

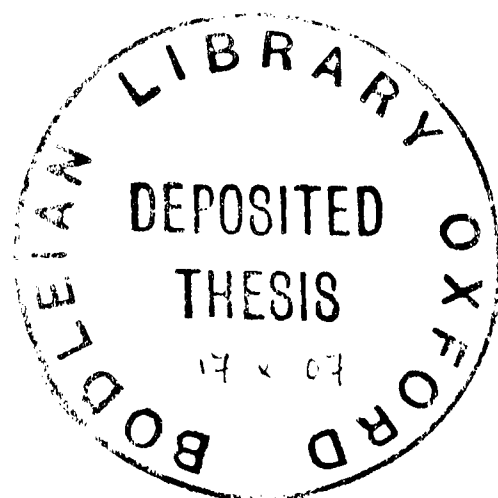
# **Topics in the Economics of Money Substitutes in Developing and Transition Countries**

Felix Martin  
University College, Oxford

A thesis submitted for the degree of Doctor of Philosophy  
in the University of Oxford.

January, 2006

Approximate word count: 85,000



# TOPICS IN THE ECONOMICS OF MONEY SUBSTITUTES IN DEVELOPING AND TRANSITION COUNTRIES

Felix Martin, University College, Oxford

A thesis submitted for the degree of Doctor of Philosophy in Michaelmas, 2005.

## ABSTRACT

Recent research has shown that money substitutes – whether in the form of foreign currency or of more exotic instruments such as privately-issued moneys – are common in developing and transition countries, and have important consequences for macroeconomic and financial sector policy. The aim of this thesis is to advance our theoretical and empirical understanding of the determinants of money substitution in developing and transition economies.

We begin in Chapter 1 by addressing the need for a general theoretical framework for the analysis of money substitutes. Reviewing both the classical and the modern theoretical literature on money, we conclude that the Credit theory of money – an ancient but until recently neglected theory which conceives of money as a unilateral financial contract between its issuer and its bearer – is a useful framework for such analysis.

In Chapter 2, we undertake an empirical analysis of non-cash settlements (NCS) in Croatia. Using time series econometric analysis, we demonstrate that the instruments used to settle NCS are at least in part substitutes for the national currency, created endogenously by the enterprise sector in response to constraints on their participation in the official monetary and banking system.

We turn to the most important form of money substitute in developing and transition countries – foreign currency – in Chapter 3, where we present a new review of the theoretical and empirical literature on dollarisation. In particular, we track the evolution of theoretical models of dollarisation in response to the increasing empirical importance of financial dollarisation relative to currency substitution.

In Chapter 4 we undertake an empirical study of the determinants of deposit dollarisation in the two transition economies of Estonia and Lithuania by building and interpreting dynamic, multiple equation, econometric models. We find that a simple, portfolio theoretic account of the dollarisation process furnishes a good explanation, but also that data availability limits the level of analytical detail that this approach can attain.

## **Acknowledgements**

I would like to thank my supervisor, Dr. Antoni Chawluk of Mansfield College, Oxford, for his guidance in all aspects of this thesis.

Chapter 2 was written whilst I was a visiting researcher at the Croatian National Bank in the summer of 2004, funded by an ESRC Overseas Institutional Visit grant. I would like to thank Dr. Evan Kraft, Director of Research at the Croatian National Bank, for facilitating this visit; and him, Maroje Lang, Vedran Šošić, and Velimir Šonje for their invaluable advice on my research. Thanks are also due to Dr. Kraft and to Dr. Timothy Lane of Balliol College, Oxford for generously providing comments on a draft of the chapter.

Chapter 4 benefitted greatly from discussions with Igor Vetlov of the Central Bank of Lithuania; as well as from comments made by participants at a Gorman Seminar in Oxford.

Needless to say, no one mentioned above should be implicated in any of my analysis or conclusions – all of which, for better or for worse, are my own.

## Contents

<b>Introduction</b>	<b>1</b>
<b>1. <i>Non Aes Sed Fides</i>: the Credit Theory of Money and its Implications for the Analysis of Money Substitutes</b>	<b>7</b>
1.1 Introduction .....	9
1.2 The Two Theories of Money .....	10
1.2.1 The Metallist Theory of Money .....	10
1.2.2 The Credit Theory of Money.....	15
1.3 Recent Models of the Micro-foundations of Money and Barter .....	23
1.3.1 Models in the Metallist Tradition .....	24
1.3.2 Models in the Credit Tradition .....	39
1.4 Towards a Credit Theory of Money Substitutes .....	51
1.4.1 The Metallist Theory of Money and the Analysis of Money Substitutes .....	52
1.4.2 The Credit Theory of Money and the Analysis of Money Substitutes .....	55
1.5 Conclusions .....	65
<b>2. Non-Cash Settlements in Transition: the Case of Croatia</b>	<b>69</b>
2.1 Introduction.....	71
2.2 Theory .....	75
2.3 Institutional Analysis .....	79
2.3.1 The primary market for credit .....	79
2.3.2 The secondary market for credit .....	86
2.4 Econometric Analysis .....	88
2.4.1 Testing H(1) and H(2) .....	88
2.4.2 Testing H(3) .....	94
2.5 Conclusions .....	97
Figures.....	99
Tables.....	113

### **3. Dollarisation in Developing and Transition Countries: a Review of the Theoretical and Empirical Literature** **121**

3.1 Introduction .....	123
3.2 A Taxonomy of Dollarisation .....	124
3.3 The Dimensions of Dollarisation .....	129
3.3.1 Financial dollarisation .....	130
3.3.2 Real dollarisation .....	133
3.3.3 Currency substitution .....	134
3.4 Theories of Dollarisation .....	137
3.4.1 Structural factors influencing dollarisation .....	139
3.4.2 Transactions dollarisation .....	143
3.4.3 Store of value dollarisation .....	151
3.5 Policy Responses to Dollarisation .....	168
3.5.1 The case in favour of a policy response to dollarisation.....	169
3.5.2 Policy options to manage dollarisation .....	176
3.6 Conclusions.....	186

### **4. Deposit Dollarisation in the Baltics: a Portfolio Balance Approach**

**191**

4.1 Introduction .....	193
4.2 Introduction to the Empirical Cases.....	195
4.2.1 Estonia and Lithuania: Currency Reform and Stabilisation Prior to the Introduction of the CBAs .....	195
4.2.2 Estonia and Lithuania: Economic and Political Developments after the Introduction of the CBAs.....	197
4.2.3 The Development of Deposit Interest Rates .....	201
4.3 Empirical Strategy .....	203
4.3.1 Choice of variables .....	203
4.3.2 Time Series Characteristics .....	205
4.3.3 Treatment of expectations.....	206
4.3.4 Modelling strategy .....	208
4.4 Results .....	210
4.4.1 Lithuania .....	210
4.4.2 Estonia .....	213
4.5 Interpretation.....	216

4.5.1 Vector versus Single Equation Methods, and Endogeneity of Variables	216
4.5.2 Empirical Application of Portfolio Balance Theory to Deposit Dollarisation	218
4.5.3 Interpretation of Numerical Estimates	221
4.6 Conclusions	225
Appendix A: Construction of Data Series	227
Appendix B: VECM models for Lithuania and Estonia	233
Appendix C: A two Currency Portfolio Balance Model of Deposit Dollarisation	238
Tables	241
Figures	253

<b>Bibliography</b>	<b>271</b>
---------------------	------------

# **Topics in the Economics of Money Substitutes in Developing and Transition Countries**

Felix Martin  
University College, Oxford

A thesis submitted for the degree of Doctor of Philosophy  
in the University of Oxford.

January, 2006

Approximate word count: 85,000

# TOPICS IN THE ECONOMICS OF MONEY SUBSTITUTES IN DEVELOPING AND TRANSITION COUNTRIES

Felix Martin, University College, Oxford

A thesis submitted for the degree of Doctor of Philosophy in Michaelmas, 2005.

## ABSTRACT

Recent research has shown that money substitutes – whether in the form of foreign currency or of more exotic instruments such as privately-issued moneys – are common in developing and transition countries, and have important consequences for macroeconomic and financial sector policy. The aim of this thesis is to advance our theoretical and empirical understanding of the determinants of money substitution in developing and transition economies.

We begin in Chapter 1 by addressing the need for a general theoretical framework for the analysis of money substitutes. Reviewing both the classical and the modern theoretical literature on money, we conclude that the Credit theory of money – an ancient but until recently neglected theory which conceives of money as a unilateral financial contract between its issuer and its bearer – is a useful framework for such analysis.

In Chapter 2, we undertake an empirical analysis of non-cash settlements (NCS) in Croatia. Using time series econometric analysis, we demonstrate that the instruments used to settle NCS are at least in part substitutes for the national currency, created endogenously by the enterprise sector in response to constraints on their participation in the official monetary and banking system.

We turn to the most important form of money substitute in developing and transition countries – foreign currency – in Chapter 3, where we present a new review of the theoretical and empirical literature on dollarisation. In particular, we track the evolution of theoretical models of dollarisation in response to the increasing empirical importance of financial dollarisation relative to currency substitution.

In Chapter 4 we undertake an empirical study of the determinants of deposit dollarisation in the two transition economies of Estonia and Lithuania by building and interpreting dynamic, multiple equation, econometric models. We find that a simple, portfolio theoretic account of the dollarisation process furnishes a good explanation, but also that data availability limits the level of analytical detail that this approach can attain.

## **Acknowledgements**

I would like to thank my supervisor, Dr. Antoni Chawluk of Mansfield College, Oxford, for his guidance in all aspects of this thesis.

Chapter 2 was written whilst I was a visiting researcher at the Croatian National Bank in the summer of 2004, funded by an ESRC Overseas Institutional Visit grant. I would like to thank Dr. Evan Kraft, Director of Research at the Croatian National Bank, for facilitating this visit; and him, Maroje Lang, Vedran Šošić, and Velimir Šonje for their invaluable advice on my research. Thanks are also due to Dr. Kraft and to Dr. Timothy Lane of Balliol College, Oxford for generously providing comments on a draft of the chapter.

Chapter 4 benefitted greatly from discussions with Igor Vetlov of the Central Bank of Lithuania; as well as from comments made by participants at a Gorman Seminar in Oxford.

Needless to say, no one mentioned above should be implicated in any of my analysis or conclusions – all of which, for better or for worse, are my own.

## Contents

<b>Introduction</b>	<b>1</b>
<b>1. <i>Non Aes Sed Fides</i>: the Credit Theory of Money and its Implications for the Analysis of Money Substitutes</b>	<b>7</b>
1.1 Introduction .....	9
1.2 The Two Theories of Money .....	10
1.2.1 The Metallist Theory of Money .....	10
1.2.2 The Credit Theory of Money.....	15
1.3 Recent Models of the Micro-foundations of Money and Barter .....	23
1.3.1 Models in the Metallist Tradition .....	24
1.3.2 Models in the Credit Tradition .....	39
1.4 Towards a Credit Theory of Money Substitutes .....	51
1.4.1 The Metallist Theory of Money and the Analysis of Money Substitutes .....	52
1.4.2 The Credit Theory of Money and the Analysis of Money Substitutes .....	55
1.5 Conclusions .....	65
<b>2. Non-Cash Settlements in Transition: the Case of Croatia</b>	<b>69</b>
2.1 Introduction.....	71
2.2 Theory .....	75
2.3 Institutional Analysis .....	79
2.3.1 The primary market for credit .....	79
2.3.2 The secondary market for credit .....	86
2.4 Econometric Analysis .....	88
2.4.1 Testing H(1) and H(2) .....	88
2.4.2 Testing H(3) .....	94
2.5 Conclusions .....	97
Figures.....	99
Tables.....	113

**3. Dollarisation in Developing and Transition Countries: a Review of the Theoretical and Empirical Literature** **121**

3.1 Introduction .....	123
3.2 A Taxonomy of Dollarisation .....	124
3.3 The Dimensions of Dollarisation .....	129
3.3.1 Financial dollarisation .....	130
3.3.2 Real dollarisation .....	133
3.3.3 Currency substitution .....	134
3.4 Theories of Dollarisation .....	137
3.4.1 Structural factors influencing dollarisation .....	139
3.4.2 Transactions dollarisation .....	143
3.4.3 Store of value dollarisation .....	151
3.5 Policy Responses to Dollarisation .....	168
3.5.1 The case in favour of a policy response to dollarisation.....	169
3.5.2 Policy options to manage dollarisation .....	176
3.6 Conclusions.....	186

**4. Deposit Dollarisation in the Baltics: a Portfolio Balance Approach**

**191**

4.1 Introduction .....	193
4.2 Introduction to the Empirical Cases.....	195
4.2.1 Estonia and Lithuania: Currency Reform and Stabilisation Prior to the Introduction of the CBAs .....	195
4.2.2 Estonia and Lithuania: Economic and Political Developments after the Introduction of the CBAs.....	197
4.2.3 The Development of Deposit Interest Rates .....	201
4.3 Empirical Strategy .....	203
4.3.1 Choice of variables .....	203
4.3.2 Time Series Characteristics .....	205
4.3.3 Treatment of expectations.....	206
4.3.4 Modelling strategy .....	208
4.4 Results .....	210
4.4.1 Lithuania .....	210
4.4.2 Estonia .....	213
4.5 Interpretation.....	216

4.5.1 Vector versus Single Equation Methods, and Endogeneity of Variables	216
4.5.2 Empirical Application of Portfolio Balance Theory to Deposit Dollarisation	218
4.5.3 Interpretation of Numerical Estimates	221
4.6 Conclusions	225
Appendix A: Construction of Data Series	227
Appendix B: VECM models for Lithuania and Estonia	233
Appendix C: A two Currency Portfolio Balance Model of Deposit Dollarisation	238
Tables	241
Figures	253

<b>Bibliography</b>	<b>271</b>
---------------------	------------

# Introduction

“As they finish their tea and croissants, two elegantly dressed ladies at a Buenos Aires café ask their waiter how they might pay. As if reciting the day’s menu from memory, the waiter gives them several options: pesos, lecops, patacones (but only Series I) and all classes of tickets – luncheon vouchers that circulate widely at restaurants and supermarkets in the city. Oh, and US dollars of course, which will be accepted at a rate of three pesos each.” *Financial Times*, April 11, 2002

“In the Ukraine, all atomic energy enterprises receive from their consumers, who do not have money, the products of those same consumers. These items could be given out to the electricity workers, who haven’t seen the national money for years. But that is bothersome. So the enterprise issues its own money, gives it a name, and workers can volunteer to accept it as wages. The known number of such private and self-accounting moneys in the Ukraine is in the hundreds, and in Russia must amount to tens of thousands.” Ryabchenko, P. (1998), quoted in Humphrey, C. (2000)

In order to celebrate the submission of this thesis, I intend to meet a friend in a pub. The price of the beer that I buy will be quoted in sterling; I will pay for my pint in sterling cash; I will have withdrawn that cash from my sterling savings; and no doubt thereby (further) depleted my prospects of contributing adequately to my sterling pension fund. In the UK, one has to be either very cosmopolitan or very rich to have a foreign currency bank account; and despite the vast number of tourists who flood into Oxford every summer, it was only last year that the Oxford Bus Company started to accept Euro cash in payment of tickets on its service from Heathrow Airport. In the UK, as in most advanced economies, the official, national money is king.

In most developing and transition countries, however, the situation is markedly different. Far from the economic functions of providing a means of payment, a unit of account, and a financial store of value being performed virtually exclusively by the official, national money, a cornucopia of alternatives flourishes. Undoubtedly the most common of these is foreign currency

– or, to be more precise, the foreign currencies of a handful of global and regional financial superpowers. Recent research has shown that the use of foreign currency in transactions and financial intermediation by households, firms, and government is now endemic in the developing and transitional world, and becoming ever more so. In 2001, over 50% of bank deposits in South America were denominated in foreign currency; the proportions for Eastern Europe and the FSU and the Middle East were 48% and 42% respectively<sup>1</sup>. In 2002, it was estimated that over two thirds of US Dollar cash was circulating abroad (indeed, US households admit to holding only 10% of the outstanding issue) – the great majority of it in developing or transition countries<sup>2</sup>.

Moreover, in addition to foreign currency the last decade has seen well-publicised episodes of the emergence of more exotic substitutes for the official, national money in several important developing and transition countries – as illustrated in the cases of Argentina and Ukraine by the quotations above. In the transition economies of the former Soviet Union (FSU) during the mid- to late 1990s, the use of such money substitutes was mainly confined to transactions within and between the enterprise and government sectors: by 1997 the share of enterprise sales settled using “non-cash settlements” (NCS) in Russia was over 40%<sup>3</sup>. In Argentina, following the financial crisis of 2001, even retail transactions came to be conducted using ‘unofficial’ moneys – in this case mainly new currencies issued by sub-national governments or the private sector: by March, 2002, the value of such substitute currencies in circulation was equivalent to 37.5% of the total peso money supply<sup>4</sup>. Nor is the emergence of such private and sub-sovereign substitutes for the official, national money only an acute pathology, driven by macroeconomic crisis – recent research has revealed that it can be a chronic condition, persisting for many years and under stable macroeconomic conditions, as well<sup>5</sup>.

There is no doubt, then, that in stark contrast to the situation in advanced economies, the state’s franchise on the institution of money in most developing and transition countries is remarkably insecure. Nor is there any doubt that the consequences of this fact for macro-

---

<sup>1</sup>De Nicoló, Honohan, and Ize (2003).

<sup>2</sup>Feige and Dean (2002).

<sup>3</sup>Aukutsionek (1998), p.179. This mean figure understates the significance of the phenomenon in individual cases: in a third of all enterprises, over 70% of all sales were settled via NCS. See Chapter 2 for further references to the extensive literature on NCS in the FSU during the 1990s; useful overviews include Woodruff (1999) and Seabright ed. (2000).

<sup>4</sup>De la Torre, Levy Yeyati, and Schmukler (2003), p.77.

<sup>5</sup>As in the case of Croatia investigated in Chapter 2 of this thesis, for example.

economic and financial sector policy are significant, whether the alternatives that substitute for the official, national money consist of foreign currency or of private money of some sort. A recent IMF review of dollarisation in low and middle income countries concluded, for example, that “the phenomenon of dollarisation poses a challenge to the pursuit of a coherent and independent monetary policy”<sup>6</sup>. The same institution called the issuance of private and sub-sovereign currencies in Argentina in the aftermath of the 2001-2 financial crisis “a practice that has complicated economic management, raised the threat of inflation, and undermined confidence in the public finances”<sup>7</sup>. At certain points in this thesis, I will discuss these consequences of the emergence of money substitutes on a large scale in more detail. The principal topic of this thesis, however, is an investigation not of the consequences, but of the causes of such widespread substitution of the official, national money in developing and transition countries. To this end, the thesis comprises four essays: two of which (Chapters 1 and 3) address theoretical issues in the economics of money substitutes, and two of which (Chapters 2 and 4) are empirical studies. The four essays are not intended as a comprehensive or unified treatment of this large and complex topic, but instead as four approaches to the question of what causes money substitutes to be so widespread in developing and transition countries, from a variety of methodological perspectives.

**Chapter 1** addresses the question of what general theory of money serves best as the framework within which to analyse the co-circulation of official, national monies and money substitutes in developing and transition economies. In particular, the purpose of this essay is to argue that the Credit theory of money – an ancient but until recently neglected tradition in the theory of money – is a useful framework for such analysis. I begin by surveying the classical expositions of the Metallist and Credit theories of money. I then proceed to review modern models of the micro-foundations of money; to demonstrate from which of the two theories they derive; and to identify both their successes and their limitations in exploring the micro-foundations of the Metallist and Credit theories of money. I find that whilst modern models have made important advances in our understanding of the Credit theory, their ability to provide an analysis of the competition between alternative monies is incomplete in several central respects. I

---

<sup>6</sup>Baliño, Bennett, and Borzenstein (1999), p.14.

<sup>7</sup>IMF (2002), p.2.

therefore conclude by further scrutinising a number of ways in which this analysis is conducted in the classical Credit theory literature, assessing their relevance to empirical examples of money substitutes in developing and transition countries, and thereby evaluating their merits as guides to future theoretical research.

In **Chapter 2**, I proceed to an analysis of an empirical case study of the emergence and persistence of inter-enterprise NCS in Croatia during the 1990s, using some of the ideas generated by the theoretical analysis in Chapter 1. Having analysed the evolution of the monetary regime and the banking system in Croatia, I proceed to demonstrate, using time series econometric models, that with respect to the private credit instruments used to settle NCS both (a) the Meltzer (1960) hypothesis that trade credit substitutes for bank credit when monetary conditions tighten and (b) the hypothesis that these instruments are substitutes not only for official credit, but for official money, are corroborated. I conclude therefore that NCS in Croatia are evidence of the endogenous creation of money substitutes by solvent but illiquid enterprises resorting to the endogenous creation of money substitutes, in the form of circulating trade credit, in the face of access to official credit being rationed; and I contrast this conclusion with the interpretation of NCS in the well-studied case of Russia – where the literature concluded that they were principally a mechanism for the delivery of fiscal subsidies to insolvent firms.

Having considered privately-issued money substitutes in Chapter 2, I turn in **Chapter 3** to the most important form of money substitute found in developing and transition countries – foreign currency. The theoretical and empirical literature on dollarisation is extremely rich and diverse, and there have been many new and important contributions over the last decade. Despite this, the most recent major surveys of the dollarisation literature were those of Giovannini and Turtelboom (1994) and Mizen and Pentecost (1996b). The purpose of this essay, therefore, is to present an up-to-date survey of the theoretical and empirical literature on dollarisation in developing and transition economies. Two particular aims of the essay are first, to present a coherent taxonomy of the dollarisation phenomenon – distinguishing the functional categories of dollarisation (transactions and store-of-value dollarisation) from actual categories (currency substitution and financial dollarisation); and second, to survey the evolution of theoretical models of the determinants of dollarisation in response to the changing nature of the empirical phenomenon, charting the shift from monetary services models of currency substitution – which

was the dominant form of dollarisation in the 1970s and early 1980s – to the modern portfolio balance model of financial dollarisation – which has become important in the 1990s and 2000s as a result of financial liberalisation around the world.

**Chapter 4** is an empirical study of deposit dollarisation in the two transition economies of Estonia and Lithuania. Sahay and Vegh (1996) argue that the use of currency board arrangements (CBAs) in transition countries during stabilisation introduces a step change in the stability of macroeconomic and political conditions which renders the deposit dollarisation process amenable to analysis in terms of simple, portfolio theoretic factors. By building dynamic, multiple equation, econometric models of the deposit dollarisation process in Estonia and Lithuania, and interpreting them in the light of modern theoretical research reviewed in Chapter 3, I seek on the one hand to test this general hypothesis, and on the other to explore the limitations of applying portfolio balance theory to dollarisation in a time series econometric context. I find that for both economies it is possible to construct robust statistical models of the deposit dollarisation process which sustain interpretation in terms of portfolio balance theory; but that constraints on data availability mean that these models cannot be used to draw more detailed inference regarding, for example, the relative contributions to this process of hedging and speculative demand.



## Chapter 1

# *Non Aes Sed Fides: The Credit Theory of Money and its Implications for the Analysis of Money Substitutes*



## 1.1 Introduction

Episodes of the emergence on a large scale of money substitutes in developing and transition countries would seem to afford us a fascinating opportunity to deepen our understanding of monetary phenomena more broadly, representing as they do a degree of variation in the sources of liquidity disposed of by households, firms, and government unprecedented in most advanced economies. But what theoretical framework are we to use to understand them? The natural candidate might seem to be the Metallist theory of money that has long been pre-eminent in the economics literature; but the empirical relevance of models derived from this theory has often been questioned. In this chapter, we consider an alternative candidate - the Credit theory of money - and identify in this ancient, but until recently neglected, theoretical tradition a promising framework for the analysis of money substitutes in developing and transition countries.

The chapter is organized as follows. In the next section, we introduce the Metallist and Credit theories of money, giving an overview of their respective concepts of money and of the nature of monetary exchange; as well as of their explanations of what determines the choice between monies. These classical expositions are general in scope; but the modern theoretical literature on the micro-foundations of monetary theory is an active one and explores the general postulates of the Metallist and Credit theories with considerable rigour. In Section 1.3, we therefore proceed to review a representative selection of modern models of the micro-foundations of money and barter. We ask from which of the two classical theoretical traditions these recent models derive, and discuss both their successes and their limitations in exploring the micro-foundations of the Metallist and Credit theories of money. We find that whilst modern models have made important advances in our understanding of the Credit theory, their ability to provide an analysis of the competition between alternative monies is incomplete in several central respects. In Section 1.4 we therefore return to the classic Credit theory literature in order to examine in more depth a number of ways in which this analysis is conducted there, and to assess their respective merits as guides to future theoretical research on the basis of their relevance to empirical examples of money substitutes in developing and transition countries. Section 1.5 gives a summary of our findings.

## 1.2 The Two Theories of Money

Two theories of money – the Metallist and the Credit theories – dominate the economics literature, ancient and modern. Each has its own concept of money in general and of the nature of monetary exchange; and each provides its own explanation of the choice between particular monies. Both theories are best known in derivative forms designed to explain the circulation of official, national currencies: in a well-known paper, Goodhart (1998), for example, makes a comparison between the ‘Optimal Currency Area’ and the ‘Cartalist’ theories of official, national currencies – the former being a derivative of the Metallist theory in which the operation of the market governs the choice between monies; and the latter being a derivative of the Credit theory in which the power of the state is identified as the chief determinant of the value of official, national money. But the Metallist and the Credit theories themselves are theories of money in general, the scope of which includes the nature of and choice between all types of money – domestic and foreign, official and private; and it is therefore these, more general versions of the two theories of money, that are of interest to us here. In this section we therefore introduce each of the two theories in turn; beginning in Section 1.2.1 with the historically more popular Metallist theory, and proceeding in Section 1.2.2 to the Credit theory of money. In each case, we survey first the concepts of money and of the nature of monetary exchange at the heart of the theory in question; and then review its theory of what determines the choice between monies.

### 1.2.1 The Metallist Theory of Money

Five claims constitute the unchanging core of the Metallist theory of money. The first and most basic of these concerns the Metallist concept of money:

1. Money is a ‘thing’ – that is, an asset amongst other assets available to be owned and exchanged by economic agents. It is an ‘organic member in the world of commodities’ in a famous phrase of Carl Menger<sup>1</sup>.

The second concerns the nature of monetary – as distinct from non-monetary – exchange:

2. Monetary exchange consists in the exchange of goods and services for the money-

---

<sup>1</sup>Menger (1892). p.240.

commodity, rather than for other goods and services (which would count as ‘barter’). Hence the money-commodity can be called a ‘medium of exchange’ – since unlike other commodities it is “an object which is taken in exchange, not for its own account, i.e. not to be consumed by the receiver or to be employed in technical production, but to be exchanged for something else within a longer or shorter period of time”<sup>2</sup>.

Together, these two claims concerning the concepts of money as a commodity, and of the nature of monetary exchange as transactions involving a medium of exchange, motivate the Metallist theory’s account of how it is that money comes to exist. This consists of two further central claims; the third being that:

3. The transactions costs inherent in barter exchange were the historical, and are the logical, precondition for the fact that one particular commodity, such as gold or silver specie, should emerge from all others to serve as the money-commodity medium of exchange.

And the fourth that:

4. Given transactions costs referred to in Claim 3, the money commodity is valued by agents, and therefore chosen to serve as the medium of exchange, on the basis of its intrinsic qualities – its durability and portability, natural scarcity, difficulty of counterfeit, and so on<sup>3</sup>.

These four claims are common to all expositions of the Metallist theory. Lest it seem as though we are erecting a straw man, however, let us briefly illustrate the venerability of these claims, the range of their appeal, and the unity of their exposition over several millennia, by the following examples.

Aristotle’s *Politics* contains the earliest example we have of an explanation of the origins of money; and it is unmistakably Metallist:

“When they had come to supply themselves more from abroad by importing things in which they were deficient and exporting those in which they had a surplus, the employment of money necessarily came to be devised. For the natural necessities are not in every case readily portable; hence for the purpose of barter men made a

---

<sup>2</sup>Wicksell (1911), quoted in Kiyotaki and Wright (1989), p.937.

<sup>3</sup>Note that this claim and Claim 3 do not necessarily imply that the money-commodity is chosen as the result of a competitive market process (although this, the ‘Mengerian’, explanation is traditionally the most popular with theorists): it is possible that, for example, the sovereign unilaterally chooses the money commodity in order to reduce the transactions costs associated with non-monetary exchange on the basis of its attractive intrinsic qualities. We discuss these alternatives in Sections 1.3.1 and 1.4.1 below.

mutual compact to give and accept some substance of such a sort as being itself a useful commodity was useful to handle in use for general life, iron for instance, silver and other metals, at the first stage defined merely by size and weight, but finally also by impressing on it a stamp in order that this might relieve them of having to measure it; for the stamp was put on as a measure of the amount”<sup>4</sup>

All three central concepts above are present in Aristotle’s account; and in addition a characteristically Metallist interpretation of the purpose of minting. Since specie coinage is, according to the Metallist theory, nothing but a particular weight and fineness of commodity in service as a medium of exchange, it is argued that the mint’s stamp has no significance for the value of money other than to relieve a potential problem of private information regarding the coin’s specie content.

John Locke advanced this analysis somewhat by stressing that it is the durability of precious metals that is the particular intrinsic quality that suits them to serve as money on the one hand, and by introducing explicitly the idea that the transition from barter to monetary exchange – the invention of a medium of exchange – is to be explained as a spontaneous agreement made between trading agents:

“The greatest part of things really useful to the life of man, and such as the necessity of subsisting made the first commoners of the world look after, as it doth the Americans now, are generally things of short duration; such as, if they are not consumed by use, will decay and perish of themselves: gold, silver and diamonds, are things that fancy or agreement has put the value on, more than real use, and the necessary support of life. . . .

And thus came in the use of money, some lasting thing that might keep without spoiling, and that by mutual consent men could take in exchange for the truly useful, but perishable supports of life.”<sup>5</sup>

---

<sup>4</sup>Aristotle (1932), I.3.13-14. This passage well illustrates Aristotle’s advocacy of the main elements of the Metallist theory as summarized above: but it also demonstrates an idiosyncrasy of his. He argues here that utility beyond their monetary role is a property that makes metals appropriate commodity monies – but elsewhere (Aristotle (1932) I.3.16) takes the more common Metallist view that it is the fact that precious metals do not have any other significant use that makes them good candidates.

<sup>5</sup>Locke (1980), §46 and §47, p.28. Locke’s introduction of the idea that the fixing of a particular commodity

It was left to Adam Smith, however, in this as in so many other fields of economic theory, to provide the first, clear, and unified exposition of the Metallist theory of money. In his chapter “Of the Origin and Use of Money” in *The Wealth of Nations*, Smith starts from the claim that money arises in order to overcome the inconvenience of barter exchange:

“But when the division of labour first began to take place, this power of exchanging must frequently have been very much clogged and embarrassed in its operations. . . . The butcher has more meat in his shop than he himself can consume, and the brewer and the baker would each of them be willing to purchase a part of it. But they have nothing to offer in exchange, except the productions of their respective trades, and the butcher is already provided with all the bread and beer which he has immediate occasion for. . . . In order to avoid such situations, every prudent man in every period of society, after the first establishment of the division of labour, must naturally have endeavoured to manage his affairs in such a manner, as to have at all times by him, besides the peculiar produce of his own industry, a certain quantity of some one commodity or other, such as he imagined few other people would be likely to refuse in exchange for the produce of their industry.”<sup>6</sup>

Money, therefore, is simply that commodity chosen to serve as a medium of exchange and store of value: and indeed in different societies at different times, different commodities have served this end:

“Many different commodities, it is probable, were successively both thought of and employed for this purpose. In the rude ages of society, cattle are said to have been the most common instrument of commerce. . . . Salt is said to be the common instrument of commerce and exchange in Abyssinia; a species of shells in some parts of the coast of India; dried cod in Newfoundland; tobacco in Virginia; sugar in some of our West India colonies; hides or dressed leather in some other countries; and

---

as the medium of exchange should have been “by mutual consent” is the grandfather of the ‘Mengerian’ approach to modeling money as an equilibrium outcome of private choices referred to above. It is informative to note that for Locke, it followed naturally from his political philosophy, of which this theory of money is just one part.

<sup>6</sup>Smith (1981), pp.37-8.

there is to this day a village in Scotland where it is not uncommon, I am told, for a workman to carry nails instead of money to the baker's shop or the alehouse."<sup>7</sup>

It is precious metals, however, that by virtue of their enjoying a unique combination of intrinsic durability and divisibility, that are generally chosen as the commodities that serve as money:

"In all countries, however, men seem at last to have been determined by irresistible reasons to give the preference, for this employment, to metals above every other commodity. Metals can not only be kept with as little loss as any other commodity, scarce any thing being less perishable than they are, but they can likewise, without any loss, be divided into any number of parts, as by fusion of those parts can easily be re-united again; a quality which no other equally durable commodities possess, and which more than any other quality renders them fit to be the instruments of commerce and circulation."<sup>8</sup>

The four central claims of the Metallist theory introduced so far provide an account of transactions money; but say nothing about the related concept of credit. Far from being an oversight, this is because the fifth core claim of the Metallist theory of money is precisely that the institution of credit is conceptually quite separate from that of money. Credit, on the Metallist theory, consists simply of the borrowing and lending of commodities; and financial credit is simply a special case of this, being the borrowing and lending of the money-commodity. The invention of credit is conceived to be a logically and historically posterior development, supervenient on the emergence of commodity-money out of the barter economy. This fifth central claim is therefore:

5. "That to economise on the use of metals and to prevent their constant transport a machinery called 'credit' has grown up in modern days, by means of which, instead of handing over a certain weight of metal at each transaction, a

---

<sup>7</sup> *Ibid.* p.38.

<sup>8</sup> *Ibid.* pp.38-9.

promise to do so is given, which under favourable circumstances has the same value as the metal itself. Credit is called a substitute for gold”<sup>9</sup>

There is one final important aspect of the Metallist theory which we must note – but this concerns not what is included in the theory, but what is omitted from it. Money serves not only as a means of payment and a store of value, but also as a unit of account. The Metallist theory offers an explanation of the first two of these roles; but leaves the third virtually untouched. Once again, this follows from the theory’s concepts of money and of monetary exchange (Claims 1 and 2 above). Money is a commodity; and the relative prices of different commodities are identical whether denominated in units of one commodity or another. The choice of numeraire is therefore irrelevant to the operation of the market; and explaining why the unit of money-commodity is chosen as the numeraire in preference to the unit of some other commodity is peripheral to the Metallist theory. As such, it is assumed, rather than explained.

### 1.2.2 The Credit Theory of Money

Although the Metallist theory of money has proved much the more popular amongst modern economists, there exists an alternative, and in fact equally distinguished, theory of money in the Credit theory<sup>10</sup>. Contemporary partisans of the Credit theory such as Ingham (2004), Smithin (2000), and Wray (2004) attribute its modern version to the little-known economic thinker Alfred Mitchell Innes – and certainly Mitchell Innes (1913, 1914) gives us one of the theory’s most enduring and elegant expositions. Nevertheless, as Keynes (1914) pointed out in his review of Mitchell Innes (1913), the modern revival of the Credit theory of money should more properly be attributed to the late nineteenth century economist Henry Dunning Macleod, whose *Principles of Economical Philosophy*, and *Theory of Credit* are all the more remarkable for having been overlooked as foreshadowing modern finance theory by generations of economists since the turn of the 19th century<sup>11</sup>. As was the case for the Metallist theory above, there is

---

<sup>9</sup>Mitchell Innes (1913) p.377.

<sup>10</sup>Numerous historical authorities for the Credit theory, starting, ironically, from Aristotle (though in his *Nicomachean Ethics*, rather than his *Politics*), are given in Macleod (1891), p.75ff.

<sup>11</sup>So Schumpeter, in his *History of Economic Analysis*, argued that whilst Thornton and Mill “failed . . . to build up a systematic credit theory of money, and on principle clung to the monetary theory of credit”, “[W]e might see the outlines of such a theory in the works of Macleod.” (Schumpeter (1954), p.718). Macleod’s contributions, however, “remained . . . completely outside of the pale of recognised economics” (*ibid.*). This was

neither the space nor the need here for a comprehensive overview of the Credit theory of money. Instead, we will present a brief account of the most important aspects of the theory by taking the five central claims of the Metallist theory set out in the previous section, and illustrating the counter-claims of the Credit theory to each one.

We may begin with Claim 3 above, that the inconvenience of barter is the central determinant of the choice of one commodity to serve as money. That barter – the swapping of commodities for commodities – was, in the distant past of advanced economies, and is, in the present of primitive economies, the archetypal form of exchange is a maintained hypothesis of the Metallist theory. As pointed out by Goodhart (1998), however, drawing on the extensive historical, anthropological, and sociological literature on the origins of money, there is in fact no record of any society at any time in history operating solely on the basis of barter exchange – this scenario, so popular amongst economic theorists, is entirely hypothetical. One modern authority on the anthropology of monetary exchange, for example, finds that:

“Barter, in the strict sense of moneyless market exchange, has never been a quantitatively important or dominant mode of transaction in any past or present economic system about which we have hard information.”<sup>12</sup>

and concludes, with respect to the hypothetical history of monetary emerging from barter economies much loved by Metallist theory, that:

“Moneyless market exchange was not an evolutionary stage in the sense of a dominant mode of transaction preceding the arrival of monetary means of market exchange. Barter occurs very widely in past and present economic systems, but always as minor, infrequent, or emergency transactions employed for special reasons by barterers who know of alternative and more important ways of transacting.”<sup>13</sup>

---

not because of any deficiency in his analysis, but because Macleod was, as Schumpeter put it, “an economist of many merits who somehow failed to achieve recognition, or even to be taken quite seriously, owing to his inability to put his many good ideas in a professionally acceptable form.” (*ibid.* p.1115). Anyone who has read Macleod’s idiosyncratic tracts, let alone his offerings to the appointing committees for the numerous chairs in political economy for which he applied (but never won), can only agree with this assessment. His 1888 submission for the Chair at Oxford, for example, includes amongst his credentials that he came third in his school mathematics exam, and had never picked up a book of political economy before the age of 26.

<sup>12</sup>Dalton (1982) p.185.

<sup>13</sup>*Ibid.* p.188.

In fact, all available historical evidence argues that the basic institution of emerging market exchange was not barter but trade credit: the story of the growth of trade is the story of the growth of networks of credit between merchants. Moreover, this is as true of primitive societies in more recent times as it is of historical cases: in every one, to a greater or lesser degree depending on the prevalence of market exchange, the institution of credit has existed almost exactly as it is known in advanced economies. That this is so was fairly accurately described as early as the beginning of the twentieth century by Mitchell Innes (1913, 1914); and recent studies of ancient monetary systems by Millet (1991), Cohen (1992), and Kim (2001); of monetary exchange in Europe during the middle ages at the local, national and international levels by Marshall (1999), de Roover (1974), and Postan (1973) respectively, and in the early modern period by Muldrew (1998); and the classic studies of money in primitive societies of Quiggin (1949) and Einzig (1966) all serve to confirm Mitchell Innes' (1913) central conclusion that "[t]here is no question but that credit is far older than cash"<sup>14</sup>, and that it is trade credit, rather than cash, which is the fundamental institution of market exchange.

If networks of trade credit are, historically speaking, the basic institution of market exchange – contradicting Claim 3 – then what becomes of Claim 2, the Metallist concept of money as a medium of exchange? The Metallist theory holds that the archetypal transaction in market-based trade consists of the exchange by an agent of commodities for a medium of exchange – the money-commodity – which may then at a later date be exchanged for other commodities. The Credit theory, by contrast, argues that this description of the archetypal transaction is misleading: the archetypal transaction in fact involves the exchange of commodities for a credit – a financial claim on the buyer, that is – which may be extinguished at a later date by the buyer's presentation of a countervailing claim on the seller. As Mitchell Innes (1913) has it:

“Exchange with money is structurally different from barter . . . in that transactions with money have two levels, as opposed to the one-dimensional exchange of goods. . . .a [monetary] sale does not involve an exchange for ‘some intermediate commodity called the “medium of exchange”’, as it would if money were no more than ‘efficient’ barter. Rather a sale is the exchange of a commodity for a ‘credit’

---

<sup>14</sup>Mitchell Innes (1913), p.396.

which, in accordance with the ‘primitive law of commerce’, represents the ‘debt’ in the next purchase.”<sup>15</sup>

A consequence of this view is the hypothesis that the great majority of market exchange involves, and has always involved, the accumulation by merchants of credits and debts, and their mutual extinction by various forms of clearing. Once again, this alternative hypothesis is corroborated by a wealth of historical and anthropological evidence. In mediaeval Europe, for example, the institutions of trade credit had already attained such a high degree of development and such a wide geographical scope that the essential role of the great merchant fairs, far from being as places to which commodities were brought and physically traded, was precisely as vast clearing-houses for merchants’ commercial credits and debts:

“At numerous centers of commerce, Lyons, Antwerp, Nuremberg, Hamburg, and many others, there were held great fairs every three months. . . . On a certain day of the fair the merchants met together and presented their acceptances to each other: and if their respective claims were equal they were balanced and paid by being exchanged against each other, by Compensation. By this means an enormous commerce was carried on and liquidated without any specie at all. Boisguillebert, one of the morning stars of modern Economics, says that at the fair of Lyons transactions to the amount of 80,000,000 livres were settled, without the use of a single coin.”<sup>16</sup>

In other words, it was as true of ancient and primitive as it is of modern, market economies that, as Hoover (1996) has it “transactions are conducted with a higher frequency than settlements: we spend money every day, but pay our bills monthly.”<sup>17</sup>

---

<sup>15</sup>Mitchell Innes (1913) p.393. Note that Hicks (1989) too acknowledged this fact (see Ch. 5 therein), preferring to define money as a means of deferred payment than a medium of exchange.

<sup>16</sup>Macleod (1891), p.333. One of the greatest twentieth century scholars of banking history, Raymond de Roover, confirms this point: “The opinion has been advanced that most local payments at the fairs of Champagne were made in coin. Such a statement overlooks that the merchants frequenting the fairs brought goods instead of specie. Where would the specie have come from in order to make all payments in hard coin? This is clearly an impossibility, and it is not surprising that the Italian money-changers who did banking at the fairs had devised an ingenious system of book transfers, which allowed them to reduce the role of specie to a minimum. In order to facilitate settlements, they were even generous in allowing overdrafts for the duration of the fair. . . . The problem of the fairs of Champagne as a clearing centre is another which requires further research, but its solution may be hampered by inadequate documentation, unless new sources are discovered.” De Roover (1954), p. 42.

<sup>17</sup>Hoover (1996) p.214, where settlements denote payment in cash, of course. A further example of the mutual

The rejection of Claims 3 and 2 above – the rejection, that is, of the concept of money as a commodity currency in favour of that of money as the liquidity provided by a network of credits and debts that may be cleared against one another – naturally casts doubts on Claim 4 above as well. The more superficial question raised is how the Credit theory explains the existence and nature of coinage. Coinage is unproblematic on the Metallist theory’s account of money – indeed, it is at the centre of that account. The Credit theory, however, holds that the basic institution of market exchange is trade credit. Coinage is therefore to be understood as merely one amongst many ways of recording and representing financial contracts amongst economic agents. Coins are nothing other than tokens used to record trade credits and debts in such a way that accounts can be settled without the need to refer at every transaction to a merchant’s general ledger of assets and liabilities. This in turn implies a fundamentally different interpretation of the role of minting: since coins are simply tokens of credit, just as, for example, bills of exchange or promissory notes are, the intrinsic properties of the physical coins themselves are of minor significance – in contradiction to Claim 4 above. Whilst these properties very likely have much to do with the choice of material used to manufacture the tokens, they have nothing to do with the value of the money these tokens represent<sup>18</sup>.

---

off-setting of debts as the central form of monetary exchange of particular importance to the development of national currencies was the use of Tallies by the mediaeval English Exchequer. Tallies were hazel rods that were squared, marked to represent a specific value of debt, and then split in half. One half (the Tally) was kept at the Exchequer; the other (the Countertally) was given to the sovereign’s creditor, and could be used to extinguish future tax or other debts to the sovereign. According to Madox (1711), “Tallies were in great and constant use in the Exchequer [sic. in the middle ages]. The use of them was very ancient; coeval, for ought I know, with the Exchequer itself in England.” (p.709); and not only the King, but corporate bodies such as the City of London used them to pay for expenditures (p.711).

<sup>18</sup>A picturesque illustration of the bankruptcy of Metallist theory on this point is provided by the monetary system of the island of Yap, which is much celebrated in monetary history and anthropology because its currency consisted of large, stone wheels. This historical case has been much abused by Metallists, who have attempted to press it into the service of the Metallist theory, arguing that these stone wheels were the optimal choice of medium of exchange on Yap as a result of their durability, scarcity, uselessness in other respects, and so on (portability could not be claimed – some of the wheels were so large as to be for all practical purposes immobile). In fact, it is clear from studies of Yap society that these stones were not currency in the Metallist sense (i.e. a medium of exchange) at all, but tokens of an underlying system of credits and debts, as the Credit theory would predict. Perhaps the best evidence of this is provided by Furness (1910), who reports a famous case in which the value associated with a particularly magnificent stone wheel still accrued to its owner, and was transferred to another party in payment of a debt, even though the wheel itself had sunk to the bottom of the sea as the result of an accident in transport.

An alternative empirical counterexample is the international circulation of some high value coins of the middle ages – such as the Byzantine Bezant, the Venetian Ducat, and the Moslem Dinar – but not others – such as the gold coins of Frederick II or of Henry III of England – despite the fact that the intrinsic properties of all these coins (i.e. ratio of specie content to nominal value; quality of mint; etc.) were practically identical (Cipolla (1956), p.24).

But this merely emphasises a more fundamental difference between the Metallist and the Credit theories. For if it is not the intrinsic qualities of the money-commodity, in conjunction with the structural transactions costs associated with non-monetary barter and agents' preferences, that determine the value of coins as per Claim 4; what is it that does? The answer given by the Credit theory is that coins – or indeed any other type of token representing money – are valued not as a commodities but as tokens of an underlying financial contract between their holder and their issuer. As a result, the value of a particular money depends on a range of institutional and proximate determinants that are irrelevant by the lights of the Metallist theory. In particular, the capacity of a money's issuer to make a binding contractual commitment – a capacity which may be affected both by the institutional environment and by the issuer's own attributes – takes centre stage; and the essential determinants of the value of coined, as of any other type of money, are therefore accurately summarised by the motto famously stamped on the coins of Malta: “*Non Aes, Sed Fides*”<sup>19</sup>.

The rejection by the Credit theory of each of Claims 2, 3, and 4 therefore amounts to a refutation of the most basic claim of the Metallist theory: that money is to be conceived of as a ‘thing’ (Claim 1). In its place, the Credit theory places the alternative claim that all money is credit – that is, a financial contract between trading agents:

“Money, then, is credit, and nothing but credit. A's money is B's debt to him, and when B pays his debt, A's money disappears. This is the whole theory of money.”<sup>20</sup>

Or rather, not quite the whole theory, since it is important at once to point out that the claim that all money is credit is not equivalent to the claim that all credit is money. In order for credit to be money, it must be transferable from one creditor to another, in settlement of debts. It is the innovation of the transferability of debts from one creditor to another which is the critical development in the history of money. As Macleod (1882) has it (somewhat melodramatically):

---

<sup>19</sup> “Not the metal, but trust”. This motto was introduced in 1565 to reassure the Maltese of what it was that guaranteed their currency's value when copper coins had to be struck for shortage of more precious metals during the siege of Valetta by the Turks. See <http://www.centralbankmalta.com/site/currency1b.html>.

<sup>20</sup> Mitchell Innes (1913) p.402.

“If we were asked – Who made the discovery which has most deeply affected the fortunes of the human race? We think, after full consideration, we might safely answer – The man who first discovered that a Debt is a Saleable Commodity.”<sup>21</sup>

Money is transferable credit. This, in contradiction to Claim 1 above, is the fundamental concept of the Credit theory of money:

“These simple considerations at once shew the fundamental nature of a Currency. It is quite clear that its primary use is to measure and record debts, and to facilitate their transfer from one person to another; and whatever means be adopted for this purpose, whether it be gold, silver, paper, or anything else, is a currency. We may therefore lay down our fundamental Conception that Currency and Transferable Debt are convertible terms; whatever represents transferable debt of any sort is Currency; and whatever material the Currency may consist of, it represents Transferable Debt, and nothing else.”<sup>22</sup>

In the language of modern economics, money is, in other words, a financial contract: and not a bilateral contract between two parties only, but a unilateral contract on the part of the creditor, to which the counterparty is whoever should happen to hold his liability at any given time<sup>23</sup>:

“Now the Duty to Pay must evidently always remain fixed in the person of the Debtor. But the Right to Demand, or the Credit, is the absolute property of the Creditor, and he may sell or transfer it, like any other property whatever, and whenever it is so transferred, a Nexus or Contract takes place between the transferee of the Right or Credit, and the original debtor, and it can only be extinguished by the revesting of the Right, or Credit, in the person of the debtor from whom it originally emanated.”<sup>24</sup>

---

<sup>21</sup> Macleod (1882), p.481. The claim is not necessarily as far fetched as it sounds: de Roover (1954) p.58, for example, argues that the development of the principle of negotiability by European merchant bankers during the sixteenth century was indeed a revolutionary moment in the history of money and finance.

<sup>22</sup> Macleod (1882), p.188.

<sup>23</sup> ‘Unilateral’ is the term used by Macleod (1882); the modern theoretical literature on financial contracting would call this a ‘multilateral’ contract. We opt for the former description, which implies more clearly the fact that the contract is not struck explicitly with any creditor subsequent to the primary issue.

<sup>24</sup> Macleod (1882), p.183-4.

As we shall discuss in the next section, this alternative conception of money as a unilateral financial contract has important consequences for the modelling of behaviour connected with money and monetary exchange.

Given its alternative concept of the nature of money as transferable credit, it is unsurprising that the Credit theory also differs from the Metallist theory in its treatment of the nature of credit (Claim 5 above), and of the importance of the numeraire. Since money itself is conceived of as simply a special case of credit – being debt that is transferable by virtue of the unilateral commitment made by its issuer – credit is taken to be a primitive concept. As Macleod puts it, “[m]oney is only the highest and most general form of credit”<sup>25</sup>. The Credit theory therefore answers to Schumpeter’s preference in his *History of Economic Analysis* for a “Creditary theory of money” over the dominant, but deficient, “monetary theory of credit”:

“Historically, this method of building up the analysis of money, currency, and banking [*sc.* by first formulating a theory of currency, and then building a theory of credit on top of this] is readily understandable . . . [B]ut logically, it is by no means clear that the most useful method is to start from the coin – even if, making a concession to realism, we add inconvertible government paper – in order to proceed to the credit transactions of reality. It may be more useful to start from these in the first place, to look upon capitalist finance as a clearing system that cancels claims and debts and carries forward the differences – so that ‘money’ payments come in only as a special case without any particularly fundamental importance. In other words, practically and analytically, a credit theory of money is possibly preferable to a monetary theory of credit.”<sup>26</sup>

The Credit theory’s adoption of credit, rather than some putative money-commodity, as a primitive concept implies a final important difference from the Metallist theory, in that a theory of the numeraire becomes not a luxury, but a necessity. In rejecting the Metallist conception of money, the Credit theory rejects the concept of a naturally countable money-commodity, units of which are by default used to denominate prices and financial contracts. On the Credit theory,

---

<sup>25</sup> Macleod (1882), p.198-9.

<sup>26</sup> Schumpeter (1954), p.717.

the numeraire is instead an abstract measure of economic value; and since the primitive concept of credit has no existence independent of this abstract measure, a theory is required of how given numeraires are invented, calibrated, and come to serve as the units of account for given networks of transferable credits and debits. The practical importance of such a theory is confirmed by the very cases of the emergence of money substitutes which are the subject of this chapter. Not only is there the question of why money substitutes come into circulation to be answered; but the question of why some money substitutes – for example, the private credit instruments used to settle non-cash settlements in transition economies during the 1990s; or emergency municipal and private currencies issued during the First World War – are denominated in units of the domestic currency, whilst others – for example, the Argentine provincial and private currencies issued after 2001; as well as, by definition, all foreign currencies – have their own denominations<sup>27</sup>.

### **1.3 Recent Models of the Micro-foundations of Money and Barter**

The classical expositions surveyed in the previous section set out the main postulates of the Metallist and Credit theories of money in general terms: for example, they set out distinct concepts of money, and of the nature of monetary exchange; and distinct criteria on the basis of which the choice between monies is made. They do not however analyse in any detail the mechanics of the choice between alternative monies. The modern theoretical literature on the micro-foundations of money and barter has, however, been extremely active over the past two decades; and has investigated in a rigorous fashion numerous questions concerning the nature of and choice between monies. In this section, we therefore review a representative selection of modern models of the micro-foundations of money and barter, with two particular aims in mind. The first is to categorise these models according to which of the two classical traditions of monetary theory - Metallist or Credit - they derive from. The second is then to analyse both where they have, and where they have not, advanced our understanding of the main postulates of those traditions.

---

<sup>27</sup>For more on the issue of denomination see Chapter 2 on NCS money substitutes in transition economies; White (1921) on World War I currencies; De la Torre, Levy Yeyati, and Schmukler (2003) on the case of Argentina; and Chapter 3 on the circulation of foreign currency and ‘real dollarisation’.

### 1.3.1 Models in the Metallist Tradition

As we found in Section 1.2.1 above, the Metallist theory conceives of money as a commodity, chosen as a medium of exchange; and has this choice made on the basis of the minimisation of transactions costs dependent on the intrinsic qualities of the money-commodity and systemic constraints that militate against barter. These concepts and criteria naturally invite modelling via the tools of neoclassical consumer demand theory; and it is indeed this framework that dominates modern models in the Metallist tradition. The observed price and quantity in circulation of a given money is modelled as an equilibrium between the supply of and demand for commodities; and the ideal model is deemed to be one which can explain the emergence and circulation of money as a process of selection by the market with recourse to the imposition of as few exogenous constraints as possible. The standard Metallist approach to modeling the mechanism whereby money is chosen as the outcome of a decentralized market process can therefore be called ‘Mengerian’ – since to Menger’s (1892) famous questions:

“Is money an organic member in the world of commodities, or is it an economic anomaly? Are we to refer its commercial currency and its value in trade to the same causes conditioning those of other goods, or are they the distinct product of convention and authority?”<sup>28</sup>

its answer is in both cases the former<sup>29</sup>.

---

<sup>28</sup>Menger (1892), p.240.

<sup>29</sup>Curiously, however, Menger’s approach to the modeling of money as a market equilibrium remained a distinctively Austrian – and hence somewhat heterodox, to Anglo-Saxon eyes – one until after the Second World War. In the first half of the twentieth century, the dominant approach in Anglo-Saxon economic circles was instead the so-called ‘Cambridge approach’ to the micro-foundations of money, which treated the demand for money in an ad hoc fashion and separately from the theory of the demand for other commodities (see Laidler (1993), Chapter 5). It was not until the seminal contribution of Hicks (1935) that Menger’s methodology was introduced to the mainstream of Anglo-Saxon theory; and only with Friedman (1956) that Hicks’ call to bring the theory of the micro-foundations of money demand within the usual structure of consumer demand theory was formalized. Even thereafter, however, much of mainstream macroeconomic theory persisted in subsuming the problem of money’s micro-foundations, having recourse to making money an argument of consumers’ utility functions (Sidrauski (1967)), or imposing on them exogenous cash-in-advance constraints (Clower (1965)). Nevertheless, it at least became axiomatic to the orthodox literature on micro-foundations that the emergence of monetary from barter exchange should be modeled as an endogenous commodity market equilibrium, in contrast to alternative approaches in which the emergence of money is to be explained, for example, by exogenous legal restrictions imposed by the authority of the sovereign (though some modern Metallist models such as that of Engineer and Bernhardt (1991) combine the two approaches by incorporating exogenous legal restrictions).

As Arrow and Debreu (1954) long ago demonstrated, however, a decentralised market process operating in a Walrasian trading environment with perfect information will not generate a demand for a money-commodity to serve as a medium of exchange: in these circumstances, a Pareto optimal allocation of commodities will be attainable by barter exchange alone. Models which analyse the choice of a money-commodity to serve as a medium of exchange on the basis of its intrinsic qualities using the tools of neoclassical consumer demand theory must therefore incorporate some sort of frictions to barter exchange of the sort assumed by Claim 3 above. Two means of introducing such frictions are widely deployed in the modern Metallist literature, corresponding to relaxations of the two main assumptions of a centralised market and perfect information that support Arrow and Debreu's (1954) results.

The first, and older, means of introducing frictions to barter exchange is in the form of specialisation in production and consumption, and the geographical dispersion of trading agents. The Walrasian assumption of a single, centralised market for the exchange of commodities is replaced with more realistic assumptions reflecting the various geographical obstacles that typically exist to the conduct of trade. Thus, Kiyotaki and Wright (1993), for example, motivate their model of the micro-foundations of money as follows:

“Since the earliest writings of the classical economists it has been understood that the essential function of money is its role as a medium of exchange. The use of monetary exchange helps to overcome the difficulty associated with barter in economies where trade is not centralised through some perfect and frictionless market.”<sup>30</sup>

The second, and more recent, method of characterising the frictions inherent in barter exchange is as asymmetries of information between buyers and sellers. If commodities are available in different qualities, but buyers are unable perfectly to distinguish between them, then trade may be inhibited as a result of buyers' fears that they will be sold a 'lemon', in the sense pioneered by Akerlof (1970). On this characterisation, the emergence of a medium of exchange is seen as an informational tool that can obviate this problem of asymmetric information.

---

<sup>30</sup>Kiyotaki and Wright (1991), p.63.

Of course, these two characterisations of the potential absence of double coincidences of wants inherent in barter exchange are not mutually exclusive; and a number of recent contributions combine the specification of non-Walrasian trading environments with imperfect information in order to analyse the existence and properties of monetary equilibria. Nevertheless, we can usefully distinguish recent Metallist models of the micro-foundations of money into two classes: those which assume full information, and hence model the lack of a double coincidence of wants as due to a non-Walrasian market structure of some description; and those which represent the frictions inherent in non-monetary exchange primarily by the introduction of asymmetric information. We review these two classes of model in turn.

### Models with Full Information

**The Turnpike Model** An early and influential attempt formally to model lacks of double coincidences of wants in the "type" dimension – that is, of different commodities at a single point in time – was made by Townsend (1980). Townsend (1980) models physical trading frictions in terms of the spatial separation of agents, whose trading opportunities are therefore limited by the physical proximity of viable partners. An infinite number of trading agents are distributed over an infinite number of locations, with one agent at each location. All agents are identical in their preferences, which are defined by:

$$\sum_{t=0}^{\infty} \beta^t U(c_t) \tag{1.1}$$

where  $0 < \beta < 1$  is a time discount factor common across agents,  $U(\cdot)$  is an instantaneous utility function, and  $c_t$  is the consumption of the commodity at time  $t$ . Agents are heterogeneous in endowments, however, and are thereby divided equally into two types: type A agents, who are endowed with one unit of the single, homogeneous consumption commodity in the initial period,  $t = 0$ , and in each subsequent even-numbered period; and type B agents, who have no endowment at  $t = 0$ , but receive a one unit endowment in every subsequent odd period.

The analysis proceeds first by deriving the optimal consumption sequences for each type of agent in the case where the trading environment is in line with the Arrow and Debreu (1954) model of a complete, competitive market organised by a central auctioneer. In this case, a

Walrasian equilibrium exists in which agents of type A and B consume at constant rates  $c^A$  and  $c^B$  respectively, which satisfy Euler equations equating the ratios of the marginal utilities of consumption in the initial period and in all subsequent periods for each of the two types of agent:

$$\frac{U'(c_t^A)}{U'(c_\tau^A)} = \frac{U'(c_t^B)}{U'(c_\tau^B)} \quad \forall t, \tau > 0 \quad (1.2)$$

The introduction of physical frictions to trade is achieved by specifying that agents are spatially separated as follows. At  $t = 0$ , each type A agent is paired with a type B agent. In each subsequent time period, every agent travels to another location where it again meets a different agent of the other type. Since it is additionally assumed that no two agents ever meet twice, this environment is sometimes referred to as a ‘turnpike’ economy: individual agents can be pictured as moving in a single direction along a turnpike, never to revisit locations and therefore trading partners having met them once. In this new set-up, an absence of double coincidences of wants may exist in both the physical and temporal dimensions. Any solution to the latter is deliberately ruled out, however, by the assumption of anonymity associated with the turnpike structure: the repayment of credit is impossible if agents never meet twice, so that its contracting is impossible. Townsend’s (1980) analysis focuses instead on the lack of double coincidences of wants in atemporal exchange, and its potential mitigation.

It is immediately obvious that the optimal consumption allocations,  $c^A$  and  $c^B$ , cannot be attained by barter exchange in the turnpike trading environment. Townsend (1980) therefore asks whether the introduction of fiat money – conceived of as “a physical commodity, say, a piece of paper, which may be carried costlessly and used in exchange by the agents as they travel between islands”<sup>31</sup> – can eliminate the trading frictions that are to blame for this<sup>32</sup>. It is demonstrated that it cannot – but that there exist nevertheless ‘monetary equilibria’ (defined

---

<sup>31</sup>Townsend (1980), p.270.

<sup>32</sup>We simplify here for the purpose of a clear exposition. In fact, Townsend (1980) first asks whether the optimal sequences of allocations can be attained with the introduction of fiat money, and without a constant sequence of interventions by the ‘government’ (i.e. sequences of lump-sum taxes and corresponding transfers from one type of agent to the other). It is demonstrated that it cannot: any monetary equilibrium that supports the optimal sequences of allocations requires intervention. It is then asked whether there exist monetary equilibria that do not require intervention – and it is the results relevant to answering this question that are described above.

to be sequences of consumptions, money balances, and positive prices that both solve the social optimisation problem and satisfy market clearing conditions) in which the sequences of consumption allocations realised Pareto-dominate the sequences that would be generated in the autarkic situation without fiat money.

**The Search-Theoretic Approach** Townsend (1980) demonstrates, therefore, that physical trading frictions and the lack of double coincidences of wants to which they give rise, can be invoked to explain the demand for money within the framework of the classical theory of consumer demand. An obvious limitation of the model is, however, that it is not designed to investigate competition between alternative monies; and hence does not meet Menger's challenge of understanding how it is that, for example, the official, national money, rather than some other thing serves as money, and in the quantity that it does. In Townsend (1980), fiat money cannot be said to dominate potential alternatives (such as a consumption-commodity money) in a market equilibrium. Rather, there are no alternatives; so that the model is, in the words of Hiyashi and Matsui (1996) "rigged so that money is effectively the only means of payment"<sup>33</sup>. The extension of Townsend's (1980) basic model in order to analyse this central issue requires a more elaborate specification of the trading environment that allows a more complex pattern of physical trading frictions<sup>34</sup>. Such a specification is provided by the search-theoretic, or 'random matching' approach introduced by an influential series of three papers by Kiyotaki and Wright (1989, 1991, 1993). In these models, the high likelihood that there will not be a double coincidence of wants in the type dimension between two trading agents is modelled by specifying a trading environment in which each individual trader is randomly matched in each period to another trader, with whom there is some less than certain probability there will be a

---

<sup>33</sup>Hiyashi and Matsui (1996) p.111

<sup>34</sup>Some modern models, however, persevere with Townsend's (1980) turnpike environment. Engineer and Bernhardt (1991), for example, adapted Townsend's framework so that alternative commodities can compete non-trivially with the fiat money commodity as the medium of exchange. Hiyashi and Matsui (1996) further extended and generalised this model by making all goods potentially desired (rather than only agents' own and their immediate neighbours' commodities as in Engineer and Bernhardt (1991)) and including capital assets that compete with fiat money. Bullard and Smith (2000) extend the spatial separation framework to include the possibility of borrowing and lending in order to analyse inside and outside money – an extension parallel to the models of Cavalcanti and Wright (1999) and Williamson (1999) within the search-theoretic framework discussed below. Nevertheless, because they do not sustain possible equilibria in which spot barter exchange (i.e. a non-monetary equilibrium) dominates the use of either fiat or commodity money as medium of exchange, they are of less general scope than the search-theoretic models reviewed below.

double coincidence of wants in the type dimension.

Kiyotaki and Wright (1989) increase the number of types of agent from two to three, each of which has heterogeneous, rather than identical, preferences (though retaining the common time discount factor of the Townsend (1980) model). Moreover, they model a production economy, and each type of agent possesses a heterogeneous production technology with which they may produce a heterogeneous consumption commodity. There is, in other words, specialisation in both production and consumption. The preferences of each of the three types of agents preclude their consuming the commodities which they themselves are able to produce, so that each must exchange the commodity that it produces with that produced by another agent in order to consume. Storage of the goods is possible at a cost which varies over the three types of goods, but is strictly dominated by consumption insofar as that is possible: in other words, storage will only be chosen when agents possess their own produced commodity, and are not matched with a partner with whom trade is possible. Townsend's (1980) assumption of anonymity is retained, so that there is once again no credit in the economy; all exchange is conducted by spot barter transactions, and potential lacks of double coincidences of wants in time are abstracted from. Once again, the derivation of the optimal sequence of allocations is the starting point for the analysis of this set-up. Analogously to the Townsend (1980) model, the specification of an Arrow and Debreu (1954) trading environment would generate a Walrasian equilibrium in which every agent would produce and consume a constant quantity (one unit) of his production and consumption commodities in every period – this outcome being achieved by centralised spot barter exchange.

