $p(M_i|Y_t)$ is the posterior probability associated with model M_i, so when this probability is high then the option price received from this model receives a larger weight in the final valuation.
- Bayes' formula gives the recursive calculation of the model posterior probabilities in the light of new data.

$$p(M_i|Y_t) = \frac{p(Y_{[s,t]}|M_i, S_s)p(M_i|Y_s)}{\sum_{i=1}^{k} p(Y_{[s,t]}|M_i, S_s)p(M_i|Y_s)} \tag{14}$$

- each model is given some informative or non-informative prior probabilities and then the calculation of posterior model probabilities

$$\theta_i \sim p(\theta|Y_s), \quad p(Y_{[s,t]}|\{M_i\}_{i=1,\dots,k}, S_s) \approx \frac{1}{n}\sum_{i=1}^{k} p(Y_{[s,t]}|\theta_i, S_s)$$

- where $p(M_i|Y_t)$ is the posterior probability associated with model M_i, so when this probability is high then the option price received from this model receives a larger weight in the final valuation.
- Bayes' formula gives the recursive calculation of the model posterior probabilities in the light of new data.

$$p(M_i|Y_t) = \frac{p(Y_{[s,t]}|M_i, S_s)p(M_i|Y_s)}{\sum_{i=1}^{k} p(Y_{[s,t]}|M_i, S_s)p(M_i|Y_s)} \tag{14}$$

- each model is given some informative or non-informative prior probabilities and then the calculation of posterior model probabilities

$$\theta_i \sim p(\theta|Y_s), \quad p(Y_{[s,t]}|\{M_i\}_{i=1,\dots,k}, S_s) \approx \frac{1}{n}\sum_{i=1}^{k} p(Y_{[s,t]}|\theta_i, S_s)$$

- The GBM model is described by the following equations in continuous-time

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma_{BS} S_t dZ_t && (15) \\ &= r S_t dt + \sigma_{BS} S_t dW_t && (16) \end{aligned}$$

where $W_t = Z_t + \lambda t$, where $\lambda = \frac{\mu - r}{\sigma_{BS}}$ is the market price of risk, and r is the constant riskfree rate. Evidently W is the Wiener process associated with the risk-neutral pricing measure while Z is the Wiener process associated with the physical probability measure.

- the CEV model is given by

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma_{CEV} S_t^{\gamma} dZ_t && (17) \\ &= r S_t dt + \sigma_{CEV} S_t^{\gamma} dW_t && (18) \end{aligned}$$

and for this model one can prove that the elasticity of the instantaneous return variance with respect to price is equal to $2(\gamma - 1)$.

- In order to avoid technical problems related to arbitrage it is usually assumed that $\gamma \in [0, 1)$.

- Bayesian analysis and Markov Chain Monte Carlo (MCMC) are going to be used, first to extract inference on the parameters of the two models and secondly to calculate no-arbitrage European call and put prices for options contingent on the FTSE100 index.
- Following Bunnin et al. (2002) we use historical data covering 50 weekly levels of the FTSE100 from 30 December 1997 to 9 December 1998.
- The dividend yield is initially ignored and the data for the options is as follows: the strike price is $K = 5500$, the initial index value is $S_{t_0} = 5669.1$, the risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.
- However, as opposed to Bunnin et al. (2002) who assumed $\mu = 0$, I have allowed this parameter to be Gaussian distributed with a very large, 10000, variance and zero mean.
- the prior distribution for the volatility parameter, is taken by Bunnin et al. (2002) as $\sigma_{BS} \sim Uniform(0.1, 0.3)$.
- I have used an inverse-gamma distribution that is very flat and covers a wide range. In other words, I will let the data decide on the most likely values for the volatility parameter.

- Bayesian analysis and Markov Chain Monte Carlo (MCMC) are going to be used, first to extract inference on the parameters of the two models and secondly to calculate no-arbitrage European call and put prices for options contingent on the FTSE100 index.
- Following Bunnin et al. (2002) we use historical data covering 50 weekly levels of the FTSE100 from 30 December 1997 to 9 December 1998.
- The dividend yield is initially ignored and the data for the options is as follows: the strike price is $K = 5500$, the initial index value is $S_{t_0} = 5669.1$, the risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.
- However, as opposed to Bunnin et al. (2002) who assumed $\mu = 0$, I have allowed this parameter to be Gaussian distributed with a very large, 10000, variance and zero mean.
- the prior distribution for the volatility parameter, is taken by Bunnin et al. (2002) as $\sigma_{BS} \sim Uniform(0.1, 0.3)$.
- I have used an inverse gamma distribution that is very flat and covers a wide range. In other words, I will let the data decide on the most likely values for the volatility parameter.

- Bayesian analysis and Markov Chain Monte Carlo (MCMC) are going to be used, first to extract inference on the parameters of the two models and secondly to calculate no-arbitrage European call and put prices for options contingent on the FTSE100 index.
- Following Bunnin et al. (2002) we use historical data covering 50 weekly levels of the FTSE100 from 30 December 1997 to 9 December 1998.
- The dividend yield is initially ignored and the data for the options is as follows: the strike price is $K = 5500$, the initial index value is $S_{t_0} = 5669.1$, the risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.
- However, as opposed to Bunnin et al. (2002) who assumed $\mu = 0$, I have allowed this parameter to be Gaussian distributed with a very large, 10000, variance and zero mean.
- the prior distribution for the volatility parameter, is taken by Bunnin et al. (2002) as $\sigma_{BS} \sim Uniform(0.1, 0.3)$.
- I have used an inverse gamma distribution that is very flat and covers a wide range. In other words, I will let the data decide on the most likely values for the volatility parameter.

- Bayesian analysis and Markov Chain Monte Carlo (MCMC) are going to be used, first to extract inference on the parameters of the two models and secondly to calculate no-arbitrage European call and put prices for options contingent on the FTSE100 index.
- Following Bunnin et al. (2002) we use historical data covering 50 weekly levels of the FTSE100 from 30 December 1997 to 9 December 1998.
- The dividend yield is initially ignored and the data for the options is as follows: the strike price is $K = 5500$, the initial index value is $S_{t_0} = 5669.1$, the risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.
- However, as opposed to Bunnin et al. (2002) who assumed $\mu = 0$, I have allowed this parameter to be Gaussian distributed with a very large, 10000, variance and zero mean.
- the prior distribution for the volatility parameter, is taken by Bunnin et al. (2002) as $\sigma_{BS} \sim Uniform(0.1, 0.3)$.
- I have used an inverse-gamma distribution that is very flat and covers a wide range. In other words, I will let the data decide on the most likely values for the volatility parameter.

- Bayesian analysis and Markov Chain Monte Carlo (MCMC) are going to be used, first to extract inference on the parameters of the two models and secondly to calculate no-arbitrage European call and put prices for options contingent on the FTSE100 index.
- Following Bunnin et al. (2002) we use historical data covering 50 weekly levels of the FTSE100 from 30 December 1997 to 9 December 1998.
- The dividend yield is initially ignored and the data for the options is as follows: the strike price is $K = 5500$, the initial index value is $S_{t_0} = 5669.1$, the risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.
- However, as opposed to Bunnin et al. (2002) who assumed $\mu = 0$, I have allowed this parameter to be Gaussian distributed with a very large, 10000, variance and zero mean.
- the prior distribution for the volatility parameter, is taken by Bunnin et al. (2002) as $\sigma_{BS} \sim Uniform(0.1, 0.3)$.
- I have used an inverse-gamma distribution that is very flat and covers a wide range. In other words, I will let the data decide on the most likely values for the volatility parameter.

- Bayesian analysis and Markov Chain Monte Carlo (MCMC) are going to be used, first to extract inference on the parameters of the two models and secondly to calculate no-arbitrage European call and put prices for options contingent on the FTSE100 index.
- Following Bunnin et al. (2002) we use historical data covering 50 weekly levels of the FTSE100 from 30 December 1997 to 9 December 1998.
- The dividend yield is initially ignored and the data for the options is as follows: the strike price is $K = 5500$, the initial index value is $S_{t_0} = 5669.1$, the risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.
- However, as opposed to Bunnin et al. (2002) who assumed $\mu = 0$, I have allowed this parameter to be Gaussian distributed with a very large, 10000, variance and zero mean.
- the prior distribution for the volatility parameter, is taken by Bunnin et al. (2002) as $\sigma_{BS} \sim Uniform(0.1, 0.3)$.
- I have used an inverse-gamma distribution that is very flat and covers a wide range. In other words, I will let the data decide on the most likely values for the volatility parameter.

- WinBUGS 1.4, simulating two chains. The convergence is very rapid and after a burn-in period of 50000 simulations I run another set of 50000 simulations from which I extract the summary inferential results in Table BS Bayesian model.
- The posterior mean and median are about 0.20, confirming the analysis detailed in Bunnin et al. (2002).
- The posterior mean and median for μ is 0.11 but the credibility interval constructed from the 2.5% and the 97.5% quantiles includes the zero value and therefore, the idea of inferring therefore that this value could equal zero is not wrong.
- Secondly the credibility interval for σ_{BS} is [0.1659,0.249].

- WinBUGS 1.4, simulating two chains. The convergence is very rapid and after a burn-in period of 50000 simulations I run another set of 50000 simulations from which I extract the summary inferential results in Table BS Bayesian model.
- The posterior mean and median are about 0.20, confirming the analysis detailed in Bunnin et al. (2002).
- The posterior mean and median for μ is 0.11 but the credibility interval constructed from the 2.5% and the 97.5% quantiles includes the zero value and therefore, the idea of inferring therefore that this value could equal zero is not wrong.
- Secondly the credibility interval for σ_{BS} is [0.1659, 0.249].

- WinBUGS 1.4, simulating two chains. The convergence is very rapid and after a burn-in period of 50000 simulations I run another set of 50000 simulations from which I extract the summary inferential results in Table BS Bayesian model.
- The posterior mean and median are about 0.20, confirming the analysis detailed in Bunnin et al. (2002).
- The posterior mean and median for μ is 0.11 but the credibility interval constructed from the 2.5% and the 97.5% quantiles includes the zero value and therefore, the idea of inferring therefore that this value could equal zero is not wrong.
- Secondly the credibility interval for σ_{BS} is [0.1659, 0.249].

- WinBUGS 1.4, simulating two chains. The convergence is very rapid and after a burn-in period of 50000 simulations I run another set of 50000 simulations from which I extract the summary inferential results in Table BS Bayesian model.
- The posterior mean and median are about 0.20, confirming the analysis detailed in Bunnin et al. (2002).
- The posterior mean and median for μ is 0.11 but the credibility interval constructed from the 2.5% and the 97.5% quantiles includes the zero value and therefore, the idea of inferring therefore that this value could equal zero is not wrong.
- Secondly the credibility interval for σ_{BS} is $[0.1659, 0.249]$.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is [708.5, 863.6], which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is [142.0, 297.1].
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is [708.5, 863.6], which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is [142.0, 297.1].
- the market price of risk has the credibility interval [−1.833, 2.196] and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is [708.5, 863.6], which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is [142.0, 297.1].
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is $[708.5, 863.6]$, which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is $[142.0, 297.1]$.
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is $[708.5, 863.6]$, which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is $[142.0, 297.1]$.
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is $[708.5, 863.6]$, which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is $[142.0, 297.1]$.
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is $[708.5, 863.6]$, which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is $[142.0, 297.1]$.
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is $[708.5, 863.6]$, which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is $[142.0, 297.1]$.
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

- The uncertainty in the parameter estimation is captured and depicted beautifully by the posterior densities in the graphs in BS Bayesian model.
  - the posterior mean of the call is 774.9,
  - the posterior median is 771.1,
  - and the credibility interval for the European call price is $[708.5, 863.6]$, which covers the value provided by Bunnin et al. (2002).
  - for the European put, the posterior mean is 208.4,
  - the posterior median is 204.6
  - and the credibility interval is $[142.0, 297.1]$.
- the market price of risk has the credibility interval $[-1.833, 2.196]$ and therefore we cannot reject the hypothesis that this might be equal to zero.
- Looking at its posterior density one can observe that zero is also the most likely value.

BS Bayesian Figures .

- the posterior density of the market price of risk defined as $\lambda = \frac{\mu - r}{\sigma_{BS}}$.
- the dividend yield equal to zero and the risk-free rate $r = 0.075$.
- The possible values for μ and σ_{BS} will generate a sample of values for λ.
- The MCMC output can be utilised to calculate the posterior densities of the Greek parameters such as Delta.
- the most likely value for the Delta parameter is 0.735 but values such as 0.7 or 0.77 are also possible.
- Likewise, for the put the most likely Delta is -0.265 but values like -0.3 or -0.22 are also feasible, albeit less likely.
- the Delta for European call and put with the Black-Scholes model assuming the estimate of volatility as $\sigma_{BS} = 0.20$ we get that $\Delta(call) = 0.7345$ and $\Delta(put) = -0.26552$.

BS Bayesian Figures.

- the posterior density of the market price of risk defined as $\lambda = \frac{\mu - r}{\sigma_{BS}}$.
- the dividend yield equal to zero and the risk-free rate $r = 0.075$.
- The possible values for μ and σ_{BS} will generate a sample of values for λ.
- The MCMC output can be utilised to calculate the posterior densities of the Greek parameters such as Delta.
- the most likely value for the Delta parameter is 0.735 but values such as 0.7 or 0.77 are also possible.
- Likewise, for the put the most likely Delta is -0.265 but values like -0.3 or -0.22 are also feasible, albeit less likely.
- the Delta for European call and put with the Black-Scholes model assuming the estimate of volatility as $\sigma_{BS} = 0.20$ we get that $\Delta(call) = 0.7345$ and $\Delta(put) = -0.26552$.

BS Bayesian Figures .

- the posterior density of the market price of risk defined as $\lambda = \frac{\mu - r}{\sigma_{BS}}$.
- the dividend yield equal to zero and the risk-free rate $r = 0.075$.
- The possible values for μ and σ_{BS} will generate a sample of values for λ.
- The MCMC output can be utilised to calculate the posterior densities of the Greek parameters such as Delta.
- the most likely value for the Delta parameter is 0.735 but values such as 0.7 or 0.77 are also possible.
- Likewise, for the put the most likely Delta is -0.265 but values like -0.3 or -0.22 are also feasible, albeit less likely.
- the Delta for European call and put with the Black-Scholes model assuming the estimate of volatility as $\sigma_{BS} = 0.20$ we get that $\Delta(call) = 0.7345$ and $\Delta(put) = -0.26552$.

BS Bayesian Figures.

- the posterior density of the market price of risk defined as $\lambda = \frac{\mu - r}{\sigma_{BS}}$.
- the dividend yield equal to zero and the risk-free rate $r = 0.075$.
- The possible values for μ and σ_{BS} will generate a sample of values for λ.
- The MCMC output can be utilised to calculate the posterior densities of the Greek parameters such as Delta.
- the most likely value for the Delta parameter is 0.735 but values such as 0.7 or 0.77 are also possible.
- Likewise, for the put the most likely Delta is -0.265 but values like -0.3 or -0.22 are also feasible, albeit less likely.
- the Delta for European call and put with the Black-Scholes model assuming the estimate of volatility as $\sigma_{BS} = 0.20$ we get that $\Delta(call) = 0.7345$ and $\Delta(put) = -0.26552$.

BS Bayesian Figures.

- the posterior density of the market price of risk defined as $\lambda = \frac{\mu - r}{\sigma_{BS}}$.
- the dividend yield equal to zero and the risk-free rate $r = 0.075$.
- The possible values for μ and σ_{BS} will generate a sample of values for λ.
- The MCMC output can be utilised to calculate the posterior densities of the Greek parameters such as Delta.
- the most likely value for the Delta parameter is 0.735 but values such as 0.7 or 0.77 are also possible.
- Likewise, for the put the most likely Delta is -0.265 but values like -0.3 or -0.22 are also feasible, albeit less likely.
- the Delta for European call and put with the Black-Scholes model assuming the estimate of volatility as $\sigma_{BS} = 0.20$ we get that $\Delta(call) = 0.7345$ and $\Delta(put) = -0.26552$.

BS Bayesian Figures.

- the posterior density of the market price of risk defined as $\lambda = \frac{\mu - r}{\sigma_{BS}}$.
- the dividend yield equal to zero and the risk-free rate $r = 0.075$.
- The possible values for μ and σ_{BS} will generate a sample of values for λ.
- The MCMC output can be utilised to calculate the posterior densities of the Greek parameters such as Delta.
- the most likely value for the Delta parameter is 0.735 but values such as 0.7 or 0.77 are also possible.
- Likewise, for the put the most likely Delta is -0.265 but values like -0.3 or -0.22 are also feasible, albeit less likely.
- the Delta for European call and put with the Black-Scholes model assuming the estimate of volatility as $\sigma_{BS} = 0.20$ we get that $\Delta(call) = 0.7345$ and $\Delta(put) = -0.26552$.

