Essays on Term Structure and Monetary Policy

Sven Skallsjö

AKADEMISK AVHANDLING

Som för avläggande av filosofie doktorsexamen
vid Handelshögskolan i Stockholm framläggs för offentlig
granskning
fredagen den 26 mars 2004, kl 15.15
i sal Ragnar, Handelshögskolan
Sveavägen 65

Essays on Term Structure and Monetary Policy

STOCKHOLM SCHOOL OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

EFI Mission
EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics.
EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

Research Organization
The research activities are organized in twenty Research Centers within eight Research Areas. Center Directors are professors at the Stockholm School of Economics.

ORGANIZATION AND MANAGEMENT	
Management and Organisation; (A)	Prof Sven-Erik Sjöstrand
Center for Ethics and Economics; (CEE)	Adj Prof Hans de Geer
Center for Entrepreneurship and Business Creation; (E)	Prof Carin Holmquist
Public Management; (F)	Prof Nils Brunsson
Information Management; (I)	Prof Mats Lundeberg
Center for People and Organization; (PMO)	Prof Jan Löwstedt
Center for Innovation and Operations Management; (T)	Prof Christer Karlsson
ECONOMIC PSYCHOLOGY	
Center for Risk Research; (CFR)	Prof Lennart Sjöberg
Economic Psychology; (P)	Prof Guje Sevón
MARKETING	
Center for Consumer Marketing; (CCM)	Acting Prof Magnus Söderlund
Center for Information and Communication Research; (CIC)	Adj Prof Bertil Thorngren
Marketing, Distribution and Industrial Dynamics; (D)	Prof Björn Axelsson
ACCOUNTING, CONTROL AND CORPORATE FINANCE	
Accounting and Managerial Finance; (B)	Prof Lars Östman
Managerial Economics; (C)	Prof Peter Jennergren
FINANCE	
Finance; (FI)	Prof Clas Bergström
ECONOMICS	
Center for Health Economics; (CHE)	Prof Bengt Jönsson
International Economics and Geography; (IEG)	Prof Mats Lundahl
Economics; (S)	Prof Lars Bergman
ECONOMIC STATISTICS	
Economic Statistics; (ES)	Prof Anders Westlund
LAW	
Law; (RV)	Prof Erik Nerep

Chairman of the Board: Prof Carin Holmquist. *Director:* Associate Prof Bo Sellstedt.

Adress
EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

Essays on Term Structure and Monetary Policy

Sven Skallsjö

STOCKHOLM SCHOOL
OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

Dissertation for the Degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics, 2004

ISBN NR 91-7258-638-9

Keywords:
Bank, Firm, Fixed Rate, Term Structure of Interest Rates, Interest Rates, Nominal Interest Rates, Monetary Policy, Policy

Printed by:
Elanders Gotab, Stockholm 2004

Distributed by:
EFI, The Economic Research Institute
Stockholm School of Economics
P O Box 6501, SE 1-113 83 Stockholm, Sweden
www.hhs.se/efi

To my father

Acknowledgements

Over the many years of writing this thesis I made acquaintance with several individuals, who in various ways contributed to my work reaching its completion.

I had four different advisers and through them I came to appreciate different aspects of the profession. I owe my gratitude to Tomas Björk, Mike Burkart, Peter Englund and Paul Söderlind. I was truly fortunate to receive such distinguished advice.

Many warm thanks go to my fellow Ph.D. students from the early years, Mats Ekelund, Rickard Eriksson, Sten Holmström, Kaj Martensen, Arvid Nilsson, Niklas Strand and Patrik Säfvenblad. With them I maintained a relation to the academic community, even after I had defected the program for an employment at Swedbank Markets. Their friendship contributed significantly to my return to the Stockholm School of Economics.

The years at Swedbank Markets were enlightening and essential for my understanding of economics. I sincerely appreciate the patience shown by Åke Gustavsson and other colleagues. I am also highly grateful for the discussions with Linus Ericsson and his colleagues at AGL Structured Finance.

At the Department of Financial Economics faculty members and Ph.D. students contributed with valuable comments at seminars and during discussions. I am greatly appreciative to the strong support and friendship of Krister Ahlersten, Magnus Blix, Raquel Gaspar, Roland Nilsson, Irina Slinko and Per Östberg. The secretaries Eva Eklund, Marita Rosing and Anneli Sandbladh deserve many thanks for they are considerate, forbearing and efficient. I gratefully acknowledge the financial support given by Bankforskningsinstitutet and the Wallander Foundation.

Finally I thank my parents, and my father in particular, with his joy and humble respect for science that I learned to appreciate that I always carry with me.

Stockholm in February 2004

Sven Skallsjö

Summary of thesis

This dissertation treats two different themes. The first, addressed in Chapter 1, regards the pricing of interest rate swaps. The second, studied in the remaining two chapters, regards the implications of monetary policy for the term structure of interest rates.

The pricing of interest rate swaps

An interest rate swap is an agreement between two parties to exchange fix for floating interest rate payments for a certain period of time. Floating rate payments are made at a floating-rate index, e.g. the three-month interbank rate, while the fixed rate payment, the *swap rate*, is determined on the market. The swap rate may include a compensation for credit risk depending on the counterparty's credit quality, but in the standard agreement there is no exchange of principal, only interest is transacted, and this effectively reduces concerns about credit risk.

The *swap spread* for a given maturity is the difference between the swap rate and the risk-free rate, measured as the yield on a government bond with similar cash flows. If the standard swap agreement entails negligible credit risk one might expect swap spreads to be low and stable, but market swap spreads vary over time. There are periods when swap spreads are low in accordance with the general theory, but there are also periods when swap spreads reach levels that seem high. Over the past five or six years the unexplained part of the swap spread has remained above 30 basis points over sustained periods of time.

Understanding the determinants of the swap spread is valuable as results can have implications for important economic decisions. For example, it is possible that a positive swap spread follows as a result of government debt buybacks. A reduced supply of government debt acts as to push down the yield on government bonds while leaving swap rates unaffected. But by this argument the reverse strategy should appear attractive for the government. If the swap spread is high the government can raise capital through long term borrowing, invest the amount in sort-term low risk securities, typically short-term bank deposits, and enter a pay float - receive fix interest rate swap. With this strategy the government is exposed to a short-term bank risk, but if the swap spread is high the compensation for taking this risk is accordingly high.

This creates strong incentives for the government to use the swap market. In spite of this we see long periods with significantly positive swap spreads.

The first chapter of this dissertation examines a setting where a positive swap spread arises as part of an equilibrium in a perfectly competitive capital market. The model is one of insurance under adverse selection. A firm that seeks debt financing can insure itself against interest rate risk either by borrowing long-term or by borrowing short-term and entering a pay fix - receive float interest rate swap. The latter alternative allows for a partial hedge as the firm can choose to swap only a fraction of the nominal amount.

In this setting, if firms' credit quality and interest rate risk tolerance are correlated creditors can use the pricing of interest rate swaps as a screening device. A low-risk firm, being a firm with favorable private information, selects short-term borrowing and partial insurance. A high-risk firm, being a firm with less favorable prospects, is by assumption also less risk tolerant. It therefore has a higher demand for insurance and the equilibrium swap spread is set such that the high-risk firm finds it more beneficial to borrow long-term at a cost that exceeds the expected cost from short-term financing, but that provides a full insurance to interest rate risk.

The positive swap spread thus separates low-risk firms from high-risk firms. With a positive swap spread the creditor makes a profit in expected terms on the swap transaction, but the assumption of perfectly competitive capital markets is still preserved as the expected profit is offset by a corresponding loss in the lending facility.

Monetary policy and the term structure of interest rates

Taken separately monetary policy and term structure modeling are two well-established research areas each comprising a substantial amount of research. But relatively few attempts have been made to integrate the two. The last two chapters of this dissertation take the view that the conduct of monetary policy is an essential element in the determination of the term structure of interest rates, and that explicitly considering the role of a monetary authority in the analysis has a potential of enhancing our understanding of term structure dynamics, and its relation to macro-economic fundamentals in particular. This approach to the term structure is supported by the fact that the analytical framework developed in the literature on optimal monetary policy translates conveniently into a setting well suited for term structure analysis.

Chapter 2 makes the point in the simplest setting. A standard model of optimal monetary policy is reformulated in continuous time. Combined with a parameterized form for the market price of risk this produces a standard term structure model with well-known characteristics. This model is estimated on US data for the period 1987 - 2002, treating state variables as latent factors of the term structure. The parameters that are estimated comprise parameters describing the monetary transmission mechanism, parameters describing the monetary authority's preferences and parameters describing the market price of risk. Our estimation technique differs from comparable estimations in the monetary policy literature as these typically take state variables to be directly observable measures of macro-economic aggregates. The results using term structure data are both similar and different to previous findings. The main difference when using term structure data is that the central bank's estimated policy is more aggressive, i.e. more responsive to changes in the underlying state variables.

This attempt to bring the two research areas together can be regarded as a first step as the model for the monetary transmission mechanism is the simplest available. Agents' expectations are formed adaptively, while the monetary policy literature emphasizes the importance of forward-looking behavior. Also, in work on monetary policy it is common to include lagged values of the state variables which results in state dynamics with longer memory. Extending the model in this direction would suggest a term structure model with more state variables.

Chapter 3 is devoted to the zero bound on nominal interest rates. While the zero bound is well recognized in the literature on term structure modeling, not much has been said about term structure dynamics under the special circumstance that the short rate is close to zero. I find the optimal monetary policy approach to be particularly well suited for this analysis.

The chapter studies a continuous time reduced form version of the monetary transmission mechanism. The monetary authority's optimization problem is formed according to two specifications, interest rate stabilization and interest rate smoothing. For the former the optimization problem is solved analytically, while numerical procedures are adopted for the latter. The chapter then turns to study implications for the term structure under risk-neutrality. Term structure equations are solved numerically and implications for the term structure are discussed. Data for a low-interest rate country like Japan for 1996 - 2003 exhibits s-shaped yield curves and yield volatility curves. This shape is found to be consistent with a smoothing objective for the short rate.

Contents

1. Financing choices and the swap spread

Abstract

The pricing of interest rate swaps is examined in the presence of asymmetric information between firm owners and creditors. A fairly standard version of Rothschild and Stiglitz (1976) is considered. A firm borrowing short-term is exposed to an interest rate risk. It can attain full insurance to this risk by borrowing long-term, or it can insure itself partially by borrowing short-term, and swapping a fraction of the nominal amount to fixed rate payments. If firms' credit quality and interest rate risk tolerance are correlated, creditors can use the swap spread as a screening device. The model is consistent with observed market rates, which are difficult to recover with standard measures of risk.

1.1 Introduction

An interest rate swap is an agreement between two parties agree to exchange fix for floating interest rate payments for a certain period of time. Floating rate payments are then made at a floating-rate index, e.g. the three-month interbank rate. The fixed rate payment is the *swap rate*. The *swap spread* for a given maturity is the difference between the swap rate and the risk-free rate, measured as the yield on a government bond with similar cash flows, taking the principal of the bond into account.

In recent years the swap spread has reached levels, which seem high considering the limited risk involved in the transaction. More precisely, the magnitude of the swap spread can be seen as a portfolio choice anomaly. The combined strategy of a roll-over investment in short-term risk-free bills and a pay float - receive fix swap produces a yield, which generally is higher than the corresponding long-term risk-free yield. The risk involved in the former strategy seems insufficient to motivate the yield differential. Therefore, given the existence of the former strategy, it is difficult to explain investors' high demand for long-term government bonds or even high-rated long-term corporate bonds.

[0] Foremost I wish to thank my adviser Mike Burkart for valuable discussions and general guidance. I also thank Michel Habib for valuable comments. My work has further benefited from numerous discussions with both academics at the Department of Financial Economics at the Stockholm School of Economics and practitioners at Swedbank Markets and AGL Structured Finance. Financial support from Bankforskningsinstitutet is gratefully acknowledged.

The present paper develops a fairly standard model of insurance under adverse selection (Rothschild and Stiglitz (1976)). A firm can attain full insurance to interest rate risk by borrowing long-term, or it can insure itself partially by borrowing short-term, and swapping a fraction of the nominal amount to fixed rate payments. If firms' credit quality and interest rate risk tolerance are correlated, creditors can use the pricing of interest rate swaps as a screening device. Similarly to Rothschild and Stiglitz there is insurance rationing. A low-risk firm, being a firm with favorable private information, selects short-term borrowing and partial insurance. This is inefficient compared to first best, where there is perfect information, since then low-risk firms attain full insurance at no cost.

The insurance part is, that a firm borrowing short-term with frequent refinancing is exposed to an interest rate risk. As debt is rolled over, interest rates may have changed, and this induces variability in the firm's cash flows. The firm can insure itself by either borrowing long-term, or by combining short-term borrowing with a pay fix - receive float interest rate swap. In this setting the swap spread can be interpreted as an insurance premium.

The adverse selection part is, that a firm's credit quality and its sensitivity to interest rate fluctuations may be correlated. A firm expecting a plunge in its credit quality, is likely to appreciate a more stable debt repayment path. A firm expecting its credit quality to improve can tolerate a higher exposure to interest rate risk. If the creditor takes the simultaneous role of a counterpart in the swap agreement, she can utilize the swap spread as a screening device. If the swap spread is set sufficiently high, the firm expecting its credit quality to fall, will find it too expensive to hedge with an interest rate swap, and will choose to borrow long-term. The firm expecting its credit quality to improve demands a less complete hedge, and will benefit from borrowing short-term and entering a limited position in a pay fix - receive float swap.

The swap spread thus separates the market. It turns out, that although the creditor makes a profit in expected terms on the swap transaction, this can be off-set by a corresponding loss on the lending facility. This secures zero profit for creditors, and makes the model consistent with perfectly competitive capital markets.

An early paper on interest rate swaps related to this one is Titman (1992). Extending an observation by Flannery (1986), Titman argues, that a firm with favorable private information has an incentive to choose short-term debt. If it borrows long-term, it risks being locked in at an unfavorable borrowing rate, even after the private information has been revealed. Short-term borrowing allows the creditor to adjust the borrowing rate as information becomes public, which secures a more correct pricing of debt. When borrowing short-term, the firm is exposed to an interest rate risk, and this induces a demand for combining short-term borrowing with a pay fix - receive float interest rate swap. Though similar to the model examined here, the motive for high-quality firms to borrow short-term differs. In my model short-term borrowing is purely a matter of screening.

Another closely related paper is Mozumdar (2001), which examines the structure of swap markets in the presence of asymmetric information between firm insiders and creditors. The main point made by Mozumdar is, that due to limited liability of equity holders some firms may use swaps for speculative purposes. This amplifies the credit component in swap agreements, and several institutional features of the swap market can be seen as methods for mitigating speculation. However, the focus of Mozumdar is not on pricing. Although pricing is examined, in equilibrium the expected profit on a swap transaction to an investor is zero. In particular it is argued, that price based methods amplifies the speculative intent of firm insiders. The present paper focuses exactly on conditions, under which price-based methods can be advocated. Nevertheless, the setup of the model and the general insurance problem studied by Mozumdar, are both similar to mine. There are two principal differences however. First, I essentially abstract from the use of swaps for speculative purposes. Second, I allow for long-term debt.

Undoubtedly both Titman's and Mozumdar's model explain several observed characteristics of the market for interest rate swaps. The Titman model is further supported by empirical findings, see e.g. Wall and Pringle (1989), Samant (1996), and Saunders (1998). However, in both the Mozumdar and the Titman model the central question of the present paper remains unanswered: Why is the swap spread so high?

The pricing of interest rate swaps has generally been studied in a different context. Using standard techniques for credit risk, counterparty risk has been investigated in Duffie and Huang (1995) and Huge and Lando (1999). The main finding is, that counterparty risk probably is not an economically significant determinant of the swap spread. In Lang, Litzenberger and Liu (1998) and Fehle (1999) imperfect competition among creditors results in a positive swap spread. However, given the size of the market for interest rate swaps, it is natural to ask if a positive swap spread can be supported even under perfect competition.

1.2 The economic setting

Consider a firm that needs to raise \$1 at date 0 to finance a project that generates a payoff at date 2. The firm is owned by an entrepreneur, who has no initial wealth.

1.2.1 Firm types

In the economy there are two types of firms, G for good and B for bad (or less efficient). The fraction of G firms is α and the fraction of B firms is

$1-\alpha$. A firm of type G generates a certain payoff G at date 2. A firm of type B generates a payoff $B < G$ with probability p, and a payoff zero with probability $1-p$, also at date 2. The zero payoff is referred to as default.[1]

The assumption $B < G$ is made to illustrate different sensitivity to interest rate risk. The idea is, that when borrowing short-term, a firm of type B is more vulnerable to interest rate fluctuations, even if it remains solvent. Thus, conditional on not defaulting a firm of type B is expected to be *less* solvent than a firm of type G. This assumption makes it possible for outside creditors to separate firms.

For a risk neutral investor both types of firms have positive net present values.

1.2.2 Information

At date 0 firm type is private information to the owner. She can not disclose any additional information about the firm to her creditors even if she wishes to do so. Further, for a B firm, the stochastic component is not realized until later, and thus at date 0 the final outcome of the project is not known to the B firm owner.

At date 1, the future payoff of the project becomes publicly known. Consequently, at date 1 firm types are also revealed. Payoffs are, however, not realized until date 2.

As it turns out, the unravel of information at date 1 is redundant. No information need to become public until date 2.

1.2.3 Utility functions

All agents maximize expected utility at date 2 with no discounting.[2] While investors are risk neutral, firm owners are risk averse. This induces a demand for firm owners to hedge interest rate risk. In particular it is assumed that $u(X) = -(c+X)^{-2}$, where c is a positive constant. A vital characteristic of this functional form is decreasing absolute risk aversion. This implies that a type B firm, which by previous assumptions expects a lower wealth in period 2 than a type G firm, has a higher demand for insurance.[3] The constant c is needed, for utility to be well defined in case final wealth is zero. However, c is assumed to be small (se Assumption 1.1 below).

[1] A zero liquidation value is made for tractability. The assumption that G firms never default is made for tractability. What is important is that the default probability is smaller for a G firm than for a B firm.

[2] We could allow for discounting, but since the relevant disount factor, the risk-free rate, will be stochastic, discounting introduces complexity without adding much economic insight.

[3] Decreasing absolute risk aversion rules out the negative exponential utility function, whose coefficient of absolute risk aversion is a constant. However, power utility is included.

1.2.4 Available financing

Firms can commit to contract with a single creditor. Available financing is debt and interest rate swaps. Debt can be short- or long-term. If the firm issues long-term debt it borrows \$1 and agrees to repay an amount N_{LT} at date 2, in case it is solvent. If the firm defaults, repayment is zero.

If the firm issues short-term debt it borrows \$1 at date 0 and agrees to roll over an amount N_{ST} at date 1. If the firm is solvent, the roll-over is made at the second period risk-free interest rate. However, this rate will be stochastic, and thus not known to participants at date 0. Denoting the second period risk-free interest rate by $\widetilde{r}$, the repayment obligation at date 2, in case the firm is solvent, is $N_{ST}(1+\widetilde{r})$. If the firm defaults, repayment is zero.

An interest rate swap, to be defined in detail below, involves swapping an amount of outstanding debt to fix or floating rate payments. The interest rate swap has the function of an insurance against interest rate fluctuations. This provides a way for a firm borrowing short-term, to hedge interest rate risk. The creditor can take the simultaneous role of a swap counterpart to the firm.

1.2.5 The interest rate process

The long-term interest rate is the yield at date 0 on a risk-free bond, maturing at date 2. The long-term interest rate, denoted by r, is normalized to 0.

The short interest rate is modeled as follows. At date 0 the short rate for the first period, between date 0 and 1, is zero, while the short rate for the second period is not known. At date 1, the second period interest rate, denoted by $\widetilde{r}$, is realized according to a uniform distribution on $[-\Delta, \Delta]$ for some $\Delta > 0$.[4] The distribution is independent of the bankruptcy component for the B firm.[5]

The long-term yield and the short-term interest rate, are related according to $1 + r = E[1+\widetilde{r}] \equiv 1$.[6]

1.2.6 Swap contracts

A swap contract specifies a unit price λ and a swap amount s. In the given setting λ is the swap spread. A pay fix -receive float swap of an amount $s > 0$ implies transferring $s(r + \lambda - \widetilde{r}) = s(\lambda - \widetilde{r})$ to the counterparty at date 2. If the quantity is negative, the counterparty is indebted to the initial part. For

[4] I allow for negative interest rates for analytical tractability.

[5] That is, $P(\widetilde{r} < \overline{r}) = P(\widetilde{r} < \overline{r} \,|B \text{ defaults}) = P(\widetilde{r} < \overline{r} \,|B \text{ does not default})$.

[6] Under no arbitrage the long-term interest rate would be determined by $1/R = E\left[1/\widetilde{R}\right]$. Since absence of arbitrage is not the focus of the paper, I use the more convenient expectation hypothesis. This is in conformity with agents' discounting of utility (see Footnote 2).

simplicity, I assume that the swap contract is such, that in case one of the parties declares bankruptcy, the contract is cancelled. Although this is not how swap contracts are written in general, the assumption keeps the analysis more tidy, while it does not affect the qualitative results.[7]

1.2.7 Sequence of moves

In the presence of adverse selection the concept of competition among creditors depends crucially on the specific game theoretic model applied. Here I model contracting as a two-stage process. In the first stage creditors offer a menu of financing options to firms. In the second stage firms choose from the options available. Creditors are obliged to meet any request.

A financing contract offered by a creditor, specifies conditions for either long-term or short-term financing. However, because firm owners are risk averse, there is no hedging motive for a firm borrowing long-term to participate in interest rate swaps. The set of financing contracts is therefore constrained, so that interest rate swaps are only made available for firms borrowing short-term.

Further it is assumed, that a short-term contract can only include a nominal amount of debt and a swap spread. Thus it is assumed, that the contract can not specify a particular swap amount. For the short-term contract the borrower can select a swap position at her own choice. One could allow short-term contracts to also include an explicit swap amount. To a certain extent creditors can and do control borrowers' use of interest rate swaps. However, it is reasonable to allow borrowers some freedom in their financing decision. Also, the proposed scheme is consistent with the interpretation of the swap spread as a marginal cost of hedging. Appendix 1.8 investigates the implications of explicitly including the swap amount in the contract.[8]

In summary, the set of admissible offers is given by

1. Long-term financing contracts N_{LT}, where $N_{LT} \in \mathbb{R}$ is the nominal amount of debt.
2. Short-term financing contracts (N_{ST}, λ), where $N_{ST} \in \mathbb{R}$ is the nominal amount, and $\lambda \in \mathbb{R}$ is the swap spread.

[7] Procedures for the handling of interest rate swaps in the event of default, are controlled by the ISDA (*International Swaps and Derivatives Association*) agreements. The market value of the swap position is added to the value of the firm's assets.

[8] An alternative to the proposed two-stage procedure is to add a third stage, where creditors are allowed to deny an observed request. However, this is effectively equivalent to letting the creditor include a specific swap amount in the contract; the creditor can announce that she will deny any loan application involving a swap amount other than s. Even if creditors can not make such announcements, firm owners can foresee that the creditor will deny any loan application, on which the expected profit is negative. So far it has been assumed, that such contracts are infeasible.

1.2.8 Additional assumption

To simplify the analysis, it is assumed, that default can not occur due to adverse interest rate movements. This abstracts from the use of interest rate swaps for speculative purposes as in Mozumdar, which is not the focus here. If the constant c in the utility function is small, low realizations of date 2 wealth are given a higher weight, implying that owners avoid such states voluntarily. A sufficient assumption is therefore the following.

Assumption 1.1. $c < \frac{1}{G-1}\left(\frac{B-1/p}{2}\right)^2$

1.3 The financing decision

We are essentially considering a standard model of insurance under adverse selection. Its structure is similar to that of Rothschild and Stiglitz (1976) and Wilson (1977). Standard results on the equilibrium apply. There are two outcomes that candidate for equilibrium, one separating and one pooling. In the separating outcome type G owners combine short-term borrowing with swaps to hedge interest rate risk, and type B owners eliminate risk completely by borrowing long-term. In the pooling outcome both types hedge interest risk entirely. As a consequence, and as a standard result, the high-risk type, i.e. a B owner, is always fully insured. The social cost of the market imperfection, i.e. the presence of adverse selection, is that in the separating outcome, a G owner carries a risk, which would be more efficiently borne by an outside investor.

It is well recognized in the literature, that in this model a perfect Bayesian equilibrium may not exist. To address this, I decompose the analysis. First I derive best contracts the from owners' perspective. These are the contracts, that maximize firm owner utility, and yield zero profit to creditors. In addition, for separating contracts I restrict attention to offers, where creditors do not cross-subsidize between firm types. Second I examine conditions, under which best contracts can be supported in equilibrium. At this stage separating contracts with cross-subsidization are also considered.

To derive best separating and best pooling contracts, the firm owners' financing decisions need to be examined. This is done in Sections 1.3.1 and 1.3.2 for long- and short-term contracts respectively. In Section 1.3.3 the creditors' zero-profit condition is investigated, and with the results obtained separating and pooling contracts can be characterized in Section 1.3.4. Section 1.3.5 pursues a graphical illustration of the separating outcome.

1.3.1 Demand for long-term contracts

Suppose a firm of type G is considering long-term financing. The repayment at date 2 is N_{LT}, and wealth at date 2 is therefore $\widetilde{X}_{G,\,LT} = G - N_{LT}$, provided $N_{LT} < G$. Expected utility prior to date 2 is

$$E\left[u\left(\widetilde{X}_{G,\,LT}\right)\right] = -\frac{1}{\left(c+G-N_{LT}\right)^2} \tag{1.1}$$

When examining the financing decision for a type B firm, it is sufficient to consider decisions conditional on the project being successful. This follows, first because the owner's wealth in case of default will always be zero. Second, because by Assumption 1.1 the default probability is unaffected by the financing decision. The expected utility from borrowing long-term is, provided $N_{LT} < B$,

$$E\left[u\left(\widetilde{X}_{B,\,LT}\right)|\text{success}\right] = -\frac{1}{\left(c+B-N_{LT}\right)^2} \tag{1.2}$$

It can be noted at this instant, that a possible pooling contract is long-term lending at the nominal amount $N_{LT} = 1/\overline{p}$, where $\overline{p} = \alpha + (1-\alpha)\,p$. For this nominal amount creditors make zero profit in expectation. In what follows, the focus is on separating contracts.

1.3.2 Demand for short-term contracts

When analyzing short-term contracts, I start by considering demand in the absence of interest rate swaps. Here Lemma 1.1 gives expected utility in the presence of interest rate risk. I then introduce interest rate swaps. For a given short-term contract firm owners' optimal swap positions are given in Lemma 1.2, and implied expected utilities are given in Lemma 1.3.

Suppose therefore initially, that swaps are not available. The repayment from a firm to its creditor at date 2 is $N_{ST}\,(1+\widetilde{r})$. For a type G firm date 2 wealth is, provided $N_{ST}\,(1+\widetilde{r}) < G$,

$$\begin{aligned}\widetilde{X}_{G,\,ST} &= G - N_{ST}\,(1+\widetilde{r}) \\ &= G - N_{ST} - \widetilde{r}N_{ST}\end{aligned}$$

The owner's financial position can be divided into two parts, a fixed quantity $G - N_{ST}$, which is expected wealth, and a quantity at risk N_{ST}, which is the amount exposed to interest rate risk. The size of N_{ST} depends on wether a separating or a pooling contract is considered. To infer expected utility, the following lemma is useful.

Lemma 1.1. *If date 2 wealth is*

$$\widetilde{X} = C - \widetilde{r}D$$

where $C > 0$, $D > 0$ and $C - \Delta D \geqslant 0$, then expected utility prior to date 2 is

$$E\left[u\left(\widetilde{X}\right)\right] = -\frac{1}{\left(c+C\right)^2 - \Delta^2 D^2}$$

Using Lemma 1.1, expected utility prior to date 2 for a type G firm owner is, provided $N_{ST}(1+\Delta) < G$,

$$E\left[u\left(\widetilde{X}_{G,ST}\right)\right] = -\frac{1}{(c+G-N_{ST})^2 - \Delta^2 N_{ST}^2}$$

Notice the resemblance with utility when borrowing long-term. The denominator is reduced by the term $\Delta^2 N_{ST}^2$ as a result of the interest rate risk exposure.

In a similar way utility for a type B owner borrowing short-term, conditional on the project being successful, is, provided $N_{ST}(1+\Delta) < B$,

$$E\left[u\left(\widetilde{X}_{B,ST}\right)|\text{success}\right] = -\frac{1}{(c+B-N_{ST})^2 - \Delta^2 N_{ST}^2}$$

Because $B < G$, and as a result of decreasing absolute risk aversion, a type B owner has higher disutility of interest rate risk exposure. This opens the possibility of separating contracts. In the separating outcome B owners borrow long-term, and G owners borrow short-term. Assuming creditors make zero profit in expectation, the long-term nominal amount is then $N_{LT} = 1/p$, while the short-term nominal amount is $N_{ST} = 1$. Type B owners have no incentive to select short-term financing, if the cost of interest rate exposure is sufficiently high. Type G owners have no incentive to borrow long-term, if the difference between short- and long-term nominal amounts is sufficiently high.

Note that the B owner is fully insured. The G owner separates from B firms at the cost of a full interest rate risk exposure. If it was possible to partially insure against interest rate risk, this could potentially reduce the cost of separation for the G owner. We are thus naturally led to consider interest rate swaps.

Suppose therefore, that interest rate swaps are made available. The creditor offers a contract (N_{ST}, λ), where N_{ST} is the nominal amount and λ is the swap spread. The firm owner chooses a swap position s at her own preference. For a swap position s, the transfer from the firm to its creditor at date 2 is

$$N_{ST}(1+\widetilde{r}) + s(\lambda - \widetilde{r}) = N_{ST} + s\lambda + \widetilde{r}(N_{ST} - s)$$

This means that an amount $N_{ST} - s$ is exposed to interest rate risk. The quantity $N_{ST} + s\lambda$ is the expected repayment. For a type G firm, the owner's financial position at date 2 is, provided the firm is solvent,

$$\widetilde{X}_{G,ST} = G - N_{ST} - s\lambda - \widetilde{r}(N_{ST} - s)$$

If the conditions of Lemma 1.1 are satisfied, the owner's expected utility prior to date 2 is

$$E\left[u\left(\tilde{X}_{G,\,ST}\right)\right] = -\frac{1}{\left(c+G-N_{ST}-s\lambda\right)^2 - \Delta^2\left(N_{ST}-s\right)^2} \tag{1.3}$$

Given N_{ST} and λ, the owner now chooses that swap amount, which maximizes (1.3). The decision for a type B owner is similar. The computations are carried out in the appendix, and the result is given by the following lemma.

Lemma 1.2. *For sufficiently small $\lambda > 0$ and $N_{ST} > 0$, the optimal swap positions for the respective firm owners are*

$$s_G = \frac{1}{1-\kappa}N_{ST} - \frac{\kappa}{1-\kappa}\frac{c+G-N_{ST}}{\lambda}$$
$$s_B = \frac{1}{1-\kappa}N_{ST} - \frac{\kappa}{1-\kappa}\frac{c+B-N_{ST}}{\lambda}$$

where $\kappa = \lambda^2/\Delta^2$.

As λ approaches zero, s_G and s_B approach N_{ST}. If $\lambda = 0$, the price of hedging is zero, and firm owners hedge interest rate risk completely. As λ increases, the optimal swap position decreases, and for sufficiently high λ it may become negative. In this case the expected profit to be made on the swap transaction, outweighs the disutility of interest rate risk exposure. Although this may be unlikely to represent a real life event, it has no fatal implications for the model.