Physical trading frictions are then introduced in a more complex manner than in the Townsend (1980) model, facilitated by the extended initial set-up. Instead of participating in a centralised market, agents are randomly matched with one another in pairs in every period, and must decide whether or not to trade bilaterally with their current counterpart. The behaviour of each type of agent is then characterised as the choice of a trading strategy: a rule determining the circumstances under which agents of that type will choose to trade. This rule is chosen by each type of agent so as to maximise net present expected utility from consumption, and on the basis of the strategies of the agents of the other two types and the distribution of the potential random matches. An equilibrium is defined as a set of three trading strategies (one for

each type of agent) which, for a given distribution of potential random matches, maximises net present expected utility, and satisfies rational expectations<sup>35</sup>. Since the number of commodities is small, it is feasible to identify equilibria of this sort using a simple algorithm for given specifications of agents' heterogeneous production technologies and consumption preferences.

This set-up, with relatively minor modifications, is used by Kiyotaki and Wright (1989, 1991, 1993) to investigate the conditions under which one or more commodities are chosen in equilibrium as a medium of exchange: that is, they are held by agents not for consumption purposes but in order to exchange for other commodities when the opportunity arises. In Kiyotaki and Wright (1989), for example, it is shown that under a very simple specification of production and consumption specialisations a unique equilibrium exists, in which one of the three commodities is chosen, on the basis of its intrinsic quality of having the lowest storage cost, to serve as a commodity money<sup>36</sup>. Alternative specifications result in the existence of multiple equilibria involving two commodities serving as media of exchange simultaneously.

The models also investigate equilibria in which 'fiat money' is chosen to serve as the (or a) medium of exchange. As in the Townsend (1980) model, 'fiat money' is represented in this framework by the introduction of an additional fixed endowment of a commodity which is not produced by any of the trading agents and from the consumption of which no type of agent derives utility. Within the random matching framework, the extrinsic beliefs of the trading agents concerning the acceptability in exchange of this new commodity become critical to its role:

"If no one believes in fiat money, then it cannot get off the ground (like the Susan B. Anthony dollar). This 'tenuousness' of fiat currency is shared by the overlapping generations model, although not by the cash-in-advance model, and we think that it is a property that a good theory of money ought to have. The value of any medium of exchange, and especially fiat money, ultimately depends at least

---

<sup>35</sup>In fact, equilibria are further divided into *fundamental* and *speculative* equilibria. The former are equilibria in which agents always prefer to hold a lower rather than a higher storage cost commodity (in the case that they hold commodity money, rather than acquire the commodity in order to consume it); the latter, equilibria in which agents sometimes trade a lower for a higher storage cost commodity, in the rational expectation that it will be more exchangeable for the commodity they wish to consume at some later date.

<sup>36</sup>Or, under a slightly modified specification, that a unique equilibrium exists in which two of the three commodities co-circulate as commodity money.

partially on faith.”<sup>37</sup>

Kiyotaki and Wright (1989) demonstrate that if all agents share the extrinsic belief that the new commodity will be accepted in exchange, then it will serve as the unique commodity money in equilibrium so long as it has intrinsic properties (i.e. in the basic set-up, a storage cost) more attractive than those of the consumable commodities. In Kiyotaki and Wright (1991), the conditions under which fiat money will be chosen as the unique medium of exchange in equilibrium even when it is dominated by some or all consumable commodities in terms of its intrinsic qualities are derived<sup>38</sup>.

### **Full Information Metallist Models: Assessment**

These results of the Kiyotaki and Wright (1989, 1991, 1993) search-theoretic models make a significant advance on those of the Townsend (1980) spatial separation model, in terms of investigating the micro-foundations of the Metallist theory of money described in Section 1.2.1 above. They demonstrate that the classical theory of consumer demand, when coupled to a trading environment with a sufficiently complex range of physical frictions, can explain the endogenous choice by rational, expected utility-maximising agents of one commodity as a medium of exchange as a result of the interaction of (a) trading agents’ specialised consumption preferences and production technologies; (b) the different intrinsic qualities of different commodities; and (c) trading frictions in general, and specifically, the need for a double coincidence of wants in the type dimension in order for spot barter exchange to take place. In addition, these models show that fiat money – in the sense of a commodity neither produced nor consumed by trading agents – may be chosen as the medium of exchange in equilibrium on the basis of the above three factors, and in addition (d) agents’ extrinsic beliefs concerning the acceptability of this fiat money by one another.

On the other hand, two important criticisms may be made of the full information search-

---

<sup>37</sup>Kiyotaki and Wright (1989) p.943. The Susan B. Anthony dollar was a dollar coin introduced in the U.S. in 1979. It was widely rejected by the general public because of its close similarity in size and weight to the quarter, and was therefore withdrawn from circulation after a short space of time.

<sup>38</sup>Since the cost of storage can be interpreted as the inverse of the rate of return to holding each commodity, the results of Kiyotaki and Wright (1991) are in turn interpreted as meeting what Hicks (1935) posed as the most important challenge to a theory of fiat money: demonstrating why agents are willing to hold fiat money when it is absolutely dominated in rate of return by interest-bearing financial assets.

theoretic approach and thereby of the results derived from it. The first of these involves its deliberate deviation from competitive, Walrasian foundations. Since orthodox, macroeconomic models are generally built on these foundations, there exists the significant drawback that the implications of the search-theoretic models of the micro-foundations of money for the more policy-relevant results derived from those models cannot be investigated<sup>39</sup>. Second – and more critically for the assessment of the models purely as contributions to the understanding of the micro-foundations of money – the essential role played by extrinsic beliefs in the models begs the question of how far the models in fact advance our understanding of why fiat money is valued and circulates. In the full information search-theoretic models reviewed so far, the existence of equilibria involving fiat money relies on agents sharing the ‘right’, extrinsic beliefs as to the future acceptability of the fiat (commodity) money. In other words, this modelling tactic merely pushes the issue of why fiat money is valued and circulates further back, onto the question of how and why agents come to share these ‘right’ beliefs. Since this aspect of agents’ behaviour is exogenously specified – the beliefs are extrinsic – this question cannot be analysed within the model. No matter how far the intervening analysis takes us, in other words, it transpires that these models in the Metallist tradition are still unable to answer the fundamental question of how the market makes its money.

### **Models with Asymmetric Information**

**The Search-Theoretic Approach** As discussed above, an alternative strategy for modelling the trading frictions that make money useful has it that they are not physical but informational. The fundamental hypothesis of this alternative strategy – that it is money’s ability to overcome problems of information asymmetry that leads to its existence in a market equilibrium – was first put forward by Alchian (1977). In Alchian’s (1977) model, it is argued that the habitual

---

<sup>39</sup>Kiyotaki and Wright (1993) recognise and attempt to meet this criticism by demonstrating how a number of welfare and other applied issues can be explored, albeit at a high level of abstraction, using extensions of the versions of their earlier random matching models that include ‘fiat money’. Given, however, that these models say nothing about the price level (i.e. the relative price of the fiat money commodity and the consumable commodities), for example, the capacity of the framework to analyse applied questions necessarily remains limited. The search-theoretic models of Shi (1995) and Trejos and Wright (1995) go some way to addressing this limitation of the approach, constructing search-theoretic models in which the produced commodities are divisible, and modelling explicitly the bilateral bargaining that can then take place and determine price levels endogenously. The results of these models allow conclusions to be drawn regarding the relative merits, in terms of welfare, of trade in such environments with and without fiat money.

lack of a double coincidence of wants is at most a marginal factor in the existence of a medium of exchange. Instead, it is the fact that the quality of a traded commodity is to a greater or lesser extent private information – that there is typically, in other words, an asymmetry of information between its seller and a potential buyer – which is decisive:

“Costs of identifying qualities of a good are what counts. If costs for some good are low and generally low across members of society, the good will become a medium through which information costs can be reduced and exchange made more economical.”<sup>40</sup>

This alternative conception of the trading frictions leading to the use of a medium of exchange was incorporated into the search-theoretic modelling framework by Williamson and Wright (1994). The basic set-up of this model is very similar to that of the Kiyotaki and Wright (1989, 1991, 1993) models discussed above, with the following important exceptions. First, rather than specialisation in production and consumption – that is, the assumption of a number of heterogeneous commodities, each produced and consumed by a different type of agent – Williamson and Wright (1994) assume a single consumption commodity that can be produced and consumed by any agent. Moreover, all agents are assumed to be homogeneous in their consumption preferences. Taken together, these assumptions rule out the possibility of the lack of double coincidence of wants during the random matching of agents that was the core of the models surveyed in the previous section. With only this adjustment to the Kiyotaki and Wright (1989, 1991, 1993) framework, therefore, the model will generate a non-monetary (spot barter) equilibrium in which all agents produce, trade, and consume in every period. It is assumed in addition, however, that the single consumption commodity in the Williamson and Wright (1994) model may be of high or low quality – the former incurring a positive cost of production, but then yielding a positive utility in consumption; the latter costing nothing to produce, but also yielding zero utility in consumption. The assumption that agents cannot consume their own production is maintained from the Kiyotaki and Wright (1989, 1991, 1993) framework, again in order to rule out equilibria without trade. Since the seller of a commodity knows its quality with certainty, whereas a potential buyer will recognise its quality with

---

<sup>40</sup> Alchian (1977).

a probability strictly less than unity, there is a problem of informational asymmetry of the sort first analysed in Akerlof (1970). This represents a source of disincentives to the first best equilibrium pattern of trade between agents alternative to the lack of a double coincidence of wants.

Analogously to the models reviewed in the previous section, fiat money is investigated in this model via the introduction of a third commodity that can be neither produced nor consumed by the trading agents. The important additional quality of fiat money in this framework, however, is that it is always identifiable: unlike the consumption commodity, information about the fiat money commodity is public. Defining equilibria as in the Kiyotaki and Wright (1989, 1991, 1993) framework discussed above, and restricting attention to so-called “active” equilibria only (that is, equilibria in which there is at least some production and trade of the consumption commodity), Williamson and Wright (1994) analyse what equilibria exist under different assumptions concerning the degree of privateness of information regarding the quality of the consumption commodity. It is demonstrated that when the private information problem is severe, the introduction of the fiat money commodity can generate active equilibria where the only non-monetary equilibria are degenerate. Since this means that under such specifications no trade would take place without the fiat money commodity serving as a medium of exchange, this result is interpreted to demonstrate that asymmetries of information can explain the demand for fiat money in equilibrium.

Introducing assumptions concerning asymmetries of information, then, is another means of generating equilibria in which one commodity, on the basis of its desirable intrinsic properties (in this case, ease of identification), circulates as a medium of exchange. This approach has also allowed the search-theoretic framework to be applied to the explanation of financial intermediaries, and more precisely to the existence of and distinction between inside and outside money. The Metallist theory originated in the analysis of specie currency – the archetype of outside money – and the search-theoretic models reviewed so far, are all, implicitly, models of outside money. Given, however, that there was already a significant theoretical tradition arguing that informational asymmetry is an important reason for the existence of banks<sup>41</sup>, the introduction

---

<sup>41</sup>Classic analyses of the role of financial intermediaries in terms of information asymmetries include Leland and Pyle (1977), Diamond (1984), Broecker (1990), etc.

of the concept of fiat money as arising out of a need to mitigate asymmetries of information naturally invited the extension of the search-theoretic framework to analyse the role of banks and inside money. A 1999 special issue of the *Journal of Money, Credit and Banking* devoted to such extensions presented two models – Cavalcanti and Wright (1999) and Williamson (1999) – that deploy information asymmetries to introduce inside money into search-theoretic models.

It is important to emphasise that in these models inside money – although it is interpreted to represent privately issued financial liabilities – is in fact modelled as privately produced commodities, rather than as financial contracts. In Cavalcanti and Wright (1999), the differences in information relate to the trading histories of agents – the information regarding the trading histories of ‘banks’ being public, whilst those of ‘non-banks’ are private. In Williamson (1999), the search-theoretic approach is combined with Townsend’s (1980) spatial separation approach. This allows the modelling of banks (in a separate sector) not only as traders with public trading histories, but as genuine collectors and diversifiers of traders’ invested commodities – inter-temporal trading is introduced, although there is no modelling of commitment (it is assumed that full, multilateral commitment by the banks is possible). This set-up is analysed with either full or private information concerning the quality of the banks’ portfolios, however – with the assumption of private information able to generate an equilibrium where private (inside) money does not circulate.

**The Walrasian Approach** The models surveyed above start from trading frictions generated by either random matching or spatial separation, upon which additional frictions associated with asymmetric information may or may not be imposed<sup>42</sup>. As noted above, however, one major problem with relying on the assumption of a non-competitive trading environment to derive the use of a medium of exchange, is that such a framework is at odds with the micro-foundations of conventional macroeconomic models of a broader scope. The recourse to search theory or spatial separation generally represents a violation of the assumption of competitive, Walrasian markets that constitutes the basis of the majority of these models<sup>43</sup>. As Williamson and Wright

---

<sup>42</sup>Although in some cases, as for example in Williamson and Wright (1994), the frictions imposed by the random matching framework are in the event negated by assuming that there is no specialisation in production or consumption.

<sup>43</sup>Of course, by the same token, these models generally treat money in an unconvincing way: either via money in the utility function, or the imposition of an exogenous Clower constraint.

(1994), however, demonstrate, the assumption of information asymmetries precisely liberates the modeller from the need to specify a non-competitive market environment in order to derive monetary equilibria. This result therefore opens up the prospect of models with Walrasian micro-foundations and imperfect information which generate monetary equilibria.

Just such a model is set out by Banerjee and Maskin (1996). Unlike Williamson and Wright (1994), Banerjee and Maskin (1996) argue explicitly that focussing on the double coincidence of wants problem is in fact inappropriate to great swathes of modern, advanced economies, in which there are well organised markets that are well-approximated by the Arrow and Debreu (1954) competitive paradigm. On the other hand, they argue that any commodity, even in competitive markets, may be subject to two types of uncertainty as to its value: the first due to imperfect knowledge of its quality (the classical Akerlof (1970) 'lemons' problem), and the second due to imperfect knowledge of the future market for the commodity if the buyer were to wish to re-sell it at a later date.

Banerjee and Maskin (1996) therefore specify a model in which a continuum of trading agents produce and consume a variety of commodities, which may be of either high or low quality. As in the Kiyotaki and Wright (1989, 1991, 1993) framework, there is specialisation in production and consumption: but since commodities are now traded in perfectly competitive markets organised by a Walrasian auctioneer, the impossibility of strictly bilateral exchange involving a coincidence of wants is obviated. It is assumed that all trade comprises anonymous exchange – in other words, as in the models above, that there is no opportunity for the development of securities markets – and takes place in separate, perfectly competitive markets for each of the commodities. It is demonstrated that within this framework that commodity with the least discrepancy in quality between its high and low quality variants will be chosen in equilibrium to serve as the unique medium of exchange. A welfare analysis of different equilibria once again allows the appraisal of fiat money and commodity money equilibria. Since the money commodity is also potentially a consumption commodity, the quantity of it produced in the monetary equilibrium will be in excess of the social optimum measured purely in terms of optimal consumption demand. As a result, an alternative 'fiat money' – defined as a commodity with a zero cost of production but limited supply which is always identifiable by all agents – will not only be chosen in equilibrium as the medium of exchange; but this equilibrium will

dominate any alternative, consumption commodity money equilibrium.

### **Metallist Models with Asymmetric Information: Assessment**

These models that adopt the assumption of asymmetric information concerning the quality of commodities therefore permit the analysis of the choice of money not only within the framework of the classical theory of commodity demand – as was the case for the full information models reviewed above – but also without the need to specify non-Walrasian trading environments. The core Metallist hypotheses discussed in Chapter 2 above are therefore maintained, and micro-foundations for them are elaborated on the basis of money serving to alleviate trading frictions associated with private information. This modelling strategy based on information asymmetries is open to its own criticisms, however: two in particular should be noted. First of all, there is a conceptual issue. The critical factor that generates trading frictions in these models, and therefore the critical factor in generating a demand for money, is the degree of public knowledge concerning the quality of different commodities. Can this reasonably be said to be an intrinsic quality of the commodities themselves, as the Metallist theory requires, rather than a product of the social environment in which the commodities are traded? It may easily be argued that it cannot; and if this point of view is adopted, then the models described above can be accused of assuming, rather than analysing, the basis on which monetary equilibria rests, in a manner similar to the reliance on extrinsic beliefs of the full information Metallist models surveyed above.

The second criticism of these Metallist models which explain money as a means of overcoming problems due to asymmetric information is that the empirical support for this, their central device, is weak. As discussed in Section 1.2 above, the entire notion of money ‘arising’ out of barter exchange – and hence the entire enterprise of modelling the trading frictions that lead to the emergence of a medium of exchange, be they characterised as the lack of a double coincidence of wants in either the type of the time dimensions, or as asymmetries of information – is highly questionable on historical and empirical grounds. In terms of explicit criticism of the modern models surveyed above, however, the deployment of asymmetric information to explain trading frictions has attracted even more scepticism. The lynchpin of Alchian’s (1977) original model of money as an instrument to overcome the informational problems of barter exchange is

the existence of one commodity for which the costs of identification are sufficiently low: specie is supposed to have been just such a commodity in the historical emergence of money. As Goodhart (1989) demonstrates, however, all evidence points to the fact that the informational difficulties of using unminted specie as money have always been considerable. The result is that:

“the costs of identifying the quality of either unworked or fabricated precious metal for the ordinary person is high. An individual could, of course, go to a money-changer for expert advice, but that would also involve costs. So, such costs were probably higher, for example, than the cost of identifying the value of items in common everyday use, e.g., salt, corn, nails or even perhaps cattle, (most people in a rural community would reckon to be able to assess the value of a cow).”<sup>44</sup>

Yet it is specie that became money on the Metallist view. The imperfect information characterisation of the frictions inherent in barter exchange is therefore not altogether convincing. Nevertheless, Banerjee and Maskin (1996) also attempt to interpret their derivation of monetary equilibria based on asymmetries of information as according with the historical record:

“This finding seems to accord with the evolution of gold and certain other metals – particularly when used for coins – as widespread media of exchange. Historically, two innovations were important to these metals’ success: Archimedes’ specific gravity test and the serrated edge. Both inventions, in effect, reduced variation in unobservable quality: the specific gravity test by making it hard to pass off base metal as gold or silver, and the serrated edge by defeating the practice of “coin-clipping”. Thus, our theoretical finding can be thought of as a formal explanation for the historical prevalence of gold as a medium of exchange.”<sup>45</sup>

This elaboration succeeds only in plumbing new depths of silliness in the field of historical speculation in support of the Metallist theory: Archimedes was born in 287 B.C., by which point specie had already been ‘successful’ – in the sense of having been used for minted coinage

---

<sup>44</sup>Goodhart (1998), p.411.

<sup>45</sup>Banerjee and Maskin (1996), p.958.

– for at least three hundred years (the earliest extant examples being Lycian coins of the seventh century B.C)<sup>46</sup>.

### 1.3.2 Models in the Credit Tradition

In Section 1.2 above, we concluded that where the Metallist theory conceives of money as a commodity which is chosen as a medium of exchange, the Credit theory conceives of money as transferable credit and regards the concept of a medium of exchange as spurious. As for its explanation of the choice between monies, the Credit theory identifies as the core criteria the relative capacities of different money-issuers to make binding unilateral commitments to accept their liabilities back in settlement of some outstanding debt. The conception of money as credit immediately makes a new demand on models in the Credit tradition: that they model exchange not only in the ‘type’ dimension – that is, of different commodities at a single point in time – but the ‘time’ dimension – that is, of the same commodity at different points in time.

By reasoning analogous to that deployed above in the case of Metallist models, Credit theoretic models of money must moreover incorporate trading frictions in the time dimension: if there were no such frictions, an equilibrium intertemporal allocation could be attained simply by the continuous contracting of bilateral (i.e. non-transferable, and hence non-monetary) credit between agents. To the possibility of intertemporal trade must be added constraints on agents’ capacities to strike fully credible intertemporal contracts, so that the absence of double coincidences of wants is possible in the time, as well as the type, dimension.

By contrast to models in the Metallist tradition, examples of models of money in the Credit tradition are relatively rare and fairly heterogeneous. Townsend (1980) and Townsend (1989) are two early departures from the Metallist tradition in the modern theoretical literature. These “communications accounting” models incorporate a number of important elements close to those of the Credit theory – eschewing the concept of money as a commodity medium of exchange, for example; and generating not only intertemporal trade in commodities, but, on one interpretation, the transfer of debts from one creditor to another – but their focus is firmly on real sector trade with frictions introduced by asymmetries of information, rather than financial

---

<sup>46</sup>See Seaford (2004) for an up-to-date discussion of the history and anthropology of money in the ancient Greek and Near Eastern worlds.

contracting *per se*; and the extent to which they illuminate the micro-foundations of the Credit theory is therefore limited. Likewise, Shubik (2001) and Shubik (2005a, 2005b) treat a number of Credit theoretic themes; but the former is intended only as an informal overview, whilst the main topic of the latter is much broader than the micro-foundations of money alone, taking in as it does the nature of social institutions and the natural scale of government and bureaucracy. One important recent contribution which undertakes a significant and detailed analysis of the micro-foundations of money in the Credit theoretic tradition, however, is Kiyotaki and Moore (2001, 2002a, 2002b) – and it is this model that we therefore review in some detail here.

### **The Kiyotaki and Moore (2001, 2002a, 2002b) Model**

**Initial set up** The Kiyotaki and Moore (2001, 2002a, 2002b) model has in common with the Metallist models reviewed above that the mechanism whereby the choice between monies is made is modeled as a decentralized market process: but unlike those models the environment for atemporal exchange is assumed to be perfectly competitive - there is neither specialisation and spatial separation nor asymmetric information. The model also incorporates intertemporal trade in dated goods, however, and here, frictions are introduced in the form of limited commitment: borrowers who issue debt are unable to pledge all of their future income, and previously-issued debt securities are less than fully re-saleable. As Kiyotaki and Moore (2001, 2002a, 2002b) put it:

“What we are doing . . . is to recast the classic idea of a lack of coincidence of wants from the type dimension to the time dimension.”<sup>47</sup>

“We [argue] that money lubricates trade when there is a lack of double coincidence of wants, but not necessarily over physically distinct commodities. The great advantage of switching to dated goods, and not having to model physical trading frictions, is that one can breathe the pure oxygen of perfect competition. The terms of trade for paper, both new and second-hand, are determined in the market place.”<sup>48</sup>

---

<sup>47</sup> Kiyotaki and Moore (2001), p.8.

<sup>48</sup> Kiyotaki and Moore (2002a), p.65.

The model begins from a deterministic, discrete time economy with a continuum of identical, infinitely-lived agents who produce and consume a single homogeneous commodity. The consumption preferences of agents are summarized by a utility function similar to Equation 1.1 above:

$$\sum_{s=0}^{\infty} \beta^s U(c_{t+s}) \quad (1.3)$$

Where  $0 < \beta < 1$  is again a discount factor common to all agents,  $U(\cdot)$  is an instantaneous utility function, and  $c_{t+s}$  denotes consumption in period  $t + s$ . Agents have access to a decreasing-returns-to-scale production technology by which  $x$  units of the commodity can be invested at time  $t$  to produce  $y$  units two periods later, according to:

$$y_{t+2} = \left( \frac{x_t}{a(1-\lambda)} \right)^{1-\lambda} \quad (1.4)$$

where  $0 < \lambda < 1$ . The associated cost function,  $G(y)$ , may therefore be defined as:

$$x = a(1-\lambda)y^{\frac{1}{1-\lambda}} \equiv G(y) \quad (1.5)$$

Storage of the commodity is also possible, at zero cost, such that one unit stored at time  $t$  remains one unit at time  $t + 1$ .

Normalising the number of agents to three, and given the assumption of a perfectly competitive trading environment, the first best, symmetric, steady state allocation is simply characterized. Agents invest at a constant level  $x^*$  to produce  $y^*$ , where  $x^* = G(y^*)$ . Consumption is constant and equal across agents at  $c^* = 1/3(y^* - x^*)$ . The marginal cost of production is equalized with the discounted marginal return on investment:

$$G'(y^*) = \beta^2 \quad (1.6)$$

There are, however, two constraints on agents' contracting. First, at the start of a new investment project, agents can commit to share only a proportion,  $\theta_1$  of its future output. In other words, they can only issue up to  $\theta_1 y_{t+2}$  units of 'two-period paper' at time  $t$ , where 'two-period paper' is interpreted as a credible promise to deliver one unit of the commodity in two periods' time.  $\theta_1$  is therefore a constraint on transactions in the primary market for

credit. Second, any agent who purchases two-period paper at time  $t$  can sell only a fraction,  $\theta_2$ , of it at time  $t + 1$ : he must wait until time  $t + 2$  in order to redeem the remaining fraction,  $(1 - \theta_2)$ , with its original issuer.  $\theta_2$  is therefore a constraint on the secondary market for credit. The value of  $\theta_1$  can be taken to summarise, in other words, a credit rationed equilibrium that arises from informational imperfections of the type explored by models of the primary market for credit in the tradition of Stiglitz and Weiss (1980). The second constraint, by contrast, has to do with the secondary market for credit, which is not analysed in such models. Kiyotaki and Moore (2002a) explain:

“The two constraints need to be thought about separately. The first constraint, the borrowing constraint, has received attention in the macroeconomics literature. A number of moral hazard stories have been invoked to rationalize why people face borrowing constraints. The second constraint, the resaleability constraint, has received much less attention in the formal literature, but we think it is just as important.”<sup>49</sup>

Defining  $q_t$  to be the price at time  $t$  of two-period paper redeemable at time  $t + 2$  (i.e. the price of a claim to one unit of commodity at time  $t + 2$ , in terms of commodity at time  $t$ );  $n_t$  to be agents’ holdings of two-period paper; and  $z_t$  to be the quantity of commodity stored at time  $t$ ; we can write out the flow of funds constraints that apply to any given agent at each of the three distinct periods in the economic cycle, as follows:

$$x_t + c_t + z_t + q_t n_t = q_t \theta_1 y_{t+2} + z_{t-1} + n_{t-2} \quad (1.7)$$

$$c_{t+1} + z_{t+1} + q_{t+1} n_{t+1} = z_t + n_{t-1} \quad (1.8)$$

$$c_{t+2} + z_{t+2} + q_{t+2} n_{t+2} = (1 - \theta_1) y_{t+2} + z_{t+1} + n_t \quad (1.9)$$

Competitive equilibrium is then defined as a sequence over  $t$ ,  $t + 1$ , and  $t + 2$  of the variables  $q_t$ ,  $x_t$ ,  $y_{t+2}$ ,  $c_t$ ,  $n_t$ , and  $z_t$  such that at time  $t$ :

a)  $x_t$ ,  $y_{t+2}$ ,  $c_t$ ,  $n_t$ , and  $z_t$  are chosen by agents so as to maximize utility subject to (i) the production function, and (ii) the three flow of funds constraints above; and

---

<sup>49</sup> *Ibid.*, p.63.

b) the markets for the commodity and for two-period paper both clear.

Since the focus is on symmetric, steady state equilibria, in which a third of the population of agents invests in each period, each agent's choices are made in an identical three period pattern, and the price  $q_t$  is constant over time, it is helpful to re-formulate the notation for these variables relative to any given agent, rather than to time.  $q$  is then the steady state price of newly-issued two-period paper;  $x$  and  $y$  denote steady state investment and output respectively;  $c$  is the consumption,  $n$  the investment in new two-period paper, and  $z$  the storage, of agents when they are investing;  $c'$ ,  $n'$ , and  $z'$  measure the same quantities in the middle of the production cycle; and  $c''$ ,  $n''$ , and  $z''$  these quantities in the output period. The flow of funds constraints can then be written as:

$$x + c + z + qn = q\theta_1 y + z'' + n' \quad (1.10)$$

$$c' + z' + qn' = z + n'' \quad (1.11)$$

$$c'' + z'' + qn'' = (1 - \theta_1)y + z' + n \quad (1.12)$$

The condition that must be met for the commodity market to clear is:

$$y = c + c' + c'' + x \quad (1.13)$$

And that for the two-period paper market is:

$$\theta_1 y = n + n' + n'' \quad (1.14)$$

Kiyotaki and Moore (2002b) analyse this set-up in two polar cases: the first in which  $\theta_2 = 0$  – the case of zero resaleability, where an agent who has purchased two-period paper at time  $t$  has no option but to wait until time  $t + 2$  to redeem it with its original issuer; and the second, in which  $\theta_2=1$  – the case of full resaleability, in which two-period paper issued by agent A and purchased by agent B at time  $t$  may be freely sold to agent C at time  $t + 1$ . We summarise their results in each of these two cases in turn.

**The case of no multilateral commitment ( $\theta_2 = 0$ )** We begin with the case in which  $\theta_2 = 0$  - two-period paper is not resaleable. In this case the characteristics of the symmetric, steady state equilibrium depend solely on the value taken by  $\theta_1$  - the borrowing constraint in the primary market for two-period paper. If agents' capacity to make bilateral commitments,  $\theta_1$ , is greater than a critical value,  $\theta_1^*$ , then it transpires, somewhat surprisingly, that the first best equilibrium characterized above is nevertheless still attainable - despite the fact that there is limited commitment. This is achieved by agents' operating two staggered, but unconnected, tracks, each of one cycle's investment in production followed by two consecutive cycles' investment in paper: "the economy's ingenious response to the constraint imposed by the non-saleability of paper, and the need to create double coincidences of wants"<sup>50</sup>. This critical value is defined as follows:

$$\theta_1^* = \frac{1}{3} \left( 2 - \frac{\lambda\beta^2(1 + 2\beta^2)}{1 + \beta^2 + \beta^4} \right) \quad (1.15)$$

- a formidable-looking expression. If, however, we set the value of  $\beta$  at approximately unity and that of  $\lambda$  at approximately zero (implying low time discounting by agents, and a production technology that exhibits near constant returns to scale), then this critical value,  $\theta_1^*$ , will equal approximately 2/3. So long as  $\theta_1 > \theta_1^*$ , so defined, then the shadow price of one period paper implicit in the actual price of two-period paper,  $\sqrt{q}$ , equals the rate of one period time preference,  $\beta$ ; and the first best equilibrium levels of consumption, savings, investment, and output are maintained.

If, however,  $\theta_1 < \theta_1^*$ , then  $\sqrt{q} > \beta$  - the shadow price of one period paper is higher than agents' time preference. Relative to the first best equilibrium, there is too little saving and investment: the constraints on borrowing are binding. Investment and output are lower than in the first best equilibrium, and consumption is not smoothed over the three periods, but instead highest in the output period and lowest at the investment period. Moreover, as  $\theta_1$  falls,  $c''/c'$ , and  $c'/c$  rise - making the path of consumption ever less smooth -  $\sqrt{q}$  rises further above  $\beta$ , and  $x$  and  $y$  fall further below the first best equilibrium values. In short, the economy 'runs too slowly', in inverse proportion to the ability of agents to make bilateral commitments. Two

---

<sup>50</sup>Kiyotaki and Moore (2000), p.15.

different patterns of agents' behaviour may be distinguished when  $\theta_1 < \theta_1^*$ , corresponding to two different regions along the support of  $\theta_1$ . In the first region, in which the value of  $\theta_1$  remains close enough to the critical value  $\theta_1^*$ , the structure of agents' saving and investment behaviour does not change from the first best equilibrium (even though its arguments are reduced). In the second region, however, in which  $\theta_1$  falls below a further threshold value,  $\tilde{\theta}_1$  (which is strictly lower than  $\theta_1^*$ ), storage is used as a means of short term saving. Storage becomes competitive with investment in two-period paper not on the basis of its own one period return (which, at zero, is necessarily dominated by the implicit one period return on two-period paper  $\sqrt{q}$ ), but because it affords the possibility of investing in the next period. As a result, in the region  $\theta_1 < \tilde{\theta}_1$ , agents pursue a two track strategy in which one of the tracks is as in the first best equilibrium – of one cycle's investment in production, followed by two tracks' investment in paper – but the other involves one cycle's investment in production followed by one period of saving via storage in preparation for investment in production again. Kiyotaki and Moore (2002b) call the former the 'slow', and the latter, the 'fast' track.

To sum up the results of the model when there is no multilateral commitment – when it is assumed that agents' paper is not resaleable in the secondary market. The character of equilibrium depends solely upon the capacity of agents for bilateral commitment – the degree to which borrowing constraints bind in the primary market. If the capacity for bilateral commitment is above a certain critical value ( $\theta_1^*$ ), then the first best equilibrium is still attainable. If it is below this critical value, then the first best is not achievable: the price of two-period paper is higher than agents' two period time discount; consumption is lower than in the first best, and is no longer smoothed; and investment and output are lower than in the first best. Depending on how far below  $\theta_1^*$  is agents' capacity to make bilateral commitments, the pattern of agents' saving and investment behaviour is different. If  $\tilde{\theta}_1 < \theta_1 < \theta_1^*$ , then agents continue to operate two, staggered but identical cycles. If, however,  $\theta_1 < \tilde{\theta}_1$ , then agents operate both a 'fast' and a 'slow' track – in the former of which, storage is used as a means of short term saving. None of these equilibria could be interpreted as 'monetary'.

**The case of full multilateral commitment ( $\theta_2 = 1$ )** Let us now turn to the case in which  $\theta_2 = 1$  – where two-period paper is perfectly resaleable. On this new assumption, agents are

able to make not only a bilateral commitment to deliver commodity in two periods' time to the original purchaser of its paper (to an extent limited by  $\theta_1$ ), but a full, multilateral commitment to do so to any other agent that acquires the claim in the meantime. In this case, storage is strictly dominated by investment in two-period paper as a means even of short term saving, since paper can be resold with certainty in the secondary market at time  $t + 1$ . The flow of funds constraints can therefore be rewritten without the terms for storage as:

$$x + c + pn = p\theta_1 y + n'' \quad (1.16)$$

$$c' + pn' = n \quad (1.17)$$

$$c'' + pn'' = (1 - \theta_1)y + n' \quad (1.18)$$

where  $p$  is the price in the secondary market at time  $t$  of two-period paper issued at time  $t - 1$ . The condition that must be met for the commodity market to clear is the same as before (Equation 1.13); but the condition for the clearing of the paper market becomes:

$$(1 + p)\theta_1 y = n + n' + n'' \quad (1.19)$$

Equation 1.19 summarises the extent to which the resaleability of two-period paper relaxes the supply constraint in the paper market, increasing its liquidity in inverse proportion to the secondary market price.

As in the case where there was no multilateral commitment, the character of equilibrium depends on the severity of the constraint on bilateral commitment in the primary market. Only two regions need be distinguished under these revised conditions, however. A competitive equilibrium now achieves the first best symmetric allocation so long as  $\theta_1$  is greater than  $\theta_1^{**}$  defined as:

$$\theta_1^{**} = \frac{1}{3} \left( 2 - \frac{\lambda\beta^2(1 + \beta)}{1 + \beta + \beta^2} \right) \quad (1.20)$$

If  $\theta_1$  is greater than this critical value  $\theta_1^{**}$ , then  $p = \beta$  just as  $\sqrt{q} = \beta$  in the case where  $\theta_1 > \theta_1^*$  above, and the optimal investment and consumption schedules are attained since  $G'(y^*) = \beta^2$ , and  $c' = c'' = c''$ . If  $\theta_1 < \theta_1^{**}$ , however, the first best equilibrium allocation cannot be attained, and there is a severe liquidity shortage, in the sense that agents' equilibrium behaviour is similar

to that where  $\theta_2 = 0$  and  $\theta_1 < \tilde{\theta}_1$ , except that the fast strategy alone is adopted.

**Comparison of the two cases** What can be gleaned from the comparison of these two sets of equilibria, with and without multilateral commitment? The threshold  $\theta_1^{**}$  can readily be compared with  $\theta_1^*$  above by again setting the value of  $\beta$  at approximately unity and that of  $\lambda$  at approximately zero. Where  $\theta_1^*$  was equal to approximately  $2/3$  at these values,  $\theta_1^{**}$  will equal approximately  $1/3$ . This is a significant result: the higher is the degree of multilateral commitment possible in the economy, the less is the need for bilateral commitment in order that trade be sufficiently liquidated. As Kyotaki and Moore (2002a) have it: “a little multilateral commitment goes a long way”<sup>51</sup>. Of even more interest for the investigation of the micro-foundations of the Credit theory, however, is the mechanism whereby this result is generated. The difference between the two set-ups that reduces the requirement for bilateral commitment is the introduction of a secondary market for two-period paper. When  $\theta_2 = 1$ , any given agent can use two-period paper as a means of short-term saving, and thereby invest in production more regularly (and with a higher volume) than if he either relied on long term saving only, or resorted to storage for short term saving. When this agent sells two-period paper in the secondary market, the identical strategy is then being implemented one period later by an agent of a different type. In other words, it is the transfer of bilateral credit from its original creditor to a third party that serves to liquidate higher equilibrium volumes of investment, trade, and consumption.

**An extension** Thus far, the analysis of the interaction between the constraints on the primary and secondary markets for credit has so far been confined to the limiting cases where  $\theta_2$  is equal to either zero or unity, and both  $\theta_1$  and  $\theta_2$  are specified exogenously. A further step in the analysis is to extend the model so that the value of  $\theta_2$  – and hence the monetization of credit – is made endogenous. Kyotaki and Moore (2002b) present such an extension, in which agents may choose the degree of resaleability of their two-period paper by investing in an exogenously-specified ‘multilateral commitment technology’, at a cost which varies positively with the degree of resaleability secured. The value of  $\theta_2$  chosen by agents in this extended specification therefore

---

<sup>51</sup>Kiyotaki and Moore (2002a), p.65.

depends in part upon the constraint on borrowing in the primary market,  $\theta_1$ ; and equilibria, differing by the value taken by  $\theta_1$ , can again be derived and compared. The first best allocation is still attainable if  $\theta_1 > \theta_1^*$ . The argument is identical to the case in which  $\theta_2$  is restricted to equal zero discussed above: there is no shortage of liquidity ( $\sqrt{q} = \beta$ ); so the first best equilibrium is attainable without a secondary market for paper. There is therefore no incentive for agents to invest in the multilateral commitment technology in order to raise  $\theta_2$  above zero. If, however,  $\theta_1 < \theta_1^*$ , then we have not only that  $\sqrt{q} > \beta$ , as in the case where  $\theta_2$  was restricted to equal zero, but also that  $p > \sqrt{q}$ . There is a liquidity shortage, and in order to mitigate this, agents invest in the multilateral commitment technology up until the marginal benefits of doing so, in terms of the facilitation of short-term saving via investment in resaleable paper, are exactly offset by its costs. Assuming, therefore, that bilateral commitment is too limited to support the first best equilibrium (i.e.  $\theta_1 < \theta_1^*$ ), transferable credit is chosen in equilibrium in order to liquidate the economy<sup>52</sup>. This, it can reasonably be said, is a model in which the market makes its money; and does so by monetizing its credit.

### **Models in the Credit Tradition: Assessment**

In focusing specifically on intertemporal financial contracts between agents, and the ability of these contracts to liquidate trade between agents by being transferred from one creditor to another, the Kiyotaki and Moore (2001, 2002a, 2002b) model adopts the core Credit theoretic concepts of money and of the nature of monetary exchange. In common with the Metallist models reviewed above, the Kiyotaki and Moore (2001, 2002a, 2002b) model adopts a decentralized market process as the mechanism whereby the choice between moneys is made. The model makes a novel contribution to the analysis of the core criterion of the choice between monies postulated by the classical Credit theory – the relative capacity of a given money-issuer to make binding unilateral commitments – by distinguishing two constraints: the first on the issuer’s ability to issue debt in the primary market; and the second on his ability to transfer such debt to new counterparties in the secondary market. In these respects the Kiyotaki and

---

<sup>52</sup>An analogous result is much more easily derived if we do not restrict attention to a specification in which the two constraints are symmetric across all the agents. If, for example, the two-period paper of one of the agents is fully resaleable ( $\theta_2 = 1$ ), but that of the other two is fully illiquid ( $\theta_2 = 0$ ), then the first agent’s paper circulates as money and liquidates trade when  $\theta_1 < \theta_1^*$ . See Kiyotaki and Moore (2002a) for an exposition of this simple example.

Moore (2001, 2002a, 2002b) model represents a rigorous exploration of the micro-foundations of the classical Credit theory.

Nevertheless, just as the Metallist models reviewed above were subject to a number of limitations in their ability to advance our understanding of the Metallist theory, so the Kiyotaki and Moore (2001, 2002a, 2002b) model is subject to a number of limitations in its ability to advance our understanding of the Credit theory. Three in particular are important. The first limitation relates to the model's elaboration of the Credit theory's account of the criteria and mechanism whereby the choice between monies is made. Unlike the intrinsic qualities of the money-commodity identified as criteria by the Metallist theory, which are by definition exogenously determined, the relative capacity of a given money-issuer to make binding unilateral commitments identified by the Credit theory, and analysed by the Kiyotaki and Moore (2001, 2002a, 2002b) model into the constraints on his borrowing in the primary market ( $\theta_1$ ) on the one hand, and on his transferring already-issued credit to a new counterparty in the secondary market ( $\theta_2$ ) on the other, begs further explanation. What is it that determines the values taken by these two constraints? Without an answer to this question, we are none the wiser as to the determinants of the choice between monies: but the Kiyotaki and Moore (2001, 2002a, 2002b) model treats both  $\theta_1$  and  $\theta_2$  as formally exogenous in its simple version; and in its extended version has  $\theta_1$  specified exogenously, whilst  $\theta_2$  is endogenised – but at the price of introducing the device of a multilateral commitment technology, the parameters of which are themselves exogenously specified.

Kiyotaki and Moore (2001, 2002a, 2002b) hint in general terms at the theoretical structure that might underpin the determination of  $\theta_1$  and  $\theta_2$  when motivating the exogenous values ascribed directly or indirectly to these parameters. As to the constraint on borrowing in the primary market, for example, it is assumed in Kiyotaki and Moore (2000) that an investing agent's human capital is both essential to his investment projects and inalienable, so that it is impossible for him to agree to contracts at time  $t$  that commit him to work until time  $t+2$ . It is also assumed, however, that a creditor is given a project-specific 'key' at time  $t$  which is required for any output to be realized by the investing agent at time  $t+2$ . Both creditor and debtor are therefore required for projects to yield output, and "[a]fter bargaining, they end up dividing

the output in the ratio  $(1 - \theta_1); \theta_1$ .”<sup>53</sup> As for the constraint on resaleability in the secondary market,  $\theta_2$ , Kiyotaki and Moore (2002b) propose that the contract-theoretic foundations might take the shape of arguments based on adverse selection (for example, the original creditor has acquired superior information on the creditworthiness of the original issuer than potential secondary purchasers), moral hazard (for example, the original issuer and creditor collude to strip the assets of the project prior to resale of the claim<sup>54</sup>), or special leverage (for example, the original creditor is a trade creditor who therefore has greater leverage over the issuer than a creditor whose relationship with him is purely financial). These general suggestions are certainly useful indications as to possible focuses for future theoretical research (and in this capacity are discussed further below); but the Kiyotaki and Moore (2001, 2002a, 2002b) model itself is not equipped to analyse these ‘micro-foundations of the micro-foundations’.

The models’s second limitation also relates to its elaboration of the Credit theoretic explanation of the choice between monies, and derives from the its central, simplifying assumption of a single, homogeneous commodity produced and consumed by all agents. As noted above, the purpose of this assumption is to abstract from the possible absence of double coincidences of wants in the type dimension in order to make tractable the possible absence of double coincidences of wants in the time dimension. In effect, however, this rules out the case that is of principal interest for the analysis of competition between alternative moneys; that is, the case in which there are multiple issuers of monetary credit, each of whom produce different commodities. In this case, claims on different issuers that are held to maturity are settled in different commodities; so that if there is any positive probability of such claims having to be redeemed by their original issuer, the possibility of there being a lack of a double coincidence of wants in the type dimension must once again play a role in agents’ choices between monies. With multiple issuers, in other words, the potential lack of a double coincidence of wants in both the time and the type dimensions needs to be considered as criteria of the choice between monies.

---

<sup>53</sup>Kiyotaki and Moore (2000), p.10. This is an argument based on limited commitment; other arguments behind rationing in the primary market for credit might be made on the basis of asymmetric information in the vein of Stiglitz and Weiss (1980) as noted above. See Chapter 2 for further discussion of this point.

<sup>54</sup>A story that assumes the divertibility of credit, and should be compared to Burkart and Ellingsen (2004). There, since there are different types of credit (bank and trade), different degrees of divertibility can be distinguished. Here, of course, there is only one type of credit (trade credit), and such distinctions are not available.

Finally, the assumption of a single, homogeneous commodity also abstracts from another critical aspect of the Credit theory identified in Section 1.2.2 above: the explanation of the numeraire. Since there is only one commodity in the Kiyotaki and Moore (2001, 2002a, 2002b) model, the assumption that present and future prices are denominated in units of this commodity is adopted naturally. Once again, the introduction of multiple issuers differentiated by the commodities that they produce would, however, upset this: an additional explanation would be required as to the units of which commodity would serve as the numeraire. More fundamentally, however, the assumption that the units of any commodity should serve as the numeraire in the Kiyotaki and Moore (2001, 2002a, 2002b) model represents an important departure from the Credit theory, and an interpolation of a central postulate of the Metallist theory. As described in Section 1.2.2 above, since credit is a primitive concept in the Credit theory, and historically coeval with or even older than money (transferable credit), the origins of the numeraire are not to be identified with the unit of some money-commodity, but are instead to be located in some autonomous process whereby an abstract unit of value is established, whether by the diktat of the sovereign or by the agreement of the community of agents. Far from attempting to describe and explain such a process – far from modelling a Credit theory of the numeraire, in other words – the Kiyotaki and Moore (2001, 2002a, 2002b) model imports the Metallist concept of the numeraire as merely the (intrinsically countable) unit of a commodity.

On close examination, in other words, the Kiyotaki and Moore (2001, 2002a, 2002b) model in fact represents not a Credit theoretic model of the emergence of money *tout court*; but of the monetisation of private credit within an economy where (a) all claims are to some homogeneous store of value, and (b) the numeraire is simply the unit of this store of value. But this amounts to nothing other than an economy which is already monetised, in the sense of having an official, national money, with its own denomination. The Kiyotaki and Moore (2001, 2002a, 2002b) model is therefore better thought of as analysing the conditions under which, given the prior existence of such an official, national currency and its associated unit of account, private credit may come to circulate as money, than as analysing the ultimate determinants of the emergence and circulation of monies in general.

## 1.4 Towards a Credit Theory of Money Substitutes

Having surveyed the classical expositions of the Metallist and Credit theories of money in Section 1.2, and reviewed a selection of recent models of the micro-foundations of money and barter which develop those theories' analyses in Section 1.3, we are now in a position to tackle the main question of this chapter: whether and how the Credit theory of money can provide a useful framework for the analysis of money substitutes in developing and transition countries. Rather than proceed directly to an evaluation of the Credit theory, we begin this final section with a brief assessment of the prospects of the Metallist theory in Section 1.4.1. Finding that although the Metallist theory, as elaborated by the models reviewed in Section 1.3, does yield some useful insights into the emergence of money substitutes in developing and transition countries, our conclusion is that in general it lacks empirical relevance, we then proceed to our assessment of the Credit theory in Section 1.4.2 .

### 1.4.1 The Metallist Theory of Money and the Analysis of Money Substitutes

A first important point that we note concerning the Metallist theory of money is that it does not distinguish between official, national currencies and any other class of money: for both, the theory postulates the same concept of money and of the nature of monetary exchange, and the same criteria driving the choice between monies. In the Metallist hypothetical history of the emergence of money set out in Claims 3 and 4 in Section 1.2.1, the emergence of an official, national currency is to be explained as the result of a process of minimising transaction costs determined by the intrinsic properties of available commodities and the frictions inherent in barter exchange which results in the selection of a money-commodity to serve as a medium of exchange<sup>55</sup>. This account of the existence of official, national currencies is weak both in conceptual and explanatory terms. First, the Metallist concept of money, whilst easily interpretable in the case of coinage, is clearly only indirectly applicable to the modern, fiat version of official, national currencies. Second, as discussed in Section 1.2 above, the Metallist account of the

---

<sup>55</sup>Note that the Metallist theory can accommodate a difference in the agent of this cost-minimisation between official, national money and other classes of money: as noted above, although in the great majority of Metallist models, it is the market which solves the cost-minimisation problem (the 'Mengerian' version), it is equally possible to have the sovereign make this choice (as, for example, in 'legal restrictions' versions).

emergence of the official, national currency as a result of a cost-minimisation exercise is historically not compelling. Furthermore, various subsidiary hypotheses implied by the Metallist theory are also easily rejected: one example, documented by Goodhart (1998), is the failure of the theory of Optimal Currency Areas – a spatial-geographical derivative of the Metallist theory – to explain the observed distribution of official, national currencies<sup>56</sup>.

The Metallist theory offers more insights when it comes to the analysis of dollarisation – the co-circulation of national and foreign, official monies. The Metallist concept of money as a commodity is of course no more persuasive for foreign than for domestic, official money; but one basic prediction of the Metallist theory – that overproduction of the domestic, official money and a subsequent decline in its value relative to (other) commodities will lead agents to prefer to hold foreign, official money if its supply, and hence its relative value, remains stable – appears to conform closely with the dollarisation experience of many high inflation economies. This is obviously a useful and compelling prediction. Savastano (1996), however, sounds a note of caution, drawing on Latin American data: whilst it has historically often been true that high and hyper-inflation has prompted currency substitution, it is also the case that other factors – such as legal restrictions and portfolio considerations – have been equally or even more important. In particular, financial intermediation in foreign currency has in numerous cases persisted in an environment of low inflation – a circumstance which is better explained by portfolio balance models of dollarisation, the microfoundations of which have little in common with the Metallist theory. Chapter 3 of this thesis analyses these issues in detail, so we omit further discussion at this point.

When the same basic prediction is confronted with cases of the emergence of the other important class of money substitutes – privately-issued monies – it is less fruitful. First, the Metallist concept of money as a commodity is not sustainable at all in this case. An illustration of this is provided by the widely studied case of non-cash settlements (NCS) in Russia in the 1990s<sup>57</sup>. Although in early analyses of the Russian case it was widely assumed that much of

---

<sup>56</sup>The Euro provides a good example of Goodhart's (1998) main argument. If the Eurozone is an optimal currency area, then OCA theory is refuted because it did not generate a single currency before January 1, 2002. If, on the other hand, it is not an optimal currency area, the theory is refuted because it does have a single currency now (as of the time of writing, at least).

<sup>57</sup>See Seabright ed. (2000) for a useful guide to the empirical literature on NCS in Russia. The introduction to Chapter 2 of this thesis contains further references.

the observed NCS were being settled by commodity-currencies or by barter, it soon transpired that this was a misconstrual; and that instead “most of what Russian firms refer to as barter is not in fact what economists would term barter – i.e. the exchange of goods for goods; it is rather payment for goods using non-monetary methods and instruments – i.e. debt.”<sup>58</sup> Barter *stricto sensu*, it transpired, accounted for a very minor part of the means whereby transactions were settled without using Rouble money<sup>59</sup>: the vast majority were conducted either by the mutual offsetting of inter-enterprise payment arrears, either bilaterally, or as a part of a multilateral circle of offsets, and by the widespread issuance and circulation of bills of exchange and promissory notes (*veksels*) by government and firms<sup>60</sup>. Likewise, in the case of the emergence of money substitutes in Argentina, the privately issued monies that blossomed in the wake of the 2001 financial crisis were fiat currencies by definition not backed by reserve money, let alone by specie<sup>61</sup>.

Perhaps unsurprisingly, given the poor fit of the Metallist concept with the empirical facts of privately-issued monies, the Metallist theory of the choice of monies does not fare well in this case either. The basic prediction that the substitution of alternative monies will follow from a depreciation in the value of the domestic, official money is not borne out by actual episodes of the emergence of NCS. Maurel and Brana (1999) illustrate this fact in the case of Russia, where NCS expanded rapidly precisely when stabilisation was producing rapidly falling inflation. Chapter 2 of this thesis is devoted to a case study of Croatia, where money substitutes emerged and persisted on a large scale in an environment of sustained low inflation.

These empirical failings of the Metallist theory may be added – or indeed perhaps ascribed – to the theoretical limitations discussed in Section 1.3 above. There, we found that on a theoretical level alone neither the full nor the imperfect information classes of Metallist models of money were entirely satisfactory. On the one hand, basing the Metallist theory on asymmetric

---

<sup>58</sup>Carlin, Fries, Schaffer, and Seabright (2000), p.238.

<sup>59</sup>Aukutsionek (1998) reports, for example, that for firms heavily involved in NCS (>60% of sales settled by NCS), 19% of in-kind payments were re-bartered with other firms. Tchaidze (1999) p.7 gives a number of examples of typical situations in which genuine barter arose (regular transactions that had previously been included in inter-Republic counter-trade arrangements; so-called ‘take-and-give’ schemes in crude oil processing; etc.).

<sup>60</sup>Tchaidze (1999) provides an invaluable overview of the nature and economics of the most important classes of these.

<sup>61</sup>See De La Torre, Levy Yeyati, and Schmuckler (2003), pp. 75-77.

information does not yield satisfactory results for two reasons: first, because it requires us to accept the idea that the degree to which knowledge about the quality of a commodity is public is an objective, and therefore exogenously specified, property of that commodity, rather than a combination of intrinsic properties and the social environment in which knowledge is generated; and second, and perhaps more seriously, because the empirical basis for this tactic is essentially very weak – there is little historical or contemporary empirical evidence to suggest that imperfect information has had an important role in determining which monies circulate. On the other hand, we found that models which ground the Metallist theory in the empirically more plausible concepts of specialisation and geographically decentralised markets are required to assume, rather than explain, extrinsic beliefs on the part of economic agents that the fiat money-commodity will be accepted as a medium of exchange. In both classes of Metallist model, therefore, the existence of equilibria in which fiat money circulates relies upon the exogenous specification of appropriate knowledge (in full information models) or beliefs (in asymmetric information models) on the part of agents, rather than being properly endogenised<sup>62</sup>.

In sum, therefore, the ability of the Metallist theory to explain the co-circulation of official, national currencies and money substitutes in developing and transition countries is equivocal at best due to shortcomings on both the theoretical and the empirical levels. Can the Credit theory fare any better?

#### **1.4.2 The Credit Theory of Money and the Analysis of Money Substitutes**

We can assess the Credit theory in respect of both the validity of its concept of money and the relevance of its theory of the choice between monies against empirical cases of money substitutes in developing and transition countries. Assessing the validity of the Credit theory's concept of money is relatively simple. By contrast to the Metallist concept of money as a commodity, the Credit theory's concept of money as a transferable financial contract is an apt characterisation of all of official, national currencies, official, foreign currencies, and privately-issued monies; since each of these classes of money is at the same time an asset on the balance sheet of its holder and

---

<sup>62</sup>In the case of foreign currency cash considered as a money substitute, other techniques have been developed with which to model the development of such beliefs. Dowd and Greenaway's (1993) model of network externalities is one of these; and Uribe (1997) and Oomes (2003) provide further examples, and even empirical applications, in this vein, which are discussed in Chapter 3 below. This is as yet a relatively unexplored branch of research, however.

a liability on the balance sheet of the domestic central bank or a dependent commercial bank; a foreign central bank or dependent commercial bank; or another private agent, respectively.

Assessing the relevance of the Credit theory's theory of the choice between monies requires a more extended analysis. We concluded in Section 1.3 above that despite making considerable advances to our understanding of the micro-foundations of the Credit theory, the Kiyotaki and Moore (2001, 2002a, 2002b) model leaves two fundamental issues in the Credit theory's account of the choice between monies unresolved. The first of these is the question of what determines the capacity of a money's issuer to make binding unilateral commitments so that its credit can be created in the primary and transferred in the secondary markets; the second, the question of how the numeraire with which that money is accounted is established and chosen. These two issues require further theoretical elucidation before the Credit theory can generate empirically testable hypotheses - and before its usefulness as a theoretical framework for the analysis of money substitutes can thereby be assessed. In the remainder of this chapter we therefore turn back from the recent models of money and barter reviewed in Section 1.3 to the classical Credit theory literature in order to review what approaches are taken in that less formal literature to these two issues. With respect to the determinants of money-issuers' capacities to make binding unilateral commitments, the classical Credit theory literature includes three well-defined traditions. We therefore review each of these - the Simmelian, the Market, and the Cartalist traditions - in turn below; evaluate what insights they provide into empirical examples of money substitutes in developing and transition countries; and use this as a basis on which to assess where future theoretical research might profitably be concentrated. With regard to the issue of what governs the choice of a money's numeraire, we identify much less guidance in the classical Credit theory literature. At the end of this section we briefly review a theory put forward by Grierson (1977); but argue that it is unfortunately not particularly useful as a theory of the numeraire that might be applied to episodes of money substitutes in modern economies.

## The Theory of Money-Issuers' Capacities to Make Binding Unilateral Commitments

**The Simmelian Tradition** Georg Simmel's monumental *Philosophy of Money* is recognised amongst sociologists as one of the cornerstones of the sociological investigation of money; but, as Laidler and Rowe (1980) pointed out in their 1980 review article for economists, "this particular work, and indeed Simmel himself, seems to be almost unknown to economists nowadays"<sup>63</sup> – and the situation has not changed significantly in the intervening twenty-five years. Simmel's philosophy of money is Credit theoretic. He explicitly rejects the Metallist concept of money as a commodity, the value of which is determined by its intrinsic properties: whereas commodities "have a specific content, from which they derive their value, . . . money derives its content from its value"<sup>64</sup>. What determines money's value – and hence its circulation – instead of its intrinsic properties is the shared faith of the economic community in the guarantee of its continued acceptability at a stable rate of exchange. Laidler and Rowe (1980) describe Simmel's characterisation of the origins of this shared faith as follows:

"Belief in this guarantee contains 'an element of socio-psychological, quasi-religious faith' which must be based upon 'confidence in the socio-political organisation and order', and is referred to by Simmel (at least in this translation) as 'trust'."<sup>65</sup>

As to the origins of this "confidence in the socio-political organisation and order", the Simmelian tradition holds that it is the product of an evolutionary process:

"In Simmel's view the monetary system is not the conscious creation of any political entity but is the unintended product of social evolution. In this, the development of money as a social institution resembles the growth of a moral code or a legal system."<sup>66</sup>

Two distinguishing features of the Simmelian tradition are therefore (a) that the principal determinant of the capacity of a money's issuer to make binding contractual commitments

---

<sup>63</sup>Laidler and Rowe (1980), p.97.

<sup>64</sup>Simmel (1978), p.121.

<sup>65</sup>Laidler and Rowe (1980), p.99.

<sup>66</sup>Laidler and Rowe (1980), p.98.

is a type of trust that is intrinsically social in nature, so that attempting to identify and measure determinants at the level of the individual issuer are futile; and (b) that the process that generates this trust is likewise social and evolutionary. These features are not unique to Simmel's own presentation, or peculiar to the German intellectual tradition. Walter Bagehot, the high priest of nineteenth century English financial economics, held a closely similar view: "Credit in business is like loyalty in Government", he wrote; "a power that may grow but cannot be constructed"<sup>67</sup>.

What guidance does this Simmelian tradition offer the Credit theoretic analysis of the competition between national and foreign official currencies, and privately-issued money substitutes? Certainly, the Simmelian tradition provides an answer of a sort to the question of why the official, national currency generally dominates foreign and privately-issued competitors. The money-issuing state is a component of the political order which is of a piece with the monetary order itself. Trust in the state's money is therefore just equivalent to trust in the state's political authority on the Simmelian view; whilst the monies of foreign or private issuers, who have by definition only subsidiary political status, enjoy a level of trust of a different (lower) order of magnitude. This is indeed a subtle and potentially persuasive answer. As a guide to further economic analysis, however, it is not particularly fruitful. It directs us to the sociological and socio-psychological roots of trust in the state and political life in general, rather than to particular institutions on the one hand, or the operation of markets on the other; and hence represents the end point of one (sociological) line of analysis, rather than a potential starting point for another (economic).

**The Market Tradition** A second tradition in the Credit theoretic literature shares with the Simmelian view the idea that the choice of money is the outcome of a decentralised, market-driven process rather than of the diktat of the state; but characterises the arguments of this choice in terms of economic factors rather than the Simmelian concept of social trust. One example of such economic factors is that provided by the information-theoretic determinants of the constraints on borrowing and transfer in the primary and secondary markets suggested by Kiyotaki and Moore (2001, 2002a, 2002b). Monies issued by debtors on whom creditors have

---

<sup>67</sup>Bagehot (1873). Lombard Street, p.68, quoted in Frankel (1977), pp.30-1.

only imperfect information compete on the basis of transferability, and the liabilities of those issuers that suffer least from problems of asymmetric information with potential counterparties enjoy the widest circulation. Another factor identified by theorists in this Market tradition (e.g. Mehrling (2000)) is the relative size of the market for the goods or services produced by competing issuers of monetary credit; since the greater the size of the issuer's market, the greater the probability of a double coincidence of wants in the event that the issuer's liability must be held to maturity by its creditor. Finally, there are structural factors such as the institutional infrastructure of the payments system associated with the liabilities of a particular money-issuer which literature in the Market tradition (e.g. Freeman (1996b)) postulates as determinants of the issuer's capacity to make a binding unilateral commitment.

Because it identifies specific, economic factors of the market's choice of its money, the potential of the Market tradition is easier to assess against empirical evidence than that of the Simmelian tradition. Certainly, a case can be made on the basis of all three of the determinants noted above for the normal dominance of the official, national money over privately-issued alternatives; as well as for the erstwhile emergence of money substitutes. With respect to information asymmetries, for example, it might be argued that it is much easier for households and firms to access information regarding the creditworthiness of their domestic government than of a foreign government or a private issuer – not least because they have first hand experience of it. Moreover, there is evidence that the currency of privately-issued monies depends heavily on the extent to which potential new counterparties to an issuer's liabilities have information regarding their creditworthiness. When, for example, during the six and a half month closure of the Associated Banks in Ireland in 1970, personal cheques irredeemable at banks circulated as money, the influence of informational factors was clear:

“Cheques drawn on the government and on well established institutions were readily accepted by the alternative banking system. . . . The negotiability of personal cheques depended on the degree of information and personal contact that the acceptor had about the issuer of the cheque.”<sup>68</sup>

As a result, it was public houses that tended to serve as the financial intermediaries of this

---

<sup>68</sup>Murphy (1978), p.45. I am indebted to the late John Flemming for introducing me to the Irish case.

“alternative banking system”: as Murphy (1978) puts it with admirable circumspection, “one does not after all serve drink to someone for years without discovering something of his liquid resources.”<sup>69</sup>

Likewise, with respect to the size of the issuer’s market, the state enjoys a unique advantage because as one adherent of the Market tradition holds, “it is the one entity with which every one of us does ongoing business. We all buy from it a variety of services, and the price we pay for those services is our taxes.”<sup>70</sup> And again, recent episodes of the emergence of money substitutes bear out the importance of this factor. Electricity and gas utilities were identified as major hubs of NCS in Russia and Ukraine, in part by virtue of their exceptionally broad bases of customers and suppliers, and the fact that many enterprises that fell into the latter category also fell into the former one: facts which allowed them to operate both as financiers and as clearing houses for inter-enterprise trade credit<sup>71</sup>. In Argentina, shopping malls were amongst the leading issuers of private currencies, again supporting the hypothesis that a large market makes private liabilities transferable since potential assignees are not overly averse to the risk of having to redeem them with the original issuer<sup>72</sup>.

Finally, with respect to the structural factor of the payments system associated with a particular money, official, national money enjoys the services of the national banking system and its associated payments clearing institutions, putting it at a structural advantage over other issuers’ monies in terms of the costs associated with the off-setting of liabilities. The services of the official payments system may or may not be extended to foreign currency money, and recent research has suggested that the exclusion of foreign currency settlements does indeed have a negative effect on the circulation of foreign currency<sup>73</sup>. In general, there is no such organised payments system for private liabilities, which conforms to the empirical observation that private liabilities are generally highly illiquid. Even more persuasively, the fact that when organised systems for the clearing of private liabilities have existed – as for example in the Irish case referred to above, private liabilities have acquired currency relatively easily. A particularly

---

<sup>69</sup>Murphy (1978). pp.44-5.

<sup>70</sup>Mehrling (2000). p.402.

<sup>71</sup>See Pinto, Drebensov, and Morozov (2000) on Russia, and Marin, Kaufmann, and Gorochowskij (2000) on Ukraine. Evidence of the same phenomenon in the Croatian case is discussed in Chapter 2.

<sup>72</sup>See Catán (2002).

<sup>73</sup>See Fritz-Krockow et al. (2001). For further discussion see Chapter 3.

relevant example of this is the case of the Soviet Union, which had special institutions for the multilateral clearing of non-monetary, inter-enterprise trade credit during several periods of its history: as a result, the liabilities of enterprises circulated as an officially sanctioned dual currency alongside Rouble money<sup>74</sup>.