.

- the posterior density of the market price of risk defined as $\lambda = \frac{\mu - r}{\sigma_{BS}}$.
- the dividend yield equal to zero and the risk-free rate $r = 0.075$.
- The possible values for μ and σ_{BS} will generate a sample of values for λ.
- The MCMC output can be utilised to calculate the posterior densities of the Greek parameters such as Delta.
- the most likely value for the Delta parameter is 0.735 but values such as 0.7 or 0.77 are also possible.
- Likewise, for the put the most likely Delta is -0.265 but values like -0.3 or -0.22 are also feasible, albeit less likely.
- the Delta for European call and put with the Black-Scholes model assuming the estimate of volatility as $\sigma_{BS} = 0.20$ we get that $\Delta(call) = 0.7345$ and $\Delta(put) = -0.26552$.

- We now proceed with the parameter estimation for the CEV model
- The prior distributions that I found to work well from all points of view were a very flat uniform distribution for σ_{CEV}, take (0,100) as an example; a uniform distribution covering (0, 0.30) for the dividend yield q and a beta distribution for the parameter γ that is constrained to be between 0 and 1 in order to avoid technical problems related to absorption at zero or explosion if other values were allowed.
- The same routine as described above for the Black-Scholes model is followed to obtain a
- a sample of 20000 values is used for posterior inference
- One cannot reject the hypothesis that the drift parameter μ is zero.
- The diffusion parameter σ_{CEV} has a posterior mean of 2 and posterior median of 1.8. The CEV elasticity parameter γ is also significant and its posterior mean is 0.29 while its posterior median is 0.2573.

BS Bayesian model

- We now proceed with the parameter estimation for the CEV model
- The prior distributions that I found to work well from all points of view were a very flat uniform distribution for σ_{CEV}, take (0,100) as an example; a uniform distribution covering $(0, 0.30)$ for the dividend yield q and a beta distribution for the parameter γ that is constrained to be between 0 and 1 in order to avoid technical problems related to absorption at zero or explosion if other values were allowed.
- The same routine as described above for the Black-Scholes model is followed to obtain a
- a sample of 20000 values is used for posterior inference
- One cannot reject the hypothesis that the drift parameter μ is zero.
- The diffusion parameter σ_{CEV} has a posterior mean of 2 and posterior median of 1.8. The CEV elasticity parameter γ is also significant and its posterior mean is 0.29 while its posterior median is 0.2573.

BS Bayesian model

- We now proceed with the parameter estimation for the CEV model
- The prior distributions that I found to work well from all points of view were a very flat uniform distribution for σ_{CEV}, take (0,100) as an example; a uniform distribution covering $(0, 0.30)$ for the dividend yield q and a beta distribution for the parameter γ that is constrained to be between 0 and 1 in order to avoid technical problems related to absorption at zero or explosion if other values were allowed.
- The same routine as described above for the Black-Scholes model is followed to obtain a
- a sample of 20000 values is used for posterior inference
- One cannot reject the hypothesis that the drift parameter μ is zero.
- The diffusion parameter σ_{CEV} has a posterior mean of 2 and posterior median of 1.8. The CEV elasticity parameter γ is also significant and its posterior mean is 0.29 while its posterior median is 0.2573.

BS Bayesian model

- We now proceed with the parameter estimation for the CEV model
- The prior distributions that I found to work well from all points of view were a very flat uniform distribution for σ_{CEV}, take (0,100) as an example; a uniform distribution covering $(0, 0.30)$ for the dividend yield q and a beta distribution for the parameter γ that is constrained to be between 0 and 1 in order to avoid technical problems related to absorption at zero or explosion if other values were allowed.
- The same routine as described above for the Black-Scholes model is followed to obtain a
- a sample of 20000 values is used for posterior inference
- One cannot reject the hypothesis that the drift parameter μ is zero.
- The diffusion parameter σ_{CEV} has a posterior mean of 2 and posterior median of 1.8. The CEV elasticity parameter γ is also significant and its posterior mean is 0.29 while its posterior median is 0.2573.

BS Bayesian model

- We now proceed with the parameter estimation for the CEV model
- The prior distributions that I found to work well from all points of view were a very flat uniform distribution for σ_{CEV}, take (0,100) as an example; a uniform distribution covering $(0, 0.30)$ for the dividend yield q and a beta distribution for the parameter γ that is constrained to be between 0 and 1 in order to avoid technical problems related to absorption at zero or explosion if other values were allowed.
- The same routine as described above for the Black-Scholes model is followed to obtain a
- a sample of 20000 values is used for posterior inference
- One cannot reject the hypothesis that the drift parameter μ is zero.
- The diffusion parameter σ_{CEV} has a posterior mean of 2 and posterior median of 1.8. The CEV elasticity parameter γ is also significant and its posterior mean is 0.29 while its posterior median is 0.2573.

BS Bayesian model

- We now proceed with the parameter estimation for the CEV model
- The prior distributions that I found to work well from all points of view were a very flat uniform distribution for σ_{CEV}, take (0,100) as an example; a uniform distribution covering $(0, 0.30)$ for the dividend yield q and a beta distribution for the parameter γ that is constrained to be between 0 and 1 in order to avoid technical problems related to absorption at zero or explosion if other values were allowed.
- The same routine as described above for the Black-Scholes model is followed to obtain a
- a sample of 20000 values is used for posterior inference
- One cannot reject the hypothesis that the drift parameter μ is zero.
- The diffusion parameter σ_{CEV} has a posterior mean of 2 and posterior median of 1.8. The CEV elasticity parameter γ is also significant and its posterior mean is 0.29 while its posterior median is 0.2573.

BS Bayesian model

- The power of the MCMC approach is that we can visualise the entire posterior distribution for each parameter. CEV Bayesian model
- the most likely values $\gamma = 0.2$ and $\sigma_{CEV} = 0.25$. Remark that this is far away from both posterior mean and posterior median.
- In Figures Bayesian call CEV the surface of option prices, calls and puts, respectively, that are obtained by combining a sample of 500 values for γ and 500 values for σ_{CEV} from the stationary part of the MCMC distribution. The dividend yield is taken as zero, for comparison with the results in Bunnin et al. (2002).
- For European calls, the maximum obtained value was 5583.7 and the smallest was 566.51, while for the puts the maximum obtained value was 5017.2 and the minimum was zero.

- The power of the MCMC approach is that we can visualise the entire posterior distribution for each parameter. CEV Bayesian model
- the most likely values $\gamma = 0.2$ and $\sigma_{CEV} = 0.25$. Remark that this is far away from both posterior mean and posterior median.
- In Figures Bayesian call CEV the surface of option prices, calls and puts, respectively, that are obtained by combining a sample of 500 values for γ and 500 values for σ_{CEV} from the stationary part of the MCMC distribution. The dividend yield is taken as zero, for comparison with the results in Bunnin et al. (2002).
- For European calls, the maximum obtained value was 5583.7 and the smallest was 566.51, while for the puts the maximum obtained value was 5017.2 and the minimum was zero.

- The power of the MCMC approach is that we can visualise the entire posterior distribution for each parameter. CEV Bayesian model
- the most likely values $\gamma = 0.2$ and $\sigma_{CEV} = 0.25$. Remark that this is far away from both posterior mean and posterior median.
- In Figures Bayesian call CEV the surface of option prices, calls and puts, respectively, that are obtained by combining a sample of 500 values for γ and 500 values for σ_{CEV} from the stationary part of the MCMC distribution. The dividend yield is taken as zero, for comparison with the results in Bunnin et al. (2002).
- For European calls, the maximum obtained value was 5583.7 and the smallest was 566.51, while for the puts the maximum obtained value was 5017.2 and the minimum was zero.

- The power of the MCMC approach is that we can visualise the entire posterior distribution for each parameter. CEV Bayesian model
- the most likely values $\gamma = 0.2$ and $\sigma_{CEV} = 0.25$. Remark that this is far away from both posterior mean and posterior median.
- In Figures Bayesian call CEV the surface of option prices, calls and puts, respectively, that are obtained by combining a sample of 500 values for γ and 500 values for σ_{CEV} from the stationary part of the MCMC distribution. The dividend yield is taken as zero, for comparison with the results in Bunnin et al. (2002).
- For European calls, the maximum obtained value was 5583.7 and the smallest was 566.51, while for the puts the maximum obtained value was 5017.2 and the minimum was zero.

- Our empirical analysis shows that the risk associated with parameter uncertainty can be
  - substantial even for vanilla products such as European call and put options,
  - asymmetric for the buyer and the seller in the contract, even when the *same* parametric model class is used by both.
- I propose a new measure of model risk related to parameter uncertainty, by analogy with the way value-at-risk was introduced for quantifying market risk.

Definition

Given a model defined unambiguously by a vector of parameters ϑ, for any contingent claim price function $\Pi(H, \vartheta)$ with payoff H, we define the parameter uncertainty model risk (PUMR) measure corresponding to $\Pi(H, \vartheta)$, at the $100(1-\alpha\%)$ level of confidence, as the $\alpha\%$ quantile in the direction of risk.

- To focus the discussion, consider the European call option with posterior densities of the fair price as represented in Fig. 17. In addition, suppose that $\alpha\% = 2.5\%$.

- Our empirical analysis shows that the risk associated with parameter uncertainty can be
  - substantial even for vanilla products such as European call and put options,
  - asymmetric for the buyer and the seller in the contract, even when the *same* parametric model class is used by both.
- I propose a new measure of model risk related to parameter uncertainty, by analogy with the way value-at-risk was introduced for quantifying market risk.

Definition

Given a model defined unambiguously by a vector of parameters ϑ, for any contingent claim price function $\Pi(H, \vartheta)$ with payoff H, we define the parameter uncertainty model risk (PUMR) measure corresponding to $\Pi(H; \vartheta)$, at the $100(1-\alpha\%)$ level of confidence, as the $\alpha\%$ quantile in the direction of risk.

- To focus the discussion, consider the European call option with posterior densities of the fair price as represented in Fig. 17. In addition, suppose that $\alpha\% = 2.5\%$.

- Our empirical analysis shows that the risk associated with parameter uncertainty can be
  - substantial even for vanilla products such as European call and put options,
  - asymmetric for the buyer and the seller in the contract, even when the *same* parametric model class is used by both.
- I propose a new measure of model risk related to parameter uncertainty, by analogy with the way value-at-risk was introduced for quantifying market risk.

Definition

Given a model defined unambiguously by a vector of parameters ϑ, for any contingent claim price function $\Pi(H;\vartheta)$ with payoff H, we define the parameter uncertainty model risk (PUMR) measure corresponding to $\Pi(H;\vartheta)$, at the $100(1-\alpha\%)$ level of confidence, as the $\alpha\%$ quantile in the direction of risk.

- To focus the discussion, consider the European call option with posterior densities of the fair price as represented in Fig. 17. In addition, suppose that $\alpha\% = 2.5\%$.

- Our empirical analysis shows that the risk associated with parameter uncertainty can be
  - substantial even for vanilla products such as European call and put options,
  - asymmetric for the buyer and the seller in the contract, even when the *same* parametric model class is used by both.
- I propose a new measure of model risk related to parameter uncertainty, by analogy with the way value-at-risk was introduced for quantifying market risk.

Definition

Given a model defined unambiguously by a vector of parameters ϑ, for any contingent claim price function $\Pi(H; \vartheta)$ with payoff H, we define the parameter uncertainty model risk (PUMR) measure corresponding to $\Pi(H; \vartheta)$, at the $100(1-\alpha\%)$ level of confidence, as the $\alpha\%$ quantile in the direction of risk.

- To focus the discussion, consider the European call option with posterior densities of the fair price as represented in Fig. 17. In addition, suppose that $\alpha\% = 2.5\%$.

- Our empirical analysis shows that the risk associated with parameter uncertainty can be
  - substantial even for vanilla products such as European call and put options,
  - asymmetric for the buyer and the seller in the contract, even when the *same* parametric model class is used by both.
- I propose a new measure of model risk related to parameter uncertainty, by analogy with the way value-at-risk was introduced for quantifying market risk.

Definition

Given a model defined unambiguously by a vector of parameters ϑ, for any contingent claim price function $\Pi(H;\vartheta)$ with payoff H, we define the parameter uncertainty model risk (PUMR) measure corresponding to $\Pi(H;\vartheta)$, at the $100(1-\alpha\%)$ *level of confidence, as the $\alpha\%$ quantile in the direction of risk.*

- To focus the discussion, consider the European call option with posterior densities of the fair price as represented in Fig. 17. In addition, suppose that $\alpha\% = 2.5\%$.

- Our empirical analysis shows that the risk associated with parameter uncertainty can be
  - substantial even for vanilla products such as European call and put options,
  - asymmetric for the buyer and the seller in the contract, even when the *same* parametric model class is used by both.
- I propose a new measure of model risk related to parameter uncertainty, by analogy with the way value-at-risk was introduced for quantifying market risk.

Definition

Given a model defined unambiguously by a vector of parameters ϑ, for any contingent claim price function $\Pi(H;\vartheta)$ with payoff H, we define the parameter uncertainty model risk (PUMR) measure corresponding to $\Pi(H;\vartheta)$, at the $100(1-\alpha\%)$ *level of confidence, as the $\alpha\%$ quantile in the direction of risk.*

- To focus the discussion, consider the European call option with posterior densities of the fair price as represented in Fig. 17. In addition, suppose that $\alpha\% = 2.5\%$.

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1-\alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1-\alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1-\alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1-\alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1 - \alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1 - \alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,
  - the PUMR for the call is 863.6 for the seller and 708.5 for the buyer,
  - the PUMR for the put is 297.1 for the seller and 142.0 for the buyer.

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1-\alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,
  - the PUMR for the call is 863.6 for the seller and 708.5 for the buyer,
  - the PUMR for the put is 297.1 for the seller and 142.0 for the buyer.

- For the seller of the derivative, the risk is represented by the right tail of the posterior distribution since if trading for a contract is done outside this area when the true price is actually in this area then the seller will incur a loss.
- In other words, the seller is exposed to feasible higher prices that he/she is not taking into consideration when choosing a point estimate of the fair price.
- Hence the PUMR for the seller is the $100(1-\alpha\%)$ quantile, or the $\alpha\%$ right quantile.
- Similarly, for the buyer of the derivative, the parameter uncertainty risk is represented by the left tail of the posterior distribution. If the real fair value of the contract is exactly equal to the $\alpha\%$ quantile level, then any trading done at a value higher than this benchmark will result in a loss.
- the 97.5% quantile quantifies the PUMR for the seller while the 2.5% quantile measures the PUMR for the buyer of the derivative.
- Under the Black-Scholes model,
  - the PUMR for the call is 863.6 for the seller and 708.5 for the buyer,
  - the PUMR for the put is 297.1 for the seller and 142.0 for the buyer.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
    - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
    - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- Under the CEV model,
  - the PUMR for the call is 716.45 for the seller and 566.51 for the buyer,
  - whereas the PUMR for the put is 149.95 for the seller and 0.0000 for the buyer.
- One way to compare different models with respect to the parameter estimation risk embedded in derivatives pricing is to consider as a discrepancy measure the PUMR for the seller and the buyer.
- Models with a smaller PUMR should be preferred because that is equivalent with posterior distributions that are narrowly spread.
- I shall call this discrepancy measure the *PUMR distance*.
- For the Black-Scholes model this distance is equal to 155 roughly for both put and call, and for the CEV model this distance is equal to 149.95.
- The two models come quite close in this particular case.
- Based on these results the CEV model shows a slight superiority despite having one extra parameter.

- The inference presented reveals that model risk due to parameter uncertainty can be quite large.
- given the skewness of the posterior densities, *the two parties in the financial contract do not have the same magnitude of exposure to model risk*.
- For the European call option in discussion the seller takes on *more* model risk of the parameter estimation type.
- This is correct since call option contracts have no downside and variation comes from the upside and also because the process used for modeling the underlying index cannot become non-positive.
- This point is very important for investment banks and financial institutions where both long and short positions may be simultaneously present on the balance sheet due to multiple counterparties.
- Model risk will not cancel out over the same contract for opposite positions. This answers a question posed by Gibson et al. (1999) as to whether model risk is symmetric. My answer is that it is not.