As functions of λ, s_G and s_B are the demand functions for the respective owners. Figure 1.1 depicts these functions for a given value of N_{ST}. Since for $\lambda > 0$ the optimal swap position for a type B firm is larger than for a type G firm, it will be less beneficial for a type B firm to borrow short-term. The mechanism, to be utilized below, is that this additional cost can be controlled by the level of the swap spread, to support separation.

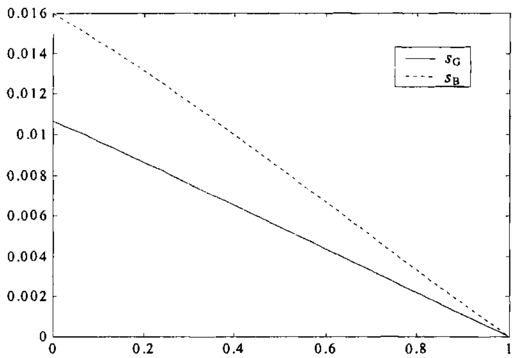

Figure 1.1. Demand for swaps.

Given Lemma 1.2, expected utility can be rewritten. Substituting for the optimal swap positions into the utility functions, gives implied expected utilities for a given swap spread. The result, again derived in the appendix, is given by the following lemma.

Lemma 1.3. *For sufficiently small* $\lambda > 0$ *and* $N_{ST} > 0$*, implied expected utilities for the respective firm owners, when swaps are available, are*

$$E\left[u\left(\widetilde{X}_{G,\,ST}\right)\right] = -\left(1-\lambda^2/\Delta^2\right)\frac{1}{\left(c+G-(1+\lambda)\,N_{ST}\right)^2}$$

$$E\left[u\left(\widetilde{X}_{B,\,ST}\right)|\text{success}\right] = -\left(1-\lambda^2/\Delta^2\right)\frac{1}{\left(c+B-(1+\lambda)\,N_{ST}\right)^2}$$

Note that as λ approaches zero, implied expected utility is similar to expected utility when borrowing long-term. This follows, because at a zero swap spread the firm attains a full hedge at zero cost. Consequently, the pooling outcome $N_{LT} = 1/\overline{p}$ has a short-term equivalent in $(N_{ST}, \lambda) = (1/\overline{p}, 0)$, where $\overline{p} = \alpha + (1-\alpha)\,p$.

1.3.3 The zero profit condition

With separating contracts type B owners choose long-term financing, and type G owners choose short-term financing. A creditor can offer both short- and long-term contracts. The expected profit on a long-term contract is

$$\pi_{LT} = N_{LT} - 1/p$$

and the expected profit on a short-term contract is

$$\pi_{ST} = N_{ST} + \lambda s_G - 1$$

If creditors offer contracts to both types of firms, the zero profit condition reads

$$\alpha\pi_{ST} + (1-\alpha)\,\pi_{LT} = 0$$

In principle one could allow for cross-subsidization between short- and long-term contracts, so that for example $\pi_{LT} > 0$ and $\pi_{ST} < 0$. However, for the moment attention is restricted to offers without cross-subsidization.[9] Thus, for the separating outcome the relevant zero-profit conditions are

$$N_{LT} = 1/p$$
$$N_{ST} + \lambda s_G = 1$$

Note, that in the short-term contract, there may be cross-subsidization between the swap position and the lending facility. If the swap spread is positive,

[9] Separating contracts with cross-subsidization are discussed in Section 1.4.

the creditor makes an expected profit on the swap transaction, which is off-set by a corresponding loss on lending.

Below it will be convenient to have the zero-profit conditions, conditional on G owners selecting their optimal swap positions s_G. Substituting for s_G from Lemma 1.2, the second zero-profit condition can be rewritten. The calculations are carried out in the appendix, and the result is given by Lemma 1.4.

Lemma 1.4. *The zero-profit conditions for long-term and short-term lending can be written as*

$$\begin{aligned} N_{LT} &= 1/p \\ (1+\lambda)\, N_{ST} &= 1 + \lambda^2/\Delta^2\,(c+G-1) \end{aligned}$$

1.3.4 The separating and pooling contracts

To start with I try to support separation. Thus I seek an outcome, where type B firms borrow long-term, and type G firms borrow short-term and use interest rate swaps. I assume that firm owners select their financing optimally according to Lemma 1.2, and that creditors make zero profit according to Lemma 1.4. For long-term financing expected utilities are obtained by substituting for the zero-profit condition $N_{LT} = 1/p$ into (1.1) and (1.2). For short-term financing, Lemmas 1.3 and 1.4 are combined.

Lemma 1.5. *For sufficiently small $\lambda > 0$ it holds that*

$$\begin{aligned} E\left[u\left(\widetilde{X}_{G,\,ST}\right)\right] &= -\left(1-\kappa\right)\frac{1}{\left(c+G-1-\kappa\left(c+G-1\right)\right)^2} \\ E\left[u\left(\widetilde{X}_{B,\,ST}\right)|\text{success}\right] &= -\left(1-\kappa\right)\frac{1}{\left(c+B-1-\kappa\left(c+G-1\right)\right)^2} \end{aligned}$$

where $\kappa = \lambda^2/\Delta^2$.

It is straightforward to show, that for a G owner borrowing short-term, utility is decreasing in λ. Thus, if a separating outcome can be supported for several values of λ, G owners will strictly prefer the contract corresponding to the minimal λ.

For a separating equilibrium the incentive compatibility constraints (IC constraints) must be satisfied. The first IC constraint states, that a type G firm owner borrowing short-term, should have no incentive to choose the long-term contract. The second IC constraint states, that a type B firm owner borrowing long-term, should have no incentive to seek short-term financing. That is

IC1 $\quad E\left[u\left(\widetilde{X}_{G,\,ST}\right)\right] \geqslant E\left[u\left(\widetilde{X}_{G,\,LT}\right)\right]$

IC2 $\quad E\left[u\left(\widetilde{X}_{B,\,LT}\right)\right] \geqslant E\left[u\left(\widetilde{X}_{B,\,ST}\right)\right]$

Heuristically, a type G owner has lower disutility of borrowing short-term. Thus, if the swap spread is set such, that a type B owner is just indifferent between long-term and short-term financing, then a type G owner should be better off borrowing short-term. A natural guess is therefore, that the separating swap spread should be such, that IC2 holds with equality. IC1 should then be satisfied automatically. This turns out to be correct, and we can state the following theorem.

Theorem 1.1. *There is a separating outcome, where firms of type B borrow long-term, and firms of type G borrow short-term. The long-term nominal amount is $N_{LT} = 1/p$, and the short-term nominal amount N_{ST} is given by*

$$N_{ST} + \lambda s_G = 1$$

where s_G was defined in Lemma 1.2, and the swap spread λ satisfies

$$\frac{\lambda^2}{\Delta^2} = a - \frac{1}{2}b^2 - \sqrt{\left(a - \frac{1}{2}b^2\right)^2 - a^2 + b^2}$$

with $\lambda > 0$ and

$$a = \frac{c + B - 1}{c + G - 1}$$

$$b = \frac{c + B - 1/p}{c + G - 1}$$

It remains to characterize the pooling contract. This has already been done in Sections 1.3.1 and 1.3.2, and we can thus formulate Theorem 1.2.

Theorem 1.2. *There are two pooling outcomes, corresponding to long- and short-term contracts respectively. In the former creditors offer both firms long-term financing at the nominal amount $N_{LT} = 1/\overline{p}$, where $\overline{p} = \alpha + (1 - \alpha)\,p$. In the latter creditors offer both firms short-term financing at $(N_{ST}, \lambda) = (1/\overline{p}, 0)$. In this case firms use interest rate swaps and hedge interest rate risk completely.*

1.3.5 A graphical representation

This section pursues a graphical illustration of the mechanism at work in the separating outcome. To start with, I examine short-term contracts in the absence of type B firms. For a type G firm, the demand function s_G, given in Lemma 1.2, is a function of λ and N_{ST}. This function thus defines a surface in (N_{ST}, λ, s)-space. For the supply side I consider those contracts, on which a creditor makes zero profit in expectation. From Section 1.3.3 the zero-profit condition is given by

$$N_{ST} + s\lambda = 1$$

which also defines a surface in (N_{ST}, λ, s)-space. The two surfaces are depicted in Figure 1.2. Since the creditor is constrained to offer contracts on the form (N_{ST}, λ), the set of admissible zero-profit contract offers is effectively constrained to the solid line in the figure. If the creditor could offer full triples (N_{ST}, λ, s), the set of admissible zero-profit contract offers would be the entire zero-profit surface.

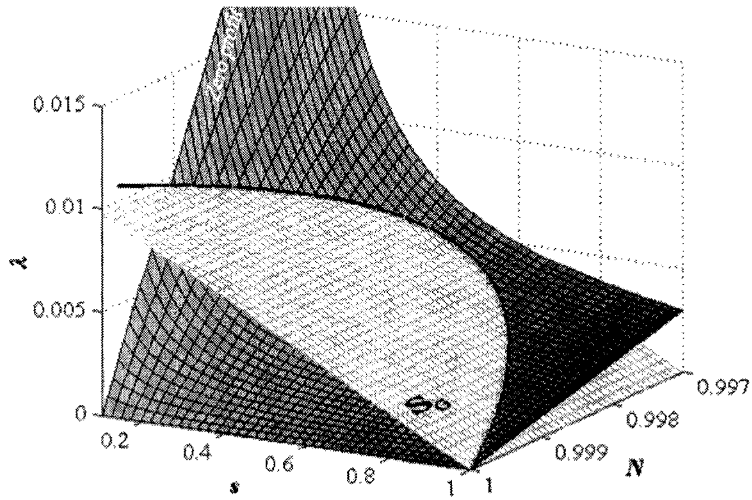

Figure 1.2. Admissible contracts.

Next allow type B owners, to choose short-term contracts also. The demand function is given by s_B, which similarly to s_G defines a surface in (N_{ST}, λ, s)-space. The alternative for B is to borrow long-term. Therefore, the iso-utility function corresponding to the implied utility of borrowing long-term is also considered. The iso-utility function represents combinations (N_{ST}, λ, s), between which B is indifferent. For a suitable utility level, this will give a representation of the incentive compatibility constraint. Comparing utility derived from long- and short-term financing respectively, the relevant iso-utility surface is given by

$$(c + B - N_{ST} - s\lambda)^2 - \Delta^2 (N_{ST} - s)^2 = (c + B - 1/p)^2$$

Figure 1.3 illustrates s_B and the iso-utility surface. By the definition of s_B, at each intersection of the two surfaces s_B characterizes the maximum of the iso-utility function along the s axis. Further, for a given N_{ST} the firm owner prefers a contract with as low swap spread as possible. Thus, a short-term contract (N_{ST}, λ) is preferred, if the corresponding s_B is below the iso-utility surface. To keep type B owners borrowing long-term, the creditor will have to offer such contracts (N_{ST}, λ), that s_B is above the iso-utility surface.

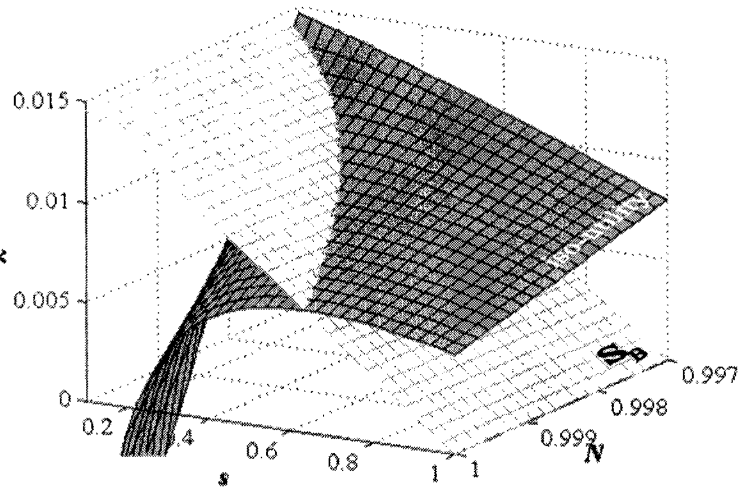

Figure 1.3. The incentive compatibility constraint.

To now illustrate the equilibrium, Figures 1.2 and 1.3 are combined. In Figure 1.4, the pale surfaces are the demand functions for the respective firms, where s_B has been 'cut off' for transparency. The dark convex surface is the zero-profit condition. Its intersection with the s_G surface, the line with the filled circles, is the set of admissible zero-profit contracts. I refer to this as the AC line (for admissible zero-profit contracts).

The dark concave surface is the firm B iso-utility function. Its intersection with the s_B surface, given by the line with the filled squares, illustrates those contract offers (N_{ST}, λ), for which a type B owner is indifferent to borrowing long-term. This line is referred to as the IC line (for incentive compatibility constraint).

The black dashed line is plotted for the equilibrium value of (N_{ST}, λ); it is parallel to the s-axis. It connects the AC line and the IC line. For a lower value of λ along AC, the IC constraint is violated. For a higher value of λ along AC, a potential outside creditor can attract all type G firms, by offering a contract along AC corresponding to a lower λ.

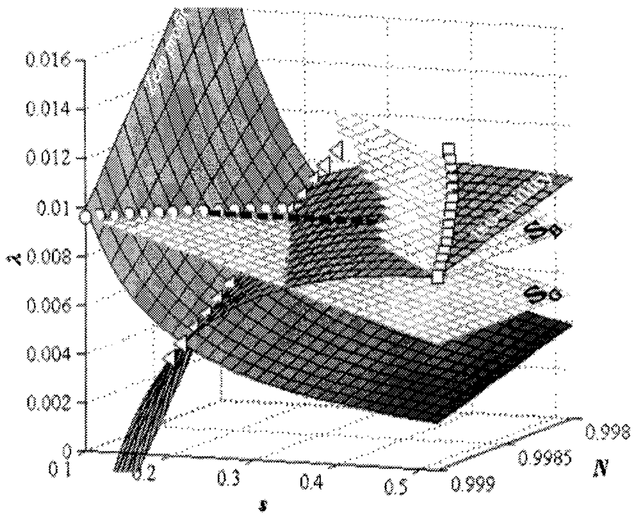

Figure 1.4. The separating contract.

It has been assumed, that creditors can only offer contracts on the form (N_{ST}, λ), and thus not full triples (N_{ST}, λ, s). If such contracts are allowed, it alters the outcome. The set of admissible zero-profit offers is then no longer constrained to the AC line, but is the entire zero-profit surface. To support separation in this case, the creditor still has to offer contracts above the iso-utility surface. Possible contracts are shown in the figure by the line with the filled triangles. However, when extending the set of admissible contracts, we do not find a unique solution - the type G firm owner is indifferent between the contracts along the triangles line. This matter is discussed further in Appendix 1.8.

1.4 Equilibrium

In the Rothschild and Stiglitz model a pooling perfect Bayesian equilibrium can never be supported. This translates directly to the current setting. Given the difference in risk tolerance, it is possible to offer a short-term contract, which leaves some risk exposure. Setting the level of risk exposure such, that type B owners are indifferent, type G owners are better off choosing the new contract. This results in expected losses for the initial contract. However, the deviating contract offer can not be part of an equilibrium itself.

By contrast the separating outcome is more robust. Although it sometimes can be broken by the pooling contract, there are parameter constellations for which the separating outcome remains intact. This will be the case if the fraction of type G firms is sufficiently small. Then the potential deviation, a long-term contract at the nominal amount $N_{LT} = 1/\overline{p}$, is close to the existing long-term contract, and the potential gain for type G owners from full insurance is small.

Theorem 1.3. *Assume that separating contracts with cross-subsidization are not allowed. If*

$$\alpha \leqslant \frac{p}{1-p} \frac{1}{1-\sqrt{1-\kappa}}$$

then a separating perfect Bayesian equilibrium [see Theorem 1.1] can be supported.

Next we allow for separating contracts with cross-subsidization between firm types. Doing that, it is for some parameter values possible, to break the separating outcome in Theorem 1.1, by a separating contract with cross-subsidization.

To see how this works, consider the separating outcome without cross-subsidization. The swap spread was set such, that type B owners were just indifferent between short- and long-term financing. With cross-subsidization, it is possible to maintain zero profit, while lowering the long-term nominal amount N_{LT} and increasing the short-term nominal amount N_{ST} accordingly. This contract is strictly preferred by type B owners, while type G owners at the initial stage are worse off. However, since N_{LT} is lowered, the IC constraint is relaxed, and it is consequently possible to lower the swap spread. This gives type G owners a higher degree of insurance. As a result the short-term contract (N_{ST}, λ) resulting under cross-subsidization, may indeed be preferred by type G firms to the initial short-term contract without cross-subsidization.

Although the new contract breaks the existing outcome, it may not in itself be part of a perfect Bayesian equilibrium. Since the creditor makes a positive profit on one type of borrower, it is possible for an outside creditor to approach only this borrower. This distorts the fraction of profitable contracts in the initial offer, and results in negative expected profits.

Somewhat comforting, there are parameter values for which the separating outcome in Theorem 1.1 is intact. If the fraction of type G firms is high, then a small increase of N_{ST}, allows for a large reduction of N_{LT}. With the implied relaxation of the IC constraint, the swap spread can be lowered more. Thus a small increase in the short-term nominal amount implies a significantly higher degree of risk coverage. On the other hand, if the fraction of type G firms is low, then a large increase of the short-term nominal amount, yields but a marginally higher degree of insurance. In this case the cross-subsidizing contract does not improve type G utility. Consequently, if the fraction of type G firms is low, then the separating outcome in Theorem 1.1 can still be supported.

To guarantee the existence of an equilibrium for general parameter constellations, it is common to alter the equilibrium concept. One way was proposed by Riley (1979). The idea is, to constrain the set of admissible deviating strategies in the equilibrium. Admissible deviations are those, which remain profitable even after a third deviating profitable outside offer is made. The intuition behind this equilibrium concept, is that possible rents from the deviating strategy, would be too short-lived to make a significant profit overall. Under the Riley equilibrium the Rothschild and Stiglitz model has a unique equilibrium, being the separating equilibrium, with no cross-subsidization between firm types. This result translates directly to the present setting.[10]

Theorem 1.4. *Assume that separating contracts with cross-subsidization are allowed. For general parameter constellations there is a unique Riley equilibrium. The equilibrium is the separating outcome given in Theorem 1.1.*

1.5 Conclusion

This paper has investigated how a positive swap spread may result on perfectly competitive capital markets. Similar to Titman (1992) and Mozumdar (2001) the model relies on financial distress costs and asymmetric information. However, in the present paper lenders use the swap spread to screen firms' credit quality. A positive swap spread puts a cost on hedging interest rate risk, which limits firms' use of interest rate swaps. As a consequence there is rationing of interest rate swaps.

A key feature of the model is that firms with lower expected cash flows have higher internal costs related to interest rate uncertainty. This makes interest rate hedging a higher concern, which may be utilized by creditors to separate the market.

One testable implication of the model is, that a higher degree of interest rate uncertainty, should imply a higher swap spread. Further it may be possible to test for insurance rationing. Extending the model to a multi-period setting, may produce implications for the term structure of swap spreads.

[10] An alternative was proposed by Wilson (1977). There the set of deviating strategies are those, which remain profitable even after existing strategies, which make a loss due to the deviation, are withdrawn. Depending on parameter values, a Wilson equilibrium may allow for cross-subsidization in separating contracts.

1.6 Appendix. Summary of variables and parameters

G	safe payoff for type G firm
B	payoff in case of success for type B firm
p	type B firm probability of success
α	fraction of type G firms
$\overline{p}$	average success probability $= \alpha + (1-\alpha)p$
$\widetilde{X}$	date 2 wealth to firm owner
$u(X)$	utility function $= -(c+X)^{-2}$
c	positive constant $< \frac{1}{G-1}\left(\frac{B-1/p}{2}\right)^2$
N_{LT}	nominal amount on long-term debt
N_{ST}	nominal amount on short-term debt
r	long-term risk-free rate $= 0$
$\widetilde{r}$	second period short risk-free rate
Δ	size of short rate distribution
λ	swap spread
π_{LT}	creditor's expected profit on long-term contract
π_{ST}	creditor's expected profit on short-term contract
κ	λ^2/Δ^2
a	$\frac{c+B-1}{c+G-1}$
b	$\frac{c+B-1/p}{c+G-1}$
ε	$\frac{1/p-1}{c+G-1}$
d	$\frac{c}{c+G-1}$
ξ	$\frac{1/\overline{p}-1}{c+G-1}$

1.7 Appendix. Proofs

1.7.1 Lemma 1.1

Suppose that $\widetilde{X} = C - \widetilde{r}D$, where $C > 0$, $D > 0$ and $C - \Delta D \geqslant 0$. The owner's expected utility can be written as

$$\begin{aligned} E\left[u\left(\widetilde{X}\right)\right] &= -E\left[\widetilde{X}^{-2}\right] \\ &= -\int_{-\Delta}^{\Delta} (C-rD)^{-2}\,\frac{dr}{2\Delta} \\ &= -\frac{1}{2\Delta D}\left[\frac{1}{C-rD}\right]_{-\Delta}^{\Delta} \\ &= -\frac{1}{2\Delta D}\left(\frac{1}{C-\Delta D} - \frac{1}{C+\Delta D}\right) \\ &= -\frac{1}{C^2 - \Delta^2 D^2} \end{aligned} \tag{1.4}$$

1.7.2 Lemma 1.2

To find the optimal swap position, I look at the first order condition. A type G firm is considered for illustration. The decision for a type B firm is similar. Throughout it is assumed that $0 < \lambda < \Delta$. The financial position at date 2 is

$$\widetilde{X}_{G,\,ST} = G - N_{ST} - \lambda s - \widetilde{r}\,(N_{ST} - s)$$

If the conditions of Lemma 1.1 are satisfied, then we have expected utility according to

$$E\left[u\left(\widetilde{X}_{G,\,ST}\right)\right] = -\frac{1}{C^2 - \Delta^2 D^2}$$

with

$$\begin{aligned} C &= c + G - N_{ST} - \lambda s \\ D &= N_{ST} - s \end{aligned} \tag{1.5}$$

In this case, differentiating expected utility with respect to s, yields

$$\frac{\partial}{\partial s} E\left[u\left(\widetilde{X}_{G,\,ST}\right)\right] = \frac{-2\lambda C + 2\Delta^2 D}{\left(C^2 - \Delta^2 D^2\right)^2}$$

Setting the first derivative equal to zero, gives the first order condition

$$-\lambda C + \Delta^2 D = 0 \tag{1.6}$$

The second derivative is

$$\frac{\partial^2}{\partial s^2} E\left[u\left(\widetilde{X}_{G,\,ST}\right)\right] = 2\frac{\lambda^2 - \Delta^2}{\left(C^2 - \Delta^2 D^2\right)^2} - 8\frac{\left(-\lambda C + \Delta^2 D\right)^2}{\left(C^2 - \Delta^2 D^2\right)^3}$$

At the optimum the second term is zero, and since $\lambda < \Delta$ the first term is negative. The first order condition thus characterizes a maximum. Substituting for C and D from (1.5) into the first order condition, we obtain

$$-\lambda\left(c + G - N_{ST} - s\lambda\right) + \Delta^2\left(N_{ST} - s\right) = 0$$

Solving for s, the optimal swap position s_G should satisfy

$$s_G = \frac{\Delta^2}{\Delta^2 - \lambda^2} N_{ST} - \frac{\lambda^2}{\Delta^2 - \lambda^2} \frac{c + G - N_{ST}}{\lambda} \tag{1.7}$$

To validate the use of Lemma 1.1, it should be verified, that for this choice of s, it holds that $C > 0$, $D > 0$ and $C - \Delta D \geqslant 0$. It is sufficient to check $D > 0$ and $C - \Delta D \geqslant 0$. For D, note that s_G can be seen as a weighted sum between N_{ST} and $\left(c + G - N_{ST}\right)/\lambda$ with a negative weight on the latter. Therefore we will have $s_G < N_{ST}$ if and only if

$$\frac{c+G-N_{ST}}{\lambda} > N_{ST}$$

which holds exactly as

$$(1+\lambda)\,N_{ST} < c+G$$

This will hold for sufficiently small N_{ST} and λ, as stated in the lemma. It remains to verify that $C - \Delta D \geqslant 0$. That is

$$C - \Delta D \equiv G - N_{ST} - \lambda s_G - \Delta\left(N_{ST} - s_G\right) \geqslant 0$$

For small values of λ, s_G is close to N_{ST}, which brings $C - \Delta D$ close to $G - N_{ST}$. For sufficiently small N_{ST} this quantity will be positive. This shows, that for the proposed value of s_G, the conditions of Lemma 1.1 are satisfied.

To close the argument, it should be verified, that when the conditions in Lemma 1.1 are not met, no other swap position attains the same level of utility. If default is possible due to adverse interest rate realizations, then there is a $\overline{r} \leqslant \Delta$, such that

$$C - c - \overline{r}D = 0$$

For a swap position sufficiently large in magnitude, provided $\lambda < \Delta$, this can always be set to hold. Further it is evident, that for $\lambda > 0$ a negative swap position will be more beneficial than a positive. Expected utility if default is possible, is given by

$$\begin{aligned}
E\left[u\left(\widetilde{X}\right)\right] &= -\int_{-\Delta}^{\overline{r}} \frac{1}{(C-rD)^2}\frac{dr}{2\Delta} - \int_{\overline{r}}^{\Delta} \frac{1}{c^2}\frac{dr}{2\Delta} \\
&= -\frac{1}{2\Delta D}\left[\frac{1}{C-rD}\right]_{-\Delta}^{\overline{r}} - \frac{1}{2\Delta}\frac{1}{c^2}\left[r\right]_{\overline{r}}^{\Delta} \\
&= -\frac{1}{2\Delta D}\left(\frac{1}{C-\overline{r}D} - \frac{1}{C+\Delta D}\right) - \frac{1}{2\Delta}\frac{1}{c^2}\left(\Delta - \overline{r}\right) \\
&= -\frac{1}{2\Delta D}\left(\frac{1}{c} - \frac{1}{C+\Delta D}\right) - \frac{1}{2\Delta}\frac{1}{c^2}\frac{1}{D}\left(\Delta D - C + c\right) \\
&= -\frac{1}{2\Delta D}\frac{1}{c}\left(1 - \frac{c}{C+\Delta D} + 1 - \frac{C-\Delta D}{c}\right)
\end{aligned}$$

It can be shown, that for sufficiently small c the derivative with respect to s is negative for all $s < N_{ST}$, which implies that a local minimum will not exist. As the swap position is set to be negative and large in magnitude, C is in the order of $-s\lambda$, and D is in the order of $-s$. Thus

$$
\begin{aligned}
\lim_{s\downarrow-\infty} E\left[u\left(\widetilde{X}\right)\right] &= \lim_{s\downarrow-\infty} -\frac{1}{2\left(-\Delta s\right)}\frac{1}{c}\left(1-\frac{c}{-s\lambda-\Delta s}+1-\frac{-\lambda s+\Delta s}{c}\right) \\
&= -\frac{1}{2c^2}\frac{\Delta-\lambda}{\Delta} \\
&= -\left(1-\lambda/\Delta\right)\frac{1}{2c^2}
\end{aligned}
$$

This is the case for speculative swaps. The owner chooses a negative swap position arbitrarily large in magnitude. The downside of date 2 wealth is bounded below by zero because of limited liability, while the upside is unbounded. However, because of the shape of the utility function, it will be possible to abstract from such behavior. In the proof of Theorem 1.1 below it is shown, that under Assumption 1.1 speculative swaps can be ruled out in equilibrium.

1.7.3 Lemma 1.3

As in the previous proof, a type G firm is considered for illustration. Implied expected utility is obtained by combining Lemmas 1.1 and 1.2. By the first order condition (1.6), $C = \Delta^2 D/\lambda$. We thus have

$$
\begin{aligned}
E\left[u\left(\widetilde{X}_{G,ST}\right)\right] &= -\frac{1}{\left(\Delta^2 D/\lambda\right)^2 - \Delta^2 D^2} \\
&= -\frac{1}{\Delta^2}\frac{1}{\left(D/\lambda\right)^2}\frac{1}{\Delta^2-\lambda^2}
\end{aligned}
\tag{1.8}
$$

From Lemma 1.2 D/λ is calculated as

$$
\begin{aligned}
D/\lambda &\equiv \frac{1}{\lambda}\left(N_{ST}-s_G\right) \\
&= \frac{1}{\lambda}\frac{\lambda^2}{\Delta^2-\lambda^2}\left(-N_{ST}+\frac{c+G-N_{ST}}{\lambda}\right) \\
&= \frac{1}{\Delta^2-\lambda^2}\left(c+G-\left(1+\lambda\right)N_{ST}\right)
\end{aligned}
$$

and substituting for D/λ into (1.8), yields the implied expected utility

$$
E\left[u\left(\widetilde{X}_{G,ST}\right)\right] = -\frac{\Delta^2-\lambda^2}{\Delta^2}\frac{1}{\left(c+G-\left(1+\lambda\right)N_{ST}\right)^2}
\tag{1.9}
$$

1.7.4 Lemma 1.4

For a separating equilibrium the zero-profit condition for short-term contracts is

$$
N_{ST}+\lambda s_G = 1
$$

Substituting for s_G from Lemma 1.2, yields

$$N_{ST} + \frac{\Delta^2}{\Delta^2 - \lambda^2}\lambda N_{ST} - \frac{\lambda^2}{\Delta^2 - \lambda^2}(c + G - N_{ST}) = 1$$

and collecting terms in N_{ST},

$$\frac{\Delta^2}{\Delta^2 - \lambda^2}(1 + \lambda) N_{ST} = 1 + \frac{\lambda^2}{\Delta^2 - \lambda^2}(c + G)$$

Finally, multiplying with $\left(\Delta^2 - \lambda^2\right)/\Delta^2$, it can be concluded, that

$$(1 + \lambda) N_{ST} = 1 + (\lambda/\Delta)^2 (c + G - 1) \tag{1.10}$$

which was to be shown.

1.7.5 Theorem 1.1

The proof is divided into four parts. The first part finds a λ, for which IC2 holds with equality. The second part shows, that for this choice of λ the IC1 automatically is satisfied. The third part shows that the solution is valid in the sense, that for an interior solution, i.e. for s_G given by (1.7), default can not occur due to adverse interest rate realizations. The last part shows, that the interior solution is the global solution.