**The Cartalist Tradition** There is, finally, a third tradition – the Cartalist tradition – which argues that in the case of official, national currency, the capacity of the issuer to make a binding, unilateral commitment, and hence the value of money, is determined not by economic factors as in the Market tradition, nor by the quality of social trust as in the Simmelian view, but by the coercive power of the state<sup>75</sup>. The Cartalist theory states that whereas in the case of money issued by private agents of one sort or another, we might allow that this capacity will be determined by the various economic factors described by the Market view, official, national currency – the money of the state – is fundamentally different. The state’s liabilities, issued in order to extinguish outstanding credits of subjects who have supplied the sovereign with commodities, are valued by those same subjects because they may be used to extinguish their liabilities to the state – the great majority of which are associated not with the market exchange of commodities, but with the non-market institution of taxation. The state is creditworthy, and its liabilities unimpeachably transferable, in other words, because they can be used to extinguish subjects’ tax liabilities. Even Smith, whose general theory of money was certainly Metallist, conceded this point:

“A prince, who should enact that a certain proportion of his taxes should be paid in a paper currency of a certain kind, might thereby give a certain value to this paper money; even though the term of its final discharge and redemption should depend altogether upon the will of the prince.”<sup>76</sup>

---

<sup>74</sup>The institutions in question were the Office for Mutual Settlements (during the 1930s), and the *Dezachet* (post-1955). In 1959, 45% of all inter-enterprise trade in the Soviet Union was settled via NCS organised by the *Dezachet* (Grossman (1966), p.215, note 27). See Grossman (1966), Grossman (1968), Garvy (1977), and Woodruff (1999) for further details.

<sup>75</sup>This Cartalist view is advocated, for example, by Mitchell Innes (1914), by Wray (2004), and by Ingham (2004).

<sup>76</sup>Smith (1981) p.328.

Numerous modern economists, whose macroeconomic monetary models are equally Metallist in inspiration, likewise admit this Cartalist argument. Buiter (2005), for example, argues in a recent meditation on the state of monetary economics that:

“Ultimately, that superior liquidity (*sc.* of official, national money) derives from the unquestioned creditworthiness of the central bank, as agent of the state. That security and creditworthiness derive partly from the legal tender nature of the central bank’s monetary liabilities. More fundamentally, it derives from the fact that the central bank is an agent of the state, the sovereign, and that behind the central bank stands the Treasury with its power to tax and other government agencies with the power to regulate, that is, to prescribe and proscribe behaviour. The monopoly of the legitimate use of force (or coercion) is what makes the state unique. The central bank trades on that.”<sup>77</sup>

The state is able to create an ‘artificial’ demand for its own liabilities - ‘artificial’ in the sense that it is not an equilibrium outcome of exchange in the market - on the basis not of voluntary exchange, but of the power of the sovereign to require that its subjects pay their taxes in whatever money it sees fit<sup>78</sup>. The positive value of official, national money is therefore determined, at least in part, by the extent of this power; not, as is the case for non-sovereign money issuers, by economic factors in the market<sup>79</sup>. In addition, the institutional infrastructure for monetary and financial policy-making, available to the sovereign for the operation of the national monetary system, but not to private issuers for the management of private monies, represent an additional source of ‘charter value’<sup>80</sup>.

Two further points are worth making regarding the Cartalist tradition. First: as demonstrated by Goodhart (1998), the Cartalist theory certainly conforms well with the historical evidence on the origins of official, national currencies: the vast majority of available evidence

---

<sup>77</sup>Buiter (2005), C23-4.

<sup>78</sup>Hence amongst a Credit theorist like Hoover’s (1996) “Suggestions for Complicating the Theory of Money” we find the acceptance of the fact that national currency is not held voluntarily. See Hoover (1996) pp.213-4.

<sup>79</sup>The State theory of money advocated by Knapp (1924) is another version of this Cartalist view, which satisfies itself with locating the power of the state in legislation *per se*, rather than identifying it with the specific institution of taxation.

<sup>80</sup>Indeed, these sources of charter value - and in particular the forms taken by the institutions of monetary policy - have recently been identified by some leading monetary economists as a focus for future research. See, for example. King (2004).

tends to suggest that the emergence of national currencies and banking systems is historically the result of deliberate fiscal policies of sovereigns<sup>81</sup>. Second: the central idea of the Cartalist theory – that power, in the sense of being able to impose the terms of exchange on trading partners, is an additional determinant of the capacity of a money-issuer to make a binding, unilateral commitment – can be applied to non-sovereign issuers as well as to states, to the extent that such issuers enjoy some degree of coercive power over economic agents. To the extent that economic factors cannot explain the dominance of official, national money as a source of liquidity, and conversely the range and dynamics of money substitutes in circulation, the Cartalist theory postulates an additional factor – the power of the issuer.

### **The Theory of the Numeraire**

In Section 1.2.2 above, we found that the Credit theory of money absolutely requires a theory of the numeraire, because it takes credit to be a primitive concept and rejects the Metallist concept of a commodity medium of exchange, the intrinsically countable unit of which serves as the numeraire by default. Remarkably, however, the classical Credit theory literature, despite its sophisticated exposition of the core of the Credit theory of money, includes no such explicit discussion of how the unit of account in which monies are denominated originates or is chosen. As Mitchell Innes (1914) admits:

“But even when we have grasped this truth (*sc.* the Credit theory of money) there remain obscurities which in the present state of our knowledge cannot be entirely eliminated. What is a monetary unit? What is a dollar? We do not know. All we do know for certain . . . is that the dollar is a measure of the value of all commodities, but is not itself a commodity, nor can it be embodied in any commodity. It is intangible, immaterial, abstract. It is a measure in terms of credit and debt.”<sup>82</sup>

---

<sup>81</sup>As well as Goodhart (1998) and the references therein, see Kim ed. (2001), Cipolla (1956), and Ferguson (2000) on the development of official, national currencies in the Ancient Greek, mediaeval, and early modern worlds respectively.

<sup>82</sup>Mitchell Innes (1914), p.159.

Historians of money have, however, been more exercised by the question of the origin and determinants of monetary units of account. One intriguing and ingenious theory has been advanced by Grierson (1977). Grierson (1977) starts from the observation that universal units of value even in the measurement of area, volume, and weight are relatively modern developments:

“A study of other standards of measurement shows their generalized character to have developed comparatively late in their history. They were devised – or at least there is good reason to suppose they were devised – for comparing different quantities of the same object, or different quantities of closely related objects.”<sup>83</sup>

Even today, units of measurement that apply only to particular commodities or activities – such as the fathom as a measure of depth, or the carat as a measure of fineness – survive in modern usage. This suggests to Grierson (1977) that the universal units of economic value represented by the numeraire of official, national monies may also be a relatively recent innovation; and indeed historical and contemporary anthropological evidence corroborates this hypothesis. Historians and economic anthropologists note that so-called ‘limited-purpose monies’ – the numeraire of which were confined to the particular sub-class of transactions for which they were used – were common features of historical and contemporary ‘primitive’ economies<sup>84</sup>.

It is clear from the historical and anthropological evidence that the use of limited-purpose monies was not necessarily connected with market exchange, but instead with various social, political, or religious practices; and Grierson (1977) argues that this makes unlikely the hypothesis that the establishment of the numeraire attaching to such monies can be explained as a market process. Rather, the determinants of the unit of value are to be sought in the particular practice for which the limited-purpose money was deployed. For example, the most important such practice for the future development of official, national money and its numeraire in Europe, according to Grierson’s (1977) argument, was the archaic institution of *wergild*, whereby judicial penalties for personal injury were established. The schedule of fines, in terms of commodities, for injuries to members of the tribe was established via the decision of the sovereign tribal assembly – an explicitly political, rather than an economic, process – and the value of an

---

<sup>83</sup>Grierson (1977). p.18.

<sup>84</sup>See Quiggin (1954), Einzig (1960), Firth (1967), and Grierson (1977) *passim*.

individual's life thereby established constituted the unit of account. The thing being accounted for was not a commodity but the individual member of the tribe – so that the problem of incommensurability was absent, insofar as all individuals were considered equal before the law. The association of this, originally limited-purpose, numeraire with the judicial process, and hence the sovereign, on the one hand, and its intrinsic universality as a measure of the value of personal injury on the other, led to its eventual adoption as the universal unit of account of official, national money. Grierson (1977) summarises the argument as follows:

“The conditions under which these laws were put together would appear to satisfy, much better than any market mechanism, the prerequisites for the establishment of a monetary system. The tariffs for damages were established in public assemblies, and based on objects of some value which a householder might be expected to possess or which he could obtain from his kinsfolk. Since what is laid down consists of evaluation of injuries, not evaluation of commodities, the conceptual difficulty of devising a common measure for appraising unrelated objects is avoided.”<sup>85</sup>

Ingenious though Grierson's (1977) theory of the numeraire is, however, it is difficult to use it as a basis on which to explain the units of account in which money substitutes in developing and transition countries are denominated. Certainly, it is an attractive aspect of the theory that it accommodates the phenomenon of multiple monies in circulation, each with its own numeraire – since this is precisely what is observed in practice in the case of, for example, the provincial currencies issued in Argentina in 2001-2. But the theory is too general to explain why in many other cases of money substitutes – for example, the money substitutes used to settle NCS in transition economies – the unit of account of the official, national money is chosen. Moreover, whilst the argument that new units of account are established via political rather than economic means is convincing in general terms, it is unclear how the theory might be applied in particular, modern cases.

In summary, therefore, there is little guidance in the Credit theoretic literature, modern or classical, on the critical questions of what determines which numeraire is associated with a

---

<sup>85</sup>Grierson (1977), p.21.

given money; and of how these numeraires are established in the first place. Nevertheless, a theory of the numeraire is essential to a comprehensive Credit theory of money substitutes. The field in this area is therefore open; and is an important avenue for future theoretical research.

## 1.5 Conclusion

Explaining the emergence on a large scale of money substitutes – be they in the form of foreign currency or privately-issued monies – presupposes a theory of money. Based on a review of the Metallist and Credit theories of money in Section 1.2, and of recent theoretical contributions to the literature on the micro-foundations of money and barter in Section 1.3, we concluded that whilst both theories provide insights into the nature of and choice between monies, it is the Credit theory which holds the greater promise as a framework for the analysis of money substitutes in developing and transition countries. Nevertheless, we also identified two critical questions which modern theoretical research in the Credit theory tradition has yet to confront; but which it must answer if there is to be a coherent Credit theory of money substitutes. The Credit theory identifies the central determinant of the choice between alternative monies as the relative capacity of their issuers to make a binding unilateral commitment: the first of these questions is therefore which factors, institutional or individual, and on the sides of both issuer and bearer, determine this capacity. Having jettisoned the concept of a money-commodity whose naturally countable units serve as the numeraire by default, the abstract monetary unit of account in which an issuer's liabilities are denominated is in need of explanation: the second major question that remains to be answered is where such units come from, and what determines the fact that some monies have their own numeraire whilst others are vicariously denominated. In Section 1.4, we surveyed three approaches taken in the Credit theory literature which provide potential guidance to modern theoretical research into the first of these questions; but found that guidance on the second, and on possible links between the two questions, was difficult to locate. We conclude that future theoretical research can concentrate profitably on attempting to model the general theoretical structure provided by the Market and Cartalist traditions; but must look elsewhere – to historical, and perhaps to experimental and psychological research – for guidance in modelling the choice of the numeraire.

In addition to these theoretical implications of the current state of the Credit theory of money for the study of money substitutes in developing and transition countries we may also ask what its implications are for empirical research. Evidently, the models reviewed in Section 1.3 above are not empirically tractable in the sense of generating solutions the parameters of which may be identified with those of empirical equations; but our conclusions concerning the Credit theory can nevertheless provide a useful guide to empirical modelling in several respects. Most fundamentally, empirical modelling based on the Credit theory should start from the concept of money and money substitutes as financial contracts – and therefore from a consideration of both borrowers and lenders – rather than as commodities. If the subject of empirical study is the choice between monies, the structural conditions underlying the use of which are similar, modelling the choice of monies as a financial equilibrium between borrowers and lenders using the theory of portfolio balance is therefore a natural framework. The choice between domestic and foreign, official currencies – for which structural factors of demand and supply such as the payments system infrastructure and norms of credit creation are often identical – is such an example; and in Chapter 4 of this thesis, we deploy just such a framework. If, on the other hand, the subject of empirical study is the choice between monies, the structural conditions underlying the use of which differ substantially – the choice, for example, between official, national currency and between privately-issued monies – then a portfolio balance approach is likely to be too demanding in terms of data to be useful. In these cases, empirical studies can nevertheless profit from the fruitful analytical distinction made by the recent Credit theoretic model of Kiyotaki and Moore (2001, 2002a, 2002b) between the primary and secondary markets for the two types of credit. In Chapter 2, for example, we undertake an empirical analysis of NCS in Croatia that makes use of this distinction.



## **Chapter 2**

# **Non-Cash Settlements in Transition: the Case of Croatia**



## 2.1 Introduction

The literature on the inter-enterprise non-cash settlements (NCS) that emerged on a large scale in many transition economies during the 1990s has been dominated by the case of Russia. As the macroeconomic data in Table 2.7 indicate, NCS in Russia emerged in the context of an exceptionally deep and prolonged recession: between 1992 and 1998 – the period during which NCS grew to their peak of over 50% of all inter-enterprise transactions – the Russian economy contracted at an average rate of 7% per year. At the same time, the Russian government was pursuing a policy of inflation stabilisation based on severe monetary restriction: a policy which achieved considerable, if gradual, success – reducing inflation from over 1,500% in 1992 to around 15% in 1997 – though at the cost of foregoing any possibility of monetary alleviation of the demand shock to the enterprise sector resulting from the collapse in aggregate demand<sup>1</sup>. The result of this concatenation of extremely negative macroeconomic conditions and restrictive monetary policy was mass insolvency in the enterprise sector, on a scale which Russia could not (or would not) in the end tolerate. The Russian government was faced with the dilemma of either giving up its stabilisation programme or imposing hard budget constraints on the enterprise sector and implementing the massive restructuring programme that this would entail. The solution, as argued in well known accounts such as those of Pinto, Drebenstov, and Morozov (2000) and Commander and Mummsen (2000), was found in NCS – which were understood as a mechanism whereby the Russian government was able to deliver massive subsidies to a loss-making enterprise sector without creating bank money<sup>2</sup>. The mechanism operated both by allowing the accumulation of a stock of net payment arrears from the enterprise to the government sector, which might then be written off, and by the offsetting of flows of payables and receivables between the two sectors at rates which transferred an implicit subsidy to the enterprise sector. In Russia, therefore, NCS and the credit instruments used to conduct them

---

<sup>1</sup>The data on inter-enterprise NCS in Table 2.7 are taken from Aukutsionek (1998) for 1992 to 1997; Commander and Mummsen (2000) for 1998; and from the World Bank-EBRD Business Environment and Enterprise Performance Surveys 1999 and 2001 for those years.

<sup>2</sup>Informal versions of this theory – such as those in Commander and Mummsen (2000), Hildebrandt (2002), or Commander, Dolinskaya, and Mumssen (2002) – found the responsibility for NCS ultimately to lie with the government's lack of commitment to its microeconomic reform programme. More formal expositions, such as that of Perotti (1998), demonstrated the possibility that the enterprises might be the motive force, using a model of the relationship between enterprises and the government as a strategic game.

were in essence deemed to be a means of bailing out an unstructured and insolvent enterprise sector, and as Commander and Mummsen (2000) observe, “[t]he Russian barter story hinges crucially on the observation that there has been a continuous infusion of net credit into the enterprise system”<sup>3</sup>.

In this paper, we analyse the case of Croatia, in which large scale NCS emerged in macroeconomic conditions very different to those that obtained in Russia. As the data in Table 2.8 illustrate, Croatia underwent an exceptionally rapid and successful stabilisation in mid-1993, since when macroeconomic conditions have been highly favourable. Inflation has been low and stable, averaging under 4% *p.a.* over the decade 1994-2003; GDP growth has been consistently high and positive, with the single exception of a slight recession in 1999, averaging 4.3% *p.a.* over the same period; and the exchange rate of the Croatian Kuna (HRK) against the Deutschmark and later the Euro has remained extremely stable, with a standard deviation of less than 5%. Yet despite these positive macroeconomic conditions, inter-enterprise NCS grew from 15% of total inter-enterprise transactions in 1994 to over 20% in 1996, and, whilst exhibiting considerable variation, continued to average over 15% in every year until 2001 – the last year for which data are available. The favourable macroeconomic conditions (as well as a superior pre-transition starting point in terms of the number of SMEs and the greater proportion of enterprises in private ownership) were reflected in a much superior profitability and performance of the Croatian enterprise sector. Although a handful of significant loss-making sectors did receive direct budgetary subsidies or tax write-offs from the Croatian government, the enterprise sector in aggregate enjoyed healthy demand growth in line with increasing GDP. As a result, far from accumulating net payment arrears to the government, as in the Russian case, the enterprise sector in Croatia was, in aggregate, subsidising the Croatian government in the late 1990s<sup>4</sup>. Furthermore, the Croatian government consistently refused to offset the

---

<sup>3</sup>Commander and Mummsen (2000), p.126. This was the dominant and best corroborated theory; but by no means the only one. Other (not mutually exclusive) theories of NCS in Russia posited that they represented means for enterprise managers illegally to extract value from their firms or to evade taxes (Poser (1998)), for the Russian government to operate a ‘virtual economy’ (Gaddy and Ickes (1998)), or – a version of the thesis that we will develop below – for liquidity-constrained firms to generate substitutes for inaccessible bank credit (Maurel and Brana (1999)).

<sup>4</sup>Survey data presented in Šonje, Faulend, and Šošić (2001) implies that the ratio of accounts receivables from, to accounts payables to, the Croatian government of the enterprise sector was approximately 1.4 in 1999 and 1.3 in 2000. Unsurprisingly, in the same surveys, a much larger proportion (55%) of enterprise managers themselves therefore reported that the problem of illiquidity could be resolved by greater fiscal discipline than that it was a

gross claims of enterprises in payment of tax or other government receivables – it did not, in other words, participate in NCS. The result of this combination of policies was that Croatia’s first non-nationalist leadership recognised net payment arrears of government to the enterprise sector of nearly 6.5% of GDP six months after its election in 2000<sup>5</sup>. Both in terms of the underlying macroeconomic conditions and in terms of government policy, the contexts in which NCS emerged in Croatia and Russia were therefore radically different; and as a result, the principal explanation of NCS in Russia – as a mechanism whereby transfers were delivered to an unstructured and insolvent enterprise sector without upsetting stabilisation – is *prima facie* implausible in the case of Croatia.

This paper therefore sets out an alternative interpretation of NCS in Croatia. We start from two basic facts about NCS in Croatia – that they consist, almost exclusively, of the off-setting of trade credit of one sort or another; and that they are solely an inter-enterprise, rather than an enterprise-government, phenomenon<sup>6</sup>. Given these two facts, a natural interpretation of the instruments used to settle NCS in Croatia is that they are a special case of trade credit: monetary trade credit. Rather than the trade credit with which NCS are conducted representing an infusion of net credit to the enterprise sector by government, and its off-setting via NCS representing a mechanism for the delivery of an implicit subsidy, NCS trade credit is to be explained on this interpretation in the same way as trade credit in developed market economies: either as a means for enterprises to secure contractual benefits in individual transactions – as ‘real’ theories of trade credit propose – or as a means of generating and allocating inter-temporal liquidity in order to facilitate value-adding production and trade – as ‘financial’ theories of trade credit postulate.

This alternative interpretation of NCS raises three important questions. The first is why

---

problem due to other firms (30%).

<sup>5</sup>See IMF (2001).

<sup>6</sup>On the former point, Dukić and Šošić (2001) report the results of a 2001 CNB survey of NCS in Croatia. Amongst the 80% of firms surveyed which reported participating in NCS, an average of 76% of NCS was reported to consist of the offsetting or endorsement of trade credit, or payment with private credit notes, and an average of 33% to be spot barter. However, as Commander and Mummsen (2000) point out, “[A] fundamental insight is that spot barter (the exchange of goods that takes place at one point in time) can indirectly replicate inter-firm trade credit. The maturity of this implicit trade credit is equal to the difference of the ‘subjective’ marketing time of the goods exchanged. When firms sell their output in a barter deal, they often receive in return goods or services that are not immediately useful to them. It takes time and effort to sell the goods or money surrogates received in barter deals for cash or use them for settling payables or purchasing inputs.”

such a large volume of trade credit emerged in the Croatian economy during the 1990s. If on the one hand NCS trade credit is essentially a mechanism whereby enterprises secure favourable transactions or reduce transactions costs – the ‘real’ theory – then the emergence of NCS indicates imperfections in Croatia’s goods markets and a need for enterprise sector restructuring. If, on the other hand, the key role of NCS trade credit is to supply liquidity to meet the needs of trade – the ‘financial’ theory – then their emergence implies imperfections in Croatia’s financial markets. A second question, if NCS trade credit is found to be a ‘financial’ rather than a ‘real’ phenomenon, is whether the negotiability of the NCS trade credit instruments is such that they are not only substitutes for public credit, but substitutes for public money. If NCS trade credit instruments provide not only inter-temporal, but atemporal exchange liquidity, to such a degree that they compete with the national currency, then the emergence of NCS has consequences not only for financial, but for monetary control. If the answers to these first two questions indicate that the instruments used to conduct NCS are indeed private monetary credit, then a third important question is whether the financial market imperfections driving NCS derive only from deficiencies in public financial intermediation and a need for banking sector reform, or whether they are in addition underpinned by more germane aspects of Croatia’s financial structure - and in particular its very high level of financial dollarisation.

In this paper, we apply this alternative theory of NCS as private monetary credit to the case of Croatia, and seek to answer these three questions that it raises. The remainder of the paper is organised as follows. In Section 2.2, we discuss the theoretical literature on the primary and secondary markets for trade credit, and derive hypotheses that allow us (a) to discriminate between the ‘real’ and a ‘financial’ theory of trade credit as a valid explanation of NCS trade credit in Croatia; (b) to identify what types of credit market imperfections drive the rationing of public credit; and (c) to determine whether NCS trade credit instruments are substitutes for transactions money as well as bank credit. In Section 2.3 we then present an institutional analysis of conditions in the Croatian financial sector over the transition; and based on this, proceed in Section 2.4 to test our two hypotheses econometrically. Section 2.5 concludes.

## 2.2 Theory

De Blasio (2003) distinguishes the two principal theories of trade credit in developed, market economies. The ‘real’ theory argues that trade credit is to be explained as a contractual instrument arrived at as the solution to a bargaining problem between firms operating in imperfect goods markets. In the simplest models derived from this theory, trade credit is just a means of reducing transactions costs (Ferris (1981)). In more sophisticated versions, it is a means of verifying product quality (Smith (1987)), or securing a more elaborate contractual advantage of some sort (Rajan and Peterson (1997)). The ‘real’ theory is the intellectual inheritor of the extensive, but now dormant, literature on countertrade. Like that literature, it presents trade credit and associated payment in kind as a solution to goods market imperfections, without broaching the issue of imperfections in financial markets.

The ‘financial’ theory, on the other hand, argues that trade credit, like bonded debt or bank credit, is a means of generating and allocating inter-temporal liquidity in order to facilitate production and trade. Just as the mainstream literature explains the existence of bank credit alongside bonded debt as a consequence of problems of imperfect information (be they the adverse selection of borrowers as in Leland and Pyle (1977), the difficulty of monitoring as in Diamond (1984), or the difficulty of screening as in Broecker (1990)); the existence and dynamics of trade credit is likewise argued to be the result of imperfections in financial markets. The simplest version of the theory consists of extending the argument for the existence for bank credit alongside bonded debt: so Biais and Gollier (1997) and Frank and Maksimovic (1998) argue that the trade credit exists because trading partners have a better ability to screen or monitor borrowing firms than banks do. Burkart and Ellingsen (2004), however, argue that such models based on information asymmetries are vulnerable to the criticism that banks are meant precisely to specialise in screening and monitoring. They therefore offer an alternative model in which inputs are less divertible (in the sense of Hart (1995)) than money, and for this reason trade credit less rationed than bank credit<sup>7</sup>.

---

<sup>7</sup>Ellingsen (2000) is a *portmanteau* model that combines elements of both the ‘real’ and the ‘financial’ theories – postulating both that firms have superior information to banks on trading partners, but also that they seek to elicit as much payment in public money as possible, and on this basis their optimal strategy when trading partners’ public money liquidity is not known with certainty is to offer a menu payment options which includes trade credit and payment in kind.

These two theories generate different predictions of the level and dynamics of trade credit that should be observed in practice. The ‘real’ theory argues that the volume of trade credit is determined by the imperfections in respective goods markets. As such, it is usually taken to imply variation across sectors and industries, in which competition and other real conditions may vary considerably; but a relatively stable level of trade credit over time. Empirical studies such as that of Ng, Smith and Smith (1999) have found support for this hypothesis. The ‘financial’ theory, on the other hand, argues that the volume of trade credit will be determined by the extent of rationing in the market for bank credit. As such, there may be variation across sectors and industries; but there will also be variation across time, as bank credit conditions become more or less restrictive. An important special case of this prediction is the Meltzer (1960) hypothesis, which postulates that trade credit substitutes for bank credit during periods of financial restriction. The Meltzer (1960) hypothesis has also found support in a number of recent empirical studies, including Fisman and Love (2002) and Nielsen (2002).

Burkart and Ellingsen (2004) have recently clarified the relation of the Meltzer (1960) hypothesis to the ‘financial’ theory of trade credit more generally. They argue that the fact that trade credit is a form of finance does not on its own entail that it is a substitute for bank credit. This conclusion requires an additional assumption, namely that the factors that restrict the primary market for bank credit do not also restrict the primary market for trade credit. Only if firms are not also trade credit constrained will increased rationing in the primary market for bank credit result in an expansion of trade credit to meet the needs of trade. If, on the other hand, firms are both trade and bank credit constrained, and rationing in both markets depends on the same factors, then more restrictive bank credit conditions will also constitute more restrictive trade credit conditions, and trade and bank credit will be complements, rather than substitutes. Thus, the Meltzer (1960) hypothesis holds true only if (a) trade credit is a financial phenomenon, and (b) firms are not trade credit constrained.

The Meltzer (1960) hypothesis can therefore be decomposed into two hypotheses for testing against the Croatian case, each of which allows different inference on the nature and dynamics

of NCS trade credit. The first of these hypotheses is:

$H(1)$  : The volume of private credit instruments used to conduct NCS varies systematically with the rationing of bank credit.

Testing this hypothesis can be used to discriminate between the ‘real’ and ‘financial’ theories as candidates to explain NCS trade credit. If  $H(1)$  is not rejected, then the ‘financial’ theory is a satisfactory explanation of NCS trade credit; if on the other hand  $H(1)$  is rejected, then it is better explained by the ‘real’ theory<sup>8</sup>.

If  $H(1)$  is not rejected, we can proceed to test the second hypothesis:

$H(2)$  : The private credit instruments used to conduct NCS are substitutes for bank credit.

Testing  $H(2)$  allows us to infer whether or not Croatian firms are trade credit as well as bank credit constrained. If  $H(2)$  is not rejected, then according to Burkart and Ellingsen’s (2004) reasoning Croatian firms are not trade credit constrained. The Meltzer (1960) hypothesis is therefore confirmed: the volume of total (private plus public) credit is determined by the needs of trade, whilst the composition of that credit between bank and trade credit varies inversely with the degree of rationing in the market for bank credit. If, on the other hand,  $H(2)$  is rejected, then NCS trade credit is a complement to, rather than a substitute for, bank credit, implying that Croatian firms are rationed in the market for trade credit as well as bank credit, and that the factors of that rationing are identical in both cases.

The theories discussed so far relate only to the issuance of trade credit in the primary market – that is, to the explanation of trade credit as a means of generating and allocating inter-temporal liquidity on a bilateral basis. The fact that NCS trade credit is used to settle payments for transactions in goods markets – its negotiability in the secondary market, in other words, and its consequent monetisation – has not been addressed. It is however precisely this fact which distinguishes the large scale NCS observed in the transition economies from trade

---

<sup>8</sup>Note that the two theories are not mutually exclusive. More precisely, therefore, non-rejection of  $H(1)$  indicates that NCS trade credit can be explained *at least in part* as a financial phenomenon - though some of it may be driven by imperfections in goods markets and associated real motives as well. Rejection of  $H(1)$  however rules out the possibility that NCS trade credit is a response to financial market imperfections.

credit as it typically exists in developed, market economies. In the latter context, trade credit is typically highly illiquid, and does not serve as a means of payment. A third, important hypothesis which we can test formally in the case of Croatia is therefore:

$H(3)$  : The private credit instruments used to conduct NCS are substitutes for transactions money.

If  $H(3)$  is rejected, then we can infer that although (by definition) NCS trade credit instruments are being offset against one another, this is happening on a predominantly bilateral basis, and so does not represent the monetisation of trade credit. If, however,  $H(3)$  is not rejected, then we can infer that NCS trade credit is not only a financial, but a monetary, phenomenon – and estimates of the elasticity of substitution between official money and NCS trade credit are then highly informative for monetary policy.

If  $H(3)$  is not rejected, and NCS trade credit is found to be private monetary credit, a final step in our analysis would be the testing of hypotheses concerning the determination of the constraints on the negotiability of NCS trade credit in the secondary market. The aim of this would be to improve our understanding of why these constraints appear to be relatively low in Croatia, but relatively high in most developed, market economies. Theories of the secondary market for private credit are, however, much less developed than theories of the primary market discussed above. In classical asset pricing theory, an analytical distinction between the two is generally not made; and even in theories of monetary credit such as that of Kiyotaki and Moore (2001, 2002a, 2002b), where constraints in the primary and secondary markets for credit are carefully distinguished and play a central role, the latter are treated as exogenous<sup>9</sup>. In the following sections, we therefore limit our formal investigation of the secondary market for NCS trade credit in Section 2.4 to the testing of  $H(3)$ . We include, however, a discussion of the

---

<sup>9</sup>Although Kiyotaki and Moore (2002b) do suggest in passing three types of information asymmetry that might influence the constraint on the negotiability of credit in the secondary market. If in the course of striking the original bilateral financial contract, the creditor acquires private information on the creditworthiness of the borrower, third parties may fear adverse selection in the acceptance of such contracts in settlement of a debt. Alternatively, the reassignment of a bilateral financial contract may give the original creditor and debtor incentives to strip the assets of the investment project to which the contract gives a claim – and this moral hazard will again discourage resale of the contract. Finally, the original creditor may have special leverage over the borrower which will not be shared by a third party to whom the claim is resold – thereby once again giving third parties a disincentive to buy.

potential determinants of the unusually high degree of monetisation of trade credit in Croatia in the institutional analysis in Section 2.3.

## **2.3 Institutional Analysis**

Empirical tests of models which involve credit rationing are inherently difficult to implement, since by assumption observed prices cannot be taken to represent equilibrium outcomes. In order to guide the econometric analysis in Section 2.4, we therefore present in this section an informal, institutional analysis of developments in the Croatian financial sector since the beginning of transition. We focus first on the primary market for credit, and analyse how the combination of a high level of financial dollarisation with delays in the structural reform of the Croatian banking sector created conditions under which the majority of the enterprise sector found it difficult to access bank credit despite underlying solvency. We then analyse the structural factors that might have supported the secondary market for NCS trade credit, including the unusual payments system infrastructure inherited from the Socialist Federal Republic of Yugoslavia (SFRY).

### **2.3.1 The primary market for credit**

#### **Financial dollarisation and the supply of bank credit**

Croatia inherited from the SFRY a very high degree of financial dollarisation, and a key feature of the supply side of the bank credit market during transition has been the predominance of foreign currency sources in the funding of credit creation. As Table 2.9 demonstrates, foreign currency deposits (FCDs) of Croatian residents and foreign currency loans (FCLs) from foreign banks have together accounted for over 60% of Croatian banks' liabilities throughout the decade between 1994 and 2003. As in many middle income countries, this high level of financial dollarisation has contributed to a 'fear of floating' induced by the potential effects of exchange rate fluctuations on the balance sheet of the financial sector (Baliño, Bennett, and Borzenstein (1999); De Nicoló, Honohan, and Ize (2003); Levy Yeyati (2003)). One important result of this has been that the process of money and credit creation has been substantially exogenous to the policies of the Croatian National Bank, being driven predominantly by capital inflows

little influenced by HRK interest rates<sup>10</sup>. The extent to which lending to the enterprise sector in particular has been driven by inflows of foreign currency sources of funds is illustrated by Figure 2-1.

In low and middle income countries, the combination of a *de facto* fixed exchange rate regime with reliance on FCDs and capital inflows is generally taken to be *prima facie* evidence of potential credit rationing deriving from an aggregate foreign currency borrowing constraint on the one hand and the fickleness of residents as providers of FCDs on the other<sup>11</sup>. Perhaps surprisingly, however, Croatia has on average enjoyed relatively rich access to foreign currency sources of funds during the transition. As is evident from Figure 2-2, the level of real FCDs rose constantly over 1994-2003, and even the blocked FCDs of residents were redeposited with Croatian banks as they were released over the mid- to late 1990s – something of a remarkable achievement, given the political events of the period<sup>12</sup>. As for FCLs, Western European banks showed few qualms about lending to Croatian banks in the early 1990s, with German banks in particular willing to take on such indirect exposure to the Croatian market, and foreign loans growing at 15% a year in 1994 and 1995.

Croatia did not entirely buck the middle income country trend in supply side liquidity problems, however. In 1996 and 1998 there were major banking crises in Croatia, the latter of which was exacerbated by the contagion effects of the Russian financial crisis of August, 1998 (Jankov (2000)). Whilst the crises themselves were asset-driven (and are therefore analysed below in greater detail), they led to supply side liquidity constraints of differing magnitudes. As Table 2.9 shows, both FCDs and FCLs were affected. Following the 1996 crisis, the rate of growth of FCLs reversed from 16% in 1995 to -18%. The supply of FCDs was less affected by the 1996 crisis, since it was resolved by recapitalisation and without the bankruptcy of any

---

<sup>10</sup>This is not to say that the Croatian National Bank has had no influence on monetary aggregates: the Croatian National Bank has deployed administrative measures – particularly changes to the reserve requirements associated with FCDs – as well as intervention in the foreign exchange market and direct lending to banks as instruments of monetary policy, as discussed further below.

<sup>11</sup>For a recent general theoretical framework for the aggregate dollar borrowing constraint, see Caballero and Krishnamurthy (2002). For discussion of both this and the problem of ‘warm capital’ represented by FCDs in the case of Croatia, see Vujcic (1998) and Kraft (2003).

<sup>12</sup>In 1991, the FCDs of Croatian residents were appropriated by the National Bank of the SFRY. In 1992, the government of newly-independent Croatia issued FC-indexed bonds against these deposits as counterpart assets to these FCDs. The FCDs themselves were then blocked until July, 1995, whereafter they were to be unfrozen over a period of ten years, or released immediately but at a discount under a number of schemes. As Figure 2.2 shows, this latter option was the one favoured by most Croatian depositors.

banks; but its growth nevertheless slowed substantially over 1996 and 1997.

The 1998 crisis, following which several banks went bankrupt, had stronger and more prolonged supply side effects. With regard to FCDs, there were runs on several banks not directly affected by the crisis, which caused the total volume of FCDs to decline in 1999 for the first time since the beginning of transition. As for the supply of FCLs, a group of German banks which suffered large losses due to exposure to bankrupted banks refused to roll over lending to the Croatian banking sector throughout 1998-2000. This so-called 'German boycott' led to a collapse in the rate of growth of FCLs to the Croatian banking sector from 17% in 1998 to 5% in 1999 and 4% in 2000. The combination of these two supply side effects contributed to a severe shortage of liquidity for Croatian banks throughout 1999 and early 2000. In the second quarter of 1998, there was a sudden rise in overnight interest rates on the Zagreb Money Market (ZMM), and a number of banks withdrew from the market altogether. The CNB responded by suspending the ZMM in June, 1998 and introducing a system of administrative allocation of free reserves across the system in place of overnight interbank credits. By the beginning of 1999, even the aggregate free reserves of the commercial banks were not sufficient to meet daily liquidity needs, and the CNB in addition introduced direct lending to illiquid banks. In the first two quarters of 1999, direct credit from the CNB was the main source of money creation in Croatia, growing by 14% whilst both the total volume of other sources, and within them the sum of FCDs and FCLs, shrank by 1%. Figure 2-3 shows the magnitude of the CNB's 1999 direct lending for liquidity provision in relation to earlier and subsequent periods.

As Figure 2-2 reveals, however, these emergency measures, though successful at the margin in alleviating the most acute liquidity crunches, could not stem the general contraction of foreign currency liquidity available to the Croatian banking sector. As Table 2.11 shows, the three years 1998, 1999, and 2000 were therefore characterised by the banking system's providing virtually no new credit to the enterprise sector - real growth was 1% over the period despite annual GDP growth of approximately 4.5%. It was not until the sale of the majority of Croatia's banking sector to foreign strategic investors was completed in 2000-2001 that that the growth in FCLs picked up again, alleviating the supply side constraint on the creation of bank credit.

## **The structure of intermediation and the allocation of bank credit**

There is good reason to believe, therefore, that supply side factors may have played an important role in constraining access to bank credit in Croatia during the periods following the banking crises in 1996 and 1998-2000 - and it is likely that such supply side constraints are the most important determinants of bank credit rationing in a financially dollarised economy. Nevertheless, there are in addition important features both of the structure of intermediation and of the allocation of bank credit in Croatia, that suggest that the majority of enterprises were bank credit constrained even in the intervening periods when there is no indication that there was an aggregate shortage of foreign currency sources of funding for bank lending to the enterprise sector.

In comparison with other countries of CEE and the FSU, the financial sector in Croatia was by some standards relatively developed, as Table 2.12 shows. Credit to the private sector averaged 29% of GDP over 1994-2003, in comparison to 13% in Russia and 17% in Poland. Moreover, the growth of real credit to the private sector was strongly positive except in the crisis periods discussed above - unlike in Russia, for example, where real credit to the private sector shrank dramatically during the years in which NCS emerged. In fact, however, such statistics reflect the special structure of the SFRY banking and government sectors inherited by Croatia on independence. Because the SFRY was not a state socialist system, but operated instead a system of 'self-management' socialism, the government had very little importance either as owner of or borrower from the banking sector. Hence, instead of state-owned banks holding a substantial amount of government debt, Croatia inherited a system of 'pocket banks' owned and managed by the enterprises that had founded them, and whose portfolios were dominated by lending to related parties<sup>13</sup>. In addition, the Croatian banking sector was highly

---

<sup>13</sup>The system as it operated at the end of the 1980s, immediately prior to Croatian independence, was summarised by a contemporary Financial Times survey as follows: "Yugoslav commercial banks are not banks in the usual sense. They have been and remain service shops for enterprises. They are provided for in the 1974 constitution which stipulates that economic enterprises and other organisations may found specialist agencies to look after the conduct of credit and other bank activities, and pool in them the resources required to pursue common interests and fund the performance, expansion and promotion of their activities. Enterprises founding banks enjoy a special status as shareholders. They manage the banks' business and receive an income from their activities, after operational costs and allocations to employees have been taken into account. Hence, Yugoslavia's banks are not independent financial institutions and, as a result, "fundamental principles of banking and in particular principles of security, profitability and liquidity, have often been sacrificed for the benefit of other objectives," as one large bank put it recently. The excessive influence of the banks' parent enterprises and of political factors,

concentrated and regionally segmented at independence – with only two banks, accounting for over 50% of the total assets of the banking system, operating nationally. Neither this system of corporate governance nor these inherently uncompetitive conditions were susceptible to reform efforts until the privatisation of the large banks was begun in earnest in 2000-2001.

There were two main consequences of this for the accessibility of bank credit to the enterprise sector as a whole. The first was that bank intermediation remained relatively uncompetitive, and therefore costly. The second was that even during the periods 1994-5, and 1997 when, as Table 2.11 shows, real credit to enterprises was growing rapidly, its allocation was heavily distorted. The distortion of enterprise credit allocation in 1994 and 1995 was principally due to the lending activities of the large ‘pocket banks’ inherited from the SFRY. Prior to Croatian independence, a 1989 independent audit found that the value of bad assets in the Croatian banking system was twice the value of total bank capital; and by 1990, only five out of a total of 19 banks in Croatia were solvent. It was widely acknowledged that the corporate governance of banks was largely to blame for this accumulation of bad loans; but the crisis came to a head just as Croatia was seceding from the SFRY, with the result that the short term, strategic concern of the Croatian government to maintain a functioning system of intermediation overrode the longer term need to reform the institutional architecture. The 1991 banking crisis was therefore resolved by a so-called ‘linear rehabilitation’ of insolvent banks, in which the banking sector was recapitalised using government bonds, and the corporate governance and lending practices of the Croatian banking sector remained substantially unchanged. At the same time, Croatia was engaged in the wars on its own territory and in neighbouring Bosnia and Herzegovina, with the result that the political direction of bank lending to loss-making enterprises with a strategic importance to the war effort – such as shipyards, the railways, and agricultural conglomerates – was increased. The result was that whilst, as Table 2.11 shows, the real credit to the enterprise sector grew at over 20% a year in 1994 and 1995, its allocation was heavily skewed towards a small number of strategic enterprises. It was this concentration of credit growth to enterprises in essentially insolvent sectors which resulted in the banking crisis of 1996. In the process of rehabilitation, over HRK 6 billion of bad loans were written off the books of four large banks,

---

have for example, reduced capital flows across regional boundaries, and encouraged non-observance of monetary and credit policy measures, poor scrutiny of investment projects, and irregular or grey money issue.” *Financial Times*, December 22, 1987, p.11.

equivalent in total to 22% of the stock of credit to enterprises at the end of 1995<sup>14</sup>. Over 90% of these bad loans were to several large enterprises under government ownership<sup>15</sup>.

Once the 1996 rehabilitation programme was underway, strong credit growth resumed in 1997: as Table 2.11 shows, credit to enterprises grew by 37% over 1997, and personal credit nearly doubled. It was now predominantly driven not by the old socially-owned banks, however, but by a group of new, private banks which pursued aggressive strategies for acquiring market share by offering high interest rates on both deposits and loans. Nevertheless, the allocation of this rapid credit creation proved to be just as distorted as the politically directed credit growth of 1994-5. The portfolios of the new, private banks were heavily skewed towards related parties and politically influential business groups; and whether by accident, or, as was suspected in some cases, by design, it quickly transpired that they were of poor quality. By the beginning of 1998, fifteen banks, accounting for approximately one quarter of the total assets of the banking system, were deemed to be financially distressed by the CNB: over the course of the year, four were selected to rehabilitation, and eleven were sent to bankruptcy, prompting the supply side problems described above<sup>16</sup>.

Since the beginning of 2001 there has again been an increase in the rate of growth of credit to the enterprise sector. Although this period is of less relevance to our analysis, since our observations on the volume of NCS extend only to the end of 2001, this development deserves some comment, because of the striking contrast it makes with the evolution of bank credit in the Czech Republic after its major banking crisis in 1997. In both cases, the aftermath of the crisis saw the substantial privatisation of the banking sector, so that subsequent lending trends were driven by large, foreign-owned, privatised banks operating in an increasingly competitive environment with improving supervision and corporate governance<sup>17</sup>. In the Czech Republic, the strategy of the foreign banks was to squeeze hidden non-performing loans out of their portfolios by avoiding new enterprise lending for a number of years, resulting in a continuous decrease in credit to the enterprise sector in every year since 1997. In the Croatian case, by contrast, enterprise lending has recovered exceptionally rapidly – despite the fact that a

---

<sup>14</sup> IMF (1998), p.9.

<sup>15</sup> Jankov (2000), p.4.

<sup>16</sup> IMF (2000), p.37.

<sup>17</sup> See Kraft (2001) for an analysis of the Croatian case.

similar problem of potentially bad assets plagued the privatised loan portfolios. A possible explanation for this contrast lies precisely in the high level of NCS, and the hypothesis of this paper that this was based on trade credit. On this argument, the privatised banks have followed the identical strategy regarding inherited portfolios as in the Czech case, but have combined this with the rapid accumulation of new portfolios through the bringing of profitable firms, previously rationed out of the market bank credit and hence financed by NCS trade credit, into the banking system. If this explanation were valid, we would therefore expect to see a reduction in NCS trade credit roughly proportionate with the increase in bank credit to enterprises in 2001-3. Since our data on NCS do not extend beyond 2001, however, we cannot subject it to such a test.

## **Summary**

Taking the previous two subsections together, therefore, there are strong grounds on which to infer that the great majority of Croatian enterprises found it costly or impossible to access the primary market for bank credit throughout the 1990s, despite the beneficial macroeconomic environment. The supply of bank credit was largely exogenous, due to the high level of financial dollarisation and the 'fear of floating' that this induced. In those periods when this supply was not constrained – such as 1994-5 and 1997 – and real credit to the enterprise sector in aggregate grew rapidly, the allocation of this credit growth was fatally skewed; during the first period towards loss-making strategic enterprises, and in the second towards related parties and pyramid schemes. Both episodes therefore led to substantial banking crises; and these led to periods in which the supply of foreign currency funding sources – and hence the supply of credit to enterprises – was constrained. Moreover, throughout the period under analysis, bank intermediation remained uncompetitive and therefore expensive as a competitor to trade credit. A final factor in the availability of bank credit was the Croatian government's increasing accumulation over the late 1990s of net payment arrears mentioned in Section 2.1, which must further have aggravated the existing shortage of monetary liquidity.

### 2.3.2 The secondary market for credit

Evidence of rationing in the primary market for credit during the 1990s can explain the emergence of high level of trade credit in Croatia. By definition, there is also an active secondary market for NCS trade credit, however – and this presents an analytically distinct puzzle, since even in developed, market economies where the use of bilateral trade credit is widespread, the secondary market for it is very limited. Whilst, as mentioned above, there are few theoretical contributions in the literature to guide an analysis of the structural determinants of the liquidity of the secondary market for trade credit, we can nevertheless propose three factors that have been important in the Croatian case.

The first of these relates once again to the institutional legacy of the SFRY. Although the SFRY had a Western-style two-tier banking system, rather than a mono-bank system like those favoured in the Eastern Bloc, payments clearing was centralized in the *Služba za Društveno Knjigovodstvo* (Social Accounting Service – SDK) rather than operated in a decentralised manner by commercial banks. Moreover, the SDK was also responsible for the central keeping of enterprise financial accounts on a real-time basis. In addition to their bank giro accounts, therefore, all enterprises were required to hold an account at the SDK, in which changes in accounts receivable and payable were registered on an accrual basis, as well as the settlement of these accounts in bank money implemented by transfers from giro accounts. Upon independence in 1991, the SDK was transformed into the *Zavod za Platni Promet* (ZPP – Trade Payments Unit). The *de iure* responsibility for financial accounting was transferred to enterprises; but until 1999, the ZPP remained in charge of payments clearing, when a decentralised wholesale payments system was introduced; and until 2001, when payments system reform was completed by the introduction of a bank-based retail payments clearing system, the ZPP continued to maintain a real-time register of all enterprises' current assets and liabilities.

The importance of the existence of this unusual payments system infrastructure for the viability of using trade credit as a means of payment should not be underestimated. As the Soviet experience of the 1970s demonstrated, an institution that combines centralised enterprise accounting and payments clearing represents the ideal forum for the development of NCS – since the costs, in terms of both information and commitment, of identifying and agreeing non-monetary giros for the multilateral offsetting of private credits and debts are enormously

reduced. In fact, these costs are reduced to the same level as the costs of settlement using bank money, since the payments system employed for both is identical. In the Croatian case, the ZPP itself was involved directly in the organisation of inter-enterprise NCS only in the first half of 1993: following the stabilisation, it was not permitted by the Croatian government to take an active role in their orchestration. Nevertheless, the complete database of enterprise accounts receivable and payable existed until 2001, and remained an important resource for the decentralised organisation of NCS from late 1993 on.

The second factor facilitating the decentralised secondary market for NCS trade credit was simply the small size of the Croatian economy, and the dominant role played by a relatively close-knit political and business elite in Croatian business and finance. As in many post-socialist countries, senior management positions continued to be held by the communist management cadre, and in Croatia this strong pre-existing network was complemented by the new network constituted by apparatchiks of the new nationalist regime. Since enterprise privatisation progressed relatively slowly in Croatia, and when it did progress tended to consist of privatisation to insiders, these strong networks in a small economy are another reason for the low identification and commitment costs associated with the multilateral offsetting of private credits and debts relative to most developed, market economies.

Finally, the structure and distribution of the non-monetary giros in which NCS trade credit is offset in Croatia indicate a third reason why it was able to become so widespread. Although there is no aggregate or survey data available on the networks of enterprises participating in NCS, interviews undertaken for this study indicated that non-monetary giros tend to cluster around large private and parastatal firms such as utilities and oil companies. One of the principal means by which multilateral offsets are effected is by the practice of *cesije* – the successive endorsement of an account payable by, for example, *Hrvatska Elektroprivreda* (HEP), the Croatian electricity utility, round a chain of trading partners, and its subsequent extinction by the last of them in payment of an electricity bill. The importance of such ‘nodes’ underlying the aggregate statistics on NCS is confirmed by the fact that HEP reports that the average monthly proportion of its sales revenue from enterprises settled via NCS in 2001 was, at around 32%, nearly twice the level reported in the ZPP dataset for the enterprise sector in aggregate,

at 17%<sup>18</sup>.

## 2.4 Econometric Analysis

Having analysed informally the structural factors of NCS trade credit and its monetisation, we now proceed to test  $H(1)$ ,  $H(2)$ , and  $H(3)$  econometrically. Previous econometric analyses of NCS have relied exclusively on microeconomic (survey) data. Even if it were available, such data would be of little use in testing our hypotheses, which postulate variation over time of the volume of private credit instruments used to conduct NCS with respect to various macroeconomic variables. Fortunately, the existence in Croatia, until 2001, of the centralised payments clearing system described above means that we are able to assemble a unique macroeconomic dataset on the volume of NCS. This ZPP dataset contains two time series: the first, which we denote  $Y_t^M$ , on the total volume of inter-enterprise transactions settled using bank money (cash and bank deposits); and the other,  $Y_t^{NCS}$  on the total volume of inter-enterprise NCS. There are 96 monthly observations on each of these two series, running from January, 1994 to December, 2001. We denote their sum  $Y_t$ : the total volume of inter-enterprise transactions, settled via either bank money or NCS.

### 2.4.1 Testing $H(1)$ and $H(2)$

Testing  $H(1)$  and  $H(2)$  can be done by imposing restrictions on an empirical equation of the general form:

$$\frac{Y_t^{NCS}}{Y_t} = \alpha' \mathbf{z}_t + \varepsilon_t \quad (2.1)$$

where  $\mathbf{z}_t$  is a vector of variables which measure the rationing of bank credit<sup>19</sup>. The restriction corresponding to  $H(1)$  is then simply that there is a stable, systematic relationship between  $\frac{Y_t^{NCS}}{Y_t}$  and the variables in  $\mathbf{z}_t$  in Equation (2.1); whilst the restriction corresponding to  $H(2)$  is that the sign of the coefficients on the variables in  $\mathbf{z}_t$  should be positive.

---

<sup>18</sup>Interviews conducted with CNB staff during July-September, 2004.

<sup>19</sup>Both  $H(1)$  and  $H(2)$  strictly concern the volume of private credit instruments used to settle NCS, rather than the volume of NCS. Direct observations on the former variable are however not available. Restrictions on Equation (1) can nevertheless be used to test  $H(1)$  and  $H(2)$  by deploying the ratio  $\frac{Y^{NCS}}{Y}$  in its stead, if we make the (reasonable) assumption that this ratio and the ratio of NCS trade credit instruments to total monetary instruments maintain a constant proportional relationship over our sample.

## Data series

The components of the series describing the dependent variable in Equation (2.1),  $\frac{Y_t^{NCS}}{Y_t}$ , are taken from the ZPP dataset. On the basis of the analysis in the previous section, the following time series were collected to fill the vector  $\mathbf{z}_t$  of variables controlling for the effects of bank credit rationing:

Variable	Description
$\frac{FC_t}{Y_t}$	Total (FCD+FCL) foreign currency sources of funds to the Croatian banking sector, as a proportion of total (monetary and NCS) inter-enterprise transactions.
$LIQU_t$	A variable that combines the overnight interest rate on the ZMM ( $MM\_ON$ ) with the volume of CNB direct lending to distressed banks ( $CNBCR$ ) into a single series <sup>20</sup> .
$SPR_t$	Spread between the average short term loan and term deposit rates available to enterprises.

These variables are chosen in order to reflect the three determinants of rationing in the primary market identified in Section 2.3 above:

Variable	Rationale
$\frac{FC_t}{Y_t}$	Controls for supply side liquidity shortage due to financial dollarisation, being a measure of the availability of foreign sources of financing for credit creation relative to total enterprise sector economic activity ( <i>n.b.</i> that $Y$ measures total, not only monetary, transactions).
$LIQU_t$	Controls for demand side liquidity shortage, being a leading indicator of the volume of bad assets in the banking system.
$SPR_t$	Controls for rationing due to the inefficiency of public intermediation, being a measure of the costliness of intermediation by banks.

<sup>20</sup> $LIQU$  is calculated according to the formula  $LIQU = MM\_ON + (\frac{\mu_M}{\mu_C}) \ln(CNBCR)$ , where  $\mu_M$  and  $\mu_C$  are the means of  $MM\_ON$  and  $\ln(CNBCR)$  over the period 1996.10 to 1998.1 when the ZMM was operating, and there was virtually no direct lending by the CNB.

The construction of the variable  $LIQU_t$  calls for some explanation. The most natural candidate to serve in this role is simply  $MM\_ON$  alone; but as can be seen in Figure 2-4, this variable is constant - and hence uninformative - during the period from June, 1998 to April, 2000 when the ZMM was suspended (see Section 2.3.1 above). We considered two ways of controlling for variation in liquidity conditions during this period. The first was simply to include a dummy variable taking the value of 1 over the period alongside  $MM\_ON$ . This is clearly a very crude control, however; and so the preferred option was to use the information conveyed by the volume of CNB direct lending ( $CNBCR$ ) over the suspension period<sup>21</sup>. This was done by combining  $MM\_ON$  and the relevant sub-sample of  $CNBCR$  into a single series, illustrated graphically in Figure 2-5, according to the formula described above.

All variables were assembled using data provided by the Croatian National Bank. There are 91 observations on  $LIQU_t$  and  $\frac{FC_t}{Y_t}$ , from June, 1994 to December, 2001; but only 77 observations on  $SPR_t$ , because the constituents of this series are available only from August, 1995 onwards. The dependent variable series  $\frac{Y_t^{NCS}}{Y_t}$  was transformed into its natural logarithm (henceforth denoted  $(y^{NCS} - y)_t$ ), in order to allow the use of a semi-logarithmic functional form. All series for estimation are illustrated graphically, along with their main time series characteristics, in Figures 2-6 (levels) and 2-7 (first differences). Visual inspection suggests that all series may be I(1). ADF tests can be used to test this inference more formally; and the results of such tests, with the appropriate augmentation chosen according to the Akaike Information Criterion, are reported in Tables 2.13 and 2.14. For all series other than  $(y^{NCS} - y)_t$ , the initial inference of integration of order 1 is strongly corroborated. The results are less conclusive for  $(y^{NCS} - y)_t$ , which the test results suggest may be stationary around a constant, or around a constant and a trend. Given, however, the low power of the ADF tests and the fact that under minimally different lag specifications evidence of stationarity disappears, we proceed on the conservative assumption that  $(y^{NCS} - y)_t$  may also have a unit root.

---

<sup>21</sup> Although we report below in footnotes specifications using  $MM\_ON$  only and  $MM\_ON$  and a dummy variable for purposes of comparison.

## Modelling strategy

Estimation of Equation (2.1) starting from an unrestricted vector auto-regression (UVAR) is not an attractive option, given the small sample size; using a General-to-Specific methodology in any case indicated that a single equation specification was appropriate<sup>22</sup>. Given the order of integration of the variables assembled, modelling Equation (2.1) in first differences is the most obvious option. However, if a cointegrating combination of the I(1) levels of some or all of the variables can be identified, then this may also enter the I(0) equation. We therefore deployed two single equation estimation methods in order to arrive at congruent empirical models of Equation (2.1) on which the restrictions corresponding to  $H(1)$  and  $H(2)$  could be imposed. The first of these is the two step dynamic modelling procedure, in which a cointegrating combination is first identified from the static solution to a dynamic equation in levels, and then included as an equilibrium correction (EC) term at one lag in the I(0) equation in first differences. The second is a one step procedure, in which the first lag of the variables in levels are themselves included in the I(0) equation in first differences. The two procedures are statistically equivalent, and so should yield statistically identical results if the cointegrating combination is valid.

## Results

The two step dynamic modelling procedure generates the following estimate of a long run cointegrating relationship between  $(y^{NCS} - y)$  and  $LIQU$ :

	<i>Cons</i>	<i>LIQU</i>
Est.	-1.29	1.37
t value	(-5.07)	(2.17)
Calc.	0.33	0.25
$\sigma = 0.09 \quad R^2 = 0.67 \quad DW = 1.97 \quad Obs = 88$		

Note: specification includes seasonal dummies but no trends and is estimated over 1994.9-2001.12

**Table 2.1:** Two Step Procedure: Static Solution to Estimated Dynamic I(1) Model

Where in the row "Calc.", we give (a) the coefficient on the constant term in the equation in levels, calculated by taking the exponent of the estimated coefficient; and (b) the elasticity

<sup>22</sup>Because only one cointegrating combination could be identified; and all variables except  $(y^{NCS} - y)_t$  were found to be weakly exogenous.

at the mean of  $\frac{Y_t^{NCS}}{Y_t}$  with respect to  $LIQU_t$  implied by the estimated coefficient<sup>23</sup>. Figure 2-8 illustrates the estimated residuals of this equation,  $\hat{\varepsilon}_t$ : these appear stationary, and the DF test reported in Table 2.15 corroborates this. The results of modelling in first differences are as follows:

	<i>Cons</i>	$\Delta \frac{FC_t}{Y_t}$	$\Delta SPR_t$	$\hat{\varepsilon}_{t-1}$
Est.	0.00	-0.37	2.01	-0.77
t value	(0.07)	(-4.81)	(1.95)	(-4.54)
$\sigma = 0.12 \quad R^2 = 0.42 \quad DW = 2.19 \quad Obs = 77$				

Note: specification estimated over 1995.8-2001.12

**Table 2.2:** Two Step Procedure: Estimated I(0) model

Table 2.2 reports the I(0) partial adjustment model arrived at after the removal of insignificant variables (with the exception of the constant term) corresponding to the first stage model in Table 2.1. The coefficient on the first difference of  $LIQU$  is not significant, and is therefore omitted. The coefficient on the EC term,  $\hat{\varepsilon}_{t-1}$ , is statistically significant and correctly signed, and its estimated value implies a half life of deviations from the long run relationship between  $\frac{Y_t^{NCS}}{Y_t}$  and  $LIQU$  of approximately two weeks.

Estimation of Equation (2.1) using the one step procedure generates the following results:

	<i>Cons</i>	$\Delta \frac{FC_t}{Y_t}$	$\Delta SPR_t$	$(y^{NCS} - y)_{t-1}$	$LIQU_{t-1}$
Est.	-0.96	-0.32	1.98	-0.52	0.44
t value	(-4.97)	(-4.22)	(1.96)	(-5.11)	(2.05)
Calc.					0.15
$\sigma = 0.12 \quad R^2 = 0.46 \quad DW = 2.21 \quad Obs = 77$					

Note: neither seasonal dummies nor trend included; estimation period 1995.8-2001.12

**Table 2.3:** One Step Procedure: Estimated Model

Here, as in Table 2.1 above, the number reported in the row "Calc." is the long run elasticity at the mean of  $\frac{Y_t^{NCS}}{Y_t}$  with respect to  $LIQU$  implied by its estimated coefficient, which in this case can be identified and calculated by imposing the assumption (supported by second stage

<sup>23</sup>Using  $MM\_ON$  instead of  $LIQU$  yields an estimated long run coefficient of 1.26 (t-value 1.96); using  $MM\_ON$  and a dummy to control for the ZMM suspension period yields a coefficients of 1.21 (2.48) on  $MM\_ON$  and 0.03 (1.65) on the dummy. Model diagnostics are only superficially different under these alternative specifications. The specification using  $LIQU$  is therefore preferred both because of the greater precision of the estimates and because it allows an additional degree of freedom in what is already a small sample.

two step estimates) that the constant term is significant only in the long run cointegrating combination<sup>24</sup>. The estimates in Table 2.3 conform well with those in Tables 2.1 and 2.2. The cointegrating combination is less precisely estimated than in the first stage of the two step procedure (as is to be expected given that the one step procedure does not explicitly model dynamic structure) but is similar in magnitude to that recovered in the first stage of the two step procedure, as can be seen by comparing the calculated elasticities in Tables 2.1 and 2.3. The estimated coefficients on the differenced variables, meanwhile, are nearly identical to those recovered in the second stage of the two step procedure; and the implied coefficient on the EC term (equal to the estimated coefficient on  $(y^{NCS} - y)_{t-1}$ ) is also of a similar magnitude to that reported in Table 2.2. The two models are compared graphically in Figure 2-9. On the basis of the estimates obtained using these two procedures, we can be satisfied, in other words, that either of these specifications can serve as a congruent empirical model of Equation (2.1).

## Inference

The restriction corresponding to  $H(1)$  - that there be a stable, systematic relationship between  $\frac{Y_t^{NCS}}{Y_t}$  and the variables in  $\mathbf{z}_t$  in Equation (2.1) - is not rejected regardless of which of the two procedures is used for estimation. Likewise, the sign restrictions corresponding to  $H(2)$  - that the sign of the coefficients on the variables in  $\mathbf{z}_t$  should be positive for those variables are increasing, and negative for those that are decreasing, in the degree of bank credit rationing - are accepted in all cases. Based on this econometric analysis, therefore, we can reject neither  $H(1)$  nor  $H(2)$ .

The econometric evidence argues, in other words, that the Meltzer (1960) hypothesis is valid for NCS trade credit in Croatia: it is a form of private inter-enterprise finance generated endogenously in response to constraints in the market for publicly intermediated finance. It shows both (a) that NCS trade credit in Croatia is a financial phenomenon - that is, a means of generating and allocating across firms inter-temporal liquidity to finance production and trade - and (b) that those factors constraining firms' access to bank credit do not also constrain access to NCS trade credit, so that the latter is a substitute for, rather than a complement to, the

---

<sup>24</sup>Given this assumption, then the estimated elasticity is calculated as  $\frac{\alpha_2}{\alpha_1}(\mu_2)$ , where  $\alpha_1$  is the estimated coefficient on  $(y^{NCS} - y)_{t-1}$ , and  $\mu_2$  and  $\alpha_2$  are the mean of and the estimated coefficient on the first lag of  $LIQU_t$  respectively.

former. This second inference is consistent with the hypotheses concerning the sources of bank credit rationing reflected in our choice of variables to measure the rationing of public credit - Croatia's high level of financial dollarisation on the one hand, and the structural deficiencies of its banking sector on the other.

### 2.4.2 Testing $H(3)$

Testing  $H(3)$  can be done by estimating a transactions demand for money function of the following form:

$$\frac{M_t}{P_t} = \beta_0 + \beta_1 \frac{Y_t^M}{P_t} + \beta_2 \frac{Y_t^{NCS}}{P_t} + \beta_3 OPCOST_t + u_t \quad (2.2)$$

The specification of Equation (2.2) uses a technique analogous to one popular in empirical analyses of currency substitution. It is conventional in that literature to include a variable measuring the opportunity cost of holding foreign currency in a conventional transactions demand for money equation, and to infer substitutability between domestic and foreign currency if its estimated coefficient is statistically significant and negative. In Equation (2.2), we take advantage of the fact that we have observations on the volumes of both monetary and non-monetary inter-enterprise transactions, and include the latter as well as the former in the transactions demand for money equation on the same reasoning<sup>25</sup>. The restriction corresponding to  $H(3)$  is therefore that the sign of the coefficient on  $\frac{Y_t^{NCS}}{P_t}$  should be negative.

### Data series

Empirical application of Equation (2.2) requires a time series on a transactions monetary aggregate and an appropriate opportunity cost variable in addition to the series  $Y^M$  and  $Y^{NCS}$ . The choice of the former is straightforward (HRK  $M1$ ); the choice of the latter, less so. Previous studies of money demand in Croatia have found that conventional opportunity cost variables are not significant for the development of transactions money (Babić (2000); Erjavec and Cota (2001)); and this is indeed what one would expect in an economy where domestic currency is barely held as a store of value, but essentially performs a purely transactional role. Visual

---

<sup>25</sup>In fact, the inclusion of a relevant transactions series is much preferable to the inclusion of a relevant opportunity cost series, since the former is according to most theories a genuine determinant of transactions demand, whereas the latter is in fact a determinant of store-of-value demand.

inspection of the evolution of real  $M1$  illustrated in Figure 2-10 nevertheless indicates clearly a rapid increase in its growth rate, above the growth rate of GDP, starting in mid-2000; and this conforms with the simultaneous relaxation of the main HRK policy instrument of the CNB - the level of the unremunerated reserve requirement (URR) for Croatian banks. We therefore include a variable,  $URR$ , designed to reflect variation in the opportunity cost of holding HRK transactions balances as the CNB progressively loosened monetary policy<sup>26</sup>.

The variables  $M1$ ,  $Y^M$ , and  $Y^{NCS}$  were all divided by the CPI to generate real money and transactions series, and all three series and the constructed series  $URR$  were transformed into their natural logarithms. The resulting series are denoted  $(m1 - p)$ ,  $(y^M - p)$ ,  $(y^{NCS} - p)$ , and  $urr$ . All four series for estimation are illustrated graphically, along with their main time series characteristics, in Figures 2-11 (levels) and 2-12 (first differences). Once again, visual inspection suggests that the series may be  $I(1)$ , and the results of ADF tests, reported in Tables 2.16 and 2.17, corroborate this inference.