- The inference presented reveals that model risk due to parameter uncertainty can be quite large.
- given the skewness of the posterior densities, *the two parties in the financial contract do not have the same magnitude of exposure to model risk*.
- For the European call option in discussion the seller takes on *more* model risk of the parameter estimation type.
- This is correct since call option contracts have no downside and variation comes from the upside and also because the process used for modeling the underlying index cannot become non-positive.
- This point is very important for investment banks and financial institutions where both long and short positions may be simultaneously present on the balance sheet due to multiple counterparties.
- Model risk will not cancel out over the same contract for opposite positions. This answers a question posed by Gibson et al. (1999) as to whether model risk is symmetric. My answer is that it is not.

- The inference presented reveals that model risk due to parameter uncertainty can be quite large.
- given the skewness of the posterior densities, *the two parties in the financial contract do not have the same magnitude of exposure to model risk*.
- For the European call option in discussion the seller takes on *more* model risk of the parameter estimation type.
- This is correct since call option contracts have no downside and variation comes from the upside and also because the process used for modeling the underlying index cannot become non-positive.
- This point is very important for investment banks and financial institutions where both long and short positions may be simultaneously present on the balance sheet due to multiple counterparties.
- Model risk will not cancel out over the same contract for opposite positions. This answers a question posed by Gibson et al. (1999) as to whether model risk is symmetric. My answer is that it is not.

- The inference presented reveals that model risk due to parameter uncertainty can be quite large.
- given the skewness of the posterior densities, *the two parties in the financial contract do not have the same magnitude of exposure to model risk*.
- For the European call option in discussion the seller takes on *more* model risk of the parameter estimation type.
- This is correct since call option contracts have no downside and variation comes from the upside and also because the process used for modeling the underlying index cannot become non-positive.
- This point is very important for investment banks and financial institutions where both long and short positions may be simultaneously present on the balance sheet due to multiple counterparties.
- Model risk will not cancel out over the same contract for opposite positions. This answers a question posed by Gibson et al. (1999) as to whether model risk is symmetric. My answer is that it is not.

- The inference presented reveals that model risk due to parameter uncertainty can be quite large.
- given the skewness of the posterior densities, *the two parties in the financial contract do not have the same magnitude of exposure to model risk*.
- For the European call option in discussion the seller takes on *more* model risk of the parameter estimation type.
- This is correct since call option contracts have no downside and variation comes from the upside and also because the process used for modeling the underlying index cannot become non-positive.
- This point is very important for investment banks and financial institutions where both long and short positions may be simultaneously present on the balance sheet due to multiple counterparties.
- Model risk will not cancel out over the same contract for opposite positions. This answers a question posed by Gibson et al. (1999) as to whether model risk is symmetric. My answer is that it is not.

- The inference presented reveals that model risk due to parameter uncertainty can be quite large.
- given the skewness of the posterior densities, *the two parties in the financial contract do not have the same magnitude of exposure to model risk*.
- For the European call option in discussion the seller takes on *more* model risk of the parameter estimation type.
- This is correct since call option contracts have no downside and variation comes from the upside and also because the process used for modeling the underlying index cannot become non-positive.
- This point is very important for investment banks and financial institutions where both long and short positions may be simultaneously present on the balance sheet due to multiple counterparties.
- Model risk will not cancel out over the same contract for opposite positions. This answers a question posed by Gibson et al. (1999) as to whether model risk is symmetric. My answer is that it is not.

MCMC Estimation of Credit Risk Measures I

- The relationship between default frequencies and rating categories has been explored in Blume et al. (1998) and Zhou (2001)) and it has been put again under scrutiny in the aftermath of the subprime crisis, with misleading ratings being blamed for inducing false investor's expectations of probabilities of default.
- Carey and Hrycay (2001) discussed an empirical examination of the major mapping methods used to estimate average default probabilities by grade. They found evidence of potential bias, instability and gaming.
- Stefanescu et al. (2009) developed a conceptual statistical calibration methodology for credit transition matrices, including probabilities of default, using a hierarchical Bayesian approach that takes into account ordinal explanatory variables.
- Bluhm et al. (2003) used Moody's ratings data to show how corporate default probabilities may be calibrated to external ratings. Their analysis is based on a log-linear model linking the observed mean default frequency (*MD*) over the period 1993 to 2000 to the credit ratings category (*Rating*) modelled as an ordinal variable.

MCMC Estimation of Credit Risk Measures II

- Default probabilities are inferred for all credit rating categories.
- use the observed mean default frequencies (MD) over the period 1993 to 2000.
- Some categories had no observed defaults and therefore there is no data available.
- The model discussed in Bluhm et al. (2003) for observed mean default frequencies (MD) of corporate companies rated by Moody's over the period 1993 to 2000, is given by the log-linear relationship

$$\ln(MD) = -5\ln(30) + 0.5075\,Rating \tag{19}$$

Table: Corporate default probabilities implied by the log-linear regression model for corporates rated by Moody's over the period 1993 to 2000.

Rating	MD	s.d.	Estimated default probability
Aaa	NA	NA	0.005%
Aa1	NA	NA	0.008%
Aa2	NA	NA	0.014%
Aa3	0.08%	0.33%	0.023%
A1	NA	NA	0.038%
A2	NA	NA	0.063%
A3	NA	NA	0.105%
Baa1	0.06%	0.19%	0.174%
Baa2	0.06%	0.20%	0.289%
Baa3	0.46%	1.16%	0.480%
Ba1	0.69%	1.03%	0.797%
Ba2	0.63%	0.86%	1.324%
Ba3	2.39%	2.35%	2.200%
B1	3.79%	2.49%	3.654%
B2	7.96%	6.08%	6.070%
B3	12.89%	8.14%	10.083%

From an econometric point of view the main problems that the analyst is facing are:

- Limited sample size; the number of observations used by the regression models is in one-to-one correspondence to the rating categories.
- The data is incomplete in the sense that the data sample used for calibration may not contain any observed defaults for obligors with some given ratings such as Aaa.
- The response variable has support in the interval [0,1]. The model described in the previous section does not satisfy this requirement and it may lead to default probabilities greater than 100%.
- Last but not least, the explanatory variable employed for calibration is ordinal with 16 categories. This issue may prove quite thorny to deal with.

From an econometric point of view the main problems that the analyst is facing are:

- Limited sample size; the number of observations used by the regression models is in one-to-one correspondence to the rating categories.
- The data is incomplete in the sense that the data sample used for calibration may not contain any observed defaults for obligors with some given ratings such as Aaa.
- The response variable has support in the interval [0,1]. The model described in the previous section does not satisfy this requirement and it may lead to default probabilities greater than 100%.
- Last but not least, the explanatory variable employed for calibration is ordinal with 16 categories. This issue may prove quite thorny to deal with.

From an econometric point of view the main problems that the analyst is facing are:

- Limited sample size; the number of observations used by the regression models is in one-to-one correspondence to the rating categories.
- The data is incomplete in the sense that the data sample used for calibration may not contain any observed defaults for obligors with some given ratings such as Aaa.
- The response variable has support in the interval [0,1]. The model described in the previous section does not satisfy this requirement and it may lead to default probabilities greater than 100%.
- Last but not least, the explanatory variable employed for calibration is ordinal with 16 categories. This issue may prove quite thorny to deal with.

From an econometric point of view the main problems that the analyst is facing are:

- Limited sample size; the number of observations used by the regression models is in one-to-one correspondence to the rating categories.
- The data is incomplete in the sense that the data sample used for calibration may not contain any observed defaults for obligors with some given ratings such as Aaa.
- The response variable has support in the interval [0,1]. The model described in the previous section does not satisfy this requirement and it may lead to default probabilities greater than 100%.
- Last but not least, the explanatory variable employed for calibration is ordinal with 16 categories. This issue may prove quite thorny to deal with.

- I propose a corrected model for calibration obtained by transform- ing the response variable MD on a different scale.

$$logit(MD) \equiv \ln\left(\frac{MD}{1-MD}\right) = \alpha + \beta \times Rating \tag{20}$$

- This is a logistic regression model that can be fitted easily to data.
- The goodness-of-fit of this model looks very good, with an adjusted R^2 of 82.3%.
- The regression coefficients are highly significant so that for prediction purposes one may use their estimates $\hat{\alpha} = -15.1673$ and $\hat{\beta} = 0.5058$, respectively.

- I propose a corrected model for calibration obtained by transform- ing the response variable MD on a different scale.

$$logit(MD) \equiv \ln\left(\frac{MD}{1-MD}\right) = \alpha + \beta \times Rating \qquad (20)$$

- This is a logistic regression model that can be fitted easily to data.
- The goodness-of-fit of this model looks very good, with an adjusted R^2 of 82.3%.
- The regression coefficients are highly significant so that for prediction purposes one may use their estimates $\widehat{\alpha} = -15.1673$ and $\widehat{\beta} = 0.5058$, respectively.

- I propose a corrected model for calibration obtained by transform- ing the response variable MD on a different scale.

$$logit(MD) \equiv \ln\left(\frac{MD}{1-MD}\right) = \alpha + \beta \times Rating \quad (20)$$

- This is a logistic regression model that can be fitted easily to data.
- The goodness-of-fit of this model looks very good, with an adjusted R^2 of 82.3%.
- The regression coefficients are highly significant so that for prediction purposes one may use their estimates $\widehat{\alpha} = -15.1673$ and $\widehat{\beta} = 0.5058$, respectively.

- I propose a corrected model for calibration obtained by transform- ing the response variable MD on a different scale.

$$logit(MD) \equiv \ln\left(\frac{MD}{1-MD}\right) = \alpha + \beta \times Rating \tag{20}$$

- This is a logistic regression model that can be fitted easily to data.
- The goodness-of-fit of this model looks very good, with an adjusted R^2 of 82.3%.
- The regression coefficients are highly significant so that for prediction purposes one may use their estimates $\widehat{\alpha} = -15.1673$ and $\widehat{\beta} = 0.5058$, respectively.

Table: Corporate default probabilities implied by the logistic linear regression model for corporates rated by Moody's over the period 1993 to 2000.

Rating	MD	s.d.	Estimated default probability
Aaa	NA	NA	0.004%
Aa1	NA	NA	0.007%
Aa2	NA	NA	0.012%
Aa3	0.08%	0.33%	0.020%
A1	NA	NA	0.032%
A2	NA	NA	0.054%
A3	NA	NA	0.089%
Baa1	0.06%	0.19%	0.148%
Baa2	0.06%	0.20%	0.245%
Baa3	0.46%	1.16%	0.407%
Ba1	0.69%	1.03%	0.675%
Ba2	0.63%	0.86%	1.119%
Ba3	2.39%	2.35%	1.855%
B1	3.79%	2.49%	3.075%
B2	7.96%	6.08%	5.098%
B3	12.89%	8.14%	8.451%

all default probabilities implied by the log-linear model are larger than the corresponding ones produced with the logistic model.

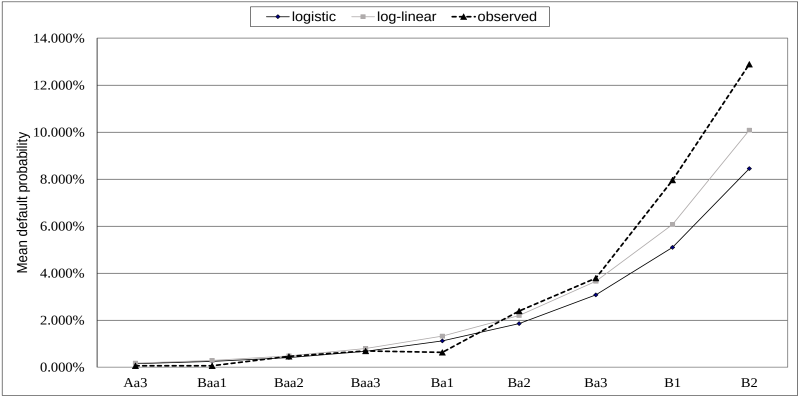


Figure: Comparison of mean default probabilities: observed versus log-linear and logistic models for corporates rated by Moody's over the period 1993 to 2000.

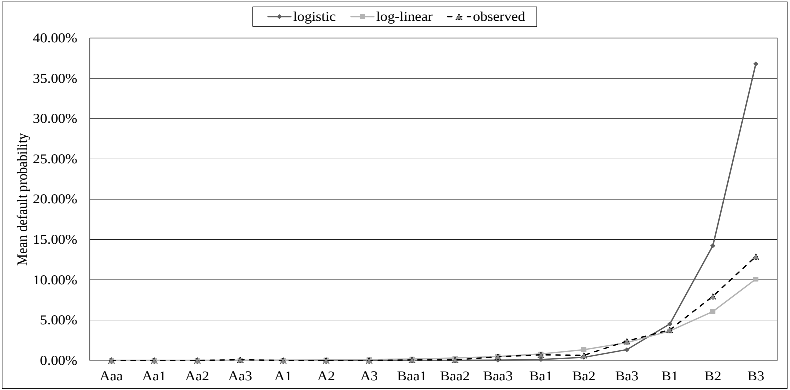


Figure: Comparison of calibration results for default probabilities: log-linear and logistic models versus observed. All credit ratings are used for corporates rated by Moody's over the period 1993 to 2000.

Further Analysis I

- What can an analyst do when data for calibration is sparse and he cannot increase the number of observations as desired?
- In addition the analyst may wish to consider some subjective information that may prove to be important and this is very difficult, if not impossible, within maximum likelihood or generalised least squares estimation and testing frameworks.
- An answer to both problems is to develop models under a Bayesian framework.
- consider the logistic regression model. Y denotes the mean default frequency between 1983 and 2000 while X represents the credit ratings, taking values from 1 to 16 in a one-to-one correspondence to the credit

Further Analysis II

ratings Aaa to B3 as used by Moody's. The Bayesian logistic model is specified here hierarchically:

$$\begin{aligned}
\ln\left(\frac{Y_i}{1-Y_i}\right)|\alpha,\beta,\tau &\sim N(\mu_i,\tau), \quad i=1,2,\ldots,16. \\
\mu_i &= \alpha+\beta X_i \\
\alpha &\sim N(0,0.001), \quad \beta \sim N(0,0.001) \\
\tau &\sim Gamma(3,1). \\
\sigma &= \sqrt{1/\tau}
\end{aligned}$$

- Expert opinion can be incorporated into this type of modelling by imposing more concentrated priors. For example, a downturn in the economy that may lead to a general increase in the defaults for all rating categories is equivalent to an upward shift of the intercept and maybe also of the slope of the logistic curve. These changes can be inserted into the model by changing the priors for α and β to plausible ranges.

Further Analysis III

- The inference is extracted here based on a sample of 10000 iterations after convergence criteria are passed.
-

	mean	s.d.	MC error	2.5%	median	97.5%
α	-16.48	1.138	0.008015	-18.7	-16.49	-14.2
β	0.9419	0.117	8.29E-4	0.7065	0.9428	1.168
σ	2.143	0.3604	0.002773	1.576	2.097	2.968

- the next Bayesian model investigated here is based on a $log(-log)$ transformation. This model is again specified hierarchically as

$$\begin{aligned} log(-log(Y_i))|\alpha,\beta,\tau &\sim N(\mu_i,\tau), \quad i=1,2,\dots,16. && (21)\\ \mu_i &= \alpha+\beta X_i && (22)\\ \alpha &\sim N(0,0.001), \quad \beta\sim N(0,0.001) && (23)\\ \tau &\sim lognormal(0,0.001) && (24)\\ \sigma &= \sqrt{1/\tau}. && (25) \end{aligned}$$

Further Analysis IV

	mean	s.d.	MC error	2.5%	median	97.5%
α	3.067	0.156	8.03E-4	2.754	3.068	3.379
β	-0.1323	0.017	7.69E-5	-0.164	-0.1323	-0.1001
σ	0.2897	0.059	3.47E-4	0.2013	0.2808	0.4312
- the two models can be compared using the Deviance Information Criterion (DIC) developed by Spiegelhalter et al. (2002) as a yardstick. This measure takes into consideration the model complexity and is based on the posterior of the deviance, that is $-2\times$likelihood, plus the effective number of parameters (pD), defined as the posterior mean of the deviance $\overline{D}$ minus the deviance of the posterior means $\hat{D}$. The model with the smallest Deviance Information Criterion is estimated to be the model that would best predict a replicate dataset of the same structure as that currently observed.

Model	$\overline{D}$	$\hat{D}$	pD	DIC
logistic	74.835	72.054	2.781	77.617
log(-log)	5.146	1.988	3.158	8.304

- it seems that the $log(-log)$ regression model provides a better fit.