Part 1. Let $\kappa = \frac{\lambda^2}{\Delta^2}$. If IC2 binds, then

$$-(1 - \kappa)\frac{1}{(c + B - 1 - \kappa(c + G - 1))^2} = -\frac{1}{(c + B - 1/p)^2} \tag{1.11}$$

Define

$$a = \frac{c + B - 1}{c + G - 1} \tag{1.12}$$
$$b = \frac{c + B - 1/p}{c + G - 1}$$

Dividing the denominator on each side of (1.11) by $(c + G - 1)^2$, we then have

$$\frac{(c + G - 1)^2}{(c + B - 1 - \kappa(c + G - 1))^2} = \frac{1}{(a - \kappa)^2}$$

$$-(1 - \kappa)\frac{1}{(a - \kappa)^2} = -\frac{1}{b^2} \tag{1.13}$$

which can be rearranged to

$$\kappa^2 - 2\left(a - \frac{1}{2}b^2\right)\kappa + a^2 - b^2 = 0 \tag{1.14}$$

Solving for κ, yields

$$\kappa = a - \frac{1}{2}b^2 \pm \sqrt{\left(a - \frac{1}{2}b^2\right)^2 - a^2 + b^2}$$

For this choice of κ IC2 is satisfied with equality. It is now verified that the solution is valid, in the sense that λ is a real number. First, κ is a real number, since

$$\left(a - \frac{1}{2}b^2\right)^2 - a^2 + b^2 = (1-a)\,b^2 + \frac{1}{4}b^4 > 0$$

where the last inequality follows from the fact that $a < 1$. Second, κ must be positive for λ to be a real number. This will clearly be the case if we select the positive root. However, since

$$a - \frac{1}{2}b^2 > a - \frac{1}{2}b > 0$$

and

$$-a^2 + b^2 < 0$$

the negative root also gives a positive κ. Choosing the negative root, secures a lower value of λ, which is strictly preferred by type G firm owners. Consequently the negative root should be chosen. Finally, since $a < 1$ it follows that $\kappa < 1$, which in turn implies that $\lambda < \Delta$. Concluding we have

$$\kappa = a - \frac{1}{2}b^2 - \sqrt{\left(a - \frac{1}{2}b^2\right)^2 - a^2 + b^2}$$

Verifying that for the implied value of λ, both s_G and s_B result in non-negative financial positions at date 2 with probability 1, is left to part 3 below.

Part 2. It is now verified, that the firm G IC constraint is satisfied at the proposed swap spread. To do this, it is shown that

$$-(1-\kappa)\frac{1}{(c+G-1-\kappa(c+G-1))^2} \geqslant -\frac{1}{(c+G-1/p)^2}$$

Multiplying with $(c+G-1)^2$ and introducing a new variable x, the following inequality should be satisfied for $x = 1$

$$-(1-\kappa)\frac{1}{(x-\kappa)^2} \geqslant -\frac{1}{(x-\varepsilon)^2}$$

where $\varepsilon = a - b$, and a and b were defined in (1.12). Since $\kappa < 1$, the inequality can be rewritten as

$$(x-\kappa)^2 - (1-\kappa)(x-\varepsilon)^2 \geqslant 0$$

Defining

$$f(x) = (x-\kappa)^2 - (1-\kappa)(x-\varepsilon)^2$$

we have that $f(a) = 0$. It will be sufficient to show that $f(1) \geqslant 0$. Now f is a second degree polynomial in x, with a positive coefficient for x^2. Thus, if $f'(a) > 0$, then $f(x) > 0$ for all $x > a$. Since $a < 1$, it will be sufficient to show that $f'(a) > 0$. The derivative of f is given by

$$f'(x) = 2(x-\kappa) - 2(1-\kappa)(x-\varepsilon)$$

and since $f(a) = 0$ we have

$$(a-\kappa)^2 = (1-\kappa)(a-\varepsilon)^2$$

Taking the square root, and recalling $\kappa < a$ and $\varepsilon < a$, this gives

$$(a-\kappa) = \sqrt{1-\kappa}\,(a-\varepsilon)$$

which, since $0 < \kappa < 1$, certifies that

$$(a-\kappa) > (1-\kappa)(a-\varepsilon)$$

ensuring $f'(a) > 0$.

Part 3. It remains to show, that for κ given by (1.13) both s_G and s_B produce non-negative financial positions at date 2. It should thus be verified that

$$\widetilde{X}_{G,\min} \equiv G - N_{ST} - \lambda s_G - \Delta(N_{ST} - s_G) \geqslant 0$$
$$\widetilde{X}_{B,\min} \equiv B - N_{ST} - \lambda s_B - \Delta(N_{ST} - s_B) \geqslant 0$$

I start by rewriting the expression for $\widetilde{X}_{G,\min}$. For $G - N_{ST} - \lambda s_G$ it holds, that

$$\begin{aligned} G - N_{ST} - \lambda s_G &= G - N_{ST} - \lambda\left(\frac{1}{1-\kappa}N_{ST} - \frac{\kappa}{1-\kappa}\frac{c+G-N_{ST}}{\lambda}\right) \\ &= G - N_{ST} - \frac{1}{1-\kappa}\left(\lambda N_{ST} - \kappa(c+G) + \kappa N_{ST}\right) \\ &= \frac{1}{1-\kappa}\left(G - (1+\lambda)N_{ST}\right) + \frac{\kappa}{1-\kappa}c \end{aligned}$$

and for $N_{ST} - s_G$, that

$$\begin{aligned} N_{ST} - s_G &= N_{ST} - \left(\frac{1}{1-\kappa}N_{ST} - \frac{\kappa}{1-\kappa}\frac{c+G-N_{ST}}{\lambda}\right) \\ &= -\frac{\kappa}{1-\kappa}\left(N_{ST} - \frac{c+G-N_{ST}}{\lambda}\right) \\ &= -\frac{\kappa}{1-\kappa}\frac{1}{\lambda}\left((1+\lambda)N_{ST} - G - c\right) \\ &= \frac{\kappa}{1-\kappa}\frac{1}{\lambda}\left(G - (1+\lambda)N_{ST}\right) + \frac{\kappa}{1-\kappa}\frac{1}{\lambda}c \end{aligned}$$

Thus

$$\widetilde{X}_{G,\min} = \frac{1}{1-\kappa}\left(G - (1+\lambda)\,N_{ST}\right)\left(1 - \kappa\frac{\Delta}{\lambda}\right) + \frac{\kappa}{1-\kappa}\left(1 - \frac{\Delta}{\lambda}\right)c$$

A similar calculation for $\widetilde{X}_{B,\min}$ yields

$$\widetilde{X}_{B,\min} = \frac{1}{1-\kappa}\left(B - (1+\lambda)\,N_{ST}\right)\left(1 - \kappa\frac{\Delta}{\lambda}\right) + \frac{\kappa}{1-\kappa}\left(1 - \frac{\Delta}{\lambda}\right)c$$

Evidently $\widetilde{X}_{B,\min} < \widetilde{X}_{G,\min}$, so it will be sufficient to show that $\widetilde{X}_{B,\min} \geqslant 0$.

Multiplying the last expression with $\frac{1-\kappa}{\kappa} > 0$, the relevant condition reads

$$\left(B - (1+\lambda)\,N_{ST}\right)\frac{\Delta}{\lambda}\left(\frac{\Delta}{\lambda} - 1\right) + \left(1 - \frac{\Delta}{\lambda}\right)c \geqslant 0$$

or, dividing by $\Delta/\lambda - 1 > 0$,

$$\left(B - (1+\lambda)\,N_{ST}\right)\frac{\Delta}{\lambda} - c \geqslant 0$$

Inserting the zero-profit condition (1.10), yields

$$\left(B - 1 - \lambda^2/\Delta^2\,(c + G - 1)\right)\frac{\Delta}{\lambda} - c \geqslant 0$$

which, after rearranging, becomes

$$\frac{B-1}{c+G-1} - \frac{\lambda^2}{\Delta^2} \geqslant \frac{\lambda}{\Delta}\frac{c}{c+G-1}$$

Using the expressions for a and b from (1.12), and defining

$$d = \frac{c}{c+G-1}$$

the condition can be rewritten as

$$a - d - \kappa \geqslant \sqrt{\kappa}d$$

Since $\kappa < 1$ it will be sufficient, that

$$a - \kappa \geqslant 2d \tag{1.15}$$

Using the expression for κ, it will be sufficient to have

$$\frac{1}{2}b^2 \geqslant 2d$$

That is

$$\left(\frac{c+B-1/p}{c+G-1}\right)^2 \geqslant 4\frac{c}{c+G-1}$$

The left hand side is increasing in c. It will be sufficient to have

$$\left(\frac{B-1/p}{G-1}\right)^2 \geqslant 4\frac{c}{G-1}$$

That is

$$c \leqslant \frac{1}{G-1}\left(\frac{B-1/p}{2}\right)^2$$

and Assumption 1.1 certifies, that the inequality is satisfied.

Part 4. It is now demonstrated that speculative swaps are suboptimal in equilibrium. In the proof of Lemma 1.2 it was shown, that as the swap position approaches $-\infty$, it holds that

$$\lim_{s\downarrow-\infty} E\left[u\left(\tilde{X}\right)\right] = -(1-\lambda/\Delta)\frac{1}{2c^2}$$

This holds for both firm G and firm B owners. It is now demonstrated that in equilibrium the owners' expected utility for $s = s_G$ and $s = s_B$ respectively are higher. It will be sufficient to show that

$$-(1-\lambda/\Delta)\frac{1}{2c^2} \leqslant -(1-\kappa)\frac{1}{(c+B-1-\kappa(c+G-1))^2}$$

With notation given above, this is equivalent to

$$\frac{1}{1-\sqrt{\kappa}}2d^2 \leqslant \frac{1}{1-\kappa}(a-\kappa)^2$$

or, after multiplying with $1-\sqrt{\kappa}$ and rearranging,

$$(a-\kappa)^2 \geqslant 2\left(1+\sqrt{\kappa}\right)d^2$$

Since $\kappa < 1$, it will be sufficient to have

$$a-\kappa \geqslant 2d$$

This is exactly condition (1.15), which in the previous section was shown to hold under Assumption 1.1. This completes the proof.

1.7.6 Theorem 1.3

This appendix derives the condition on α, under which the separating outcome can be supported as a perfect Bayesian equilibrium. The separating outcome can be sustained, if it can not be upset by the pooling contract. The pooling contract is strictly preferred by type B owners, but to break the separating outcome, it will be necessary and sufficient, that it is preferred by type G owners also. The separating contract can thus be supported as an equilibrium if and only if

$$-(1-\kappa)\frac{1}{(c+G-1-\kappa(c+G-1))^2} \geqslant -\frac{1}{(c+G-1/\overline{p})^2}$$

Dividing the denominator on each side by $(c+G-1)^2$, this is equivalent to

$$-(1-\kappa)\frac{1}{(1-\kappa)^2} \geqslant -\frac{1}{(1-\xi)^2}$$

where $\xi = \frac{1/\overline{p}-1}{c+G-1}$. Rearranging we get the condition

$$(1-\xi)^2 \leqslant (1-\kappa)$$

or equivalently, since $\xi < 1$ and $\kappa < 1$,

$$\xi \geqslant 1-\sqrt{1-\kappa}$$

Using

$$\overline{p} = \alpha + (1-\alpha)p = p + \alpha(1-p)$$

this can be written as

$$p + \alpha(1-p) \leqslant \frac{1}{1-\sqrt{1-\kappa}}$$

and solving for α we obtain the necessary and sufficient condition

$$\alpha \leqslant \frac{p}{1-p}\frac{1}{1-\sqrt{1-\kappa}}$$

1.8 Appendix. An alternative contracting scheme

This appendix investigates the implications of letting the swapped amount be explicitly incorporated in the contract. Thus for short-term contracts, in stead of just offering (N_{ST}, λ) and letting the firm owner choose her optimal s, creditors now offer a full triples (N_{ST}, λ, s).

The previous comparison with the Rothschild and Stiglitz framework still applies. There is a separating equilibrium, where firms of type G borrow short-term, and firms of type B borrow long-term.

However, explicitly incorporating s in the contract, affects the properties of the equilibrium. A larger set of contracts is made available, and this gives a more efficient outcome. At the same time multiple equilibria arise. In particular the swap spread becomes indeterminate.

To see this, recall from Section 1.3.2, that the financial position at date 2 of a type G firm borrowing short-term, is given by $\widetilde{X}_{G,ST} = G - N_{ST} - s\lambda - \widetilde{r}\,(N_{ST} - s)$. In equilibrium creditors make zero profit on each type of borrower, which for short-term financing contracts implies that $N_{ST}+s\lambda = 1$. Substituting for N_{ST}, the financial position at date 2 is then

$$\begin{aligned}\widetilde{X}_{G,\,ST} &= G - N_{ST} - s\lambda - \widetilde{r}\,(N_{ST} - s)\\ &= G - 1 - \widetilde{r}\,(1 - s\,(1+\lambda))\end{aligned}$$

To interpret this, note that the fair expected debt repayment from firm G to the creditor is unity. Thus $s\,(1+\lambda)$ defines that fraction of the repayment, which is set to be at a fixed rate. In the previous analysis the swap spread effectively played the role of the fixed rate, which could be varied to control the firm owner's behavior. Here this fixed rate is constrained to be zero, and the creditor controls the firm's interest rate risk exposure by varying the level of $s\,(1+\lambda)$. However, the firm owner is indifferent between those combinations of s and λ, for which $s\,(1+\lambda)$ is constant. The equilibrium will specify a level of $s\,(1+\lambda)$, such that type B firm owners are just indifferent to mimicing the short-term financing strategy. While this defines a unique value of $s\,(1+\lambda)$, there will be an indeterminacy regarding the swap spread.

To find the separating value of $s\,(1+\lambda)$, set the IC2 to hold with equality. That is,

$$-\frac{1}{(c+B-1)^2 - \Delta^2\,(1-s\,(1+\lambda))^2} = -\frac{1}{(c+B-1/p)^2}$$

Solving for $(1-s\,(1+\lambda))^2$ yields

$$(1-s\,(1+\lambda))^2 = \frac{1}{\Delta^2}\left((c+B-1)^2 - (c+B-1/p)^2\right)$$

and[11]

$$s\,(1+\lambda) = 1 - \frac{1}{\Delta}\sqrt{(c+B-1)^2 - (c+B-1/p)^2}$$

We can now state the following theorem.

[11] The solution is given with a negative sign of the square root, which ensures that the date 2 financial position is non-negative.

Theorem 1.3.1. *There is a separating Riley equilibrium, where firms of type G borrow short-term and use swaps to hedge interest rate risk, and firms of type B borrow long-term and stay away from the swap market. The long-term contract is given at the nominal amount $N_{LT} = 1/p$. The short-term contract is given by (N_{ST}, λ, s), where s and λ are related according to*

$$s\,(1+\lambda) = 1 - \frac{1}{\Delta}\sqrt{(c+B-1)^2 - (c+B-1/p)^2}$$

and the short-term nominal amount is determined by $N_{ST} + \lambda s = 1$.

Though Theorem 1.3.1 does not specify a particular swap spread, it shows that the model is consistent with *any* swap spread. A positive swap spread does not necessarily imply that markets are incompetitive. Creditors making expected profits on swap contracts, can make corresponding losses on lending.

However, allowing creditors to include s in the contract, affects the efficiency of the equilibrium. In the previous analysis, when s could not be explicitly incorporated, the expected utility for a type G owner in equilibrium was

$$E\left[u\left(\widetilde{X}_{G,ST}\right)\right] = -\frac{1}{1-\lambda^2/\Delta^2}\frac{1}{(c+G-1)^2}$$

where λ/Δ was defined in Theorem 1.1. When allowing for full triples, implied utility is

$$E\left[u\left(\widetilde{X}_{G,ST}\right)\right] = -\frac{1}{(c+G-1)^2 - \left((c+B-1)^2 - (c+B-1/p)^2\right)}$$

It is now demonstrated, that utility in the latter case is higher. It will be sufficient to show, that

$$(c+G-1)^2 - \left((c+B-1)^2 - (c+B-1/p)^2\right) \geqslant \left(1-\lambda^2/\Delta^2\right)(c+G-1)^2$$

This is equivalent to

$$\lambda^2/\Delta^2 \geqslant \left(\frac{c+B-1}{c+G-1}\right)^2 - \left(\frac{c+B-1/p}{c+G-1}\right)^2$$

or in terms of a and b from Theorem 1.1,

$$\lambda^2/\Delta^2 \geqslant a^2 - b^2$$

Substituting further for λ^2/Δ^2 from Theorem 1.1, yields

$$a - \frac{1}{2}b^2 - \sqrt{\left(a - \frac{1}{2}b^2\right)^2 - a^2 + b^2} \geqslant a^2 - b^2$$

Isolating the square on the right and side and squaring, it will be sufficient to show

$$\left(a-\frac{1}{2}b^2\right)^2-2\left(a-\frac{1}{2}b^2\right)(a^2-b^2)+(a^2-b^2)^2 \geqslant \left(a-\frac{1}{2}b^2\right)^2-(a^2-b^2)$$

Cancelling the first term on each side, and dividing by $a^2-b^2>0$, yields the equivalent

$$-2\left(a-\frac{1}{2}b^2\right)+(a^2-b^2) \geqslant -1$$

Cancelling b^2, this can be written as

$$(a-1)^2 \geqslant 0$$

Since $a<1$, this inequality will be sharp. Tracking the argument backwards it follows, that allowing for full triples (N_{ST}, λ, s), implies a strictly higher degree of efficiency in the separating equilibrium.

Graphically the loss in efficiency can be illustrated by revisiting Figure 1.4. When creditors could only offer pairs (N_{ST}, λ), they choose the offer corresponding to the dashed black line - that value of (N_{ST}, λ) where the corresponding values of s_G and s_B intersect with the AC and IC lines respectively. When creditors are allowed to include s, and offer contracts on the form (N_{ST}, λ, s), it is possible to offer contracts along the line with the filled triangles. This gives firm G owners a higher risk coverage at a lower price, and is thus strictly preferred to the previous outcome. It is evident from the figure, that the triangles line by construction is parallel to the firm B iso-utility surface. For the same reason it is parallel to firm G iso-utility surfaces. We can therefore not pin down a unique value of the swap spread.

References

1. Degrér, Henrik, 1999, The influence of credit risk on swap agreements, Working paper, School of economics, University of Lund, Sweden
2. Duffie, Darrel D. and Ming Huang, 1996, Swap rates and credit quality, Journal of Finance, 51(3), 921-949
3. Fehle, Frank, 1999, Market structure and swap spreads: International evidence, PhD thesis, University of South Carolina
4. Flannery, Mark J., 1986, Asymmetric information and risky debt maturity choice, Journal of Finance 41, 19-38
5. Hellwig, Martin, 1987, Some recent developments in the theory of competition in markets with adverse selection
6. Huge, Brian and David Lando, 1999, Swap pricing with two-sided default risk in a rating-based model, Working paper, Centre for analytical finance, University of Aarhus, Denmark
7. Kreps, D. and R. Wilson, 1982, Sequential equilibria, Econometrica 50, 863-894
8. Larry, H. P. Lang, Robert H Litzenberger and Andy Luchuan Liu, 1998, Journal of Banking and Finance, 22, 1507-1532
9. Mozumdar, Abon, 2001, Corporate hedging and speculative incentives: Implications for swap market default risk, Journal of Financial and Quantitative Analysis 36(2), 221-250
10. Myers, Stewart C., 1977, Determinants of corporate borrowing, Journal of Financial Economics 5, 147-176
11. Riley, John G., 1979, Informational equilibrium, Econometrica 47(2), 331-359
12. Ross, Stephen A., 1976, The determination of financial structure: the incentive signalling approach, The Bell Journal of Economics, 23-40
13. Rothschild, Michael and Joseph Stiglitz, 1976, Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, Quarterly Journal of Economics 90:4, 629-649
14. Samant, Ajay, 1996, An empirical study of interest rate swap usage by nonfinancial corporate business, Journal of Financial Research 10(1), 43-57
15. Saunders, Kent T., 1999, The interest rate swap: Theory and evidence, Journal of Corporate Finance 5, 55-78
16. Titman, Sheridan, 1992, Interest rate swaps and corporate financing choices, Journal of Finance 47(4), 1503-1516
17. Wall, Larry D., 1989, Interest rate swaps in an agency theoretic model with uncertain interest rates, Journal of Banking and Finance 13, 261-270
18. Wall, Larry D. and John .J. Pringle, 1989, Alternative explanations of interest rate swaps: A theoretical and empirical analysis. Financial Management 18, 59-73
19. Wilson, Charles A., 1977, A model of insurance markets with incomplete information, Journal of Economic Theory 16(2), 167-207

2. Monetary policy and the yield curve

Abstract

This paper integrates two branches of research, optimal monetary policy and no-arbitrage yield curve modeling. Transforming a standard model of optimal monetary policy to a continuous time setting produces a standard interest rate model with well-known characteristics. The model is estimated to US data November 1987 to April 2002. Results align with previous findings on monetary policy.

2.1 Introduction

In term structure modelling a vital part of the analysis is how the dynamics of the short rate is specified. Given that the short rate is so closely linked to the monetary authority, it is rather surprising that more sophisticated measures of this linkage have not received greater attention. In contrast there is a separate class of models that explicitly focus on the design of monetary policy. In its most simple form the central bank sets its instrument rate according to a Taylor (1993) rule, an approach that has since been extended by Svensson (2000), Rudebusch (2001) and others.

Usually these studies are conducted in a discrete time setting. The current paper reformulates a standard model on optimal monetary policy in continuous time, which directly translates into a dynamic system for short rate of affine constant volatility type. Complemented with assumptions on investors' risk preferences this completely characterizes the term structure of interest rates. If the market price of risk is an affine function of the state variables, the implied term structure model is affine in the state variables, and we obtain a multi-factor version of Vasicek (1977). This facilitates the analysis of term structure dynamics, and in particular we can relate term structure dynamics to the stance of monetary policy. Thus the model may be helpful for understanding differences in term structure dynamics under different monetary regimes.

[0] I am greatly appreciative to my adviser Paul Söderlind for excellent guidance throughout this work. I would also like to thank colleagues at the Department of Financial Economics and seminar participants at the Swedish Riksbank. In particular the paper improved significantly from the comments by Jesper Lindé, Ulf Söderström and Anders Vredin. I have also benefitted from valuable comments by Pierre Collin-Dufresne, Frank de Jong and Monika Piazzesi. Financial support from Bankforskningsinstitutet is gratefully acknowledged.

A standard model on monetary policy starts out by assuming that the dynamics of inflation and the output gap is governed by a Markov system, with a structure motivated by usual economic arguments. If the output gap is positive inflation has a positive drift, thus reflecting the traditional Philips curve. In a similar way the output gap follows a process with a drift that depends on the short-term real rate. If the real rate is high, output growth is slowed down. Due to the dependence on the short rate, it is possible for a monetary authority to control the system. In the current paper we follow Svensson (2000) by assuming that the monetary authority sets its instrument rate, the short rate, to control the system in a way that is optimal with respect to a particular objective function. The objective is to minimize unconditional variances in inflation and output growth, each with a certain weight.

It is not uncommon for papers on optimal monetary policy to look at the implied term structure of interest rates. The approach is usually to consider the yield on a long-term bond as the expected average of the short rate plus a term premium. A contribution of reformulating the model in terms of a standard term structure model, is that it facilitates a more elaborate treatment of the term premium. The term premium is related to risk compensation with respect to the each of the model's state variables. This in turn requires that we specify investors' attitude towards risk with respect to each state variable. In the current paper we draw on a recent result by Cochrane and Piazzesi (2001), who find that excess returns on bonds of various maturities are well described by a one-factor model. This allows us to reduce parameters in the specification of the market price of risk.

Given a term structure model, estimating it to data is an endeavor in itself. In the current situation, where the model's state variables are familiar economic quantities, there is a trade-off. On one hand one would like the variables to be in resemblance with regularly published figures such as the consumer price index or quarterly GDP. On the other hand one would like the state variables to be in accordance with observed term structure dynamics. In Ang and Piazzesi (2001) observed measures are used to explain as much term structure variation as possible, and then unobserved variables are incorporated to explain the remainder. Another approach is to use a Kalman filter, and along this branch we find Wu (2001) and Brennan, Wang and Xia (2002).

In the current paper the technique is somewhat different. For a given instantaneous covariance matrix of yields, there is only a limited set of three-factor models that are consistent with the observed covariance matrix. In a first step the set of covariance consistent three-factor realizations is characterized. This allows us to extract that realization of the observed term structure that best matches some publicly observed series of the state variables. In a second step the extracted factor series are taken to be the true observations, and the model for the monetary transmission mechanism is estimated with usual estimation techniques. In summary the two steps produce a model for the factor dynamics. In a third and final step the model obtained is taken to be the true model, and the market price of risk is estimated as a residual between term structure dynamics and factor dynamics.

The paper is organized as follows. In Section 2.2 the model for the monetary transmission mechanism is presented. This section also gives some general theory on term structure modeling and specifies a functional form for the market price of risk. Section 2.3 describes the estimation techniques. This includes how factors series are extracted from the yield curve, how the monetary model is estimated and how risk premia are estimated. Section 2.4 describes the data set, Section 2.5 presents the results and Section 2.6 summarizes.

2.2 The model

The description of the model is divided into two parts. The first part, in Section 2.2.1, defines the monetary transmission mechanism, and is the continuous time equivalent of standard papers on monetary policy. To study the implications for market prices, investors' risk preferences also need to be considered. Section 2.2.2 studies market prices and specifies a simple factor model for the market price of risk. Section 2.2.3 concludes with a numerical example.

2.2.1 The monetary transmission mechanism

The model is similar to Clarida, Galí and Gertler (1999). The state of the economy is represented by three state variables, the short nominal rate r_t, the instantaneous rate of inflation π_t, and the output gap y_t. For notational convenience inflation and the output gap are both assumed to have zero unconditional means. Their dynamics is governed by

$$\begin{aligned} d\pi_t &= a_{23} y_t dt + \sigma_\pi dw_{\pi t}, \\ dy_t &= a_{31}\left(r_t - \pi_t\right) dt + a_{33} y_t dt + \sigma_y dw_{yt}. \end{aligned} \tag{2.1}$$

The processes $w_{\pi t}$ and w_{yt} are standard Wiener processes with instantaneous correlation ρ_{23}. The volatilities σ_π and σ_y are positive constants. As for the drift term, the coefficient a_{23} is positive, which means that a positive output gap gives a positive drift in inflation. This captures the traditional Philips curve. The quantity $r_t - \pi_t$ is a measure of the instantaneous real rate. The coefficient a_{31} is negative, and consequently a positive instantaneous real rate contributes negatively to the output gap. The dependence on the short interest rate opens the way for a monetary authority to control the system. By adjusting the short nominal rate, the instantaneous real rate is altered. The coefficient a_{33} can have either sign. A standard utility based model would typically imply $a_{33} = 0$, so we here allow for richer dynamics. Parameter are "deep" in the sense that they are policy invariant.

The monetary authority controls the short nominal rate, but can not do so perfectly. The short nominal rate follows a stochastic process according to[1]

$$dr_t = u_t dt + \sigma_r dw_{rt},$$

where w_{rt} is a standard Wiener process with an instantaneous correlation to $w_{\pi t}$ and w_{yt} of ρ_{12} and ρ_{13} respectively. The volatility σ_r is an exogenous positive constant, and the drift u_t is an arbitrary adapted process. While the central bank has no means of influencing the volatility σ_r, it can specify exactly the drift u_t.

For u_t we follow Svensson (2000) by assuming that the monetary authority is equipped with a loss function according to

$$L = q_u \operatorname{Var}(u_t) + q_\pi \operatorname{Var}(\pi_t) + q_y \operatorname{Var}(y_t), \tag{2.2}$$

where q_u, q_π and q_y are positive weights, normalized to sum to unity.[2] This specification of the loss function implies that the monetary authority is concerned with variations in inflation and output, but also considers a smoothing criterion for the short rate. Rapid changes of the short rate in either direction are penalized through the weight q_u.

The monetary authority chooses the control u_t to solve

$$u_t^* = \arg\min_u L.$$

The problem is set up as a standard exercise of optimal control, a version of the linear-quadratic regulator. It is well-known (see e.g. Björk (1998)) that the optimal control law u_t is a linear function of the state variables. That is, there are constants a_{11}, a_{12}, and a_{13} such that

$$u_t^* = a_{11} r_t + a_{12} \pi_t + a_{13} y_t.$$

With this control the unconditional mean of the short rate is zero.[3] For convenience we do not distinguish between u_t and u_t^*.

All state variables and all parameters are publicly observable.

[1] The short rate is modeled as a continuous process. This contrasts to the typical behavior of central banks, who adjust the short rate discontinuously. A continuous process is assumed for tractability.

[2] The variance operator for the respective state variables need not be a well-defined quantity. For a mathematically correct treatment of the subsequent control problem one can replace the respective variance operators by $E_t\left[\int_t^\infty e^{-\rho(s-t)} x_{is}^2 ds\right]$, where ρ is some positve constant. The optimal control used in the paper corresponds to the limiting optimal control as $\rho \to 0$.

[3] Analytical formulas for a_{11}, a_{12} and a_{13} as functions of the underlying parameters exist, but are complicated. In the estimation numerical procedures provided in Matlab were used.

2.2.2 Market prices

The model for the monetary transmission mechanism gives the state of the economy in terms of a three-dimensional state vector $x_t = (r_t \; \pi_t \; y_t)'$ with dynamics

$$dx_t = Ax_t dt + V dw_t, \tag{2.3}$$

where A and V are constant 3×3 matrices. Although this constitutes a system for the general dynamics of the economy, as a model for pricing financial securities it is incomplete, since this requires an assessment of investors' attitude towards risk. A naive approach is to disregard risk concerns, and consider the yield on a long-term bond as the expected average of the short rate over the bond's maturity. This is the expectation hypothesis in its most rudimentary form, and it is usually rejected. Here I follow the tradition in term structure modelling as set out in Black and Scholes (1973) and Merton (1973).

Let $P_t = P(t, x_t)$ denote the price of a financial asset in the economy. The Itô differential of P_t is given as

$$dP_t = P_t \mu_t dt + P_t \sigma_t dw_t,$$

where the drift μ_t and the volatility $\sigma_t = (\sigma_{1t} \; \sigma_{2t} \; \sigma_{3t})'$ may depend on P_t. If P_t is fully immunized against variations in x_t then $\sigma_t = \mathbf{0}$. In this case the asset is locally risk-free and economically equivalent to a short risk-free bond. Since such a bond has an instantaneous return equal to the short risk-free rate r_t, it is a natural requirement that the drift μ_t equal r_t.

If P_t is partially immunized so that $\sigma_{it} = \sigma_{jt} = 0$ but $\sigma_{kt} \neq 0$, for distinct i, j and k, then the holder of the asset is exposed to only one risk factor. In this case the drift μ_t need not equal to the short rate, but can in a similar way be interpreted as a compensation for holding risk k. Taking account for the degree of exposure σ_{kt} it can be shown that the appropriate relation between μ_t, σ_{kt} and r_t is that at each point in time there is a number λ_{kt} such that

$$\frac{\mu_t - r_t}{\sigma_{kt}} = \lambda_{kt}. \tag{2.4}$$

This conjectures that for a given risk factor, assets exposed to only this factor share a common ratio for the excess return per unit volatility. Given one asset exposed to risk k, all other such assets can be seen as economically equivalent to a combined position in the first asset and a risk-free bond. The ratio λ_{kt} is called the market price of risk with respect to factor k, and it has a natural interpretation as the factor's Sharpe ratio. The relation (2.4) must hold at any instant in time, although λ_{kt} very well may change over time.

Since there are three sources of uncertainty, market prices of risk are completely characterized by the three-dimensional vector $\lambda_t = (\lambda_{rt} \; \lambda_{\pi t} \; \lambda_{yt})'$, which accordingly reflects investors' attitude towards risk. To price financial securities, we need to know λ_t. It is not sufficient that we learn λ_t at a given point in time — we need to know the entire dynamic structure of λ_t. However, if the dynamics of λ_t is known, this is sufficient to price financial securities.

To illustrate, suppose that $\lambda_t = \mathbf{0}$ for all t. In this case investors are risk-neutral with respect to each factor, and the expected excess return on any financial asset is zero. This corresponds to the expectations hypothesis. The natural extension of this is the case when λ_t is constant over time. In the constant volatility model studied here, this implies that the risk premium on a constant maturity bond is constant over time. In other words, the term premium for a fixed time to maturity is constant over time.