### Modelling strategy

As in the case of Equation (2.1), starting from a UVAR reveals a single cointegrating vector and the weak exogeneity of  $(y^M - p)_t$  and  $(y^{NCS} - p)_t$ , so that proceeding to estimate Equation (2.2) as a single equation is justified<sup>27</sup>. Also as in the case of Equation (2.1), we proceed to do so using both the two step dynamic modelling and the one step procedures.

### Results

The two step dynamic modelling procedure generated the following estimated static solution corresponding to Equation (2.2):

Figure 2-13 illustrates the estimated residuals of this equation: DF tests reported in Table 2.18 suggest that these are stationary. Denoting these estimated residuals  $\hat{u}_t$ , the result of modelling in first differences is as follows:

---

<sup>26</sup>The variable  $URR$  takes a value of zero until 2000.4, and is equal to the level of the URR on HRK deposits (scaled so that its value in 2000.3 equals zero) thereafter. Since it is in the nature of a dummy variable, inference should not be made on its estimated coefficient.

<sup>27</sup>The  $\chi^2(2)$  statistic for the test of excluding  $(y^M - p)$  and  $(y^{NCS} - p)$  from the matrix of adjustment loadings in the corresponding CVAR was 4.38 (5% critical value 5.99).

	<i>Cons</i>	$(y^M - p)$	$(y^{NCS} - p)$	<i>urr</i>
Coefficient	0.47	0.55	-0.24	-0.67
t value	(1.05)	(8.80)	(-2.46)	(-8.24)
	$\sigma = 0.03$	$R^2 = 0.96$	$DW = 2.07$	$Obs = 72$

Note: specification includes seasonal dummies but no trend and is estimated over 1996.1-2001.12

**Table 2.4:** Two Step Procedure: Static Solution to Estimated Dynamic Model

	<i>Cons</i>	$\Delta(m1 - p)_{t-1}$	$\Delta(y^M - p)_t$	$\Delta(y^{NCS} - p)_{t-2}$	$\Delta\Delta urr_t$	$\hat{u}_{t-1}$
Coefficient	0.00	0.59	0.32	0.06	-0.41	-0.61
t value	(0.50)	(5.67)	(9.29)	(3.80)	(3.01)	(-3.50)
	$\sigma = 0.03$	$R^2 = 0.65$	$DW = 2.06$	$Obs = 71$		

**Table 2.5:** Two Step Procedure: Estimated I(0) Model

Once again we arrive at a partial adjustment equation. Acceleration, rather than the change, in the variable  $urr_t$  is found to effect the development of  $\Delta(m1 - p)_t$ . The estimated coefficient on the EC term,  $\hat{u}_{t-1}$ , is statistically significant and correctly signed, confirming that our inference of cointegration between the I(1) series in levels made on the basis of the first stage static solution was valid, and indicating a half life of stochastic deviations from the long run equilibrium of a little less than two months.

Estimation of Equation (2.2) via the one step procedure gives the following results:

	<i>Cons</i>	$\Delta(y^M - p)_t$	$\Delta(y^{NCS} - p)_t$	$\Delta urr_t$
Est.	0.72	0.16	-0.08	-0.42
t value	(3.82)	(3.11)	(-2.29)	(-2.61)
	$(m1 - p)_{t-1}$	$(y^M - p)_{t-1}$	$(y^{NCS} - p)_{t-1}$	$urr_{t-1}$
Est.	-0.41	0.20	-0.11	-0.31
t value	(-3.84)	(3.14)	(-2.79)	(-4.37)
Calc.		0.49	-0.27	-0.75
	$\sigma = 0.03$	$R^2 = 0.81$	$DW = 1.74$	$Obs = 72$

**Table 2.6:** One Step Procedure: Estimated Model

Here as in Table 2.3 above the numbers in the row "Calc." are the long run elasticities of the dependent variable with respect to the independent variable in question implied by the

estimated coefficient<sup>28</sup>. The estimates in Table 2.6 conform well with those in Tables 2.4 and 2.5. The dynamic specification arrived at via the one step procedure is somewhat different to that recovered in the second stage of the two step procedure, making comparison of the short run coefficients impossible - although the coefficients on the EC terms are of similar magnitudes ( $-0.61$  and  $-0.41$ ). The short run coefficients are however of less interest to us than the coefficients of the estimated long run cointegrating relationships, since it is a restriction on these that tests  $H(3)$ . These are very similar in both cases, giving us confidence that either is a good empirical model of the long run transactions demand for money described in Equation (2.2) to which this test can be applied. The two models are compared graphically in Figure 2-14.

### Inference

The restriction corresponding to  $H(3)$  - that the private credit instruments used to conduct NCS are substitutes for transactions money - is that the coefficient on  $(y^{NCS} - p)$  should be negative. As Tables 2.4 and 2.6 show, this restriction is accepted when imposed on either of the empirical models estimated above. We can therefore infer from the econometric analysis above that NCS trade credit in Croatia is not only a financial phenomenon ( $H(1)$ ) and a substitute for bank credit ( $H(2)$ ), but that it is also a substitute for transactions money ( $H(3)$ ).

## 2.5 Conclusions

NCS in Croatia emerged not in the context of collapsing real output and compromised government priorities, as in the case of Russia, but in a situation of rapidly growing real output and a government policy not to engage in NCS. The conventional explanation of NCS in transition as a means for government to deliver subsidies to an unstructured and insolvent enterprise sector without compromising stabilisation is therefore implausible in the case of Croatia.

NCS can instead be explained as private monetary credit, generated endogenously by the Croatian enterprise sector in order to meet the needs of trade. Econometric evidence presented

---

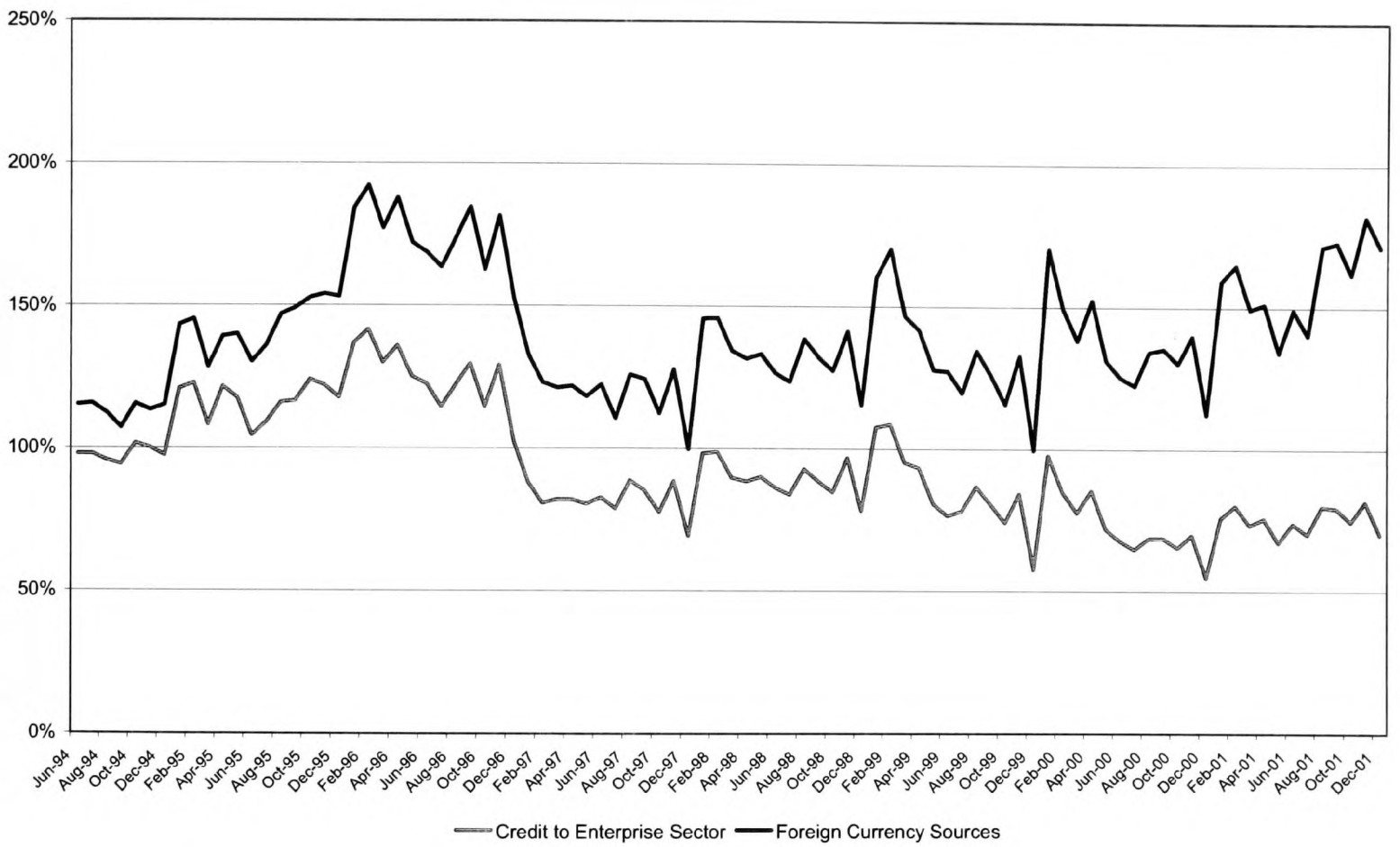
<sup>28</sup>In this case, the recovered elasticities are constant however, due to the double-log functional form of specification (2.2).

in this paper shows (a) that the credit instruments used to conduct NCS in Croatia are (at least in part) a financial, rather than a real, phenomenon; and (b) that they are substitutes not only for bank credit, but for bank money, so that the NCS have important consequences for both financial and monetary policy.

Since NCS in Croatia are at least in part a financial phenomenon, their emergence is to be explained as a result of financial as well as goods market imperfections, and 'solutions' to NCS are therefore to be sought in financial and monetary reform as well as in enterprise sector restructuring. With regard to the secondary market for credit, we have argued that the existence of a centralised payments clearing institution during the 1990s, the small size of the Croatian economy, and the participation of natural 'nodes' of NCS such as the electricity utility facilitate the negotiability of private credit in Croatia. Insofar as Croatia introduced a decentralised, bank-based payments system in 2001, and the Croatian electricity utility HEP issued an operational directive banning NCS in 2003, those secondary market imperfections amenable to policy measures have been addressed, and one might expect to see a substantial reduction in the liquidity of NCS trade credit as a result.

With regard to the primary market for credit, our analysis has shown that two main factors lie behind the rationing of public credit to the Croatian enterprise sector over the 1990s. The first factor is the structure and corporate governance of financial intermediation, which led to large distortions in the allocation of credit to the enterprise sector and to high costs of public intermediation. The second factor is Croatia's very high level of financial dollarisation. The implicit financial fragility this involves has induced a 'fear of floating' which in turn has made the supply of funding for enterprise sector lending dependent on foreign currency funding sources, which have proved fickle in times of crisis and contagion. Of these two factors affecting the primary market for credit in Croatia, policy-makers have succeeded in addressing the former by the privatisation of Croatia's main banks in 2000-01 and the ever-improving prudential and regulatory framework for the Croatian banking sector. The latter factor, however, remains; and illustrates an important link between these two forms of money substitute in Croatia.

**Figure 2-1:** Foreign Currency Sources of Funding and Credit to Enterprise Sector (both as % of total inter-enterprise transactions)



**Figure 2-2: Liabilities of the Croatian Banking Sector**

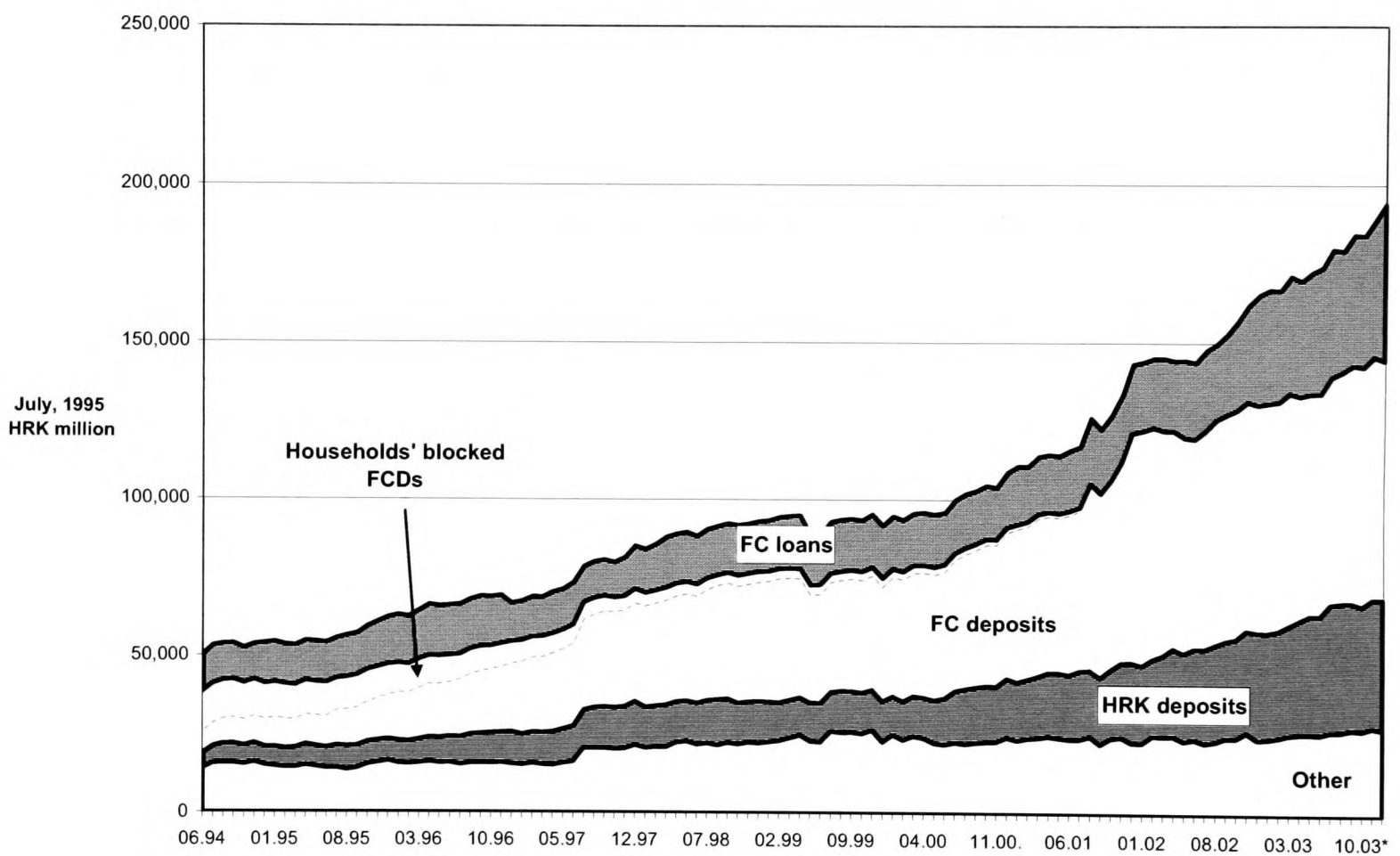
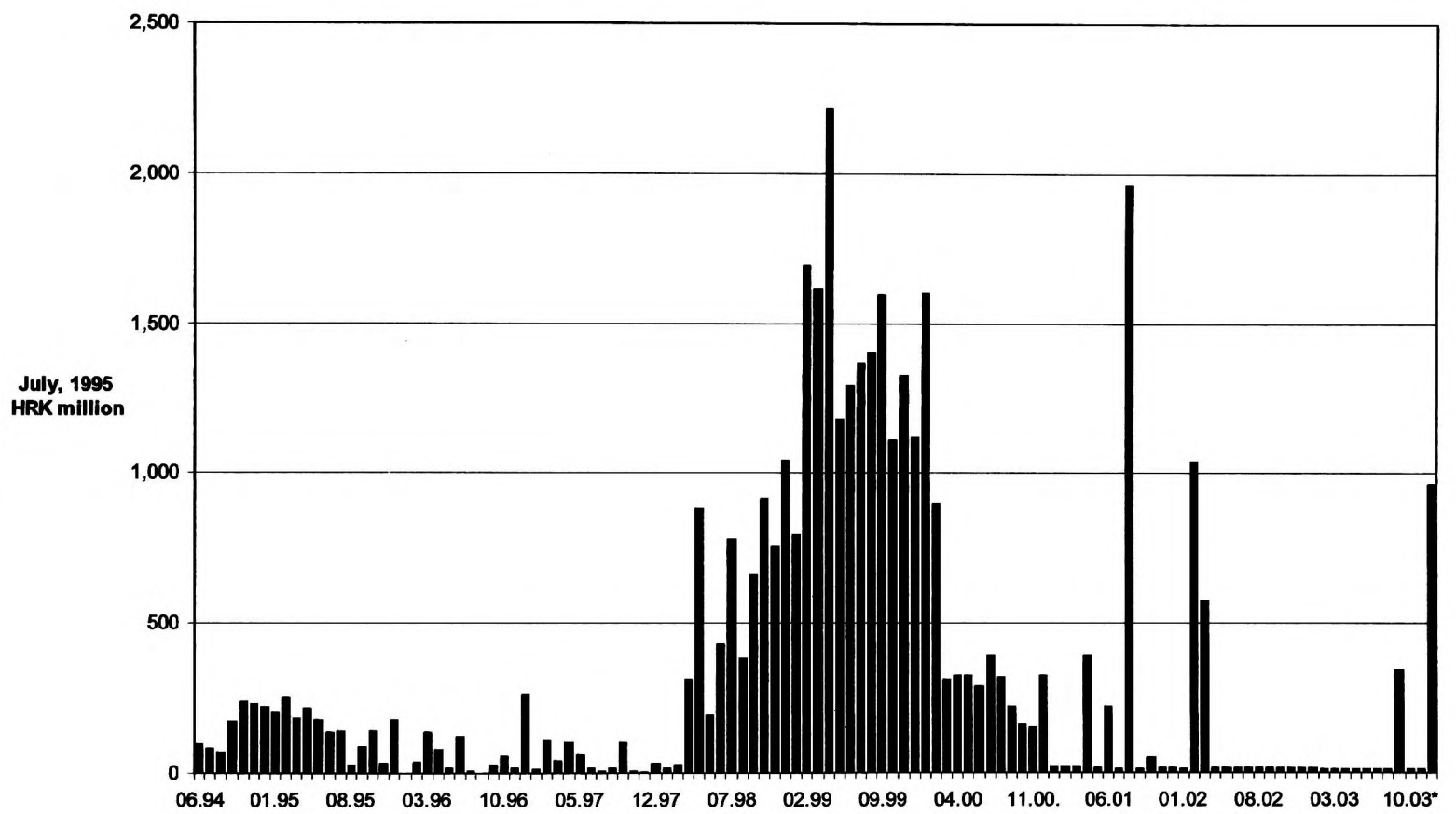


Figure 2-3: CNB Direct Lending to the Croatian Banking Sector



**Figure 2-4: Overnight rate on the Zagreb Money Market**

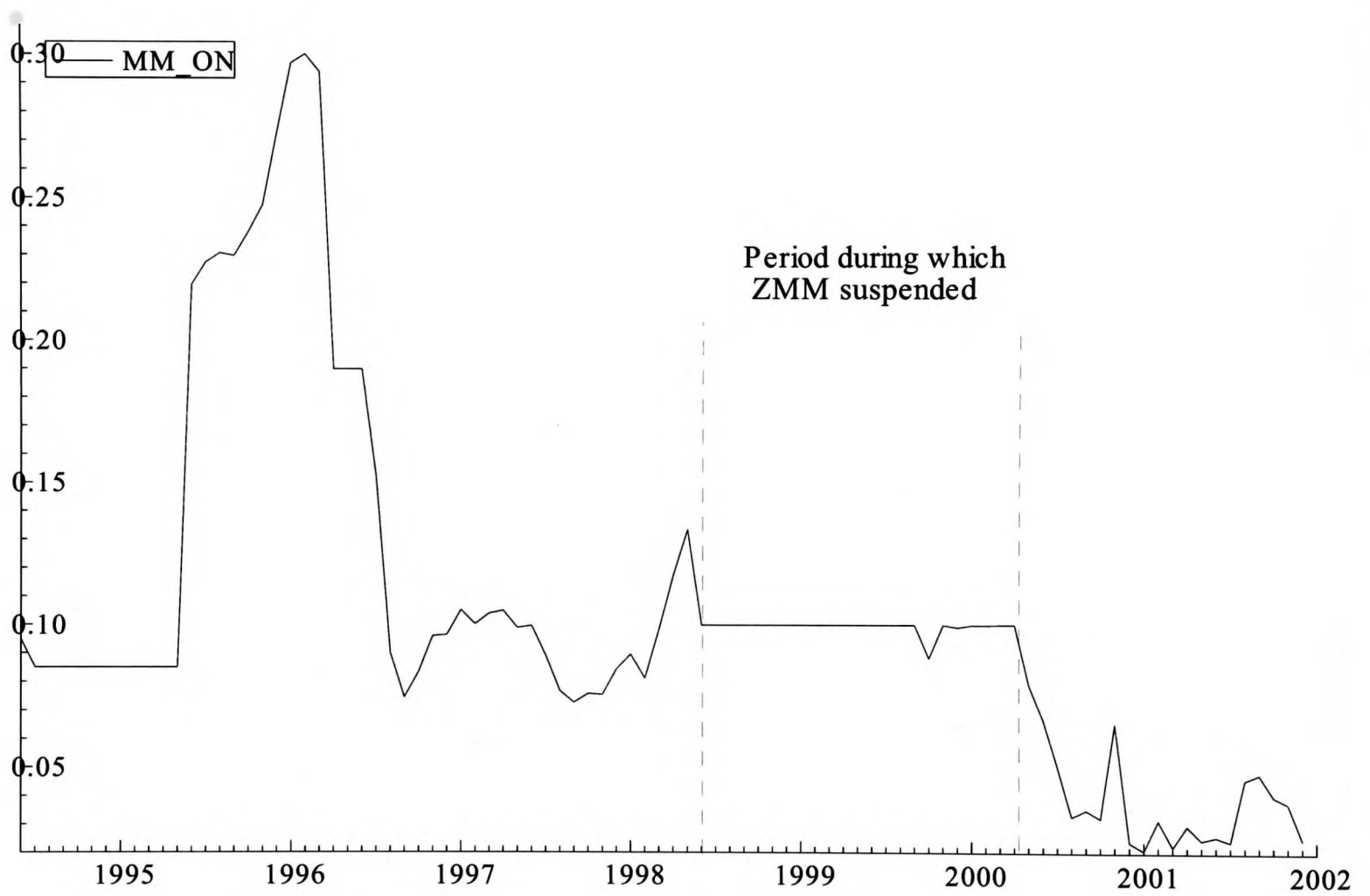
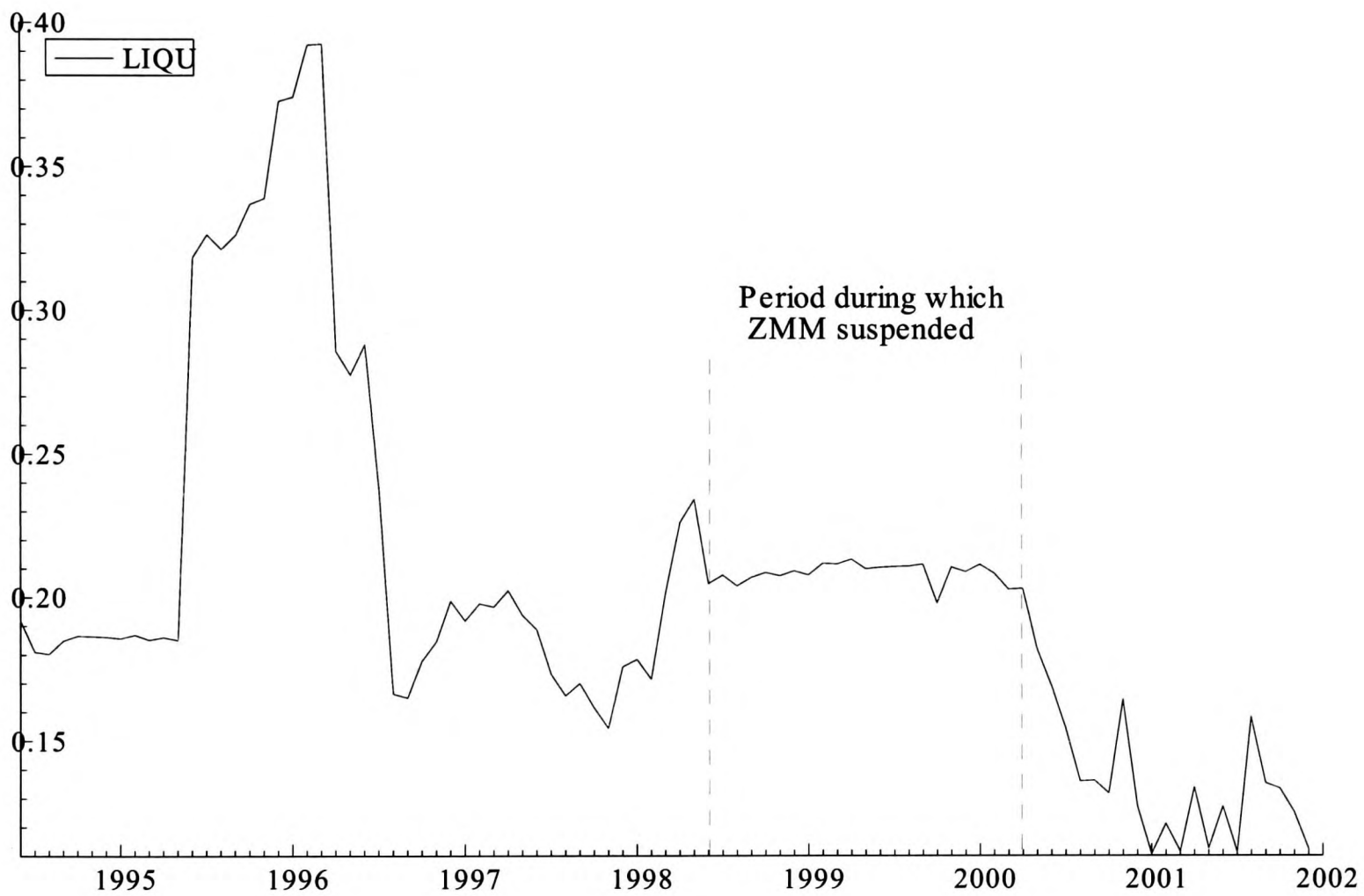
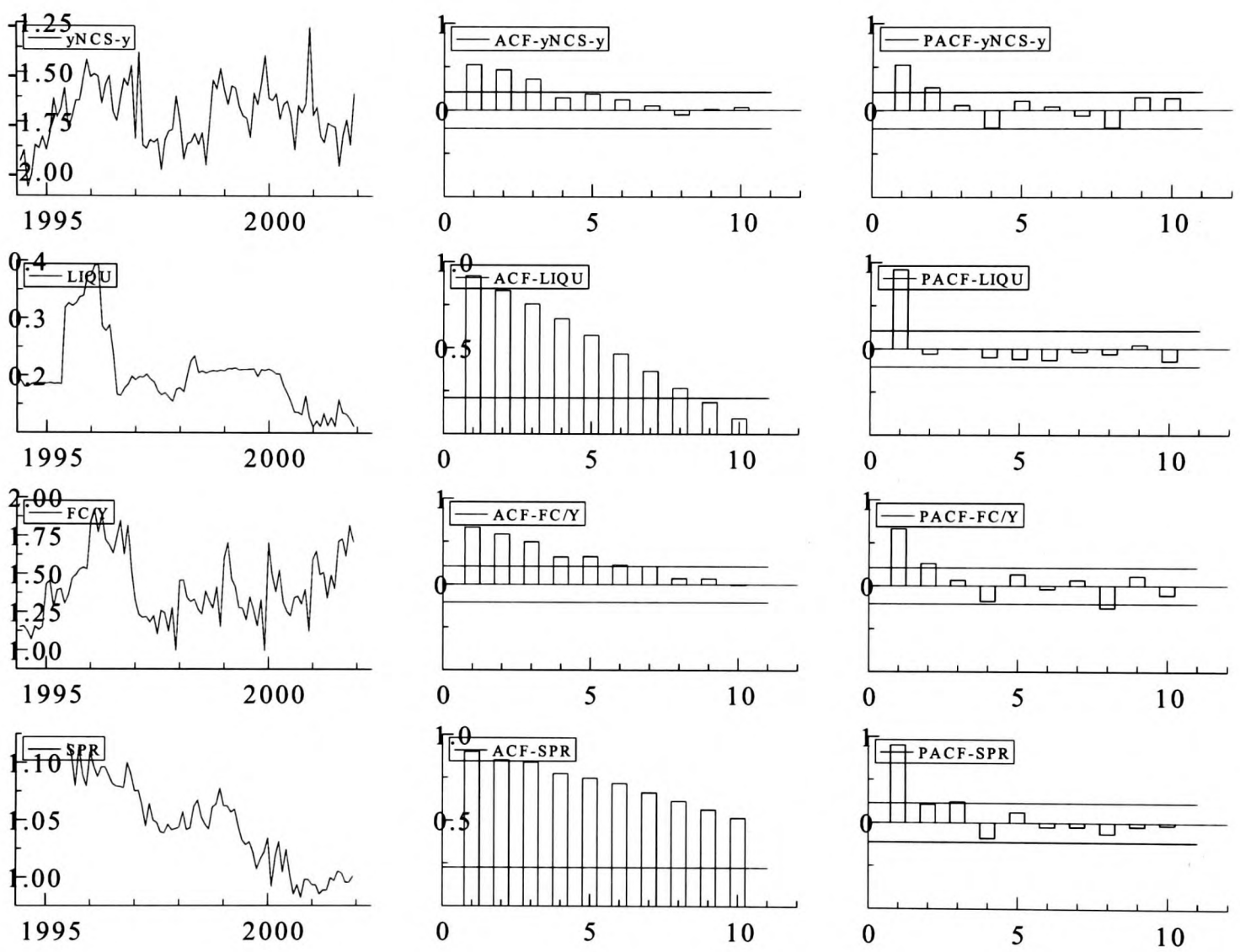


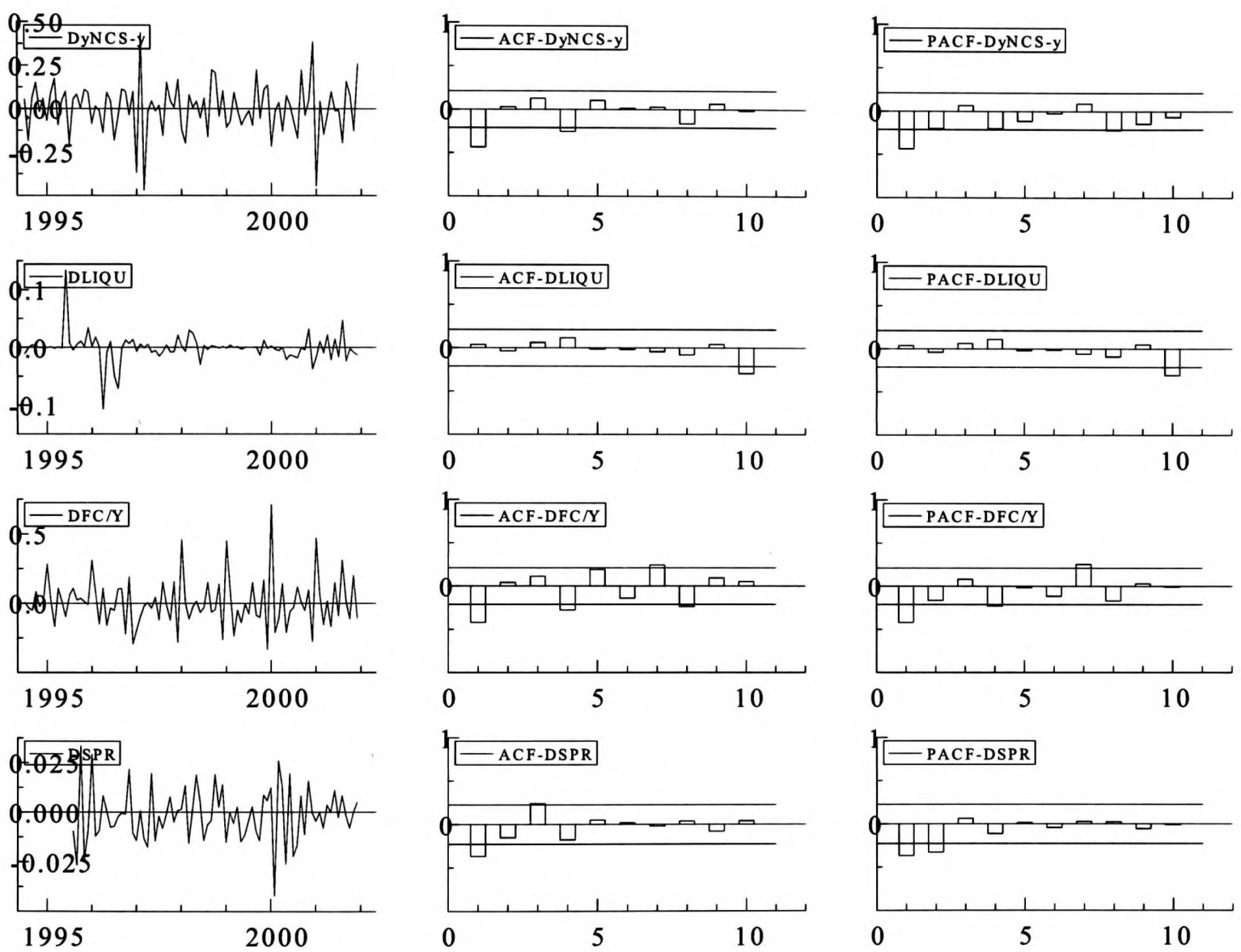
Figure 2-5: Graph of the variable *LIQU*



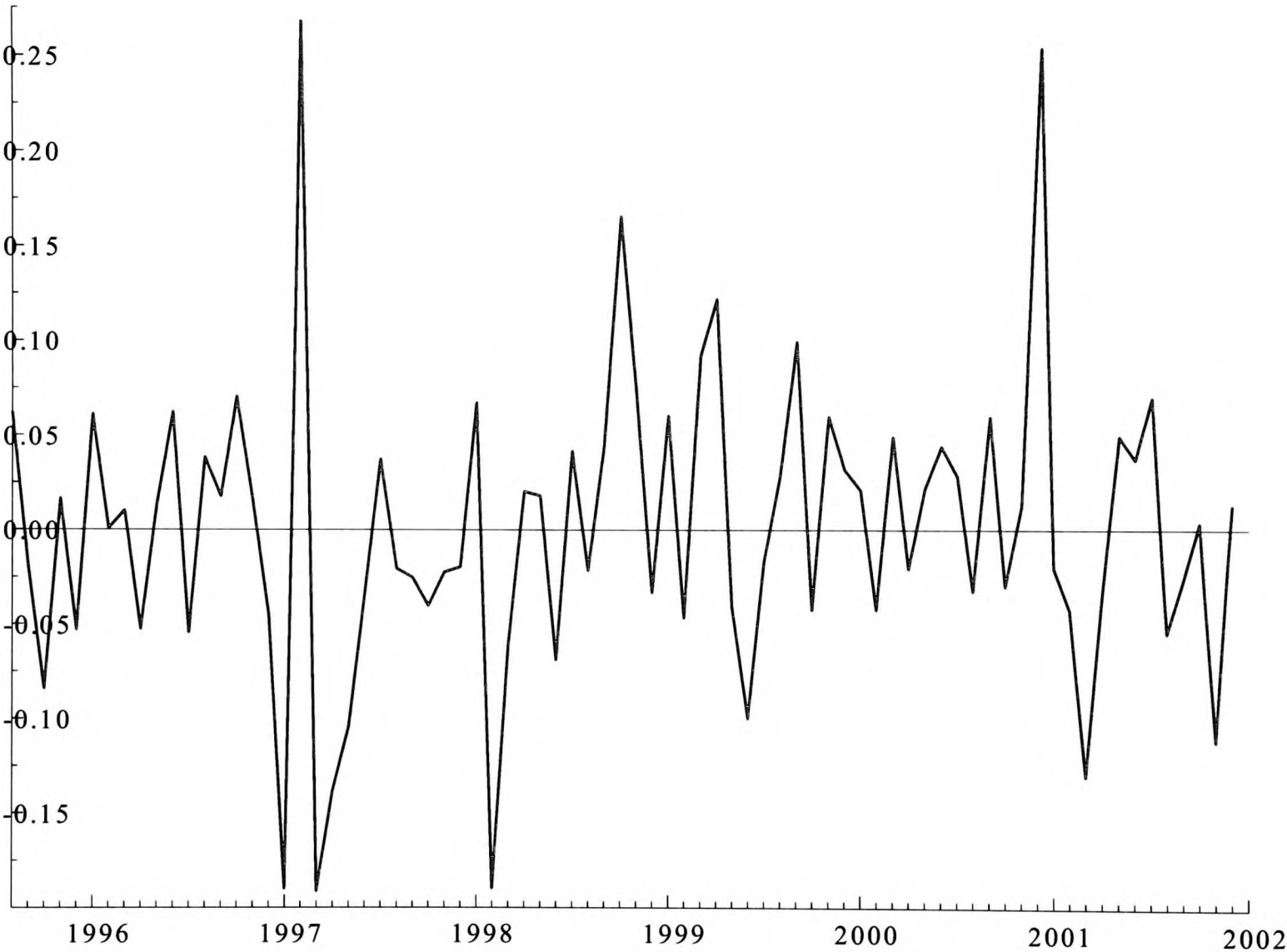
**Figure 2-6: Meltzer Equation: Time Series Characteristics of Variables in Levels**

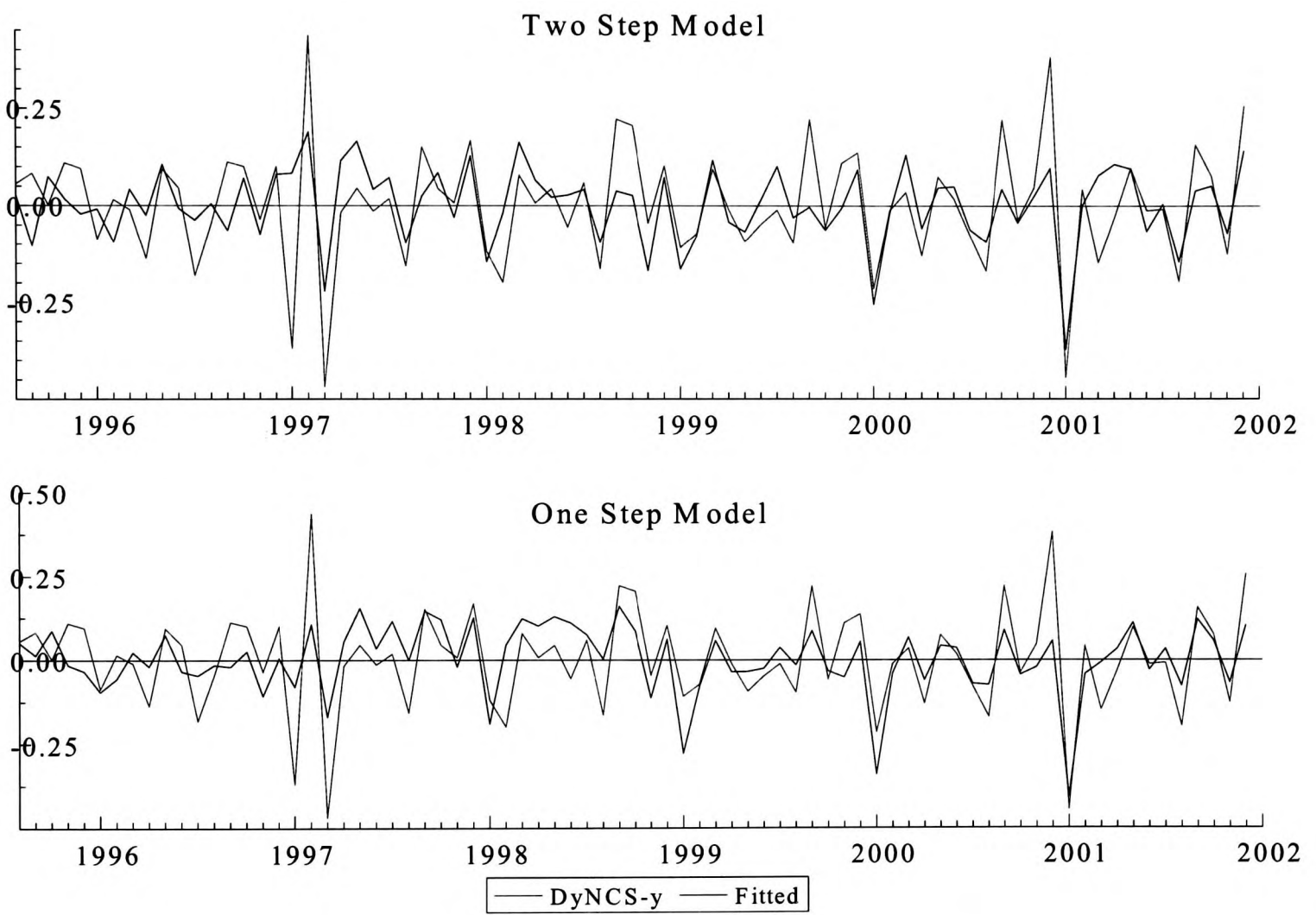


**Figure 2-7:** Meltzer Equation: Time Series Characteristics of Variables in First Differences



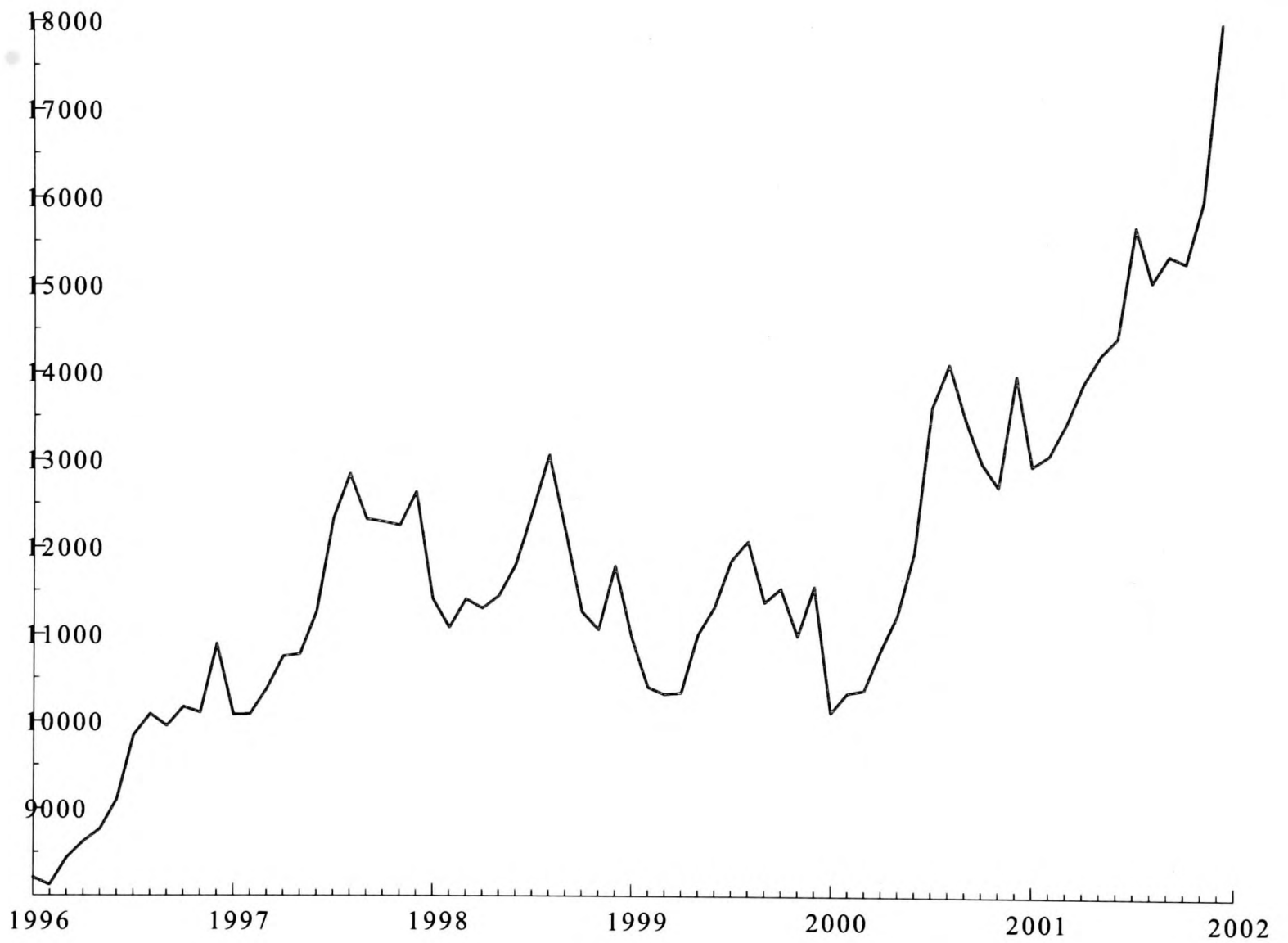
**Figure 2-8:** Meltzer Equation: Estimated Residuals of First Stage Dynamic Model



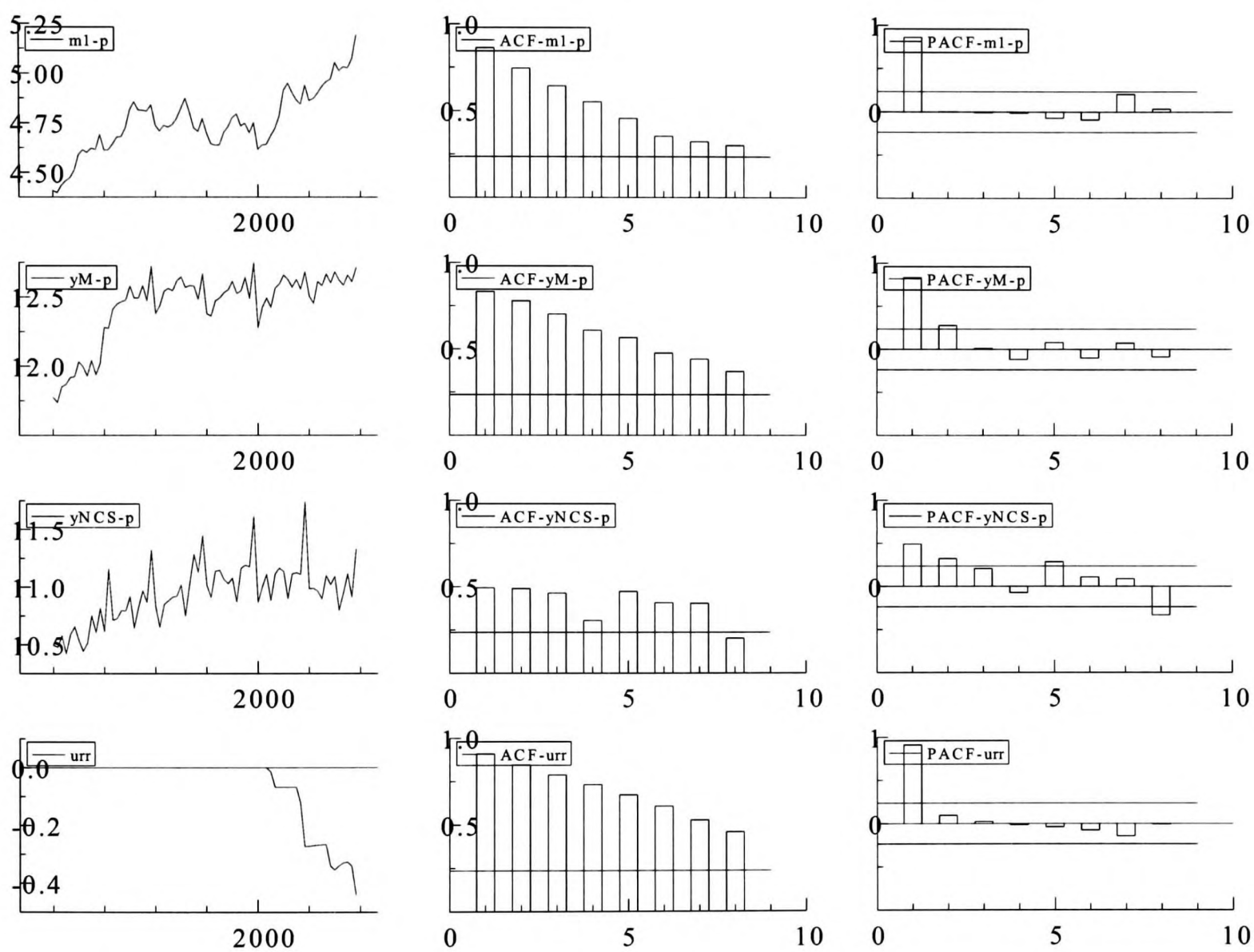


**Figure 2-9:** Meltzer Equation: Graphical Comparison of Models

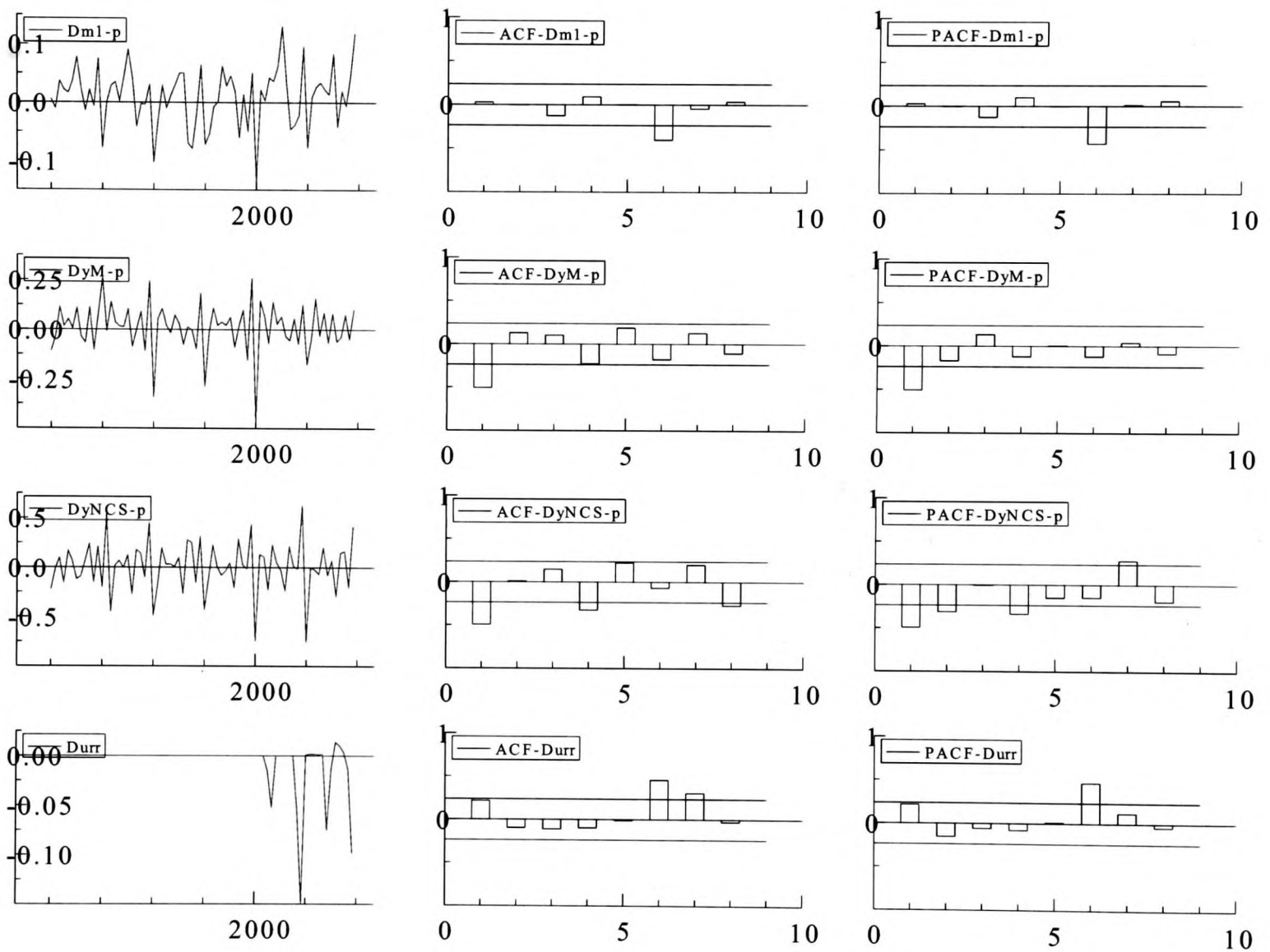
**Figure 2-10:** Evolution of Real  $M1$  (January, 1994 HRK)



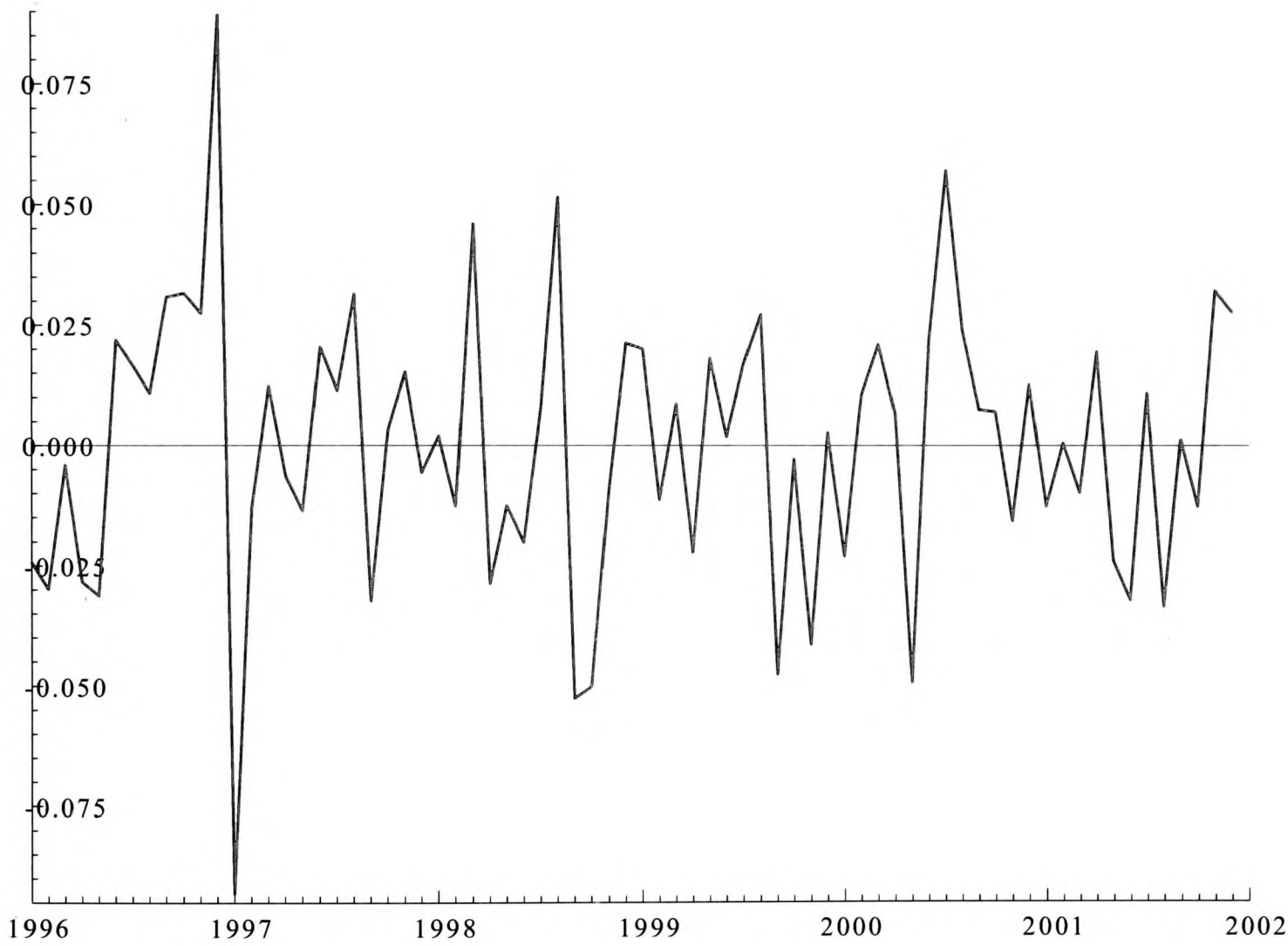
**Figure 2-11:** Transactions Demand for Money Equation: Time Series Characteristics of Variables in Levels



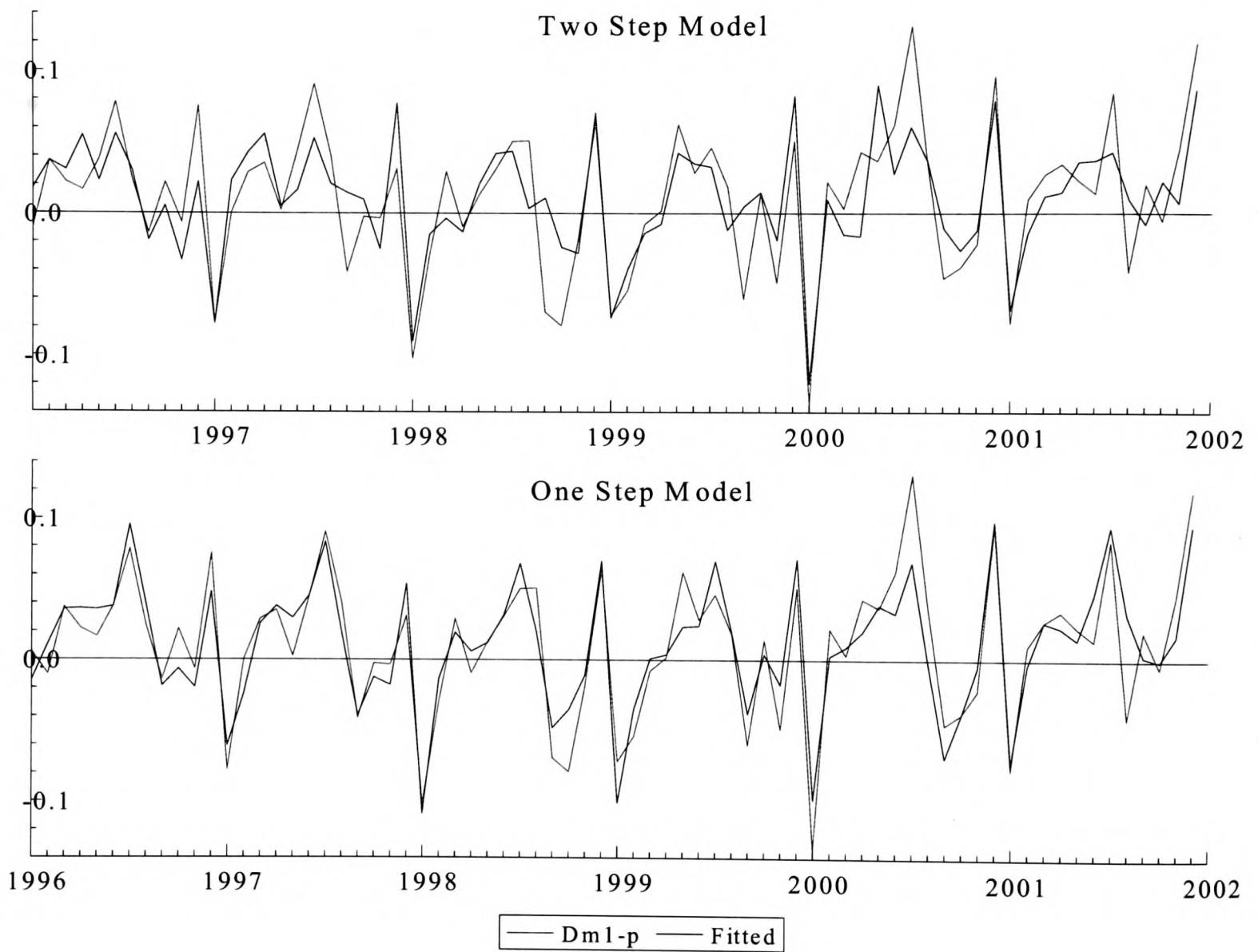
**Figure 2-12:** Transactions Demand for Money Equation: Time Series Characteristics of Variables in Levels



**Figure 2-13:** Transactions Demand for Money Equation: Residuals of First Stage Dynamic Model



**Figure 2-14:** Transactions Demand for Money Equation: Graphical Comparison of Models



	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
GDP growth, %	-14.5%	-8.7%	-12.7%	-4.1%	-3.5%	0.7%	-5.3%	6.4%	10.0%	5.1%
Central government deficit, % of GDP	-42.6%	-7.3%	-10.4%	-6.0%	-8.9%	-7.9%	-8.2%	-3.1%	3.1%	2.7%
Inflation, %, year average	1526.0%	875.0%	311.4%	197.7%	47.8%	14.7%	27.6%	86.1%	20.8%	21.6%
RUB/USD	0.42	1.25	3.55	4.64	5.56	5.96	20.65	27.00	28.16	30.14
NCS, % of total enterprise sales	6.0%	9.0%	17.0%	22.0%	35.0%	41.0%	51.0%	23.7%	-	5.7%

Source: EBRD, IMF, Russian Economic Barometer, EBRD-World Bank Business Environment and Enterprise Surveys

**Table 2.7:** Russia, main macroeconomic indicators, 1992-2001

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
GDP growth, %	5.9%	6.8%	6.0%	6.5%	2.5%	-0.9%	2.9%	3.8%	5.2%	4.3%
Central government deficit, % of GDP	1.2%	-1.4%	-1.0%	-1.9%	-1.0%	-8.2%	-7.5%	-6.7%	-5.0%	-6.3%
Inflation, %, year average	97.6%	2.0%	3.5%	3.6%	5.7%	4.2%	6.2%	4.9%	2.2%	1.8%
HRK/EUR	7.07	6.75	6.81	6.96	7.16	7.59	7.63	7.48	7.42	7.58
NCS, % of total enterprise sales	14.7%	19.0%	20.7%	15.7%	16.2%	18.3%	18.3%	15.6%	-	-

Source: Croatian National Bank, Croatian Bureau of Statistics, FINA

**Table 2.8:** Croatia, main macroeconomic indicators, 1994-2003

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Foreign currency deposits	16.1%	22.5%	32.3%	36.7%	40.7%	39.6%	42.7%	50.4%	43.5%	38.9%
Foreign currency loans	24.0%	24.2%	18.5%	16.2%	17.3%	18.5%	16.2%	15.3%	21.1%	25.6%
Households' blocked FCDs	21.0%	15.7%	10.6%	5.4%	3.7%	2.9%	1.5%	0.5%	0.2%	0.1%
HRK deposits	11.0%	11.2%	15.4%	16.4%	14.5%	14.3%	17.3%	17.8%	20.6%	21.3%
Other sources	27.9%	26.4%	23.2%	25.4%	23.9%	24.7%	22.3%	16.0%	14.5%	14.1%

Source: Croatian National Bank

**Table 2.9:** Sources of Funds of Croatian Banks, per cent of total, 1994-2003

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Foreign currency deposits	60.4%	61.6%	54.1%	44.4%	21.3%	-3.9%	27.6%	55.5%	-0.4%	5.1%
Foreign currency loans	7.6%	16.1%	-18.0%	11.5%	17.0%	5.0%	4.1%	24.6%	59.1%	42.0%
Households' blocked FCDs	-18.2%	-13.9%	-27.2%	-35.8%	-25.3%	-20.8%	-37.8%	-53.9%	-58.8%	-47.7%
HRK deposits	84.3%	17.7%	47.3%	35.9%	-3.9%	-2.8%	44.1%	35.4%	33.6%	21.5%
Other sources	17.7%	9.6%	-6.0%	39.3%	2.9%	2.2%	7.0%	-5.8%	5.0%	13.7%

Source: calculated from Croatian National Bank data

**Table 2.10:** Sources of Funds of Croatian Banks, annual real growth, 1994-2003

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
<i>% of total real domestic credit to</i>										
Government	4.0%	3.4%	1.1%	2.3%	3.4%	5.6%	7.0%	5.8%	7.0%	7.6%
Enterprises	52.8%	50.3%	44.4%	38.4%	35.6%	31.3%	28.5%	27.9%	25.7%	22.9%
Households	15.2%	16.8%	22.0%	28.2%	31.4%	35.8%	39.4%	41.3%	44.2%	48.3%
Others	28.0%	29.5%	32.5%	31.0%	29.6%	27.3%	25.2%	25.0%	23.1%	21.2%
<i>Annual growth of real domestic credit to</i>										
Government	-4.7%	5.5%	-63.9%	209.0%	81.3%	53.0%	39.7%	3.0%	60.5%	27.2%
Enterprises	24.6%	22.1%	2.2%	36.6%	17.0%	-15.6%	1.5%	23.1%	21.8%	4.7%
Households	82.0%	36.4%	38.9%	94.8%	38.3%	7.2%	21.8%	31.2%	42.0%	27.2%
Others	145.6%	75.8%	44.6%	88.5%	-31.3%	47.9%	-18.5%	101.2%	104.8%	30.2%

Source: calculated from Croatian National Bank data

**Table 2.11:** Assets of Croatian DMBs, 1994-2003

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
<b>Croatia</b>										
Credit to private sector (% of GDP)	21.2%	22.9%	21.4%	25.3%	26.6%	22.1%	27.8%	33.7%	44.0%	48.5%
Real growth of credit to private sector	-60.5%	19.0%	-0.5%	42.4%	17.1%	-10.7%	1.5%	17.9%	27.8%	13.3%
<b>Russia</b>										
Credit to private sector (% of GDP)	12.1%	8.5%	7.4%	9.4%	13.2%	10.8%	11.9%	15.5%	17.3%	20.9%
Real growth of credit to private sector	-610.0%	-230.5%	-172.8%	1.9%	49.5%	26.0%	-32.5%	31.1%	8.5%	29.0%
<b>Poland</b>										
Credit to private sector (% of GDP)	12.0%	12.7%	15.9%	18.1%	17.5%	18.7%	18.1%	18.0%	17.6%	17.8%
Real growth of credit to private sector	-4.4%	4.5%	13.3%	11.2%	10.3%	13.4%	4.9%	1.7%	2.3%	5.6%

Source: EBRD, author's calculations using IMF data

**Table 2.12:** Selected financial sector indicators for Croatia, Russia, and Poland



Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$(y^{NCS} - y)_t$	None	2	0.1157
	C	1	-3.583**
	C&T	1	-3.519*
	C&SD	2	-1.384
	C,T, & SD	2	-1.494
$LIQU_t$	None	0	2.881**
	C	0	-2.863
	C&T	0	-2.136
	C&SD	0	-1.577
	C,T, & SD	0	-1.967
$\frac{FC_t}{Y_t}$	None	1	-0.4058
	C	1	-2.311
	C&T	1	-2.192
	C&SD	0	-1.692
	C,T, & SD	0	-1.470
$SPR_t$	None	2	-1.597
	C	2	-0.9226
	C&T	2	-2.216
	C&SD	2	-0.6294
	C,T, & SD	2	-1.739

Note: sample is 1995.08 - 2001.12.