Further Analysis V

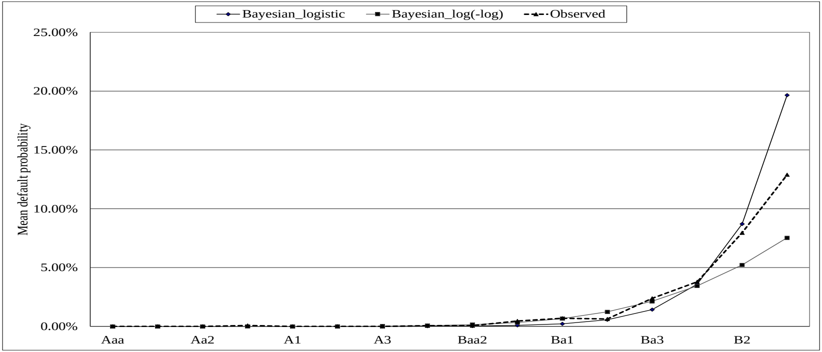


Figure: Comparison of calibration results for default probabilities: Bayesian log-log and Bayesian logistic models versus observed. All credit ratings used are for corporates rated by Moody's over the period 1993 to 2000.

Hierarchical Bayesian Models for Credit Risk

- The hierarchical Bayesian models presented below continue the line of the non–Bayesian specifications from Kao and Wu (1990), Terza (1987) and Hsiao (1983).
- The ordered probit model has been used, among others, by Nickell et al. (2000) to explore the dependence of rating transition probabilities on business cycles and on other characteristics of the borrowers, and by Cheung (1996) to explain rating levels based on indebtedness.
- Gossl (2005) proposed an extension of Merton's model for credit default fitted in a Bayesian framework, being able to capture correlations between default probabilities of obligors from different rating classes.
- Bayesian models,while offering similar benefits by estimating the entire joint posterior distribution of default probabilities, are different in that they model *explicitly* the impact of rating on default probabilities.

- Consider a population of B borrowers indexed by $j = 1, \ldots, B$. Let Z_j be a binary variable taking the value 0 if borrower j defaulted, and 1 otherwise. Let $\mathbf{B}(X)_j \in \mathbb{R}^d$ be a covariate vector for borrower j.
- The covariate information is usually borrower specific (for example, ratings), but it could also consist of general economic indicators.
- the first component of each $\mathbf{B}(X)$ will typically be one, denoting the presence of an intercept term. We shall denote the probability of default by $p(\mathbf{B}(X)) = \Pr(Z = 0; \mathbf{B}(X))$.

- Consider a population of B borrowers indexed by $j = 1, \ldots, B$. Let Z_j be a binary variable taking the value 0 if borrower j defaulted, and 1 otherwise. Let $\mathbf{B}(X)_j \in \mathbb{R}^d$ be a covariate vector for borrower j.
- The covariate information is usually borrower specific (for example, ratings), but it could also consist of general economic indicators.
- the first component of each $\mathbf{B}(X)$ will typically be one, denoting the presence of an intercept term. We shall denote the probability of default by $p(\mathbf{B}(X)) = \Pr(Z = 0; \mathbf{B}(X))$.

- Consider a population of B borrowers indexed by $j = 1, \ldots, B$. Let Z_j be a binary variable taking the value 0 if borrower j defaulted, and 1 otherwise. Let $\mathbf{B}(X)_j \in \mathbb{R}^d$ be a covariate vector for borrower j.
- The covariate information is usually borrower specific (for example, ratings), but it could also consist of general economic indicators.
- the first component of each $\mathbf{B}(X)$ will typically be one, denoting the presence of an intercept term. We shall denote the probability of default by $p(\mathbf{B}(X)) = \Pr(Z = 0; \mathbf{B}(X))$.

- A common approach is to assume the existence of an underlying (or latent) continuous variable for the ordinal indicator — for example, this gives rise to the probit model when the ordinal indicator is the dependent variable.
- In the case when the ordinal indicator is the explanatory variable, this is often either replaced by a set of dummy variables, or used itself as a regressor.
- Kukuk (2002) shows that both approaches could lead to wrong answers when assessing whether the corresponding continuous latent variable has a significant influence on the dependent variable or not.
- Hsiao and Mountain (1985) proposed a linear model with ordinal covariates based on latent variables with known thresholds, such as is the case, for example, with grouped income data.
- Ronning and Kukuk (1996) relaxed this strong assumption and further discussed a model where both dependent and explanatory variables are ordinal, where the unknown thresholds are estimated jointly with the structural parameters in a two-stage procedure.

- A common approach is to assume the existence of an underlying (or latent) continuous variable for the ordinal indicator — for example, this gives rise to the probit model when the ordinal indicator is the dependent variable.
- In the case when the ordinal indicator is the explanatory variable, this is often either replaced by a set of dummy variables, or used itself as a regressor.
- Kukuk (2002) shows that both approaches could lead to wrong answers when assessing whether the corresponding continuous latent variable has a significant influence on the dependent variable or not.
- Hsiao and Mountain (1985) proposed a linear model with ordinal covariates based on latent variables with known thresholds, such as is the case, for example, with grouped income data.
- Ronning and Kukuk (1996) relaxed this strong assumption and further discussed a model where both dependent and explanatory variables are ordinal, where the unknown thresholds are estimated jointly with the structural parameters in a two–stage procedure.

- A common approach is to assume the existence of an underlying (or latent) continuous variable for the ordinal indicator — for example, this gives rise to the probit model when the ordinal indicator is the dependent variable.
- In the case when the ordinal indicator is the explanatory variable, this is often either replaced by a set of dummy variables, or used itself as a regressor.
- Kukuk (2002) shows that both approaches could lead to wrong answers when assessing whether the corresponding continuous latent variable has a significant influence on the dependent variable or not.
- Hsiao and Mountain (1985) proposed a linear model with ordinal covariates based on latent variables with known thresholds, such as is the case, for example, with grouped income data.
- Ronning and Kukuk (1996) relaxed this strong assumption and further discussed a model where both dependent and explanatory variables are ordinal, where the unknown thresholds are estimated jointly with the structural parameters in a two–stage procedure.

- A common approach is to assume the existence of an underlying (or latent) continuous variable for the ordinal indicator — for example, this gives rise to the probit model when the ordinal indicator is the dependent variable.
- In the case when the ordinal indicator is the explanatory variable, this is often either replaced by a set of dummy variables, or used itself as a regressor.
- Kukuk (2002) shows that both approaches could lead to wrong answers when assessing whether the corresponding continuous latent variable has a significant influence on the dependent variable or not.
- Hsiao and Mountain (1985) proposed a linear model with ordinal covariates based on latent variables with known thresholds, such as is the case, for example, with grouped income data.
- Ronning and Kukuk (1996) relaxed this strong assumption and further discussed a model where both dependent and explanatory variables are ordinal, where the unknown thresholds are estimated jointly with the structural parameters in a two–stage procedure.

- A common approach is to assume the existence of an underlying (or latent) continuous variable for the ordinal indicator — for example, this gives rise to the probit model when the ordinal indicator is the dependent variable.
- In the case when the ordinal indicator is the explanatory variable, this is often either replaced by a set of dummy variables, or used itself as a regressor.
- Kukuk (2002) shows that both approaches could lead to wrong answers when assessing whether the corresponding continuous latent variable has a significant influence on the dependent variable or not.
- Hsiao and Mountain (1985) proposed a linear model with ordinal covariates based on latent variables with known thresholds, such as is the case, for example, with grouped income data.
- Ronning and Kukuk (1996) relaxed this strong assumption and further discussed a model where both dependent and explanatory variables are ordinal, where the unknown thresholds are estimated jointly with the structural parameters in a two–stage procedure.

- This approach relies on a set of distributional assumptions for the variables, and estimates can be biased if these assumptions are not met.
- Here latent variables with unknown thresholds are used for modelling the ordinal covariate indicators in a hierarchical Bayesian framework.
- Without loss of generality it is assumed that the covariate information is solely given by the rating category.
- Let n be the number of rating categories, and let C_j be the rating category for borrower j, with $j = 1, \ldots, B$. The random variables C_j are ordinal and observable for each borrower, so that the covariate vector for j is given by $\mathbf{B}(X)_j = (1, C_j)$. Our goal is to model the probability of default in each rating category i, defined by $p_i = Pr(Z = 0 | C = i)$, for $i = 1, \ldots, n$.
- The main assumption is that the category variable C is an indicator of the event that some unobservable continuous variable, say R, lies between certain thresholds. Specifically, let $\gamma_1, \gamma_2, \ldots, \gamma_{n+1}$ be a set of unknown thresholds.

- This approach relies on a set of distributional assumptions for the variables, and estimates can be biased if these assumptions are not met.
- Here latent variables with unknown thresholds are used for modelling the ordinal covariate indicators in a hierarchical Bayesian framework.
- Without loss of generality it is assumed that the covariate information is solely given by the rating category.
- Let n be the number of rating categories, and let C_j be the rating category for borrower j, with $j = 1, \dots, B$. The random variables C_j are ordinal and observable for each borrower, so that the covariate vector for j is given by $\mathbf{B}(X)_j = (1, C_j)$. Our goal is to model the probability of default in each rating category i, defined by $p_i = \Pr(Z = 0 | C = i)$, for $i = 1, \dots, n$.
- The main assumption is that the category variable C is an indicator of the event that some unobservable continuous variable, say R, lies between certain thresholds. Specifically, let $\gamma_1, \gamma_2, \dots, \gamma_{n+1}$ be a set of unknown thresholds.

- This approach relies on a set of distributional assumptions for the variables, and estimates can be biased if these assumptions are not met.
- Here latent variables with unknown thresholds are used for modelling the ordinal covariate indicators in a hierarchical Bayesian framework.
- Without loss of generality it is assumed that the covariate information is solely given by the rating category.
- Let n be the number of rating categories, and let C_j be the rating category for borrower j, with $j = 1, \ldots, B$. The random variables C_j are ordinal and observable for each borrower, so that the covariate vector for j is given by $\mathbf{B}(X)_j = (1, C_j)$. Our goal is to model the probability of default in each rating category i, defined by $p_i = \Pr(Z = 0|C = i)$, for $i = 1, \ldots, n$.
- The main assumption is that the category variable C is an indicator of the event that some unobservable continuous variable, say R, lies between certain thresholds. Specifically, let $\gamma_1, \gamma_2, \ldots, \gamma_{n+1}$ be a set of unknown thresholds.

- This approach relies on a set of distributional assumptions for the variables, and estimates can be biased if these assumptions are not met.
- Here latent variables with unknown thresholds are used for modelling the ordinal covariate indicators in a hierarchical Bayesian framework.
- Without loss of generality it is assumed that the covariate information is solely given by the rating category.
- Let n be the number of rating categories, and let C_j be the rating category for borrower j, with $j = 1, \ldots, B$. The random variables C_j are ordinal and observable for each borrower, so that the covariate vector for j is given by $\mathbf{B}(X)_j = (1, C_j)$. Our goal is to model the probability of default in each rating category i, defined by $p_i = \Pr(Z = 0 | C = i)$, for $i = 1, \ldots, n$.
- The main assumption is that the category variable C is an indicator of the event that some unobservable continuous variable, say R, lies between certain thresholds. Specifically, let $\gamma_1, \gamma_2, \ldots, \gamma_{n+1}$ be a set of unknown thresholds.

- This approach relies on a set of distributional assumptions for the variables, and estimates can be biased if these assumptions are not met.
- Here latent variables with unknown thresholds are used for modelling the ordinal covariate indicators in a hierarchical Bayesian framework.
- Without loss of generality it is assumed that the covariate information is solely given by the rating category.
- Let n be the number of rating categories, and let C_j be the rating category for borrower j, with $j = 1, \ldots, B$. The random variables C_j are ordinal and observable for each borrower, so that the covariate vector for j is given by $\mathbf{B}(X)_j = (1, C_j)$. Our goal is to model the probability of default in each rating category i, defined by $p_i = \Pr(Z = 0 | C = i)$, for $i = 1, \ldots, n$.
- The main assumption is that the category variable C is an indicator of the event that some unobservable continuous variable, say R, lies between certain thresholds. Specifically, let $\gamma_1, \gamma_2, \ldots, \gamma_{n+1}$ be a set of unknown thresholds.

- A corporate bond issue belongs to a certain risk category, say $C = i$, if the latent variable R falls in the interval (γ_i, γ_{i+1}).
- It is expected that the issuers in a given risk category i will exhibit roughly the same expected default risk. The widths of the risk category intervals need not be equal, and in practice the interval for Aaa bonds may have a different length than the interval for Bb bonds.
- For $i = 1, \ldots, n$, let m_i be the number of issuers and Y_i the number of defaults in rating category i. We shall consider the following model:

$$\begin{aligned} Y_i \mid m_i, p_i &\sim \mathrm{Binomial}(m_i, p_i), \quad i = 1, \ldots, n \\ \mathrm{logit}(p_i) &= \beta_0 + \beta_1 R_i + b_i \\ R_i \mid \gamma_i, \gamma_{i+1} &\sim U(\gamma_i, \gamma_{i+1}) \\ b_i \mid \sigma^2 &\sim N(0, \sigma^2) \\ \gamma_{i+1} &= \gamma_i + z_i \end{aligned} \tag{26}$$

- A corporate bond issue belongs to a certain risk category, say $C = i$, if the latent variable R falls in the interval (γ_i, γ_{i+1}).
- It is expected that the issuers in a given risk category i will exhibit roughly the same expected default risk. The widths of the risk category intervals need not be equal, and in practice the interval for Aaa bonds may have a different length than the interval for Bb bonds.
- For $i = 1, \dots, n$, let m_i be the number of issuers and Y_i the number of defaults in rating category i. We shall consider the following model:

$$
\begin{aligned}
Y_i \mid m_i, p_i &\sim \mathrm{Binomial}(m_i, p_i), \quad i = 1, \dots, n \qquad (26)\\
\mathrm{logit}(p_i) &= \beta_0 + \beta_1 R_i + b_i \\
\\
R_i \mid \gamma_i, \gamma_{i+1} &\sim U(\gamma_i, \gamma_{i+1}) \\
b_i \mid \sigma^2 &\sim N(0, \sigma^2) \\
\gamma_{i+1} &= \gamma_i + z_i
\end{aligned}
$$

- A corporate bond issue belongs to a certain risk category, say $C = i$, if the latent variable R falls in the interval (γ_i, γ_{i+1}).
- It is expected that the issuers in a given risk category i will exhibit roughly the same expected default risk. The widths of the risk category intervals need not be equal, and in practice the interval for Aaa bonds may have a different length than the interval for Bb bonds.
- For $i = 1, \ldots, n$, let m_i be the number of issuers and Y_i the number of defaults in rating category i. We shall consider the following model:

$$\begin{aligned} Y_i \,|\, m_i, p_i &\sim \mathrm{Binomial}(m_i, p_i), \quad i = 1, \ldots, n \\ \mathrm{logit}(p_i) &= \beta_0 + \beta_1 R_i + b_i \\ R_i \,|\, \gamma_i, \gamma_{i+1} &\sim U(\gamma_i, \gamma_{i+1}) \\ b_i \,|\, \sigma^2 &\sim N(0, \sigma^2) \\ \gamma_{i+1} &= \gamma_i + z_i \end{aligned} \tag{26}$$