Although the assumption of a constant risk premium provides a natural benchmark, several studies suggest that risk premia vary over time. For the bond market early papers include Fama and Bliss (1987) and Campbell and Shiller (1991). In these studies excess returns on bonds of various maturities are regressed on contemporaneous forward rates or yields. In a more recent study Cochrane and Piazzesi (2001) extend Fama and Bliss' study to include several forward rates. A main finding in that paper is that the excess return on bonds of various maturities is well described by a one factor model, a factor Cochrane and Piazzesi refer to as the *return forecasting factor.*

Here I adopt a factor model for the market price of risk similar to that in Cochrane and Piazzesi, except that a constant term is added. More precisely, I assume that the market price of risk can be written as

$$\begin{aligned} \lambda_t &= k_0 + k_1 \gamma_t, \\ \gamma_t &= h' x_t, \end{aligned} \tag{2.5}$$

where k_0, k_1 and h are constant 3×1 vectors, and where the first element of h for normalization is set equal to one. The factor γ_t captures the dynamic part of the risk premium, and is comparable to the return forecasting factor in Cochrane and Piazzesi.

Taken together with the dynamics of the state variables (2.3) the market price of risk characterizes the pricing of financial assets. Furthermore, because the dynamics of x_t is of linear constant volatility type, and because the functional form for λ_t is affine in x_t, we obtain a term structure model that is a multi-factor version of Vasicek (1977). This is convenient, since formulas for bond prices take a particular simple form.

Let $F^{\tau}(t)$ denote the price on a nominal risk-free discount bond with remaining maturity τ, and let $Y^{\tau}(t)$ denote its yield so that

$$F^{\tau}(t) = \exp\left(-Y^{\tau}(t) \cdot \tau\right).$$

With assumptions as above there are functions $a : [0, \infty) \to \mathbb{R}$ and $b : [0, \infty) \to \mathbb{R}^3$ such that

$$Y^{\tau}(t) = a(\tau) + b(\tau)' x_t. \tag{2.6}$$

The functions a and b are given in Appendix 2.7.

2.2.3 An example

The following example is provided to illustrate how the model works. Parameter values are chosen to match the estimation results in Section 2.5 below. Inflation and the output gap are governed by

$$d\pi_t = 0.22 y_t dt + 0.55 dw_{\pi t},$$
$$dy_t = -1.46\,(r_t - \pi_t)\,dt + 0.5 y_t dt + 1.44 dw_{yt},$$

and the short rate follows the process

$$dr_t = u_t dt + 1.14 dw_{1t}.$$

The instantaneous correlation between the Wiener processes is $(\rho_{12},\ \rho_{13},\ \rho_{23}) = (0.64,\ 0.25,\ -0.34)$. The drift u_t is optimal with respect to a certain set of weights in the monetary authority's objective function (2.2). We consider two set of weights. The first is given by $(q_r,\ q_\pi,\ q_y) = (0.24,\ 0.76,\ 0)$, in accordance with subsequent estimation results, and the second is given by $(q_r,\ q_\pi,\ q_y) = (0.1,\ 0.9,\ 0)$. In each case the weight on output targeting is zero, and the difference lies solely in the relative weight on interest rate smoothing. For the latter example the weight on interest rate smoothing is lower, which allows for a more flexible rule for the interest rate.

We study implications for market prices in a scenario where the short rate and the output gap are at their long-run means, but where inflation is one percentage point below the target. For expositional reasons the market price of risk is set equal to zero. Figure 2.1 depicts yield curves for each example. Both yield curves display a negative hump. This reflects expectations that the short interest rate will first be lowered as inflation is stabilized, and then successively adjusted back toward its long-run mean. For the second set of weights the hump on the yield curve is more pronounced. The lower weight on interest rate smoothing is reflected in a more aggressive policy against inflation.

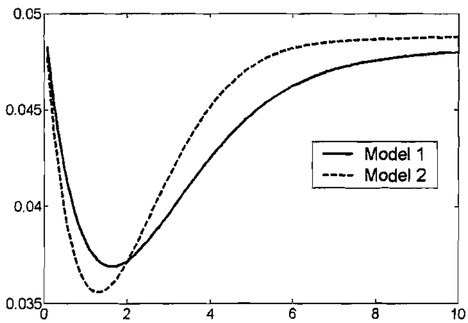

Figure 2.1. Yield curves. Model 1 corresponds to $(q_r,\ q_\pi,\ q_y) = (0.24,\ 0.76,\ 0)$. Model 2 corresponds to $(q_r,\ q_\pi,\ q_y) = (0.1,\ 0.9,\ 0)$. Curves are shifted upward 0.05. The y-axis is percentage points, and the x-axis is time to maturity in years.

In this context it is also illustrative to look at yield volatilities. Using (2.6) for yields together with (2.3) for the state dynamics, we obtain the instantaneous volatility of a yield for a given maturity τ as

$$v(\tau) = \sqrt{b(\tau)' VRR'V'b(\tau)}$$

where RR' is the instantaneous correlation matrix of the state variables. Consequently, yield volatilities are independent of the current values of the state variables. In other words the yield volatility curve is time homogenous. Figure 2.2 depicts yield volatilities for the given sets of weights. Each curve exhibits a hump in a similar way as the yield curves. The same reasoning also applies. Unexpected changes in the state variables cause the monetary authority to adjust the short rate. However, because of the smoothing condition adjustments are delayed, and as a consequence volatility curves are upward sloping for low maturities, albeit a small curl is observed at the very short end due to the instantaneous volatility of the short rate. Over time the short interest rate is adjusted back, and the volatility curves are therefore downward sloping at higher maturities. Similarly to the yield curves, a lower weight on interest rate smoothing is associated with a hump that is more pronounced and located at lower maturities.

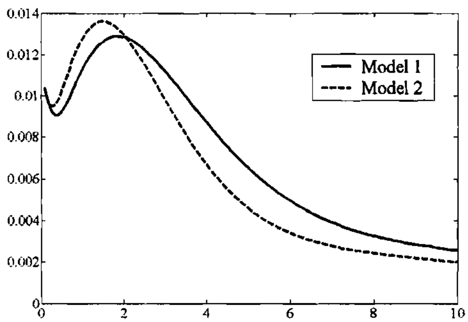

Figure 2.2. Yield volatility curves. Model 1 corresponds to $(q_r,\ q_\pi,\ q_y) = (0.24,\ 0.76,\ 0)$. Model 2 corresponds to $(q_r,\ q_\pi,\ q_y) = (0.1,\ 0.9,\ 0)$. The y-axis is percentage points, and the x-axis is time to maturity in years.

2.3 Estimation

The model for the monetary transmission mechanism as defined in Section 2.2.1 taken together with the specification for the market price of risk as defined in Section 2.2.2, specifies a model for the term structure of interest rates. In this aspect the current paper is very typical. Three-factor affine interest rate models have been studied in e.g. Babbs and Nowman (1999) and Dai and Singleton (2001).

In additional resemblance with previous work, the current paper argues that the determinants of the term structure can be associated with macro economic aggregates. Along this branch we find Ang and Piazzesi (1999) and Wu (2001) who study affine interest rate models and relate factors to measures of inflation and output.

Our contribution to the literature lies in the specification of the factor dynamics. The model for the monetary transmission mechanism adds structure to the dynamics, which allows us to relate model parameters to more familiar economic relationships, for instance the Philips curve and the weights in the monetary authority's objective function. As such the model may be helpful for understanding differences in term structure dynamics under different monetary regimes.

An important part of this agenda is to estimate the model using historical data. This can validate the use of the model, and can in itself be a way of identifying differences between monetary regimes. Estimating the parameters of the model amounts to estimating the underlying state equation (2.1), the weights for the loss function (2.2), as well the parameters for the market price of risk (2.5). In the literature on monetary policy various versions of the model (2.1) – (2.2) have been estimated, see e.g. Söderström, Söderlind and Vredin (2002). The current paper deviates from these estimations in that term structure data is used to infer the state processes. Using financial prices is attractive, since these tend to be forward looking and possibly more representative of the monetary authority's information set.

Using financial prices is however also problematic since state variables are not directly observable. More precisely there is an identification problem. The economic model, taken together with a specification of the market price of risk, gives rise to certain term structure dynamics. However, on observing term structure dynamics, the same dynamics is generated by an infinite number of other factor models, only one of which is consistent with the original economic model.

In the current paper the identification problem is treated by adding an initial step to the estimation procedure. In this first step factor series are extracted from the yield curve in a manner that is consistent with the general affine three-factor structure. The extracted factor series are selected to match some publicly observed measures of the original state variables optimally in a usual least squares sense.

In a second step the extracted factor series are interpreted as true observations of the state variables, and accordingly referred to as the short rate, inflation and the output gap. In this step the model's parameters are estimated with usual techniques.

In the third and final step the economic model is reassociated with the yield curve by estimating the parametric form for the market price of risk (2.5).

The remainder of the section covers each step in some more detail, with selected parts deferred to the appendix.

2.3.1 Extracting factor series

Suppose we observe yields on n zero coupon bonds with maturities $\tau_1, \ldots, \tau_n$ and let the observations be denoted by $Y_t^o = (Y_t^{\tau_1}, \ldots, Y_t^{\tau_n})'$. The first maturity τ_1 is taken to be a short date so that the first yield can be interpreted as the short rate. According to the model there is an $n \times 1$ vector a and an $n \times 3$ matrix B such that

$$Y_t^o = a + Bx_t,$$

where $x_t = (r_t \; \pi_t \; y_t)'$. The respective entries of a and B are in accordance with equation (2.6). Letting $Y_t = Y_t^o - E[Y_t^o]$, i.e. yields detrended to mean zero, we obtain the linear factor model

$$Y_t = Bx_t. \tag{2.7}$$

If x_t is observable, we are in a good position of estimating B. We can for example estimate the dynamics of x_t with usual techniques, and then estimate B by choosing appropriate parameters for the market price of risk. If on the other hand x_t is not observable, there is a problem in that the representation of yields in terms of an unobservable state vector is not unique. Any linear transformation of the state variables $\widetilde{x}_t = Cx_t$, where C is an invertible 3×3 matrix, yields a new representation. In this manner we obtain an entire family of term structure realizations, each of which allows a representation (2.7). Within this family we are interested in one particular realization, namely that one for which the corresponding factor series can be interpreted as the original state variables the short rate, inflation and the output gap.

In order to extract this particular realization, two matters need to be settled. First, given set of yield curve realizations, we require a criterion for selecting a particular realization within this set. Second, we require a procedure for characterizing the relevant set of yield curve realizations.

Starting with the first problem, how to select a particular realization, different approaches are conceivable. One plausible selection criterion is, to select that realization x_t for which the constraints that the economic model imposes on the dynamics of x_t make most sense. This has however turned out to be problematic. A possible explanation to the difficulties is, that although the economic model puts constraints on the parameters, the same set of restrictions may apply for several yield curve realizations. If the constraints are satisfied, this does not in itself guarantee that implied series can be interpreted as those basic economic quantities we seek. Consequently, finding a model where the economic constraints make most sense, does not guarantee that corresponding parameter estimates make sense economically.

An alternative selection criterion is to make use of publicly observed values of the state variables. The short rate is naturally observed since it is equal to the first element of Y_t. For inflation and the output gap we can take some regularly published figures (see Section 2.4 below). We can then select that realization for which the sum of squared residuals between observed series and extracted factors is minimized. This is the selection criterion used in the sequel.

Proceeding with the second problem, how to characterize the set of factor realizations, the method is described in some detail in Appendix 2.8. To give the general idea, suppose we are given one particular factor realization, and suppose this can be written as

$$Y_t = B^0 \widetilde{x}_t$$

with the instantaneous covariance matrix of $\widetilde{x}_t$ equal to the identity matrix.[4] By defining new factors x_t according to

$$x_t = D\widetilde{x}_t$$

for some invertible 3×3 matrix D we obtain a new factor model,

$$Y_t = B^0 D^{-1} x_t.$$

By inverting this equation, we obtain the corresponding factors x_t. The family of factor realizations is therefore parameterized via the matrix D. Because the matrix D has nine elements, the dimension of the factor realizations set is equal to nine. The appendix shows how to select the matrix D optimally to minimize the sum of squared residuals between the observations and the extracted factors.

Note that the approach is independent of the market price of risk, as long as this is affine in the state variables (as it is in e.g. (2.5)). A different specification of the market price of risk is likely to induce different factor loadings for the yields, i.e. a different matrix B in equation (2.7), and accordingly different yields. But given yields originally generated by the factors x_t, there will always be matrices B^0 and D that recover these factors.

2.3.2 The monetary transmission mechanism

To estimate the basic model, we take as given the extracted factor series x_t of the previous section, and interpret it as representing the short rate, inflation and the output gap respectively. We thus have the model

$$dx_t = Ax_t dt + V dw_t.$$

[4] This initial representation will be taken to be the principal components representation (see Appendix 2.8 for details).

Note that the instantaneous covariance matrix of x_t is known. In terms of the previous section we have that

$$dx_t \, dx_t' = DD'dt,$$

where we have used that the instantaneous covariance matrix for the initial factors $\widetilde{x}_t$ is equal to the identity matrix. In the estimation there is consequently no need to estimate V or the correlation structure of w_t. The parameters to be estimated are

$$\omega = \left(a_{23} \; a_{31} \; a_{33} \; q_\pi \; q_y\right)'.$$

The model is estimated with maximum likelihood, and the likelihood function is given in Appendix 2.9.

2.3.3 The market price of risk

In the third and final step we take as given the dynamics estimated in the previous section. From Section 2.3.1 we have the affine model for yields

$$Y_t^o = a + Bx_t.$$

The vector a and matrix B are taken to be the empirical counterparts. To estimate the market price of risk, recall the model (2.5),

$$\begin{aligned}\lambda_t &= k_0 + k_1 \gamma_t, \\ \gamma_t &= h' x_t.\end{aligned}$$

Since the dynamics for the system is given, for a given choice of parameters k_0, k_1 and h we obtain pricing formulas $a(\tau)$ and $b(\tau)$. Denoting the corresponding estimated vectors for the factor model by $\widehat{a}$ and $\widehat{B}$ we select that set of parameters which minimizes the sum of squared residuals

$$\|a - \widehat{a}\|^2 + \left\|B - \widehat{B}\right\|^2.$$

2.4 Data

The data set consists of observed state variables and observed yields. For the basic model, the short rate is set to the one-month interbank rate. Inflation is measured as quarterly changes in the GDP implicit price level. The output gap is measured as real GDP minus potential GDP as estimated by the Congressional Budget Office (2002). The measures of inflation and GDP are in accordance with Rudebusch (2001).

Term structure data is obtained from swap rates of maturities 2, 3, 4, 5, 7, and 10 years. The short rate is set to the one-month interbank rate. In the estimation it is more convenient to work with zero coupon yields, and the swap rates were therefore stripped. This was done by for each date fitting Nelson - Siegel parameters to the swap curve. The parameters were constrained to set the short rate equal to the one-month interbank rate. For each swap rate the fitted curve was used to strip coupon payments, and the residual was stored as the zero coupon yield for the corresponding maturity.

The sample period is November 1987 to April 2002, approximately covering the Greenspan period.

2.5 Results

The extracted factor series are depicted in Figure 2.3. Extracted and observed series display similar patterns, but at any given point in time values can be quiet different. Series need not match perfectly. For example, there may be times when observed inflation is influenced by temporary effects, thus making it less representative of the state variable relevant for the monetary authority's objective function. Also the monetary authority's assessment of potential GDP need not perfectly align with the current estimates.

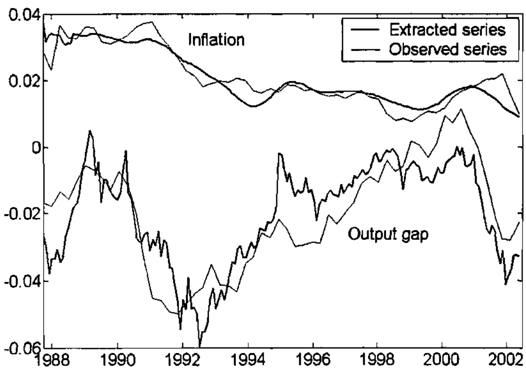

Figure 2.3. Estimated and observed state processes. Inflation is detrended and then shifted upward by 0.03. The output gap is detrended and then shifted downward by 0.03. Inflation is plotted as a one-year moving average.

The volatilities and correlations of the state variables are given by

$$\frac{\sigma_r \quad \sigma_\pi \quad \sigma_y}{1.14\% \quad 0.55\% \quad 1.44\%},$$

$$\frac{\rho_{12} \quad \rho_{13} \quad \rho_{23}}{0.64 \quad 0.25 \quad -0.34}.$$

The yields' loadings on the factors are depicted in Figure 2.4. The lines correspond to the respective column vectors of B in equation (2.7). The figure suggests that long term yields are quiet sensitive to changes in inflation, and negatively related to unexpected changes in the short rate. Note however that in the estimation there is no control to keep factor loadings moderate, and results should therefore be interpreted with care. For example, the high sensitivity to inflation and the negative dependence on the short rate may relate to the seemingly high correlation ρ_{12} between these state variables.

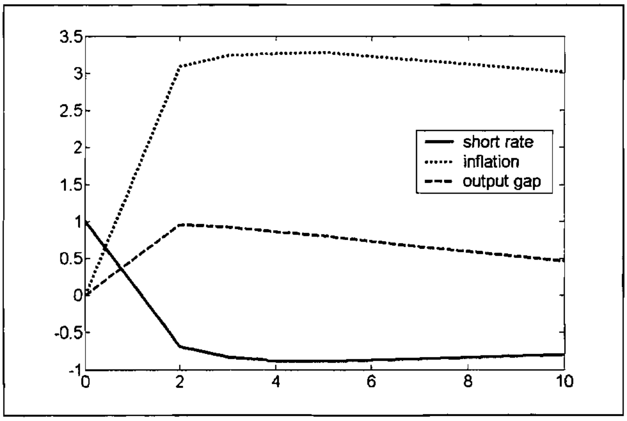

Figure 2.4. Factor loadings. Each line corresponds to a row vector of the matrix B in (2.7). Time axis is time to maturity in years.

For the estimation of the monetary transmission mechanism, the model is written as

$$dx_t = Ax_t dt + V dw_t.$$

The estimate of the drift matrix A is[5]

$$A = \begin{pmatrix} -2.49 & 4.28 & 2.12 \\ 0 & 0 & \underset{(0.09)}{0.22} \\ -\underset{(0.37)}{1.46} & \underset{(0.37)}{1.46} & \underset{(0.29)}{0.50} \end{pmatrix}.$$

The estimate of a_{23} of 0.22 is low in comparison with previous studies — Rudebusch (2001) estimates a value of 0.13×4 (quarterly observations) on US data 1968:Q3 to 1996:Q4. The estimate of a_{31} of 1.46 is higher than Rudebusch' 4×0.09.

The corresponding estimates of the weights in the monetary authority's objective function are

[5] The eigenvalues of A are -1.04 and $-0.48 \pm 0.57i$.

q_u	q_π	q_y
0.24	0.76	0
(0.07)	(0.07)	(–)

The weight on output stabilization vanished and was therefore constrained to zero in the estimation. The estimates are roughly in line with previous studies, although these vary considerably. Söderström, Söderlind and Vredin (2002) report that generally the interest rate smoothing parameter is estimated so that q_u/q_π is in the range 0.5 to 2, while the weight on output stabilization usually is such that q_y/q_π is less than 0.1. The corresponding numbers for the current estimation are 0.31 and 0. Consequently the current estimation suggests a lower weight on interest rate smoothing.

Impulse-response functions are depicted in Figure 2.5. The graphs are similar to those obtained by Söderström, Söderlind and Vredin (2002) when calibrating the model to US data 1987-1999.

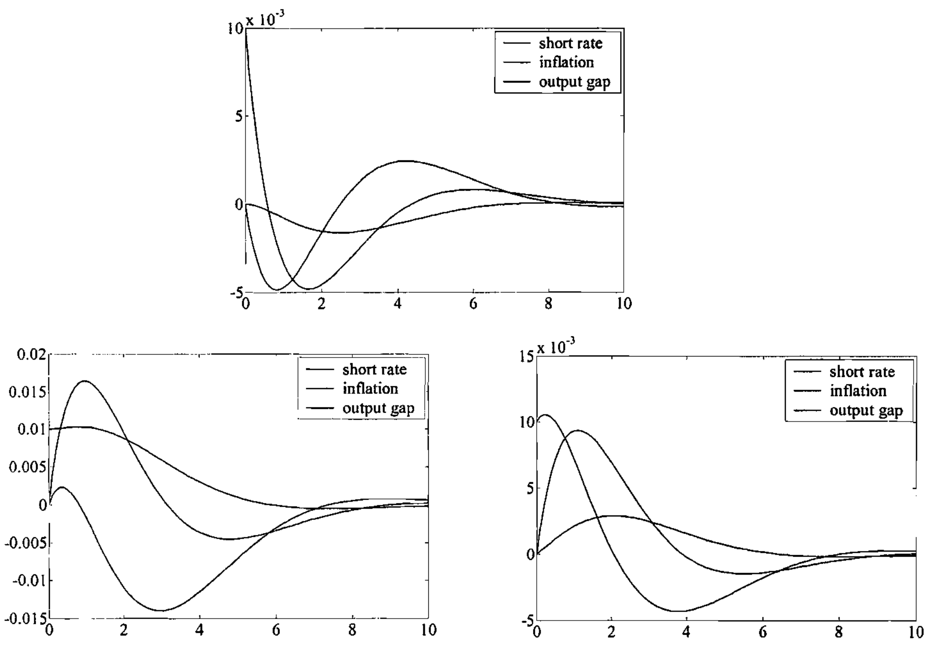

Figure 2.5. Impulse-response functions. The upper graph shows expected paths of the state variables starting with the short rate at 0.01. The lower left graph depicts expected paths starting with inflation at 0.01 and the graph at the lower right expected paths starting with the output gap at 0.01. Time axis in years.

It remains to estimate the market price of risk. The original data has been through two manipulations, first as state variables were extracted, and second as the economic model was imposed on the dynamics. Results should therefore be interpreted cautiously. We follow the procedure set out in Section 2.3.3 and estimate the market price of risk as

$$\lambda_t = \begin{pmatrix} 0.544 \\ 0.1523 \\ 0.6907 \end{pmatrix} + \begin{pmatrix} -0.0365 \\ -0.1204 \\ 4.0739 \end{pmatrix} \gamma_t,$$
$$\gamma_t = \begin{pmatrix} -0.4080 & 1 & 0.2475 \end{pmatrix} x_t.$$

The constant term of λ_t have all entries positive, which suggests that if the state variables are all at their long-run means, all risk factors have a positive price. As for γ_t, the time varying part of λ_t, a time series is given in Figure 2.6. Evidently γ_t is mainly related to inflation. A possible explanation is that in the constant volatility model studied here, an increase in inflation does not a priori affect the level of financial risk. However, if volatility in inflation is positively related to the level of inflation, an increase in inflation may increase the perceived risk in a financial security. Further, the time variation in λ_t is moderate – the unconditional standard deviation of γ_t is merely 0.28%. Both findings are in contrast with Cochrane and Piazzesi (2001) who, also in a constant volatility model, find the return forecasting factor, the time varying part of the market price of risk, to mainly relate to measures of the business cycle and also economically significant.

The time variation in λ_t goes primarily through the market price of risk with respect to the output gap. Thus, if inflation is high, investors require a high compensation for carrying risk related to output.

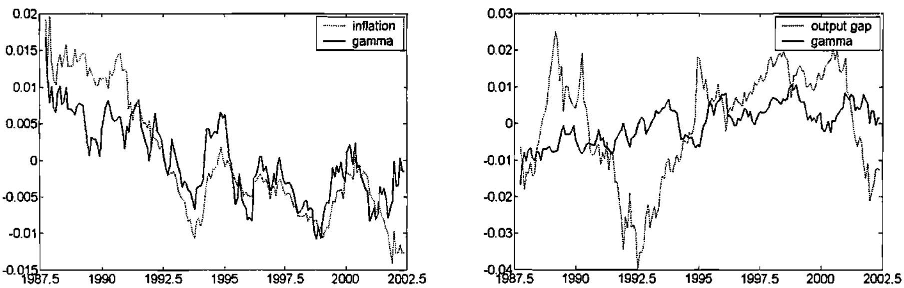

Figure 2.6. Time variation in γ_t. The left hand graph shows the extracted series for inflation and the estimated γ_t, where γ_t has been multiplied by a factor 2. The right hand graph shows the extracted series for the output gap and γ_t, with γ_t multiplied by a factor -2.

The implied vector a and matrix B are depicted in Figure 2.7 together with their sample counterparts. For a, the constant part of λ_t, observed and model implied values are hardly distinguishable. Also, the one factor model for the time varying part of λ_t seems to work well in that observed and model implied vectors are close.

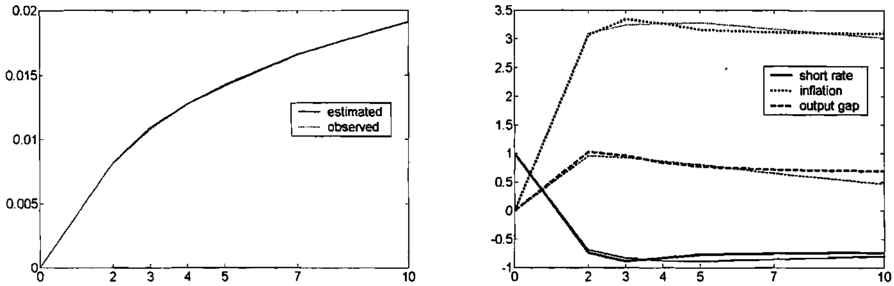

Figure 2.7. Observed and model implied vales of a and B. Time axis is time to maturity in years.

2.6 Summary

This paper has investigated how term structure dynamics may be related to monetary policy. The finding is that a standard model on monetary policy can be associated with observed term structure dynamics in a way that aligns with previous results. This gives support to the idea that financial prices can be used to infer the market's perception of the monetary regime.

2.7 Appendix. Bond prices

Let

$$\begin{aligned} A^Q &= A + V k_1 h' \\ m &= V k_0 \end{aligned}$$

The spot rate for maturity τ is given by

$$Y^\tau (t) = a(\tau) + b(\tau)' x_t$$

where

$$\begin{aligned} b(\tau) &= e_1' \mathbb{A}(\tau) \\ a(\tau) &= \overline{r} - e_1' (I + \mathbb{A}(\tau)) m \\ &\quad - \frac{1}{2} \operatorname{trace} \left(V \left\{ M - M\mathbb{A}(\tau) - \mathbb{A}(\tau)' M + \frac{1}{\tau} \left(e^{A'\tau} K e^{A\tau} - K \right) \right\} V' \right) \end{aligned}$$

and $e_1 = (1\ 0\ 0)'$. The function $\mathbb{A}(\tau)$ is given by

$$\mathbb{A}(\tau) = (I - \exp(A\tau))(A\tau)^{-1}$$

and the matrices M and K are defined according to

$$\begin{aligned} M &= (A')^{-1} e_1 e_1' (A)^{-1} \\ A'K + KA &= M \end{aligned}$$

2.8 Appendix. Extracting factor series

From the observed yields we obtain an empirical covariance matrix of yield innovations. Given monthly observations let

$$\overline{\Sigma}_Y = 12 \times \operatorname{cov} \Delta Y_t .$$

A Cholesky decomposition of $\overline{\Sigma}_Y$ gives an orthonormal matrix C and a non-negative diagonal matrix Λ such that

$$\overline{\Sigma}_Y = C \Lambda C' .$$

The diagonal entries of Λ equal the eigenvalues of $\overline{\Sigma}_Y$. If the original model is correct, it is true for the theoretical counterpart of $\overline{\Sigma}_Y$ that all but three of the eigenvalues equal zero. For the empirical covariance matrix $\overline{\Sigma}_Y$ we generally have all the eigenvalues positive, but for an estimation it is natural to focus on the three greatest.

This allows us to formulate a first candidate factor model of the yield curve. Suppose Λ is chosen such that its diagonal entries satisfy $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_n \geq 0$. Let

$$B^0 = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & c_{n3} \end{bmatrix} \begin{bmatrix} \sqrt{\ell_1} & 0 & 0 \\ 0 & \sqrt{\ell_2} & 0 \\ 0 & 0 & \sqrt{\ell_3} \end{bmatrix},$$

where each c_{ij} is the corresponding element of C. The first factor model is defined as

$$Y_t = B^0 \widetilde{x}_t$$

with $\widetilde{x}_t = (\widetilde{x}_{1t}\ \widetilde{x}_{2t}\ \widetilde{x}_{3t})'$ and

$$d\widetilde{x}_t = M_t^0 dt + dw_t^0,$$

where w_t^0 is a three-dimensional standard Wiener process. The drift term M_t^0 has not been specified since it is not in our immediate interest. Note that the factors $\widetilde{x}_t$ have been normalized to unit variance. Since the principal components are independent, this implies that the instantaneous covariance matrix of $\widetilde{x}_t$ equals the identity matrix.

Given this particular realization of the term structure, the complete set of linear three-factor constant volatility realizations can be characterized. Each such realization x_t corresponds to a matrix D for which

$$x_t = D\widetilde{x}_t.$$

Because D is a 3×3 matrix, the set of factor realization has dimension nine.

Mathematically it is convenient to think of the matrix D in terms of its row vectors. We can then parameterize D by factoring it into three distinct matrices so that $D = D_1 \times D_2 \times D_3$. The first matrix D_1 is diagonal and normalizes each row vector to unit length. We can thus think of $D_2 \times D_3$ as comprised of three unit vectors in $\mathbb{R}^3$. The second matrix D_2 defines the relative placement of these vectors, or equivalently the angels between any two vectors. Because there are three vectors, there are three two-vector combinations, and thus three angels that should be defined. Consequently the parametrization of D_2 has dimension three. We can choose D_2 lower triangular. The last matrix D_3 is any orthonormal matrix. Again we can parameterize D_3 using three angels. The first two angels define the locus of the first vector. The first angel is a rotation in the xy-plane, and the second angel is an elevation along the z-axis. Given the locus of the first vector, the remaining two vectors must be located on the two-dimensional plane orthogonal to the first vector. The third angel defines the locus of the second vector in terms of a rotation in this plane. This also defines the locus of the third vector as orthogonal to the other two.

The first two matrices relate to the covariance between the state variables. The matrix D_1 defines the volatilities, and the matrix D_2 the correlation structure. The matrix D_3 does not give any information about the covariance structure. Consequently, for a given covariance matrix, the set of covariance consistent factor realizations has dimension three.

Each choice of D produces a factor model for yields according to

$$Y_t = B^D x_t,$$

where $B^D = B^0 D^{-1}$.

Since we assume that the short rate is observable and equal to the first element of Y_t, we can restrict the set of admissible matrices D. Variations in inflation and the output gap do not affect the short rate directly, but do so implicitly via the drift term. Therefore it must hold for the first row of B^D that

$$\left(b_{11}^D \; b_{12}^D \; b_{13}^D\right) = \left(\sigma_r \; 0 \; 0\right),$$

where b_{ij}^D is the corresponding element of B^D and where σ_r is the volatility of the short rate. This imposes two additional restrictions on D, and in effect the set of admissible matrices has dimension seven. The restrictions are set to be restrictions on the matrix D_3.