**Table 2.13:** ADF Tests of Variables of Meltzer Equation in Levels

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$\Delta(y^{NCS} - y)_t$	None	3	-6.046**
	C	1	-8.714**
	C&T	3	-5.959**
	C&SD	0	-14.31
	C,T, & SD	1	-9.385**
$\Delta LIQU_t$	None	0	-8.428**
	C	0	-8.518**
	C&T	0	-8.472**
	C&SD	0	-8.178**
	C,T, & SD	0	-8.129**
$\Delta \frac{FC_t}{Y_t}$	None	0	-13.51**
	C	0	-13.42**
	C&T	0	-13.40**
	C&SD	0	-9.113**
	C,T, & SD	0	-9.386**
$\Delta SPR_t$	None	1	-9.365**
	C	1	-9.587**
	C&T	1	-9.517**
	C&SD	1	-9.775**
	C,T, & SD	1	-9.666**

Note: sample is 1995.09 - 2001.12.

**Table 2.14:** ADF Tests of Variables of Meltzer Equation in First Differences

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$\widehat{\varepsilon}_t$	None	0	-9.261**

**Table 2.15:** Meltzer Equation - DF Test of Estimated Residuals of First Stage Dynamic Model

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$(m1 - p)_t$	None	0	1.819
	C	1	-0.8393
	C&T	0	-1.823
	C&SD	0	-0.7834
	C,T, & SD	0	-1.392
$(y^M - p)_t$	None	2	1.536
	C	2	-1.937
	C&T	1	-1.977
	C&SD	1	-2.220
	C,T, & SD	1	-1.633
$(y^{NCS} - p)_t$	None	2	0.5747
	C	2	-2.112
	C&T	2	-3.242
	C&SD	2	-2.442
	C,T, & SD	2	-1.831
$urr_t$	None	2	2.753
	C	2	2.157
	C&T	2	0.4333
	C&SD	2	1.964
	C,T, & SD	2	0.4906

Note: sample is 1996.01 - 2001.12.

**Table 2.16:** ADF Tests of Variables of Money Demand Equation in Levels.

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$\Delta(m1 - p)$	None	0	-7.552**
	C	0	-7.856**
	C&T	0	-7.806**
	C&SD	0	-7.842**
	C,T, & SD	1	-4.163**
$\Delta(y^M - p)$	None	0	-14.48**
	C	1	-8.566**
	C&T	1	-8.643**
	C&SD	0	-12.15**
	C,T, & SD	0	-12.36**
$\Delta(y^{NCS} - p)$	None	1	-10.29**
	C	3	-7.746**
	C&T	3	-7.722**
	C&SD	1	-9.815**
	C,T, & SD	3	-7.260**
$\Delta urr$	None	0	-5.284**
	C	1	-5.380**
	C&T	3	-5.704**
	C&SD	0	-5.581**
	C,T, & SD	1	-5.807**

Note: sample is 1996.01 - 2001.12

**Table 2.17:** ADF Tests of Variables of Money Demand Equation in First Differences

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$\hat{u}_t$	None	0	-9.078**

**Table 2.18:** Transactions Demand for Money Equation - DF Test of Estimated Residuals of First Stage of Dynamic Model

## **Chapter 3**

# **Dollarisation in Developing and Transition Countries: a Review of the Theoretical and Empirical Literature**



## 3.1 Introduction

In contrast to other kinds of money substitution, the use by residents of one country of financial instruments denominated in the currency of another – the phenomenon known as ‘dollarisation’<sup>1</sup> – is the subject of a rich and diverse theoretical and empirical literature, and has been an active area of research in recent years. Despite these facts, the most recent major surveys of the dollarisation literature – those of Giovannini and Turtelboom (1994) and Mizen and Pentecost (1996b) – are now nearly a decade old. The purpose of this chapter, therefore, is to present a new and coherent survey of the theoretical and empirical literature on dollarisation in developing and transition economies. To these ends, the chapter is organised as follows.

The question of how dollarisation is defined is more complicated in theory – and less trivial in practice – than one might at first think. We begin in Section 3.2, therefore, by defining our terms and setting out a clear taxonomy of the dollarisation phenomenon. We distinguish between different categories of dollarisation along several axes – according to the function of money as a store of value, means of payment, or unit of account, that is dollarised; according to the kinds of actual financial instruments, such as cash, bank deposits, and bonded debts, that are dollarised; according to whether the economy is fully or partially dollarised; according to whether the impetus for dollarisation is primarily due to supply or demand side factors; and according to the location of both the dollarised instruments themselves and their owners. Having defined dollarisation, we proceed in Section 3.3 to ask how significant a phenomenon it is. We discuss the various methods of estimating dollarisation that have been deployed in the recent literature; and survey the estimates thereby obtained. The clearest trend that emerges is the increasing importance of financial dollarisation over the past two decades.

Section 3.4 forms the centrepiece of the review. In it, we survey the theories that have been advanced in the literature to explain the determinants of dollarisation. We review the evolution of the principal models, and assess their relative merits in the light of recent empirical applications and more general empirical evidence. Three types of theory to explain transactions dollarisation – the dollarisation of money in its role as a means of payment – are surveyed: the

---

<sup>1</sup>The name ‘dollarisation’ has come to cover the use of financial instruments denominated in any currency other than the official national currency of the country in question, regardless of whether or not the US Dollar is the relevant foreign currency or not.

monetary services approach, which explains transactions dollarisation in terms of conventional money demand analysis, and two other approaches – ratchet variable methods, and models that employ network externalities - which aim to explain the persistence, rather than the origins, of transactions dollarisation. Theories aimed at explaining store of value dollarisation are more homogeneous in form, being by and large variations on the portfolio balance model of asset and liability allocation. The most important recent contributions, which together constitute what can be called the modern portfolio balance theory of dollarisation, are reviewed; and an alternative approach, aimed at explaining financial dollarisation in terms of governments' attempts to make their macroeconomic policy more credible, is also surveyed.

The final substantive part of the review, Section 3.5, proceeds to answer the question of whether there is a case for proactive economic policy in the face of dollarisation. This requires first a review of the broader macroeconomic and financial sector consequences of dollarisation that are identified in the literature; then an assessment of whether these consequences, given the theories of dollarisation surveyed in Section 3.4, justify policy intervention; and finally, a review of specific policy responses in the case that they do.

Section 3.6 concludes with a summary of the review's findings.

## **3.2 A Taxonomy of Dollarisation**

The theoretical and empirical literature on dollarisation is voluminous and diverse; and since different branches of it adopt different definitions of dollarisation, great care must be taken in interpreting and comparing theoretical and empirical results. We begin our survey, therefore, by making an analysis of dollarisation along several axes, and relating this to various definitions commonly found in the literature. The aim is to arrive at a perspicuous overview of the dollarisation phenomenon, in preparation for a review of what is known about its dimensions in the next section.

Although it is also the most abstract method of categorisation, a functional definition of dollarisation is perhaps the most useful place to start. According to the classical theory, monies serve three purposes in economic life: as stores of value, as means of payment, and as units of account. A first distinction can be made, therefore, between the use of financial instruments

denominated in foreign currency in each of these three roles:

- *Store of value dollarisation* is the use of foreign currency denominated financial instruments to transform wealth through time. Store of value dollarisation may take different forms, depending on the degree of financial sophistication of the economy in question. In financially underdeveloped economies, foreign currency cash is likely to be the main foreign currency financial instrument used as a store of value. In more developed financial systems, the use of foreign currency bank deposits and loans will be an important example of store of value dollarisation. In economies still more financially sophisticated, the holding of foreign currency pension or mutual funds, or of foreign currency government or corporate bonds may constitute a major aspect of the dollarisation phenomenon.
- *Transactions dollarisation* is the use of foreign currency denominated financial instruments as a means of payment to settle transactions. The assets most often used in this way are foreign currency cash, and foreign currency demand deposits on which cheques may be drawn.
- *Real dollarisation* is the use of a foreign currency to denominate wage contracts, other factor prices, or wholesale or retail prices. These prices need not generally (or ever) be settled by payment of foreign currency for there to be real dollarisation – it is enough that the prices themselves are denominated in a foreign currency, even if payment is in the domestic currency. Indexation – the practice of stipulating a particular mark-up of prices denominated in domestic currency over some index such as a price index, or even the exchange rate – is not, however, real dollarisation.

The critical point to be clear on is that this division of the dollarisation phenomenon into three core categories – of store of value, transactions, and real dollarisation – is made on the basis of the three functions that any money can perform. These three functions are analytically distinct and therefore convenient for theoretical purposes – and for this reason, we shall have particular recourse to them in Section 3.4 below, when we come to review theories of dollarisation. In practice, however, there do not exist separate financial instruments that correspond to each function. Instead, most financial instruments can perform more than one of them. A

demand deposit, for example, is both a store of value and a means of payment; and cash functions as both of these and a unit of account as well. These functional categories of dollarisation are not much use for empirical analysis as a result – and a different categorization, in terms of actual financial instruments, is required. Fortunately, there are standard and well-known distinctions of this sort that can serve this purpose: between cash, or narrow money, on the one hand, and all types of interest-bearing financial instrument, or broad money, on the other. A second important division of the dollarisation phenomenon, and the one that is central to its empirical analysis, is therefore into the use of foreign currency cash – *currency substitution* – and the use of foreign currency interest-bearing instruments of one sort or another – *financial dollarisation*.

There is a well-known asymmetry between cash and interest-bearing financial assets generally which points to the next useful definition. Cash, whether it be domestic or foreign currency cash, can appear on the asset side of the balance sheet of any economic agent – public or private sector, individual, firm, or other institution. It can appear on the liability side of the balance sheet of the monetary authorities (domestic or foreign, as it is domestic or foreign currency) alone, however. This asymmetry does not hold for most interest-bearing financial assets – they can be either assets or liabilities of any economic agent, and must in aggregate be the assets of some and the liabilities of others simultaneously. Another useful distinction is therefore between the dollarisation of financial assets – or *asset dollarisation*, for short – and the dollarisation of financial liabilities – or *liability dollarisation*, for short. There is an obvious (though far from inelastic) link between these two phenomena, since, for example, a firm's obtaining a foreign currency bank loan from a domestic bank – and hence contributing to liability dollarisation in the economy as a whole – implies that the bank holds a corresponding foreign currency asset – hence adding to the degree of asset dollarisation of the same economy, at least insofar as we are considering gross measures. Whether or not the levels of asset and liability dollarisation are equal in aggregate, and how they are distributed between the public and private sectors, and between households, firms, banks, the monetary authorities, and the government within these, will clearly affect the degree of overall net financial dollarisation of the economy in aggregate, and in particular sectors – an important statistic for a number of policy issues, to which we will return in Section 3.5 below.

This line of reasoning immediately suggests another pair of axes along which financial dollarisation can be analysed: those of the locations of both the financial instruments in question, and the agents that own them. A country might become more financially dollarised either because residents come to hold a higher proportion of foreign currency deposits in the domestic banking sector, or because they hold the same proportion of foreign currency deposits in the domestic banking sector, but have acquired a larger value of cross-border foreign currency deposits in a foreign banking sector. An analogous situation might apply with respect to domestic and cross-border foreign currency loans and liability dollarisation. Moreover, since the maximum geographical scope of our analysis will generally be the dollarisation of a country's balance sheet, consideration of the location of the owners of domestic financial assets – dollarised or not – may be useful. We may, for example, be interested only in the degree of financial dollarisation insofar as the assets and liabilities in question are owned by domestic residents.

There are two further dimensions in which different types of dollarisation can usefully be distinguished. These have to do not with the actual objects of dollarisation; nor with the functions that these objects perform; but with the extent of dollarisation and the types of factors driving it. The first of these further dimensions in which to distinguish different categories of dollarisation is that between full and partial dollarisation. This is a simple distinction to draw. A country is fully dollarised if its economy exhibits complete (or perhaps more loosely, overwhelming) financial and real dollarisation, and currency substitution. Anything less, and it is partially dollarised. Although some cases of full dollarisation have achieved notoriety, as a group they are in fact very few<sup>2</sup>. Partial dollarisation – where a country exhibits one or more of financial dollarisation, real dollarisation, or currency substitution without the complete displacement of its national currency – is at once a more widespread and a less well understood phenomenon. It is the literature on partial, rather than full, dollarisation that is therefore the focus of our attention in this review<sup>3</sup>.

---

<sup>2</sup>Cohen (2004) lists 13 fully dollarised countries and territories (Andorra; Northern Cyprus; East Timor; Kosovo; Lichtenstein; the Marshall Islands; Micronesia; Monaco; Montenegro; Nauru; Palau; San Marino; and the Vatican City), and 5 'near-dollarised' countries, which rely primarily on a foreign currency, but also issue a token domestic currency (Ecuador, since 2000; El Salvador, since 2001; Kiribati; Panama; and Tuvalu). To this second group should be added Cambodia, which, according to recent evidence (de Zamaróczy and Sa (2002)), is essentially fully dollarised.

<sup>3</sup>Though there are significant overlaps, there is a similarly large and fairly distinct theoretical and empirical literature on full dollarisation. A good way into this literature is via one of two recent essay collections: Levy

The distinction between full and partial dollarisation is a useful introduction to a third important division of the dollarisation phenomenon into *supply side* and *demand side* dollarisation. With rare exceptions, full dollarisation usually occurs at the instigation of the national monetary authorities. The government of the country in question decides to adopt a foreign currency for use as its national money, as Montenegro chose the Deutschmark in 1999, or Ecuador the US Dollar in 2000. Since the monetary authorities are the monopoly suppliers of the national currency, such cases of full dollarisation can be also be called ‘supply side’ dollarisation. The much more common phenomenon of partial dollarisation, on the other hand, usually occurs without – and sometimes actually against – any explicit policy of the national monetary authorities. Partial dollarisation is therefore better characterised as a choice of the country’s residents, or as ‘demand side’ dollarisation. These categories of full and partial, and supply and demand side dollarisation are by no means exactly coextensive, however. Supply side policies to promote partial dollarisation have been advocated in the past in a number of contexts. Hayek’s (1976) proposal to allow the co-circulation of currencies is a well known, if radical, example which never made it onto the policy-makers’ desks. The UK Treasury’s (1989, 1991) original proposals for the transition to the Euro by making all EMS member states’ currencies legal tender within one another’s national markets, on the other hand, is a practical example of a supply side policy to encourage partial dollarisation that did. Lately, a similar proposal – to introduce the Euro alongside national currencies in the EU accession countries during the interim period before their full membership of the single currency – has been put forward by, for example, Buitter and Grafe (2002). Of particular interest is recent empirical work that has shown that supply side factors favouring both partial and full dollarisation need not be confined to the national monetary authorities or the national government. The example of Cambodia analysed in Rumbaugh et al. (2000) and de Zamaróczy and Sa (2002) shows how a massive influx of foreign currency associated with a large scale international peace-building mission can be a major factor in a transition to either partial or full dollarisation. The cases of Bosnia and Herzegovina between 1992 and 1997, and of Afghanistan and Iraq today, suggest that this supply side channel is not unique to the Cambodian case. Nevertheless, these cases are unusual: demand side dollarisation is much the more common type, and it is therefore on

---

Yeyati and Sturzenegger (2003) and Salvatore, Dean, and Willett (2003).

this category of dollarisation that the recent literature, and this review, concentrates.

To summarise the analysis so far: the focus of our review is the literature on partial, demand side dollarisation, in all of its three functional manifestations – financial, real, and transactions. There remains one final definition that it is useful to set out, in order to make the discussion in the rest of our review as simple and comprehensive as possible; and this is the analysis of different types of financial dollarisation by the nationality and legal residency of the economic agents responsible for them. Financial assets may be held domestically or offshore, for example, and in domestic or foreign institutions. To give the example of the type of foreign currency financial asset most widely held in developing and transition countries: residents may hold foreign currency bank deposits in a domestic bank onshore; in a foreign bank onshore; in a foreign bank in a foreign country; or even in a domestic bank offshore – though this is much the least common option. A similar matrix of nationality and legal residence can be used to analyse foreign currency liabilities.

### **3.3 The Dimensions of Dollarisation**

Having organised the variety of definitions of dollarisation in the literature in Section 3.2 above, we proceed in this Section to review the evidence regarding the quantitative significance of dollarisation, in its different forms. We begin, in Section 3.3.1, by surveying the evidence on the level of financial dollarisation in developing and transition countries. It is with regard to this category of dollarisation that empirical data on the major objects are easiest to come by, and estimates are as a result the most precise. In Section 3.3.2, we turn to real dollarisation. The measurement of this type of dollarisation is more elusive, and less research has been done on it as a result. Finally, we survey recent research into the level of currency substitution in Section 3.3.3. Here again, one critical kind of data – on holdings of foreign currency – is generally missing; but in contrast to the study of real dollarisation, the last few years have seen a flourishing of different approaches to its estimation via indirect methods.

### 3.3.1 Financial dollarisation

In comparison to real dollarisation and currency substitution, it is relatively easy to measure important components of financial dollarisation, and it is as a result on the basis of measurements of financial dollarisation that most developing and transition countries are today characterised as “addicted to dollars”, as Reinhart, Rogoff, and Savastano (2003) put it<sup>4</sup>. As discussed above, one basic distinction to be drawn within financial dollarisation is between asset and liability dollarisation.

#### Asset dollarisation

On the asset side, central banks typically require domestic banks to report data on domestic and foreign currency bank deposits held in domestic banks, and since data on other important categories of financial assets – such as holdings of pension or mutual funds, or direct holdings of equities – are not so simply available, it is these that serve as a proxy for total financial dollarisation on the asset side in most recent studies of dollarisation. De Nicoló, Honohan, and Ize (2003), for example, have collected a dataset on bank deposit dollarisation, with annual observations on 100 countries over the period 1990-2001<sup>5</sup>. A summary of their data, disaggregated by region for the period 1996-2001, is given in Table 3.1.

Region	No. of countries	1996	1997	1998	1999	2000	2001
South America	8	45.8	46.1	49.4	53.2	54.0	55.9
Eastern Europe and FSU	26	37.3	38.9	43.5	44.3	46.9	47.7
Middle East	7	36.5	37.2	37.7	37.5	38.2	41.9
Africa	14	27.9	27.3	27.8	28.9	32.7	33.2
Asia	13	24.9	28.0	26.8	28.8	28.7	28.2
Central America and Mexico	7	20.6	20.8	22.0	22.1	22.5	24.7
Caribbean	10	6.3	7.6	6.8	6.7	6.1	6.2
Developed Countries	14	7.4	7.5	7.5	6.7	7.0	6.6

Source: de Nicoló, Honohan, and Ize (2003), p.6.

**Table 3.1:** Evolution of Average Foreign Currency Deposits to Total Deposits - per cent

Table 3.1 demonstrates that bank deposit dollarisation has been on the increase in most areas of the world in the last few years. The fastest rates of growth, as well as the highest

<sup>4</sup>Reinhart, Rogoff, and Savastano (2003) – this is the title of the paper.

<sup>5</sup>Data for the full period are not available for all 100 countries.

levels, of deposit dollarisation, have occurred in three regions dominated by middle income countries – South America, Eastern Europe and the FSU, and the Middle East. All regions outside the developed countries, except for the Caribbean, have seen high and rising deposit dollarisation. There are a number of examples of countries that have experienced very high levels of deposit dollarisation, with individual maxima over the period of over 90% in Bolivia, Cambodia and the Lao P.D.R.; more than 80% in Angola, Armenia, Azerbaijan, Bosnia and Herzegovina, Georgia, and Uruguay; and over 70% in the D.R.C., Croatia, Lebanon, Nicaragua, Peru, and Tajikistan. A small number of countries that have actively implemented policies to restrict deposit dollarisation have enjoyed ratios that are exceptionally low relative their respective regions – for example, Brazil and Colombia in South America, and Thailand in Asia. The number of countries in the dataset that have experienced any significant decline in deposit dollarisation over the period was limited: Egypt among the low income countries; and Hungary, Israel, Mexico, Latvia, Poland, and Slovenia in the middle income group. Overall, de Nicoló, Honohan, and Ize (2003) find that “with some exceptions, de facto dollarisation has gained ground across the board in recent years”<sup>6</sup>.

### **Liability dollarisation**

De Nicoló, Honohan, and Ize (2003) also present some data on liability dollarisation. They have collected data on foreign currency loans made by domestic banks – an important (for the private sector in most developing and transition countries, the most important) component of liability dollarisation. They find that the pattern of domestic bank loan dollarisation generally follows that of domestic bank deposit dollarisation in individual countries – but that across countries, the relationship between loan and deposit dollarisation is less than proportional, with a 10% higher deposit dollarisation ratio associated with only a 7.3% higher loan dollarisation ratio across their sample. The authors interpret this finding as corroborating the conclusion drawn by Honohan and Shi (2003), that dollarised banking systems generally operate as a channel for the export of savings, because domestic banks typically choose not to recycle the total quantity of foreign currency funds deposited with them in the form of foreign or domestic currency lending, but instead to redeposit them in liquid correspondent accounts offshore or to buy

---

<sup>6</sup>De Nicoló, Honohan, and Ize (2003), p.6.

sovereign assets denominated in the foreign currency in question<sup>7</sup>.

Domestic bank loan dollarisation is however a less reliable proxy for overall liability dollarisation in developing and transition countries than is the domestic bank foreign currency deposit dollarisation ratio for overall financial asset dollarisation. Although the latter measure does not capture at least two important categories of foreign currency financial assets relatively commonly held in developing and transition countries – foreign currency bank deposits held in banks offshore, and foreign currency cash – as well as several more minor (but possibly important in certain cases) categories such as foreign currency trade receivables, it is nevertheless not a bad measure of the dollarisation of financial assets. The dollarisation of liabilities in developing and transition countries takes place more strongly, in wider variety of classes of instruments.

Aside from loans made by domestic banks, cross-border loans denominated in foreign currency are prominent features of most developing and transition countries' aggregate balance sheets – particularly short term loans to finance international trade in the case of the private sector, and loans from international financial institutions and other official donors in the case of the public sector. Levy Yeyati (2003), for example, finds that the aggregate value of foreign currency cross-border loans (CBLs) as a proportion of GDP across a sample of 160 countries increased from 18.6% in 1995 to 25.3% in 2001; and that foreign currency Long Term Official Debt as a proportion of GDP increased from 12.7% to 15.0% over the same period. Beyond bank debt, bonds issued by domestic private sector companies, or by the public sector, may be denominated in a foreign currency. Where bonds are issued in an international financial centre in order to attract foreign investors, this is almost always the case: Levy Yeyati (2003), for example, calculates that between 1995 and 2001 foreign currency bonded debt contracted by the private and public sectors as a proportion of GDP increased from 3.4% to 10.2% and from 4.8% to 5.7% respectively; and that 72% of private sector, and 77% of public sector bonded debt was therefore contracted in foreign currency by 2001. An increasing number of developing and transition country governments, however, denominate even domestic debt issues in foreign currency: Reinhart, Rogoff and Savastano (2003) find that in 1996, only 11 countries

---

<sup>7</sup>This immediately points to an example of the potentially negative consequences of financial dollarisation, and one to which we will return in Section 3.5 below.

had issued such bonds; by 2001, this number had doubled, and the total value of outstanding foreign currency debt issued on domestic markets by developing and transition countries had grown from USD 60 to USD 250 billion<sup>8</sup>.

### 3.3.2 Real dollarisation

Real dollarisation – the degree to which retail or wholesale prices, wages, or other factor prices are denominated in foreign currency – is a rather more elusive concept than either financial dollarisation or currency substitution, since there need not be any documentary evidence of its taking place. Where it is important, however, there is in principle no obstacle to its direct estimation by extrapolation from survey data. The current standard for this type of analysis has been set by Murray and Powell (2003), who study the case of real dollarisation in Canada. The authors find that although Canada exhibits a moderately high level of financial dollarisation – over 30% of financial assets held by Canadian households being denominated in foreign currencies in 2000, for example – survey evidence suggests that real dollarisation is much less common – with only 6% of Canadian firms quoting prices in US Dollars in domestic transactions, and only 7% preparing their financial statements in US Dollars only.

Unfortunately, no such direct estimates of real dollarisation based on specially designed survey data, have been reported in the literature for developing or transition countries. There is, on the other hand, a general consensus that real dollarisation has at various times and in various countries been fairly widespread. Casual observation suggests that real dollarisation often takes off during episodes of high and hyperinflation, when the so-called ‘menu costs’ of quoting prices in the domestic currency tend to assume some importance; and that in many developing and transition countries, particular classes of transactions (for instance, the sale of cars or of real estate) frequently exhibit real dollarisation even in low inflation environments. The indexation of collective wage contracts to foreign currencies in order to protect their value in the event of a sudden depreciation of the domestic currency is a practice that has been used from time to time in a number of Latin American countries, and constitutes a kind of halfway house to real dollarisation; but once again, evidence has not been assembled with which to

---

<sup>8</sup>Reinhart, Rogoff, and Savastano (2003), p.11.

analyse this phenomenon systematically<sup>9</sup>.

### 3.3.3 Currency substitution

The essential problem for the quantification of currency substitution is the difficulty of measuring the private sector's holdings of foreign currency cash. Since the domestic monetary authorities enjoy the monopoly on issuing domestic currency cash, they are in a position to know at any given time how much of it is in circulation<sup>10</sup>. Domestic holdings of foreign currency cash, outside of holdings of the government and other public sector bodies, are much less easily monitored. Since, however, foreign currency cash is typically believed to be amongst the most important (if not the most important) component of the stock of foreign currency money in most developing and transition countries – and as such potential evidence of either substantial transactions or substantial financial dollarisation – considerable efforts have been devoted to the development of methods of estimating the quantity of foreign currency cash in circulation.

Some of these methods are indirect – that is, they calculate the unobservable volume of foreign currency cash in circulation by construing some deterministic relationship between it and other, observable macroeconomic statistics. Examples of such methods are those deployed by Feige et al. (2000), Beaugrand (2003), and De Zamorócný and Sa (2002). Intuitively more attractive, however, are methods which attempt to assemble direct estimates of the volume of foreign currency cash in circulation. One means of doing this is by the use of survey or foreign exchange office data. Mongardini and Mueller (2000) report estimates for Kyrgyzstan using such data; Feige and Dean (2002) do likewise for five Central and Eastern European countries<sup>11</sup>. The direct method most widely used in the literature is however the extrapolation

---

<sup>9</sup>The use of the speed of pass-through from nominal exchange rate movements to domestic prices as an indirect measure of the degree of real dollarisation is suggested by some recent authors (see, for example, Ize and Parrado (2002)). The extent to which the pass-through is indeed a good proxy for real dollarisation is questionable, however, and simply conflating the two concepts is not defensible. These issues are treated in more detail in our discussion of the relationship between financial dollarisation and the pass-through in Section 3.5.

<sup>10</sup>Though not, of course, how much of it is in circulation within the domestic economy. This problem is only significant for the handful of countries with currencies that serve as means of payment abroad, however – really only the US and the Eurozone countries to any substantial degree – and is probably not important for any developing or transition countries.

<sup>11</sup>Mongardini and Mueller (2000) use data from a 1998 Central Bank of Kyrgyzstan survey and data reported on a daily basis by officially licensed foreign exchange offices; Feige and Dean (2002) compare estimates made using data from Austrian National Bank surveys with those made using US Treasury Department survey and interview data.

of estimates of US Dollar cash in circulation from US Customs Service data. Since 1970, US domiciled individuals and institutions importing or exporting US Dollar cash in value greater than a certain threshold (under the terms of the original legislation, US\$ 5,000; since 1980, US\$ 10,000) have been legally obliged to register the details of these transactions with the US Customs Service by submitting a Currency and Monetary Instrument Report (CMIR). These CMIRs can be analysed to generate aggregate data on the origins, destinations, and value of a significant share of the flows of US Dollar cash into and out of the US. Since public sector institutions, banks, firms, and individuals all register their cross-border transfers of US Dollar cash using CMIRs, it is only a relatively small proportion of cash imports and exports – essentially only illegal flows, and legal flows composed of individual transfers below the US\$ 10,000 threshold – that are not captured by such aggregate analysis.

A number of recent studies have employed CMIR data in order to estimate the volume of US Dollar cash circulating in certain developing and transition countries. The essence of the method is the calculation of net inflows of US Dollar cash to the country in question from the US on the basis of CMIRs, and then the summation of these over time. On the assumption that flows originating in the US are neither significantly defrayed by subsequent transshipment to third countries, nor significantly supplemented in an analogous manner, this yields an estimate of the cumulative net flow of US Dollar cash into or out of the country in question. All that is then required to arrive at an estimate of the total quantity of US Dollar cash in circulation at any given point in time is an estimate of the initial stock of US Dollars in circulation in the year from which CMIR records are available. Since for many countries, this base year is in the 1970s or early 1980s, and it can safely be assumed that the stock of US Dollars in circulation in these periods must by recent standards have been relatively small, the accuracy of this estimate of the initial stock is not likely to be of great importance.

This type of approach was first deployed by Kamin and Ericsson (1993, 2003) to estimate US Dollar cash in circulation in Argentina on a quarterly basis over the period 1988 to 1993. Based on an indirect estimate of the initial stock of US Dollars in circulation at the end of 1987 of US\$ 5 billion, the authors concluded that US Dollar cash was “the single most important monetary aggregate in Argentina at the end of 1992”, totalling approximately US\$

26 billion, or 11% of Argentina's GDP in that year<sup>12</sup>. Feige et al. (2000) applied the same methodology to estimate the value of foreign currency cash per capita in circulation in 28 developing and transition countries in 1997. The results of their exercise are reported in Table 3.2 below, which presents for each country in their sample the value of an overall dollarisation index (ODI) calculated as the ratio of estimated foreign currency cash in circulation, plus foreign currency bank deposits, to broad money. As can be seen by comparing the values of this index with the values of a conventional asset side financial dollarisation index (FDI), calculated as the ratio of foreign currency deposits to broad money, the estimated value of foreign currency cash in circulation over Feige et al. (2000)'s sample of developing and transition countries is on average equal to 9% of broad money. In four countries (Argentina, Latvia, Russia, and Yemen), it is equal to 20% or more.

An additional interesting point that can be inferred from Feige et al.'s (2000) estimates relates to the relationship between financial dollarisation and currency substitution. The Spearman rank correlation coefficient between the values of the overall dollarisation index, ODI, and the asset side financial dollarisation index, FDI, in Table 3.2, is 0.91; and the simple correlation coefficient between the values of the two indices is 0.87. These values suggest that overall dollarisation and asset side financial dollarisation are typically closely related in developing and transition countries. The simple correlation coefficient between the financial dollarisation index, FDI, and the figures in the final column of Table 5 – equal to the estimates of foreign currency cash in circulation as a proportion of broad money, and hence a kind of currency substitution index – is much lower, at 0.57. This implies that asset side financial dollarisation is not necessarily a good predictor of currency substitution.

The most recent example of the CMIR method applied to a middle income country are Oomes' (2003) estimates of US Dollar cash in circulation in Russia. The author constructs CMIR-based estimates – assuming, like Feige et al. (2000) a zero stock of US Dollars in circulation in the period prior to estimation – and adjusts them using balance of payments statistics in order to control for holdings of US Dollars used to settle transactions abroad<sup>13</sup>. These estimates suggest very high levels of transactions and asset side financial dollarisation

---

<sup>12</sup>Kamin and Ericsson (1993) p.5.

<sup>13</sup>For example, by individuals on holiday, or for the purchase of football players – as is well known in the UK, such holdings are apparently non-negligible in Russia.

Country	ODI	Rank by ODI	FDI	Rank by FDI	ODI-FDI
Angola	6.3	26	1.3	25	5
Argentina	68.5	4	48.9	4	19.6
Armenia	36.1	12	33.6	8	2.5
Bolivia	83.5	1	78.8	1	4.7
Bulgaria	49.8	9	42.2	6	7.6
Costa Rica	40.8	11	35.5	7	5.3
Estonia	18.6	21	16	21	2.6
Hungary	26.7	15	26.6	13	0.1
Indonesia	20.9	18	20.7	16	0.2
Israel	18	23	17.9	19	0.1
Jordan	19.3	20	16.4	20	2.9
Kuwait	16.9	22	15.3	22	1.6
Kyrgyzstan	18	24	15.4	23	2.6
Latvia	56.8	5	31.1	9	25.7
Mozambique	45.5	10	29.9	11	15.6
Nicaragua	77.3	2	62	2	15.3
Peru	54	6	53.9	3	0.1
Poland	19.5	19	15.2	24	4.3
Romania	28.9	14	28.5	12	0.4
Russia	77.2	3	30.4	10	46.8
Saudi Arabia	36	13	17.9	18	18.1
Tanzania	50	17	46.4	17	3.6
Turkey	23.3	8	18	5	5.3
Ukraine	26.3	16	21.1	15	5.2
Venezuela	9.2	25	0.2	26	9
Yemen	50.7	7	25.8	14	24.9

Source: Feige et al. (2000), p.14.

**Table 3.2:** Estimates of dollarisation with and without estimated foreign currency cash (per cent of M2 unless otherwise indicated)

in Russia over the 1990s, driven largely by exceptionally large holdings of US Dollar cash – amounting to approximately 40% of the broad money stock between 1995 and July, 1998, and rising to nearly 70% of M2 after the financial crisis in August, 1998.

### 3.4 Theories of dollarisation

In this section we review the recent theoretical literature on the determinants of dollarisation. This literature is extensive; and as a result, we have not set out to provide a comprehensive overview, let alone a synthesis, of the whole of it. The reader is referred to Giovannini and

Turtelboom (1994) and Mizen and Pentecost (1996b) for detailed surveys of the early literature, and in particular for more thorough treatments of the theoretical literature on transactions dollarisation. In this Section, we give only a relatively brief overview of these early theories, based on these two previous surveys. The last decade has, however, seen a surge of interest amongst theorists in the determinants of dollarisation related to the growth of the empirical importance of financial dollarisation described in Section 3.3 above; so that there is now a substantial body of theory that was not covered by the two surveys of the mid-1990s. The bulk of this section is therefore devoted to an overview of this more recent theoretical literature.

Before turning to specific theories, we first review a number of structural factors that are generally acknowledged to be important determinants of the level and dynamics of both transactions and store of value dollarisation, but which are often excluded from formal modelling. Openness to trade and labour flows, the presence of real dollarisation, the influence of the payments system for transactions dollarisation, and the structural and regulatory characteristics of the financial sector are all discussed in Section 3.4.1.

In Section 3.4.2, we proceed to survey modelling strategies designed to explain transactions dollarisation. The first, and undoubtedly the most influential, of these is the monetary services approach, which explains transactions dollarisation in terms of a representative agent's motivation to hold different monies as a result of their relative effectiveness in reducing the frictional costs of conducting market transactions. This is essentially a generalisation to a multicurrency context of the conventional method of analysing the domestic transactions demand for money; and as such has the benefit of being intuitive and familiar. At the same time, it is ill-suited to explaining an important empirical regularity observed of dollarisation – the tendency of shifts in the demand for foreign currency to persist after the factors thought to generate them have subsided. A separate branch of transactions dollarisation theory has emerged with the aim of explaining this phenomenon of 'dollarisation hysteresis'; and models of this sort are also therefore reviewed.

Due to the increasing prominence of financial dollarisation over the past two decades, the focus of the most recent theoretical literature has shifted towards the explanation of store of value dollarisation.; and models in this category are reviewed in Section 3.4.3. The core of models of this type is the theory of portfolio balance, whereby agents' holdings of both monies

and interest bearing assets are explained by reference to financial motives – that is, by the relative capacity of each to transform agents’ wealth over time. We begin with a brief overview of the original unrestricted portfolio balance, and its application to the analysis of store of value dollarisation. In the next subsection, we consider the Thomas (1985) model, which forms an important bridge between this earlier approach to modelling store of value dollarisation and the modern portfolio balance approach, which dominates the most recent theoretical literature. We then present the model that best represents this approach: that of Ize and Levy Yeyati (2003). Ize and Levy Yeyati’s (2003) model is an important theoretical contribution that has been influential in much of the recent theoretical and empirical work on financial dollarisation. It has also proved to be highly flexible model, which a number of other recent authors have successfully employed as the basis of their own theoretical investigations of further aspects of the determinants of store of value dollarisation. In the final subsection, we review a selection of these other models in the modern portfolio balance approach.

Finally, in Section 3.4.3, we introduce a macroeconomic approach to explaining store of value dollarisation that is not based on the theory of portfolio balance and aimed primarily at explaining the dollarisation of public sector financial liabilities in terms of a government’s attempt to make its macroeconomic policy more credible. Though much less elaborated in the recent literature than the modern portfolio balance models reviewed in Section 3.4.3, this is an important alternative approach given the implications of dollarisation for the independence of monetary and exchange rate policy.

In order to organise the discussion of what are rather involved theoretical issues, we adopt a uniform structure wherever it is necessary to discuss the details of the models. In those sections in which this is the case, we begin by describing the empirical motivation for each class of models; then provide a simplified exposition; and conclude with an assessment of the theory of dollarisation provided, both in terms of the theory alone, and against the results of empirical applications of it in the recent literature.

### **3.4.1 Structural factors influencing dollarisation**

Underpinning the models of individual, bank, and government choice choices discussed in Sections 3.4.2 and 3.4.3 below are a variety of structural features of developing and transition

countries that promote or hinder dollarisation. In this subsection, we discuss a number of these that the recent literature has emphasised.

### **Trade openness**

The degree to which a country engages in foreign trade is an obvious candidate as a structural determinant of dollarisation. Foreign trade transactions require foreign currency with which they may be settled – a factor leading to transactions dollarisation. Moreover, trade finance and precautionary balances will tend to encourage a wider range of foreign currency financial instruments to supplement the cash and demand deposits demanded by importing and exporting firms or individuals for settlement purposes – a source of store of value dollarisation. A good case study example is given by Cowen et al. (2002), who argue that Vietnam's strong export performance over 1999 to 2001 was a major factor behind rising financial dollarisation over the same period, and cite as evidence of this the fact that 40% of Vietnam's foreign currency deposits at the end of this period were held in the country's main bank for international trade. The tourism industry is one export sector that is a particularly likely channel for the import of foreign currency direct to domestic households. Mazzaferro et al. (2003) show how tourism has contributed to financial dollarisation in a sample of middle income countries; Kraft (2003) discusses the case of Croatia.

### **Immigration and emigration**

A substantial level of emigration of domestic residents to foreign countries from where they then send remittances to their families and friends, or eventually themselves return with foreign currency incomes or savings, is a second structural feature that can contribute to dollarisation. Given the importance of remittances in many Latin American and Eastern European countries, and the likelihood that emigration to North America and the EU will increase in the coming decades, it is also a feature likely to become more and more prominent in many middle income countries. Amongst recent studies, Fritz-Krockow et al. (2001) adduces evidence of changing expenditure, savings, and denomination patterns in Haiti associated with the return of a professional class from the US between 1994 and 1997 to argue that the immigration of recent emigrés had an important role in initiating and then sustaining dollarisation. Another, though more un-

usual, way in which the literature notes that movements of people can be a structural stimulus in dollarisation is afforded by countries where large international peace-building missions have been introduced for a number of years. Both Rumbaugh et al. (2000) and de Zamaróczy and Sa (2001) present evidence for this supply side impetus in the case of Cambodia; Afghanistan, and most of the countries of the former Yugoslavia are also cases in point<sup>14</sup>.

### **Real dollarisation**

It would seem to be intuitive that extensive real dollarisation may induce or sustain dollarisation – and especially transactions dollarisation. Explanations in such terms would certainly lend some much-needed structure to the modelling of dollarisation hysteresis using ratchet variables, discussed below in Section 3.4.2. One attempt to do just this was made by Fritz-Krockow et al. (2001) in the case of Haiti's lingering dollarisation after 1994 to 1995<sup>15</sup>. However, as discussed in Section 3.3.2 above, data on real dollarisation is so limited that it is difficult to draw any robust conclusions on this link, or even to say anything categorical concerning the strong possibility of the simultaneity of its relationship with transactions and financial dollarisation.

### **The payments system**

For transactions dollarisation that generates not only currency substitution but also the dollarisation of checkable deposits within banks, the characteristics of the payments system constitute another set of structural factors that may act as an important influence on the level and dynamics of dollarisation. This influence may be either simply permissive – if clearing in foreign currency exists on identical terms to clearing in domestic currency – or active – if the tax or other efficiency conditions are different between the two. Baliño, Bennett, and Borzenstein (1999) give examples of both official (typically established, supervised, and regulated by the domestic monetary authorities) and unofficial (typically established by consortia of private banks) foreign currency payments systems in a variety of developing and transition countries, and discuss their

---

<sup>14</sup>Bosnia and Herzegovina is perhaps the exception that proves the rule: despite having the largest military and civilian peace-building presence of all the former Yugoslav republics, Bosnia and Herzegovina is one of the few transition economies in which the level of financial dollarisation has fallen substantially and continuously over the past few years.

<sup>15</sup>See Fritz-Krockow et al. (2001), p.10f.

role in facilitating dollarisation. Haiti and Cambodia are again interesting low income countries cases. The creation of a clearing houses for US Dollar transactions in 1995 in the former is reported by Fritz-Krockow et al. (2001) to have played a role in promoting dollarisation; whereas de Zamaróczy and Sa (2001) report that the National Bank of Cambodia's establishment of a similar institution in 2001 came only after much deliberation, and its judgment that the new US Dollar clearing house would improve payments stability without promoting dollarisation. Oomes' (2003) model of transactions dollarisation driven by network externalities, intended as a stylized representation of the determinants of dollarisation in Russia, gives a prominent role to the differential turnover tax charged on US Dollar transactions, as well as to the enforcement of regulations banning such transactions<sup>16</sup>.

### **Structural and regulatory characteristics of the financial sector**

Finally, with regard to financial dollarisation, the structural and regulatory environment of the financial sector is clearly the stage on which any degree of dollarisation emerges. Savastano (1996) presents a deservedly well-known analysis of how such structural aspects of the financial sector institutional framework determined the level and dynamics of financial dollarisation in the late 1980s and early 1990s in Latin America. The author shows how a distinction could be drawn in these cases between countries with strict foreign exchange and capital controls, and those where residents had been allowed to maintain foreign currency deposits in domestic banks. In the former groups of countries, macroeconomic instability generally led to capital flight in the form of the conversion of domestic banking sector liabilities into cross-border deposits and foreign currency cash, putting pressure on monetary authorities' international reserves (in the case of a fixed exchange rate) and the overall level of financial intermediation. In the latter group, on the other hand, macroeconomic instability sometimes led to increased intermediation as residents attempted to protect monetary wealth against inflation by diversifying into foreign currency or indexed instruments within the domestic financial system. Since a number of structural and regulatory issues will emerge from the presentation of the formal theories of store of value dollarisation in Section 3.4.3 below, however, and since many of the policy options to manipulate the level of dollarisation discussed in Section 3.5 will also have to do with this

---

<sup>16</sup>Oomes' (2003) model is surveyed below in Section 3.4.2.

category of structural determinants, we will postpone further analysis of their role until then.

### 3.4.2 Transactions dollarisation

#### The monetary services approach

**Motivation** The development of models of transactions dollarisation has followed closely the evolution of models of money demand more generally. Three principal approaches to the modelling of money demand dominate the post-war theoretical literature. The first is the ‘money in the utility function’ (MIU) approach due originally to Sidrauski (1967), in which real money balances simply appear as an argument of agents’ utility functions in the same way that consumption does. Since this approach essentially assumes away the motivation behind money demand, let alone specifying any structural relationship between economic exchange and money demand, it has not been widely deployed by theorists seeking to model transactions dollarisation – though Imrohoroğlu (1994) and Selcuk (1997) do provide recent examples of models developed in this tradition. A second method, stemming from Tobin (1958) and Samuelson (1958) is to model money simply as a zero-interest financial asset amongst a number of interest bearing financial assets, with all assets demanded on the same terms for their capacity to transform wealth over time. This is the theory of money demand underlying the portfolio balance models of dollarisation, reviewed in Section 3.4.3 below. The third approach to modelling money demand involves the introduction of transactions costs associated with consumption, and to assume that money is demanded because it provides liquidity services, which reduces these costs. We consider models of transactions dollarisation that are based on this so-called ‘monetary services’ approach to the modelling of money demand in this section.

**Exposition** Two main variants of the monetary services approach have been employed in the modelling of transactions dollarisation. The first variant, in which money is treated as an input into a transactions costs production function, is due originally to Baumol (1952) and Tobin (1956). In the case where there are just two monies, the transactions costs productions function may be written as:

$$T = Ch \left( \frac{M}{C}, \frac{M^*}{C} \right) \quad h \geq 0, h_1 \leq 0, h_2 \leq 0, h_{11} > 0, h_{22} > 0, h_{12} > 0 \quad (3.1)$$

where  $C$  is consumption of a single homogenous good, and  $M$  and  $M^*$  are domestic and foreign currency denominated nominal money balances respectively. According to Equation 3.1, holding higher real balances of either domestic or foreign monies reduces transactions costs incurred by a given level of consumption, though at a decreasing rate. The sign of the cross partial derivative  $h_{12}$  indicates that the two monies are substitutes, rather than complements, in the production of liquidity services. Agents maximise utility, which is defined over the single consumption good  $C$ , by choosing the level of  $C$  and the level of their holdings of the two monies so as to reduce transactions costs to the minimum possible. The degree of currency substitutability is determined by the precise form and parameters of the function  $h$ , in accordance with conventional production theory.

In the second variant of the monetary services approach, the notion of a transactions cost production function is replaced by the imposition of exogenous constraints on the holding of real money balances in order to facilitate consumption of different goods. In this type of so-called ‘cash in advance’ (CIA) model, due originally to Clower (1967), agents maximise utility defined over a number of goods, subject both to the usual budget constraint, and to additional CIA constraints regarding the necessity of holding of given minimum real money balances in order to purchase consumption. For example, in the case where the representative agent’s utility function is defined over two consumption goods,  $C$  and  $C^*$ , which the CIA constraints dictate may only be purchased with domestic and foreign currency monies respectively, the additional constraints may be written as:

$$\frac{M}{P} \geq C, \frac{M^*}{P^*} \geq C^* \quad (3.2)$$

where  $P$  and  $P^*$  are the prices of  $C$  and  $C^*$  respectively.

The main difference between these two types of model consists of the mechanism whereby domestic and foreign money are substituted for one another. In the first type of model, the substitutability of the two moneys is determined endogenously according to the interaction of the agent’s preferences and the form of the transactions cost production function described by Equation 3.1 above. The substitutability of consumption goods in the agent’s utility function is not an issue, since there is but a single, homogenous consumption good. In the class of CIA models, however, the substitutability of the two moneys in the purchasing of the two goods is determined by the specification of the CIA – which is exogenous to the model – but there

is an additional, indirect channel of substitutability represented by the substitutability of the two consumption goods in the agent's utility function. As a result, the degree of transactions dollarisation that results depends jointly on the both. If only foreign money may be used to buy foreign consumption, and vice versa, then the degree of currency substitutability will simply be equal to the coefficient of substitutability between the different goods in the agent's utility function. If, however, this restriction is relaxed, so that foreign money may be used to purchase either foreign or domestic consumption, and vice versa, then the currencies become infinitely substitutable.

**Assessment** In both transactions cost production function and CIA models, the channel of transactions dollarisation is a real one – foreign money is acquired in order to buy consumption, rather than for financial (speculative or hedging) motives. Conceived of as tools of international macroeconomic analysis, these classes of models are therefore most appropriate to small, open economies with underdeveloped financial systems, where the dominant channel for the accumulation of foreign currency money is a surplus on current account. There are two main methods of their empirical application. The first, of which Miles (1978) is the *locus classicus*, is to specify a transactions cost production function, and to derive from it an equation for estimation of the following form:

$$\frac{M_t}{S_t M_t^*} = \rho \lambda + \rho \left( \frac{i_t^*}{i_t} \right) + \varepsilon_t \quad (3.3)$$

which defines the ratio of domestic to foreign currency in circulation, measured in domestic currency ( $\frac{M_t}{S_t M_t^*}$ , where  $S_t$  is the period  $t$  nominal exchange rate) as a function of the elasticity of substitution of domestic and foreign currency in demand ( $\rho$ ); the parameters of the transactions cost production function  $t$  (denoted  $\lambda$ ); the respective nominal interest rates on interest bearing assets (IBAs) denominated in domestic and foreign currency ( $i_t, i_t^*$ ), and a stochastic error term ( $\varepsilon_t$ ). One advantage of such an equation is that it enables the direct estimation of the parameter  $\rho$  – the elasticity of currency substitution; a disadvantage is that it requires data on both domestic and foreign currency cash in circulation, the difficulty of assembling which we reviewed in Section 3.3.3 above. Heimonian (2000) is a rare example of a recent empirical study in this tradition, where the author estimates the elasticity of substitution between Euro and US Dollar money in Estonia over the period 1997 to 2000. Predictably, however, the study

is forced to employ a ratio of broad money, rather than cash, aggregates, and (perhaps as a result – it is not possible to tell) does not find any significant transactions motives at work in the substitution of the two foreign currencies for one another by Estonian residents.

An alternative, and more common, approach to the empirical application of monetary services models is to derive what amounts to a conventional domestic money demand equation augmented by including a nominal rate of return on foreign currency denominated IBAs amongst the variables describing the opportunity cost of holding domestic currency cash. This nominal rate of return on foreign currency denominated IBAs is typically captured by either a foreign currency interest rate, the foreign currency inflation rate, or the rate of nominal depreciation of the domestic against the foreign currency. In terms of its data requirements, this approach is much less demanding than the Miles (1978) method, since the dependent variable of the equation for estimation measures a domestic currency monetary aggregate only. It does not, however, enable direct estimation of the elasticity of currency substitution in demand denoted  $\rho$  above. Instead, the estimated coefficient on the nominal rate of return on foreign currency denominated IBAs is interpreted as a measure of the substitutability of the two currencies in question. As a variant on the empirical analysis of domestic money demand that has been popular since the 1950s, this approach has been employed by numerous studies. Recent applications have been somewhat less frequent, however; and what empirical results they have generated are equivocal. Carruth and Sanchez-Fung (2001) and Bahmani-Oskooee and Techaratenachai (2001) find that the domestic currency – US Dollar exchange rate has been a significant factor for domestic money demand in the Dominican Republic and Thailand respectively, for example; whereas Cuthbertson and Bredin (2001) find that the expected depreciation of the Koruna against the Deutschmark was not a significant determinant of Czech money demand between 1992 and 1997.

### **Models of dollarisation hysteresis**

**Motivation** The monetary services approach, then, is intuitive, and has even scored some empirical successes in explaining currency substitution. There is, however, one important respect in which it is unable adequately to characterise the actual experience of numerous dollarised developing and transition economies. It has been very widely observed that the process of

currency substitution often exhibits hysteresis – that is, that once an economy has arrived at a significant level of currency substitution, this level tends to be sustained even though the factors which originally induced it, such as a high level of domestic inflation or of nominal depreciation against the substitute foreign currency, may have subsided<sup>17</sup>. Standard monetary services models of the sort reviewed in the previous subsection are symmetrical – the actions of their determinants on the relative demands for foreign and domestic currency are assumed to be identical whether transactions dollarisation is increasing or decreasing – and are, as a result, ill-equipped to explain the dynamics of hysteresis. Alternative, or augmented, approaches are needed in order to explain hysteresis in transactions dollarisation.

In recent empirical studies, the usual practice in order to account for hysteresis in transactions dollarisation is simply to keep faith with one of the two monetary services approaches above as the underlying model of the process generating the choice of currencies, whilst augmenting the equation for estimation with some kind of ‘ratchet’ variable selected on an ad hoc basis. So, for example, Kamin and Ericsson (1993) construct a variable that takes the value in each period of the highest inflation in the sample level up to that period and find this to be significant in explaining the evolution of currency substitution in Argentina within a monetary services framework. Mueller (1994) and Mongardini and Mueller (1999), on the other hand, choose a variable that takes the value in each period of the highest level of the dollarisation ratio itself in the sample level up to that period<sup>18</sup>. The successes of empirical studies that employ a ratchet variable do not by themselves, however, contribute much to our understanding of currency substitution hysteresis. As Oomes (2003) points out, “it is tautological to interpret the finding of a ratchet effect (i.e., a significant ratchet variable) as an explanation for hysteresis”<sup>19</sup>. Inclusion of an *ad hoc* ratchet variable may help an empirical model adequately to characterise the data in statistical terms; but adequate explanation of the hysteresis phenomenon requires

---

<sup>17</sup>Hysteresis has also been detected in many cases of financial dollarisation; but it is studies that have employed estimates of foreign currency cash in circulation and hence measure currency substitution that hysteresis has been best documented: see, for example, Kamin and Ericsson (1993) on Argentina, Muller (1994) on Lebanon, Peiers and Wrase (1997) on Bolivia, Mongardini and Mueller (1999) on the Kyrgyz Republic, and Oomes (2003) on Russia.

<sup>18</sup>Strictly speaking, of these two studies only Mongardini and Mueller (1999) falls into the category of applications of the monetary services approach augmented with a ratchet variable; since the empirical models in Muller (1994) are of financial dollarisation ratios.

<sup>19</sup>Oomes (2003), p.9.

the extension of the underlying theoretical model so that such equations for estimation, perhaps incorporating ratchet variables, can be derived from them.

**Exposition** The chief theoretical innovation in this direction in the recent literature is the introduction of the reduced transactions costs associated with an increased size of the network of users of a given currency as arguments in agents' choices. The idea that such 'network externalities' are important in the choice of fiat money in preference to any other transactions media is at the heart of much recent monetary theory<sup>20</sup>. Dowd and Greenaway (1993), however, were the first to bring some formal structure to the discussion of how they might be important in explaining the level of transactions dollarisation; and it was left to Uribe (1997) to demonstrate specifically that they could provide an explanation for dollarisation hysteresis when incorporated into a two currency CIA model. Uribe (1997) explains hysteresis in transactions dollarisation as the result of agents getting used to transacting in a foreign currency – a process which is subject to both initial inertia, and subsequent momentum. The author models this asymmetric process of getting used to transacting in foreign currency by recourse to the notion of the 'dollarisation capital' of an economy – a theoretical artefact which reflects the extent of the network of users of foreign currency money as a means of payment in the domestic economy. The level of the stock of dollarisation capital is then included as an argument of a transactions costs production function as part of a model built according to the traditional monetary services approach, so as to reflect the fact that the greater is the number of agents in the economy settling transactions using foreign currency money, the easier it is for any individual agent to do so. By postulating a particular law of motion for the stock of dollarisation capital, Uribe (1997) then finds that the CIA model generates multiple steady state solutions for the level of transactions dollarisation: hysteresis can then be explained as the transition from a lower to a higher steady state level.

Whilst the empirical literature discussed above delivers clear evidence of hysteresis but little theoretical structure, Dowd and Greenaway (1993) and Uribe (1997) provide interesting theories of how network externalities may cause dollarisation hysteresis, but ones from which empirically testable hypotheses cannot easily be derived. A recent study that has attempted to combine substantial theoretical and empirical substance is Oomes (2003). Oomes (2003) presents a

---

<sup>20</sup>It is the central feature of Kiyotaki and Wright's (1989) model, for example.

model incorporating network externalities, as well as a variety of other typical transactions costs associated with the use of foreign currencies in a dollarised economy – such as the cost associated with possible confiscation of foreign currency, should its use for the settlement of domestic payments be illegal; and the cost associated with purchasing foreign currency transactions if it is not. The model can be solved for a (nonlinear) law of motion governing the evolution of the ratio of foreign to domestic currency means of payment in circulation – the transactions dollarisation ratio – and this linearised by a logistic transformation to give:

$$\ln \left( \frac{1 - d_t}{d_t} \right) = -\frac{1}{\varphi} (s_t^e - \sigma_t + (2\sigma_t + \tau_t - q_t) d_{t-1}) \quad (3.4)$$

where  $d_t$  is the transactions dollarisation ratio;  $\varphi$  is a coefficient capturing the aversion of agents to the use of foreign currency for ethical or other exogenous reasons;  $s_t^e$  is the expected rate of depreciation of the domestic against the foreign currency;  $\sigma_t$  measures the switching cost associated with transacting in a combination of domestic and foreign currency, and having to find counterparties who will accept one or the other, or intermediaries who will exchange them;  $\tau_t$  is the foreign exchange tax incurred in purchasing foreign currency; and  $q_t$  is the probability that foreign currency will be confiscated because it is illegally used, or that it is found to be counterfeit. An important feature of this solution (prior to linearization) is that it converges on a steady state solution that is not monotonically increasing in  $s_t^e$  – so that there may exist multiple steady state transactions dollarisation ratios, allowing for a characterisation of hysteresis as the shift from a lower to a higher equilibrium as in the Uribe (1997) model discussed above.

Two assumptions are then made concerning the functional forms of the processes generating the network externality  $\sigma_t$  and the expected rate of depreciation of the domestic currency respectively. First, it is assumed that the network externality  $\sigma_t$  depends in a linear fashion on the previous period's transactions dollarisation ratio:

$$\sigma_t = 1 - \gamma d_{t-1} \quad (3.5)$$

Second, it is assumed that agents form their expectation concerning the rate of depreciation of the domestic currency in such a way that it is equal to the actual rate of depreciation with

probability  $\alpha$ , and equal to the maximum rate recorded in the recent past with probability  $(1 - \alpha)$ :

$$s_t^e = \alpha s_t + (1 - \alpha) s_t^{\max} \quad s_t^{\max} = \max \{s_t, \dots, s_{t-n}\} \quad (3.6)$$

By substituting Equations 3.5 and 3.6 into Equation 3.4, Oomes (2003) is able to derive an equation for estimation which includes both network externalities, in the form of a nonlinear influence of the previous period's transactions dollarisation ratio, and a ratchet effect, in the form of the highest rate of depreciation of the domestic currency in the sample up to time  $t$ , as follows:

$$\ln \left( \frac{1 - d_t}{d_t} \right) = \beta_0 + \beta_1 s_t + \beta_2 s_t^{\max} + \beta_3 \tau_t d_{t-1} + \beta_4 d_{t-1} + \beta_5 d_{t-1}^2 + \varepsilon_t \quad (3.7)$$

where the coefficients to be estimated have the following interpretations:

$$\beta_0 = \frac{1}{\varphi}, \beta_1 = -\frac{1}{\varphi} \alpha, \beta_2 = -\frac{1}{\varphi} (1 - \alpha), \beta_3 = -\frac{1}{\varphi}, \beta_4 = -\frac{1}{\varphi} (2 + \gamma - q), \beta_5 = \frac{1}{\varphi} (2\gamma) \quad (3.8)$$

and appropriate identifying restrictions can be imposed accordingly.

**Assessment** The model presented by Oomes (2003) is undoubtedly an advance on previous contributions on the monetary services approach. It supplements the modelling of the real motives for the holding of different monies by CIA constraints and the transactions services production function with the incorporation of switching costs, network externalities, and ratchet effects in a formal framework. It is a moot point as to how deep is the theoretical structure that it provides for the ratchet effect incorporated – it is really no more than an artefact of the specification of agents' expectations formation mechanism, and as such in the final analysis not so very different from the ad hoc variables deployed by earlier empirical studies. Nevertheless, Oomes's (2003) own empirical results from fitting the model to Russian data suggest that it is a useful theory which allows more detailed analysis of the hysteresis phenomenon, and of transactions dollarisation more generally than did these previous studies. Using estimates of foreign currency cash held by Russian residents, Oomes (2003) finds first, that network externalities were an important determinant of the level of currency substitution in Russia between 1992 and 1998, and second, that a characterisation of transactions dollarisation hysteresis as a phase transition from a low-level to a high-level equilibrium is supported by the data.

### 3.4.3 Store of value dollarisation

#### Early portfolio balance models

**Motivation** One major shortcoming of the monetary services approach remains, however: in confining their explanation of the demand for different monies to their roles as a means of payment, monetary services models avoid the additional complexity implied by money being simultaneously a store of value and unit of account. By the same token, however, they lose explanatory power when applied to empirical questions, and might lead on their own to erroneous conclusions. Since no analytical distinction is made between the transactions and store of value components of demand, all variation will be explained as changes in transactions demand, when changes in the relative demand for monies as stores of value may in fact be to blame. What is missing from the monetary service models is the idea that money might be held for financial (hedging and speculative) motives as well as real (transactions) ones.

The principal alternative approach to modelling dollarisation, therefore, is a version of conventional portfolio balance theory, which is designed specifically to model these financial motives. The original, unrestricted portfolio balance model was due in its closed economy form to Tobin (1958), and in its open economy variant, allowing for foreign as well as domestic currency denominated money and IBAs, to McKinnon and Oates (1966). Its use for the analysis of dollarisation was pioneered by Cuddington (1983), whose model thereafter served as the theoretical workhorse for the bulk of the empirical literature in this tradition. We begin, in this section, with a brief survey of this theory, which is the ancestor of all that follows. Thomas' important (1985) model forms the basis for most of the models developed in the last decade, and indeed represents the beginning of the modern portfolio balance approach to dollarisation: we review it next. The *locus classicus* of the modern portfolio balance approach itself, however, is the model of Ize and Levy Yeyati (2003): most of the important elements of this, the most popular approach to modelling the determinants of store of value dollarisation, are evident in the Ize and Levy Yeyati (2003) model, and it is reviewed in the final subsection.

**Exposition** In the unrestricted portfolio balance model, the representative agent's holding of foreign currency denominated money is modelled as being chosen simultaneously with his holdings of foreign currency denominated IBAs and domestic currency denominated money

and IBAs. This choice of overall portfolio is made on purely financial grounds – the risks and expected returns associated with each class of assets – in order to maximise the agent’s expected utility. The motivation to hold any quantity of money for transactions purposes is not included: instead, liquidity preference is modelled as behaviour towards risk.

Taking first the closed economy case, domestic money ( $M$ ) is a riskless asset with zero expected return, whereas domestic bonds ( $B$ ) generate a positive expected return, but with an associated degree of risk. If the representative agent is risk-averse, then domestic bonds will not absolutely dominate domestic money in the portfolio; though for any given degree of risk, an increase in the interest rate on domestic bonds will induce him to shift the composition of his portfolio towards bonds and away from money. The demand for money, in other words, is explained by agents’ aversion to risk, rather than by their need to hold money balances in order to buy consumption. The model can be extended to the open economy case by introducing foreign money ( $M^*$ ) and foreign bonds ( $B^*$ ). Immediately, we are forced to make an assumption about how our agent values income in each of the two currencies. In the monetary services models, this would have been decided by reference to the optimal allocation of consumption between goods that require foreign currency for their purchase, and those that require domestic currency; and to the (exogenously specified) substitutability of the two currencies in transactions services. In the unrestricted portfolio balance model, however, the analysis stops at the accumulation of purchasing power *per se*. The implicit assumption is made, therefore, that all demanded consumption is available for domestic money. The return to foreign bonds is therefore valued at the associated interest rate plus the expected rate of depreciation of the domestic currency. The return to foreign money is by the same reasoning equal to the rate of depreciation of the domestic currency.