- The random variables R_i have a uniform distribution on (γ_i, γ_{i+1}) and represent the latent effect of category ratings.
- The uniform is a special case of the generalized beta distribution;
- The random effects b_i are assumed to have a Gaussian distribution with mean zero and standard deviation σ.
- The increments z_i between the unknown thresholds must be positive, and in practice they will be given a gamma prior distribution as described in the following section.
- for the probit link we replace $\text{logit}(p_i) = \beta_0 + \beta_1 R_i + b_i$ in (26) with

$$p_i = \Phi(\beta_0 + \beta_1 R_i + b_i), \tag{27}$$

- for the log(-log) link,

$$\log(-\log(p_i)) = \beta_0 + \beta_1 R_i + b_i. \tag{28}$$

- The random variables R_i have a uniform distribution on (γ_i, γ_{i+1}) and represent the latent effect of category ratings.
- The uniform is a special case of the generalized beta distribution;
- The random effects b_i are assumed to have a Gaussian distribution with mean zero and standard deviation σ.
- The increments z_i between the unknown thresholds must be positive, and in practice they will be given a gamma prior distribution as described in the following section.
- for the probit link we replace $\text{logit}(p_i) = \beta_0 + \beta_1 R_i + b_i$ in (26) with

$$p_i = \Phi(\beta_0 + \beta_1 R_i + b_i). \tag{27}$$

- for the log(-log) link,

$$\log(-\log(p_i)) = \beta_0 + \beta_1 R_i + b_i. \tag{28}$$

- The random variables R_i have a uniform distribution on (γ_i, γ_{i+1}) and represent the latent effect of category ratings.
- The uniform is a special case of the generalized beta distribution;
- The random effects b_i are assumed to have a Gaussian distribution with mean zero and standard deviation σ.
- The increments z_i between the unknown thresholds must be positive, and in practice they will be given a gamma prior distribution as described in the following section.
- for the probit link we replace $\mathrm{logit}(p_i) = \beta_0 + \beta_1 R_i + b_i$ in (26) with

$$p_i = \Phi(\beta_0 + \beta_1 R_i + b_i), \tag{27}$$

- for the log(-log) link,

$$\log(-\log(p_i)) = \beta_0 + \beta_1 R_i + b_i. \tag{28}$$

- The random variables R_i have a uniform distribution on (γ_i, γ_{i+1}) and represent the latent effect of category ratings.
- The uniform is a special case of the generalized beta distribution;
- The random effects b_i are assumed to have a Gaussian distribution with mean zero and standard deviation σ.
- The increments z_i between the unknown thresholds must be positive, and in practice they will be given a gamma prior distribution as described in the following section.
- for the probit link we replace $\text{logit}(p_i) = \beta_0 + \beta_1 R_i + b_i$ in (26) with

$$p_i = \Phi(\beta_0 + \beta_1 R_i + b_i), \tag{27}$$

- for the log(-log) link,

$$\log(-\log(p_i)) = \beta_0 + \beta_1 R_i + b_i. \tag{28}$$

- The random variables R_i have a uniform distribution on (γ_i, γ_{i+1}) and represent the latent effect of category ratings.
- The uniform is a special case of the generalized beta distribution;
- The random effects b_i are assumed to have a Gaussian distribution with mean zero and standard deviation σ.
- The increments z_i between the unknown thresholds must be positive, and in practice they will be given a gamma prior distribution as described in the following section.
- for the probit link we replace $\mathrm{logit}(p_i) = \beta_0 + \beta_1 R_i + b_i$ in (26) with

$$p_i = \Phi(\beta_0 + \beta_1 R_i + b_i), \tag{27}$$

- for the log(-log) link,

$$\log(-\log(p_i)) = \beta_0 + \beta_1 R_i + b_i. \tag{28}$$

- The random variables R_i have a uniform distribution on (γ_i, γ_{i+1}) and represent the latent effect of category ratings.
- The uniform is a special case of the generalized beta distribution;
- The random effects b_i are assumed to have a Gaussian distribution with mean zero and standard deviation σ.
- The increments z_i between the unknown thresholds must be positive, and in practice they will be given a gamma prior distribution as described in the following section.
- for the probit link we replace $\mathrm{logit}(p_i) = \beta_0 + \beta_1 R_i + b_i$ in (26) with

$$p_i = \Phi(\beta_0 + \beta_1 R_i + b_i), \tag{27}$$

- for the log(-log) link,

$$\log(-\log(p_i)) = \beta_0 + \beta_1 R_i + b_i. \tag{28}$$

- The hierarchical model may be estimated from sample data in a MCMC framework.
- Let us denote by θ the vector of parameters and hyperparameters of the model, given by $\theta = (\gamma_1, z_1, \ldots, z_n, \beta_0, \beta_1, \sigma^2)$.

$$p(\theta \mid \mathscr{Y}) \propto p(\mathscr{Y} \mid \theta) \cdot p(\theta) \tag{29}$$

- The first model with a logistic link function is described by the following general system of equations. For all $i = 1, 2, \ldots, n$

$$\begin{aligned}
Y_i &\sim Binomial(m_i, p_i) \qquad (30)\\
\mathrm{logit}(p_i) &= \beta_0 + \beta_1 R_i + b_i\\
b_i &\sim N(0, \sigma^2), \quad \sigma^2 = 1/\tau\\
R_i &= \gamma_i + z_i U_i, \quad \gamma_{i+1} = \gamma_i + z_i\\
z_i &\sim \mathrm{Gamma}(\delta, \delta), \quad U_i \sim U(0,1)\\
\gamma_1 &\sim \mathrm{Gamma}(\alpha, \alpha), \quad \tau \sim \mathrm{Gamma}(u, v)\\
\beta_0 &\sim N(0, \tau_0), \quad \beta_1 \sim N(0, \tau_1)
\end{aligned}$$

- For all parameters, 95% credible intervals can be computed from the samples of observations generated from the posterior densities, and these can be then used in testing specific hypotheses about the parameters.

- The hierarchical model may be estimated from sample data in a MCMC framework.
- Let us denote by θ the vector of parameters and hyperparameters of the model, given by $\theta = (\gamma_1, z_1, \ldots, z_n, \beta_0, \beta_1, \sigma^2)$.

$$p(\theta \,|\, \mathscr{Y}) \propto p(\mathscr{Y} \,|\, \theta) \cdot p(\theta) \tag{29}$$

- The first model with a logistic link function is described by the following general system of equations. For all $i = 1, 2, \ldots, n$

$$\begin{aligned}
Y_i &\sim Binomial(m_i, p_i) \\
\mathrm{logit}(p_i) &= \beta_0 + \beta_1 R_i + b_i \\
b_i &\sim N(0, \sigma^2), \quad \sigma^2 = 1/\tau \\
R_i &= \gamma_i + z_i U_i, \quad \gamma_{i+1} = \gamma_i + z_i \\
z_i &\sim \mathrm{Gamma}(\delta, \delta), \quad U_i \sim U(0,1) \\
\gamma_1 &\sim \mathrm{Gamma}(\alpha, \alpha), \quad \tau \sim \mathrm{Gamma}(u, v) \\
\beta_0 &\sim N(0, \tau_0), \quad \beta_1 \sim N(0, \tau_1)
\end{aligned} \tag{30}$$

- For all parameters, 95% credible intervals can be computed from the samples of observations generated from the posterior densities, and these can be then used in testing specific hypotheses about the parameters.

- The hierarchical model may be estimated from sample data in a MCMC framework.
- Let us denote by θ the vector of parameters and hyperparameters of the model, given by $\theta = (\gamma_1, z_1, \ldots, z_n, \beta_0, \beta_1, \sigma^2)$.

$$p(\theta \,|\, \mathscr{Y}) \propto p(\mathscr{Y} \,|\, \theta) \cdot p(\theta) \tag{29}$$

- The first model with a logistic link function is described by the following general system of equations. For all $i = 1, 2, \ldots, n$

$$\begin{aligned}
Y_i &\sim Binomial(m_i, p_i) \\
\text{logit}(p_i) &= \beta_0 + \beta_1 R_i + b_i \\
b_i &\sim N(0, \sigma^2), \ \ \sigma^2 = 1/\tau \\
R_i &= \gamma_i + z_i U_i, \ \ \gamma_{i+1} = \gamma_i + z_i \\
z_i &\sim \text{Gamma}(\delta, \delta), \ \ U_i \sim U(0,1) \\
\gamma_1 &\sim \text{Gamma}(\alpha, \alpha), \ \ \tau \sim \text{Gamma}(u, v) \\
\beta_0 &\sim N(0, \tau_0), \ \ \beta_1 \sim N(0, \tau_1)
\end{aligned} \tag{30}$$

- For all parameters, 95% credible intervals can be computed from the samples of observations generated from the posterior densities, and these can be then used in testing specific hypotheses about the parameters.

- The hierarchical model may be estimated from sample data in a MCMC framework.
- Let us denote by θ the vector of parameters and hyperparameters of the model, given by $\theta = (\gamma_1, z_1, \ldots, z_n, \beta_0, \beta_1, \sigma^2)$.

$$p(\theta \mid \mathscr{Y}) \propto p(\mathscr{Y} \mid \theta) \cdot p(\theta) \tag{29}$$

- The first model with a logistic link function is described by the following general system of equations. For all $i = 1, 2, \ldots, n$

$$\begin{aligned}
Y_i &\sim Binomial(m_i, p_i) \\
\mathrm{logit}(p_i) &= \beta_0 + \beta_1 R_i + b_i \\
b_i &\sim N(0, \sigma^2), \quad \sigma^2 = 1/\tau \\
R_i &= \gamma_i + z_i U_i, \quad \gamma_{i+1} = \gamma_i + z_i \\
z_i &\sim \mathrm{Gamma}(\delta, \delta), \quad U_i \sim U(0,1) \\
\gamma_1 &\sim \mathrm{Gamma}(\alpha, \alpha), \quad \tau \sim \mathrm{Gamma}(u, v) \\
\beta_0 &\sim N(0, \tau_0), \quad \beta_1 \sim N(0, \tau_1)
\end{aligned} \tag{30}$$

- For all parameters, 95% credible intervals can be computed from the samples of observations generated from the posterior densities, and these can be then used in testing specific hypotheses about the parameters.

Hierarchical model for aggregated data

- The first analysis of the Standard& Poor's data considers the aggregate number of defaults over the horizon 1981–2004.
- The credit risk industry is divided on the issue whether ratings are cross–sectionally independent. While this can be a matter for debate, here the view of Credit Suisse Financial Products (1997), Wilson (1997) and Nickell et al. (2000) is taken that independence can be assumed, at least in the first instance.
- Using standard Bayesian applied statistical modelling I utilise diffuse but proper priors for all parameters. Hence, $N(0, 10^3)$ priors are taken for the regression parameters β_0 and β_1. I also specified a gamma prior with large variance $Gamma(1, 0.1)$ for z_i, $i = 1, \ldots, 7$, and a diffuse inverse gamma prior $Inv - Gamma(1, 0.1)$ for the random effects variance σ^2.
- For each model two parallel chains were started with different sets of initial values, and the Gibbs sampler was run for 50,000 iterations with the first 20,000 iterations discarded as a burn–in period. Gelman and Rubin's diagnostic (Gelman et al. (1995)) indicated satisfactory convergence of the chains.

Hierarchical model for aggregated data

- The first analysis of the Standard& Poor's data considers the aggregate number of defaults over the horizon 1981–2004.
- The credit risk industry is divided on the issue whether ratings are cross–sectionally independent. While this can be a matter for debate, here the view of Credit Suisse Financial Products (1997), Wilson (1997) and Nickell et al. (2000) is taken that independence can be assumed, at least in the first instance.
- Using standard Bayesian applied statistical modelling I utilise diffuse but proper priors for all parameters. Hence, $N(0, 10^3)$ priors are taken for the regression parameters β_0 and β_1. I also specified a gamma prior with large variance $Gamma(1, 0.1)$ for z_i, $i = 1, \ldots, 7$, and a diffuse inverse gamma prior $Inv - Gamma(1, 0.1)$ for the random effects variance σ^2.
- For each model two parallel chains were started with different sets of initial values, and the Gibbs sampler was run for 50,000 iterations with the first 20,000 iterations discarded as a burn–in period. Gelman and Rubin's diagnostic (Gelman et al. (1995)) indicated satisfactory convergence of the chains.

Hierarchical model for aggregated data

- The first analysis of the Standard& Poor's data considers the aggregate number of defaults over the horizon 1981–2004.
- The credit risk industry is divided on the issue whether ratings are cross–sectionally independent. While this can be a matter for debate, here the view of Credit Suisse Financial Products (1997), Wilson (1997) and Nickell et al. (2000) is taken that independence can be assumed, at least in the first instance.
- Using standard Bayesian applied statistical modelling I utilise diffuse but proper priors for all parameters. Hence, $N(0, 10^3)$ priors are taken for the regression parameters β_0 and β_1. I also specified a gamma prior with large variance $Gamma(1, 0.1)$ for z_i, $i = 1, \ldots, 7$, and a diffuse inverse gamma prior $Inv - Gamma(1, 0.1)$ for the random effects variance σ^2.
- For each model two parallel chains were started with different sets of initial values, and the Gibbs sampler was run for 50,000 iterations with the first 20,000 iterations discarded as a burn–in period. Gelman and Rubin's diagnostic (Gelman et al. (1995)) indicated satisfactory convergence of the chains.

Hierarchical model for aggregated data

- The first analysis of the Standard& Poor's data considers the aggregate number of defaults over the horizon 1981–2004.
- The credit risk industry is divided on the issue whether ratings are cross–sectionally independent. While this can be a matter for debate, here the view of Credit Suisse Financial Products (1997), Wilson (1997) and Nickell et al. (2000) is taken that independence can be assumed, at least in the first instance.
- Using standard Bayesian applied statistical modelling I utilise diffuse but proper priors for all parameters. Hence, $N(0, 10^3)$ priors are taken for the regression parameters β_0 and β_1. I also specified a gamma prior with large variance $Gamma(1, 0.1)$ for z_i, $i = 1, \ldots, 7$, and a diffuse inverse gamma prior $Inv - Gamma(1, 0.1)$ for the random effects variance σ^2.
- For each model two parallel chains were started with different sets of initial values, and the Gibbs sampler was run for 50,000 iterations with the first 20,000 iterations discarded as a burn–in period. Gelman and Rubin's diagnostic (Gelman et al. (1995)) indicated satisfactory convergence of the chains.

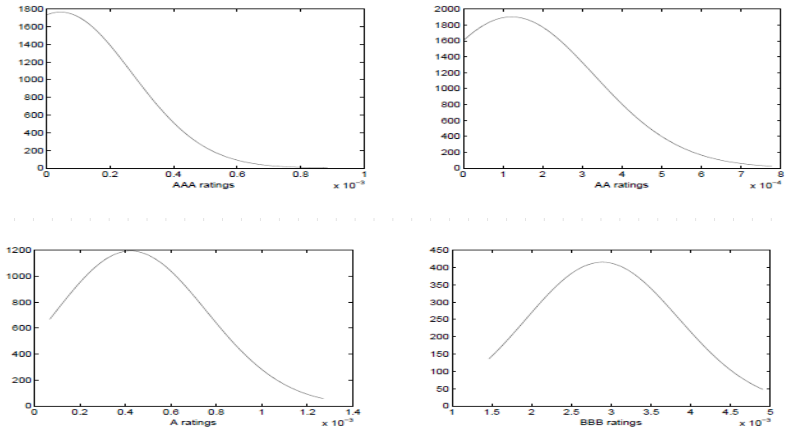


Figure: Posterior kernel density estimates for investment grade default probabilities using the logistic link model and the S&P data for the aggregate number of defaults over the horizon 1981–2004.

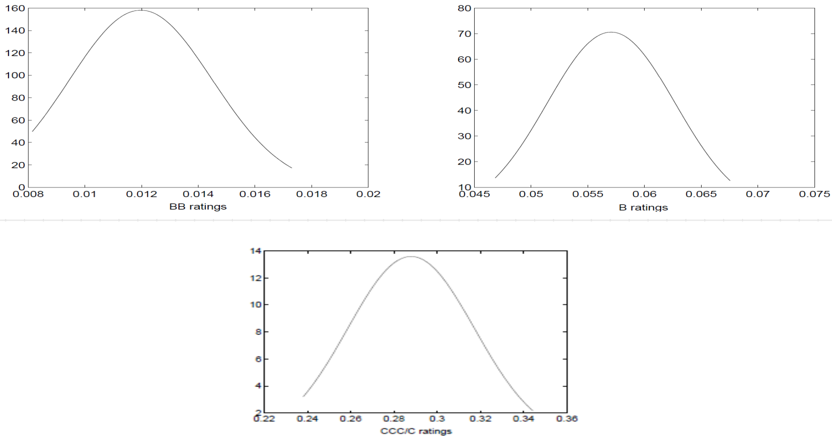


Figure: Posterior kernel density estimates for non-investment grade default probabilities using the logistic link model and the S&P data for the aggregate number of defaults over the horizon 1981–2004.