For a given admissible matrix D we can back out factor series according to

$$x_t = \left(\left(B^D\right)' B^D\right)^{-1} \left(B^D\right)' Y_t.$$

We can then choose that matrix D for which the extracted factor series look reasonable.

By construction the first element of x_t is equal to the short rate. We seek that choice of D for which the remaining two factors resemble inflation and the output gap respectively. Suppose therefore that we have observations of inflation $\widehat{x}_{2t}$ and the output gap $\widehat{x}_{3t}$. For each D define the loss function[6]

$$f(D) = \sum_{t=1}^{T} \left\{ \left(x_{2t} - \widehat{x}_{2t}\right)^2 + \left(x_{3t} - \widehat{x}_{3t}\right)^2 \right\}.$$

The function f is minimized using numerical procedures in Matlab.

Given D we have the instantaneous covariance of the state variables according to

$$\begin{aligned} dx_t\, dx_t' &= DD'dt \\ &= D_1 D_2 D_2' D_1' dt, \end{aligned}$$

[6] Introducing a weightning of the respective residuals does not affect the resulting implied factor series.

where the last equality follows from the fact that D_3 is orthonormal. Conversely however, for a given covariance matrix of the state variables, we do not have a unique matrix D. Because there are three volatilities and three correlations, and because the set of admissible D has dimension seven, there is one free element of D. The free element corresponds to the fact that on only observing yield dynamics, we can not directly distinguish between covariance independent linear transformations of inflation and the output gap.

2.9 Appendix. The likelihood function

Given

$$dx_t = Ax_t dt + V dw_t$$

we have

$$x_t \sim N\left(e^{At} x_0, M_t\right)$$

where

$$M_t = K - e^{At} K e^{A't}$$

and K solves

$$AK + KA' = VV'$$

If we are given observations $(\widehat{x}_t)_{t=1,\dots,T}$ the log-likelihood function is given by

$$L = -\frac{1}{2}(T-1)\ln |M_{\Delta t}| - \frac{1}{2}\sum_{t=2}^{T} \Delta\widehat{x}_t M_{\Delta t}^{-1} \Delta\widehat{x}_t'$$

where

$$\Delta\widehat{x}_t = \left(\widehat{x}_t - e^{A\Delta t}\widehat{x}_{t-1}\right)$$

References

1. Ang, Andrew and Monika Piazzesi, "A no-arbitrage vector auto-regression of term structure dynamics with macroeconomic and latent variables," *Working Paper*, Columbia University.
2. Babbs, Simon H., and K. Ben Nowman, "Kalman filtering of generalized Vasicek term structure models", Journal of Financial and Quantitative Analysis, 34:1, 115-130.
3. Björk, Tomas, 1998, "*Arbitrage theory in continuous time*," Oxford University Press.
4. Brennan, Michael J., Ashley W. Wang and Yihong Xia, "Estimation and test of a simple model of intertemporal capital asset pricing," *Working Paper*, Anderson Graduate School of Management, UCLA.
5. Campbell, John Y. and Robert J. Shiller, 1991, "Yield spreads and interest rate movements: a bird's eye view," *Review of Economic Studies* 58, 495-514.
6. Clarida, Richard, Jordi Galí and Mark Gertler, 1999, "The science of monetary policy: A new Keynesian perspective," *Journal of Economic Literature 37*, 1661-1707.
7. Cochrane, John H. and Monika Piazzesi, 2001, "Bond risk premia," *Working Paper.*
8. Dai, Qiang and Kenneth Singleton, 2001, "Expectation Puzzles, time-varying risk premia, and affine models of the term structure," *Journal of Financial Economics* 63, 415-441.
9. Duffee, Gregory R., 2002, "Term premia and interest rate forecasts in affine models," *Journal of Finance* 57(1), 405-443.
10. Ellingsen, Tore and Ulf Söderström, 2001, "Monetary policy and market interest rates," *American Economic Review* 91(5), 1594-1607.
11. Fama, Eugene F. and Robert R. Bliss, 1987, "The information in long-maturity forward rates," *American Economic Review* 77, 680-692
12. Fama, Eugene F. and Kenneth R. French, 1989, "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics* 25, 23-49.
13. Nelson, C., and A. Siegel, 1987, "Parsimonious modeling of yield curves," *Journal of Business* 60, 473-489.
14. Rudebusch, Glenn D., 1995, "Federal Reserve interest rate targeting, rational expectations, and the term structure," *Journal of Monetary Economics* 35, 245-274.
15. Rudebusch, Glenn D., 2001, "Is the Fed too timid? Monetary policy in an uncertain world," *Review of Economics and Statistics* 83(2), 203-217.
16. Svensson, Lars E. O., 2000, "Open economy inflation targeting," *Journal of International Economics* 50 (February), 155-183
17. Söderström, Ulf, Paul Söderlind and Anders Vredin, 2002, " Can a calibrated New-Keynesian model of monetary policy fit the facts?," *Working Paper*, Sveriges Riksbank.

18. Söderlind, Paul, 2001, "Monetary policy and the Fisher effect," *Journal of Policy Modeling* 23, 491-495.
19. Vasicek, O., 1977, "An equilibrium characterization of the term structure", *Journal of Financial Economics* 5, 177-188.
20. Wu, Tao, 2001, "Macro factors and affine term structure models", *Working Paper*, Federal reserve Bank of San Francisco.

3. Optimal monetary policy, the zero bound and the term structure of interest rates

Abstract

The paper studies optimal monetary policy and its implication for the term structure of interest rates when the nominal short rate is bounded at zero. We state the monetary authority's optimization problem in continuous time according to two specifications, interest rate stabilization and interest rate smoothing. For the former the optimization problem is solved analytically, while numerical procedures are adopted for the latter. The paper then turns to study implications for the term structure of interest rates under risk-neutrality. Term structure equations are solved numerically and implications for yield curves and yield volatility curves are discussed. Data for a low-interest rate country like Japan for 1996 - 2003 exhibits s-shaped yield curves and yield volatility curves. According to our results this shape is consistent with a smoothing objective for the short rate.

3.1 Introduction

The conduct of monetary policy is undoubtedly an important determinant for the term structure of interest rates. The present paper studies the relation with a particular focus on the zero bound for the short rate. In this process we extend existing research on two frontiers.

First we bring the two research areas monetary policy and term structure modelling closer together. This is an endeavor that has received attention in recent years by Ang and Piazzesi (2003), Dewachter and Lyrio (2002), Hördahl, Tristani and Vestin (2003), Rudebusch and Wu (2003) and others. Our approach differs from the papers mentioned in that we consider monetary policy as implied from an optimization problem for a monetary authority, similar to Ellingsen and Söderström (2001), (2003) and Skallsjö (2003). In a study of the zero bound this distinction becomes essential as it otherwise can be difficult to assess an appropriate functional form for the monetary policy rule.

[0] I am deeply indebted to my advisers Tomas Björk and Peter Englund for their guidance and support. I am also grateful to Ulf Jönsson for help with issues in optimal control, to Anders Szepezzy for help with the numerical solution to the control problem in Section 3.5, and to Paul Söderlind for insightful comments on my work. Financial support from Bankforkningsinstitutet and the Wallander Foundation is gratefully acknowledged.

Second our approach to the zero bound is a contribution to the literature on term structure modelling. The zero bound is well recognized and many popular models exclude negative interest rates almost surely, but this property does not in itself guarantee realistic term structure dynamics when the short rate is close to zero. Examples are Cox, Ingersoll and Ross (1985), the squared Gaussian (see e.g. Pelsser (2000)) and Black and Karasinski (1991) which all combine analytical tractability and non-negative interest rates with sound short rate dynamics under normal monetary conditions. An alternative is Black (1995) who suggests that the short rate be modelled as the non-negative part of an underlying stochastic process, which itself may be negative. Black provides a discussion of the rationale for his proposal, noting that the zero bound applies to nominal interest rates while its determinants may be negative. In the current paper we put more structure on this approach by deriving the short rate process from an optimal monetary policy. Underlying state dynamics relate to the short rate in a non-trivial way and this specifies term structure dynamics in a way that would not be possible in a more direct approach.

The model for the monetary transmission mechanism is a reduced form of the standard framework as set out in e.g. Clarida, Galí and Gertler (1999) and Svensson (2000). In these papers the state of the economy is determined by two state variables, inflation and the output gap. The state variables are governed by a Markov system, in which one of the determinants is the short rate as set by a monetary authority. The setup thus takes the form of a control problem with the short rate as a control variable. The monetary authority is then equipped with an objective function formed as expected quadratic deviations from target levels in inflation, the output gap and (possibly) the short rate. This results in a version of the linear-quadratic regulator, for which it is possible to solve explicitly for the optimal control law minimizing the value of the objective function. The optimal policy for the short rate is linear in the state variables, thus similar to a Taylor (1993) rule. From a term structure perspective the result is tractable since the model translates into a multi-factor version of Vasicek (1977).

The above framework has become standard for monetary policy theory due to its combined simplicity, realism and flexibility. However, the basic model allows for negative interest rates. It is then natural to analyze the same model under the constraint that the short rate remains non-negative. This has been done by Orphanides and Wieland (1999) and Kato and Nishiyama (2001), who both study a setting similar to Svensson (1999). It is shown that the presence of a lower bound leads to nonlinear policy rules. In particular, monetary policy becomes more expansive close to zero (though trivial at zero) relative to the corresponding unconstrained problem. The monetary authority follows a preemptive strategy. To compensate for the probability that monetary policy is constrained in the next period, it becomes more expansive in the current period.

The tradition in the literature on optimal monetary policy is to conduct the analysis in discrete time. In the present paper the model is set up in continuous time since term structure modelling is better suited for this framework. To set focus on short rate dynamics the two state variables inflation and output are merged into one, referred to as the target variable. This assumption keeps the analysis simple, and because we focus on implications for the term structure of interest rates the target variable can be regarded as a latent factor of the yield curve.

We study two specifications for the monetary authority's objective function, in the literature commonly known as interest rate stabilization and interest rate smoothing. In the case of stabilization the monetary authority's objective depends on the expected quadratic deviations in the short rate from a target level. In the case of smoothing it depends on the expected squared *changes* in the short rate. The latter assumption is more common in the literature as it seems to better capture the tendency among central banks to adjust the short rate in several consecutive moves in the same direction, and to change the direction only infrequently. The literature has also given some attention to the origins of this behavior and several possible explanations have been offered. For a review together with some empirical evidence see Sack and Wieland (1999).

In both specifications, without the zero bound the control problem becomes a version of the linear-quadratic regulator with an optimal control law that is linear in the state variables. With the zero bound the analysis becomes more involved. In the case of interest rate stabilization it is possible to solve for the optimal control law analytically, but in the case of interest rate smoothing we do not have an analytical expression. We then use the Pontryagin minimum principle to study the special case where the target variable is non-stochastic, and given this solution we apply a numerical procedure to approximate the solution to the stochastic problem.

The implied control laws exhibit similarities as well as differences. In both specifications increased volatility in the target variable is associated with a more expansive policy. This reestablishes results in Orphanides and Wieland, and Kato and Nishiyama – increased uncertainty induces the monetary authority to act more preemptively. An implication of this is that the steady state in the economy is reached at a value of the target variable that exceeds its equivalent in the unconstrained problem.

The vital difference between the two specifications is the behavior of the short rate in the vicinity of zero. With interest rate stabilization the short rate is as most responsive to changes in the target variable just before the zero bound is reached. With interest rate smoothing the opposite holds. In particular, with smoothing the short rate always approaches zero in a smooth manner, i.e. with a time derivative equal to zero.

The two specifications result in two alternative models for the short rate. Given these we turn our attention to the term structure of interest rates.

In general this would require that we take investors' risk preferences into account. Here, in order to make the analysis more transparent we assume risk-neutrality. Bond prices are solved for using numerical methods. We study model implied yield curves and yield volatility curves and compare shapes with observations from the term structure in Japan 1996 - 2003. These tend to exhibit s-shaped curves. A general observation is that the specification with a smoothing objective has a stronger tendency to generate s-shaped curves. This characteristic relates directly to the difference in short rate behavior in the vicinity of zero reported above. A lower responsiveness in the short rate acts as to compress the short end of the yield curve, resulting in lower volatility in short to medium term yields. For long term yields the difference between the two specifications is smaller.

The paper is organized as follows. Section 3.2 presents the general model. Sections 3.3 and 3.4 are devoted to interest rate stabilization, with theory in the former and numerical examples in the latter. Interest rate smoothing is treated similarly in Sections 3.5 and 3.6. For the analysis of the term structure of interest rates some empirical background is given in Section 3.7, while Sections 3.8 and 3.9 examine implications for the term structure in the case of stabilization and smoothing respectively. Section 3.10 concludes.

3.2 The model

The model is a reduced form of the standard model. To set the focus on short rate dynamics, the two state variables inflation and output are merged into one, referred to as the target variable. This assumption keeps the analysis simple, and because we focus on implications for the term structure of interest rates, the target variable can be regarded as a latent factor of the yield curve.

Thus there is one state variable x_t, the target variable. It is governed by the process

$$\begin{aligned} dx_t &= -\left\{a\left(x_t - \bar{x}\right) + b\left(r_t - \bar{r}\right)\right\} dt + \sigma dw_t, \\ x_0 &= x. \end{aligned} \tag{3.1}$$

Here w_t is a standard Wiener process, $\bar{x}$ and $\bar{r}$ are centering constants with $\bar{x}$ the target level and with $\bar{r} > 0$. For the parameters a, b and σ we assume that a and σ are non-negative and that b is positive. We deviate from previous studies in the restriction on a. If $a < 0$ the process for x_t is "mean fleeing," which is normally compensated for by an appropriate process for r_t ensuring that overall dynamics remains stable. In the present case however, when we impose a lower bound on r_t, the analysis has to be restricted to the case $a \geq 0$.

All parameters are given exogenously and unaffected by any actions taken by the monetary authority. This exposes the formulation to the Lucas critique. However, endogenity of parameters is not in our present focus.

The short rate r_t is set by the monetary authority, and since b is positive this makes it possible to control the dynamics of x_t. The objective is by assumption to keep x_t close to $\bar{x}$. If controlling x_t was the monetary authority's sole objective the optimization problem would be ill-posed, since the implied short rate rule would be arbitrarily aggressive. It is necessary to include measures that ensure a more moderate policy. We investigate two specifications that are common in the literature, interest rate stabilization and interest rate smoothing.

Starting with the case of interest rate stabilization the monetary authority's preferences are formed according to a loss function

$$L_1(x) = \rho E_x \left[\int_0^\infty \exp(-\rho t) \left\{ q (x_t - \bar{x})^2 + (r_t - \bar{r})^2 \right\} dt \right],$$

where $q > 0$ and $\rho \geq 0$ are parameters with $\rho = 0$ taken to be the limit as $\rho \downarrow 0$. Here the monetary authority has two objectives, it values low deviations from the target in x_t but is also concerned with keeping r_t stable around $\bar{r}$. This gives rise to a control problem where the objective is to find a rule for the short rate r as a function of x such that the loss function L_1 is minimized.

In the case of interest rate smoothing the monetary authority follows a policy that makes the short rate evolve smoothly over time. The process for r_t can then be written as

$$\begin{aligned} dr_t &= u_t dt, \\ r_0 &= r, \end{aligned}$$

where u_t is the control law. With this specification the short rate is locally deterministic, and with this we mean that there is no volatility term in the Itô differential of r_t. The loss function is formed according to

$$L_2(r, x) = \rho E_{x,r} \left[\int_0^\infty \exp(-\rho t) \left\{ q (x_t - \bar{x})^2 + u_t^2 \right\} dt \right],$$

where $q > 0$ and $\rho \geq 0$ are parameters. The control is no longer the short rate itself, but rather the time derivative of the short rate. In effect r_t becomes a state variable, and the formulation assumes in total two state variables, r_t and x_t. As a consequence this specification allows for richer dynamics although there is still only one driving Wiener process and the number of parameters is the same.

In the literature on optimal monetary policy it is most common to adopt the latter form of the loss function, i.e. with a smoothing objective for the short rate. This seems to better capture the tendency among central banks to adjust the short rate in several consecutive moves in the same direction, and to change the direction only infrequently. The literature has given some attention to the origins of this behavior and several possible explanations have been suggested. For a review together with some empirical evidence see Sack and Wieland (1999).

Below we study the monetary authority's control problem for each of the cases interest rate stabilization and interest rate smoothing. In both specifications we constrain the study of the optimal control law to the limiting case $\rho \downarrow 0$. This is not uncommon in the literature (see e.g. Svensson (2000)) and in the case of interest rate stabilization it also admits an analytical solution. In case the state variables have ergodic distributions L_1 and L_2 can be written more conveniently as

$$\begin{aligned} L_1 &= q \operatorname{var}(x_t) + \operatorname{var}(r_t), \\ L_2 &= q \operatorname{var}(x_t) + \operatorname{var}(u_t), \end{aligned}$$

where $\operatorname{var}(x_t)$, $\operatorname{var}(r_t)$ and $\operatorname{var}(u_t)$ denote stationary variances. Note in particular that L_1 and L_2 lose their dependence on x and (x, r).

3.3 Interest rate stabilization

In the case of interest rate stabilization a statement of the monetary authority's minimization problem is

$$V_1(x) = \min_r L_1(x),$$

where

$$L_1(x) = \lim_{\rho \downarrow 0} \rho E_x \left[\int_0^\infty \exp(-\rho t) \left\{ q(x_t - \bar{x})^2 + (r(x_t) - \bar{r})^2 \right\} dt \right],$$

where x_t follows the process

$$\begin{aligned} dx_t &= -\{a(x_t - \bar{x}) + b(r(x_t) - \bar{r})\} dt + \sigma dw_t, \\ x_0 &= x, \end{aligned}$$

and where the control r is subject to the constraint

$$r(x_t) \geq 0, \qquad \forall t > 0.$$

Had it not been for the constraint this minimization problem would have been a standard application of the linear-quadratic regulator, in which case the optimal control law is linear (or affine) in the state variables. This control law can be seen as the limiting law as $\bar{r} \longrightarrow \infty$.[1] For reference this is stated as a lemma.

[1] That is for a fixed value of x, as $\bar{r} \longrightarrow \infty$ the control laws for the constrained and the unconstrained problems converge. Letting $\bar{r}$ approach infinity is conceptually equivalent to a relaxation of the lower bound.

Lemma 3.1. *For the unconstrained problem the optimal control law is given by*

$$r_{free}(x) = A(x_t - \bar{x}),$$

where

$$A = -a/b + \sqrt{(a/b)^2 + q}.$$

Proof. See Appendix 3.11.

As a standard result in linear-quadratic optimal control the unconstrained law is independent of the volatility σ. This implies in particular that the control law for the corresponding deterministic problem, i.e. with $\sigma = 0$, is the same. This will not be the case when the constraint is taken into account. For $\sigma = 0$ the solution to the constrained problem is the non-negative part of the unconstrained law, that is

$$r_{det}(x) = \max\{r_{free}(x), 0\}.$$

For $\sigma > 0$ the optimal law is given by Theorem 3.1.

Theorem 3.1. *For the constrained problem the optimal control law can be written as*

$$r(x) = \max\{\psi(x), 0\},$$

where

$$\psi(x) = r_{free}(x) + \lambda(x - \bar{x}) \frac{\frac{d}{dz}\mathcal{M}(1/2, \alpha; z)}{\mathcal{M}(1/2, \alpha; z)},$$

$$z = \kappa(x - \bar{x})^2,$$

where M is a version of the confluent hypergeometric function and where the determination of the constants α, λ and κ are deferred to Appendix 3.11.

Proof. See Appendix 3.11.

The control law $r(x)$ for $\sigma > 0$ is below or equal to $r_{det}(x)$ for all values of x with equality exactly where $r_{det}(x) = 0$. Thus volatility in x_t induces a more expansive policy. One consequence of this is that with $\sigma > 0$ the zero bound is hit earlier, that is for a higher value of x. For reference we define x^* to be the value of x where the constraint activates. This is given implicitly by $\psi(x^*) = 0$.

Another consequence of the more expansive policy is that the steady state in the economy is reached at a value of x that exceeds $\bar{x}$. The steady state is defined as the value x_{ss} for which the drift term of x_t equals zero, i.e. where

$$a(x_{ss} - \bar{x}) + b(r(x_{ss}) - \bar{r}) = 0.$$

Because the left hand side is monotone in x_{ss} and because r is strictly below r_{free} whenever $u > 0$, it follows that equality obtains for an x_{ss} that is strictly higher than in the corresponding unconstrained problem.

3.4 Numerical examples on stabilization

This section presents numerical examples to illustrate the mechanics of the model. Our benchmark setting for the parameters is as follows.

$$\begin{array}{cccccc} a & b & \sigma & \bar{r} & \bar{x} & q \\ \hline 0.01 & 0.15 & 1.5\% & 3.5\% & 0 & 3 \end{array} \tag{3.2}$$

The values for a, b, σ and $\bar{r}$ are in line with the simple regression below, except that σ and $\bar{r}$ have been somewhat adjusted to make the graphical presentation more accessible. The value for $\bar{x}$ is arbitrary and therefore set equal to zero. The value for q is set to give a reasonable volatility of the short rate.

To motivate the parameter constellation in (3.2) and in order to acquire some intuition for the state variable x_t we pursue a simple estimation of parameters based on US data 1987 - 2002. We consider the short rate r_t as the three-month treasury rate, inflation π_t as the implicit GDP deflator and the output gap y_t as estimated by the Congressional Budget Office (2002). All data is quarterly. We detrend inflation and the output gap by deducting their sample means, and we then regress r_t on π_t and y_t. This yields

$$r_t = \underset{(0.15\%)}{5.5\%} + \underset{(0.14)}{1.17}\pi_t + \underset{(0.10)}{0.54}y_t$$

with asymptotic standard errors in parentheses. In terms of the model a natural definition of the state variable x_t is then

$$x_t = \omega\pi_t + (1-\omega)\,y_t$$

with $\omega = 1.17/\,(1.17 + 0.54) = 0.68$. This suggest that the monetary authority follows the rule $u_t = \bar{r} + Ax_t$ with $A = 1.71$ and $\bar{r} = 5.5\%$. Assuming that x_t is governed by the process (3.1) we then regress yearly increments in x_t on x_{t-1} and r_{t-1}. Without a constraint on a the estimate for b is negative, and we therefore impose $a = 0$ so that if left uncontrolled x_t would be unstationary. This gives

$$x_t - x_{t-1} = -\underset{(0.065)}{0.15}\,(r_{t-1} - \bar{r}) + 0.011\varepsilon_t.$$

A similar estimation using data for Japan 1980-2003 gave similar results, except that the sample mean of the treasury rate was lower, 3.3%.

For the benchmark constellation (3.2) Figure 3.1 depicts the optimal control law. Whenever $r > 0$ the constrained law is below the unconstrained, so taking account of the zero bound results in a more expansive policy. When $x = 0$ the constrained law is some 0.15% below the unconstrained. As x becomes large the two laws converge; asymptotic convergence is at the rate $1/x$.

The constrained law exhibits a kink at $x^* = -3.19\%$ as the constraint activates. It is worth noting how the constraint activates. As x falls the short rate accelerates towards zero hitting the barrier with positive speed. This appears to be in contrast with monetary policy in Japan over the last decade. The zero bound was approached gradually with small adjustments. In particular short rate volatility fell, indeed in contrast with the model which suggests that short rate volatility should increase.

The constrained law corresponds well to the results of Orphanides and Wieland (1999) and Kato and Nishiyama (2001). These papers consider a setting with two state variables, inflation and the output gap, which gives the optimal control law as a function of two variables, and they solve the control problem numerically. Still, their short rate appears to hit the zero bound in a similar manner as in our model, and their control law also appears to converge to the unconstrained law in a similar way.

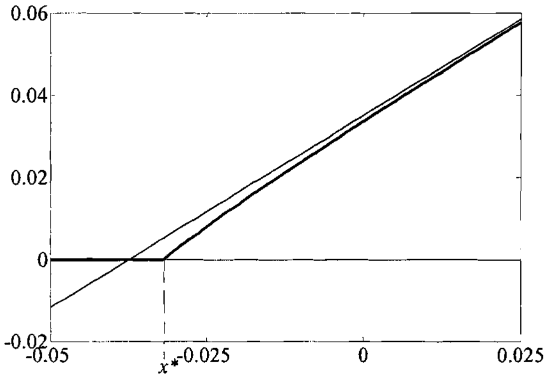

Figure 3.1. Optimal control law in the constrained (solid) and the unconstrained (dashed) problem. Parameter assessments as in (3.2).

Figure 3.2 illustrates the effect of changing volatility. As σ approaches zero the control law approaches $r_{det}(x)$. For $\sigma = 1\%$ the difference between the two is hardly economically significant.

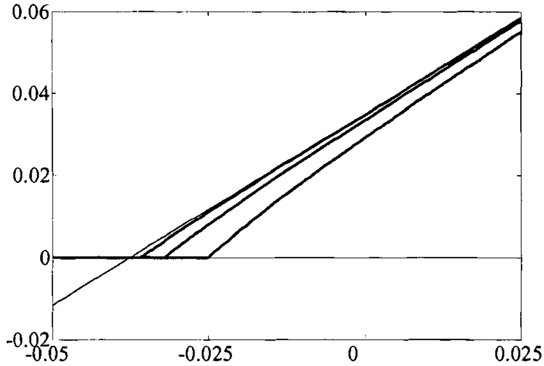

Figure 3.2. Optimal control law for $\sigma = 2\%$ (bottom thick), $\sigma = 1.5\%$ (middle thick) and $\sigma = 1\%$ (top thick). Apart from σ parameter assessments as in (3.2).

To gain some further perspective on the mechanics Figure 3.3 depicts what is similar to an impulse-response graph. Starting at $x_0 = -5\%$ we study one particular path for x_t, the one corresponding to $w_t = 0$ for all t. Because $x_0 < x^*$ the short rate is initially at the lower bound. As x_t passes $x^* = -3.19\%$, which occurs at approximately $t = 3$ years, monetary policy immediately becomes significantly more active.

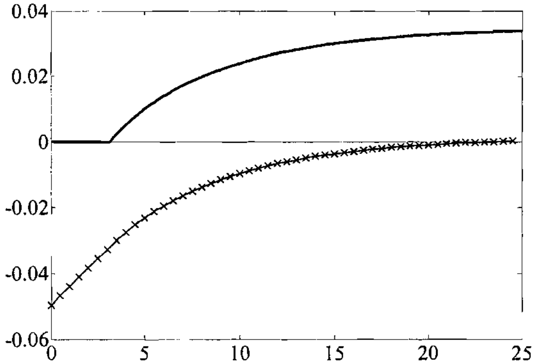

Figure 3.3. Optimal control law as a function of time when x_t follows $dx_t = -\{\ a(x_t - \bar{x}) + b(r_t - \bar{r})\ \}dt$. Solid line r_t and crossed line x_t. Time axis in years. Parameter settings as in (3.2).

3.5 Interest rate smoothing

In the case of interest rate smoothing the control problem becomes more involved and for the full constrained problem we do not have an analytical solution. We begin the study with a formal statement of the minimization problem and give its solution in absence of the constraint. Proceeding with the constrained problem we then consider the deterministic case first since this can be solved with the method of Pontryagin, and we finish with an outline of the numerical procedure to approximate the solution to the stochastic problem.

A statement of the monetary authority's minimization problem is

$$V_2(r,x) = \min_u L_2(r,x),$$

where

$$L_2(r,x) = \lim_{\rho \downarrow 0} \rho E_{r,x}\left[\int_0^\infty \exp(-\rho t)\left\{q(x_t-\bar{x})^2 + u_t^2\right\} dt\right],$$

as r_t and x_t follow

$$\begin{cases} dr_t = u_t dt, \\ dx_t = -\{a(x_t-\bar{x}) + b(r_t-\bar{r})\} dt + \sigma dw_t, \end{cases}$$

with initial values

$$\begin{cases} r_0 = r, \\ x_0 = x, \end{cases}$$

and where the state variable r_t is subject to the constraint

$$r_t \geq 0, \qquad \forall t > 0.$$

When working with a constrained control problem it is highly desirable to have the constraint on the control rather than on one of the state variables. Fortunately we can rephrase the constraint as a constraint on the control since we simply require that $u_t \geq 0$ whenever $r_t = 0$. We define the set of admissible controls as

$$\mathcal{U} = \{u_t : u_t \geq 0 \text{ whenever } r_t = 0\}.$$

Without the constraint the problem is a version of the linear-quadratic regulator with an optimal control law that is affine in the state variables. Lemma 3.2 gives the optimal unconstrained law.

Lemma 3.2. *For the unconstrained problem the optimal control law is given by*

$$u_{free}(r, x) = B(r - \bar{r}) + C(x - \bar{x}),$$

where

$$B = a - \sqrt{a^2 + 2b\sqrt{q}}, \qquad C = (1/2b)\, B^2.$$

Proof. See Appendix 3.12.

As a consequence of the linear-quadratic form the unconstrained law is independent of the volatility σ, and in particular the control law for the corresponding deterministic problem is the same. This will not hold for the corresponding constrained problem.

For the constrained problem the deterministic case $\sigma = 0$ can be solved with the Pontryagin minimum principle. This solution technique is inherently different from that of optimal control as it for given initial values of the state variables characterizes the solution in terms of an optimal path for u, i.e. u as a function of time. We are usually interested in the solution on feed-back form, i.e. u as a function of r and x, and then the optimal control approach is more natural. If this is not successful Pontryagin's principle is however worth pursuing.

The implementation of the program was rather straightforward although it involves a numerical one-dimensional optimization. The methodology is outlined in Appendix 3.13. Because the integral in the loss function remains well-defined even as $\rho \downarrow 0$ there is no need to pre-multiply by ρ, and we define the value function corresponding to the deterministic problem, $V_{det}(r, x)$, according to

$$V_{det}(r, x) = \min_{u \in \mathcal{U}} \int_0^\infty \left\{ q\,(x_t - \bar{x})^2 + u_t^2 \right\} dt,$$

as

$$\begin{cases} \dot{r}_t = u_t, \\ \dot{x}_t = -\left\{ a\,(x_t - \bar{x}) + b\,(r_t - \bar{r}) \right\}. \end{cases}$$

The corresponding control law is denoted $u_{det}(r, x)$.

For the case $\sigma > 0$ we adopt a numerical procedure. We are interested in the case $\rho = 0$ and in the literature on optimal control this is known as average cost minimization, see e.g. Arapostathis, Borkar, Fernández-Gaucherand, Ghosh and Marcus (1993) for a review. It is common to rephrase the problem with a finite horizon T and consider the limiting law as $T \to \infty$. For a numerical implementation this approach is also more natural since we can iterate the value function with increasing values of T.

Next, in comparison with many other control problems in economics volatility is likely to be small and consequently one can expect the control law to be close to $u_{det}(r, x)$. We therefore use the deterministic value function V_{det} as a final value function. Thus we consider

$$V^T(r, x) = \min_u L_2^T$$

with

$$L_2^T = \frac{1}{T} E_{r,x}\left[\int_0^T \left\{q(x_t - \bar{x})^2 + u_t^2\right\} dt + V_{det}(r_T, x_T)\right].$$

For this problem the optimal control law is no longer time homogenous and $u_t = u^T(r, x, t)$ with calender time t as an additional argument. The idea is that as T approaches infinity the initial control law $u^T(r, x, 0)$ should converge to the control law of the original problem, $u(r, x)$. This will hold provided that

$$\lim_{T\to\infty} V^T(r, x) = V_2(r, x).$$

It has not been shown that this convergence holds, but it will hold if x_t and u_t are sufficiently well-behaved under the optimal law. This is discussed in Arapostathis et al.