An optimal demand function can therefore be derived for each of the four assets available to the representative agent. The arguments of these demand functions are identical for each asset, with their partial derivatives with respect to each argument representing respective demand

elasticities, as follows<sup>21</sup>:

$$\frac{M}{P} = m(y, r, r^* + s, s) \quad m_1 > 0, m_2 < 0, m_3 < 0, m_4 < 0 \quad (3.9)$$

$$\frac{B}{P} = b(y, r, r^* + s, s) \quad b_1 < 0, b_2 > 0, b_3 < 0, b_4 > 0 \quad (3.10)$$

$$\frac{SM^*}{P^*} = n(y, r, r^* + s, s) \quad n_1 > 0, n_2 < 0, n_3 < 0, n_4 > 0 \quad (3.11)$$

$$\frac{SB^*}{P^*} = f(y, r, r^* + s, s) \quad f_1 < 0, f_2 < 0, f_3 > 0, f_4 > 0 \quad (3.12)$$

where  $P$  and  $P^*$  are the domestic and foreign price levels,  $S$  is the nominal exchange rate;  $y$  is the level of real income,  $r$  and  $r^*$  are the real interest rates on domestic and foreign bonds, and  $s$  is the rate of depreciation of the domestic currency. Currency substitutability in this model is measured by the coefficient  $m_4$ : an expected increase in the rate of depreciation of domestic money is assumed to reduce its risk adjusted expected return relative to that of foreign money, and hence to induce a shift in portfolio composition towards foreign money.

**Assessment** As discussed above, the unrestricted portfolio balance model has generally been superseded as the theoretical basis for empirical studies of dollarisation. Two recent studies have however bucked this trend. Komárek and Melecký (2003) estimate equations derived from a Cuddington (1983) type model for the Czech Republic over the period 1994 to 2001; Rodríguez and Turner (2003) make a similar analysis of dollarisation in Mexico over the period 1978 to 2000. Nevertheless, the principal objection to the unrestricted portfolio balance approach is on the level of theory, rather than of empirical application. Just as the monetary services approach reviewed in Section 3.4.2 above is deficient because it is exclusively focused on real determinants of dollarisation – those related to transactions dollarisation – so the unrestricted portfolio balance model is deficient because it explains currency substitution and financial dollarisation purely in terms of store of value dollarisation factors. There is no room in the model for variation in the ability of different monies to procure desired consumption to affect the demand for them. This class of determinants of currency substitution and financial dollarisation related to transactions dollarisation – which makes up the whole subject of the monetary services models – is omitted from the unrestricted portfolio balance model.

---

<sup>21</sup>This formulation is borrowed from Mizen and Pentecost (1996b), pp.27-8

### Restricted portfolio balance models: the Thomas (1985) model.

**Motivation** The Thomas (1985) model seeks to marry the monetary services and portfolio balance approaches to modelling dollarisation, and in so doing to incorporate denomination substitutability both in transactions and in store of value demand. Like the unrestricted portfolio balance model, we begin from a single, representative agent and four assets – domestic ( $M$ ) and foreign ( $M^*$ ) money, and domestic ( $B$ ) and foreign ( $B^*$ ) bonds. The agent's choice problem is to apportion his wealth between consumption ( $C$ ) and investment, and within investment between investment in each of these four classes of assets.

**Exposition** The real determinants of money holding are captured by specifying that consumption involves transactions costs. These transactions costs may be reduced by holding real money balances (domestic or foreign); and the level of transactions costs incurred may therefore be summarised by a transactions costs production function identical to that in Equation 3.1 above:

$$T = Ch \left( \frac{M}{C}, \frac{M^*}{C} \right) \quad h \geq 0, h_1 \leq 0, h_2 \leq 0, h_{11} > 0, h_{22} > 0, h_{12} > 0 \quad (3.13)$$

Increasing real balances of domestic or foreign monies reduces transactions costs incurred by a given level of consumption, though at a decreasing rate; and domestic and foreign monies are imperfect substitutes in this regard.

The investment part of the agent's problem is modelled as in the unrestricted portfolio balance model. The expected real return to each asset depends upon its respective nominal interest rate (equal to zero for both monies), and the levels and variances of inflation and the depreciation of the domestic currency. The agent is assumed to have a von Neumann – Morgenstern utility function that is strictly increasing and concave in its sole argument, consumption ( $C$ ). The proportions of his portfolio held in each of the four assets,  $M$ ,  $M^*$ ,  $B$ , and  $B^*$ , are denoted  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ , and  $\theta_4$ ; together, these proportions sum to unity. His dynamic problem is then to maximise lifetime expected utility by choosing his rate of real consumption,  $C$ , and the allocation of his portfolio across the four assets over time. Assuming, finally, that complete sets of domestic and foreign nominal bond markets are available to the agent in which he may buy and sell at zero cost, the results of Merton's (1971) continuous time portfolio allocation model can be used to derive the following first order conditions that will be satisfied

by his optimal choice:

$$i = -h_1 \quad (3.14)$$

$$i^* = -h_2 \quad (3.15)$$

where  $i$  and  $i^*$  are the nominal interest rates on domestic and foreign currency denominated bonds; and

$$\lambda = \Phi - \zeta\Psi(E[r] - E[r^*]) \quad (3.16)$$

where

$$\Phi = \left( \frac{S_{\pi\pi} + S_{\pi s}}{S_{\pi\pi} + S_{ss} + 2S_{\pi s}} \right)$$

$$\Psi = \left( \frac{1}{S_{\pi\pi} + S_{ss} + 2S_{\pi s}} \right);$$

$\lambda = (\theta_2 + \theta_4)$  is the proportion of the agent's net asset portfolio denominated in foreign currency;  $S_{ss}$  is the variance of the rate of depreciation of the domestic currency;  $S_{\pi\pi}$  is the variance of the rate of inflation of the domestic currency;  $S_{\pi s}$  is the covariance of the depreciation and inflation rates of the domestic currency;  $\zeta > 0$  is the inverse of the agent's degree of risk aversion;  $E$  is the expectations operator; and  $r$  and  $r^*$  the real returns to holding domestic and foreign currency denominated bonds.

According to Equations 3.14 and 3.15, real transactions motives will induce the agent to choose each money until the liquidity services yielded by the marginal unit exactly equal the opportunity cost of holding it, represented by the nominal interest rate. Meanwhile, financial motives will determine the agent's overall net foreign currency denominated asset demand, as described by Equation 3.16. These financial motives are of two sorts. The first term,  $\Phi$ , in Equation 3.16 describes hedging demand, which aims to absolutely minimize the portfolio's purchasing power risk. The second represents demand due to speculative motives, and describes the proportion of the portfolio the agent will desire to hold in foreign currency denominated money and bonds in order to maximize the portfolio's overall expected return, given the spread between real returns on domestic and foreign currency denominated bonds, their variances, and the agent's degree of aversion to risk.

In effect, the Thomas (1985) model has the agent construct his portfolio in two stages. In the first stage, purely real factors determine how much domestic and foreign money the agent will choose to hold – his transactions demand for domestic and foreign monies, as it were. In the second stage, however, the agent costlessly borrows and lends in the bond markets in order to attain an optimal portfolio allocation across asset classes (money and bonds) and currencies of denomination (domestic and foreign) determined by purely financial factors – embedded in which are his asset demand for domestic and foreign monies. The key feature of the first order conditions for his optimal choice is that the decisions in these two stages are separable:  $h_1$  and  $h_2$  do not appear in Equation 3.16. The complete set of nominal bond markets available to the agent allow him completely to diversify the risks associated with money holdings demanded for transactions purposes, and the ultimate currency composition of the portfolio is therefore determined exclusively on financial grounds. This two stage construction allows the model to distinguish between what Thomas designates currency substitutability ‘proper’ – that is the degree of domestic and foreign money’s substitutability in the transactions costs production function in Equation 3.13 – and the substitution of domestic and foreign money for one another (and for domestic and foreign currency denominated bonds) in response to risk and return considerations implicit in arriving at the optimal overall currency composition of the agent’s portfolio described by Equation 3.16.

**Assessment** Despite its elegant modelling of both transactions and store of value motives for holding different currencies, the Thomas (1985) model too suffers from a major drawback in practice, in that its central assumption of complete sets of domestic and foreign currency denominated nominal bond markets available to all agents is unlikely to be plausible in most empirical cases. If this is so, and agents cannot as a result costlessly diversify the risks associated with money holdings, then the separability of their transactions and asset demands for foreign currency denominated assets will no longer hold, and the substitutability of domestic and foreign money in the transactions cost production function will influence the currency composition of the equilibrium portfolio of all four assets. But in any event, the theoretical merits or shortcomings of the Thomas (1985) model’s attempt to combine the modelling of transactions and store of value motives for holding moneys denominated in different currencies were

overtaken by rise to prominence of financial dollarisation in precedence to currency substitution described in Section 3.3 above, and by the consequent shift in the policy issues associated with dollarisation. As described there, the major monetary and exchange rate policy issue associated with dollarisation is no longer so much the instability of money demand and the nominal exchange rate due to transactions dollarisation, but the potential of financial dollarisation – assumed to be predominantly due to store of value motives – to undermine the credibility and range of monetary authorities’ policy options, to induce fragility in the financial sector, and to deteriorate public finances.

As a result, the latest models in the portfolio balance tradition place little emphasis on the first stage of Thomas’ (1985) two stage portfolio balance model (the choice of cash holdings); and instead develop the second stage (the choice of financial instruments for the purposes of hedging and speculation) in order to account for the more politically important and empirically tractable phenomenon of store of value dollarisation on both sides of agents’ balance sheets. In addition, the contemporary importance of on- and offshore allocation decisions is reflected in their incorporation into this latest generation of models.

### **The modern portfolio balance approach: the Ize and Levy Yeyati (2003) model.**

**Motivation** The essence of this modern portfolio balance approach to modelling the determinants of dollarisation is best exemplified by the model of Ize and Levy Yeyati (2003), which has formed the basis of the majority of the recent theoretical work on the determinants of dollarisation, and also underpinned much recent empirical research. As described above, the purpose of models on this approach is primarily to explain financial dollarisation, rather than currency substitution – hence the modelling of transactions motives for holding domestic or foreign currency cash or IBAs are omitted, in favour of a more detailed treatment of the interaction between agents and financial intermediaries, allocation decisions, and other choices of interest. In this subsection, we survey the Ize and Levy Yeyati (2003) model in some detail, since an understanding of it conveys the essential theoretical results of the modern portfolio balance approach. In the following subsection, we then proceed to discuss a number of other recent models, also formulated on the modern portfolio balance approach, which effectively modify the Ize and Levy Yeyati (2003) to analyse further policy issues.

**Exposition** The basic structure of the Ize and Levy Yeyati (2003) model is of two representative agents, one a depositor facing an asset portfolio allocation problem, and the other a borrower facing a liability portfolio allocation problem. The equilibrium dollarisation ratio is then defined by a financial equilibrium which satisfies both agents' optimisation problems. Exposition can begin with the depositor, whose problem is not essentially different from that that faced by the representative agent in the Thomas (1985) model, except that money holdings are excluded, so that the 'first stage' of the Thomas (1985) model's problem is dispensed with. Three classes of assets are available to the depositor – domestic currency denominated deposits, and foreign currency denominated deposits held in either domestic or foreign banks. In addition to the risks associated with inflation and depreciation incorporated in the Thomas (1985) model, there is an additional element of country risk associated with deposits in either currency that are held onshore; so that real returns to the three different assets can be differentiated as follows:

$$\begin{aligned} r &= E(r) - \varepsilon_\pi + \varepsilon_c \\ r^* &= E(r^*) + \varepsilon_s + \varepsilon_c \\ r^{**} &= E(r^{**}) + \varepsilon_s \end{aligned}$$

where  $r^{**}$  is the real rate of return on foreign currency denominated deposits held off-shore, and  $\varepsilon_z$  is a stochastic disturbance associated with inflation ( $z = \pi$ ), depreciation ( $z = s$ ), and country ( $z = c$ ), risk respectively. All three risks are assumed to be independent of one another.

Since money holdings are not considered, there is no need to include the transactions cost minimisation aspect of Thomas' (1985) model. In other respects, however, including the depositor's utility function, the set-up is identical, and yields the equivalent solution for the optimal share of foreign currency denominated assets (here, foreign currency denominated deposits, held either on- or offshore) in the overall net asset portfolio. Denoting the proportions of the depositor's net asset portfolio held as domestic currency deposits, foreign currency deposits held onshore, and foreign currency deposits held offshore, as  $\gamma_1$ ,  $\gamma_2$ , and  $\gamma_3$  respectively, this solution is given by:

$$\mu_D = \Phi - \zeta_D \Psi(E[r] - E[r^*]) \quad (3.17)$$

where  $\mu_D = (\gamma_2 + \gamma_3)$ ;  $\zeta_D > 0$  is the inverse of the degree of the depositor's risk aversion; and all other variables are as in Equation 3.16 above. A solution can also be derived for  $\gamma_3$  alone – the proportion of the net depositor's net asset portfolio held in offshore foreign currency denominated deposits:

$$\gamma_3 = 1 - \zeta_D \Xi (E[r^*] - E[r^{**}]) \quad (3.18)$$

where

$$\Xi = \frac{1}{S_{cc}}$$

and  $S_{cc}$  is the variance of country risk.

$\Phi$  in Equation 3.17 is therefore once again the proportion of the depositor's portfolio held in foreign currency that absolutely minimises the riskiness of its real return, and Ize and Levy Yeyati (2003) therefore denote it the minimum variance portfolio (MVP) dollarisation ratio. As is intuitive, this ratio increases with inflation volatility, and decreases with the volatility of the real exchange rate – since these measure the purchasing power risks associated with domestic and foreign currency respectively. So long as  $\mu_D > \gamma_3$  (so long, in other words, as country risk conditions are not such that all foreign currency denominated deposits are held offshore), and controlling for the depositor's risk aversion and the relevant real return spread, the depositor's choice of the currency composition and location of his portfolio is determined only by the second moments of the relevant variables: these being inflation and real depreciation in the case of currency composition, and country risk with respect to location.

The borrower's problem is effectively the inverse of the depositor's. The borrower is assumed to have access to onshore loans denominated in domestic or foreign currency (unlike the depositor, he does not have direct access to offshore finance). He is assumed to use his loan to finance a project with a known return in units of the domestic price index, so that the real return to the project is riskless, and the borrower's problem reduces to that of minimising his risk-adjusted cost of borrowing. Assuming that the borrower shares a similar utility function with the depositor, a solution for his optimal share of foreign currency denominated liabilities,  $\mu_L$ , may be derived. This has the same form as the solution to the depositor's problem, but with the relevant interest rate spread (now between onshore foreign and domestic currency loans,

rather than deposits) entering with the opposite sign, as follows:

$$\mu_L = \Phi + \zeta_L \Psi(E[r_L] - E[r_L^*]) \quad (3.19)$$

where  $\zeta_L > 0$  is the inverse of the borrower's risk aversion; and  $r_L$  and  $r_L^*$  are the domestic and foreign currency denominated loan interest rates.

Financial equilibrium is identified by making three additional assumptions. The first is that the real interest rate spreads between onshore domestic and foreign currency denominated deposits and loans do not differ. On this assumption, we can define a single onshore interest rate spread,  $\delta$ , such that  $\delta = (r - r^*) = (r_L - r_L^*)$ . The second is that the depositor's and borrower's degrees of aversion to risk are identical:  $\zeta_D = \zeta_L$ . Given these two assumptions, the currency compositions of the depositor's and the borrower's portfolios will always be either at the MVP dollarisation ratio, or equidistant from the MVP ratio on opposite sides of it. The third assumption is that financial intermediaries are restricted to balanced open foreign exchange positions, so that depositor's and borrower's portfolios must be identical:  $\mu_D = \mu_L$ . Given this additional assumption, the MVP dollarisation ratio  $\Phi$  alone determines the only possible financial equilibrium in the case of a closed financial economy.

If financial intermediaries can borrow abroad to finance the demand for onshore foreign currency denominated loans, it is possible for the equilibrium dollarisation ratio to diverge from the MVP dollarisation ratio even under the three assumptions imposed above, to the extent that the financial system's net foreign position diverges from zero. The argument can be illustrated as follows. Maintaining the three assumptions described above, we define  $X$  to represent financial intermediaries' total net cross border borrowing, and  $D$  and  $L$  total deposits (including cross border deposits) and total loans respectively. The balance sheet of the financial system can therefore be characterized as meeting the following condition:

$$(1 - \gamma_3)D + X = L \quad (3.20)$$

Equation 3.20 implies that the difference between deposits and loans must be financed by the

financial intermediaries' net foreign position:

$$D - L = \gamma_3 D - X \quad (3.21)$$

Since by assumption financial intermediaries maintain a balanced open foreign exchange position, an expression can also be derived for the domestic currency denominated part of the balance sheet of the financial system alone:

$$(1 - \mu_D)D = (1 - \mu_L)L \quad (3.22)$$

By combining Equations 3.17 and 3.19, the solutions to the depositor's and borrower's portfolio allocation problems, we can deduce that deviations from the MVP dollarisation must be symmetric:

$$\zeta_L(\mu_L - \Phi) = \frac{1}{\Psi} \delta = \zeta_L(\Phi - \mu_D) \quad (3.23)$$

and substituting Equations 3.17 and 3.19 and the identity 3.21, into Equation 3.22, we obtain:

$$\delta = -\frac{\frac{1}{\Psi}(1 - \Phi)(\gamma_3 D - X)(\zeta_D \zeta_L)}{(\zeta_D D + \zeta_L L)} \quad (3.24)$$

from which it is clear that these deviations depend upon the financial system's net foreign position,  $(\gamma_3 D - X)$ .

**Assessment** The Ize and Levy Yeyati (2003) model makes two important advances on the Thomas (1985) model. First of all, it generates the unambiguous hypothesis that observed dollarisation driven by store of value motives will gravitate around the MVP dollarisation ratio, allowing for deviations induced by a non-zero net foreign asset position of the financial sector. This hypothesis (if it is not rejected) is important for policy, because the MVP dollarisation ratio is a function exclusively of the expected volatilities, rather than levels, of inflation and real depreciation. A particularly attractive implication of it is therefore, for example, that it provides a simple explanation for the phenomenon of dollarisation hysteresis alternative to those discussed in Section 3.4.2. If stabilisation involves targeting a stable real exchange rate, then domestic inflation, despite being reduced to a low level, may well be relatively more

volatile than the real exchange rate; with the result that the MVP dollarisation ratio, and hence the level of dollarisation actually observed, will remain at high levels even after a successful disinflation. The second important advance of the Ize and Levy Yeyati (2003) model is its emphasis on observed dollarisation as a financial equilibrium between borrowers and lenders - and its analysis of the decisions of both sides of this relationship. This marks the model out not only from the models of transactions dollarisation reviewed in Section 3.4.2 above, but from the earlier versions of portfolio balance approach reviewed above.

These advances, however, come at a cost. First, the unequivocal identification of the MVP dollarisation ratio as a determinant of observed dollarisation is achieved by abstracting entirely from transactions motives. This is clearly in one sense a regression from the Thomas (1985) model: and one that Ize and Levy Yeyati (2003) justify by confining the empirical ambitions of the model to the explanation of financial dollarisation only. Second, the assumptions required to identify a unique dollarisation ratio that solves both the borrower's and the lender's optimization problems may not hold in empirical cases. The interest rate spreads between onshore domestic and foreign currency denominated deposits and loans may not be identical<sup>22</sup>; banks may be less risk averse than depositors (or *vice versa*); and perhaps most relevantly in developing and transition countries, banks may not maintain balanced foreign currency positions - or even if they do, significant credit risk may attach to banks' foreign currency loan portfolios due to the probability of borrowers' defaulting in the event of a devaluation, and so induce banks to hold a smaller proportion of their loan portfolios in foreign currency than otherwise<sup>23</sup>.

Despite these potential violations of its assumptions, empirical tests of the two main hypotheses generated by the modern portfolio balance model in cross-country contexts have been relatively successful. Ize and Levy Yeyati (2003), De Nicoló, Honohan, and Ize (2003), and Levy Yeyati (2003) have all found that the MVP dollarisation ratio (calculated from historical variances and covariances) is a significant factor in explaining the observed level of financial dollarisation, and that a non-zero net foreign position of the financial sector is significant in explaining deviations from it. These are undoubtedly important results, signifying the promise

---

<sup>22</sup>Indeed, Ize and Levy Yeyati (2003) consider two reasons why this might be so - differential reserve requirements at the Central Bank and public domestic debt - in extensions to their basic model.

<sup>23</sup>It should be noted, however, that these three assumptions are required only to identify a unique financial equilibrium between depositors and borrowers. Solving the depositor's or the borrower's optimisation problems on their own does not require any of them.

of the modern portfolio balance approach, even if there remain a number of open questions concerning the empirical methodology of the cross-country studies<sup>24</sup>. Nevertheless, case study and time series econometric evidence in support of the modern portfolio balance theory of financial dollarisation in developing and transition countries remains scarce. One notable exception in this sense is Vetlov's (2001) study of Lithuania, where the author finds strong evidence that much of the variation in Lithuania's deposit dollarisation ratio between 1992 and 2000 can be explained by simple portfolio considerations. In Chapter 4 of this thesis we attempt to carry forward such application of the modern portfolio balance model using time series econometric methods by investigating the determinants of deposit dollarisation in countries with currency board arrangements (CBAs).

### **The modern portfolio balance approach: extensions.**

A number of recent theoretical contributions have extended or adapted the basic framework established by Ize and Levy Yeyati (2003) in order further to analyse financial dollarisation in terms of a financial equilibrium between agents and financial intermediaries, and deviations from this. Two main types of innovation have been introduced. The first is the introduction of factors related to financial sector structure into the determination of the financial dollarisation ratio. The second is the extension of the basic model to analyse the effects of moral hazard induced by government policy on the equilibrium financial dollarisation ratio.

A representative model embodying the first of these innovations is outlined in Catão and Terrones (2000). The authors present a model that incorporates structural conditions in both the borrowing and the banking sectors<sup>25</sup>. With respect to the first, their model assumes a credit market segmented by the tradability of potential borrowers' output: producers of internationally tradable goods and services are able to contract loans denominated in foreign currency, whereas producers of internationally non-tradable goods and services cannot. With respect to the second, the model incorporates limited competition in the domestic banking sector that enables financial intermediaries to exert market power over borrowers in domestic

---

<sup>24</sup>For example, whether the *ad hoc* calculation of the MVP dollarisation ratio from historical data series makes sense; and particularly if such data series are non-stationary. See Chapter 4 for a discussion of this and other questions.

<sup>25</sup>Ize and Levy Yeyati (2003) was originally published in 1998 as IMF Working Paper 98/28: hence the apparent paradox that Catão and Terrones (2000) is an extension of Ize and Levy Yeyati (2003).

currency. Introducing these structural factors as additional arguments in the choice functions of borrowers, lenders, and financial intermediaries modifies the results of the Ize and Levy Yeyati (2003) model. In particular, the availability of internationally tradable collateral and the precise degree of market power enjoyed by financial intermediaries in the market for domestic credit assume considerable importance in explaining the equilibrium level of liabilities dollarisation. Overall, financial dollarisation is determined by the interaction of structural factors, prices in the markets for loans and deposits, and macroeconomic shocks, rather than the first and second moments of a limited menu of macroeconomic variables identified by the Ize and Levy Yeyati (2003) model.

Examples of models introducing the effects of moral hazard induced by aspects of government policy are provided by Burnside, Eichenbaum, and Rebelo (2001) and Broda and Levy Yeyati (2003). Burnside, Eichenbaum, and Rebelo (2001) adopt the identical approach (though using an original model) to Ize and Levy Yeyati (2003) of modelling the degree of financial dollarisation as an equilibrium outcome in the banking sector. The authors show that both currency mismatches between banking sector gross assets and liabilities, and the decision not to hedge the resulting exchange rate risk, can be explained as optimal choices of the banking sector in response to implicit or explicit guarantees from government or the monetary authorities. A common type of such a guarantee analysed in the authors' model is the blanket external price guarantee furnished by a fixed exchange rate regime. The authors show that a government may opt to maintain a fixed exchange rate for the benefits it confers in terms of a higher or more stable level of output, and in doing so induce moral hazard which results in increased financial fragility due to a subsequent unhedged currency mismatch between banking sector assets and liabilities.

Broda and Levy Yeyati (2003) extend this analysis in two different directions. Their model enables a more detailed analysis of banks' choices of the currency composition of their liabilities, and demonstrates first, that the failure of bankruptcy legislation or institutions to distinguish between the claims of domestic and foreign currency creditors on illiquid or insolvent banks may induce banks to accept foreign currency deposits in excess of the social optimum; and second, that in addition to the sort of government guarantees analysed by Burnside, Eichenbaum, and Rebelo (2001), deposit insurance that likewise fails to distinguish between domestic and foreign

currency deposits represents an analogous source of moral hazard that may encourage ‘excess’ dollarisation of deposits. The root of the moral hazard problem in both cases is the existence of costs associated with the exchange rate risk introduced by unhedged foreign currency deposits, which are not internalised by banks and depositors. In the former case, these costs could be internalised in the form of higher provisioning (and therefore lower deposit interest rates and a relatively smaller volume of foreign currency deposits), but are not as a result of the implicit, free subsidy represented by a currency-blind liquidation regime; in the latter, the costs could be internalised by higher deposit insurance premiums to be paid on foreign currency deposits, but are not as a result of the presence of a currency-blind deposit guarantee scheme.

These two types of extension to the Ize and Levy Yeyati (2003) model do not therefore alter its essential results. The observed level of financial dollarisation is still explained as the equilibrium outcome of the interaction between borrowers, lenders, and financial intermediaries seeking to maximise the risk-adjusted returns to their financial portfolios. The arguments on the basis of which this maximisation is undertaken still include the first and second moments of macroeconomic variables and prices in the markets for loans and deposits. A non-zero net foreign position of the financial sector may still allow the observed financial dollarisation ratio to deviate from the MVP dollarisation ratio. Two further important sources of divergence are also identified in the extended models, however. The first is the structure of the markets for loans and deposits – how competitive they are on the supply side, and to what degree different categories of borrowers and lenders enjoy identical conditions on the demand side. The second is the existence of externalities of one sort or another, possibly generated by government policies such as a fixed exchange rate regime or a deposit insurance scheme, which induce moral hazard on the part of agents or financial intermediaries.

### **The intertemporal public finance approach**

The dollarisation of public sector liabilities has generated a special theoretical literature primarily aimed at explaining its determinants that may be called the ‘intertemporal public finance’ approach to explaining dollarisation, and which is only indirectly related to the modern portfolio balance approach surveyed in the previous sections. Despite what might seem to be its relatively limited explanatory ambitions, relative to the theories reviewed above, this is an

important alternative theoretical approach, given the potential consequences of public sector liability dollarisation for the independence of monetary and exchange rate policy<sup>26</sup>.

The hallmark of the intertemporal public finance approach, and its similarity to the modern portfolio balance approach to explaining store of value dollarisation more generally, is that it interprets the public sector's decisions concerning the currency composition of its liabilities as the endogenous outcome of its solving a portfolio balance problem. The government, for example, has a public sector borrowing requirement; and it seeks to finance this requirement by selling bonded debt as cheaply as it can. The public finance approach starts from the empirical observation, however, that the great majority of developing and transition countries are either simply unable to float debt securities denominated in the domestic currency – in which case we might say that there is a missing market problem – or can only do so at a much higher cost than they can float securities with identical maturities but denominated in some foreign currency (typically, the US Dollar, Euro, or Yen). Two different theories as to why this is the case are proposed in the recent theoretical literature.

The first, originally due to Calvo and Giudotti (1989), and most recently restated by de la Torre, Levy Yeyati, and Schmukler (2002), offers a forward-looking explanation. The governments of developing and transition countries face a moral hazard in public finance. Once they have succeeded in selling their domestic currency denominated debt securities, it is in their interest to repudiate the repayments in real terms by engineering the domestic inflation or external depreciation of the domestic currency. However, this moral hazard is understood by potential lenders (both domestic and international), who therefore either refuse to purchase these governments' securities, or do so only at the price of demanding a very high yield. The second theory, proposed by Eichengreen and Hausman (1999), for example, offers instead a backward-looking explanation. Potential lenders (again, domestic or international) are unwilling to purchase the debt securities of the majority of low and middle because of their poor track record of macroeconomic management. The governments of these countries, and the currencies that they manage, are cursed with 'original sin', as the metaphor common in the literature has it.

Both theories present the developing or transition country government's inability to finance

---

<sup>26</sup>See Section 3.5 below.

its requirements via domestic currency denominated borrowing as in essence the outcome of a lack of the credibility of its macroeconomic policy and institutional framework – and are therefore in tune with the modern nominal anchor – credibility paradigm for international macroeconomics. Potential lenders are not convinced, either on the basis of past experience, or on the basis of reasoning about the government’s future incentives, that the real value of the debt repayments they will receive will not deteriorate as a result of the government’s macroeconomic management. Since credibility is the root of the problem, an attractive solution for the government is to borrow credibility from a foreign monetary authority by denominating its securities in foreign currency instead. Doing so functions either as a powerful commitment mechanism, in the case where the time consistency of future monetary policy is at issue; or absolves the government of its original sin, in the case where a profligate past is the problem – since it rules out the possibility of using monetary policy to repudiate the debt once the issue has been sold. The result is the dollarisation of public sector liabilities; which may then itself be interpreted in turn as a useful commitment mechanism whereby the developing or transition country government in question binds itself to the pursuit of low inflation monetary and stable exchange rate policy regime, so that it does not risk escalating or volatile foreign currency denominated debt repayments, and the associated risks to its solvency or liquidity<sup>27</sup>.

Given its focus on the dollarisation of public sector liabilities, what capacity does this public finance approach have to explain store of value dollarisation in developing and transition countries more generally, and particularly, financial dollarisation in the private sector? The implications of the theory are ambiguous. If the dollarisation of public sector liabilities tends to retard the development of domestic currency financial intermediation more generally, then the dollarisation of public sector liabilities will tend to increase the dollarisation of private sector finance as well. An opposite view is also easily accommodated, however. In many developing and transition countries government borrowing in thin domestic currency capital markets is commonly believed to crowd out the private sector. If markets for domestic currency denominated financial instruments are already at least moderately well-developed, therefore, it might be argued that the factors identified by the public finance approach actually alleviate this crowding out effect, and mean that fewer private sector residents are forced to resort to foreign

---

<sup>27</sup>This is the line of argument developed by Mizen and Pentecost (1994) with respect to the EMS.

currency denominated borrowing – so countervailing the effect of the government’s borrowing in foreign currency on the overall level of net financial dollarisation. Finally, if the view that the dollarisation of public sector liabilities may itself serve as a commitment mechanism that increases the credibility of the government’s macroeconomic policy is accepted, then it can be argued on the public finance approach that the government’s borrowing in foreign currency will make residents more willing to purchase non-indexed domestic currency denominated financial instruments as a result of the lower level of purchasing power risk attached; and that again, the net effect of the factors driving public sector liabilities dollarisation on the overall net level of financial dollarisation will therefore be ambiguous.

### **3.5 Policy Responses to Dollarisation**

In Section 3.3, we established that recent research shows that currency substitution and financial dollarisation are increasingly defining features of many developing and transition countries. In Section 3.4, meanwhile, we reviewed a variety of contributions to the modern theoretical literature that hypothesise why households, firms, and governments choose to use foreign currency to transform wealth through time (store of value dollarisation) and to settle transactions (transactions dollarisation) – why, in other words, this is the case. On the basis of these previous sections, we now proceed to answer two questions concerning the appropriate policy response to dollarisation: first, "Should governments implement proactive policies in the face of dollarisation at all?" and second, "If so, what should those policies be?"

Partial answers to both these questions can be given based on the theories of dollarisation set reviewed in Section 3.4 alone. Both the monetary services models of transactions dollarisation surveyed in Section 3.4.2 and the portfolio balance models of store of value dollarisation surveyed in Section 3.4.3 are based on utility-maximising micro-foundations, so that their solutions can be characterised as optimal levels of dollarisation in terms of social welfare. Insofar as currency substitution induced by transactions dollarisation exceeds or falls short of the optimal level determined by agents’ demand for foreign currency means of payment in order to purchase demanded goods and services, and insofar as currency substitution or financial dollarisation induced by store of value dollarisation exceeds or falls short of the optimal level determined by

portfolio balance considerations, a case in favour of policy intervention can evidently be made based on the models reviewed in Section 3.4 alone. Since those models are limited in scope to the dollarisation phenomenon alone, however, a comprehensive analysis of the appropriate policy responses to dollarisation requires in addition that we take into consideration a variety of potential macroeconomic and financial sector consequences of dollarisation that are not tractable within them. The recent literature has identified several such potential consequences – such as instability of money demand and the exchange rate; the loss of seignorage revenues; and increased financial fragility – which can be evaluated only in the context of more general macroeconomic models. An assessment of the desirability and design of policy interventions in the face of dollarisation must therefore weigh the costs and benefits of these consequences as well as those implied by the models in Section 3.4 alone.

In order to answer the first of our two questions – Should governments make proactive policy in the face of dollarisation at all? – in Section 3.5.1, we begin by reviewing what broader economic consequences of dollarisation are identified in the literature; and then proceed to ask whether the combination these consequences with the theories set out in Section 3.4 justifies a policy response. Overall, we find that the case for a policy response to currency substitution induced by transactions dollarisation is ambiguous; whereas the modern portfolio balance approach to explaining store of value dollarisation, and the potential consequences of financial dollarisation for financial sector fragility in developing and transition countries, combine to make a strong argument in favour of proactive policy to manage financial dollarisation in particular. In Section 3.5.2, we then proceed to the second question: What kinds of proactive policy are appropriate for the management of dollarisation?

### **3.5.1 The case in favour of a policy response to dollarisation**

#### **Macroeconomic and financial sector consequences of dollarisation**

The potential consequences of dollarisation for the broader macroeconomy and the financial sector identified in the literature can be organised into five categories as follows.

**Currency substitution and nominal instability.** In the immediate post-Bretton Woods era, the instability of money demand and of exchange rates was a prominent topic of research

and currency substitution was empirically the most important form of dollarisation. A number of prominent analysts - for example, Brittain (1981) and McKinnon (1982) - sought to connect the two, arguing that the failure of previously stable money demand equations observed from the mid 1970s onwards was due to demand-side currency substitution. A high degree of substitutability between internationally convertible currencies could result in sudden changes in national monetary demand as a result of expected exchange rate movements via two channels: the first a direct one, involving international banks and trading firms exchanging working cash balances between currencies; and the second an indirect one, involving the domestic interest rate changes required to maintain Fisher parity affecting the yields on domestic currency interest-bearing assets, and hence the demands for domestic cash<sup>28</sup>. Currency substitution was also linked to the unexpected volatility of nominal exchange rates. Girton and Roper (1981), for example, argued that if two co-circulating currencies were perfect substitutes for one another, their relative price would be indeterminate; and with less than perfect substitutability, the adjustments in the relative price that would clear the money market would increase with the degree of currency substitution. Recent versions of these arguments connecting currency substitution to nominal instability are found in Berg and Borzenstein (2000) and Baliño, Bennett, and Borzenstein (1999).

**Financial dollarisation and the pass-through** Over the 1990s and 2000s, financial dollarisation has displaced currency substitution as the most significant empirical form of dollarisation in developing and transition countries: and it in turn has been linked to nominal instability via its effect on the rate of pass-through from changes in the exchange rate to domestic prices. The basic channels through which financial dollarisation might be expected to increase the pass-through are intuitive. One direct channel is provided by the calculation of the price indices themselves. If some share of the total stock of domestic financial assets or liabilities on the one hand, or some proportion of all retail or wholesale prices on the other, is either denominated in foreign currency or indexed to changes in the exchange rate, then the rate of pass-through for this sector of the economy will necessarily be equal to unity; and the average rate of pass-through for the economy as a whole will be higher than if there were no dollarisation. An

---

<sup>28</sup>In terms of the definitions introduced in Section 3.2 above, the direct effect here results in store of value dollarisation, and the indirect one in transactions dollarisation.

indirect channel is provided by price-setting on the part of domestic producers and retailers whose liabilities are denominated in foreign currency. Such producers and retailers will have a natural incentive to adjust as quickly as possible the domestic currency prices at which they sell into the domestic market in line with exchange rate changes which instantaneously affect their effective cost of capital. Ize and Levy Yeyati (2003) and Ize and Parrado (2002) investigate these and other potential links between financial dollarisation and an increased rate of pass-through in theory. The empirical evidence is, however, mixed. Hohohan and Shi (2003), De Nicoló, Honohan and Ize (2003), and Reinhart, Rogoff, and Savastano (2003) all find empirical evidence of a systematic link between financial dollarisation and the rate of pass-through. Empirical studies with a narrower geographical scope, however – such as Gonzalez-Arraya's (2000) study of a sample of Latin American economies and Kraft's (2002) study of Croatia – do not. Moreover, the recent, general macroeconomic literature on the pass-through has typically found it to be determined primarily by the inflationary environment, either as a simple result of 'persistence' – as in Taylor (2000) – or because it serves as a signal of the credibility of the monetary authorities – as in Chaudhri and Hakura (2001).

**Financial sector fragility in the short term.** A further potential impact of financial dollarisation is that it may increase financial fragility. The literature postulates a number of avenues whereby this may happen. The most common argument, made in numerous recent contributions – Baliño, Bennett, and Borzenstein (1999), de Nicoló, Honohan, and Ize (2003), and Levy Yeyati (2003), for example – is that domestic banks which accept a significant proportion of liabilities in foreign currency, are liable to liquidity or solvency shocks in the event of a rapid and sizeable depreciation of the real exchange rate either directly, if they lend predominantly in domestic currency, or indirectly (via credit risk) if they lend in foreign currency to borrowers whose revenues are predominantly in domestic currency. The financial fragility so generated is not only undesirable in itself, since it increases the likelihood of banking and currency crises as documented in Baliño, Bennett, and Borzenstein (1999) and Gulde et al. (2003), but may also be an important cause of the 'fear of floating' which Calvo and Reinhart (2002) show is exhibited by so many developing and transition country monetary authorities – and may thereby lead to the development of other macroeconomic imbalances over the longer

term<sup>29</sup>. Another avenue whereby financial dollarisation may increase financial fragility that is identified in the literature is via the disabling of standard mechanisms for the prevention and remedy of financial crises. Gulde et al. (2003), for example, argue that both the lender of last resort facility provided by the monetary authorities and conventional deposit insurance schemes, are generally ineffective in a financially dollarised banking sector because they rely ultimately on the state's prerogative to issue domestic currency.

**Financial sector development over the medium to long term.** In addition to this short term effect on the stability of the financial sector, the literature also argues that financial dollarisation may have a significant influence on its longer term development. On the one hand, it is argued that financial dollarisation may be a positive factor of the overall depth of financial intermediation (regardless of currency of denomination), because it offsets the impact of, for example, episodes of high inflation that might otherwise lead to a flight from money *tout court*. This argument is made by Baliño, Bennett, and Borzenstein (1999), for example, and is corroborated by empirical evidence presented in De Nicoló, Honohan and Ize (2003), Ize and Levy Yeyati (2003), Levy Yeyati (2003), and Reinhart, Rogoff, and Savastano (2003). On the other hand, financial dollarisation may have a negative medium term effect on the structure of financial intermediation – in that it may retard the development of capital markets by encouraging the government either to issue its debt offshore and in foreign currency rather than to build up a domestic currency yield curve – and also on conditions in domestic financial markets – in terms of their completeness and competitiveness, and the spreads between deposit and loan rates of interest. Using a cross-country dataset, for example, Honohan and Shi (2003) find that that on average roughly 50% of an increase in deposit dollarisation is used to finance foreign currency assets held abroad, rather than recycled within the domestic banking system, and that deposit dollarisation tends to widen the spread between the average deposit and loan rates of interest.

---

<sup>29</sup>Note that the combination of this and the previous potential consequence of financial dollarisation may indeed represent something of a Catch-22 situation. If financial dollarisation does not increase the pass-through, then monetary and exchange rate policy will not be constrained by fears of importing inflation, but the economy will be financially fragile because domestic firms borrowing in foreign currency will represent a significant credit risk to their banks. If, on the other hand, financial dollarisation does increase the pass-through, this source of financial fragility is mitigated - since rapid re-pricing will leave the real exchange rate unchanged - but it may be impossible to combine an inflation-targeting monetary policy with any substantial exchange rate flexibility.

**Loss of seignorage revenue.** Finally, the loss of government revenue from seignorage is identified in the literature as an important potential consequence of both currency substitution and financial dollarisation. In a classic paper discussing the costs and benefits of maintaining an official, national currency, Fischer (1982) argued that since seignorage revenue is earned by taxing domestic base money, any category of dollarisation that affects the quantity of domestic base money must clearly undermine it. Since domestic base money is created in one of two ways – either by the monetary authorities issuing domestic currency cash for circulation in the economy, or by commercial banks making deposits against unremunerated reserve requirements (URRs) at the Central Bank as a result of attracting the savings of residents – both currency substitution and financial dollarisation may have a role in this. Currency substitution will unequivocally reduce the demand for that part of base money required to settle transactions, and hence the tax base for seignorage. Financial dollarisation, on the other hand, may reduce the creation of base money via the accumulation of commercial bank reserves, so long as the URRs on domestic currency bank deposits are higher than those on equivalent foreign currency bank deposits – a feature that Fritz-Krockow et al. (2001) show to be characteristic of a number of developing and transition countries’ banking sector prudential regulations.

### **Dollarisation: a case for policy intervention?**

Does the combination of these potential consequences of dollarisation with the models surveyed in Section 3.4 above provide grounds for policy intervention? With respect to currency substitution, the potential consequences of identified above – increased macroeconomic instability, and less effective public revenue generation – are fairly unequivocally negative. The consequences of financial dollarisation admit of a wider variety of normative assessments. The potential for financial dollarisation to impair the long run development of the structure and conditions of financial intermediation must be accounted an undesirable consequence; but it must be weighed against the empirical evidence that financial dollarisation has often sustained the level of on-shore financial intermediation in the face of unstable macroeconomic circumstances. Likewise, if financial dollarisation does indeed increase the rate of pass-through, it would seem that this would be a bad thing, and especially so if the monetary authorities of the country in question were pursuing a strategy of inflation targeting. However, increased pass-through may in fact be

a desirable outcome in the presence of financial dollarisation – since it will reduce the likelihood of a liquidity or solvency crisis in the event of rapid exchange rate movements if banks maintain balanced foreign exchange positions, and the exchange rate risk is borne by their borrowers. In either case, however, the financially dollarised country risks losing a substantial degree of autonomy in macroeconomic management – either as a result of less control over the domestic price level, if financial dollarisation does raise the rate of pass-through; or as a result of a fear of floating in the face of balance sheet currency mismatches in the financial, enterprise, or public sectors, if it does not. This loss of policy autonomy might seem to be unequivocal grounds for policy intervention – but even that is not necessarily so. As we discovered above, there is an alternative view which argues that it is a good thing that governments be deprived of some degree of freedom in macroeconomic policy-making, since this can invest their macroeconomic policies with a credibility which they would not otherwise enjoy. It is therefore unclear in which direction the potential broader economic consequences of dollarisation tip the scales with respect to a rationale for policy intervention.

Insofar as observed currency substitution or financial dollarisation can be assigned to store of value motives, however, the modern portfolio balance theory of store of value dollarisation reviewed in Section 3.4.3 does offer on its own a transparent case for policy proactivity in some circumstances. The modern portfolio balance theory's characterisation of the level of store of value dollarisation in an economy as a financial equilibrium determined by the interaction of the optimal choices of borrowers and lenders provides a natural normative benchmark against which to assess actual cases of financial dollarisation. The equilibrium level of financial dollarisation is the optimal response by residents to the perceived risks and returns to holding financial assets and liabilities denominated in different currencies, and deviations from this equilibrium can be analysed in terms of the standard microeconomic theory of market failure. To this end, the theoretical literature explores a number of common features of the institutional and policy frameworks of developing and transition country financial sectors, and illustrates how these may induce residents to choose a level of financial dollarisation that will deviate from the social optimum that would obtain if there were a complete and perfect set of markets for every class of financial asset, denominated in every currency. The presence of such features may therefore make the observed level of dollarisation super-(or sub-)optimal – and hence argue for proactive

policy in order to reduce (or increase) it.

In Section 3.4.3, we found that the recent literature separates these features into three basic categories – the first, of features relating to the lack of perfect competition financial intermediation; the second, of missing markets for certain classes of financial assets denominated in domestic currency; and the third, of divergences between the private and social costs of holding certain classes of foreign currency denominated financial assets generated by macroeconomic or financial sector regulatory policy. The first of these categories was discussed in Section 3.4.3, where we reviewed arguments that less than perfectly competitive conditions in the markets for deposits and credit, or conditions that are not symmetrical for borrowers and lenders, might induce deviations from the optimal financial dollarisation ratio. The second category was discussed in the context of the dollarisation of public sector liabilities: since many developing and transition countries are unable to float bonds denominated in domestic currency – since, in other words, the market for domestic currency denominated debt is missing – they are forced to denominate a larger proportion of their liabilities in foreign currency than they would do based on risk-adjusted return considerations alone. The third category was also discussed in Section 3.4.3: currency blind deposits insurance, implicit guarantees from government against solvency or liquidity risk, and differential unrestricted reserve requirements on deposits denominated in domestic and foreign currencies were cited as common examples of institutions and policies that allow the private cost of banks’ or residents’ holding financial assets or liabilities denominated in foreign currency to diverge from the cost of doing so to society as a whole.

The case for policy intervention is therefore clearest in the case of store of value dollarisation. Although the overall social welfare impact of its potential consequences is moot, the modern portfolio balance theory of dollarisation provides a familiar rationale for policy intervention to ensure competition in financial intermediation, to complete missing markets, and to bring into equality the private and social costs of domestic agents’ decisions to allocate their portfolios between domestic and foreign currencies. In the next section, therefore, we proceed to review what particular policies might be deployed to do this.

### 3.5.2 Policy options to manage dollarisation

#### Macroeconomic policy

The first, and perhaps most obvious, set of measures recommended in the literature to manage dollarisation is that of macroeconomic policies. The scope and consequences of macroeconomic policy goes far beyond the financial sector, of course, and the management of dollarisation will be only one amongst many competing goals that a government will be seeking to attain in designing and implementing it. Nevertheless, the recommended policy stance that emerges is far from controversial by more general standards. Leaving aside the issue of dollarisation hysteresis, the clearest policy implication of the theory surveyed in Section 3.4.2 is that the maintenance of a low and stable inflationary environment is important if monetary authorities are to avoid a flight from domestic currency to foreign currency cash as a means of payment. As to the key question of how the monetary authorities are to achieve such an environment – in other words, of what monetary policy regime should be adopted – much will depend on the particular context. However, given that the choice of and adherence to a credible nominal anchor is now the regime of choice in the mainstream literature, there is one additional general point that may be made on the basis of our discussion of the modern portfolio balance theory of financial dollarisation in Section 3.4.3. Choosing the level of the nominal exchange rate, rather than the domestic price level or inflation rate, as the nominal anchor for stabilisation carries with it two important risks. The first of these emerges from Ize and Levy Yeyati's (2003) model. If, by successfully targeting the nominal exchange rate, the variance of the real exchange rate is reduced below the variance of the domestic inflation rate – with the result that the purchasing power of the real cash flow from financial assets denominated in foreign currency is more stable than that of the real cash flow from those denominated in domestic currency – a stabilisation programme may in fact induce a higher level of financial dollarisation rather than a lower one. The second risk is that analysed by Burnside, Eichenbaum, and Rebelo (2001): that the commitment to target a particular level of the nominal exchange rate may serve as an implicit blanket guarantee to banks lending domestically in foreign currency, and so may induce excess financial dollarisation. These risks imply, therefore, that, absent contextual arguments to the contrary, the ideal policy stance will be of a credible domestic inflation target with a

flexible nominal exchange rate – the former serving to discourage transactions dollarisation, and the latter, excess financial dollarisation.

Nevertheless, the recent literature acknowledges several potential difficulties in implementing this combination of macroeconomic policies in many developing and transition countries. First of all, there is the possibility, discussed in Section 3.5.1, that dollarisation induces a faster rate of pass-through. If this is the case, then the combination of an inflation-targeting monetary policy regime with a genuinely flexible nominal exchange rate may be difficult for monetary authorities to tolerate. In fact, the essence of this problem is really the more general issue of credibility, as discussed in Ize and Parrado (2002) and Levy Yeyati (2003); since, if the monetary authorities' policy stance were fully credible it could be argued that monetary authorities would have no fear of floating even in the face of a high rate of pass-through and a genuinely flexible nominal exchange rate. As Levy Yeyati (2003) argues, however, many developing and transition countries are likely to be in a situation where their past macroeconomic history militates strongly against the credibility of any sort of macroeconomic policy they might adopt – whether this be mainly as a result of perceived time inconsistency as in Burnside, Eichenbaum, and Rebelo (2001), financial fragility as in De Nicoló, Honhan, and Ize (2003), or so-called 'original sin', as in Eichengreen and Hausman (1999). Ize and Parrado (2002) add that it is difficult to see how the 'clean break' with past policy and institutions that is implied in recommending macroeconomic policy reform in order to reverse dollarisation could ever be achieved in practice.

Overall, therefore, the general conclusion to be drawn from the recent theoretical literature must be that macroeconomic policies aimed at low and stable inflation – though clearly a necessary condition of a desirable degree of dollarisation, as of so many other things – are unlikely in practice to be the most effective policy interventions to manage dollarisation proactively. Indeed, they may even be counterproductive under the circumstances described by Ize and Levy Yeyati (2003), in that they can lead to dollarisation hysteresis. Two recent empirical studies examine the experiences of countries in which policy-makers have made a deliberate attempt at 'de-dollarisation', and both have found this general theoretical conclusion to be borne out. Gulde et al. (2003) have studied a number of Latin American countries which have undergone financial crises – many of them at least in part due to financial dollarisation – and thereby

experienced a steep decline in the level of financial intermediation. The authors conclude that where stabilisation has been successful and led to a substantial recovery in the level of financial sector assets, this reintermediation has often been skewed towards foreign currency denominated financial instruments – leading, in other words, to an increase, rather than a decrease, in financial dollarisation. Peru following the reform programme in 1991-93, and Argentina after the introduction of the Convertibility Plan in 1991, are cited as recent cases in point. Likewise, Reinhart, Rogoff, and Savastano (2003) found that only four out of a total of 143 developing and transition countries in their sample (Israel, Mexico, Pakistan, and Poland) had managed to achieve a substantial and sustained reduction in the ratio of foreign currency denominated deposits to total broad money over the past two decades; and of these, only two (Israel and Poland) had done so without the overall level of financial intermediation shrinking as a result of concomitant capital flight. Given that many more developing and transition countries than these four have implemented successful macroeconomic stabilisation programmes over the same period, the hypothesis put forward in the theoretical literature that stabilisation policy, though a necessary condition, is not sufficient for de-dollarisation (let alone de-dollarisation without disintermediation as well), is given significant support by these data.

### **Financial sector regulatory policy**

If macroeconomic stabilisation policy alone is not enough, then the most obvious additional type of proactive policy suggested by the theories of dollarisation reviewed in Section 3.4 is financial sector regulatory policy. We concluded at the end of Section 3.4.3 above that the modern portfolio balance theory of dollarisation identifies three types of feature, typical of many developing and transition country financial sectors, which may induce deviations from the optimal level of financial dollarisation. One of these was the lack of perfect and symmetrical competition in the markets for deposits and credit. This is an area where competition regulatory policy – whether generic or specific to the financial sector – may clearly have an important role to play. In either case, however, the principles governing the policy response have less to do with dollarisation *per se*, and more with the promotion of competition in financial intermediation in general: since their exposition and appraisal would require reviewing a separate and substantial body of theory and evidence, we do not survey them here. Another was the absence of

markets for certain classes of domestic currency financial instrument. Remedies for this will involve policies to develop domestic currency financial markets: these are discussed in the next subsection. The third type of feature identified by the modern portfolio balance theory was the existence of institutions and policies that allow the private cost of banks' or residents' holding financial assets or liabilities denominated in foreign currency to diverge from the cost of doing so to society as a whole. These too are features which financial sector regulatory policy is clearly suited to address, and we survey the main suggestions in the literature here.

The reform of deposit insurance policies is a proposal that arises from Broda and Levy Yeyati (2003), who argue that insurance schemes that do not distinguish between deposits denominated in domestic and foreign currency effectively encourage excess financial dollarisation by failing to internalise the greater social costs or risks associated with foreign currency liabilities. An obvious policy to mitigate excess dollarisation is therefore the reform of deposit insurance schemes either to exclude foreign currency denominated deposits from their coverage (in which case it is possible that some form of privately organised insurance would arise, or that banks would self-insure, and therefore offer lower deposit rates of interest in order to charge implicit premiums), or to include foreign currency denominated deposits, but to charge a higher premium in order to reflect the systemic real exchange rate risk introduced by the acceptance of unhedged foreign currency denominated liabilities. As Garcia (1999) has shown, such policies are not widely pursued in developing and transition countries: in fewer than 20 of a sample of 72 dollarised countries were foreign currency denominated deposits found to be ineligible for the general deposit insurance scheme.

Analogous arguments can be made for the introduction of currency-specific prudential regulation in a number of other areas, with the common feature of either higher prices or quantitative restrictions imposed on foreign currency denominated financial instruments in order to internalise the social costs of balance sheet currency mismatches to the economy as a whole. Ize and Levy Yeyati (2003), for example, advocate the raising of unremunerated reserve requirements for foreign currency denominated deposits in order to reflect their greater systemic riskiness. Fritz-Krockow et al. (2001) examine a sample of 21 developing and transition countries, and find that in only five of them are URRs higher on foreign currency denominated deposits than on those denominated in domestic currency. In nine countries, URRs are identical regardless

of the currency of denomination; and in seven, they are in fact lower on foreign currency denominated deposits than on those denominated in domestic currency – exactly the opposite of what the theoretical argument of the modern portfolio balance approach would recommend.

Ize and Levy Yeyati (2003) also propose that the calculation of capital adequacy ratios for potential borrowers by banks be subject to differential thresholds depending of the currency of denomination of the loan under consideration – with higher standards applied in the case of foreign currency denominated lending. Such a policy internalises the external cost of foreign currency intermediation on the account of the potential borrower. An alternative (or complementary) way of dealing with the identical problem, that instead internalises the external cost on the account of the lender, is proposed by Del Mar Cacha and Morales (2003), who recommend that specific-to-group provisioning be required of banks. Which of these two policies is preferable in a given case will have much to do with whether circumstances favour a more centralised solution involving monitoring by the regulatory authorities (the provisioning approach), or a decentralised, compliance-based one based on audit (the capital adequacy approach); as well as with the prevalence of foreign currency denominated collateral amongst potential borrowers – a characteristic of many developing and transition country exporters, for example, which would tend to favour the capital adequacy approach.

There are other such policies discussed in the recent literature and tried in practice – for example, the currency specific treatment of liquid asset holdings required of banks prompted by the different dynamics of bank runs in the presence of financial dollarisation, or the imposition of currency-specific deposit-loan ratios, as Fritz-Krockow et al. (2001) reports to have been done in Haiti. All, however, share the vital characteristic of distinguishing the regulation of financial instruments by their currency of denomination, in order to internalise the external costs that attach to foreign currency denominated instruments according to the modern portfolio balance theory of financial dollarisation surveyed in Section 3.4.3. One final important practical aspect of such policies is how precise levels of the differential pricing or quantitative restrictions involved should be chosen. In an ideal scenario, financial regulatory authorities would presumably deploy sophisticated value-at-risk techniques in order to assess the additional systemic cost involved in the sale or purchase of any given foreign currency denominated financial instrument, and tax or limit such transactions accordingly. Honohan and Stiglitz (2001) however sound a note of

caution on this front, arguing that in small financial sectors with a limited ranges of financial instruments – such as the financial sectors of the majority of developing and transition countries – regulatory rules of thumb may in practice be just as effective as any other method.

### **Policies to develop markets for domestic currency financial instruments**

We concluded above that in addition to imperfect competition in financial intermediation and the mis-pricing of externalities, the modern portfolio balance theory also points to the simple absence of markets for some classes of domestic currency financial instruments as a third important reason why residents' portfolio allocation decisions may be skewed towards foreign currency denominated assets and liabilities. Moreover, one of the major disincentives to proactive de-dollarisation policy that has emerged out of recent experience is the risk of concomitant capital flight and disintermediation. The most recent policy literature therefore lays particular stress on the need for what Levy Yeyati (2003) calls a 'carrot and stick' approach. Levy Yeyati (2003) argues that implementation of a set of de-dollarisation policies that combines the 'stick' of reformed financial sector prudential regulation described in the previous sub-section needs to be complemented by the 'carrot' of the development of markets for domestic currency denominated financial instruments. Policy to encourage the development of such markets is important both because the persistence of missing markets is itself a potential cause of excess dollarisation, and because the imposition of taxes and quantitative restrictions on foreign currency denominated financial instruments without providing an alternative outlet for hedging and speculative demand in the form of domestic currency denominated instruments will likely lead to disintermediation.

If it is clear from the modern portfolio balance theory that policies to encourage the development of markets for domestic currency financial instruments constitute the third major class of proactive de-dollarisation policies recommended in the recent literature, it is much less clear precisely what such policies should look like. Evidently, the development of domestic currency financial markets is desirable for many reasons other than de-dollarisation alone, and the literature and experience on policies to promote it is accordingly large, and largely outside the scope of this review. A number of aspects of such policies are addressed in the recent literature on dollarisation, however. The first of these is the importance of the public sector in

pioneering the development of domestic currency securities markets. In most developing and transition countries, the public sector – perhaps even just the government – is the major issuer of bonded debt, for example; and as a result, the currency of denomination of the securities that it issues more or less defines the range of instruments available to residents for hedging and speculative purposes. Gulde et al. (2003) present Mexico as an interesting example of a middle income country where a proactive policy of domestic currency securities issuance has begun to succeed in creating markets for financial instruments such as fixed-rate mortgages where foreign currency instruments have traditionally denominated. A successful macroeconomic stabilisation programme following the 1995 financial crisis has allowed the Mexican government more recently to begin the issuance of long-term, peso-denominated, fixed-rate bonds. The markets for these domestic currency denominated securities have in turn created the conditions for a general deepening of financial intermediation, with medium term (eight year), fixed-rate peso mortgages beginning to be offered by Mexican banks.

The public sector, and particularly the government, is likely to play a vital role in the creation of domestic currency denominated financial instruments, then. In conjunction with the intertemporal public finance theory, however, this proposition harbours something of a paradox. After all, the recent literature reviewed there argued that it is precisely the difficulty that most developing and transition country governments face in persuading domestic residents to purchase their debt securities that leads them to issue foreign currency denominated debt in the first place. The central place in the story was taken by the credibility of such governments' macroeconomic policy commitments: it was argued that for many governments, the issuance of foreign currency denominated debt – an important driver of financial dollarisation in the economy as a whole – was the simplest available method of importing credibility, and of putting in place a commitment mechanism whereby potential lenders could be comforted as to the time-consistency of macroeconomic policy, and the marketability of the government's debt thereby assured.

Levy Yeyati (2003) discusses an alternative method of overcoming the low credibility of the macroeconomic policy regime, that does not lead at the same time to the financial dollarisation of the economy. This is the introduction of domestic currency denominated financial instruments the nominal returns on which are indexed to an appropriate domestic macroeconomic

indicator – most commonly an index of domestic price or wage inflation. This indexation effectively safeguards the real value of the returns to the domestic currency denominated financial instrument in question – and thereby, in theory, removes the incentive for residents to demand foreign currency denominated public sector debt in preference to securities denominated in domestic currency. The obvious risk inherent in price- or wage-indexed financial instruments is of increased inflation, however; and the experience of the several middle income countries that have attempted to deploy indexation in order to assist in de-dollarising the economy, without inducing capital flight and financial disintermediation, is mixed. Levy Yeyati (2003) argues that the examples of Chile (also discussed by Gulde et al. (2003)) and Israel demonstrate that CPI-indexed financial assets can indeed compete with foreign currency denominated substitutes; but acknowledges that the similar schemes introduced in Argentina and Uruguay in the late 1970s had to be abandoned because they led to accelerating inflation.

Levy Yeyati (2003) argues that given this equivocal empirical evidence on the merits of the indexation of domestic currency denominated financial instruments, an alternative policy that deserves consideration is so-called ‘synthetic de-dollarisation’ – whereby the government does not attempt to promote markets for domestic currency denominated instruments, but instead to neutralise the negative potential consequences of public liabilities dollarisation by indexing their returns to some price that is closely correlated with the government’s own capacity to repay. A common example of such an indexation is the practice of issuing foreign currency denominated bonds for which the level and schedule of repayment is dependent upon the level of GDP. The advantage of this type of arrangement is that it hedges the solvency and liquidity risks associated with a balance sheet currency mismatch – since in the event of a rapid depreciation of the domestic currency, GDP will decline, and payments on the foreign currency denominated debt will be suspended or rescheduled automatically. Evidently, the price chosen by the government when attempting to ‘synthetically de-dollarise’ its liabilities in this manner would need to be outside its own control – otherwise, the moral hazard identified by the intertemporal public finance theory of dollarisation would emerge here too, and potential lenders would once again prefer foreign currency denominated securities to the risk that the government would deliberately repudiate its debt at some point during the repayment period.

One final policy proposal in this area, concerning the currency of denomination of interna-

tional financial institution (IFI) intermediation, is made by Levy Yeyati (2003). In Section 3.3 above, we showed that in many developing and transition countries official lending accounts for an important share of public liabilities dollarisation. Since this is so – and since, in contrast to debt issuance on the international capital markets, the currency composition of official lending is subject to the exogenous control of the IFIs and bilateral donors – it seems natural to ask whether the IFIs and bilateral donors might not play a supporting role in developing and transition countries' efforts to de-dollarise by intermediating development finance in such countries domestic currencies, rather than, as is presently much more usual, in their own currencies, in the US Dollar or Euro, or in SDRs. Such intermediation would require IFIs and bilateral donors to issue their own securities denominated in the domestic currency of the recipient country, and then to onlend or donate the domestic currency liabilities so accumulated to the country in question. If the recipient country either enjoys a reputation for a credible macroeconomic policy regime, or the domestic currency denominated bonds in question are indexed to a domestic price or wage level as a commitment mechanism, Levy Yeyati (2003) argues that such an arrangement with IFIs and bilateral donors might succeed in defraying all three of the most important risks associated with domestic currency denominated financial instruments in developing and transition countries. Any risk associated with currency depreciation or domestic inflation would be hedged by the indexation mechanism (or absent in the case of a credible regime). Sovereign risk associated with the jurisdiction under which the securities were being issued would be eliminated by the IFIs or bilateral donors being located in international financial centres with trusted regulatory frameworks. Default risk would be reduced to the level associated with the issuing IFI or government of the bilateral donor country – typically a very low level indeed. As a result, the incentives for potential lenders to prefer foreign currency denominated financial instruments to those denominated in domestic currency would be much reduced; and in consequence, the level of financial dollarisation as well.

### **Radical responses**

For some developing and transition countries, however, it is possible that none of the three types of de-dollarisation policy options reviewed in the previous three sub-sections will be viable, for one reason or another. For such countries, two further, but more radical, policies

to combat dollarisation are discussed in the recent literature. The first of these is the use of capital controls and legal restrictions on the holding of foreign currency denominated financial instruments domestically. The second is recourse to full, supply-side dollarisation, as defined in Section 2 above.

Both Reinhart, Rogoff, and Savastano (2003), in the case of developing and transition countries generally, and Buitier and Grafe (2002), in the case of the EU accession countries of Central and Eastern Europe, have recently made the case for limited derogations from full capital account convertibility in order to manage dollarisation and its consequences. The logic for doing so is the reverse of that set out above for the stimulation of markets for domestic currency denominated financial instruments. If the consequences of the resulting excess financial dollarisation are considered too severe to be borne, and it is impossible that such missing markets can be quickly established (as is likely), then instead of attempting to attain symmetry in the markets for domestic and foreign currency denominated financial assets by attempting to furnish the missing markets for domestic currency denominated instruments, it may be wiser to do so by eliminating the extra markets for foreign currency denominated ones. Capital controls, and legal restrictions on the holding of certain classes of foreign currency denominated financial instruments domestically, can effect this. These types of administrative policy instruments have much broader economic ramifications than only to constrain dollarisation, however: and because of this, they are the object of a number of regulations of the IMF, the EU, and so on – both of which bodies seek strongly to encourage capital account convertibility. As a result, this radical policy option is unlikely to be open to many developing and transition countries except in emergency circumstances, and at the cost of expending considerable reserves of political capital.

One final option open to dollarised developing and transition countries whose governments wish to escape the attendant risks is to implement full, supply-side dollarisation in the manner of Panama, Montenegro, or, most recently, Ecuador, as discussed in Section 3.2 above. This option has been debated particularly widely with respect to the ten countries that acceded to the EU in May, 2004<sup>30</sup>. It should by now be clear from the preceding discussions, however,

---

<sup>30</sup> A special issue of *Economics of Transition* in July, 2002, presented papers from a conference on this topic, for example, in which Nuti (2002) and Coricelli (2002) analyse the costs and benefits of unilateral euroisation by accession countries. In addition to the consequences mentioned here, the standard Optimum Currency Area

that though full, supply-side dollarisation may eliminate immediately a number of the negative consequences of partial, demand-side dollarisation – such as the financial fragility associated with balance sheet currency mismatches, and the acceleration of the rate of pass-through – it will not only not eliminate others – such as the loss of seignorage revenue and of monetary policy independence – but may in fact make them worse. What is more, recent theoretical and empirical studies have analysed a variety of consequences peculiar to full, supply-side dollarisation, that we have not surveyed in this review, which would also need to be taken into account<sup>31</sup>.