Table: Comparison of different models after Bayesian fitting

		Mean	standard deviation	Median	Credibility intervals $(2.5\% - -97.5\%)$
Logit link					
(DIC = 42.607)	β_0	-11.930	0.640	-11.930	(-13.060, -10.770)
	β_1	1.630	0.128	1.617	(1.391, 1.877)
	σ^2	0.389	0.195	0.3395	(0.164, 0.899)
Probit link					
(DIC = 42.974)	β_0	-5.031	0.375	-5.103	(-5.622, -4.261)
	β_1	0.654	0.051	0.647	(0.573, 0.759)
	σ^2	0.275	0.107	0.251	(0.140, 0.544)
Log–log link					
(DIC = 39.913)	β_0	-11.560	0.720	-11.61	(-13.140, -10.130)
	β_1	2.118	0.199	2.079	(1.809, 2.499)
	σ^2	0.373	0.177	0.330	(0.162, 0.819)

	Default prob	Mean	Standard deviation	Median	Credibility intervals $(2.5\% - -97.5\%)$
Logit link	p_1	5.453e-5	5.419e-5	3.974e-5	(4.22e-6, 1.85e-4)
	p_2	1.295e-4	8.628e-5	1.091e-4	(2.42e-5, 3.42e-4)
	p_3	4.295e-4	1.452e-4	4.113e-4	(1.93e-4, 7.46e-4)
	p_4	0.002888	4.327e-4	0.002868	(0.0021, 0.0038)
	p_5	0.01197	0.001034	0.01194	(0.0100, 0.0140)
	p_6	0.05708	0.002281	0.05704	(0.0528, 0.0616)
	p_7	0.2879	0.01258	0.2876	(0.2637, 0.3122)
Probit link	p_1	2.118e-5	4.185e-5	6.174e-6	(5.77e-8, 1.34e-4)
	p_2	1.135e-4	9.24e-5	9.023e-5	(9.76e-6, 3.45e-4)
	p_3	4.212e-4	1.428e-4	4.029e-4	(1.92e-4, 7.48e-4)
	p_4	0.002932	4.374e-4	0.002912	(0.0021, 0.0038)
	p_5	0.01202	0.001055	0.01196	(0.0100, 0.0142)
	p_6	0.05706	0.002284	0.05702	(0.0526, 0.0617)
	p_7	0.2876	0.01287	0.2874	(0.2625, 0.313)
Log–log link	p_1	4.161e-5	3.393e-5	3.225e-5	(6.46e-6, 1.31e-4)
	p_2	1.226e-4	7.706e-5	1.041e-4	(2.73e-5, 3.22e-4)
	p_3	4.67e-4	1.535e-4	4.508e-4	(2.156e-4, 8.13e-4)
	p_4	0.002871	4.253e-4	0.002854	(0.0021, 0.0037)
	p_5	0.01197	0.001038	0.01195	(0.0099, 0.0141)
	p_6	0.05708	0.002223	0.05705	(0.0528, 0.0615)
	p_7	0.2878	0.01263	0.2877	(0.2634, 0.3126)

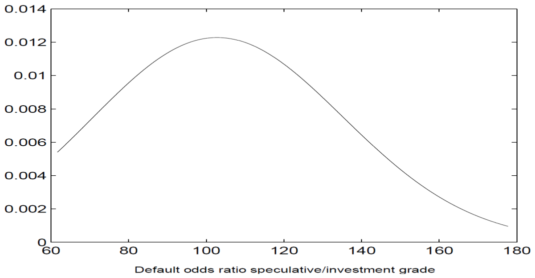


Figure: Posterior kernel density estimate for the ratio between the cumulative default probability in the speculative grade categories and the cumulative default probability in the investment grade categories. Aggregated 1981–2004 data, logistic model.

Hierarchical time-series model I

- Since the yearly frequency of defaults in each rating category are available in the S&P ratings data, it is relevant to attempt an analysis that can take into account any serial correlation over consecutive years.
- Let T be the length of the time horizon (here $T = 24$), and let $t = 1, ..., T$ be the yearly observation times. I extend the notation to let Y_{it}, m_{it} and p_{it} denote respectively the number of defaults, number of issuers, and probability of default in rating category i at time t.

$$
\begin{aligned}
Y_{it} \,|\, m_{it}, p_{it} &\sim \mathrm{Binomial}(m_{it}, p_{it}), \quad i = 1, \ldots, n, \quad t = 1, \ldots, T \\
\mathrm{logit}(p_{it}) &= \beta_0 + \beta_1 R_i + b_{it} \\
b_{it} &= a b_{i(t-1)} + \varepsilon_{it} \qquad (31) \\
\\
R_i \,|\, \gamma_i, \gamma_{i+1} &\sim U(\gamma_i, \gamma_{i+1}) \\
\gamma_{i+1} &= \gamma_i + z_i \\
\varepsilon_{it} \,|\, \sigma^2 &\sim N(0, \sigma^2)
\end{aligned}
$$

Hierarchical time-series model II

- This model uses the parameter $a \in \mathbb{R}$ to account for a possible autoregressive correlation structure of the random terms b_{it}.
- The rating variables R_i do not depend on time and, as previously, have a uniform distribution on (γ_i, γ_{i+1}).
- The model was fit to the yearly data using non–informative prior distributions $N(0, 10^3)$ for β_0, β_1, and a, and inverse Gamma$(0.1, 0.1)$ for σ.

Table: Bayesian estimates for S&P yearly rating data on all corporates between 1981–2004; parameters of the time series model.

Parameter	Mean	SD	Median	95% Credibility Interval
β_0	-13.020	0.781	-13.000	(-14.42, -11.21)
β_1	1.834	0.120	1.829	(1.619, 2.085)
a	0.159	0.147	0.160	(-0.111, 0.452)
σ^2	0.697	0.077	0.694	(0.557, 0.857)

Hierarchical model for disaggregated data

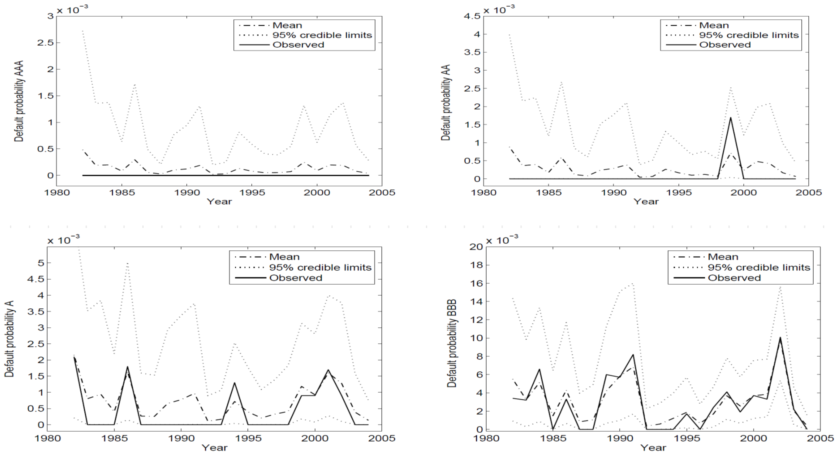


Figure: Observed values and posterior means with credible intervals for investment grade default probabilities, Standard & Poor's yearly rating data on all corporates between 1981–2004.

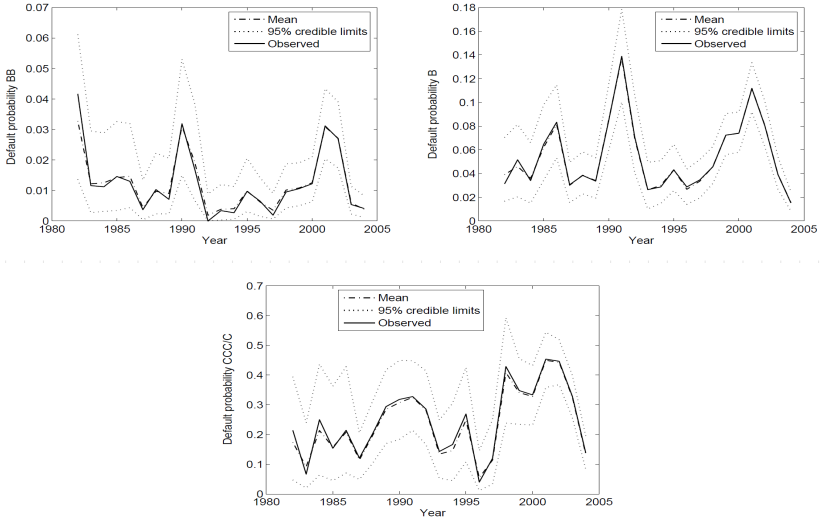


Figure: Observed values and posterior means with credible intervals for non-investment grade default probabilities, Standard & Poor's yearly rating data on all corporates between 1981–2004.

Table: Bayesian MCMC posterior estimates of correlations of probabilities of default, based on the logistic link model, Standard& Poor's yearly rating data on all corporates between 1981–2004.

	p_1	p_2	p_3	p_4	p_5	p_6	p_7
p_1	1.000	0.968	0.939	0.584	0.673	0.316	0.152
p_2	–	1.000	0.942	0.576	0.669	0.398	0.241
p_3	–	–	1.000	0.679	0.779	0.488	0.363
p_4	–	–	–	1.000	0.766	0.579	0.583
p_5	–	–	–	–	1.000	0.598	0.482
p_6	–	–	–	–	–	1.000	0.647
p_7	–	–	–	–	–	–	1.000

Further Credit Modelling with MCMC Callibration

- the Bayesian Credit Portfolio Model developed by Gossl (2005).
- Let X_i be the generic one-year asset return. For a portfolio of N credit risky instruments

$$\begin{aligned} X_i &= \sqrt{\rho} Y + \sqrt{1-\rho} Z_i, \quad i = 1, \ldots, N \\ \rho &\in [0,1], \quad Y \sim N(0,1), \quad Z_i \overset{\text{i.i.d.}}{\sim} N(0,1) \end{aligned}$$

ρ accounts for the intra-portfolio dependencies within the portfolio.

$$\mathbb{P}(X_i < k_i) = p_i \quad \Longrightarrow \quad k_i = \Phi^{-1}(p_i)$$

$$\mathbb{P}(X_i < k_i | Y = y) = p_{i|y} \quad \Longrightarrow \quad p_{i|y} = \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho} \times y}{\sqrt{1-\rho}}\right)$$

Further Credit Modelling with MCMC Callibration

- the Bayesian Credit Portfolio Model developed by Gossl (2005).
- Let X_i be the generic one-year asset return. For a portfolio of N credit risky instruments

$$\begin{aligned} X_i &= \sqrt{\rho}\,Y + \sqrt{1-\rho}\,Z_i, \quad i = 1, \ldots, N \\ \rho &\in [0,1], \quad Y \sim N(0,1), \quad Z_i \overset{\text{i.i.d.}}{\sim} N(0,1) \end{aligned}$$

ρ accounts for the intra-portfolio dependencies within the portfolio.

$$\begin{aligned} \mathbb{P}(X_i < k_i) = p_i \quad &\Longrightarrow \quad k_i = \Phi^{-1}(p_i) \\ \mathbb{P}(X_i < k_i | Y = y) = p_{i|y} \quad &\Longrightarrow \quad p_{i|y} = \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho} \times y}{\sqrt{1-\rho}}\right) \end{aligned}$$

For K rating classes and T years of data define

- $\mathbf{L}_t = (L_{t,1}, \dots, L_{t,K})$ is the vector of defaults at the end of year t
- $\mathbf{n}_t = (n_{t,1}, \dots, n_{t,K})$ is the vector of rated issuers at the beginning of year t
- $\mathbf{p} = (p_1, \dots, p_K)$ is the vector of the probability of defaults from each rating category

$$p(\mathbf{L}_t = \mathbf{l}_t | \mathbf{n}_t, \mathbf{p}, \rho, y_t) = \prod_{j=1}^{K} \mathrm{Binomial}(n_{t,j}, p_{j|y_t}, l_{t,j}) \tag{32}$$

$$p(\mathbf{p}, \rho, \mathbf{y} | \mathbf{n}, \mathbf{l}) \propto \prod_{t=1}^{T} p(\mathbf{L}_t = \mathbf{l}_t | \mathbf{n}_t, \mathbf{p}_{|\mathbf{y}}, \rho, y_t) p(\mathbf{p}) p(\rho) p(\mathbf{y})$$

$$p_j \overset{\text{i.i.d.}}{\sim} U(0,1), \qquad \rho \sim U(0,1), \quad y_t \overset{\text{i.i.d.}}{\sim} N(0,1) \tag{33}$$

For K rating classes and T years of data define

- $\mathbf{L}_t = (L_{t,1}, \ldots, L_{t,K})$ is the vector of defaults at the end of year t
- $\mathbf{n}_t = (n_{t,1}, \ldots, n_{t,K})$ is the vector of rated issuers at the beginning of year t
- $\mathbf{p} = (p_1, \ldots, p_K)$ is the vector of the probability of defaults from each rating category

$$p(\mathbf{L}_t = \mathbf{l}_t | \mathbf{n}_t, \mathbf{p}, \rho, y_t) = \prod_{j=1}^{K} \mathrm{Binomial}(n_{t,j}, p_{j|y_t}, l_{t,j}) \tag{32}$$

$$p(\mathbf{p}, \rho, \mathbf{y} | \mathbf{n}, \mathbf{l}) \propto \prod_{t=1}^{T} p(\mathbf{L}_t = \mathbf{l}_t | \mathbf{n}_t, \mathbf{p}_{|\mathbf{y}}, \rho, y_t) p(\mathbf{p}) p(\rho) p(\mathbf{y})$$

$$p_j \overset{\text{i.i.d.}}{\sim} U(0,1), \qquad \rho \sim U(0,1), \quad y_t \overset{\text{i.i.d.}}{\sim} N(0,1) \tag{33}$$

For K rating classes and T years of data define

- $\mathbf{L}_t = (L_{t,1}, \ldots, L_{t,K})$ is the vector of defaults at the end of year t
- $\mathbf{n}_t = (n_{t,1}, \ldots, n_{t,K})$ is the vector of rated issuers at the beginning of year t
- $\mathbf{p} = (p_1, \ldots, p_K)$ is the vector of the probability of defaults from each rating category

$$p(\mathbf{L}_t = \mathbf{l}_t | \mathbf{n}_t, \mathbf{p}, \rho, y_t) = \prod_{j=1}^{K} \mathrm{Binomial}(n_{t,j}, p_{j|y_t}, l_{t,j}) \quad (32)$$

$$p(\mathbf{p}, \rho, \mathbf{y} | \mathbf{n}, \mathbf{l}) \propto \prod_{t=1}^{T} p(\mathbf{L}_t = \mathbf{l}_t | \mathbf{n}_t, \mathbf{p}_{|\mathbf{y}}, \rho, y_t) p(\mathbf{p}) p(\rho) p(\mathbf{y})$$

$$p_j \overset{\text{i.i.d.}}{\sim} U(0,1), \qquad \rho \sim U(0,1), \quad y_t \overset{\text{i.i.d.}}{\sim} N(0,1) \quad (33)$$

Another model is a Bayesian Panel Count Data model

$$
\begin{aligned}
L_{t,j}|\mathbf{n},\mathbf{p_{t,j}} &\sim \mathrm{Binomial}(n_{t,j},p_{t,j}) &(34)\\
p_{t,j} &= \Phi\left(\frac{\Phi^{-1}(p_j)-\sqrt{\rho}\times y_t}{\sqrt{1-\rho}}\right) &(35)\\
p_j &= \Phi(\alpha+\beta\times j),\;\; y_t\sim N(0,1) &(36)\\
\rho\sim U(0,1),\;\; \alpha &\sim N(0,0.0001),\;\; \beta\sim N(0,0.0001) &(37)
\end{aligned}
$$

Table: Bayesian MCMC posterior inference for the Bayesian Panel Count Data model using S&P yearly rating data on all corporates between 1981–2004.

node	mean	s.d.	2.50%	median	97.50%
α	-5.727	0.1383	-5.996	-5.73	-5.451
β	0.6989	0.02067	0.6581	0.6989	0.7395
p[1]	2.99E-07	2.00E-07	7.18E-08	2.44E-07	8.40E-07
p[2]	8.30E-06	4.05E-06	3.00E-06	7.37E-06	1.86E-05
p[3]	1.49E-04	5.17E-05	7.38E-05	1.40E-04	2.74E-04
p[4]	0.0017	4.17E-04	0.001072	0.0017	0.0027
p[5]	0.0129	0.002154	0.009276	0.0127	0.0178
p[6]	0.0629	0.00745	0.04949	0.0623	0.0792
p[7]	0.2024	0.01778	0.1687	0.2018	0.2393
ρ	0.07335	0.02403	0.03841	0.06918	0.1328

- The evolution of the yearly factor y_t between 1981 and 2004, negative values of the yearly factor are associated with an overall increase of credit risk whereas positive values of the yearly factor correspond to a view less risky credit environment.
- as of 2004 the increasing positive local trend between 2002 and 2004 may increase for few more years but history shows that there will be a mean reversion not far away.