When applying a numerical procedure to this problem, the fact that one of the state variables is locally deterministic complicates the issue of numerical stability. Noise in the state variables tend to smooth irregularities, and it is therefore common to consider a slightly perturbed problem with noise in all state variables, see e.g. Goodman, Moon, Szepessy, Tempone and Zouraris (2002). Adding a noise term to the process for r_t state dynamics is given by

$$\begin{cases} dr_t = u_t dt + \sigma_\varepsilon dw_{\varepsilon t}, \\ dx_t = -\{a(x_t - \bar{x}) + b(r_t - \bar{r})\} dt + \sigma dw_t, \end{cases}$$

with $w_{\varepsilon t}$ a standard Wiener process independent of w_t, and $\sigma_\varepsilon > 0$ a small constant. With this process for r_t the constraint $r_t \geq 0$ is no longer feasible, so we redefine the set of admissible controls as

$$\mathcal{U}' = \{u_t : u_t \geq 0 \text{ whenever } r_t \leq 0\}.$$

To now solve the problem numerically the value function is updated recursively according to

$$V^{[k]}(r, x) = \min_{u_t \in \mathcal{U}'} L^{[k]}(r, x),$$

where

$$L^{[k]}(r, x) = E_{r,x}\left[\int_0^{\Delta t} \left\{q(x_t - \bar{x})^2 + u_t^2\right\} dt + V^{[k-1]}(r_{\Delta t}, x_{\Delta t})\right],$$

for $k = 1, 2, \ldots$, with Δt a small time step and with the initial value for $V^{[k]}$ as

$$V^{[0]}(r, x) = V_{det}(r, x).$$

The recursion is carried out on a grid over r and x. At each gridpoint $V^{[k-1]}$ is locally approximated with a quadratic function, which produces a standard linear-quadratic control problem, for which standard solution techniques exist. The method works well except close to the constraint where the quadratic approximation is worse. In this region $V^{[k-1]}$ is approximated with a different functional form and the local control problem is solved numerically. The above procedure works but is slow and requires some 30 hours on a 1000MHz computer. The procedure is outlined in Appendix 3.14.

3.6 Numerical examples on smoothing

For our numerical examples we use the following benchmark setting for the parameters.

a	b	σ	$\bar{r}$	$\bar{x}$	q
0.01	0.15	1.5%	3.5%	0	3

(3.3)

For the numerical approximation we set $\sigma_\varepsilon = 0.1\%$. The weight q does not have the same meaning as in the case of interest rate stabilization and it is not immediate how to assess an appropriate value. Previous studies usually consider two target variables, inflation and the output gap, and results are therefore not directly comparable. Still, in this instance the objective function assumes three weights, q_π, q_y and q_u corresponding to inflation, the output gap and smoothing respectively. Söderström, Söderlind and Vredin (2002) report that for the US typical estimates are such that q_u/q_π is in the range 0.5 to 2 and q_y/q_π is less than 0.1. Using term structure data Skallsjö (2002) finds $q_u/q_\pi = 0.32$ and $q_y/q_\pi = 0$, also for the US. If comparable the corresponding figure for the parameter constellation (3.3) is $1/q = 0.33$.

In principle it should be possible to estimate the model with maximum likelihood, but this is not so straightforward due to the numerical computation of the control law. In ongoing research, not reported here, the control law is approximated and the model is estimated using term structure data for Japan. This leads to a value of q equal to 3.7 with a standard error of 1.35.

For the parameter constellation (3.3) the optimal control law $u(r, x)$ is depicted in Figure 3.4. Along the line $r = 0$ we have the constraint $u \geq 0$. There is a critical value x^* such that for $x < x^*$ the monetary authority would like to set $u < 0$ but is constrained to $u = 0$. For $x > x^*$ the monetary authority has no desire to lower r_t and selects an admissible control $u > 0$. For current parameter settings $x^* = -1.48\%$.

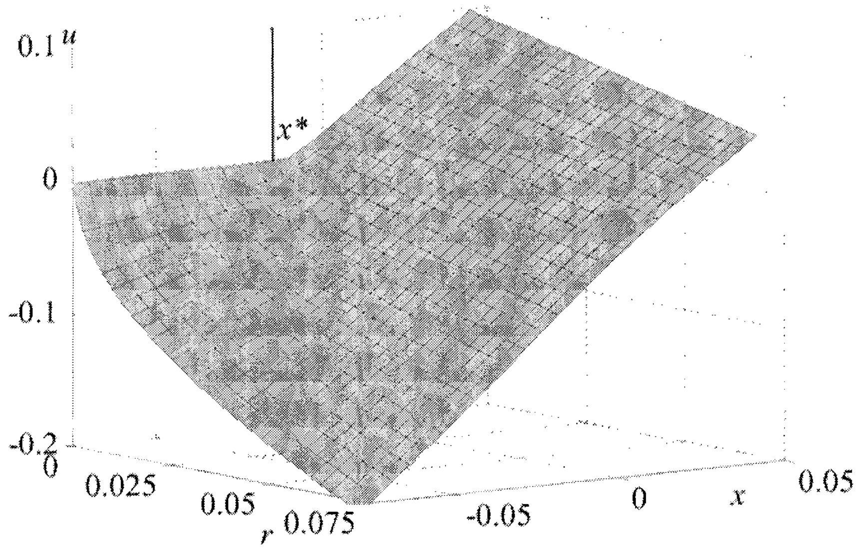

Figure 3.4. Optimal control law. Parameter settings as in (3.3).

Figure 3.5 displays u in comparison with u_{free}, this time at a more distant view. Similarly to the case of interest rate stabilization the control law can be both above and below its unconstrained counterpart. We have $u < u_{free}$ whenever x is sufficiently high, which appears to be $x > x^*$. As $x \to \infty$ the controls converge. We also have $u < u_{free}$ for any x provided r is sufficiently high, and as $r \to \infty$ the solutions diverge. For a fixed x as $r \to \infty$ the drift term for x_t becomes negative and high in magnitude, which pushes future expected values of x_t below $\bar{x}$. It is then urgent to bring r_t to the zero bound more quickly and as a consequence the constrained control law is below the unconstrained.

We have $u > u_{free}$ whenever x is sufficiently low and r is sufficiently close to zero. In this region the problem exhibits an important difference to the case of interest rate stabilization. For a fixed $x < x^*$ letting $r \downarrow 0$, u approaches zero smoothly. Consequently one would expect r_t to approach zero smoothly with small adjustments, in contrast to the case of stabilization where the optimal control law was as most responsive when r_t was close to zero. Recalling the discussion on the Bank of Japan, in this regard a smoothing objective for the short rate seems to better capture their policy over the last decade.

A natural question at this point is wether the zero bound will ever be reached, and it can be shown that with the smoothing objective r_t will reach zero in finite time with probability one.

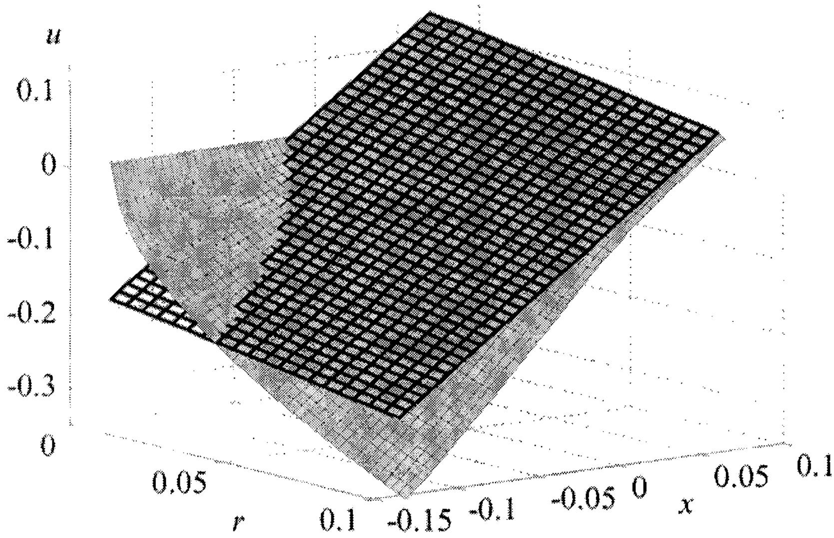

Figure 3.5. Optimal control law in the constrained (grey) and unconstrained (black wireframe) Parameter settings as in (3.3).

To assess the role of volatility Figure 3.6 depicts u in comparison with u_{det}, and Figure 3.7 the difference $u - u_{det}$. From Figure 3.7 it appears that $u \leq u_{det}$ for all values of r and x with equality exactly as $x \leq x^*$ and $r = 0$. This is in line with the results for interest rate stabilization where introducing volatility resulted in a more expansive rule whenever r was above zero. Further it appears from the graph that the value of x^* does not vary with σ. Loosely speaking one would expect the stochastic control law to be a smoothed version of the deterministic law. It appears that smoothing at $x = x^*$ works by flattening out the control law to the right of x^*, keeping the value of x^* unaffected.

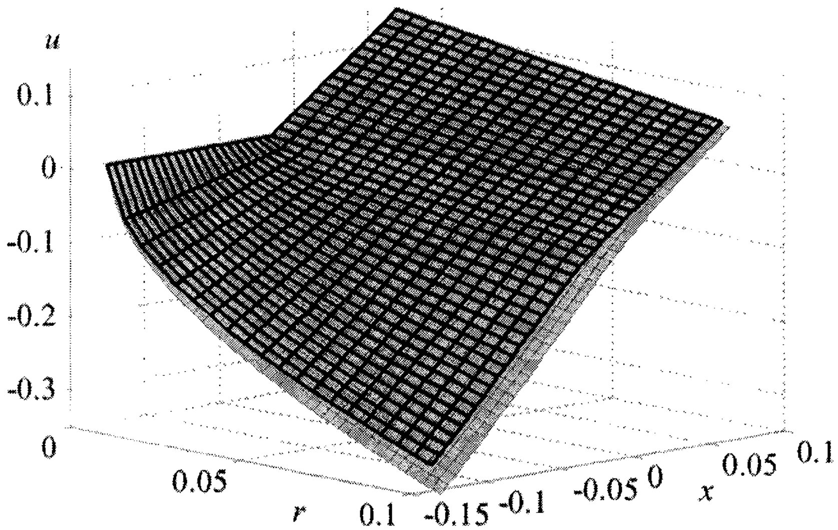

Figure 3.6. Optimal control law in the stochastic (grey) and deterministic (black wireframe) problem. Parameter settings as in (3.3).

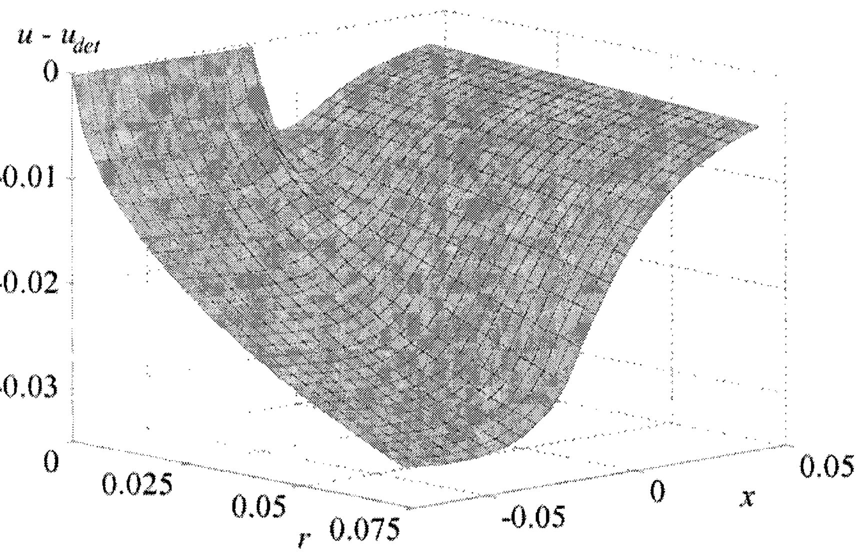

Figure 3.7. The difference $u - u_{det}$. Parameter settings as in (3.3).

Figure 3.8 depicts the equivalent to Figure 3.7 for various values of σ. As in the case of stabilization, for $\sigma = 1\%$ the difference between the stochastic and the deterministic control laws is small.

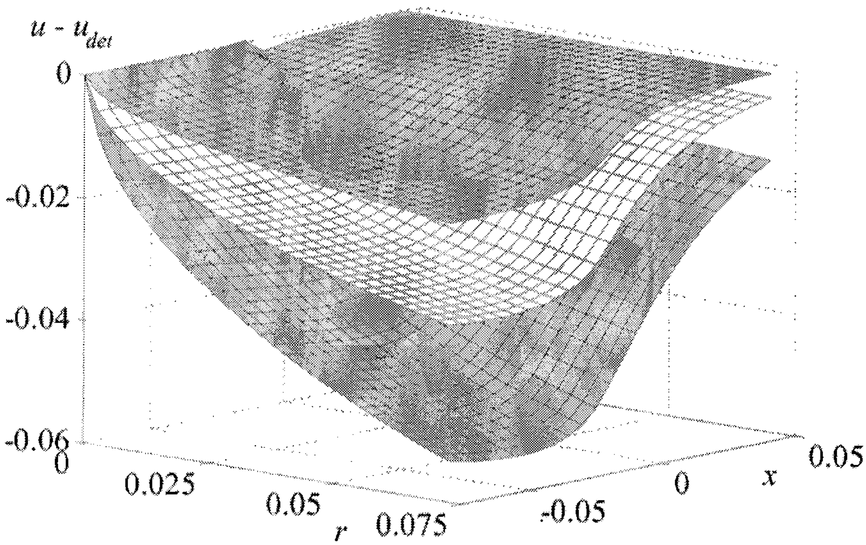

Figure 3.8. The difference $u - u_{det}$ for σ=1% (top), σ=1.5% (middle) and σ=2% (bottom). Parameter settings as in (3.3).

In Figure 3.9 we give the corresponding item to Figure 3.3, the impulse-response graph. Starting the system at $(r, x) = (1.5\%, -2.5\%)$ we study paths corresponding to $w_t = 0$ for all t (as well as $w_{\varepsilon t} = 0$ for all t). Initially $u_t < 0$ and r_t falls, hitting the zero bound after $t = 0.9$ years. The short rate then remains at zero until x_t stabilizes to $x^* = -1.48\%$ which in the figure occurs after $t = 2$ years. Then r_t is successively adjusted back towards the steady state, though initially very slowly. For current parameter settings as $t \rightarrow \infty$ paths are oscillating with diminishing amplitude. This holds true also for the unconstrained problem.

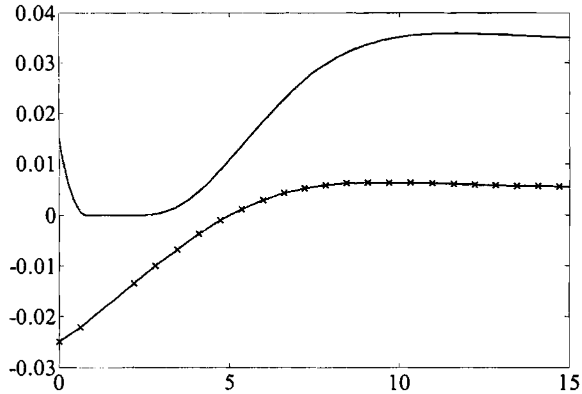

Figure 3.9. Optimal control law as a function of time when state dynamics is $dr_t = u_t dt$ and $dx_t = -\{a(x_t - \bar{x}) + b(r_t - \bar{r})\}dt$. Solid line r_t and crossed line x_t. Time axis in years. Parameter settings as in (3.3).

We conclude this section with a note on the steady state. Figure 3.10 depicts a vector field of the state dynamics. The steady state obtains at approximately $(r_{ss}, x_{ss}) = (3.47\%, 0.58\%)$ to be compared to the free problem's counterpart $(3.5\%, 0)$. For x_{ss} we note that the inclusion of the constraint implies an x_{ss} that is higher than its unconstrained counterpart, which is directly comparable to the case of stabilization. For current parameter assessments the difference, 0.58%, is somewhat higher than the corresponding 0.15% in the case of interest rate stabilization. However, since the parameter q has a different meaning in the two models, the result should not be overemphasized.

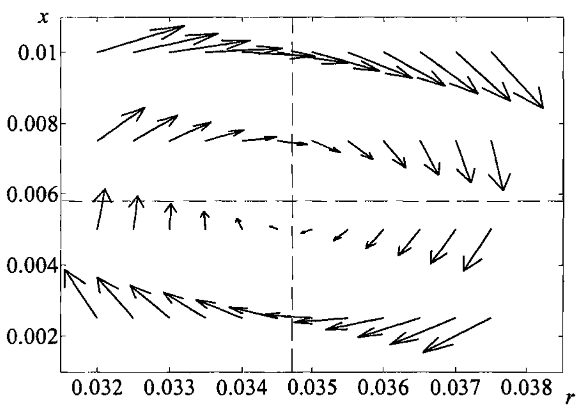

Figure 3.10: Vector field of dynamics in the stochastic constrained problem. Motion in x (along the $y-$axis) multiplied by a factor 10. Parameter settings as in (3.3).

3.7 Some empirical background

Before looking at model implications for the term structure of interest rates we give some empirical background. Japan with its prolonged period of deflation and interest rates close to zero makes a natural example.

Figure 3.11 depicts time series of the short rate r_t, measured by the central bank's discount rate, inflation π_t, measured by the implicit GDP deflator and the output gap y_t as estimated by IMF.

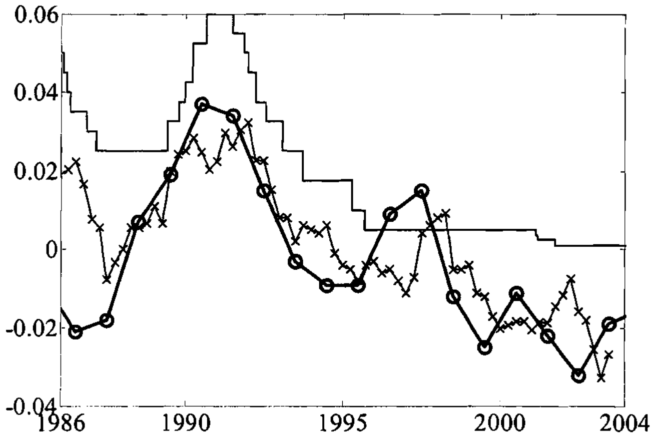

Figure 3.11. The Bank of Japan discount rate (solid), inflation (crosses), and the output gap (circles).

Figures 3.12 and 3.13 depict a series of yen yield curves and yield volatility curves over the period 1996 - 2003. Supplementary graphs for intermediate dates are given in Appendix 3.16. Data refer to swap rates, which are not directly comparable to the analysis in the remainder of the paper but patterns should be similar.

Over the period yields fall rather steadily except for a rebound in 1998 - 1999. For the initial dates, in 1996 and 1997, we infer from Figure 3.11 that inflation is around zero while the output gap is slightly positive. For these dates yield curves and yield volatility curves are both concave. The initial volatility curve, in 1996, exhibits the characteristic hump familiar from other currencies and more normal monetary circumstances. In 2001 and 2003 inflation and the output gap have both fallen to around -2%. For these dates yield curves and yield volatility curves assume a different shape. In 2001 the yield curve is slightly s-shaped and in 2003 the s appears to have flattened out to the right. Volatility curves exhibit a similar pattern.

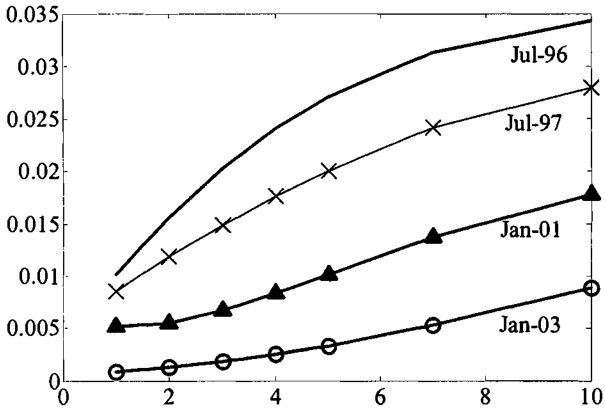

Figure 3.12. Swap rate curves. Swap rates refer to the beginning of the indicated month.

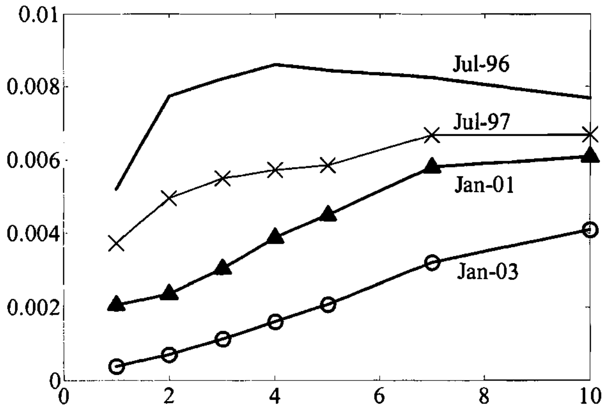

Figure 3.13. Swap rate volatility curves. Volatility curves depict the daily standard deviation in the swap rate for a given maturity over the six months following the indicated date.

3.8 The term structure in the case of stabilization

In this section we take as given the state dynamics implied by the model in Section 3.3 and investigate implications for the term structure of interest rates. We follow the standard approach in term structure modelling, forming the price of a unit discount bond with maturity τ as

$$F(x;\tau) = E_x^Q \left[\exp\left(-\int_0^\tau r_s ds \right) \right],$$

where superindex Q indicates that the expectation is taken under the risk-neutral measure. In general the risk-neutral measure will be different from the objective measure, a fact that could be accounted for by assuming a parameterized form for the market price of risk. Here, in order to make the analysis more transparent we take the simplified route and assume risk-neutrality. We can then write

$$F(x;\tau) = E_x \left[\exp\left(-\int_0^\tau r_s ds \right) \right].$$

To solve for bond prices a numerical procedure is required. Fixing a small time step Δt bond prices are calculated recursively according to

$$\begin{aligned} F(x;0) &= 1, \\ F(x;k\Delta t) &= \exp\left(-\int_0^{\Delta t} \mathbf{r}_s(x)\, ds \right) E_x \left[F\left(x_{\Delta t}, (k-1)\Delta t \right) \right], \\ \mathbf{r}_t(x) &\equiv E_x \left[r(x_t) \right], \end{aligned}$$

for $k = 1, 2, \ldots$ For $\mathbf{r}_t(x)$ the expectations are approximated using local approximations of the dynamics of x_t as well as of the functional form for r. Local time has been applied to take account of the kink in r. The procedure is outlined in Appendix 3.15, together with a proof that the recursion indeed converges to the true prices as $\Delta t \downarrow 0$.

We focus our analysis on yields and yield volatilities. Given a bond price F the yield is defined as

$$y(x;\tau) = -(1/\tau) \times \ln F(x;\tau).$$

The yield volatility is then obtained as

$$v(x;\tau) = \frac{d}{dx} y(x;\tau) \times \sigma.$$

Figure 3.14 depicts yield curves for various values of the target variable x with parameter settings as in (3.2). Note that for $x < x^*$, in this case -3.19%, the initial short rate is constrained to zero. One implication of this is that for $x < x^*$ curves assume a convex shape for low maturities, moving to concave for higher maturities. The change in curvature produces an s-shape, though for current parameter settings this is weak – there are extended segments around the inflection point where curves are close to linear. The change in curvature obtains only if the initial short rate is zero. This need not be a serious limitation, but it may demand an open mind regarding the interpretation of the zero bound.

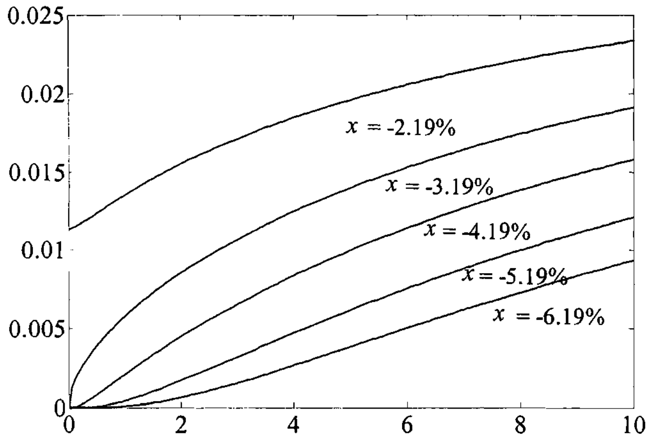

Figure 3.14. Yield curves. Maturity axis in years. Parameter settings as in (3.2).

For the parameter constellation (3.2) used in Figure 3.14, to obtain a yield curve below 1% for all maturities up to ten years x must be some 6% below its target value. From Section 3.4 we recall that a rough interpretation of x is deviation from target inflation. To explain a yield curve below 1%, the market should thus perceive inflation to be 6% below its target value. This is certainly possible but it would be comforting if more moderate values would suffice. Furthermore, we have calculated yield curves under risk-neutrality. With a positive market price of risk yield curves should be more upward sloping, which in turn would require an even lower value of x.

We conclude that model implied yield curves are reasonable, possibly with two sources of concern. First, curvature in model implied curves is weak and second, the value of x required to explain a low level of the yield curve, below 1% for maturities up to ten years, may be excessive. Both these results may in part be due to the high value of σ (equal to 1.5%) and for this reason Figure 3.15 displays yield curves corresponding to the more modest $\sigma = 1\%$. In this case the control law is close to $\max\{u_{free}(x), 0\}$ and $x^* = -3.56\%$ (cf. Figure 3.2). The lower σ helps to resolve both issues, though not to full satisfaction and in particular curvature is still weak. As noted in Section 3.4 it is possible that the short rate inhibits an excessive sensitivity to x.

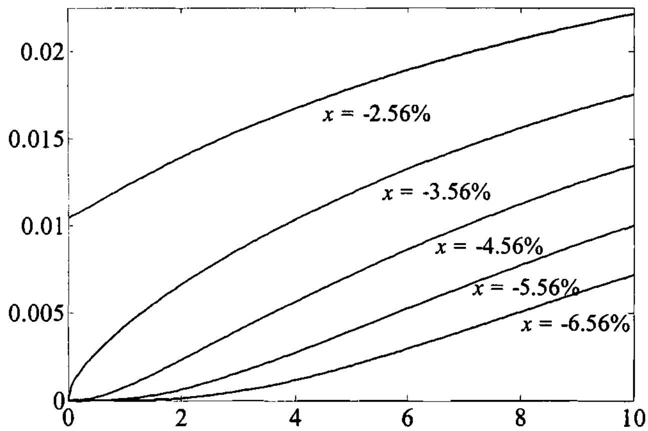

Figure 3.15. Yield curves. Maturity axis in years. Parameter settings as in (3.2) except $\sigma = 1\%$.

The yield curves' dependence on volatility is depicted in Figure 3.16. Changing volatility affects the yield curve in several ways. As in most term structure models higher volatility in the state variable implies a higher expected capital gain on long term bonds. If investors are risk-neutral they then require a lower yield, and increasing volatility therefore tends to decrease long term yields. In the current constrained problem there are other effects also that tend to dominate when the short rate is close to zero.

First, the short rate is a non-linear function of the underlying state variable. Keeping the functional form fixed, when r is low increasing volatility increases the expected value of future short rates, thus elevating long term yields. This is comparable to the effect volatility has on regular options. A second effect is that as we increase volatility the monetary policy rule adapts and becomes more expansive. This works in the opposite direction. In the first graph of Figure 3.16, when x is relatively high, the latter effect (the adjustment of the monetary rule) dominates and higher volatility implies a lower yield curve. In the remaining graphs, when x is low, the former (option-like) effect dominates and higher volatility implies a higher yield curve.

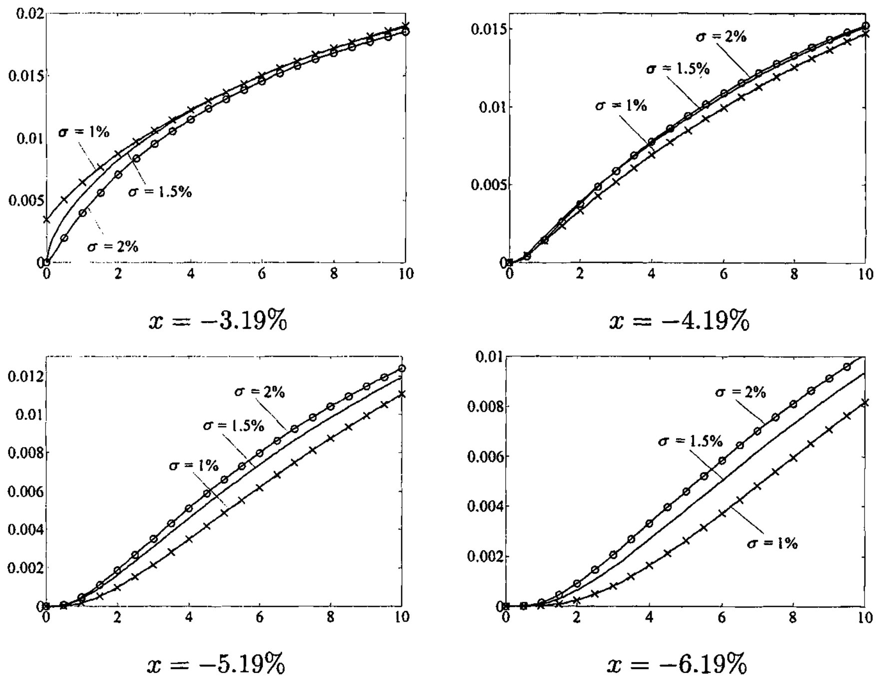

Figure 3.16. Yield curves for different values of σ and x. Maturity axis in years. Parameter settings apart from σ as in (3.2).

Figure 3.17 depicts yield volatilities for various values of x. The short rate volatility is of particular interest and is given by

$$v(x;0) = \begin{cases} 0 & x < x^* \\ \frac{1}{2}\sigma r'_+(x^*) \ , & x = x^* \ , \\ \sigma r'(x) & x > x^* \end{cases}$$

where r' is the derivative of the control law from Theorem 3.1 and where subindex '+' indicates limit from the right. Thus the short rate volatility is discontinuous in x. At $x = x^*$ it is simply the average of volatility to the left and to the right. In Figure 3.17 this occurs at $x = -3.19\%$.

Comparing the volatility curves for $x > x^*$ we see that they intersect. This follows as the optimal rule is as most responsive when x is above but close to x^*. In this region for x yield volatilities for short maturities are high, while the probability of hitting zero dampens the volatility in higher maturity yields. As $x \to \infty$ the volatility curve converges to that of the unconstrained problem, which in turn is a version of Vasicek with an exponential decline. The limiting curve is close to the one corresponding to $x = -0.19\%$ in the figure.

For $x < x^*$ we have $r = 0$ and the initial volatility is zero. Volatility curves then exhibit a hump. With lower values of x the hump is transferred to the right and becomes less pronounced. Similarly to the yield curves, when $x < x^*$ volatility curves are initially convex moving to concave for higher maturities.