### 3.6 Conclusions

As financial and capital account liberalisation has progressed over the 1990s and 2000s, financial dollarisation has increasingly become a feature of developing and transition countries. Recent empirical research has found that asset side financial dollarisation is high and following an increasing trend in all developing and transition countries outside of the Caribbean; and very high in South America, Eastern Europe, and the FSU. Liability side financial dollarisation – domestic and cross-border bank borrowing, bonded debt, and official borrowing – is following a similar increasing trend in most areas; though the degree of domestic bank loan dollarisation is generally slightly lower than that of bank deposit dollarisation due to the fact that banks generally use a substantial share of foreign currency deposits to finance foreign assets rather than domestic lending. This increasing importance of financial dollarisation should not obscure the enduring popularity of currency substitution in the developing and transitional world, however. Although by its nature more difficult to quantify exactly, recent empirical research has shown that foreign currency cash is a highly significant monetary aggregate in many developing and transition countries, in many cases dwarfing the quantity of domestic currency cash in circulation.

---

hypothesis that adopting the Euro would boost regional trade is presented as an important argument in favour of euroisation in this debate; and the costs of accumulating sufficient Euro reserves and cash, and the difficulty of choosing a sustainable parity, as important arguments against.

<sup>31</sup>For a guide to the recent literature on full, supply-side dollarisation, see footnote 3 above. Another relevant recent contribution is that of Edwards and Magendzo (2003), who review the empirical evidence on the macroeconomic record of fully dollarised economies.

The growing empirical importance of dollarisation has led to a revival of theoretical research into its determinants. With respect to transactions dollarisation, the monetary services approach provides an intuitive explanation in terms of the relative usefulness of domestic and foreign currency financial instruments as means of payment. It suggests that the relative levels of determinants familiar from the conventional analysis of the transactions demand for money – inflation, interest rates, the rate of depreciation of the domestic currency, and the volume of transactions to be settled – influence the degree of transactions dollarisation. The monetary services approach does not however offer an adequate explanation for the widely observed phenomenon of dollarisation hysteresis. Two types of theory have therefore been put forward in the recent literature to explain this. The first type, designed principally to be applicable to cases of currency substitution hysteresis, extends the monetary services approach by incorporating network externalities and switching costs to generate multiple equilibrium levels of transactions dollarisation – and then characterises hysteresis as the transition from a lower to a higher equilibrium. The second type, designed more with financial dollarisation hysteresis in mind, draws on the modern portfolio balance theory of dollarisation, and explains hysteresis as a consequence of stabilisation policies that target the real exchange rate, thereby making the purchasing power of the cash flow from foreign currency financial instruments less volatile than that from those denominated in domestic currency.

As a result of the increasing importance of financial dollarisation - which is driven primarily by store of value motives - the modern portfolio balance approach has become dominant in the most recent theoretical literature. This approach characterises the financial dollarisation ratio as the equilibrium outcome of the interaction of borrowers and lenders attempting to minimise the variance of the purchasing power of the cash flows from their portfolios of financial assets and liabilities. The main determinants of the observed financial dollarisation ratio is therefore the MVP ratio, which is itself a function of the variances of various macroeconomic variables. The modern portfolio balance approach allows for deviations of the observed financial dollarisation ratio from the optimum to be explained in a variety of ways. Important examples identified by recent models include: the existence of a non-zero net foreign asset position of the financial sector; less than perfectly competitive conditions in the markets for deposits and credit, or conditions that are not symmetrical for borrowers and lenders; the maintenance of implicit or

explicit insurance schemes that do not price the premiums associated with financial instruments denominated in different currencies to reflect their respective social costs; and the absence of markets for certain financial instruments denominated in certain currencies.

Finally, the intertemporal public finance approach is designed to explain one important sub-category of store of value dollarisation – the dollarisation of public sector liabilities. It suggests that this may result from a missing market for domestic currency-denominated government debt, itself due to the lack of credibility of developing and transition countries' macroeconomic policies or institutional frameworks, and their consequent need to borrow credibility from foreign monetary authorities via dollarisation. The public finance approach does not venture to suggest what may be the effect of this on the level of dollarisation of the economy as a whole.

Both currency substitution and financial dollarisation can have potentially serious consequences for the broader macroeconomy. In particular, currency substitution may cause nominal instability and the loss of seignorage revenue; whilst financial dollarisation, which may also deplete seignorage revenue, may increase the rate of pass-through and cause financial fragility. The longer term effects of dollarisation may, however, be either positive - in terms of sustaining the depth of intermediation in times of crisis - or negative - in terms of deteriorating the level, structure, and conditions of domestic financial intermediation. Given these various potential consequences, it is difficult to evaluate rigourously the desirability of proactive policy in the face of dollarisation - and particularly the desirability or otherwise of efforts to de-dollarise. Nevertheless, the modern portfolio balance theory implies that there are three features commonly found in developing and transition country financial sectors that induce deviations from the optimal dollarisation ratio, and which are therefore eligible candidates for policy responses. The first such feature is imperfect competition in the markets for deposits and credit. Generic or financial sector specific competition regulatory policies should be deployed to diffuse the effects of such problems. The second such feature is the existence of institutions and policies that generate externalities in financial markets by allowing the private cost of banks' or residents' holding financial assets or liabilities denominated in foreign currency to diverge from the cost of doing so to society as a whole. A variety of currency-specific financial sector prudential policies, including the differential pricing or quantitative restriction of deposit insurance premiums, reserve requirements, capital adequacy ratios, provisioning, and deposit-loan ratios,

may be appropriate to mitigate these market failures. The third common distortionary feature is the absence of markets for certain classes of domestic currency financial instruments – particularly, for example, for domestic currency government debt. The introduction of financial instruments indexed to the domestic price level can assist in the development of such missing markets, as might intermediation by IFIs and bilateral donors in the domestic currencies of recipient countries. An alternative policy response that aims to neutralise the risks associated with dollarisation without necessarily reducing its level is that of ‘synthetic de-dollarisation’. Where none of these policies are viable, there remain two, more radical options aimed at avoiding the risks associated with dollarisation: administrative controls on capital flows or holding foreign currency financial instruments domestically; and full, supply-side dollarisation. The latter policy, whilst eliminating some of the consequences of partial, demand-side dollarisation, serves to strengthen others.



## Chapter 4

# **Deposit Dollarisation in the Baltics: a Portfolio Balance Approach**



## 4.1 Introduction

Empirical studies explaining the determinants of dollarisation in cases of particular national economies are rare. No doubt this is because, as discussed in Chapter 3 of this thesis, the factors that drive dollarisation are manifold, typically comprising both transactions and store of value motives, as well the constraints imposed by a changing structural and regulatory environment. Identifying the independent effects of these various factors is typically impossible, both because no theory exists which incorporates them all in a single conceptual framework, and because the data that would be required to test it if one did is generally not available. As a result, what empirical studies of dollarisation – and especially of dollarisation in developing and transition countries – there are in the literature have tended to rely on informal and qualitative, rather than formal and econometric, analysis.

Nevertheless, in their pioneering review of dollarisation in transition economies, Sahay and Vegh (1996), argued that for one particular group of transition economies – those that stabilised under CBAs – a simple explanation of deposit dollarisation in terms of relative real rates of return to holding DC and FC deposits was available. Using short samples of quarterly data, they demonstrated that in two such countries – Estonia and Lithuania – a clear correlation could be observed over time between the deposit dollarisation ratio on the one hand, and the relative real rates of return to holding DC and FC deposits implied by actual nominal interest rates adjusted for inflation or depreciation on the other. In the cases of both Estonia and Lithuania, deposit dollarisation had begun to decrease rapidly as soon as the real rate of return on DC deposits had turned positive after the introduction of the CBAs<sup>1</sup>.

Sahay and Vegh (1996) argued that a major precondition of this surprisingly simple explanation of deposit dollarisation in terms of rudimentary portfolio balance considerations was the presence of CBAs. Prior to stabilisation, the uncertainty concerning the political and economic future of both countries was of a radical nature which would not be captured by simple measures of the financial risks and returns to alternative assets. The introduction of CBAs signalled an intense commitment to inflation and exchange rate stabilisation and thereby to the newly-won monetary independence from Russia, and as such transformed this radical uncertainty into a

---

<sup>1</sup>Indeed, Sahay and Vegh (1996) called Estonia and Lithuania examples of “successful stabilisation with de-dollarisation”: a premature assessment, as we shall see.

level of risk that, whilst high, was not qualitatively different from that attaching to financial assets in less volatile and more developed economies. Depositors' choices concerning the currency composition of their deposit portfolios in this environment could therefore be explained in terms of their consideration of conventional measures of relative financial risk and return. As a result, deposit dollarisation in transition countries with CBAs could be explained empirically as a function of the small vector of financial risk and return variables postulated by portfolio balance theory.

As we saw in Chapter 3, much recent theoretical research since Sahay and Vegh (1996) has elaborated significantly on the portfolio theoretic explanation of deposit dollarisation as the result of store of value motives. Meanwhile, attempts to model deposit dollarisation econometrically as a function of risk and return factors have met with success in a variety of both cross-country (e.g. De Honohan, Nicolo, and Ize (2003), and Ize and Levy Yeyati (2003)) and time series (e.g. Heimonian (2001), Korhonen (1996), Sarajevs (2000), and Vetlov (2001)) contexts. In this chapter, we therefore set out to test in greater detail and with the benefit of significantly longer time series Sahay and Vegh's (1996) hypothesis that deposit dollarisation in transition economies with CBAs can essentially be explained using portfolio balance theory. We do so by building dynamic, multiple equation, econometric models of the deposit dollarisation process in Estonia and Lithuania; and then interpreting them in the light of modern theoretical research. In the course of this process, we discuss in detail a number of generic econometric and theoretical challenges in the empirical application of portfolio balance theory in a time series econometric context.

To this end, the chapter is structured as follows. We begin, in Section 4.2, by introducing the two empirical cases of Estonia and Lithuania, providing an analytical narrative of their experiences of stabilisation and dollarisation over the transition, and making a preliminary evaluation of whether Sahay and Vegh's (1996) hypothesis is credible on the basis of this informal analysis. Concluding that it is, we then proceed to econometric analysis. First, in Section 4.3, we describe the strategy we deploy in econometric modelling; and in Section 4.4, the results arrived at by implementing this strategy. Then, in Section 4.5, we present an economic interpretation of these results in terms of portfolio balance theory. On the basis of this interpretation, we then both discuss what we can infer concerning the determinants of

deposit dollarisation in Estonia and Lithuania, and assess the advantages and shortcomings of the portfolio balance approach to explaining dollarisation in empirical cases in light of this. Section 4.6 concludes.

## 4.2 Introduction to the Empirical Cases

### 4.2.1 Estonia and Lithuania: Currency Reform and Stabilisation Prior to the Introduction of the CBAs

In the Union of Soviet Socialist Republics (USSR), foreign currency (FC) deposits were forbidden to households and virtually all firms - so deposit dollarisation was, by definition, virtually zero. With the reform of the USSR banking system in 1989, and the introduction of independent Central Banks in the constituent republics in 1990, restrictions on the holding of FC deposits were relaxed, however - with the result that deposit dollarisation began to rise rapidly in the uncertain political and economic environment that preceded and accompanied the break-up of the Union.

Currency reform was perceived by the Baltic republics to be an especially critical step in the assertion of independence from the USSR - indeed, it had been discussed openly in Estonia since 1987 - so that once political independence had been declared in August, 1991, both Estonia and Lithuania moved quickly to withdraw from the rouble zone<sup>2</sup>. Quite different strategies were however implemented by the two countries. In Estonia, the introduction of a CBA constituted the centrepiece of both the currency reform and the stabilisation strategy from the outset. A new Estonian currency, the kroon (EEK), was introduced in June, 1992, and its value pegged to the Deutschemark (DEM) in October of the same year. The result was a reduction of inflation from an annualised rate of nearly 500% in the second quarter of 1992 to double figures by the first quarter of 1993. In Lithuania, by contrast, an interim currency, the *talonas* ('coupon'), was first introduced alongside the rouble in May, 1992, with a permanent currency, the litas (LIT), introduced as sole legal tender only in June, 1993. The Bank of Lithuania then achieved

---

<sup>2</sup>An interesting indication of the political significance of the currency reforms is furnished by the fact, revealed only in September, 1994, that Estonia, having originally agreed to return the roubles collected during the reform to the Central Bank of Russia, in fact sold them at a steep discount to an unknown buyer - reputed to be the rebel Russian republic of Chechnya (see Lainela and Sutela (1994), p.47).

notable success in stabilising inflation around a conventional fixed exchange rate in the second half of 1993. This led to a considerable debate within Lithuania - primarily between the Bank of Lithuania and the Government - over the desirability of introducing a CBA. Eventually, the Government decided that a CBA would be introduced in April, 1994 - initially against the protests of the Central Bank<sup>3</sup>. By contrast to the Estonian CBA, the US Dollar was initially chosen as the anchor currency for the Lithuanian CBA<sup>4</sup>.

The two countries also took different paths as regards the convertibility of the reformed currencies. In Estonia, the EEK was fully convertible for current account transactions from its introduction; but in order to prevent runs on the new currency, it was decided to impose surrender requirements on the FC receipts of firms and to forbid the accumulation of new FC deposits by households in December, 1992. These restrictions were repealed in the course of the following two years - with surrender requirements abolished in 1993, and both households and firms permitted to accumulate new FC deposits from March, 1994 onwards. In Lithuania, on the other hand, all existing FC deposits were convertible from the beginning of the reform, and there were neither surrender requirements nor restrictions on the accumulation of new FC deposits by firms or households from the introduction of the LIT.

To sum up: the period between the beginnings of financial reform in the Soviet Union in 1989 and the introduction of CBAs in the newly-independent Estonia and Lithuania was characterised by internal and external macroeconomic instability on an unprecedented scale; by far-reaching changes to the institutions of the monetary and financial system and to the regulatory framework governing FC deposits; and by radical uncertainty as to the political, and therefore the economic, future of the two countries. It is *a priori* hardly likely that any simple theoretical account of deposit dollarisation, such as that provided by portfolio balance theory, would be able to provide a satisfactory explanation of the level and dynamics of dollarisation under these conditions.

---

<sup>3</sup>Hence the verdict of two close observers of the Baltic transitions that “[a]s to Lithuania, what could tell more of the role of political chance than the decision to adopt the currency board, after stabilization was succeeding without it?” (Lainela and Sutela (1994), p.66).

<sup>4</sup>Alonso-Gama, Fabrizio, Kramarenko, and Wang (2002) give three reasons for this choice: first because 90% of Lithuania’s external trade was denominated in USD; second, because most domestic FC assets and liabilities were denominated in USD; and third, because most cash transactions were being conducted using USD cash by 1993. Presumably, all three of these circumstances had also obtained in Estonia, however: so it seems that here again political chance played a lead role in Lithuanian economic history.

## 4.2.2 Estonia and Lithuania: Economic and Political Developments after the Introduction of the CBAs

Tables 4.1 and 4.2 report the main macroeconomic indicators for Estonia and Lithuania respectively during the decade 1993-2003. In both Estonia and Lithuania, deposit dollarisation, which had persisted at a high level and with significant variation during the period of the currency reforms made a rapid transition to a lower level with the introduction and consolidation of the CBAs. In Estonia, the deposit dollarisation ratio fell from a peak of approximately 60% in the second quarter of 1992 to less than 10% a year later. In Lithuania, deposit dollarisation peaked at 56% in April, 1993, and fell to 34% in the month following the successful introduction of the CBA a year after that. It was on the basis of these data that Sahay and Vegh (1996) categorised these countries (along with Latvia and Mongolia) as examples of "stabilisation with de-dollarisation". The accolade proved to be only half-deserved. As Tables 4.1 and 4.2 report, whilst neither Estonia nor Lithuania saw any significant revival of inflation and the CBAs held firm, the drop in dollarisation proved to be transitory in both countries, and the deposit dollarisation ratio remained high and variable throughout the decade 1993 to 2003.

Both in pure macroeconomic terms, and in terms of broader institutional development, stabilisation proceeded remarkably smoothly in both countries. Following substantial transitional recessions, real GDP returned to growth in both countries in 1995; neither fiscal nor current account deficits expanded unsustainably (the latter, though large, were easily financed by some of the highest rates of *per capita* capital inflows in the transition countries)<sup>5</sup>; the reorientation of trade towards the EU was rapid<sup>6</sup>; and there were no major reversals of financial sector regulatory reforms<sup>7</sup>. This is not to say that there were not substantial macroeconomic and political instability: as in many transition countries, governments in both countries tended to alternate regularly between main political parties and generally did not see out their full terms, with elections precipitated on a number of occasions (for example, in January and October, 1996 in Lithuania; and in January 2002 in Estonia) by political crises. But it is not unreasonable to argue that after the introduction of the CBAs, the highly successful stabilisation

---

<sup>5</sup>For a detailed discussion of this point, see Chapters 2 and 3 of OECD (2000).

<sup>6</sup>See OECD (2000), Chapter 7.

<sup>7</sup>Notably, for example, there has been no re-imposition of restrictions on the holding or convertibility of FC deposits in either Lithuania and Estonia since 1993 and 1994 respectively.

associated with this, and the rapid integration of the two countries with European markets, this type of instability was of a different order of magnitude to the radical uncertainty which had preceded it - and therefore of a sort which does not necessarily incapacitate simple and empirically tractable theoretical explanations to the same extent.

There were, for example, banking crises in Estonia in 1992 and in Lithuania in 1996 - but these did not affect holders of domestic currency (DC) and FC deposit in domestic banks differentially, and hence we should not expect them to have had significant effects on the onshore deposit dollarisation ratio (though they very likely influenced depositors' allocation of FC deposits between domestic and cross-border deposits). The pace of privatisation and restructuring in the banking sector was likewise fitful, and - perhaps because of the blanket external price guarantee provided by the CBAs - markets for DC financial assets that might serve to compete with FC deposits as stores of value in addition to plain DC deposits hardly developed at all<sup>8</sup>. Generally speaking, Estonia integrated more quickly and thoroughly with EU and international capital markets - mainly through the high volume of direct foreign investment in the financial and other sectors - whilst Lithuania remained, in relative terms, financially closed. The result of this dissimilarity is evident in the differential effect of the Russian financial crisis that started in August, 1998 on the two economies. Both countries suffered a fall in external demand due to the contraction of export markets in Russia - but Estonia also experienced a minor financial crisis of its own. Neither country was very much exposed financially to Russia in a direct sense - 2% of Estonia's total banking assets were Russian at the time of the crisis; and only 1.4% of Lithuania's - so that this run on the EEK and related spikes in money market and deposit interest rates are generally interpreted as an example of

---

<sup>8</sup>Lithuania, for example, has developed the largest bill and bond markets of all three ex-Soviet Baltic republics - but Vetlov (2001) reports that these are dominated by government paper, which does not serve as an alternative wealth-holding instrument for the vast majority of investors (Vetlov (2001), p.18). Likewise, the National Stock Exchange of Lithuania (NSEL) - the largest equity market in the region - was essentially an artefact of the voucher privatisation process, and far from furnishing an expanding menu of alternative DC assets, has contracted significantly in value as the stocks of the great majority of companies and investment funds initially listed were discovered to be illiquid. Of 1,200 privatised companies initially listed, only 46 remain, of which only 6 are on the primary official list: of 200 investment funds, only 11 are even operating, and only one is listed on the NSEL (IMF (2003), p.19).

Nevertheless, it is the case that in Lithuania Treasury Bills, following their introduction in late 1994, did constitute a relatively important alternative to term bank deposits (at least for the enterprise sector). For this reason, we include a dummy variable to capture the effects of the opening of the T Bill market at the modelling stage (see below).

contagion in the pure sense from the Russian default<sup>9</sup>.

---

<sup>9</sup>See OECD (2000) p.114 ff.

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
GDP growth, %	-9.0	-2.0	4.3	3.9	9.8	4.6	-0.6	7.3	6.5	6.0	4.7
Inflation, %, year average	89.8	47.7	29.0	23.1	11.1	8.2	3.5	3.9	5.8	3.6	1.3
Central government balance, % of GDP	0.8	1.3	-1.3	-1.9	2.2	-0.3	-4.7	-0.7	0.4	1.2	2.6
Current account balance, % of GDP	1.3	-7.2	-4.4	-9.2	-12.1	-9.2	-4.7	-6.4	-6.5	-12.4	-13.7
Growth of M2, %	85.9	30.9	27.8	37.2	37.8	4.2	23.5	25.1	23.8	11.2	10.9
Deposit dollarisation ratio, %, year average	-	7.5	23.2	20.4	25.8	28.9	38.0	42.9	38.6	34.2	24.1

Source: Bank of Estonia, Estonian Statistics Office.

**Table 4.1:** Estonia: main macroeconomic indicators, 1993-2003

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
GDP growth, %	-16.2	-9.8	3.3	4.7	7.0	7.3	-1.8	4.0	6.5	6.8	7.9
Inflation, %, year average	410.2	72.2	39.7	24.6	8.9	5.0	0.7	1.0	1.4	0.3	-1.2
Central government balance, % of GDP	-4.0	-5.0	-4.5	-4.5	-1.8	-5.8	-8.2	-2.8	-1.9	-1.2	-
Current account balance, % of GDP	-3.1	-2.1	-10.2	-9.2	-10.2	-12.1	-11.2	-6.0	-4.8	-5.3	-5.9
Growth of M2, %	100.2	63.0	28.9	-3.5	34.1	14.5	7.7	16.5	21.4	16.9	18.2
Deposit dollarisation ratio, %, year average	63.2	33.0	43.7	33.1	32.9	30.6	37.7	48.9	55.7	42.3	28.3

Source: Bank of Lithuania, Lithuanian Statistics Office.

**Table 4.2:** Lithuania: main macroeconomic indicators, 1993-2003

Finally, in both Estonia and Lithuania, the CBA was eventually re-pegged to a new anchor currency. In the case of Estonia, the re-pegging had few practical consequences: with the introduction of the Euro in January, 1999, the CBA commitment was re-defined with respect to the Euro at parity with the previous DEM peg. In Lithuania, however, the re-pegging had more substantial ramifications, and was consequently more controversial. As trade was reoriented away from Russia and towards the EU during the 1990s, Lithuania's prospects of joining the EU blossomed, and the significant appreciation of the USD against the Euro in the late 1990s threatened to undermine Lithuanian competitiveness, a change of the CBA anchor currency to the Euro was mooted. The re-peg was originally planned to take place in 1999 - but in the wake of political turmoil, it was postponed first until 2001; and then until February 2, 2002, when it was finally implemented<sup>10</sup>.

### 4.2.3 The development of deposit interest rates

On the basis of the brief summary above, we draw the conclusion that Sahay and Vegh's (1996) argument that the introduction of CBAs and the attendant stabilisation of macroeconomic conditions and depositors' expectations can reasonably be maintained as the basis for the econometric modelling of deposit dollarisation in Lithuania and Estonia within a simple, portfolio theoretic framework. As discussed in the previous section, this implies an empirical analysis that will focus on a small vector of explanatory variables - with deposit interest rates chief amongst them. A brief discussion of the development of deposit interest rates in Estonia and Lithuania is therefore also of interest.

Table 4.3 reports the annual average interest rate on the most representative category of DC and FC time deposits (1-3 months maturity) for both Estonia and Lithuania, for the longest period that data is available, over the CBA periods. In addition, Figures 4-3 and 4-4 illustrate the interest rate spreads on this category of deposits in the two countries, using monthly data<sup>11</sup>. Several notable points concerning the pattern of deposit interest rates emerge from these two presentations of the data.

---

<sup>10</sup>Details of the technical aspects of the re-pegging of the LIT from the USD to the Euro - interesting in their own right, since such an operation has rarely been done - are given in Alonso-Gama, Fabrizio, Kramarenko, and Wang (2002), pp.8-9.

<sup>11</sup>Unfortunately, deposit interest rate data disaggregated by currency and maturity are available only from 1997 onwards. See Appendix A below for further details.

		1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Lithuania	DC	93.9	46.9	20.1	13.6	8.1	6.5	7.4	6.7	4.6	2.7	2.0
	FC	27.7	24.4	15.4	8.9	5.2	4.0	4.2	4.5	3.6	1.8	1.6
	Spread	66.2	22.5	4.7	4.7	2.9	2.5	3.2	2.2	1.0	0.9	0.4
Estonia	DC	—	—	—	—	6.3	7.9	3.9	3.6	3.9	2.6	2.3
	FC	—	—	—	—	4.6	4.9	3.5	4.8	3.6	2.2	1.6
	Spread	—	—	—	—	1.7	3.0	0.4	-1.2	0.3	0.4	0.7

Sources: Bank of Estonia, Bank of Lithuania.

**Table 4.3:** Average annual interest rates on 1-3 month domestic (DC) and foreign currency (FC) time deposits in commercial banks

First of all, as is evident from the Lithuanian data, nominal interest rates not only on DC but also on FC deposits were very high in the early years of transition. The fact that rates on FC deposits diverged so significantly from comparable "world" interest rates may seem surprising, given the absence of foreign exchange controls from 1993 onwards: but this is in fact a pattern that was not uncommon in the early stages of transition in a number of Eastern European and former Soviet Union (FSU) countries. The situation was similar in the early years on Estonia's transition, and the wedge between world and domestic rate on FC deposits can be explained by two main factors. The first factor is country risk: as argued by Serven and Schmukler (2002), both macroeconomic and regulatory risks peculiar to the domestic country can cause interest rates on FC deposits in domestic banks to diverge from deposits of otherwise identical maturity and default risk profiles held abroad even under a CBA. The second factor is the nature and profitability of the financing transactions being undertaken by Lithuanian and Estonian banks in the early transition period. This factor is discussed by Lainela and Sutela (1994), who argue that the exceptional profitability and riskiness of providing scarce FC liquidity to Russia in these years enabled banks in the Baltic republics to pay these high deposit interest rates on the one hand, and required them to do so in order to attract deposits from risk averse depositors on the other.

Second, the spread between the interest rates on DC and FC deposits is generally positive and declining over time. Once again this might at first seem surprising against a dogmatic interpretation of interest parity, given the presence of CBAs in both countries. An obvious explanation is that expectations of a possible devaluation persisted (and varied) in both Lithuania and Estonia, despite the *ex post* fact that their CBAs were maintained without adjustment

throughout the period. Credence is added to this informal interpretation of the DC-FC interest rate spread as a measure of the expected depreciation of the DC by Figure 4-3, where the impact of contagion from the Asian and Russian financial crises can be clearly traced in the behaviour of the deposit interest rate spread in Estonia between late 1997 and early 1999<sup>12</sup>.

The fact that the DC-FC interest rate spread sometimes falls below zero (for example, in late 1999 and 2000 in Estonia) suggests that this interpretation in terms of expected depreciation alone may be too simplistic, however; and that there may be other factors driving the development of deposit interest rates as well. Despite their rapid financial development during the 1990s relative to other transition economies, the financial sectors of both Estonia and Lithuania remained dominated by banks and highly concentrated relative to those of more developed market economies. At the end of 2001, for example, total banking assets were valued at approximately 250% of GDP in the Euro area banking system; the same statistic stood at approximately 72% of GDP for Estonia, and only 32% of GDP in Lithuania<sup>13</sup>. It is reasonable to suppose, therefore, that the interest rates offered by banks on DC and FC deposits might have been influenced not only by expectations of depreciation, but also by banks' incentives to manage the currency composition of their liabilities. More precise analysis of the importance of this type of factor in the evolution of the relative levels of the interest rates on DC and FC deposits would however require detailed analysis not only of the composition of deposits, but of bank's assets.

## 4.3 Empirical strategy

### 4.3.1 Choice of variables

In accordance with basic theoretical priors and the precedents set by previous econometric studies of deposit dollarisation in the Baltics (Korhonen (1996); Sarajevs (2000); Vetlov (2001)), the main variables constructed for the purpose of econometric modelling were series measuring the deposit dollarisation ratio ( $DR_t$ ), the interest rates on DC and FC deposits ( $I_t^{DCD}$  and

---

<sup>12</sup>Ideally, this interpretation could be checked against data on discounts in the markets for forward contracts: unfortunately, in neither Estonia nor Lithuania are these markets developed enough for their respective Central Banks to collect data on them.

<sup>13</sup>IMF (2003), p.5.

$I_t^{FCD}$ ), and the rate of depreciation of the DC against non-CBA anchor foreign currencies ( $N_t$ ). A detailed description of the construction of the series used in modelling is given in Appendix A: since a number of elements are relevant to the interpretation of the models, we give a brief summary here.

The most intuitive measure of the deposit dollarisation ratio is calculated as  $DDR_t = \left( \frac{FCD}{FCD+DCD} \right)_t$ , where  $FCD$  are FC and  $DCD$  DC deposits respectively. This variable is bounded between 0 and 1, however, and hence not suitable for OLS estimation; so that we use instead the variable  $DR_t = \left( \frac{FCD}{DCD} \right)_t$ , which is bounded below only. The interest rate series  $I_t^{DCD}$  and  $I_t^{FCD}$  were then constructed as weighted averages of the interest rates on the component classes of deposit in  $DCD$  and  $FCD$ .

Typically, it is real rather than nominal returns that are considered important to depositors' allocation decisions in portfolio theory; so empirical controls for expected inflation and depreciation are usually considered important. In the context of a narrow focus on deposit dollarisation, where domestic depositors ultimately value returns in DC, expected inflation is probably irrelevant, however - since it deteriorates the nominal returns to DC and FC deposits equally. Expected depreciation, by contrast, has an obvious differential effect, and is therefore important. In the absence of either survey data on devaluation expectations, or evidence from the forward market (neither of which are available in our two cases), it is conventional to capture depreciation expectations in empirical studies by using the series for actual depreciation, making the assumption either of data-based prediction (in which case the contemporaneous value of actual depreciation is used) or rational expectations (in which case the next period's value of actual depreciation is used) on the part depositors<sup>14</sup>.

With respect to the expected depreciation of the DC against the CBA anchor currency, the existence of CBAs in Lithuania and Estonia is therefore a handicap on this point. The actual depreciation of the LIT and EEK against their respective CBA anchor currencies was by definition zero over our samples - despite the fact that expectations of devaluation most likely were not. No empirical variable is therefore available to measure expectations of depreciation against the CBA anchor currency. The same difficulty does not arise in respect of the expected depreciation of the DC against non-CBA anchor foreign currencies, however. Since, as described

---

<sup>14</sup>The rationale for, and relative merits of, these two options are considered in more detail below.

in Appendix A, significant shares of  $FCD$  were denominated in USD - which was not the EEK CBA anchor currency throughout the sample; and not the LIT CBA anchor currency from February 2002 onwards - we constructed series for the EEK-USD and LIT-USD rates of depreciation ( $N_t$ ) for use in modelling.

In the course of the General-to-Specific modelling strategy employed to build the empirical models, a functional form in which  $DR_t$ ,  $I_t^{DCD}$ ,  $I_t^{FCD}$  were transformed into natural logarithms, whilst  $N_t$  was measured in levels was discovered to be both statistically and theoretically the most appropriate. In reporting the models below, we write  $\ln(DR_t)$ ,  $\ln(I_t^{DCD})$ , and  $\ln(I_t^{FCD})$  as  $dr_t$ ,  $i_t^{DCD}$ , and  $i_t^{FCD}$  to simplify presentation.

### 4.3.2 Time series characteristics

Figures 4-7 to 4-10 present graphs of the main time series characteristics of the levels and first differences of the variables  $dr_t$ ,  $i_t^{DCD}$ ,  $i_t^{FCD}$ , and  $N_t$  for both of our cases. Visual inspection of these graphs suggests that in both cases, all four series appear to be nonstationary in their levels, and stationary in their first differences. The graphs of the Partial Autocorrelation Functions (PACFs) of the series in levels, all of which suggest AR(1) series with a coefficient on the first lag close to unity, appear to corroborate this. The Autocorrelation Functions (ACFs) of the levels series suggest that in both cases all four are autoregressive processes of long memory. The ACFs and PACFs of the series in first differences are all consistent with their stationarity, and suggest that some of the series (both interest rate series in the case of Lithuania; the dollarisation ratio and FC deposit interest rate series in the case of Estonia) are not AR processes, whilst the rest are likely to be at least AR(1) processes, though with a coefficient on the first lag significantly lower than unity.

For both cases, the degrees of integration of the levels and first difference of the four series were investigated more formally using Augmented Dickey-Fuller (ADF) tests. The results of these tests, where the appropriate lag length for the ADF test statistic was chosen to minimise the Akaike Information Criterion, are reported in Tables 4.9 to 4.12. The null hypothesis of non-stationarity is strongly rejected for all eight series in first differences under all specifications (i.e. with and without combinations of seasonal dummies, a trend, and a constant term). With regard to the series in levels, the null cannot be rejected except (a) in the case of the deposit

interest rate series for Lithuania, when no deterministic terms are included, and (b) in the case of the domestic currency deposit interest rate series for Estonia, when a constant and trend, or a constant, trend, and seasonals are included. The former case can be ignored, however, since it is clear from scrutinising their graphs that these series do not have a zero mean; and the latter results are inconclusive, given the non-rejection of the null under minimally different lag-specifications.

On the basis of these ADF tests, therefore, and bearing in mind their low power, we can make the tentative inference that in both cases, all four series are  $I(1)$ <sup>15</sup>.

### 4.3.3 Treatment of expectations

The variable  $N_t$  is constructed in order to allow us to control for the possible effect on depositors' allocation decisions of their expectations regarding the depreciation of the DC against non-CBA anchor foreign currencies. The inclusion of expectational variables in dynamic empirical models has implications both the choice of the estimation procedure and the identification of estimates with theoretical parameters of interest. The two main empirical strategies available each involve different assumptions concerning the nature of agents' expectations formation process, and we review both briefly here.

The most popular theoretical characterisation of the expectations formation process is that embodied in the assumption of 'rational expectations'. The observed outcome of a variable at  $t + 1$  is supposed to be equal to the sum of agents' expectations of its value at time  $t$  and some zero-mean stochastic forecast error. The forecast errors imply serial correlation – in particular, the error has a moving average representation – with the result that the Ordinary Least Squares (OLS) estimator is not valid. The preferred solution to this problem is that pioneered in the applied macroeconomic context by Clarida, Gali, and Gertler (1998). This involves persevering with the use of the observed value of the expectational variable at time  $t + 1$  in place of its expected value at time  $t$ ; but deploying the Generalised Method of Moments (GMM), rather than the OLS, estimator. As many instruments for the expected variables as

---

<sup>15</sup>This inference should be attended by the qualification that it applies over the available sample - no doubt because it represents the period of transition from a centrally-planned to a market economy in both countries. It is highly unlikely that the statistical characterisation of these series as  $I(1)$  could continue indefinitely into the future.

are available are used, and any over-identifying restrictions required as a result are tested using Hansen's (1982) test.

The use of the GMM estimator circumvents the problem of serial correlation implied by the assumption of rational expectations. An additional assumption must also be made, however, concerning the actual process whereby rational expectations are formed – even if only implicitly – and the nature of this assumption will have implications for identification. In particular, if the rational expectations formation process is assumed to be parametric, then the parameters of the empirical model (however estimated), will not be identifiable with those of the structural theoretical model, but instead with convolutions of those structural parameters and the parameters of the rational expectations formation process.

Hendry (1995) argues that the assumption of rational expectations is in any case dubious on theoretical grounds; and that a more reasonable assumption is that agents form expectations according to simple rules of thumb. Such rules are most likely non-parametric and data-based; a general example for economic time series being, for example, that the rate of change is equal to zero -  $\Delta E_t[x_{t+1}] = 0$  - so that  $E_t[x_{t+1}] = x_t$ , and the current, rather than the next, period's observed value of the relevant variable is used in place of its expected value. With this option, there are problems neither of auto-correlation (since there are no forecast errors) nor of identification (since the expectations formation process is non-parametric).

How the choice between these two assumptions is made depends upon the general modelling strategy being implemented<sup>16</sup>. Researchers of the "structural modelling" school will make this choice on a priori theoretical grounds; those following the "LSE" approach, and hence deploying a General-to-Specific methodology, will regard the alternative assumptions as (at least potentially) testable.

Our General-to-Specific methodology – and the avoidance of the identification problem thereby – argues in favour of data-based prediction rather than rational expectations as the basis for modelling, and the use therefore of  $N_t$  to stand for the expected depreciation of the domestic currency against non-CBA anchor foreign currencies. Nevertheless, we present results using  $N_{t+1}$  as well in the interests of generality – though by deploying OLS estimation with

---

<sup>16</sup>See Favero (2004) for a recent analysis of contemporary macroeconometric modelling into three main schools, of which the "structural modelling" and the "LSE" approaches are the most widely known.

an appropriate correction to standard errors rather than a GMM estimator, since the latter is feasible only in  $I(0)$  applications.

#### 4.3.4 Modelling Strategy

We deploy a General-to-Specific modelling strategy as described, for example, in Hendry (1995). We proceed via the sequential marginalisation and reduction of an unrestricted vector auto-regression (UVAR) to a model of the form:

$$\begin{aligned}\Delta \mathbf{y}_t &= \alpha \beta' (1 : \mathbf{y}'_{t-1} : N_{t-1}) + \sum_{i=1}^{j-1} \delta_i \Delta \mathbf{y}_{t-i} + \Theta \mathbf{q}_t + \boldsymbol{\nu}_t \\ \boldsymbol{\nu}_t &\sim \mathbf{IN}_n(\mathbf{0}, \boldsymbol{\Omega})\end{aligned}\tag{4.1}$$

where  $\mathbf{y}'_t = \begin{bmatrix} dr_t & i_t^{DCD} & i_t^{FCD} \end{bmatrix}$ ,  $\mathbf{q}_t$  is a vector of deterministic terms, and  $j$  is the number of lags that are found to be needed to whiten the error terms of the UVAR. Since we have inferred that our empirical data series are all  $I(1)$ , we then test whether the product matrix  $\beta' (1 : \mathbf{y}_{t-1} : N_{t-1})$  in (4.1) contains (at least) one cointegrating combination using Johansen's (1988) procedure and Doornik's (2003) critical values; and ascertain whether this combination can be interpreted as an empirical equation describing the evolution of the deposit dollarisation ratio as a function of the nominal interest rate and expected depreciation series.

If such a cointegrating combination is found, then we can proceed to estimate a cointegrated vector auto-regression (CVAR), and to impose and test the validity of restrictions that either identify estimated coefficients with structural parameters of interest or correspond to hypotheses derived from theory. Of particular interest in this respect are two sets of restrictions.

The first of these relates to the exogeneity of the interest rate series. *A priori*, it would seem to be natural to postulate that whereas the interest rate on DC deposits may very well be endogenous, the interest rate on FC deposits will be an exogenous variable, since it is usual to assume that interest rates on FC assets must be equal to equivalent 'world' interest rates for financially small and open economies such as are Lithuania's and Estonia's. There are considerations, described in Section 4.2 above, which might give us pause on this front; but we can nevertheless test the hypothesis that  $i_t^{FCD}$  is weakly exogenous by imposing the exclusion restriction that the short run adjustment coefficient on  $i_t^{FCD}$  in the matrix  $\alpha$  in (4.1) is equal

to zero<sup>17</sup>. We therefore have:

$$H(1) : i_t^{FCD} \text{ is exogenous to the deposit dollarisation process.}$$

and the corresponding restriction:

$$R(1) : \hat{\alpha} = \begin{bmatrix} \hat{\alpha}^{11} & \hat{\alpha}^{12} & 0 \end{bmatrix}$$

The second hypothesis of particular interest is that the coefficients on  $i_t^{DCD}$  and  $i_t^{FCD}$  are of identical magnitude, but opposite sign. If valid, this hypothesis allows the interpretation of the estimated relationship as between  $dr$  and  $(i_t^{DCD} - i_t^{FCD})$  - and hence as evidence of deposit dollarisation being a function of the relative, rather than the absolute, interest rates on DC and FC deposits, as predicted by portfolio balance theory. This second hypothesis is therefore that:

$$H(2) : \text{The coefficients on } i_t^{DCD} \text{ and } i_t^{FCD} \text{ are of identical magnitude, but opposite sign}$$

to which the corresponding restriction is:

$$R(2) : \hat{\beta}' = \begin{bmatrix} \hat{\beta}^{11} & 1 & \pm \hat{\beta}^{31} & \mp \hat{\beta}^{31} & \hat{\beta}^{51} \end{bmatrix}$$

Finally, we investigate the dynamics of the short run relationships between the endogenous variables in  $\mathbf{y}_t$  by estimating (4.1) as a VECM. Of particular interest are the coefficients on the equilibrium correction term in the VECM, and the corresponding elements of  $\hat{\alpha}$  in the CVAR, which generate estimates of the speed of adjustment of the endogenous variables in  $\mathbf{y}_t$  to their long run equilibrium values.

---

<sup>17</sup>Note that maintaining the hypothesis that  $i_t^{FCD}$  is exogenous does not require that it be treated as exogenous in the UVAR. Although it is conventional practice to interpret the long run cointegrating combination(s) found in the matrix  $\beta$  of a CVAR as expressing one of its elements as a function of the others, this is done in a statistically arbitrary manner based on theoretical priors and without recourse to the notions of endogeneity and exogeneity, since as pointed out by, e.g. Johansen (2005), "[I]n the cointegrated VAR model, there are no exogenous and endogenous variables and there is no natural way in which we can pick out some variables and interpret the others as functions thereof." (Johansen (2005), p.98).

## 4.4 Results

In this section, we present the results of the modelling strategy set out in the previous section as applied to the cases of Lithuania and Estonia in turn. We discuss in detail the statistical analysis of the models constructed; but defer until Section 4.5 their theoretical interpretation. Given the General-to-Specific methodology followed in the building of the models, this separation of the two discussions is inevitably somewhat artificial - but it improves clarity.

Given that our principal interest is in estimates of the long run cointegrating combinations recoverable from CVARs, we relegate the detailed reporting of the VECM models for both Estonia and Lithuania to Appendix B. The estimates recovered from both versions of the models are however discussed when we come to discuss their interpretation in Section 4.5.

### 4.4.1 Lithuania

#### UVAR(8)

In the case of Lithuania, analysis started from a UVAR model with the following specification:

$$\begin{aligned}
 \mathbf{y}_t &= \sum_{i=1}^8 \pi_i \mathbf{y}_{t-i} + \Theta \mathbf{q}_t + \boldsymbol{\nu}_t & (4.2) \\
 \mathbf{y}_t &= \begin{bmatrix} dr_t \\ i_t^{DCD} \\ i_t^{FCD} \end{bmatrix} \\
 \mathbf{q}_t' &= \left[ 1 \quad N_s \quad CSeasonal1 \quad \dots \quad CSeasonal11 \right] \\
 \boldsymbol{\nu}_t &\sim \mathbf{IN}_n(\mathbf{0}, \boldsymbol{\Omega})
 \end{aligned}$$

where all variables are as defined in Section 4.3 above, and  $s$ , the time-subscript on  $N$ , was equal either to  $t$  on the assumption of data-based prediction by depositors, or  $t + 1$  on the assumption of rational expectations.

The minimum number of lags required to generate white noise residuals in estimating the UVAR model was found to be 8, and the Wald test statistic for the reduction from a lag length of 12 indicates no significant loss of likelihood ( $\chi^2(36) = 45.57$ ). The results of standard

diagnostic tests applied to both the residual vector  $\nu_t$  as a whole, and to its elements, the residuals of the individual equations of the UVAR(8), are reported in Table 4.13. There is no evidence of abnormality, serial correlation, heteroskedasticity, or autoregressive conditional heteroskedasticity (ARCH), except in the case of the residuals of the equation for  $i^{FCD}$ , where there is evidence of abnormality. This does not translate into abnormality of the system residuals considered jointly, however, and we proceed on that basis. The cross-correlations and standard deviations of the equation residuals are reported in Table 4.14. The cross-correlations are low, with the largest value being 0.38, between the two deposit interest rate series. The standard deviations, which can be interpreted as percentages given the log linear transformations of the variables, are below 10 per cent for  $dr$ , and marginally above 10 per cent for  $i^{DCD}$  and  $i^{FCD}$ , indicating that in statistical terms, the model is less relatively good at explaining the development of the deposit interest rate series. The correlations between the actual and fitted values of the elements of  $\mathbf{y}_t$ , given in Table 4.15, and are high, as one would expect given that the series appear to be non-stationary. Graphs illustrating the actual and fitted values of the variables in  $\mathbf{y}_t$ , and the values and distributions of the residuals of the individual equations of the UVAR(8) are included in Figure 4-11.

### Cointegration analysis

Dynamic analysis of the estimated UVAR(8) model appears to confirm the non-stationarity of the variables of  $\mathbf{y}_t$  indicated by the ADF tests, and to suggest the possibility of a single cointegrating relationship between them. Table 4.16 lists the eigenvalues of the matrix of long run equilibrium coefficients on  $\mathbf{y}_t$ , as calculated from the estimated coefficients of the UVAR(8) model. The single obviously non-zero eigenvalue suggests a rank of 1 for  $\alpha\beta'$  in (4.1).

Restricting the constant term and  $N_t$  to lie within the cointegrating space, this inference can be tested formally using Johansen's (1988) procedure. The results of both the Trace and Maximum Eigenvalue (ME) tests, using the critical values reported in Doornik (2003) are given in Table 4.4.

Table 4.4 indicates that on neither test can we reject the hypothesis that there is a single cointegrating vector. We therefore proceed by imposing a single cointegrating vector on the CVAR(8).

$H_0$	Trace	ME
$r < 1$	48.02*	26.78 * *
$r < 2$	21.24	13.90
$r < 3$	7.34	7.34

Note: \*\* indicates rejection at the 1% and \* at the 5% level of significance

**Table 4.4:** Lithuania: Tests for Cointegrating Vectors

### CVAR(8)

Having inferred the existence of a single cointegrating combination of the variables in  $\mathbf{y}_t$ , we proceed, as described in Section 4.3.4 above, to estimate (4.2) as a CVAR, maintaining the hypothesis of a single cointegrating vector on the matrix of long run coefficients  $\beta$ , and imposing cumulatively the restrictions  $R(1)$  and  $R(2)$  corresponding to the two hypotheses  $H(1)$  and  $H(2)$ . Table 4.5 presents the results of this procedure, under the assumptions both of data-based prediction (and therefore including  $N_t$  in  $\mathbf{y}_t$ ) and rational expectations (and therefore including  $N_{t+1}$  in  $\mathbf{y}_t$ ).

Estimates are reported in the four rows (1) to (4), with the five columns marked  $\beta$  reporting the estimated coefficients of the long run cointegrating relationship, and the three columns marked  $\alpha$  reporting the estimated coefficients of the short run adjustment matrix. Rows (1) and (2) report the results of estimating the CVAR(8) under  $R(1)$ , using the assumptions of DGP and RE respectively. Rows (3) and (4) then report the results under  $R(1)$  and  $R(2)$  together, again assuming DGP and RE respectively. The final column of Table 4.5, headed LR  $\chi^2$ , reports the value taken by the likelihood ratio test of the restrictions imposed.

		<i>Cons</i>	$i_t^{DCD}$	$\beta$ $i_t^{FCD}$	$N_t$	$N_{t+1}$	$dr_t$	$\alpha$ $i_t^{DCD}$	$i_t^{FCD}$	LR $\chi^2$
$R(1)$	DBP (1)	0.47 (0.08)	0.83 (0.17)	-0.87 (0.20)	-4.16 (0.84)		-0.08 (0.03)	-0.18 (0.06)	0.00 -	1.90 [0.17]
	RE (2)	0.46 (0.07)	0.84 (0.17)	-0.86 (0.19)		-4.37 (0.82)	-0.08 (0.03)	-0.18 (0.06)	0.00 -	0.94 [0.33]
$R(2)$	DBP (3)	0.44 (0.05)	0.81 -	-0.81 (0.17)	-4.45 (0.80)		-0.07 (0.03)	-0.17 (0.06)	0.00 -	2.46 [0.29]
	RE (4)	0.44 (0.05)	0.83 -	-0.83 (0.16)		-4.58 (0.77)	-0.08 (0.02)	-0.17 (0.06)	0.00 -	1.17 [0.56]

**Table 4.5:** Lithuania: CVAR(8), under R(1) and R(2)

The LR tests of the restrictions imposed reported in the last row of Table 4.5 indicate that regardless of our assumption concerning the process whereby depositors form expectations, neither H(1) nor H(2) can be rejected. Given the greater precision of the estimates when both  $R(1)$  and  $R(2)$  are imposed, and the theoretical attractions of these restrictions which will be discussed in Section 4.5 below, we therefore proceed on the basis that specification (3) in Table 4.5 is the preferred model<sup>18</sup>. The stability of the coefficient estimates over the sample is illustrated by Figure 4-12, which graphs recursive estimates of the elements of  $\hat{\beta}$  in specification (3)<sup>19</sup>.

#### 4.4.2 Estonia

##### UVAR(6)

Modelling of the Estonia case started from a UVAR model with the following specification:

$$\begin{aligned} \mathbf{y}_t &= \sum_{i=1}^6 \pi_i \mathbf{y}_{t-i} + \gamma N_s + \vartheta q + \boldsymbol{\nu}_t & (4.3) \\ \mathbf{y}_t &= \begin{bmatrix} dr_t \\ i_t^{DCD} \\ i_t^{FCD} \end{bmatrix} \\ q &= 1 \\ \boldsymbol{\nu}_t &\sim \mathbf{IN}_n(\mathbf{0}, \boldsymbol{\Omega}) \end{aligned}$$

where all variables are as defined in Section 4.3 above, and  $s$ , the time-subscript on  $N$ , was equal either to  $t$  on the assumption of data-based prediction by depositors, or  $t + 1$  on the assumption of rational expectations..

The minimum number of lags required to generate white noise residuals in estimating the UVAR model was found to be 6. The results of standard diagnostic tests applied to both

---

<sup>18</sup> Additional restrictions, excluding other elements of both the short and long run matrices using specification (3) were all rejected, as reported in Table 4.17.

<sup>19</sup> Evidently, a shortcoming with Figure 4-12 as a demonstration of the stability of these estimates is that it omits the early part of the sample. Unfortunately, neither repeating this recursive estimation in reverse, nor alternative and more formal tests of parameter stability such as Hansen's (1992) test or other tests of the CUSUM and CUSUMSQ family are available: the former because of the dynamic structure of the VAR, and the latter because of the presence of non-stationary variables (see Hansen (1992)).

the residual vector  $\boldsymbol{\nu}_t$  as a whole, and to its elements, the residuals of the individual equations of the UVAR(6), are reported in Table 4.21. There is no evidence of abnormality, serial correlation, heteroskedasticity, or autoregressive conditional heteroskedasticity (ARCH). The cross-correlations and standard deviations of the equation residuals are reported in Table 4.22. The cross-correlations are low, with the largest value being 0.21, between the two deposit interest rate series. The standard deviations, which can be interpreted as percentages given the log linear transformations of the variables, are below 10 per cent for  $dr$  and  $i^{FCD}$ , and marginally above 10 per cent for  $i^{DCD}$ . The correlations between the actual and fitted values of the elements of  $\mathbf{y}_t$ , given in Table 4.23. Graphs illustrating the actual and fitted values of the variables in  $\mathbf{y}_t$ , and the values and distributions of the residuals of the individual equations of the UVAR(6) are included in Figure 4-15.

### Cointegration analysis

As was the case with the Lithuanian model, the evidence points to a single cointegrating vector. Table 4.24 lists the eigenvalues of the matrix of long run equilibrium coefficients on  $\mathbf{y}_t$ , as calculated from the estimated coefficients of the UVAR(6) model: one would seem to be definitely non-zero, and another possibly so.

Once again, we use Johansen's (1988) Trace and Maximum Eigenvalue (ME) tests to test this inference formally against the critical values reported in Doornik (2003). The results are given in Table 4.6.

$\mathbf{H}_0$	Trace	ME
$r < 1$	54.81 **	37.14 **
$r < 2$	17.67	14.58
$r < 3$	3.09	3.09

Note: \*\* indicates rejection at the 1% and \* at the 5% level of significance

**Table 4.6:** Estonia: Tests for Cointegrating Vectors

As for Lithuania, Table 4.6 indicates that on neither test can we reject the hypothesis that there is a single cointegrating vector. As in the case of Lithuania, we therefore proceed by imposing a single cointegrating vector on the CVAR(6).

## CVAR(6)

Having inferred the existence of a single cointegrating combination of the variables in  $\mathbf{y}_t$ , we proceed, as described in Section 4.3.4 above, to estimate (4.3) as a CVAR, maintaining the hypothesis of a single cointegrating vector on the matrix of long run coefficients  $\beta$ , and imposing sequentially the restrictions  $R(1)$  and  $R(2)$  corresponding to the two hypotheses  $H(1)$  and  $H(2)$ . Table 4.7 presents the results of this procedure, under the assumptions both of data-based prediction (and therefore including  $N_t$  in  $\mathbf{y}_t$ ) and rational expectations (and therefore including  $N_{t+1}$  in  $\mathbf{y}_t$ , and correcting the estimated standard errors accordingly). The presentation of the estimates is identical to that used for Lithuania above.

			$\beta$					$\alpha$			
			<i>Cons</i>	$i_t^{DCD}$	$i_t^{FCD}$	$N_t$	$N_{t+1}$	$dr_t$	$i_t^{DCD}$	$i_t^{FCD}$	LR $\chi^2$
$R(1)$	DBP	(1)	1.26	0.95	-1.22	0.31		-0.29	-0.29	0.00	7.02
			(0.12)	(0.13)	(0.18)	(0.39)		(0.07)	(0.09)	-	[0.01] **
$R(1)$	RE	(2)	1.25	0.94	-1.19		0.23	-0.29	-0.30	0.00	6.85
			(0.12)	(0.13)	(0.18)		(0.38)	(0.07)	(0.09)	-	[0.01] **
$R(2)$	DBP	(3)	1.00	0.77	-0.77	-0.71		-0.31	-0.34	-0.23	1.20
			(0.03)	-	(0.11)	0.25		(0.08)	(0.09)	(0.06)	[0.27]
$R(2)$	RE	(4)	0.98	0.79	-0.79		-0.58	-0.32	-0.33	-0.21	2.55
			(0.03)	-	(0.11)		(0.26)	(0.08)	(0.10)	(0.07)	[0.11]

**Table 4.7:** Estonia: CVAR(6) under  $R(1)$  and  $R(2)$

Note: \*\* indicates rejection at the 1% level of significance

Unlike in the models of Lithuania, the LR tests reported in the last row of Table 4.7 indicate that  $R(1)$  is rejected at the 1% level of significance: it is not legitimate, in other words, to exclude  $i_t^{FCD}$  from the matrix of short run adjustment coefficients, and the corresponding  $I(0)$  model must be a system of three, rather than two, equations<sup>20</sup>. Specifications (3) and (4) therefore involve the imposition of one restriction ( $R(2)$ ) only, therefore, rather than the two restrictions that were imposed on the equivalent model for Lithuania.  $R(2)$  cannot be rejected under the assumption of either DBP or RE. As in the case of Lithuania, we therefore proceed

<sup>20</sup>We note that the long run coefficients on the expected depreciation term can be excluded from the  $\beta$  matrix in both specifications (1) and (2). This has no effect on the rejection of  $R(1)$  however (and little effect on the estimates of the other elements of  $\beta$ ); so we report the model with  $N_t$  and  $N_{t+1}$  included in  $\beta$  for the sake of consistency.

using specification (3) as our preferred model<sup>21</sup>. Figure 4-16 illustrates graphically recursive estimates of the elements of  $\hat{\beta}$  under this specification.

## 4.5 Interpretation

Using the preferred specifications of the two models presented in the previous section, we have found that, in the long run, the following robust statistical relationships hold in Lithuania and Estonia:

$$\text{Lithuania: } dr_t = -0.44 - 0.81i_t^{DCD} + 0.81i_t^{FCD} + 4.45N_t \quad (4.4)$$

$$\text{Estonia: } dr_t = -1.00 - 0.77i_t^{DCD} + 0.77i_t^{FCD} + 0.71N_t \quad (4.5)$$

In addition, we have constructed congruent statistical models of the short run dynamics of the deposit dollarisation ratio and deposit interest rates around these long run equilibrium relationships, as reported in Appendix B.

In this section, we ask whether a sensible economic interpretation can be given to these results within the framework of portfolio balance theory. We begin, in Section 4.5.1 with the theoretical interpretation of the most basic features of our models: the single long run cointegrating relationship identified in both cases, and the exogeneity status of the variables used in modelling. We then proceed, in Section 4.5.2, to the interpretation of the long run relationships (4.4) and (4.5) themselves. Finally, in Section 4.5.3, we turn to a discussion of the numerical estimates of the parameters of (4.4) and (4.5) and their associated VECMs.

### 4.5.1 Vector versus single equation methods, and endogeneity of variables

In the cases of both Lithuania and Estonia, both the Johansen (1988) test and the subsequent estimation of the CVAR and VECM indicate the existence of a single, long run, cointegrating combination of the variables  $dr_t$ ,  $i_t^{DCD}$ ,  $i_t^{FCD}$ ,  $N_t$ , and a constant term; and these combinations are *prima facie* amenable to interpretation as the equilibrium solutions to portfolio balance

---

<sup>21</sup>As in the model for Lithuania, additional restrictions excluding other elements of the  $\alpha$  and/or  $\beta$  matrices from specification (3) were all rejected, as reported in Table 4.25.

problems, with  $dr_t$  interpreted as the 'dependent variable' within the cointegrating vector<sup>22</sup>. The finding that the coefficients not only on  $dr_t$ , but on also on  $i_t^{DCD}$ , and in the case of Estonia on  $i_t^{FCD}$  also, cannot be excluded from the matrix  $\alpha$  of short run adjustment coefficients indicates however that estimation using single equation methods would not be legitimate - a system of two I(0) equations, in  $\Delta dr_t$  and  $\Delta i_t^{DCD}$  in the case of Lithuania, and of three equations, in  $\Delta dr_t$ ,  $\Delta i_t^{DCD}$ , and  $\Delta i_t^{FCD}$  in the case of Estonia, is necessary.

The endogeneity of the interest rate on DC deposits in both models is an intuitive result. A more challenging aspect of the models is the finding that in the case of Lithuania, the restriction  $R(1)$ , corresponding to the hypothesis that the interest rate on FC deposits is exogenous, is accepted; whereas in Estonia, it is not. The former result is no doubt the less surprising - since as noted above, in the absence of capital controls or other significant frictions, we might *a priori* expect there to be strong pressures on interest rates on domestic FC deposits to converge with those on FC deposits of identical maturities abroad. However, it is equally clear from the account of the evolution of the interest rate on FC deposits given in Section 4.2 above that it took a considerable time for this convergence to occur in both countries; and that this was at least in part due to the relatively gradual decline of country risk over the transition. The theoretical case for the exogeneity of  $i_t^{FCD}$ , though it must remain strong over the long term, and once the transition is complete, is therefore not so clear cut for the transitional samples over which our models have been estimated.

Even if it is conceded, however, that in the circumstances of Lithuania and Estonia the *a priori* case for the exogeneity of  $i_t^{FCD}$  does not necessarily apply, it is still necessary to explain the difference between the models for the two countries - with  $i_t^{FCD}$  found to be exogenous in Lithuania, but endogenous in Estonia. Reverting our attention to the significant difference between the two countries in the currency composition of the overall stock of FC deposits discussed in Appendix A furnishes a possible explanation. We infer there that whereas practically no FC deposits were exposed to nominal exchange rate risk up until late in the sample (after the re-peg in February, 2002) in the Lithuanian case, a very large share of FC deposits were exposed to fluctuations in the nominal exchange rate throughout the sample in the case of Estonia. Since endogeneity implies the deposit dollarisation ratio's being an important factor in

---

<sup>22</sup>Though see footnote 17 and Johansen (2005) for provisos on this terminology.

banks' decisions to set interest rates, then we would much less expect to find  $i_t^{FCD}$  endogenous in Lithuania than in Estonia - since for over 80% of the period covered by our sample deposit dollarisation in Lithuania carried relatively much less risk to banks' balance sheets in Lithuania than it did in Estonia.

#### 4.5.2 Empirical Application of Portfolio Balance Theory to Deposit Dollarisation

As discussed by Schmukler and Servén (2002), the simplest level of portfolio theoretic analysis at which (4.4) and (4.5) can be interpreted is in terms of the basic currency premium,  $cp_t$ , where  $cp_t = (I^{DCD} - I^{FCD})_t$ . At this level of analysis, our ambitions are limited to giving an interpretation to (4.4) and (4.5) that is consistent with the core result of the portfolio balance theory of deposit dollarisation that the deposit dollarisation ratio can be explained primarily by depositors' reactions to the relative risks and returns associated with DC and FC deposits rather than by real, structural, or institutional factors. Since there are no empirical variables designed explicitly to control for risk terms in the econometric models reported above, the basic currency premium must be assumed, for even this level of analysis to be possible, to be a *portmanteau* statistic, capturing both the expected risks and the returns to holding DC and FC deposits in one. Evidently, this limits significantly the extent to which the models can be identified with theoretical results that distinguish between the effects of these two concepts.

Two features of the models are important for an interpretation at this level: first, that in the preferred specification, both  $dr_t$  and the interest rate series  $i_t^{DCD}$  and  $i_t^{FCD}$  are measured in natural logarithms, and that the *moduli* of the estimated long run coefficients on the interest rate terms are less than unity; and second, that the restriction  $R(2)$ , that the long run coefficients on  $i_t^{DCD}$  and  $i_t^{FCD}$  are of identical magnitude and opposite sign, is accepted in both models.

The first of these two features implies a non-linear relationship between the deposit dollarisation ratio and the interest rates on DC and FC deposits in which an increase (decrease) in the interest rate offered on DC (FC) deposits is associated, *ceteris paribus*, with a fall in the deposit dollarisation ratio - though at a declining rate, the higher (lower) the absolute level of that offered interest rate. The higher the nominal interest rate offered on deposits denominated in a given currency, in other words, the greater the incentive for depositors to hold them;

though the incentive generated by a given increase becomes smaller as the absolute level of that interest rate rises. This implication of the models has an intuitive interpretation in terms of portfolio balance theory if we make the assumption that higher absolute levels of nominal returns are associated with higher degrees of risk attached to those returns. This is a plausible assumption - especially given that within the samples for Lithuania and Estonia, high absolute levels of interest rates occur in the earlier parts of the samples, when risks of all sorts could reasonably be assumed to be higher. If, then, this assumption is maintained, then the positive but decreasing marginal incentive to hold deposits of a given currency as the interest rate on those deposits rises is consistent with the basic postulate of portfolio balance theory that asset holding is motivated positively by expected returns, but negatively by expected risk.

It is also a basic result of portfolio balance theory, however, that it is the relative, rather than the absolute, returns and risk to holding different assets which determines the composition of the optimal portfolio chosen by investors. Interpreting the interest rate series as individual regressors has no particular bearing on the validity or otherwise of this postulate; but because the restriction  $R(2)$ , that the long run coefficients on  $i_t^{DCD}$  and  $i_t^{FCD}$  are of identical magnitude and opposite sign, is not rejected in either model, the additive combination of the two separate series in logs can also be interpreted as the ratio of the two series in levels. This interpretation of (4.4) and (4.5) as:

$$\text{Lithuania: } dr_t = -0.44 - 0.81 \ln \left( \frac{I_t^{DCD}}{I_t^{FCD}} \right) + 4.45N_t \quad (4.6)$$

$$\text{Estonia: } dr_t = -1.00 - 0.77 \ln \left( \frac{I_t^{DCD}}{I_t^{FCD}} \right) + 0.71N_t \quad (4.7)$$

recasts the two separate interest rate terms in logs as a single interaction term in levels. This alternative interpretation of (4.4) and (4.5) implies that it is indeed the relative, rather than the absolute, levels of the interest rate series that matters for the determination of the deposit dollarisation ratio - though in a non-linear manner approximated by the divisive interaction term, rather than the linear manner more usually assumed (and captured by the difference). The higher the level of the interest rate on FC (DC) deposits relative to that on DC (FC) deposits, the smaller is the fall in the deposit dollarisation ratio with which a given marginal increase (decrease) in that rate is associated.

It is not possible to discriminate empirically between these two interpretations of (4.4) and (4.5) in levels - as log-linear in their individual interest rate terms, or as linear in a divisive interaction term - given the particular evolution of the interest rate series over the samples in our two cases. So long as the incentives to allocate deposits to a particular currency are interpreted as diminishing as a result of greater associated riskiness, however - be that greater riskiness associated with a higher absolute level of nominal returns to the currency in question as in the first interpretation or to its substitute as in the second - then both interpretations are consistent with a portfolio theoretic explanation of deposit dollarisation.

It is natural to ask whether we can use (4.4) and (4.5) to analyse the motives driving dollarisation in more detail than this; and in particular, whether we can make any inference concerning the relative contributions of hedging and speculative demand to deposit dollarisation, given the significance of this distinction for financial sector regulatory policy, as discussed in Chapter 3 of this thesis. The basis for more detailed interpretation of this sort would be a more fully and explicitly specified portfolio balance model of the type reviewed in Chapter 3.4.3. For illustrative purposes, we set out a simple version of such a model, adapted to the cases under consideration in this chapter by allowing for two currencies of denomination (domestic and foreign), but abstracting from the location decision (all deposits are assumed to be held in the domestic banking system), in Appendix C. The solution for the equilibrium deposit dollarisation ratio in this model is:

$$x_t^* = \Phi_t - \zeta \Psi_t(i_t - i_t^* - E_t[e_{t+1}]) \quad (4.8)$$

where  $x_t^*$  is the deposit dollarisation ratio;  $\Phi_t$  and  $\Psi_t$  are functions of the variances and covariances of the real returns to holding DC and FC assets; and  $E_t[e_{t+1}]$  is the expected depreciation of the domestic against the foreign currency. As discussed in Chapter 3.4.3, an advantage of this specification is that the first term on the right hand side of (4.8),  $\Phi_t$ , can be interpreted directly as the gross proportion of deposit dollarisation determined by hedging demand, whilst the second term can be interpreted as gross dollarisation determined by speculative demand.