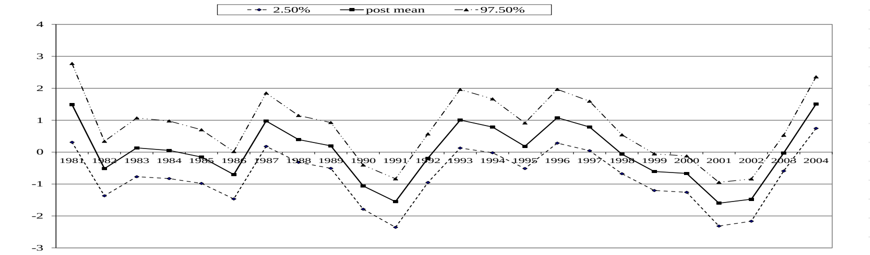


Figure: The posterior mean and credible interval for the yearly factor y_t of the Bayesian Panel Count Data model and Standard&Poor's data between 1981 and 2004.

- the posterior correlation across the estimated probabilities of default.
- there is strong correlation between the adjacent parameters as indicated by the dark region around the main diagonal.
- There is also stronger correlation between superior ratings, that is between probabilities of default p_1, p_2, p_3.
- There is also very weak correlation of estimates of probabilities of default across the second diagonal.

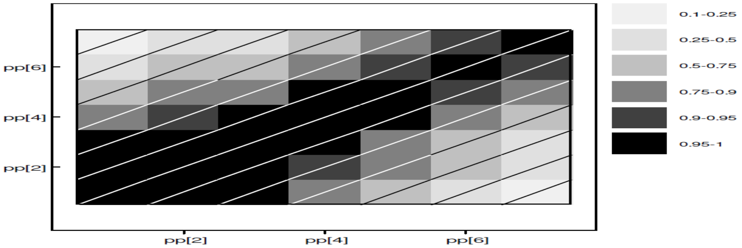


Figure: Correlation matrix of probabilities of default $p_1, p_2, \ldots, p_7$ corresponding to the Standard&Poor's seven rating categories: AAA, AA, A, BBB, BB, B, and CCC/C. The data used is on all corporates between 1981–2004.

Estimating the Transition Matrix

Estimating the rating transition matrix is important in credit markets and it is notoriously difficult, see Nickell et al. (2000), Berd (2005) and Engelman and Ermakov (2011). Denoting by

- $\mathbf{L}_{t,i} = (L_{t,i,1}, \ldots, L_{t,i,K})$ the vector of the number of assets that moved over year t from rating i, and
- $\mathbf{p}_{t,i} = (p_{t,i,1}, \ldots, p_{t,i,K})$ the vector of the probability of defaults from each rating category:

$$p(\mathbf{L}_{t,i} = \mathbf{l}_{t,i} | \mathbf{n}_{t,i}, \mathbf{p}_{t,i}) = \mathrm{Multi}(\mathbf{n}_{t,i}, \mathbf{p}_{t,i}) \tag{38}$$

$$\begin{aligned} p_{t,i,j} &= \frac{\alpha_{t,i,j}}{\sum_{j=1}^{K} \alpha_{t,i,j}}, \quad \alpha_{t,i,j} = e^{a_{t,i,j}} & (39)\\ a_{t,i,j} &= b_{i,j} \frac{\delta_{i,j}}{(1+|i-j|)}, \quad a_{t,i,i} = 0 & (40)\\ b_{i,j} &\sim N(0, 0.001) & (41) \end{aligned}$$

Table: Posterior means of parameters b. Their value on the main diagonal is not relevant because of the identifiability constraint.

$b_{i,j}$	AAA	AA	A	BBB	BB	B	CCC	D
AAA		-4.952	-15.82	-28.23	-37.35	-62.25	-68.68	-74.78
AA	-9.979		-4.821	-15.06	-29.52	-33.69	-50.97	-63.77
A	-22.69	-7.494		-5.523	-15.97	-25.29	-39.82	-45.95
BBB	-34	-18.04	-6.171		-5.905	-14.08	-24.47	-28.34
BB	-38.4	-27.83	-16.35	-5.332		-4.664	-13.17	-16.6
B	-66.94	-35.29	-23.74	-16.72	-5.28		-5.694	-7.646
CCC	-45.65	-54.2	-25.56	-19.55	-10.8	-3.145		-0.9696

Table: Posterior medians of transition probabilities using data from S&P between 1981-2004.

$p_{i,j}$	AAA	AA	A	BBB	BB	B	CCC
AAA	0.9160	0.0771	0.0047	8.3E-04	5.6E-04	4.4E-05	6.9E-05
AA	0.0062	0.9045	0.0812	0.0059	5.8E-04	0.0011	1.9E-04
A	4.8E-04	0.0216	0.9133	0.0577	0.0045	0.0016	3.2E-04
BBB	1.9E-04	0.0022	0.0409	0.8964	0.0468	0.0082	0.0019
BB	4.0E-04	8.1E-04	0.0036	0.0579	0.8327	0.0809	0.0103
B	1.7E-05	7.2E-04	0.0022	0.0031	0.0587	0.8229	0.0477
CCC	9.1E-04	9.8E-05	0.0034	0.0042	0.0148	0.1111	0.5348
Default	1.0E-04	1.1E-04	4.4E-04	0.0031	0.0131	0.0644	0.3292

MLE estimation

Bangia et al. (2002) and Hu et al. (2002) describe the derivation of the MLE for the credit transition matrix leading to the following formulae:

$$\widehat{p}_{i,j} = \sum_{t=1}^{T} w_i(t) \frac{l_{t,i,j}}{n_{i,t}} = \frac{\sum_{t=1}^{T} l_{t,i,j}}{\sum_{t=1}^{T} n_{i,t}} \qquad (42)$$

$$w_i(t) = \frac{n_{t,i}}{\sum_{t=1}^{T} n_{t,i}} \neq \frac{1}{T}, \quad \widehat{p}_{i,j} = \frac{1}{T} \sum_{t=1}^{T} \frac{l_{t,i,j}}{n_{t,i}} \qquad (43)$$

Table: MLE of transition probabilities on all corporates between 1981-2004.

$p_{i,j}$	AAA	AA	A	BBB	BB	B	CCC
AAA	0.91646	0.07721	0.00484	0.00091	0.00060	0	0
AA	0.00621	0.90460	0.08117	0.00600	0.00060	0.00110	0.0002
A	0.00049	0.02156	0.91330	0.05775	0.00448	0.00167	0.00034
BBB	0.00021	0.00223	0.04101	0.89638	0.04682	0.00824	0.00200
BB	0.00041	0.00083	0.00363	0.05794	0.83268	0.08096	0.01036
B	0	0.00074	0.00223	0.00317	0.05876	0.82296	0.04775
CCC	0.00088	0	0.00353	0.00441	0.01501	0.11132	0.53534
Default	0	0.00010	0.00044	0.00311	0.01316	0.06437	0.32950

References I

Avellaneda, M., A. Levy, and A. Paras (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance 2*, 73–88.

Bangia, A., F. Diebold, A. Kronimus, C. Schagen, and T. Schuermann (2002). Rating migration and the business cycle, with application to credit portfolio stress testing. *Journal of Banking and Finance 26*, 445–474.

Barrieu, P. and G. Scandolo (2013, July). Assessing financial risk model. arXiv.org. arXiv:1307.0684.

Basel (2006, June). International convergence of capital measurement and capital standards. Technical report, Basel Committee on Banking Supervision.

Basel (2011, February). Revisions to the Basel II market risk framework. Technical report, Basel Committee on Banking Supervision.

Berd, A. M. (2005). Dynamic estimation of credit rating transition probabilities. working paper 0912.4621, ArXiv.org.

Bernardo, J. and A. Smith (1994). *Bayesian Theory*. Chichester, UK: Wiley.

Bharadia, M.A., C. N. and G. Salkin (1996). Computing the Black-Scholes implied volatility. *Advances in Futures and Options Research 8*, 15–29.

Bluhm, C., L. Overbeck, and C. Wagner (2003). *An Introduction to Credit Risk Modeling*. Financial Mathematics Series. London: Chapman & Hall/CRC.

Blume, M., F. Lim, and A. Mackinlay (1998). The declining credit quality of U.S. corporate debt: Myth or reality. *The Journal of Finance 53*(4), 1389–1413.

Boucher, C. M., J. Danielsson, P. S. Kouontchou, and B. B. Maillet (2014). Risk models at risk. *Journal of Banking and Finance 44*, 72–92.

Brenner, M. and M. Subrahmanyam (1988). A simple formula to compute the implied standard deviation. *Financial Analysts Journal 5*, 80–83.

Bunnin, F., Y. Guo, and Y. Ren (2002). Option pricing under model and parameter uncertainty using predictive densities. *Statistics and Computing 12*, 37–44.

Buraschi, A. and F. Corielli (2005). Risk management implications of time-inconsistency: Model updating and recalibration of no-arbitrage models. *Journal of Banking and Finance 29*, 2883–2907.

Cairns, A. J. (2004). *Interest rate models*. Princeton: Princeton University Press.

Carey, M. and M. Hrycay (2001). Parameterizing credit risk models with rating data. *Journal of Banking and Finance 25*, 197–270.

Cheung, S. (1996). Provincial credit ratings in Canada: An ordered probit analysis. working paper 96-6, Bank of Canada.

Corrado, C. and T. Miller (1996). A note on a simple, accurate formula to compute implied standard deviations. *Journal of Banking and Finance 20*, 595–603.

Derman, E. (1996). Model risk. Quantitative strategies research notes, Goldman Sachs, New York.

Dowd, K. (2002). *An Introduction to Market Risk Measurement*. Chichester: Wiley.

Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society Series B 57*(1), 45–97.

Engelman, B. and K. Ermakov (2011). *The Basel II Risk Parameters*, Chapter Transition Matrices: Properties and Estimation Methods, pp. 103–116. Berlin: Springer.

References II

Filipovic, D. (1998). A note on the Nelson-Siegel family. *Mathematical Finance 9*(4), 349–359.

Filipovic, D. (2009). *Term-Structure Models*. Berlin: Springer.

Gelman, A., J. Carlin, H. Stern, and D. Rubin (1995). *Bayesian Data Analysis*. New York: Chapman and Hall.

Gibson, R. (Ed.) (2000). *Model Risk - Concepts, Calibration and Pricing*. Risk Books.

Gibson, R., F. Lhabitant, N. Pistre, and D. Talay (1999). Interest rate model risk: An overview. *Journal of Risk 3*, 37–62.

Gossl (2005). Predictions based on certain uncertainties -a Bayesian credit portfolio approach. working paper, HypoVereinsbank AG.

Hsiao, C. (1983). *Studies in Econometrics, Time Series, and Multivariate Statistics*, Chapter Regression Analysis With a Categorized Explanatory Variable, pp. 93–129. New York: Academic Press.

Hsiao, C. and D. Mountain (1985). Estimating the short run income elasticity for demand of electricity by using cross-sectional categorized data. *Journal of American Statistical Association 80*, 259–265.

Hu, Y., R. Kiesel, and J. Perraudin (2002). The estimation of transition matrices for sovereign credit ratings. *Journal of Banking and Finance 26*, 1383–1406.

Jarrow, R. A. (2009). The term structure of interest rates. *Annual Review of Financial Economics 1*.

Kao, C. and C. Wu (1990). Two-step estimation of linear models with ordinal unobserved variables: The case of corporate bonds. *Journal of Business & Economic Statistics 8*(3), 317–325.

Kukuk, M. (2002). Indirect estimation of (latent) linear models with ordinal regressors. A Monte Carlo study and some empirical illustrations. *Statistical Papers 43*, 379–399.

Li, S. (2005). A new formula for computing implied volatility. *Applied Mathematics and Computation 170*(1), 611–625.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management*. Princeton Series in Finance. Princeton and Oxford: Princeton University Press.

Nickell, P., W. Perraudin, and S. Varotto (2000). Stability of rating transitions. *Journal of Banking and Finance 24*, 203–227.

Ronning, G. and M. Kukuk (1996). Efficient estimation of order probit models. *Journal of American Statistical Association 91*(435), 1120–1129.

Spiegelhalter, D., N. Best, B. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of Royal Statistical Society, Series B 64*, 583–640.

Stefanescu, C., R. Tunaru, and S. Turnbull (2009). The credit rating process and estimation of transition probabilities: A Bayesian approach. *Journal of Empirical Finance 16*(2), 216–234.

Terza, J. (1987). Estimating linear models with ordinal qualitative regressions. *Journal of Econometrics 34*, 275–291.

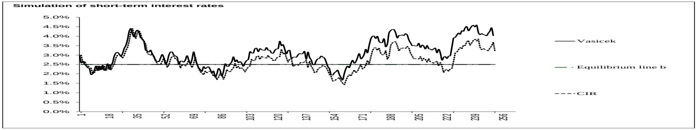
Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–188.

Williams, D. (1999). Models vs. the market: Survival of the fittest. report FIN514, Meridien Research.

Wilson, T. (1997). Credit risk modelling: A new approach. Technical report, McKinsey Inc., New York. unpublished mimeo.

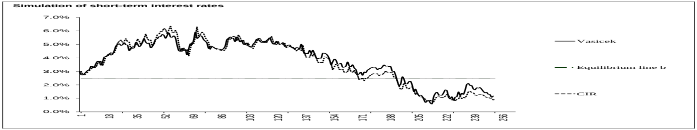
Zhou, C. (2001). Credit rating and corporate defaults. *Journal of Fixed Income 3*(11), 30–40.

Return From



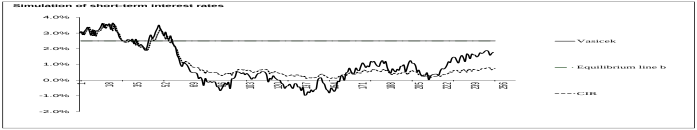
(a) expected

(b) all rates above long-run level

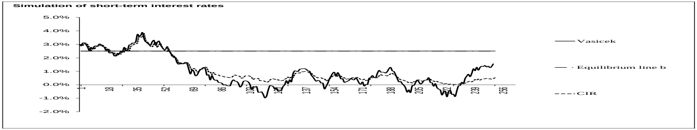


(c) downward trend

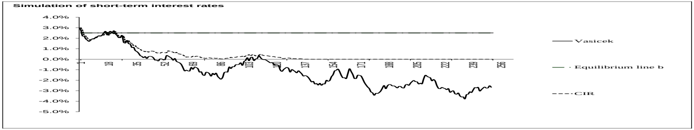
Figure: First comparison of simulated paths for the Vasicek model and the CIR model



(a) Vasicek negative and then recovers



(b) Vasicek negative and recover



(c) Vasicek negative and CIR goes to zero

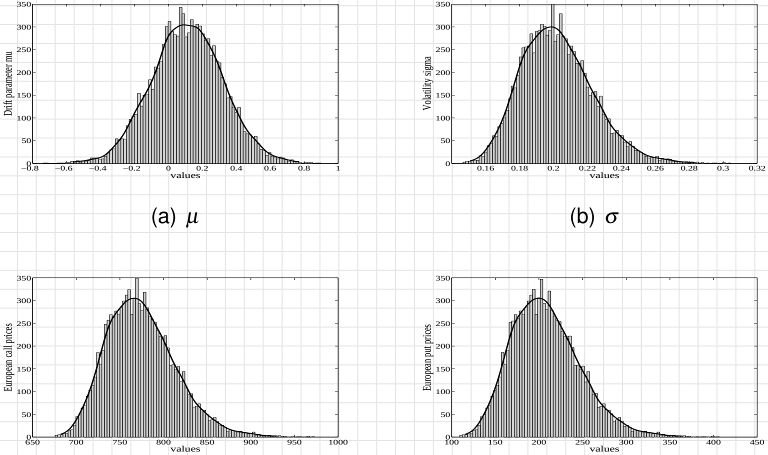
Figure: Second comparison of simulated paths for the Vasicek model and the CIR

MCMC Analysis of Bayesian option pricing under the Black-Scholes GBM model. Posterior inference statistics for mean, standard deviation, median and 2.5% and 95% quantiles from a sample of 50000 values.

Variable	mean	s.d.	MC error	2.5%	median	97.5%
μ	0.1108	0.2094	6.648E-4	-0.3001	0.1107	0.5237
σ_{BS}	0.2023	0.02119	7.036E-5	0.1659	0.2006	0.249
call	774.9	39.63	0.1315	708.5	771.1	863.6
put	208.4	39.63	0.1315	142.0	204.6	297.1
λ	0.1783	1.029	0.003299	-1.833	0.1788	2.196

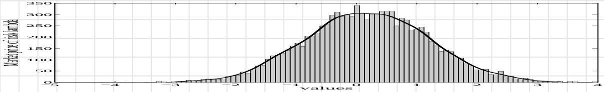
Return From

Return From

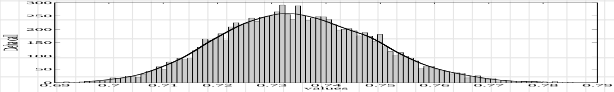


(a) μ (b) σ (c) call (d) put

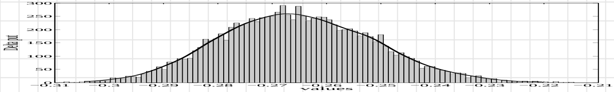
Figure: Posterior densities of the Black-Scholes parameters and the European call and put option price for the FTSE100 index. The strike price is $K = 5500$, initial index value is $S_{t_0} = 5669.1$, risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.