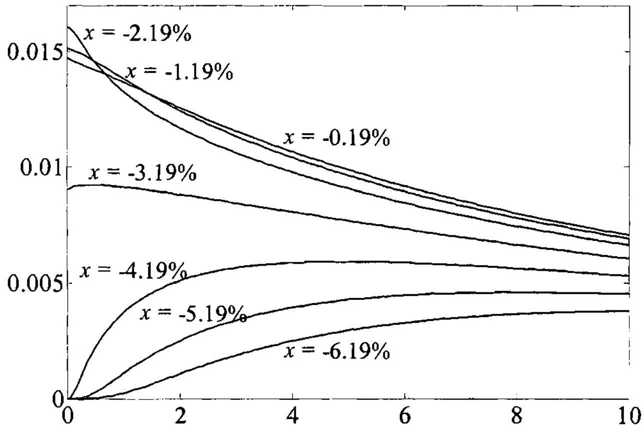

Figure 3.17. Yield volatilities in the constrained problem for different values of x. Maturity axis in years. Parameter settings as in (3.2).

3.9 The term structure in the case of smoothing

We now turn to interest rate smoothing and study the implications for the term structure of interest rates, again under the assumption of risk-neutrality. We calculate bond prices numerically and adopt a recursion similar to that in the previous section. Letting $F(r, x; \tau)$ denote the price of a unit discount bond with maturity τ for initial values (r, x) we fix a small time step Δt and apply the following procedure for $k = 1, 2, \ldots$

$$F(r, x; 0) = 1,$$

$$F(r, x; k\Delta t) = \exp\left(-\int_0^{\Delta t} \mathbf{r}_s(r, x)\, ds\right) E_{r,x}\left[F(r_{\Delta t}, x_{\Delta t}, (k-1)\Delta t)\right],$$

$$\mathbf{r}_t(r, x) = E_{r,x}[r_t].$$

The recursion is performed on a grid over r and x. Because we calculate prices under risk-neutrality $\mathbf{r}_t(r, x)$ is immediately available from the numerical approximation to the control problem. For the expectation of F we make a local approximation of $\ln F$ according to a quadratic functional form over r and x. The distribution for $(r_{\Delta t}, x_{\Delta t})$ is approximated according to a bivariate normal, and with these approximations the expectation of F can be calculated. Close to the zero bound the quadratic approximation for $\ln F$ is worse and a different functional form is used.

As in the previous section given a bond price F we define its yield according to

$$y(r, x; \tau) = -(1/\tau) \ln F(r, x; \tau).$$

For the yield volatility we define

$$v(r, x; \tau) = \frac{d}{dx} y(r, x; \tau) \times \sigma.$$

For v we thus neglect the term σ_ε that was imposed in the numerical approximation.

Yield curves now take two arguments, r and x. In Figure 3.18 we set $r = 0$ and depict yield curves for various values of x. For current parameter settings $x^* = -1.48\%$. Since the initial short rate is zero for all the curves, low maturity yields do not vary much with x. For high maturity yields the dependence on x is stronger.

As in the case of stabilization curves exhibit a convex and a concave segment, though with smoothing convexity is more pronounced. The reason is that with a smoothing objective the control law is close to zero in a region around $(r, x) = (0, x^*)$. Therefore, even if x_t should stabilize quickly the control law remains close to zero for some time. This implies that r_t leaves zero smoothly and slowly. Further away from zero the control law is more sensitive to variations in x, and as a consequence in this region the short rate is more responsive.

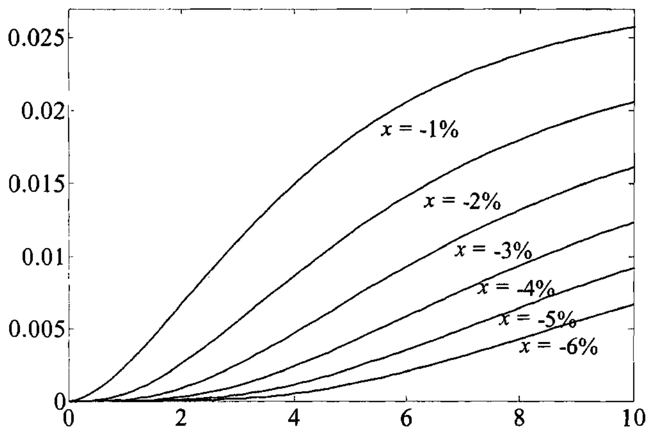

Figure 3.18. Yield curves for different x with r fixed at $r = 0$. Maturity axis in years. Parameter settings as in (3.3).

In the previous section we raised a question regarding the value of x needed to obtain a low level of the yield curve, below 1% for all maturities up to ten years. The required value of x seemed excessive, though this partly was resolved by considering a lower value of σ, equal to 1%. Figure 3.19 depicts yield curves under smoothing for $\sigma = 1\%$. The curve corresponding to $x = -5\%$ is below 1% with some margin, and in comparison with Figure 3.15 the value of x required to obtain a low level yield curve is more moderate. Again the result follows from the fact that the short rate leaves zero smoothly. However, as the parameter q has a different meaning in the two specifications the results should not be overemphasized.

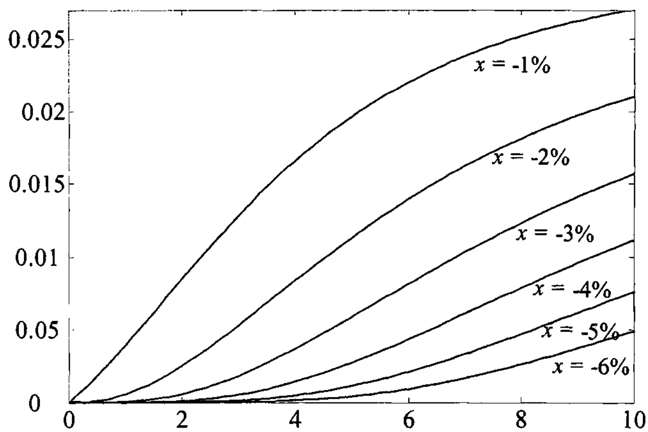

Figure 3.19. Yield curves for different x with r fixed at r =0. Maturity axis in years. Parameter settings as in (3.3) except σ= 1%.

The yield curves' sensitivity to σ is depicted in Figure 3.20. As in the case of stabilization there are two major counterweighing effects. First, the control law is a convex function of x. Keeping this functional form fixed there is an option-like effect, by which an increase in volatility gives higher yields. The second effect is that with an increase in volatility the control law adjusts and becomes more expansive, and this works in the opposite direction. When x is relatively high the latter effect (the adjustment of the monetary rule) dominates, and in the first exhibit in Figure 3.20 higher volatility is associated with a lower yield curve. When x is low the former (option-like) effect dominates, and in the two lower exhibits of Figure 3.20 higher volatility is associated with a higher yield curve.

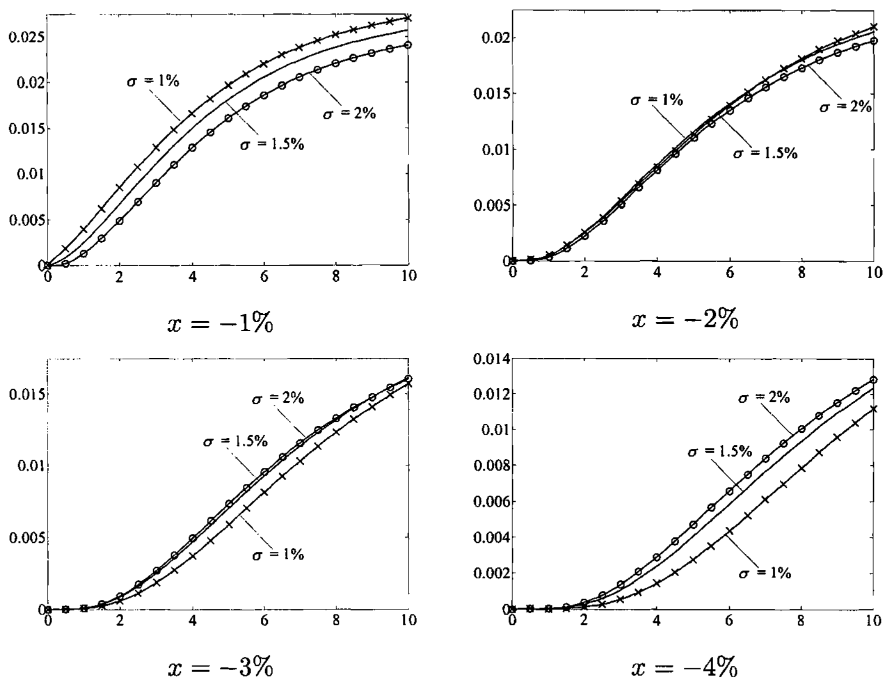

Figure 3.20. Yield curves for different values of σ and x when r =0. Maturity axis in years. Parameter settings apart from σ as in (3.3).

Now turning to yield volatilities we again set $r = 0$ and study the dependence on x. Curves are depicted in Figure 3.21. As $x \to \infty$ the volatility curve converges to that of the unconstrained problem, which is close to the one corresponding to $x = 0$ in the figure. The well-pronounced hump follows from monetary policy inertia. Because the short rate is locally deterministic volatility in the target variable does not transmit directly into volatility in the short rate, but induces uncertainty about the average short rate over a given horizon. This causes the volatility curve to be initially upward sloping. Over the longer horizon however the system is stabilized, and this causes curves to be downward sloping at the long end.

The volatility curves exhibit both convex and concave segments. For low values of x curves are initially convex moving to concave for higher maturities. This is also what we found in the case of interest rate stabilization. A difference between the two specifications is however that with smoothing the convex segment is more pronounced. More specifically, the inflection point generally obtains at a higher maturity. For example, the curve corresponding to $x = -4\%$ has an inflection point at a maturity of approximately 4 years, while in Figure 3.17 the location of the inflection point does not exceed 2 years for any of the curves. This finding is similar to what we found for yield curves above.

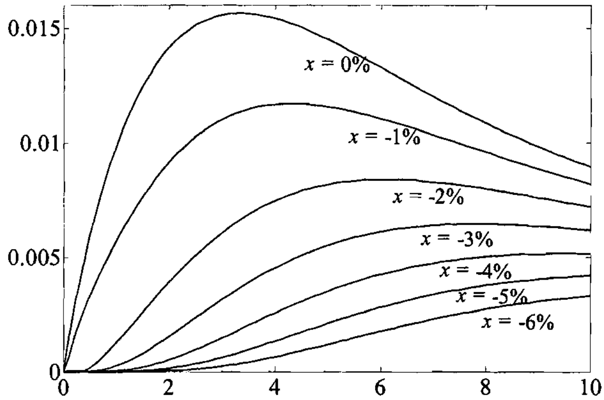

Figure 3.21. Yield volatilities for different values of x when r =0. Maturity axis in years. Parameter settings as in (3.3).

So far we have set the initial short rate $r = 0$. In Figure 3.22 we set $r = 2\%$ and let x vary. Again we focus on low values of x. In each scenario the short rate is expected to first decrease as x_t is stabilized, and then successively adjust back towards the steady state (cf. Figure 3.9). The depicted values of (r, x) are chosen for illustrative purposes. For example, conditional on $x = -5\%$ it is unlikely to observe an r as high as 2%, as in this region for x the short rate adjusts very quickly. The figure nevertheless illustrates a wide range of possible yield curves.

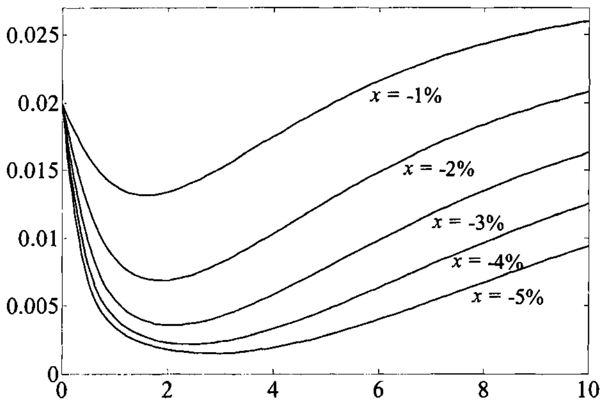

Figure 3.22. Yield curves for different x when r =2%. Maturity axis in years. Parameter settings as in (3.3).

Figure 3.23 depicts the volatility curves corresponding to the previous figure. A striking characteristic is the double hump assumed for the lower values of x. Curves are initially upward sloping but then follows a set of maturities where curves are downward sloping. In this range the short rate is expected to be close to the zero bound. For higher maturities the usual argument with monetary policy inertia applies, and hump number two obtains, though for the lower values of x this is located outside the picture.

We can conclude that the model allows for a broad collection of both yield curves and yield volatility curves when the short rate is close to zero.

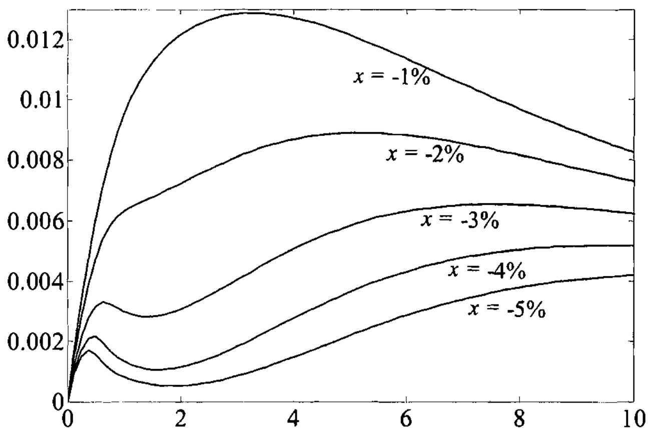

Figure 3.23. Yield volatilities for different values of x with r fixed at r =2%. Maturity axis in years. Parameter settings as in (3.3).

3.10 Conclusion

The paper has investigated two formulations of the objective function of monetary policy – interest rate stabilization and interest rate smoothing – and their implications for term structure dynamics. We have put particular emphasis on situations when the short rate is close to zero. An important difference between the two formulations is that with smoothing the short rate is locally deterministic. With regard to term structure dynamics this translates into a lower volatility in short-term yields. We find that when the short rate is close to zero this effect is amplified. Indeed, in one aspect the zero bound affects the optimal control law diametrically different in the two specifications. With interest rate stabilization the short rate is as most responsive to changes in the target variable just above the zero bound, whereas with smoothing responsiveness is at its weakest in this region. The implications for the term structure is that with smoothing the lower volatility obtained for short term yields transmits into higher maturity yields. This results in a more pronounced convexity in the initial segment of the yield curve. The same holds for yield volatility curves.

Using data for Japan it is in principle possible to put the two specifications to a formal test against each other. This exercise is complicated by the numerical approximations, which makes a more exact computation of bond prices time-consuming. Less accurate numerical approximations result in irregularities in the likelihood function, which implies that standard automatic optimization routines do not apply. However, because the number of parameters is small a manual search algorithm may be feasible.

An interesting extension of the model concerns the monetary transmission mechanism. The literature on monetary policy emphasizes the importance of forward-looking behavior, whereas the model studied in this paper is purely adaptive. The continuous time framework may however be well suited for studying forward-looking behavior.

3.11 Appendix. Control laws in the case of stabilization

This appendix presents the solutions to the control problems in Section 3. We may without loss of generality set $\bar{x} = 0$. We begin with the unconstrained problem and then turn to the constrained.

3.11.1 The unconstrained problem

We begin with the unconstrained problem. For $\rho > 0$ we consider

$$V(x;\rho) = \min_r E_x \left[\int_0^\infty e^{-\rho t} \left\{ q x_t^2 + (r(x_t) - \bar{r})^2 \right\} dt \right].$$

Note that the expectation term has not been pre-multiplied by ρ. The HJB equation for this problem is

$$0 = \min_r \left\{ -\rho V + q x^2 + (r - \bar{r})^2 - (ax + b(r - \bar{r})) V_x + \frac{1}{2} \sigma^2 V_{xx} \right\},$$

where subindices indicate derivatives.

We are interested in the limiting problem as $\rho \downarrow 0$. In the literature on optimal control this is referred to as average cost minimization, see e.g. Arapostathis, Borkar, Fernández-Gaucherand, Ghosh and Marcus (1993) for a review. Although the value function in this case becomes unbounded, a limiting control law may still exist. This is obtained from a limiting form of the HJB equation known as the average cost optimality equation. Even if V becomes unbounded the term ρV may still have a limit that is finite. We denote this limit by c and assume it is independent of x. A sufficient condition for the existence of such a c is that $(x_t, r(x_t))$ has an ergodic distribution under the optimal law. In that case c is given by

$$c = \lim_{t \to \infty} E_x \left[q x_t^2 + (r(x_t) - \bar{r})^2 \right]. \tag{3.4}$$

Letting $\rho \downarrow 0$ the HJB equation becomes

$$0 = \min_r \left\{ -c + q x^2 + (r - \bar{r})^2 - (ax + b(r - \bar{r})) v_x + \frac{1}{2} \sigma^2 v_{xx} \right\}, \tag{3.5}$$

which is the average cost optimality equation for the problem. Finding a function $v(x)$ and a scalar $c > 0$ such that (3.5) holds we obtain a candidate solution to the problem, although actually establishing the implied control law as the optimal law for the original problem can be more involved.

Proceeding with the equation (3.5) the optimal r is given by

$$r = \bar{r} + (b/2) v_x.$$

Substituting into (3.5) v should solve

$$0 = -c + qx^2 - (b/2)^2 v_x^2 - axv_x + \frac{1}{2}\sigma^2 v_{xx}. \tag{3.6}$$

A particular solution is given by

$$v(x) = (A/b)\, x^2,$$
$$A = -a/b + \sqrt{(a/b)^2 + q},$$

which works for

$$c = \sigma^2 A/b.$$

The candidate control law is thus given by

$$r(x) = \bar{r} + Ax.$$

For this control law it can be verified that (3.4) holds. It remains however to verify that $r(x)$ is indeed the optimal law, and this causes some problems as we do not have a verification theorem for the limiting case $\rho \downarrow 0$. We do however expect v to be asymptotically quadratic, and one way to establish $r(x)$ as the optimal control law is to calculate the general solution to the differential equation (3.6) and rule out solutions that grow at a rate higher than x^2 as $x \to \pm\infty$. It can be verified that this indeed results in the control law $r(x)$ above. This argument is also pursued below.

3.11.2 The constrained problem

For the constrained problem the average cost optimality equation becomes

$$0 = \min_{r \geq 0} \left\{ -c + qx^2 + (r - \bar{r})^2 - (ax + b(r - \bar{r}))\, v_x + \frac{1}{2}\sigma^2 v_{xx} \right\}, \tag{3.7}$$

where c is a constant, not the same as above. In light of (3.4) we expect c this time to be greater than $\sigma^2 A/b$.

The optimal r is given by

$$r = \max\left\{\bar{r} + (b/2)\, v_x, 0\right\}.$$

Substituting for r we obtain two versions of (3.7),

$$0 = -c + qx^2 - (b/2)^2 v_x^2 - axv_x + \frac{1}{2}\sigma^2 v_{xx}, \tag{3.8}$$

$$0 = -c + qx^2 + \bar{r}^2 - (ax - b\bar{r})\, v_x + \frac{1}{2}\sigma^2 v_{xx}. \tag{3.9}$$

where (3.8) holds when $r = \bar{r} + (b/2)\, v_x$ and (3.9) holds when $r = 0$.

The solution will be such that v_x is monotone, and this implies in particular that there is one value of x, in the sequel denoted x^*, such that (3.8) holds to the right of x^* and (3.9) to the left of x^*.

Starting with Equation (3.8) its general solution can be written as

$$v(x) = (A/b)\,x^2 + \lambda \ln \mathcal{M}\left(z; \frac{1}{2}, \alpha\right) + C_1,$$
$$z = \kappa x^2,$$

where the constants A, λ, α and κ are given by

$$A = -a/b \pm \sqrt{(a/b)^2 + q},$$
$$\lambda = -2\sigma^2/b^2,$$
$$\kappa = b\,(A + a/b)\,/\sigma^2,$$
$$\alpha = \left(c - \sigma^2 A/b\right) / \left(4\sigma^4 \kappa\right).$$

The constant C_1 is an integration constant, which for our purposes is irrelevant. The function $\mathcal{M}$ is a version of the confluent hypergeometric function to which we return shortly. For A we select the positive root since this gives the optimal control law in the absence of the constraint. This implies that κ and α are both positive.

The confluent hypergeometric function has two linearly independent solutions and we follow the notation in Abromowitz and Stegun (1972) denoting them by M and U. With this convention the function M grows at the rate of $\exp\left(z^2\right)$ as $z \to \infty$ while the function U declines at the rate $1/z$ as $z \to \infty$. Because we expect the optimal control law to converge to that of the unconstrained problem as $x \to \infty$, we select the solution corresponding to U as $z \to \infty$. To ensure that $\mathcal{M}$ remains continuous special care must however be taken at $z = 0$. We then obtain

$$\mathcal{M}\left(z; \frac{1}{2}, \alpha\right) = \begin{cases} C_2 M\left(z; \frac{1}{2}, \alpha\right) - U\left(z; \frac{1}{2}, \alpha\right) & z \leq 0 \\ U\left(z; \frac{1}{2}, \alpha\right) & z > 0 \end{cases},$$

with $C_2 = 2\sqrt{\pi}/\Gamma(1/2 + \alpha)$ and with M and U the two independent solutions as in Abromowitz and Stegun.

For the Equation (3.9) we have in terms of v_x a first order linear equation with the general solution

$$v_x = \left(q/a^2\right)(ax + b\bar{r}) + (\bar{c}/a)\sqrt{\pi} \times \exp\left(z^2\right)\{\operatorname{erf}(z) + C_3\},$$
$$z = (ax - b\bar{r})\,/\sqrt{a\sigma^2},$$
$$\bar{c} = c - q\sigma^2/(2a) - \left(1 + q\,(b/a)^2\right)\bar{r}^2,$$

where C_3 is an integration constant. We expect the optimal value function to be asymptotically quadratic as $x \to -\infty$ and because $\operatorname{erf}(z)$ has the asymptotic expansion $-1 - (z\sqrt{\pi})^{-1}\exp\left(-z^2\right)$ as $z \to -\infty$ this requires that we set $C_3 = 1$.

It remains to determine the constant c and the value x^* where the constraint activates. For this value of x we have the two conditions that v_x and v_{xx} should be continuous. Thus we have the two equations

$$v_{x-}(x^*) = v_{x+}(x^*),$$
$$v_{xx-}(x^*) = v_{xx+}(x^*),$$

and the two unknowns x^* and c. Subindices '$-$' and '$+$' indicate limits taken from the left and right. From the latter equation it is possible to solve for c as a function of x^*, and substituting into the former equation this yields an equation for x^* that can be solved numerically.

3.12 Appendix. Smoothing: The unconstrained problem

We proceed as in Appendix 3.11 and form the average cost optimality equation for the problem. This is given by

$$0 = \max_u \left\{ -c + qx^2 + u^2 + uv_r - (ax + br) v_x + \frac{1}{2}\sigma^2 v_{xx} \right\}.$$

The optimal u is given by

$$u = -(1/2) v_r,$$

and inserting v should solve

$$0 = c + qx^2 - (1/4) v_r^2 - (ax + br) v_x + \frac{1}{2}\sigma^2 v_{xx}.$$

A particular solution is given by

$$v(r, x) = -Br^2 - 2Crx + Dx^2$$

with

$$B = a - \sqrt{a^2 + 2b\sqrt{q}}, \qquad C = \frac{1}{2b}B^2, \qquad D = \frac{1}{2b^2}B^2(a - B),$$

which works for

$$c = \sigma^2 D.$$

This gives a candidate control law according to

$$u(r, x) = Br + Cx.$$

We lack a verification theorem to establish u as the solution to the original problem. In the previous appendix this was resolved by considering the general solution for v and then rule out control laws that were not asymptotically quadratic, but here we do not have the general solution for v. It is however possible to solve the problem for the case with positive discounting, i.e. with $\rho > 0$, and for this problem the verification theorem applies. It is then readily verified that the control law for $\rho > 0$ converges to u as $\rho \downarrow 0$.

3.13 Appendix. Smoothing: The deterministic case

Here we present the solution to the control problem in Section 5 in the special case that $\sigma = 0$. Without loss of generality we set $\bar{x} = 0$. Following the Pontryagin minimum principle we form the Hamiltonian

$$H(r, x) = qx^2 + u^2 + \lambda_{1t} u + \lambda_{2t}(-ax - b(r - \bar{r})) + \mu_t \, r,$$

where λ_1 and λ_2 are the multipliers associated with the state variables r_t and x_t, and where μ_t is the multiplier associated with the constraint. The conditions for optimality are

$$\begin{aligned} u_t &= -\frac{1}{2}\lambda_{1t}, \\ \dot{\lambda}_{1t} &= -H_r, \\ \dot{\lambda}_{2t} &= -H_x. \end{aligned}$$

Let $z_t = [(r_t - \bar{r}) \;\; x_t \;\; \lambda_{1t} \;\; \lambda_{2t}]'$. In case the constraint is inactive z_t satisfies

$$\dot{z}_t = \mathbf{M} z_t, \tag{3.10}$$

with

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & -1/2 & 0 \\ -b & -a & 0 & 0 \\ 0 & 0 & 0 & b \\ 0 & -2q & 0 & a \end{bmatrix}.$$

Thus

$$z_t = \exp(\mathbf{M}t)\, z_0,$$

where z_0 is determined by boundary conditions. The initial values give us the two conditions

$$\begin{aligned} \mathbf{e}_1' z_0 &= r_0 - \bar{r}, \\ \mathbf{e}_2' z_0 &= x_0. \end{aligned}$$

In absence of the constraint the two remaining conditions are that the multipliers λ_{1t} and λ_{2t} should vanish as $t \to \infty$,

$$\begin{aligned} \lim_{t\to\infty} \mathbf{e}_3' \exp(\mathbf{M}t)\, z_0 &= 0, \\ \lim_{t\to\infty} \mathbf{e}_4' \exp(\mathbf{M}t)\, z_0 &= 0. \end{aligned}$$

Solving for z_0 solves the problem in the unconstrained case.

If the unconstrained solution results in a path that is admissible in that $r_t \geq 0$ for all t, then we also have the solution to the constrained problem. If however $r_t < 0$ for some $t > 0$ we must find a different path. In this case the solution is characterized by two instances in time, $t = \bar{t}$ where r_t reaches the lower bound for the first time, and $t = t^* > \bar{t}$ where r_t leaves the lower bound. For $t \in (0, \bar{t})$ the state vector z_t evolves according to (3.10) with r_t hitting the zero bound at $t = \bar{t}$. For $t \in (\bar{t}, t^*)$ the short rate r_t remains at zero while x_t stabilizes according to the equation

$$\dot{x}_t = -\left(ax_t - b\bar{r}\right).$$

At $t = t^*$ the target variable x_t reaches the value x^* where the unconstrained solution again becomes admissible, and for $t \in (t^*, \infty)$ the state vector z_t evolves according to the unconstrained solution with initial values $(r, x) = (0, x^*)$.

The problem can now be solved numerically as a one-dimensional minimization problem over $\bar{t}$. It can be shown that the optimal path r_t is such that r_t reaches the zero bound smoothly, i.e. with $\dot{r}_t = 0$. Thus, if the zero bound is reached at $t = \bar{t}$ then z_0 should satisfy

$$\begin{aligned}
\mathbf{e}_1' z_0 &= r_0 - \bar{r}, \\
\mathbf{e}_2' z_0 &= x_0, \\
\mathbf{e}_1' \exp\left(\mathbf{M}\bar{t}\right) z_0 &= -\bar{r}, \\
\mathbf{e}_3' \exp\left(\mathbf{M}\bar{t}\right) z_0 &= 0,
\end{aligned}$$

which is a linear equation for z_0. For a given guess $\bar{t}$ we can calculate the implied value function, and we can proceed numerically to find the optimal $\bar{t}$.

3.14 Appendix. Smoothing: The numerical approximation

This appendix describes the numerical method to approximate the solution to the stochastic control problem in Section 5. Without loss of generality we set $\bar{x} = 0$. We apply a recursive procedure, where we in each step $k = 1, 2, \ldots$ consider the problem

$$V^{[k]}\left(r, x\right) = \min_{u_t \in \mathcal{U}'} L^{[k]}\left(r, x\right), \tag{3.11}$$

with

$$L^{[k]}\left(r, x\right) = E_{r,x}\left[\int_0^{\Delta t} \left(q{x_t}^2 + u_t^2\right) dt + V^{[k-1]}\left(r_{\Delta t}, x_{\Delta t}\right)\right],$$

as the state variables have dynamics according to

$$\begin{cases} dr_t = u_t dt + \sigma_\varepsilon dw_{\varepsilon t}, \\ dx_t = -\left(ax_t + b\left(r_t - \bar{r}\right)\right) dt + \sigma dw_t. \end{cases}$$

The initial value for $V^{[k]}$ is given as

$$V^{[0]}(r,x) = V_{det}(r,x).$$

We construct a grid for r and x on $[r_{\min}, r_{\max}] \times [x_{\min}, x_{\max}]$ equidistant with $\Delta r = 0.0005$ and $\Delta x = 0.0025$. The same values for Δr and Δx were used for all the parameter constellations considered in the main text. The endpoints $r_{\min}$, $r_{\max}$, $x_{\min}$ and $x_{\max}$ should be set such that the state variables' dynamics along the boundaries point *into* the grid. For $r_{\min}$ and $r_{\max}$ this requires respectively that $u(r_{\min}, x) \geq 0$ and $u(r_{\max}, x) \leq 0$ hold for all $x \in [x_{\min}, x_{\max}]$. The constraint gives $r_{\min} = 0$, but $r_{\max}$ will in general depend on parameter settings.

For $x_{\min}$ and $x_{\max}$ it is not possible to ensure drift into the problem for all values of r. But fortunately along the x-axis the problem is less severe because we can utilize the fact that the value function is asymptotically quadratic in x as $x \to \pm\infty$. For all the parameter constellations considered in the main text we used

$$r_{\min} = 0, \;\; r_{\max} = 0.3, \;\; x_{\min} = -0.2, \;\; x_{\max} = 0.1.$$

For a given k and for a given gridpoint (r,x) we approximate the solution to (3.11) by making a local approximation of the value function $V^{[k-1]}$ according to a quadratic function. Locally the problem then assumes a linear-quadratic form with a quadratic endpoint function. The nature of the problem is such that r_t in some regions takes on a very high drift, moving a considerable distance even over the short time interval Δt. We therefore apply the local approximation of $V^{[k-1]}$ not around (r,x) but around a point where r_t and x_t are expected to be an instant Δt later. Defining

$$\mathbf{r}_t^k(r,x) = E_{r,x}[r_t],$$
$$\mathbf{x}_t^k(r,x) = E_{r,x}[x_t],$$

the quadratic approximation of $V^{[k-1]}$ is made around $\left(\mathbf{r}_t^{k-1}(r,x), \mathbf{x}_t^{k-1}(r,x)\right)$.

Although the local problem is linear-quadratic, the standard solution technique involves solving a continuous time matrix Riccati equation, and this requires numerical procedures. Since one such problem has to be solved at each point of the grid, the technique becomes overly time-consuming. We therefore approximate the problem by considering

$$L^{[k]}(r,x) = \int_0^{\Delta t} \left(q{x_t}^2 + u_t^2\right) dt + E_{r,x}\left[V^{[k-1]}\left(\widetilde{r}_{\Delta t}, \widetilde{x}_{\Delta t}\right)\right],$$

and

$$\begin{cases} \dot{r}_t = u_t dt, \\ \dot{x}_t = -\left(a x_t + b\left(r_t - \bar{r}\right)\right), \end{cases}$$

where

$$\begin{pmatrix} \widetilde{r}_{\Delta t} \\ \widetilde{x}_{\Delta t} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} r_{\Delta t} \\ x_{\Delta t} \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \Delta t\right). \tag{3.12}$$

Because the endpoint function is quadratic this problem can be solved using the method of Pontryagin for deterministic problems.