(a) λ



(b) Δ(call)



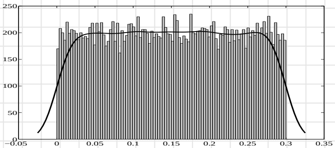
(c) Δ(put)

Figure: The Posterior densities of the Black-Scholes market price of risk $\frac{\mu - r}{\sigma}$, and the Greek delta parameter for the European call and put option prices for the FTSE100.

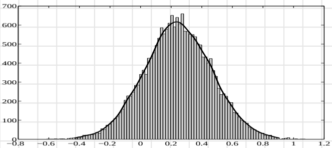
Posterior estimates of the parameters of the CEV model from the FTSE100 data. Inference is obtained with MCMC from a sample of 20000 values.

Variable	mean	s.d.	MC error	2.5%	median	97.5%
γ	0.2904	0.1801	0.002925	0.03987	0.2573	0.7229
μ	0.2441	0.2216	0.001651	-0.1903	0.2431	0.6824
q	0.1503	0.08642	6.688E-4	0.007853	0.1502	0.2919
σ_{CEV}	2.0	1.241	0.01737	0.2403	1.801	4.71

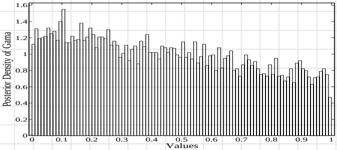
Return From



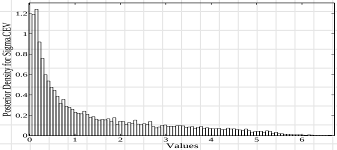
(a) dividend yield q



(b) drift parameter μ



(c) elasticity γ



(d) σ_{CEV}

Figure: The posterior densities for the parameters of the CEV model calculated using data on FTSE100 index with MCMC from a sample of 20000 values. The strike price is $K = 5500$, initial index value is $S_{t_0} = 5669.1$, risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.

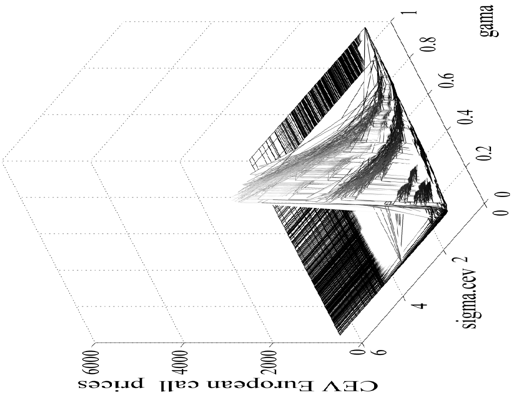


Figure: Posterior surface for the European call price on the FTSE100 generated by the parameter uncertainty on γ and σ_{CEV}. The strike price is $K = 5500$, initial index value is $S_{t_0} = 5669.1$, risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.

Show put CEV

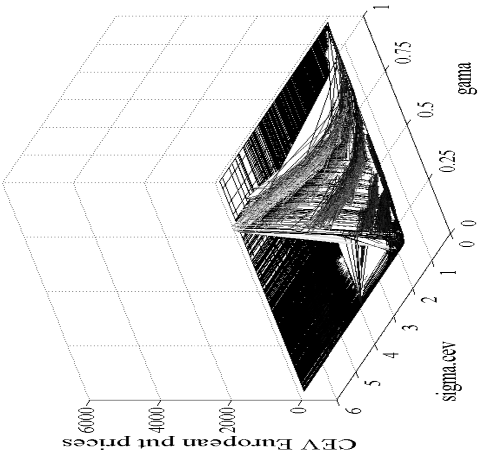
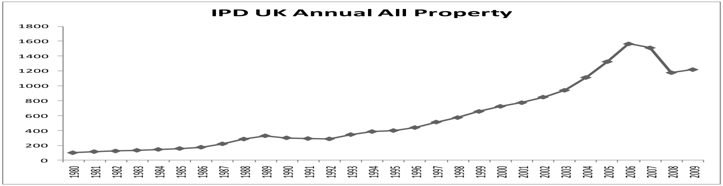
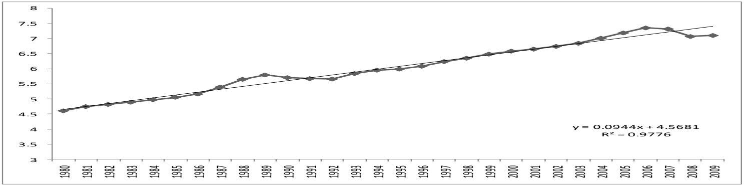


Figure: Posterior surface for the European put price on the FTSE100 generated by the parameter uncertainty on γ and σ_{CEV}. The strike price is $K = 5500$, initial index value is $S_{t_0} = 5669.1$, risk-free rate is $r = 0.075$ and time to maturity is $T = 1$ year.

Return From



(a) actual levels



(b) log scale; linear trend fitted by OLS with R-squared 97.76% suggesting a “very good” fit.

Approximation formulae for implied volatility

- Brenner and Subrahmanyam (1988) derived a formula based on the assumption that $S_0 = Ke^{-rT}$. $\widehat{\sigma} \approx \sqrt{\frac{2\pi}{T}} \frac{C^{mkt}}{S_0}$
- Bharadia and Salkin (1996) used a general strike price,

$$\widehat{\sigma} \approx \sqrt{2\pi} T \frac{C^{mkt} - (S_0 - Ke^{-rT})/2}{(S_0 + Ke^{-rT})/2} \tag{44}$$

- Corrado and Miller (1996) derived the following approximation

$$\begin{aligned} \widehat{\sigma} \quad &\approx \quad \sqrt{\frac{2\pi}{T}} \frac{1}{S_0 + Ke^{-rT}} \left[C^{mkt} - \frac{S_0 - Ke^{-rT}}{2} \right] \\ &+ \sqrt{\frac{2\pi}{T}} \frac{1}{S_0 + Ke^{-rT}} \sqrt{\frac{(C^{mkt} - S_0 + Ke^{-rT})^2}{4} - \frac{(S_0 - Ke^{-rT})^2}{\pi}} \end{aligned} \tag{45}$$

- Li (2005) provided an approximation formula that is valid regardless of the market option moneyness. If $\eta = \frac{Ke^{-rT}}{S_0}$, $\alpha = \frac{\sqrt{2\pi}}{1+\eta}\left[\frac{2C^{mkt}}{S_0} + \eta - 1\right]$

$$\widehat{\sigma} \approx \begin{cases} \frac{2\sqrt{2}}{\sqrt{T}} z - \frac{1}{\sqrt{T}} \sqrt{8z^2 - \frac{6\alpha}{z\sqrt{2}}}, & \text{if } \frac{S_0 |Ke^{-rT} - S_0|}{(C^{mkt})^2} \leq 1.4 \\ \frac{\alpha + \sqrt{\alpha^2 - \frac{4(\eta-1)^2}{1+\eta}}}{2\sqrt{T}}, & \text{otherwise.} \end{cases} \tag{46}$$

where $z = \cos\left[\frac{1}{3}\cos^{-1}\left(\frac{3\alpha}{\sqrt{32}}\right)\right]$.

Return From

Approximation formulae for implied volatility

- Brenner and Subrahmanyam (1988) derived a formula based on the assumption that $S_0 = Ke^{-rT}$. $\widehat{\sigma} \approx \sqrt{\frac{2\pi}{T}} \frac{C^{mkt}}{S_0}$
- Bharadia and Salkin (1996) used a general strike price,

$$\widehat{\sigma} \approx \sqrt{2\pi T} \frac{C^{mkt} - (S_0 - Ke^{-rT})/2}{(S_0 + Ke^{-rT})/2} \tag{44}$$

- Corrado and Miller (1996) derived the following approximation

$$\begin{aligned} \widehat{\sigma} \quad &\approx \quad \sqrt{\frac{2\pi}{T}} \frac{1}{S_0 + Ke^{-rT}} \left[C^{mkt} - \frac{S_0 - Ke^{-rT}}{2} \right] \\ &+ \sqrt{\frac{2\pi}{T}} \frac{1}{S_0 + Ke^{-rT}} \sqrt{\frac{(C^{mkt} - S_0 + Ke^{-rT})^2}{4} - \frac{(S_0 - Ke^{-rT})^2}{\pi}} \end{aligned} \tag{45}$$

- Li (2005) provided an approximation formula that is valid regardless of the market option moneyness. If $\eta = \frac{Ke^{-rT}}{S_0}$, $\alpha = \frac{\sqrt{2\pi}}{1+\eta}\left[\frac{2C^{mkt}}{S_0} + \eta - 1\right]$

$$\widehat{\sigma} \approx \begin{cases} \frac{2\sqrt{2}}{\sqrt{T}} z - \frac{1}{\sqrt{T}} \sqrt{8z^2 - \frac{6\alpha}{z\sqrt{2}}}, & \text{if } \frac{S_0 |Ke^{-rT} - S_0|}{(C^{mkt})^2} \leq 1.4 \\ \frac{\alpha + \sqrt{\alpha^2 - \frac{4(\eta-1)^2}{1+\eta}}}{2\sqrt{T}}, & \text{otherwise.} \end{cases} \tag{46}$$

where $z = \cos\left[\frac{1}{3}\cos^{-1}\left(\frac{3\alpha}{\sqrt{32}}\right)\right]$.

Return From

Approximation formulae for implied volatility

- Brenner and Subrahmanyam (1988) derived a formula based on the assumption that $S_0 = Ke^{-rT}$. $\widehat{\sigma} \approx \sqrt{\frac{2\pi}{T}}\frac{C^{mkt}}{S_0}$
- Bharadia and Salkin (1996) used a general strike price,

$$\widehat{\sigma} \approx \sqrt{2\pi T}\,\frac{C^{mkt} - (S_0 - Ke^{-rT})/2}{(S_0 + Ke^{-rT})/2} \tag{44}$$

- Corrado and Miller (1996) derived the following approximation

$$\begin{aligned} \widehat{\sigma} \quad &\approx \quad \sqrt{\frac{2\pi}{T}}\frac{1}{S_0 + Ke^{-rT}}\left[C^{mkt} - \frac{S_0 - Ke^{-rT}}{2}\right] \\ &+ \sqrt{\frac{2\pi}{T}}\frac{1}{S_0 + Ke^{-rT}}\sqrt{\frac{(C^{mkt} - S_0 + Ke^{-rT})^2}{4} - \frac{(S_0 - Ke^{-rT})^2}{\pi}} \end{aligned} \tag{45}$$

- Li (2005) provided an approximation formula that is valid regardless of the market option moneyness. If $\eta = \frac{Ke^{-rT}}{S_0}$, $\alpha = \frac{\sqrt{2\pi}}{1+\eta}\left[\frac{2C^{mkt}}{S_0} + \eta - 1\right]$

$$\widehat{\sigma} \approx \begin{cases} \frac{2\sqrt{2}}{\sqrt{T}}z - \frac{1}{\sqrt{T}}\sqrt{8z^2 - \frac{6\alpha}{z\sqrt{2}}}, & \text{if } \frac{S_0|Ke^{-rT} - S_0|}{(C^{mkt})^2} \leq 1.4 \\ \frac{\alpha + \sqrt{\alpha^2 - \frac{4(\eta-1)^2}{1+\eta}}}{2\sqrt{T}}, & \text{otherwise.} \end{cases} \tag{46}$$

where $z = \cos\left[\frac{1}{3}\cos^{-1}\left(\frac{3\alpha}{\sqrt{32}}\right)\right]$.

Approximation formulae for implied volatility

- Brenner and Subrahmanyam (1988) derived a formula based on the assumption that $S_0 = Ke^{-rT}$. $\widehat{\sigma} \approx \sqrt{\frac{2\pi}{T}\frac{C^{mkt}}{S_0}}$
- Bharadia and Salkin (1996) used a general strike price,

$$\widehat{\sigma} \approx \sqrt{2\pi T}\,\frac{C^{mkt} - (S_0 - Ke^{-rT})/2}{(S_0 + Ke^{-rT})/2} \tag{44}$$

- Corrado and Miller (1996) derived the following approximation

$$\begin{aligned}\widehat{\sigma} &\approx \sqrt{\frac{2\pi}{T}}\frac{1}{S_0 + Ke^{-rT}}\left[C^{mkt} - \frac{S_0 - Ke^{-rT}}{2}\right] \\ &+ \sqrt{\frac{2\pi}{T}}\frac{1}{S_0 + Ke^{-rT}}\sqrt{\frac{(C^{mkt} - S_0 + Ke^{-rT})^2}{4} - \frac{(S_0 - Ke^{-rT})^2}{\pi}}\end{aligned} \tag{45}$$

- Li (2005) provided an approximation formula that is valid regardless of the market option moneyness. If $\eta = \frac{Ke^{-rT}}{S_0}$, $\alpha = \frac{\sqrt{2\pi}}{1+\eta}\left[\frac{2C^{mkt}}{S_0} + \eta - 1\right]$

$$\widehat{\sigma} \approx \begin{cases} \frac{2\sqrt{2}}{\sqrt{T}}z - \frac{1}{\sqrt{T}}\sqrt{8z^2 - \frac{6\alpha}{z\sqrt{2}}}, & \text{if } \frac{S_0|Ke^{-rT} - S_0|}{(C^{mkt})^2} \leq 1.4 \\ \frac{\alpha + \sqrt{\alpha^2 - \frac{4(\eta-1)^2}{1+\eta}}}{2\sqrt{T}}, & \text{otherwise.} \end{cases} \tag{46}$$

where $z = \cos\left[\frac{1}{3}\cos^{-1}\left(\frac{3\alpha}{\sqrt{32}}\right)\right]$.

Return From

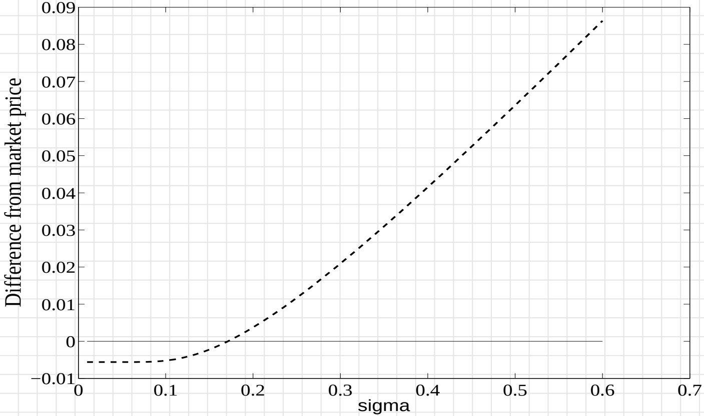


Figure: Calculating the implied volatility for a stock with current value \$34.14 from the market price \$4.7 of a European call option with maturity $T = 0.45$ years and a strike price of $K = \$30.00$, assuming that the risk-free rate is 2.75%.

Return From

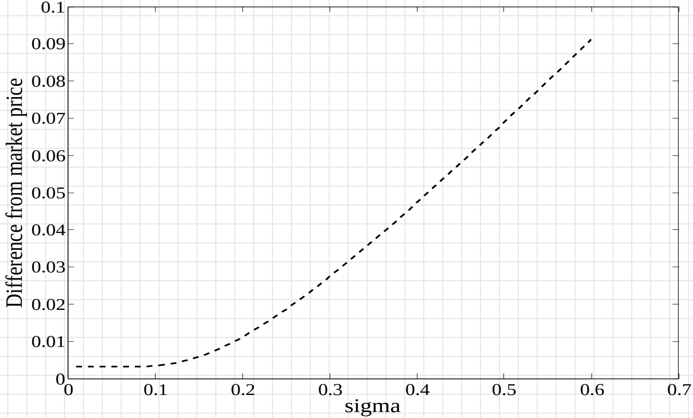


Figure: Calculating the implied volatility for a stock with current value \$34.14 from the market price \$4.7 of a European call option with maturity $T = 0.45$ years and a strike price of $K = \$30.00$, assuming that the risk-free rate is 5%.

Return From