In the region where $x < x^*$ and r is close to zero the quadratic approximation is worse and we then use $V^{[k-1]}(r, x) \approx v(r, x)$ with

$$v(r, x) = g(x) + c_1 \left(r^+\right)^{3/2} + c_2 \left(r^+\right)^2,$$

where $r^+ = \max(r, 0)$, g is a quadratic function and where c_1 and c_2 are constants. The functional form for v is motivated by a local examination of the HJB-equation for the problem in the vicinity of the constraint. This problem is solved numerically with Pontryagin.

After each iteration k the value function still exhibits some irregularities close to the constraint and in this region we therefore apply a smoother. For each $x < x^*$ we use ordinary least squares to fit the function

$$h(r) = \sum_{n=1}^{4} c_n \left(r^+\right)^{(n+2)/2},$$

to $V^{[k]}(x, r)$ for r close to the constraint. In the estimation in the main text this was done for $x < x^*$ and $r \in [0, 0.075)$. The function $h(r)$ was used to replace $V^{[k]}$ for $r \in [0, 0.0325)$ while linear interpolation between h and $V^{[k]}$ was used for $r \in [0.0325, 0.075)$.

3.15 Appendix. Bond prices in the case of stabilization

This appendix presents the numerical approximation of bond prices in the case of stabilization. Fixing a small time step Δt bond prices are calculated recursively according to

$$\begin{aligned} F(x; 0) &= 1, \\ F(x; k\Delta t) &= \exp\left(-\int_0^{\Delta t} \mathbf{r}_s(x)\, ds\right) E_x\left[F\left(x_{\Delta t}, (k-1)\Delta t\right)\right], \\ \mathbf{r}_t(x) &\equiv E_x\left[r\left(x_t\right)\right], \end{aligned}$$

for $k = 1, 2, \ldots$ We begin in Section 3.15.1 by proving that the procedure converges to the true bond prices as $\Delta t \downarrow 0$. This is then followed by the numerical computation of $\mathbf{r}_t(x)$ in Section 3.15.2 and we finish with the numerical approximation of the expectation of F in Section 3.15.3.

3.15.1 Convergence to true prices

For a given maturity τ the true bond price satisfies

$$f(x,\tau) = E_x\left[\exp\left(-\int_0^\tau r_s ds\right)\right],$$

and we would like to show that $D_{\Delta t}$ as defined by

$$D_{\Delta t}(x,\tau) = f(x,\tau) - F(x,\tau)$$

approaches zero as $\Delta t \downarrow 0$ for all τ and x. We can write D as

$$\begin{aligned} D_{\Delta t}(x,\tau) = \exp&\left(-\int_0^{\Delta t} \mathbf{r}_s ds\right) \times \\ E_x&\left[\left\{\exp\left(-\int_0^{\Delta t}(r_s - \mathbf{r}_s)\,ds\right) - 1\right\}\exp\left(-\int_{\Delta t}^\tau r_s ds\right)\right. \\ &\left. + \left\{\exp\left(-\int_{\Delta t}^\tau r_s ds\right) - F(x_{\Delta t}, \tau - \Delta t)\right\}\right]. \end{aligned}$$

where the dependence on x for $\mathbf{r}_s$ has been suppressed. By the triangle inequality we have

$$\begin{aligned} D^2_{\Delta t}(x,\tau) \le 2E_x&\left\{\left[\exp\left(-\int_0^{\Delta t}(r_s - \mathbf{r}_s)\,ds\right) - 1\right]\right\}^2 \\ &+2\left\{E_x\left[\exp\left(-\int_{\Delta t}^\tau r_s ds\right) - F(x_{\Delta t}, \tau - \Delta t)\right]\right\}^2, \end{aligned}$$

where we have used the fact that $r_t \ge 0$. For the first term we have

$$\left\{E_x\left[\exp\left(-\int_0^{\Delta t}(r_s - \mathbf{r}_s)\,ds\right) - 1\right]\right\}^2 = o(\Delta t)^2.$$

where $o(\Delta t)^2 / (\Delta t)^2 \to 0$ as $\Delta t \downarrow 0$. For the second term we have

$$\left\{E_x\left[\exp\left(-\int_{\Delta t}^\tau r_s ds\right) - P(x_{\Delta t}, \tau - \Delta t)\right]\right\}^2 \le E_x\left[D^2_{\Delta t}(x_{\Delta t}, \tau - \Delta t)\right].$$

Thus

$$\begin{aligned} D^2_{\Delta t}(x,\tau) &\le o(\Delta t)^2 + E_x\left[D^2_{\Delta t}(x_{\Delta t}, \tau - \Delta t)\right] \\ &\le o(\Delta t)^2 + E_x\left[o(\Delta t)^2 + E_x\left[D^2_{\Delta t}(x_{\Delta t}, \tau - 2\Delta t)\right]\right] \end{aligned}$$

and if Δt is an even fraction of τ we obtain

$$\begin{aligned} D^2_{\Delta t}(x,\tau) &\le E_x\left[o(\Delta t)^2 \times \tau/\Delta t\right] \\ &= E_x\left[o(\Delta t)^2/\Delta t\right] \end{aligned}$$

and the term inside the expectation converges to zero pointwise.

3.15.2 Approximation of $\mathbf{r}_t(x)$

By a generalization of Itô's rule (see e.g. Karatzas and Shreve (1988)) $r(x_t)$ can be written as

$$r(x_t) = r(x_0) + \int_0^t r'_-(x_s)\,dx_s + \int_{-\infty}^{\infty} \Lambda_t(a)\,r''(da)$$

where $\Lambda_t(a)$ is the local time of r_t in the vicinity of a, r'_- is the left limit of the derivative and r'' is the second derivative measure. Making a local approximation of r according to

$$u(x) = \max(c_0 + c_1 x, 0)$$

we obtain

$$E_{0,x}[r(x_t)] = r(x_0) + E_{0,x}\left[\int_0^t 1_{(\bar{x},\infty)}(x_s)\,c_1 dx_s\right] + c_1 E_{0,x}[\Lambda_t(\bar{x})], \quad (3.13)$$

where $\bar{x} = -c_0/c_1$. To approximate the expectations we make a linear approximation of the dynamics for x_t. If

$$dx_t = (\mu_0 - \mu_1 x_t)\,dt + \sigma dw_t$$

for constants μ_0 and μ_1 we have

$$\begin{aligned} x_t &\sim \mathcal{N}\left(m_t(x), v_t^2\right), \\ m_t &= (x - \mu_0/\mu_1)\exp(-\mu_1 t) + \mu_0/\mu_1, \\ v_t^2 &= \sigma^2\{1 - \exp(-2\mu_1 t)\}/2\mu_1. \end{aligned}$$

This yields for the former expectation in (3.13) that

$$\begin{aligned} E_{0,x}\left[\int_0^t 1_{(\bar{x},\infty)}(x_s)\,c_1 dx_s\right] &= E_{0,x}\left[\int_0^t 1_{(\bar{x},\infty)}(x_s)\,c_1(\mu_0 - \mu_1 x_s)\,ds\right] \\ &= c_1\mu_0 t - c_1\mu_1\int_0^t E_{0,x}\left[x_s^+\right] ds, \end{aligned}$$

and using the normal approximation for x_t the expectation inside the integral can be calculated analytically. For the latter expectation in (3.13) we have, again using the normal approximation for x_t, that

$$E_{x,0}[\Lambda_t(\bar{x})] = \frac{\sigma^2}{2}\int_0^t \varphi\left(\bar{x}; m_s(x), v_s^2\right) ds$$

where $\varphi\left(x; m, v^2\right)$ is the density function for a normal variable with mean m and variance v^2. Concluding we have

$$\mathbf{r}_t(x) \approx r(x) + c_1\left(\mu_0 t - \mu_1\int_0^t E_{0,x}\left[x_s^+\right] ds\right) + \frac{\sigma^2}{2}\int_0^t \varphi\left(\bar{x}; m_s(x), v_s^2\right) ds.$$

The integrals are computed numerically.

3.15.3 Approximation of bond prices

For the expectation of future bond prices we make a local approximation of $\ln F$ according to

$$\ln F\left(x;(k-1)\,\tau\right) \approx c_0 + c_1 x + c_2 x^2$$

and with the normal approximation for x_t we can use

$$E_x\left[F\left(x_{\Delta t},(k-1)\,\Delta t\right)\right] \approx \int_{-\infty}^{\infty} \exp\left(c_0 + c_1\xi + c_2\xi^2\right)\varphi\left(\xi; m_t\left(x\right), v_t^2\right) d\xi.$$

The integral can be calculated analytically.

3.16 Appendix. Supplementary graphs

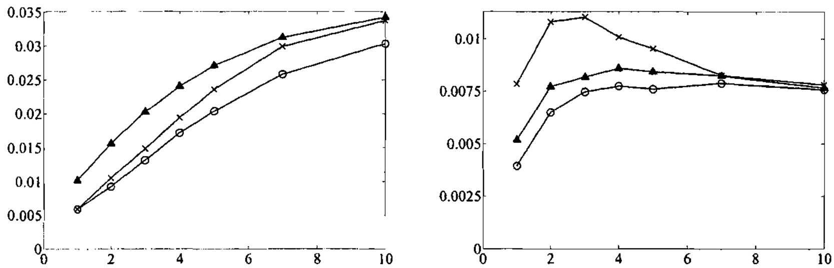

Figure 3.24.a: Swap rate curves (left) and swap rate volatility curves (right) in Jan-1996 (crosses), Jul-96 (triangles) and Jan-97 (circles). Swap rates refer to the beginning of the indicated month. Volatility curves depict the standard deviation in the swap rate for a given maturity over the six months following the indicated date.

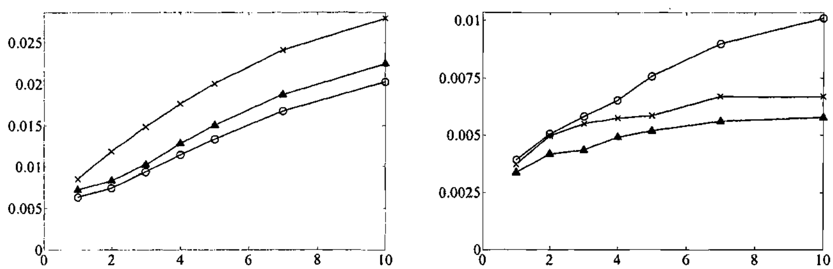

Figure 3.24.b: Corresponding item to Figure 3.24.a for the dates Jul-1997- (crosses), Jan-98 (triangles) and Jul-98 (circles).

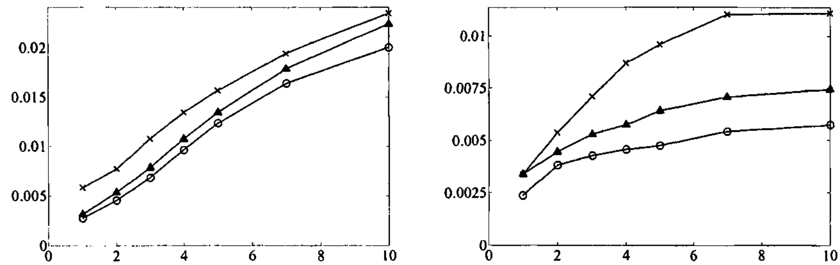

Figure 3.24.c: Corresponding item to Figure 3.24.a for the dates Jan-99- (crosses), Jul-99 (triangles) and Jan-00 (circles).

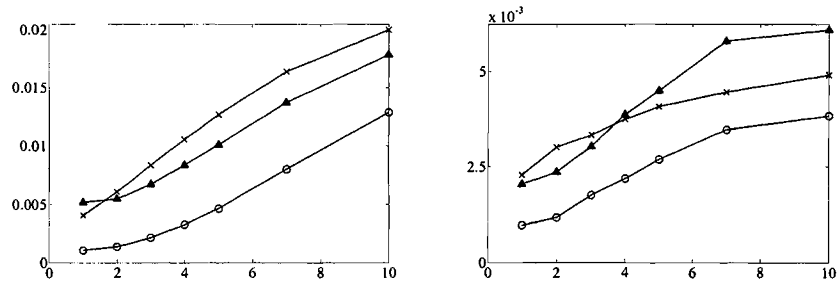

Figure 3.24.d: Corresponding item to Figure 3.24.a for the dates Jul-00- (crosses), Jan-01 (triangles) and Jul-01 (circles).

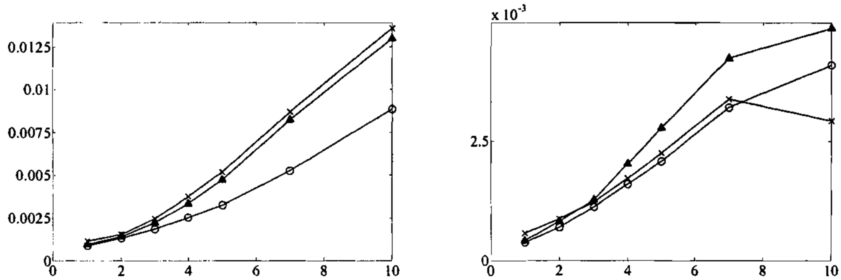

Figure 3.24.e: Corresponding item to Figure 3.24.a for the dates Jan-02- (crosses), Jul-02 (triangles) and Jan-03 (circles).

References

1. Abramowitz, Milton and Irene A. Stegun, 1972, "Handbook of mathematical functions with formulas, graphs and mathematical tables," Government Printing Office.
2. Ang, Andrew and Monika Piazzesi, 2003, "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables," *Journal of Monetary Economics* 50(4), 745-787.
3. Arapostathis, Aristotle, Vivek S. Borkar, Emmanuel Fernández-Gaucherand, Mrinal K. Gosh, and Steven I. Marcus, 1993, "Discrete-time controlled Markov processes with average cost criterion: A survey," *SIAM Journal of Control and Optimization*, 31(2), 282-344.
4. Black, Fisher, 1995, "Interest rates as options," *Journal of Finance* 50(7), 1371-1376.
5. Black, F. and P Karasinski, 1991, "Bond and option pricing when short rates are log-normal," *Financial Analysts Journal* 47, 52-59.
6. Bomfim, Antulio, 2003, "Monetary policy and the yield curve," manuscript, Finance and Economics Discussion Series, Federal Reseve Board.
7. Bomfim, Antulio, 2003, "Interest rates as options," manuscript, Finance and Economics Discussion Series, Federal Reserve Board.
8. Braun, R. Anton and Etsuro Shioji, 2003, "Monetary policy and the term structure of interest rates in Japan ," manuscript, Center of International Research on the Japanese Economy.
9. Cadenillas, Abel and Fernando Zapatero, 2000, "Classical and impulse stochastic control of the exchange rate using interest rates and reserves," *Mathematical Finance* 10(2), 141-156.
10. Clarida, Richard, Jordi Galí and Mark Gertler, 1999, "The science of monetary policy: A new Keynesian perspective," *Journal of Economic Literature* 37, 1661-1707.
11. Cox, John C, Jonathan E. Jr. Ingersoll and Stephen A. Ross, 1985, "A theory of the term structure of interest rates," *Econometrica*, March 53(2), 385-407.
12. Dewachter, Hans and Marco Lyrio, 2002, "Macro factors and the term structure of interest rates," manuscript, Catholic University of Leuven.
13. Diebold, Francis X., Glenn D. Rudebusch and S. Boragan Aruoba, 2003, "The macroeconomy and the yield curve: A nonstructural analysis," manuscript, Federal Reserve Bank of San Fransisco.
14. Eggertson, Gauti and Michael Woodford, 2003, "The zero bound on interest rates and optimal monetary policy," *Conference paper*, Brookings Panel on Economic Activity.
15. Ellingsen, Tore and Ulf Söderström, 2001, "Monetary policy and market interest rates," *American Economic Review* 91(5), 1594-1607.
16. Ellingsen, Tore and Ulf Söderström, 2003, "Why are long rates sensitive to monetary policy?" manuscript.

17. Estrella, Arturo, 2003, "Why does the yield curve predict output and inflation?" manuscript, Federal Reserve Bank of New York.
18. Evans, Charles L. and David Marshall, 2002, "Economic determinants of the nominal treasury yield curve," manuscript, Federal Reserve Bank of Chicago.
19. Favero, Carlo A., 2002, "Taylor rules and the term structure," manuscript, Universitá Bocconi.
20. Fuhrer, Jeffrey and Brian F. Madigan, 1997, "Monetary policy when interest rates are bounded at zero," *Review of Economics and Statistics*, November, 573-585.
21. Godstein, Robert S. and William P. Keirstad, 1997, "On the term structure of interest rates in the presence of reflecting and absorbing boundaries," *Working Paper*.
22. Goodman, Jonathan, Kyoung-Sook Moon, Anders Szepessy, Raul Tempone and Georgios Zouraris, 2002, "Stochastic and partial differential equations with adapted numerics," *Lecture Notes*, Royal Institute of Technology, Stockholm.
23. Hördahl, Peter, Oreste Tristani and David Vestin, 2002, "A joint econometric model of macroeconomic and term structure dynamics," manuscript, European Central Bank.
24. Kato, Ryo and Shinichi Nishiyama, 2001, "Optimal monetary policy when interest rates are bounded at zero," *Working Paper* #01-12, August, The Ohio State University.
25. Karatzas, Ioannis and Steven E. Shreve, 1988, "Brownian Motion and Stochastic Calculus," Springer-Verlag New York Inc.
26. Kocicki, Sharon and Peter Tinsley, 2001, "Shifting endpoints in the term structure of interest rates," *Journal of Monetary Economics* 47, 613-652.
27. Kotlán, Viktor, 2001, "Monetary policy and the term structure of interest rates in a small open economy – a model framework approach," manuscript, Czech National Bank.
28. Marumo, Kohei, Takashi Nakayama, Shinichi Nishioka and Toshihiro Yoshida, 2003, "Extracting market expectations on the duration of the zero interest rate policy from Japan's bond prices," manuscript, Bank of Japan.
29. Nagayasu, Yun, 2003, "The term structure of interest rates and monetary policy during a zero-interest rate period," manuscript, International Monetary Fund.
30. Nelson, C., and A. Siegel, 1987, "Parsimonious modeling of yield curves," *Journal of Business* 60, 473-489.
31. Orphanides, Anthanasios and Volker Wieland, 1999, "Efficient monetary policy design near price stability," *Finance and Economics Discussion Series*, 1999-67, Board of Governors Federal Reserve System.
32. Piazzesi, Monika, 2003, "Bond yields and the Federal Reserve," manuscript, *Journal of Political Economy*, forthcoming.
33. Pelsser, Antoon, 2000, "Efficient Methods for Valuing Interest Rate Derivatives," Springer Finance.
34. Rudebusch, Glenn D., 2001, "Is the Fed too timid? Monetary policy in an uncertain world," *Review of Economics and Statistics* 83(2), 203-217.
35. Rudebusch, Glenn D. and Tao Wu, 2003, "A macro-finance model of the term structure, monetary policy, and the economy," manuscript, Federal Reserve Bank of San Francisco.
36. Ruge-Murcia, Francisco J., 2002, "Some implications of the zero lower bound on interest rates for the term structure and monetary policy," manuscript, Université de Montréal.
37. Skallsjö. Sven, 2003, "Monetary policy and the yield curve," manuscript, Stockholm School of Economics.

38. Svensson, Lars E.O., 1997, "Inflation forecast targeting: Implementing and monitoring inflation targets," *European Economic Review*, 41, 1111-1146.
39. Svensson, Lars E. O., 2000, "Open economy inflation targeting," *Journal of International Economics* 50 (February), 155-183.
40. Söderström, Ulf, Paul Söderlind and Anders Vredin, 2003, "New-Keynesian models and monetary policy: A reexamination of the stylized facts," manuscript.
41. Taylor, John B., 1993, "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy* 39 (December), 195-214.
42. Wu, Tao, 2001, "Macro factors and affine term structure models," manuscript, Federal Reserve Bank of San Francisco.

EFI
The Economic Research Institute

Reports since 1999
A complete publication list can be found at www.hhs.se/efi

Published in the language indicated by the title

2004
Larsson, P., Förändringens villkor. En studie av organisatoriskt lärande och förändring inom skolan.
Salabasis, M., Bayesian Time Series and Panel Models – Unit Roots, Dynamics and Random Effects.
Söderström, J., Från Produkt till Tjänst. Utveckling av affärs- och miljöstrategier i produktorienterade företag.

2003
Andersson, H., Valuation and Hedging of Long-Term Asset-Linked Contracts.
Bergman, M., Essays on Human Capital and Wage Formation.
Damsgaard, N., Deregulation and Regulation of Electricity Markets.
Eklund, B., Four Contributions to Statistical Inference in Econometrics.
Hakkala, K., Essays on Restructuring and Production Decisions in Multi-Plant Firms.
Holgersson, C., Rekrytering av företagsledare. En studie i hoimosocialitet.
Ivaschenko, I, Essays on Corporate Risk, U.S. Business Cycles, International Spillovers of Stock Returns, and Dual Listing.
Lange, F., Brand Choice in Goal-derived Categories - What are the Determinants?
Le Coq, C., Quantity Choices and Market Power in Electricity Market.
Magnusson, P.R., Customer-Oriented Product Development - Experiments Involving Users in Service Innovation.
Meisiek, S., Beyond the Emotional Work Event Social Sharing of Emotion in Organizations.
Mårtensson, A.. Managing Mission-Critical IT in the Financial Industry.
Nilsson, G., Processorientering och styrning – Regler, mål eller värderingar?
Sandberg, R., Corporate Consulting for Customer Solutions Bridging Diverging Business Logics.
Sturluson, J.T., Topics in the Industrial Organization of Electricity Markets.
Tillberg, U., Ledarskap och samarbete – En jämförande fallstudie i tre skolor.
Waldenström, D., Essays in Historical Finance.
Wallén, U., Effektivitet i grundskolan i anslutning till en stadsdelsnämndsreform.
Ögren, A., Empirical Studies in Money, Credit and Banking - The Swedish Credit Market in Transition under the Silver and the Gold Standards, 1834 – 1913.

2002
Barinaga, E., Levelling Vagueness – A study of cultural diversity in an international project group.
Berglund, J., De otillräckliga - En studie av personalspecialisternas kamp för erkännande och status.
Bolander, P., Anställningsbilder och rekryteringsbeslut.
Damjanovic, T., Essays in Public Finance.

Ekman, M., Studies in Health Economics – Modelling and Data Analysis of Costs and Survival.
Företagerskan – Om kvinnor och entreprenörskap. Holmquist, C. och Sundin, E (red)
Heyman, F., Empirical Studies on Wages, Firm Performance and Job Turnover.
Kallifatides, M., Modern företagsledning och omoderna företagsledare.
Kaplan, M., Acquisition of Electronic Commerce Capability - The Cases of Compaq and Dell in Sweden.
Mähring, M., IT Project Governance.
Nilsson, M., Essays in Empirical Corporate Finance and Governance.
Rekrytering av koncernstyrelsen – Nomineringsförfaranden och styrelsesammansättning med focus på kvinnors ställning och möjligheter. Sjöstrand, S-E. och Petrelius, P.,(red)
Scener ur ett företag – Organiseringsteori för kunskapssamhället. Löwstedt, J. Stymne, B.,(red).
Schenkel, A., Communities of Practice or Communities of Discipline - Managing Deviations at the Øresund Bridge.
Schuster, W., Företagets Valutarisk – En studie av horisontella och vertikala styrprocesser.
Skogsvik, S., Redovisningsmått, värderelevans och informationseffektivitet.
Sundén, D., The Dynamics of Pension Reform.
Ternström, I., The Management of Common-Pool Resources - Theoretical Essays and Empirical Evidence.
Tullberg, J., Reciprocitet – Etiska normer och praktiskt samarbete.
Westling, G., Balancing Innovation and Control – The Role of Face-to-face Meetings in Complex Product Development Projects.
Viklund, M., Risk Policy – Trust, Risk Perception, and Attitudes.
Vlachos, J., Risk Matters - Studies in Finance, Trade and Politics.

2001
Adolfson, M., Monetary Policy and Exchange Rates – Breakthrough of Pass-Through.
Andersson, P., Expertise in Credit Granting: Studies on Judgment and Decision-Making behavior.
Björklund, C., Work Motivation - Studies of its Determinants and Outcomes.
Center for Management and Organization 50 (1951-2001).
Charpentier, C., Uppföljning av kultur- och fritidsförvaltningen efter stadsdelsnämnds-reformen.
Dahlén, M., Marketing on the Web - Empirical Studies of Advertising and Promotion Effectiveness.
Eckerlund, I., Essays on the Economics of Medical Practice Variations.
Ekelund, M., Competition and Innovation in the Swedish Pharmaceutical Market.
Engström, S., Success Factors in Asset Management.
Ericsson, D., Kreativitetsmysteriet – Ledtrådar till arbetslivets kreativisering och skrivandets metafysik.
Eriksson, R., Price Responses to Changes in Costs and Demand.
Frisell, L., Information and Politics.
Giordani, P., Essays in Monetary Economics and Applied Econometrics.
Gustavsson, P., Essays on Trade, Growth and Applied Econometrics.
Hedlund, A., Konsumentens erfarenhet – och dess inverkan på livsmedelsinköp på Internet.

Hill, M., Essays on Environmental Policy Analysis: Computable General Equilibrium Approaches Applied to Sweden.
Hvenmark, J., Varför slocknar elden? Om utbrändhet bland chefer i ideella organisationer.
Hägglund, P.B., Företaget som investeringsobjekt – Hur placerare och analytiker arbetar med att ta fram ett investeringsobjekt.
Höök, P., Stridspiloter i vida kjolar, om ledarutveckling och jämställdhet.
Johansson, C., Styrning för samordning.
Josephson, J., Evolution and Learning in Games.
Kjellberg, H., Organising Distribution - Hakonbolaget and the efforts to rationalise food distribution, 1940-1960.
Lange, F. och Wahlund, R., Category Management – När konsumenten är manager.
Liljenberg, A., Customer-geared competition – A socio-Austrian explanation of Tertius Gaudens.
Lindkvist, B., Kunskapsöverföring mellan produktutvecklingsprojekt.
Ljunggren, U., Nyckeltal i grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen.
Läkemedel – Kostnad eller resurs för sjukvården? Jönsson, B.,(red).
Löf, M., On Seasonality and Cointegration.
Martensen, K., Essays on Entry Externalities and Market Segmentation.
Matros, A., Stochastic Stability and Equilibrium Selection in Games.
Mårtensson, P., Management Processes – An Information Perspective on Managerial Work.
Nilsson, A., Market Transparency.
Norberg, P., Finansmarknadens amoralitet och det kalvinska kyrkorummet – En studie i ekonomisk mentalitet och etik.
Persson, B., Essays on Altruism and Health Care Markets.
Rech, G., Modelling and Forecasting Economic Time Series with Single Hidden-layer Feedforward Autoregressive Artificial Neural Networks.
Skoglund, J., Essays on Random Effects Models and GARCH.
Strand, N., Empirical Studies of Pricing.
Thorén, B., Stadsdelsnämndsreformen och det ekonomiska styrsystemet - Om budgetavvikelser.

2000

Berg-Suurwee, U., Styrning före och efter stadsdelsnämndsreform inom kultur och fritid – Resultat från intervjuer och enkät.
Bergkvist, L., Advertising Effectiveness Measurement: Intermediate Constructs and Measures.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Koncernchefen och ägarna.
Bornefalk, A., Essays on Social Conflict and Reform.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform.
Edman, J., Information Use and Decision Making in Groups.
Emling, E., Svenskt familjeföretagande.
Ericson, M., Strategi, kalkyl, känsla.
Gunnarsson, J., Wahlund, R., Flink, H., Finansiella strategier i förändring: segment och beteenden bland svenska hushåll.
Hellman, N., Investor Behaviour – An Empirical Study of How Large Swedish Institutional Investors Make Equity Investment Decisions.
Hyll, M., Essays on the Term Structure of Interest Rates.

Håkansson, P., Beyond Private Label – The Strategic View on Distributor Own Brands.
I huvudet på kunden. Söderlund, M (red). *EFI och Liber Förlag.*
Karlsson Stider, A., Familjen och firman.
Ljunggren, U., Styrning av grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen – Resultat från intervjuer och enkät.
Ludvigsen, J., The International Networking between European Logistical Operators.
Nittmar, H., Produktutveckling i samarbete – Strukturförändring vid införande av nya Informationssystem.
Robertsson, G., International Portfolio Choice and Trading Behavior.
Schwarz, B., Weinberg, S., Serviceproduktion och kostnader – att söka orsaker till kommunala skillnader.
Stenström, E., Konstiga företag.
Styrning av team och processer – Teoretiska perspektiv och fallstudier. Bengtsson, L., Lind, J., Samuelson, L.A., (red).
Sweet, S., Industrial Change Towards Environmental Sustainability – The Case of Replacing Chloroflouorocarbons.
Tamm Hallström, K., Kampen för auktoritet – standardiseringsorganisationer i arbete.

1999
Adler, N., Managing Complex Product Development.
Allgulin, M., Supervision and Monetary Incentives.
Andersson, P., Experto Credite: Three Papers on Experienced Decision Makers.
Ekman, G., Från text till batong – Om poliser, busar och svennar.
Eliasson, A-C., Smooth Transitions in Macroeconomic Relationships.
Flink, H., Gunnarsson, J., Wahlund, R., Svenska hushållens sparande och skuldsättning– ett konsumentbeteende-perspektiv.
Gunnarsson, J., Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
Hamrefors, S., Spontaneous Environmental Scanning.
Helgesson, C-F., Making a Natural Monopoly: The Configuration of a Techno-Economic Order in Swedish Telecommunications.
Japanese Production Management in Sunrise or Sunset. Karlsson, C., (red).
Jönsson, B., Jönsson, L., Kobelt, G., Modelling Disease Progression and the Effect of Treatment in Secondary Progressive MS. Research Report.
Lindé, J., Essays on the Effects of Fiscal and Monetary Policy.
Ljunggren, U., Indikatorer i grundskolan i Stockholms stad före stadsdelsnämndsreformen – en kartläggning.
Ljunggren, U., En utvärdering av metoder för att mäta produktivitet och effektivitet i skolan – Med tillämpning i Stockholms stads grundskolor.
Lundbergh, S., Modelling Economic High-Frequency Time Series.
Mägi, A., Store Loyalty? An Empirical Study of Grocery Shopping.
Mölleryd, B.G., Entrepreneurship in Technological Systems – the Development of Mobile Telephony in Sweden.
Nilsson, K., Ledtider för ledningsinformation.
Osynlig Företagsledning. Sjöstrand, S-E., Sandberg, J., Tyrstrup, M., (red).
Rognes, J., Telecommuting – Organisational Impact of Home Based – Telecommuting.
Sandström, M., Evaluating the Benefits and Effectiveness of Public Policy.
Skalin, J., Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.
Spagnolo, G., Essays on Managerial Incentives and Product-Market Competition.

Strauss, T., Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.
Svedberg Nilsson, K., Effektiva företag? En studie av hur privatiserade organisationer konstrueras.
Söderström, U., Monetary Policy under Uncertainty.
Werr, A., The Language of Change The Roles of Methods in the Work of Management Consultants.
Wijkström, F., Svenskt organisationsliv – Framväxten av en ideell sektor.