Empirical analysis of the motives driving dollarisation at the more detailed theoretical level represented by a solution such as (4.8) would proceed in either of two ways. The first, and preferable, method would be to collect observations on  $E_t[e_{t+1}]$ ,  $\Phi_t$ , and  $\Psi_t$ , and then to build

a new set of econometric models including these time series. These series would ideally be constructed using data from the forward foreign exchange market; or failing this by extrapolating from observations of the actual (non-expected) depreciation rate as described above in the case of  $N_t$ . The alternative, and less direct, method would be to retain the empirical models (4.4) and (4.5), and to interpret these in light of (4.8) by imposing appropriate identifying restrictions based on assumptions concerning the evolution of  $E_t[e_{t+1}]$ ,  $\Phi_t$ , and  $\Psi_t$ . This method has the advantage of not requiring any new data; but the disadvantage that the assumptions are likely to be *ad hoc*, and the precision of the inference permitted by the identifying restrictions potentially limited.

Unfortunately, neither method proves satisfactory in the cases at hand. With regard to the direct estimation method, the presence of CBAs makes the extrapolation of  $E_t[e_{t+1}]$ ,  $\Phi_t$ , and  $\Psi_t$  from observed variation in the actual depreciation rate impossible, and the rudimentary state of the financial markets in Lithuania and Estonia mean that forward foreign exchange market data is not available. With regard to the use of identifying restrictions, the assumptions required to identify the empirical estimates (4.4) and (4.5) with structural parameters of interest are inherently implausible (for example, that  $\Phi_t$  and  $\Psi_t$  are constant over the sample)<sup>23</sup>.

It is not feasible, therefore, to extend our economic interpretation of (4.4) and (4.5) further than is done above; and although this means that we cannot, for example, discriminate between hedging and speculative motives in the determination of the deposit dollarisation ratio, we are nevertheless able to interpret (4.4) and (4.5) as evidence that the dollarisation process is indeed driven by agents' attempts to balance risk and return in their portfolios.

### 4.5.3 Interpretation of numerical estimates

#### Parameter estimates - long run

**Constant term** Since the dependent variable in (4.4) and (4.5) is measured in logs, taking the exponent of the estimated coefficients on the constant term transforms them into levels commensurable with the deposit dollarisation ratio  $DR_t$ ; and taking exponents after first ap-

---

<sup>23</sup>It might also be objected to either of the two methods that (a) the necessity of assuming a particular form for the utility function of depositors, and (b) the narrow conception of risk as the expected second moment of the relevant stochastic variable, embody further unbelievable restrictions which render interpretation at this level of a fully specified theoretical model unwarranted.

appropriately adjusting the estimates transforms them into levels commensurable with the more intuitive ratio  $DDR_t$ <sup>24</sup>. Both adjustments, for both Lithuania and Estonia, are presented in Table 4.8.

	Dependent variable	
	$DR_t$	$DDR_t$
Lithuania	0.64	0.39
Estonia	0.37	0.27

**Table 4.8:** Adjusted estimates of the coefficient on the constant term

Clearly, the magnitude of these estimates are economically significant; and if it were possible to interpret (4.4) and (4.5) in terms of the solution (4.8), they would be liable to a convenient interpretation in terms of the minimum variance portfolio - that part of deposit dollarisation driven by hedging demand alone. Since such an interpretation would, however, rely on implausible identifying restrictions, as discussed above, the valid interpretation of these coefficients is the more pedestrian one conventionally given to coefficients on constant terms - that they capture the effect of omitted variables and other (relatively minor) mis-specifications.

**Deposit interest rate terms** The most striking property of the numerical values of the estimated coefficient on the deposit interest rate terms - apart from their *moduli* being less than unity, and hence their implying diminishing, rather than constant or increasing, incentives to allocate deposits in a given currency as the interest rate on deposits in that currency rises - is that they are approximately identical in the models for both countries ( $\pm 0.81$  in the case of Lithuania and  $\pm 0.77$  in the Estonia). Our ability to analyse the implications of this interesting fact is, however, limited. As discussed, it is not possible to identify separately the contributions of the degree of depositors' aversion to risk on the one hand, and the actual riskiness of the returns to each class of deposits on the other. The estimated coefficients therefore measure only the gross elasticity of deposit dollarisation with respect to the interest rate in question - with the most that can be said about the relative contributions of risk and return factors to this elasticity derived simply from the non-linear nature of the relationship analysed above. In

---

<sup>24</sup>The appropriate adjustment, evaluated at the sample means, is  $\widehat{\beta}^{11} - \ln(x)$  where  $x = \frac{1}{T} \sum_{t=1}^T \frac{(DCD_t + FCD_t)}{(DCDt)}$ , since  $dr_t = \ln\left(\frac{(DCD_t + FCD_t)}{(DCDt)}\right) + \ln(FDR_t)$ .

this context, it is important to note that the elasticities of the more intuitive measure of the deposit dollarisation ratio,  $DDR_t$ , with respect to the deposit interest rate series implied by the estimated coefficients are significantly lower than the values of the coefficients estimated on  $DR_t$ . Using the analogous adjustment to that deployed above, we can calculate approximate equivalent values for these elasticities of 0.30 and 0.47 for Lithuania and Estonia respectively. These lower elasticities imply that the degree to which associated risk is perceived to increase with the offered nominal interest rate in both countries is quite high (since the concavity of the implied relationship in levels is correspondingly pronounced).

**Expected depreciation against non-CBA anchor currencies** The inclusion of the variable  $N_t$  in the models above is intended to control for the expected depreciation of DC against non-CBA anchor currencies - in practice, the US Dollar throughout the sample in the case of Estonia, and the US Dollar after February, 2002 in the case of Lithuania, as discussed in Appendix A. The numerical values of the estimated coefficients on this variable in the two models are of interest because they are so different - that in the model for Lithuania being some six times larger than that in the model for Estonia.

Unlike the interest rate series, which are in logs, the expected depreciation series  $N_t$  is in levels in both models. Its estimated coefficients therefore represent constant semi-elasticities, rather than constant elasticities. Interpretation is made easier if we evaluate the elasticities that these estimates imply at the mean values of the series<sup>25</sup>. This gives us estimates of the long run elasticities of deposit dollarisation with respect to depositors' expectations of depreciation against non-CBA anchor currencies of approximately 12% and 7% for Lithuania and Estonia respectively<sup>26</sup>. The difference in implied elasticities between the two models is therefore much less than the difference between the estimated semi-elasticities - due to the greater variability of actual depreciation over the sample in Estonia than in Lithuania. A substantial difference remains, of course, and even though comparisons of numerical estimates between the two models are, as mentioned above, perhaps invidious, given the short sample lengths and consequent imprecision of the estimates, an interpretation of why this is so is asked

---

<sup>25</sup>Note that we use the mean values of the *moduli* of the series.

<sup>26</sup>The calculated values are identical (to 0 decimal places) in both cases whether the mean of  $|N_t|$  (DBP) or of  $|N_{t+1}|$  (RE) is used.

for.

The simplest such interpretation reverts our attention to the fact that, as described in Appendix A, the proportion of FC deposits exposed to the expected exchange rate changes captured by  $N_t$  is substantially different in the two countries. In Estonia, we deduced that perhaps close to half of all FC deposits might have been so exposed over the sample; whereas in Lithuania, nearly all were suddenly exposed following the re-peg of February, 2002 - but nearly none before that time. Since, as described in Appendix A, we do not know with certainty the stocks of CBA anchor and non-CBA anchor FC deposits for each time period, it is not possible to interpret the different elasticities recovered with any precision in light of this. The general point that different elasticities would be recovered depending on the proportion of FC deposits affected by variation in  $N_t$  is, however, self-evident: and the fact that the elasticity derived for Lithuania is significantly higher than that for Estonia conforms with what we can say about the order of magnitude, if not the actual levels, of this proportion.

#### Parameter estimates - short run

Although short run dynamics, of the sort estimated both in the  $\alpha$  matrix of the CVARs reported above and in the VECMs reported in Appendix B, are not considered explicitly at any of the three analytical levels set out above, economic interpretation of these estimates is nevertheless possible, and we discuss two aspects.

First, Table 4.9 presents the estimated coefficients on the equilibrium correction terms in the two short run systems, which represent the speed of adjustment of the relevant variables to their long run equilibrium values discussed in the previous section.

Equation	Lithuania	Estonia
$\Delta dr_t$	0.08	0.34
$\Delta i_t^{DCD}$	0.20	0.41
$\Delta i_t^{FCD}$	—	0.16

**Table 4.9:** Estimated Coefficients on ECM terms

The principal point of interest in these estimates is the slow speed of adjustment of the deposit dollarisation ratio in the case of Lithuania. Whereas in Estonia, the estimates imply that the half life of shocks to the deposit dollarisation ratio is approximately three months, in

Lithuania the same statistic is estimated at approximately a year. We should not necessarily expect the speed of adjustment of depositors to changes in their desired equilibrium deposit dollarisation ratio to be quick however. Despite the fact that these are financial assets, and hence (in theory) highly liquid, it may be that Lithuanian depositors prefer to bide the time in anticipation of future readjustments of their equilibrium deposit dollarisation ratio rather than to arbitrage currencies on a running basis.

Second, we can usefully give an economic interpretation to the short run coefficients on the dummy variables in the Lithuanian ((4.11) and (4.12)), and Estonian ((4.15) to (4.17)) systems respectively. In the case of Lithuania, we can infer from the estimates of the coefficients on *DummyEL* and *DummyTB* that the October, 1996 general election had a significant, positive impact on the short term development of the financial dollarisation ratio, but no significant impact on the weighted average interest rate paid on domestic currency deposits; whereas the opening of the Treasury Bill market had a significantly negative short term effect on the weighted average interest rate paid on domestic currency deposits; and a small positive effect on the financial dollarisation ratio. Both sets of coefficients are consistent with the hypothesis that these events, the one political, and the other financial, stimulated short term uncertainty both on the part of depositors and banks causing the former to increase the share of their portfolios held in foreign currency, and the latter to shore up their margins against the possibility of crisis or competition. In the case of the Estonian short run system, the estimated coefficients on *Dummy0112* indicate that the resignation of the government in December, 2001 had a small negative effect on the interest rate offered by banks on domestic currency deposits. Once again, this is consistent with the hypothesis that banks wanting to hedge against the uncertainty generated by the political crisis increased margins as a form of self-insurance.

## 4.6 Conclusions

In this paper we have reported the results of building two new, dynamic econometric models of the deposit dollarisation process in Lithuania and Estonia - two transition economies with CBAs. In addition to recovering numerical estimates of the parameters of the deposit dollarisation process in Lithuania and Estonia during their transitions, we have been able to draw two main

conclusions.

First, and most generally, our econometric models corroborate the Sahay and Vegh (1996) hypothesis that the presence of CBAs renders investigation of the deposit dollarisation process using portfolio balance theory analytically tractable. By deploying a General-to-Specific modeling strategy, we were able to construct dynamic, cointegrating, multiple equation models of the deposit dollarisation process as a function of a small vector of variables which may be interpreted as capturing the expected risks and returns associated with DC and FC deposits; and on the basis of a battery of diagnostic tests to conclude, allowing for the short samples available for estimation, that these models are congruent in the sense of Hendry (1995).

Second, we have however illustrated that the level of analytical depth that can be achieved using time series econometric models such as those we have constructed is constrained by data availability on the one hand, and by the plausibility of potential identifying restrictions on the other. More detailed analysis of the portfolio balance motives involved in deposit dollarisation - and in particular, analysis of the relative contributions of hedging and speculative demand - require either the construction of new empirical models incorporating variables that measure expected depreciation and its volatility, or the reinterpretation of the existing models after the imposition of strong identifying restrictions. The former approach is impossible in the case of countries with CBAs, and without forward foreign exchange market data. The latter requires introducing assumptions that are inherently implausible, and which therefore make Pyrrhic victories of the inferential gains that might in theory result. We therefore concluded that the additional analytical tractability introduced by the presence of a CBA is paid for in reduced empirical tractability; with the result that the time series econometric analysis of deposit dollarisation at the level of the fully specified portfolio balance models discussed in Chapter 3 of this thesis is not feasible in the cases of Lithuania and Estonia.

## Appendix A: Construction of Data Series

**Deposit dollarisation ratios** Three issues arise in the construction of empirical deposit dollarisation ratio series. The first issue concerns the formula used to calculate the dollarisation ratios. The most intuitive measure of the deposit dollarisation ratio is no doubt that illustrated in Figures 4-1 and 4-2:

$$DDR_t = \frac{FCD_t}{(FCD_t + DCD_t)}$$

where  $FCD_t$  denotes deposits denominated in foreign, and  $DCD_t$  deposits denominated in domestic currency at time  $t$ . For the purposes of econometric analysis, however, this definition of the deposit dollarisation ratio presents potential problems for estimation by OLS since it is bounded between 0 and 1. An alternative definition that avoids these problems is the following:

$$DR_t = \frac{FCD_t}{DCD_t}$$

since  $DR_t$  has only a lower bound at zero.

A second issue concerns what categories of deposits should be included in  $FCD_t$  and  $DCD_t$ , and has two parts. The first part is whether FC deposits denominated in foreign currencies other than the CBA anchor currency should be included in  $FCD_t$ . Evidently, including only FC deposits denominated in the CBA anchor currency would make for the most convenient measure of the deposit dollarisation ratio. Unfortunately, neither the Bank of Lithuania nor the Bank of Estonia maintain series on the outstanding stock of FC deposits disaggregated by currencies of denomination - so that our empirical measure of  $FCD_t$  by necessity includes FC deposits denominated in both the CBA anchor currency and in other foreign currencies. The second part is whether both demand and term deposits should be included in  $FCD_t$  and  $DCD_t$ . As discussed in Chapter 3, the usefulness of empirical models of dollarisation based on portfolio balance theory is predicated on the hypothesis that store of value (financial), rather than transactions (real), motives dominate the demand for and supply of all financial assets. This hypothesis is most naturally maintained when the deposit dollarisation ratio used includes only those categories of deposits which are not customarily used as means of payment. As a

rule of thumb, this would include term deposits, but exclude demand deposits<sup>27</sup>. As it happens, this dilemma is also settled by data availability in our two cases. In the case of Estonia, we were able to use series of foreign ( $FCTD_t$ ) and domestic ( $DCTD_t$ ) currency time and savings deposits only in calculating the deposit dollarisation ratio ( $DR_t$ ). For our analysis of Lithuania, where separate series on demand, time, and savings deposits denominated in foreign currency are not reported, we instead calculated the deposit dollarisation ratio ( $DR_t$ ) as the ratio of total FC deposits (demand plus time and savings -  $FCDD_t + FCTSD_t$ ) to total domestic currency deposits ( $DCDD_t + DCTSD_t$ ).

A final issue concerns the effects of the Euro-In process on the volume of FC deposits, and thereby the deposit dollarisation ratio, in the two cases<sup>28</sup>. Whilst in Lithuania, this process did not cause any noticeable change in the volume of FC deposits in the banking system, in the case of Estonia it resulted in an obvious and temporary spike, as Euro-In currencies were first deposited in the banking system, and then withdrawn again in kroon or foreign currency cash or demand deposits. We chose to interpolate values of the foreign currency time and savings deposits over the five month period July to November, 2002 as a result. A linear interpolation was applied, as being the most transparent, resulting in the series  $EXTRFCTSD_t$  to replace  $FCTSD_t$ .

All data series on litas, kroon and foreign currency denominated demand, time, and savings deposits were taken from the websites or printed publications of the Bank of Lithuania ([www.lbank.lt](http://www.lbank.lt)) and the Bank of Estonia ([www.eestipank.info](http://www.eestipank.info)). For Lithuania, 132 monthly observations could thereby be generated on  $DR_t$ , covering the period January, 1993 to December, 2003; and for Estonia, 81 monthly observations on  $DR_t$ , covering the period April, 1997 to December, 2003. The series describing the deposit dollarisation ratio in Lithuania and Estonia

---

<sup>27</sup>It should be noted, however, that this simple guideline conceals a number of complications. If (as has been the case in Estonia and Lithuania) demand deposits in either currency bear interest, it is likely that they will be at least in part the objects of store of value demand. Moreover, even if the holding of DC demand deposits is likely to be motivated predominantly by transactions demand, the holding of FC demand deposits might be driven predominantly by store of value demand. Indeed, as discussed in Chapter 3 of this thesis, it is often argued that in most developing countries even FC cash should be viewed predominantly as a financial asset, and hence the object of store of value demand, rather than a medium of exchange, and hence the object of transactions demand.

<sup>28</sup>The Euro-In process was the exchange of holdings of Euro-area member currency cash for Euros. In both Lithuania and Estonia, exchanges above a certain nominal value had to be conducted via deposit in the banking system.

for use in estimation were therefore calculated as follows:

$$\begin{aligned} \text{Lithuania} & : DR_t = \frac{FCDD_t + FCTSD_t}{DCDD_t + DCTSD_t} \\ \text{Estonia:} & DR_t = \frac{EXTRFCTSD_t}{DCTSD_t} \end{aligned}$$

### Deposit interest rates

Since the dollarisation ratio series were calculated using the total stocks of DC and FC deposits outstanding in a given month, it was necessary to transform the multiple interest rate series associated with demand deposits and term deposits of different maturities, denominated in domestic and foreign currencies, into just two series for each country – one reflecting the weighted average interest rate offered on FC deposits, and the other, the weighted average interest rate offered on DC deposits.

Data availability once again limited the range of options regarding the weightings that might be implemented. In the case of time deposits, neither the Bank of Lithuania nor the Bank of Estonia requires deposit money banks to report their stocks of time deposits disaggregated by maturity; but instead requires that they report the volume of time deposits, disaggregated by maturity, that they collect in each calendar month. For demand deposits, the requirement is the opposite: banks report the total stock of demand deposits outstanding at the end of each calendar month, but not the gross inflows in that month. The lack of data on stocks clearly rules out one obvious method for weighting the average interest rates on time deposits of different maturities – by each maturity’s share in the total outstanding stock of time deposits. However, the average interest rate on time deposits of a given maturity applies only to deposits opened in that month, and not to time deposits of the same maturity opened previously. The most appropriate weighting is therefore by that maturity of deposit’s share in the total gross inflow, rather than in the total outstanding stock, of time deposits. Nor is the absence of gross inflow data on demand deposits an obstacle to this approach: because demand deposits may be withdrawn without notice, it is reasonable to interpret the stock held at the end of each month as the gross inflow in that month.

Given the formulae used to calculate their respective dollarisation ratios given above, weighted average interest rates offered by banks on foreign and domestic currency deposits were therefore

calculated for our two cases according to the following formulas:

$$\begin{aligned} \text{Lithuania:} \quad I^c &= \frac{\alpha^c}{\left(\alpha^c + \sum_{i=1}^6 \beta_{i,t}^c\right)} DDI_t^c + \sum_{j=1}^6 \frac{\beta_{j,t}^c}{\left(\alpha^c + \sum_{i=1}^6 \beta_{i,t}^c\right)} TDI_{j,t}^c \\ \text{Estonia:} \quad I^c &= \sum_{j=1}^4 \frac{\beta_{j,t}^c}{\left(\alpha^c + \sum_{i=1}^4 \beta_{i,t}^c\right)} TDI_{j,t}^c \end{aligned}$$

where for currency of deposit denomination  $c$  (=domestic or foreign), and month  $t$ ;  $\alpha$  and  $DDI_t$  are the total stock of, and the average interest rate offered on, demand deposits; and  $\beta_j$  and  $TDI_{j,t}$  are the total gross inflows of, and average interest rate offered on, time deposits of maturity  $j$ . In the case of Lithuania, the calculation was made using data on six maturities of time deposits (1=up to one month; 2=from one to three months; 3=from three to six months; 4=from six to twelve months; 5=from one to two years; and 6=over two years); in the case of Estonia, four (1=up to three months; 2=from three to six months; 3=six to twelve months; 4=over twelve months).

### Expected depreciation

In the case of a country with a CBA, conventional empirical counterparts to  $E_t[e_{t+1}]$  where  $e$  represents the depreciation of the DC against the CBA anchor currency are not available. However, the fact that for both Estonia and Lithuania,  $DR_t$  - the empirical measure of the deposit dollarisation ratio - includes in its numerator FC deposits denominated in non-CBA anchor currencies means that it is not only expectations of changes in the exchange rate of the DC against the CBA anchor currency, but against other, non-CBA foreign currencies in which deposits are held which play a part in the portfolio decision to be modelled. Since observations on the rate of depreciation of the EEK and LIT against these other, non-CBA anchor currencies are available and exhibit variation over time, either of the two conventional strategies above can be used to generate empirical counterparts to the expected change in their exchange rate against the domestic currency. We therefore constructed two series to control for expected changes in the exchange rate of the EEK and LIT respectively against non-anchor foreign currencies.

Given the construction of  $DR_t$ , these series evidently need to be calculated as weighted

averages of the depreciation of the EEK and LIT against a range of currencies, in a manner analogous to the calculation of the deposit interest rate series described above. Once again, data on the composition of outstanding stocks of deposits (this time the currency composition of FC deposits, rather than the maturity composition as above) are not reported, so that available data on the currency composition of gross inflows of FC deposits has to be used as a guide. Fortunately these data are informative for both Lithuania and Estonia.

The simpler of the two cases is Estonia. Here, the data shows that inflows of FC deposits denominated in foreign currencies other than the DEM / Euro or the US Dollar have been negligible over the period. As Figure 4-5 demonstrates, the proportion of gross FC deposit inflows denominated in US Dollars (a non-CBA anchor currency) was close to 100% at the beginning of our period under investigation, and despite a substantial decline, remains at approximately 50% even at its end. Even without knowing the share of the stock of FC deposits held in US Dollars, therefore, we can be sure that it is important to control for the rate of expected EEK-USD depreciation in our empirical modelling over the whole sample period.

The case of Lithuania is somewhat different. As discussed above, the Lithuanian CBA was initially pegged to the US Dollar; but in February, 2002 was re-pegged to the Euro at the prevailing nominal exchange rate. By contrast to the case of Estonia, FC deposit inflows denominated in currencies other than the anchor currency were scarce up until the re-peg. As can be seen in Figure 4-6, until mid-2001, nearly all FC deposit inflows in Lithuania continued to be denominated in US Dollars. We can reasonably infer from this that approximately the whole stock of FC deposits remained denominated in US Dollars, so that expected fluctuations in, for example, the LIT-Euro exchange rate would have played little role in depositors portfolio allocation decisions prior to February, 2002. By the same reasoning, however, this situation is dramatically reversed following the re-peg of the CBA. Immediately after that date, the vast majority of FC deposits - being still denominated in US Dollars - were suddenly exposed to the nominal fluctuation of the litas - now pegged to the Euro - against the US Dollar. In the case of Lithuania, therefore, it is important to control for the expected devaluation of the LIT-USD once it started to vary after February, 2002.

We therefore constructed series describing EEK-USD and LIT-USD nominal depreciation by taking the twelve month difference of the natural logarithm of the EEK-USD ( $NER_t^{EEK/USD}$ )

and LIT-USD ( $NER_t^{LIT/USD}$ ) nominal exchange rates reported by the Bank of Estonia and Bank of Lithuania, as follows:

$$\text{Lithuania: } N_t = \ln(NER_t^{LIT/USD}) - \ln(NER_{t-12}^{LIT/USD})$$

$$\text{Estonia : } N_t = \ln(NER_t^{EEK/USD}) - \ln(NER_{t-12}^{EEK/USD})$$

## Appendix B: VECM models for Lithuania and Estonia

### Lithuania

The single cointegrating combination estimated under  $R(1)$  and  $R(2)$  was defined, as indicated in Table 4.5, as follows:

$$CIV_t = dr_t + 0.44 + 0.81i_t^{DCD} - 0.81i_t^{FCD} - 4.45N_t \quad (4.9)$$

Since  $i^{FCD}$  was excluded from the matrix of short run adjustment coefficients in the I(1) analysis described above, the modelling of the analogous I(0) system started from a reduced  $\mathbf{y}_t$  vector including only  $dr$  and  $i^{DCD}$ . Seven lags were included to begin with; and two impulse dummies designed to control for the impact of particular events that might be expected to have had an effect on the short run development of either the deposit dollarisation ratio or the interest rates offered on domestic currency deposits. Our investigation of the short run dynamics of the system therefore started from the following VECM(7):

$$\begin{aligned} \Delta \tilde{\mathbf{y}}_t &= \alpha\beta'(1 : \tilde{\mathbf{y}}_{t-1} : N_t) + \sum_{i=1}^7 \delta_i \Delta \mathbf{y}_{t-i} + \Xi \mathbf{z}_t + \nu_t \\ \nu_t &\sim \mathbf{IN}_n(\mathbf{0}, \Omega) \end{aligned} \quad (4.10)$$

$$\tilde{\mathbf{y}}_t' = \begin{bmatrix} dr_t & i_t^{DCD} \end{bmatrix}$$

$$\mathbf{z}_t' = \begin{bmatrix} 1 & DummyTB_t & DummyEL_t \end{bmatrix}$$

where the place of  $\alpha\beta'(1 : \tilde{\mathbf{y}}_{t-1} : N_{t-1})$  was taken by the one period lag of the combination (4.9);  $DummyTB$ , taking a value of 1 in September and October, 1994, and 0 otherwise, controls for the effects of the opening of the Lithuanian Treasury Bill market in September, 1994; and  $DummyEL$ , taking a value of 1 in October, 1996 and zero otherwise, controls for the effects of the general election in that month.

This initial specification was then reduced sequentially, excluding lags upon which the coefficients were found not to be significant. Substantial simplification of the lag structure of the short run system was possible using this method. The two equation system eventually arrived

at, and chosen on the basis of its combination of parsimony with relatively good fit, is the following:

$$\begin{aligned}
\Delta dr_t = & 0.005 + 0.22 \Delta dr_{t-1} + 0.14 \Delta dr_{t-4} + 0.03 \Delta i_{t-3}^{DCD} \\
& (0.005) \quad (0.08) \quad (0.08) \quad (0.04) \\
& - 0.08 \Delta i_{t-5}^{DCD} + 0.10 \Delta i_{t-7}^{DCD} + 0.08 \Delta i_t^{FCD} \\
& (0.04) \quad (0.04) \quad (0.04) \\
& - 0.11 \Delta i_{t-4}^{FCD} + 0.05 \Delta_{t-6}^{FCD} - 0.03 \Delta i_{t-7}^{FCD} \\
& (0.03) \quad (0.03) \quad (0.03) \\
& + 0.08 \textit{DummyTB}_t + 0.13 \textit{DummyEL}_t - 0.08 \textit{CIV}_{t-1} \\
& (0.04) \quad (0.05) \quad (0.02)
\end{aligned} \tag{4.11}$$

$$\begin{aligned}
\Delta i_t^{DCD} = & -0.013 - 0.10 \Delta dr_{t-1} - 0.20 \Delta dr_{t-4} - 0.14 \Delta i_{t-3}^{DCD} \\
& (0.01) \quad (0.16) \quad (0.16) \quad (0.08) \\
& - 0.06 \Delta i_{t-5}^{DCD} - 0.15 \Delta i_{t-7}^{DCD} + 0.18 \Delta i_t^{FCD} \\
& (0.08) \quad (0.07) \quad (0.07) \\
& + 0.10 \Delta i_{t-4}^{FCD} - 0.08 \Delta_{t-6}^{FCD} + 0.14 \Delta i_{t-7}^{FCD} \\
& (0.07) \quad (0.06) \quad (0.06) \\
& - 0.27 \textit{DummyTB}_t + 0.14 \textit{DummyEL}_t - 0.20 \textit{CIV}_{t-1} \\
& (0.07) \quad (0.10) \quad (0.05)
\end{aligned} \tag{4.12}$$

The first lag of the long run combination  $\textit{CIV}_t$  is strongly significant for the development of both of the variables in  $\Delta \tilde{\mathbf{y}}_t$ , confirming, by the Engle-Granger representation theorem, that the single vector identified and estimated under  $R(1)$  and  $R(2)$  does indeed represent a cointegrating combination of the variables  $dr_t$ ,  $i_t^{DCD}$ ,  $i_t^{FCD}$ , and  $N_t$ .

The estimated coefficients on the equilibrium correction term are consistent with the estimates of the elements of the short run adjustment matrix recovered during the I(1) analysis above and reported in Table 4.5. The significance of the variables retained after reduction for

the system as a whole is indicated by the results of  $F$  tests reported in Table 4.18. The joint significance of all the variables included in the system was likewise confirmed using an  $F$  test ( $F(24, 220) = 5.64[0.00]$ ).

The results of the standard diagnostic tests applied to the residuals of this short run system are reported in Table 4.19: there is no evidence of abnormality, serial correlation, heteroskedasticity, or ARCH. The cross-correlations and standard deviations of the equation residuals are reported in Table 4.20. The former are low; the latter are acceptable (less than 5%) in the case of the equation for the first difference of  $dr$ ; less so (roughly 9%) in the case of the first difference of  $i^{DCD}$ . The fit of the model, summarised by the correlations between actual and fitted values of the two endogenous variables reported in Table 4.20, the  $R^2$  ratio of 0.62, and its graphical presentation in Figure 4-13, is reasonable. The stability of the parameter estimates is demonstrated graphically by Figure 4-14, which illustrates the results of Break Point, one step, and N down Chow tests for each of the individual equations and for the system as a whole.

## Estonia

In the Estonian case, the specification of the VECM(5) from which investigation of the short run dynamics of the system started was the following:

$$\begin{aligned} \Delta \mathbf{y}_t &= \alpha \beta' (1 : \mathbf{y}_{t-1} : N_{t-1}) + \sum_{i=1}^5 \delta_i \Delta \mathbf{y}_{t-i} + \Xi \mathbf{z}_t + \boldsymbol{\nu}_t \\ \boldsymbol{\nu}_t &\sim \mathbf{IN}_n(\mathbf{0}, \boldsymbol{\Omega}) \end{aligned} \quad (4.13)$$

$$\mathbf{z}'_t = \begin{bmatrix} 1 & Dummy0112 \end{bmatrix}$$

where the place of  $\alpha \beta' (1 : \mathbf{y}_{t-1} : N_{t-1})$  was taken by the one period lag of the  $I(0)$  combination:

$$CIV_t = dr_t + 1.00 + 0.77i_t^{DCD} - 0.77i_t^{FCD} - 0.71N_t \quad (4.14)$$

and  $Dummy0112$ , taking a value of 1 in December, 2001 and zero otherwise, was included to control for the resignation of the Estonian government in that month.

Once again, substantial simplification of the lag structure of the short run system was

possible, to arrive at the following three equation system:

$$\begin{aligned}
\Delta dr_t = & \underset{(0.011)}{0.006} + \underset{(0.12)}{0.07} \Delta dr_{t-4} + \underset{(0.09)}{0.09} \Delta i_{t-1}^{DCD} + \underset{(0.08)}{0.08} \Delta i_{t-2}^{DCD} \\
& + \underset{(0.08)}{0.17} \Delta i_{t-3}^{DCD} + \underset{(0.09)}{0.20} \Delta i_{t-4}^{DCD} - \underset{(0.16)}{0.18} \Delta i_{t-1}^{FCD} \\
& - \underset{(0.16)}{0.37} \Delta i_{t-2}^{FCD} - \underset{(0.13)}{0.16} \Delta i_{t-3}^{FCD} - \underset{(0.14)}{0.34} \Delta i_{t-4}^{FCD} \\
& - \underset{(0.09)}{0.04} \textit{Dummy0112}_t - \underset{(0.06)}{0.27} \textit{CIV}_{t-1}
\end{aligned} \tag{4.15}$$

$$\begin{aligned}
\Delta i_t^{DCD} = & - \underset{(0.012)}{0.001} + \underset{(0.13)}{0.39} \Delta dr_{t-4} + \underset{(0.10)}{0.32} \Delta i_{t-1}^{DCD} + \underset{(0.09)}{0.21} \Delta i_{t-2}^{DCD} \\
& + \underset{(0.09)}{0.30} \Delta i_{t-3}^{DCD} - \underset{(0.10)}{0.02} \Delta i_{t-4}^{DCD} - \underset{(0.17)}{0.23} \Delta i_{t-1}^{FCD} \\
& - \underset{(0.17)}{0.35} \Delta i_{t-2}^{FCD} - \underset{(0.14)}{0.56} \Delta i_{t-3}^{FCD} - \underset{(0.16)}{0.41} \Delta i_{t-4}^{FCD} \\
& - \underset{(0.10)}{0.34} \textit{Dummy0112}_t - \underset{(0.07)}{0.34} \textit{CIV}_{t-1}
\end{aligned} \tag{4.16}$$

$$\begin{aligned}
\Delta i_t^{FCD} = & - \underset{(0.009)}{0.006} + \underset{(0.10)}{0.02} \Delta dr_{t-4} - \underset{(0.07)}{0.022} \Delta i_{t-1}^{DCD} + \underset{(0.07)}{0.28} \Delta i_{t-2}^{DCD} \\
& + \underset{(0.07)}{0.06} \Delta i_{t-3}^{DCD} + \underset{(0.08)}{0.11} \Delta i_{t-4}^{DCD} - \underset{(0.13)}{0.20} \Delta i_{t-1}^{FCD} \\
& - \underset{(0.13)}{0.06} \Delta i_{t-2}^{FCD} + \underset{(0.11)}{0.02} \Delta i_{t-3}^{FCD} - \underset{(0.12)}{0.18} \Delta i_{t-4}^{FCD} \\
& - \underset{(0.08)}{0.10} \textit{Dummy0112}_t - \underset{(0.05)}{0.16} \textit{CIV}_{t-1}
\end{aligned} \tag{4.17}$$

The first lag of the long run combination  $CIV_t$  is strongly significant for the development of all three variables in  $\Delta \mathbf{y}_t$ , confirming that the single vector identified and estimated under  $R(2)$  in the previous section does indeed represent a cointegrating combination. The estimated coefficients on the equilibrium correction term are consistent with the estimates of the elements of the short run adjustment matrix recovered during the analysis of the  $I(1)$  model under the same restriction. The significance of the variables retained after reduction for the system as a whole is indicated by the results of  $F$  tests reported in Table 4.26. The joint significance of all the variables included in the system was likewise confirmed using an  $F$  test ( $F(33, 180) = 3.45$ ).

The results of the standard diagnostic tests applied to the residuals of this short run system are reported in Table 4.27: there is no evidence of abnormality, serial correlation, heteroskedasticity, or ARCH. The cross-correlations and standard deviations of the equation residuals are reported in Table 4.28. The former are low; the latter below 10% for all three equations. The fit of the model, summarised by the correlations between actual and fitted values of the two endogenous variables reported in Table 4.28, the  $R^2$  ratio of 0.76, and its graphical presentation in Figure 4-17, is, like that of the Lithuanian  $I(0)$  model, reasonable. The parameter estimates are stable over the sample, as demonstrated graphically by Figure 4-18.

## Appendix C: A Two Currency Portfolio Balance Model of Deposit Dollarisation

In Chapter 3 of this thesis, we reviewed the modern portfolio balance model of Ize and Levy Yeyati (2003) which explains observed financial dollarisation as a financial equilibrium between lenders (depositors) and borrowers (banks). Here we present a simplified version of this model which analyses the choices of depositors only, and abstracts from decisions concerning the location (on- or off-shore) of FC deposits, in order to motivate the empirical investigation in the main body of this chapter.

We begin with a single, representative depositor, and consider his optimal choice of the composition of his deposit portfolio between two currencies.

### Portfolio shares

We define the share of deposits denominated in foreign and domestic currency in total deposits at time  $t$  as  $x_t$  and  $1 - x_t$  respectively.

### Real returns

We define the real returns to holding deposits denominated in foreign and domestic currency as follows:

$$\begin{aligned} r_t^F &= E(r_t^F) + \mu_{s,t} \\ r_t^D &= E(r_t^D) - \mu_{\pi,t} \end{aligned}$$

where  $\mu_{i,t}$  is a stochastic, zero mean disturbance with variance  $S_{ii,t}$  associated with unexpected changes in the domestic currency real exchange rate against the foreign currency ( $i = s$ ) and the domestic inflation rate ( $i = \pi$ ).

We define the spread between foreign and domestic currency deposits as follows:

$$\rho_t^F = (r_t^F - r_t^D)$$

## Distribution of average portfolio returns

The first two moments of average portfolio returns  $r$  can therefore be defined as:

$$\begin{aligned} E(r_t) &= x_t E(\rho_t^F) + E(r_t^D) \\ \text{Var}(r_t) &= \text{Var}(x_t \rho_t^F) + \text{Var}(r_t^D) + 2\text{Cov}(x_t \rho_t^F, r_t^D) \\ &= x_t^2 \text{Var}(\rho_t^F) + \text{Var}(r_t^D) + 2x_t \text{Cov}(\rho_t^F, r_t^D) \end{aligned}$$

## Depositors' utility

The representative depositor has a utility function of the form:

$$U_t(r_t) = E(r_t) - \gamma \frac{\text{Var}(r_t)}{2}$$

where  $\gamma$  is a coefficient of risk aversion.

## Utility maximisation

The depositor's problem is therefore:

$$\text{Max}_{x^F} U_t(r_t) = x_t E(\rho_t^F) + E(r_t^D) - \gamma \frac{(x_t^2 \text{Var}(\rho_t^F) + \text{Var}(r_t^D) + 2x_t \text{Cov}(\rho_t^F, r_t^D))}{2}$$

Assuming that  $U(r)$  is continuously differentiable, we have that:

$$\begin{aligned} \frac{\delta U_t}{\delta x_t} &= E(\rho_t^F) - \gamma \frac{(2x_t \text{Var}(\rho_t^F) + 2\text{Cov}(\rho_t^F, r_t^D))}{2} \\ &= E(\rho_t^F) - \gamma(x_t \text{Var}(\rho_t^F) + \text{Cov}(\rho_t^F, r_t^D)) \end{aligned}$$

from which the first order condition for maximisation is:

$$E(\rho_t^F) - \gamma(x_t \text{Var}(\rho_t^F) + \text{Cov}(\rho_t^F, r_t^D)) = 0$$

so that the optimal choice of  $x_t^F$  is:

$$x_t^* = \frac{(\zeta E(\rho_t^F) - Cov(\rho_t^F, r_t^D))}{Var(\rho_t^F)} \quad (4.18)$$

where  $\zeta = \frac{1}{\gamma}$ .

Equation (4.18) can be written more simply as:

$$x_t^* = \Phi_t - \zeta \Psi_t (E[r_t^D] - E[r_t^F]) \quad (4.19)$$

where:

$$\begin{aligned} \Phi_t &= \left( \frac{S_{\pi\pi} + S_{\pi s}}{S_{\pi\pi} + S_{ss} + 2S_{\pi s}} \right)_t \\ \Psi_t &= \left( \frac{1}{S_{\pi\pi} + S_{ss} + 2S_{\pi s}} \right)_t \end{aligned}$$

Equation (4.19) is the solution used in (4.8) above.

**Table 4.9: Estonia: ADF Tests of Variables in Levels.**

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
<i>dr</i>	None	0	-0.1229
	C	0	-1.501
	C&T	0	-1.511
	C&SD	0	-1.131
	C,T, & SD	0	-1.126
<i>i<sup>DCD</sup></i>	None	1	-1.033
	C	4	-1.398
	C&T	3	-4.611 **
	C&SD	1	-1.143
	C,T, & SD	3	-3.639*
<i>i<sup>FCD</sup></i>	None	0	-1.327
	C	2	0.1056
	C&T	0	-1.860
	C&SD	2	0.06775
	C,T, & SD	2	-1.937
<i>N</i>	None	1	-1.588
	C	1	-1.497
	C&T	1	-2.141
	C&SD	1	-1.314
	C,T, & SD	2	-1.658

Note: sample is 1997.10 - 2003.12.

Table 4.10: Estonia: ADF Tests of Variables in First Differences.

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$\Delta dr$	None	0	-9.852 **
	C	0	-9.791 **
	C&T	0	-10.04 **
	C&SD	0	-8.875 **
	C,T, & SD	0	-9.153 **
$\Delta i^{DCD}$	None	3	-4.885 **
	C	4	-4.982 **
	C&T	4	-4.934 **
	C&SD	0	-6.161 **
	C,T, & SD	0	-6.124 **
$\Delta i^{FCD}$	None	1	-4.758 **
	C	1	-4.945 **
	C&T	1	-5.194 **
	C&SD	1	-3.897 **
	C,T, & SD	2	-3.549*
$\Delta N$	None	0	-6.379 **
	C	0	-6.432 **
	C&T	0	-6.382 **
	C&SD	0	-5.930 **
	C,T, & SD	2	-5.879 **

Note: sample is 1997.11 - 2003.12.

**Table 4.11: Lithuania - ADF Tests of Variables in Levels.**

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
<i>dr</i>	None	4	-0.1025
	C	4	-1.560
	C&T	4	-1.680
	C&SD	4	-1.763
	C,T, & SD	4	-1.865
<i>i<sup>DCD</sup></i>	None	0	-3.571 **
	C	0	-1.639
	C&T	0	-2.044
	C&SD	0	-1.559
	C,T, & SD	0	-1.878
<i>i<sup>FCD</sup></i>	None	0	-2.436*
	C	0	-1.277
	C&T	0	-1.792
	C&SD	0	-1.192
	C,T, & SD	2	-1.895
<i>N</i>	None	0	-0.002230
	C	4	-0.3629
	C&T	4	-1.410
	C&SD	4	-0.3260
	C,T, & SD	4	-1.344

Note: sample is 1993.09 - 2003.12.

Table 4.12: Lithuania - ADF Tests of Variables in First Differences.

Variable	Deterministic terms	Augmentation	Value of ADF test statistic
$\Delta dr$	None	3	-4.264 **
	C	0	-8.653 **
	C&T	3	-4.307 **
	C&SD	3	-4.009 **
	C,T, & SD	0	-7.533 **
$\Delta i^{DCD}$	None	0	-9.836 **
	C	1	-7.489 **
	C&T	0	-10.62 **
	C&SD	0	-9.822 **
	C,T, & SD	0	-9.846 **
$\Delta i^{FCD}$	None	4	-4.308 **
	C	0	-11.42 **
	C&T	4	-4.725 **
	C&SD	4	-4.275 **
	C,T, & SD	4	-4.289 **
$\Delta N$	None	4	-4.697 **
	C	4	-4.862 **
	C&T	4	-5.166 **
	C&SD	4	-4.587 **
	C,T, & SD	4	-4.898 **

Note: sample is 1993.10 - 2003.12.

**Table 4.13. Lithuania: diagnostic tests of on the residuals of UVAR(8).**

Variable	Test	Test Statistic	Value
<i>dr</i>	AR(1-7)	$F(7, 80)$	0.35035[0.9278]
	Normality	$\chi^2(2)$	0.48025[0.7865]
	ARCH(1-7)	$F(7, 73)$	1.0022[0.4368]
	Heteroskedasticity	$F(50, 36)$	0.41983[0.9977]
<i>i<sup>DCD</sup></i>	AR(1-7)	$F(7, 80)$	0.87598[0.5293]
	Normality	$\chi^2(2)$	1.8881[0.3890]
	ARCH(1-7)	$F(7, 73)$	0.65095[0.7124]
	Heteroskedasticity	$F(50, 36)$	0.53495[0.9797]
<i>i<sup>FCD</sup></i>	AR(1-7)	$F(7, 80)$	0.58419[0.7669]
	Normality	$\chi^2(2)$	12.471[0.0020] **
	ARCH(1-7)	$F(7, 73)$	0.82178[0.5724]
	Heteroskedasticity	$F(50, 36)$	1.5732[0.0782]
System	AR(1-7)	$F(63, 191)$	1.0048[0.4769]
	Normality	$\chi^2(6)$	8.0629[0.2335]
	Heteroskedasticity	$F(300, 193)$	0.59310[1.0000]

**Table 4.14. Lithuania: cross-correlations and standard deviations of residuals, UVAR(8).**

	<i>dr</i>	<i>i<sup>DCD</sup></i>	<i>i<sup>FCD</sup></i>	s.d.
<i>dr</i>	1	-0.051727	0.072219	0.043363
<i>i<sup>DCD</sup></i>	-0.051727	1	0.37527	0.10715
<i>i<sup>FCD</sup></i>	0.072219	0.37527	1	0.12471

**Table 4.15. Lithuania: cross-correlations of actual and fitted values of UVAR(8).**

	Fitted values
<i>dr</i>	0.98582
<i>i<sup>DCD</sup></i>	0.99678
<i>i<sup>FCD</sup></i>	0.99376

**Table 4.16. Lithuania: Eigenvalues of long run matrix  $P_0$ .**

Real	Imaginary	Modulus
-0.3335	0.000	0.3335
-0.06289	0.000	0.06289
-0.02480	0.000	0.02480

**Table 4.17. Lithuania: tests of additional restrictions imposed on CVAR(8).**

Additional Restrictions Not Accepted	Test statistic	Value
$\alpha: dr = 0$	$\chi^2(3)$	9.8264[0.0201]*
$\alpha: i^{DCD} = 0$	$\chi^2(3)$	10.453[0.0151]*
$\beta: dr = 0$	$\chi^2(3)$	13.687[0.0034] **
$\beta: i^{DCD} = 0$	$\chi^2(2)$	16.659[0.0002] **
$\beta: i^{FCD} = 0$	$\chi^2(2)$	16.027[0.0003] **

**Table 4.18. Lithuania: tests for the significance of regressors, VECM.**

Variable	Test statistic	Value
$\Delta dr_{t-1}$	$F(2, 110)$	3.86555[0.024]*
$\Delta dr_{t-4}$	$F(2, 110)$	2.35931[0.099]
$\Delta i_{t-3}^{DCD}$	$F(2, 110)$	1.57813[0.211]
$\Delta i_{t-5}^{DCD}$	$F(2, 110)$	2.27711[0.107]
$\Delta i_{t-7}^{DCD}$	$F(2, 110)$	5.30533[0.006] **
$\Delta i_t^{FCD}$	$F(2, 110)$	5.66602[0.005] **
$\Delta i_{t-4}^{FCD}$	$F(2, 110)$	6.23790[0.003] **
$\Delta i_{t-6}^{FCD}$	$F(2, 110)$	1.87735[0.158]
$\Delta i_{t-7}^{FCD}$	$F(2, 110)$	2.87730[0.061]
<i>DummyTB</i>	$F(2, 110)$	8.70559[0.000] **
<i>DummyEL9610</i>	$F(2, 110)$	4.99331[0.008] **
<i>Constant</i>	$F(2, 110)$	1.02240[0.363]
$CIV_{t-1}$	$F(2, 110)$	16.9400[0.000] **
Joint test, all regressors	$F(24, 220)$	5.6385[0.0000] **

**Table 4.19. Lithuania: diagnostic tests of on the residuals of VECM.**

Variable	Test	Test Statistic	Value
$\Delta dr$	AR(1-7)	$F(7, 104)$	0.28693[0.9578]
	Normality	$\chi^2(2)$	1.3611[0.5063]
	ARCH(1-7)	$F(7, 97)$	0.90171[0.5085]
	Heteroskedasticity	$F(22, 88)$	0.58489[0.9235]
$\Delta i^{DCD}$	AR(1-7)	$F(7, 104)$	0.30441[0.9505]
	Normality	$\chi^2(2)$	4.7799[0.0916]
	ARCH(1-7)	$F(7, 97)$	0.30709[0.9492]
	Heteroskedasticity	$F(22, 88)$	0.83510[0.6752]
System	AR(1-7)	$F(28, 192)$	0.68707[0.8803]
	Normality	$\chi^2(4)$	6.2016[0.1846]
	Heteroskedasticity	$F(66, 257)$	0.57201[0.9961]

**Table 4.20. Lithuania: cross-correlations and standard deviations of residuals, and cross-correlations of actual and fitted values, PVECM.**

	$dr$	$i^{DCD}$	s.d.	Fitted values
$dr$	1	-0.061774	0.045862	0.60949
$i^{DCD}$	-0.061774	1	0.091873	0.62269

### 4.6.1 Estonia

Table 4.21. Estonia: diagnostic tests of on the residuals of UVAR(6).

Variable	Test	Test Statistic	Value
$dr$	AR(1-5)	$F(5, 50)$	0.29469[0.9136]
	Normality	$\chi^2(2)$	3.7913[0.1502]
	ARCH(1-5)	$F(5, 45)$	0.19123[0.9644]
	Heteroskedasticity	$F(38, 16)$	0.63292[0.8774]
$i^{DCD}$	AR(1-5)	$F(5, 50)$	0.94396[0.4610]
	Normality	$\chi^2(2)$	1.4240[0.4907]
	ARCH(1-5)	$F(5, 45)$	0.55957[0.7303]
	Heteroskedasticity	$F(38, 16)$	0.33699[0.9970]
$i^{FCD}$	AR(1-5)	$F(5, 50)$	0.89925[0.4890]
	Normality	$\chi^2(2)$	1.7389[0.4192]
	ARCH(1-5)	$F(5, 45)$	1.0074[0.4245]
	Heteroskedasticity	$F(38, 16)$	0.36122[0.9950]
System	AR(1-5)	$F(45, 113)$	1.2642[0.1616]
	Normality	$\chi^2(6)$	6.1390[0.4078]
	Heteroskedasticity	$F(228, 73)$	0.44936[1.0000]

Table 4.22. Estonia: cross-correlations and standard deviations of residuals, UVAR(6).

	$dr$	$i^{DCD}$	$i^{FCD}$	s.d.
$dr$	1	-0.047346	0.20781	0.087702
$i^{DCD}$	-0.047346	1	0.21335	0.10950
$i^{FCD}$	0.20781	0.21335	1	0.072621

**Table 4.23. Estonia: cross-correlations of actual and fitted values of UVAR(6).**

	Fitted values
$dr$	0.94590
$i^{DCD}$	0.97934
$i^{FCD}$	0.98974

**Table 4.24. Estonia: Eigenvalues of long run matrix  $P_0$ .**

Real	Imaginary	Modulus
-0.4262	0.000	0.4262
-0.1955	0.000	0.1955
-0.05297	0.000	0.05297

**Table 4.25. Estonia: tests of additional restrictions imposed on CVAR(6).**

Additional Restrictions Not Accepted	Test statistic	Value
$\alpha: dr = 0$	$\chi^2(2)$	18.543[0.0001] **
$\alpha: i^{DCD} = 0$	$\chi^2(2)$	14.720[0.0006] **
$\alpha: i^{FCD} = 0$	$\chi^2(2)$	14.275[0.0008] **
$\beta: dr = 0$	$\chi^2(1)$	22.533[0.0000] **
$\beta: i^{DCD} = 0$	$\chi^2(1)$	17.947[0.0000] **
$\beta: i^{FCD} = 0$	$\chi^2(1)$	17.947[0.0000] **

Table 4.26. Estonia: tests for the significance of regressors, VECM.

Variable	Test statistic	Value
$\Delta dr_{t-4}$	$F(3, 61)$	3.15168[0.031]*
$\Delta i_{t-1}^{DCD}$	$F(3, 61)$	4.32688[0.008] **
$\Delta i_{t-2}^{DCD}$	$F(3, 61)$	7.10553[0.000] **
$\Delta i_{t-3}^{DCD}$	$F(3, 61)$	4.99041[0.004] **
$\Delta i_{t-4}^{DCD}$	$F(3, 61)$	1.89711[0.140]
$\Delta i_{t-1}^{FCD}$	$F(3, 61)$	1.54218[0.213]
$\Delta i_{t-2}^{FCD}$	$F(3, 61)$	3.48258[0.021]*
$\Delta i_{t-3}^{FCD}$	$F(3, 61)$	5.92551[0.001] **
$\Delta i_{t-4}^{FCD}$	$F(3, 61)$	4.56988[0.006] **
<i>Dummy0112</i>	$F(3, 61)$	3.98667[0.012]*
<i>Constant</i>	$F(3, 61)$	0.280262[0.839]
$CIV_{t-1}$	$F(3, 61)$	17.6237[0.000] **
Joint test, all regressors	$F(33, 180)$	3.45328[0.0000] **

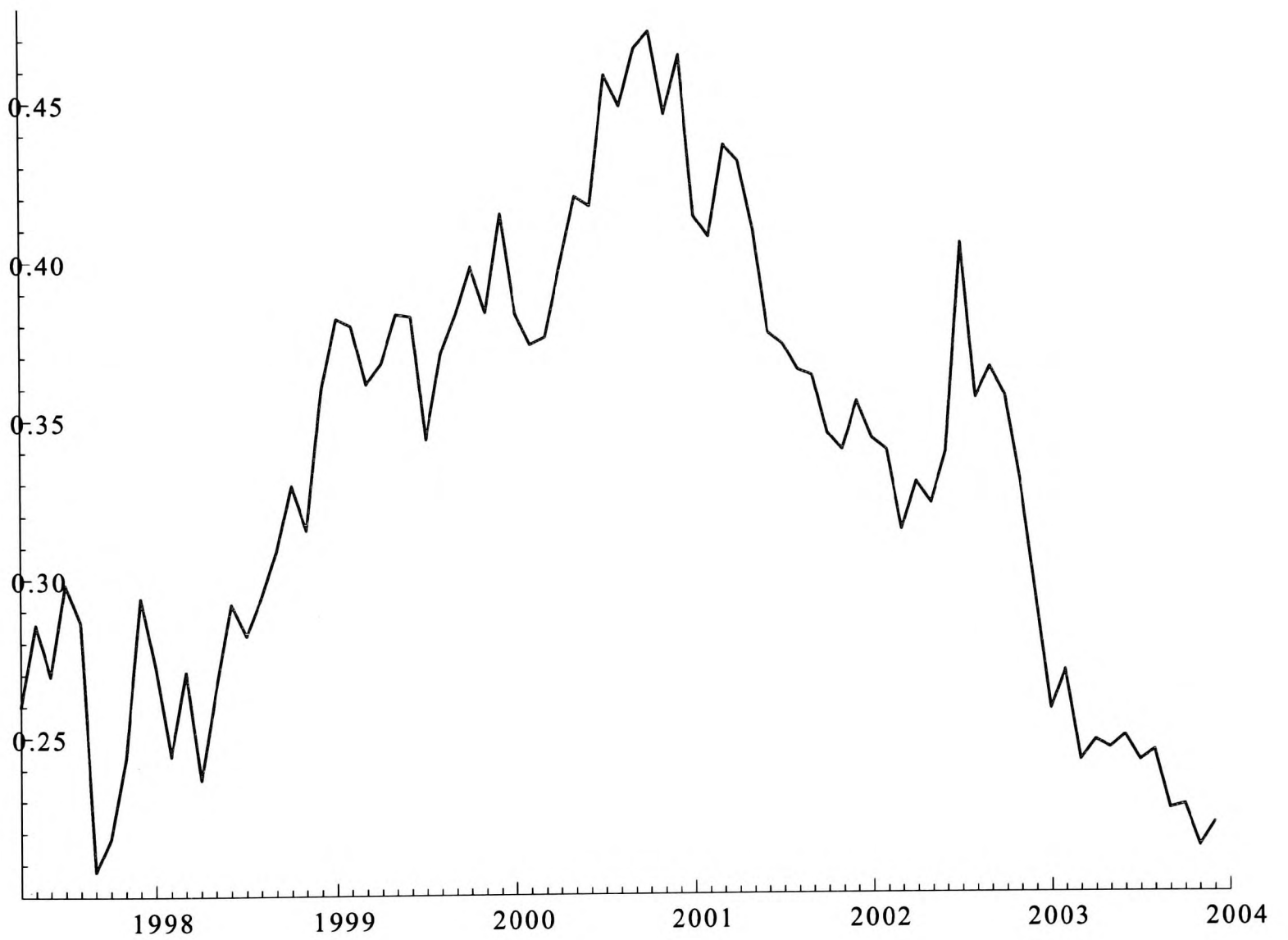
**Table 4.27. Estonia: diagnostic tests of on the residuals of VECM.**

Variable	Test	Test Statistic	Value
$\Delta dr$	AR(1-5)	$F(5, 58)$	0.34898[0.8809]
	Normality	$\chi^2(2)$	4.6505[0.0978]
	ARCH(1-5)	$F(5, 53)$	0.34840[0.8810]
	Heteroskedasticity	$F(21, 41)$	0.88920[0.6040]
$\Delta i^{DCD}$	AR(1-5)	$F(5, 58)$	0.27987[0.9223]
	Normality	$\chi^2(2)$	1.0820[0.5822]
	ARCH(1-5)	$F(5, 53)$	0.90930[0.4821]
	Heteroskedasticity	$F(21, 41)$	0.88727[0.6062]
$\Delta i^{FCD}$	AR(1-5)	$F(5, 58)$	2.0007[0.0920]
	Normality	$\chi^2(2)$	0.46956[0.7907]
	ARCH(1-5)	$F(5, 53)$	0.52746[0.7544]
	Heteroskedasticity	$F(21, 41)$	0.51551[0.9473]
System	AR(1-5)	$F(45, 137)$	0.86093[0.7141]
	Normality	$\chi^2(6)$	6.1747[0.4039]
	Heteroskedasticity	$F(126, 216)$	0.79889[0.9170]

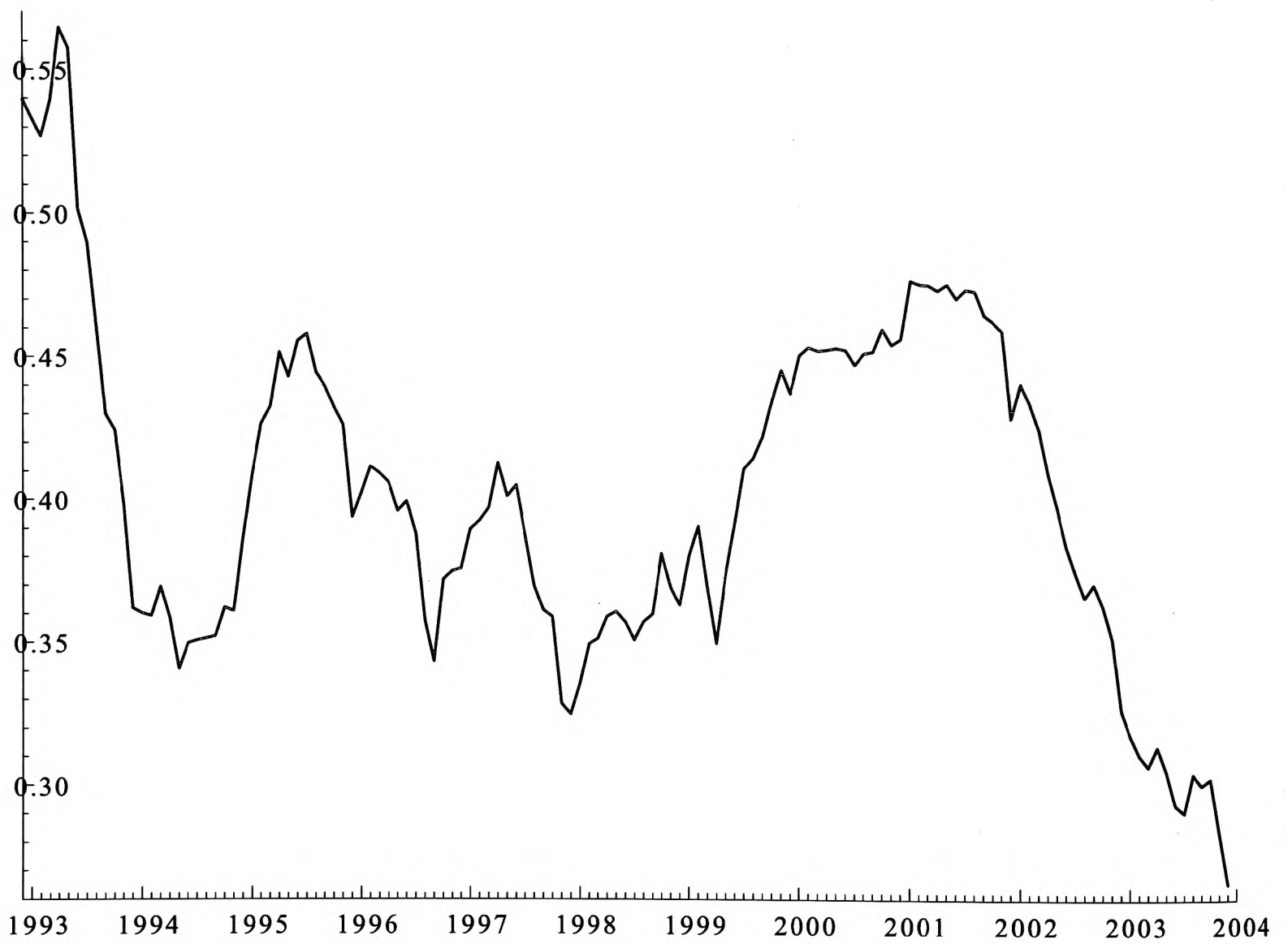
**Table 4.28. Estonia: cross-correlations and standard deviations of residuals, and cross-correlations of actual and fitted values, PVECM.**

	$dr$	$i^{DCD}$	$i^{FCD}$	s.d.	Fitted values
$dr$	1	-0.070976	0.17446	0.086463	0.53646
$i^{DCD}$	-0.070976	1	0.095077	0.095861	0.75402
$i^{FCD}$	0.17446	0.095077	1	0.072686	0.57715

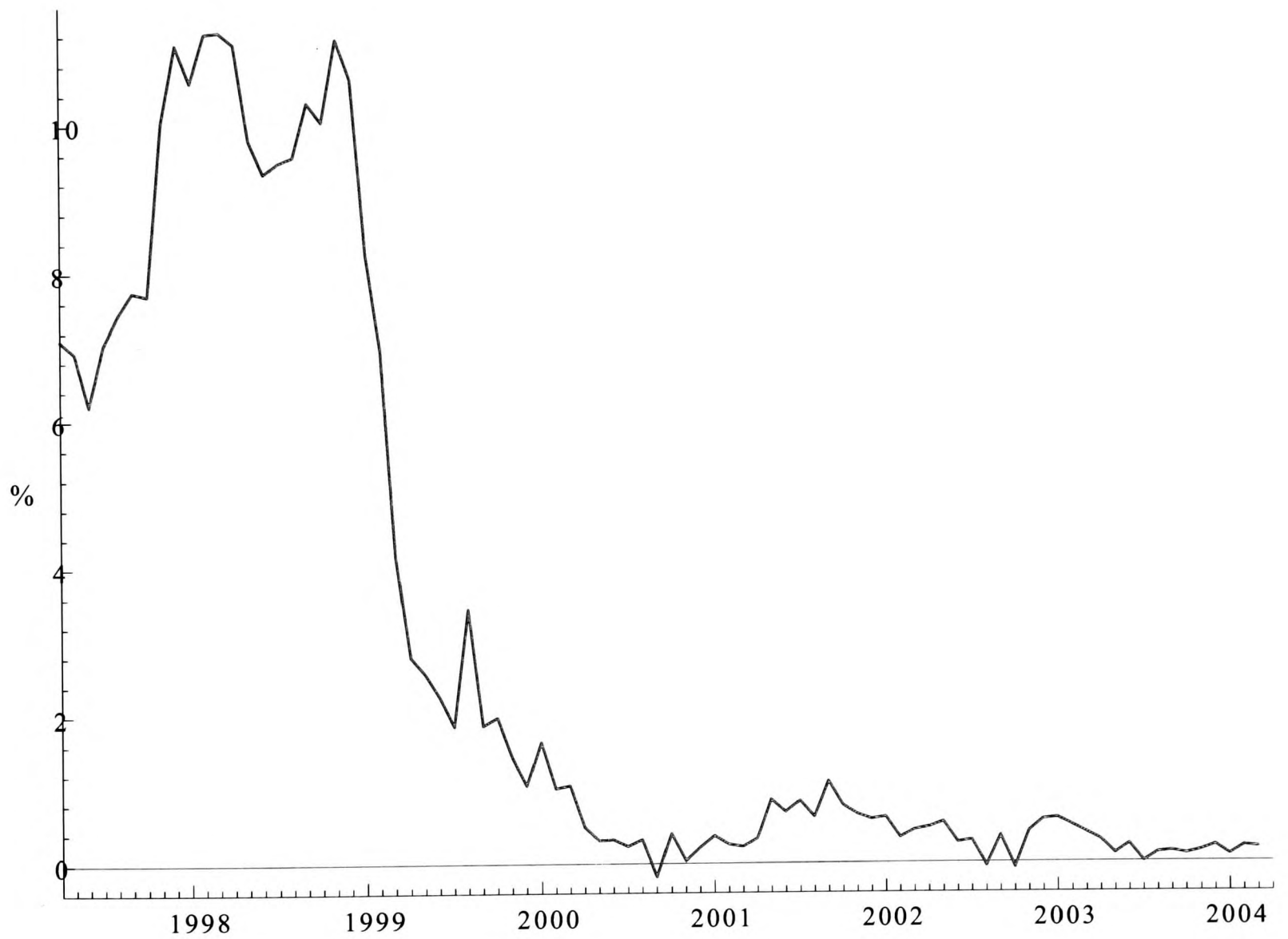
**Figure 4-1:** Estonia: Deposit Dollarisation Ratio ( $DDR = \frac{FCD}{DCD+FCD}$ )



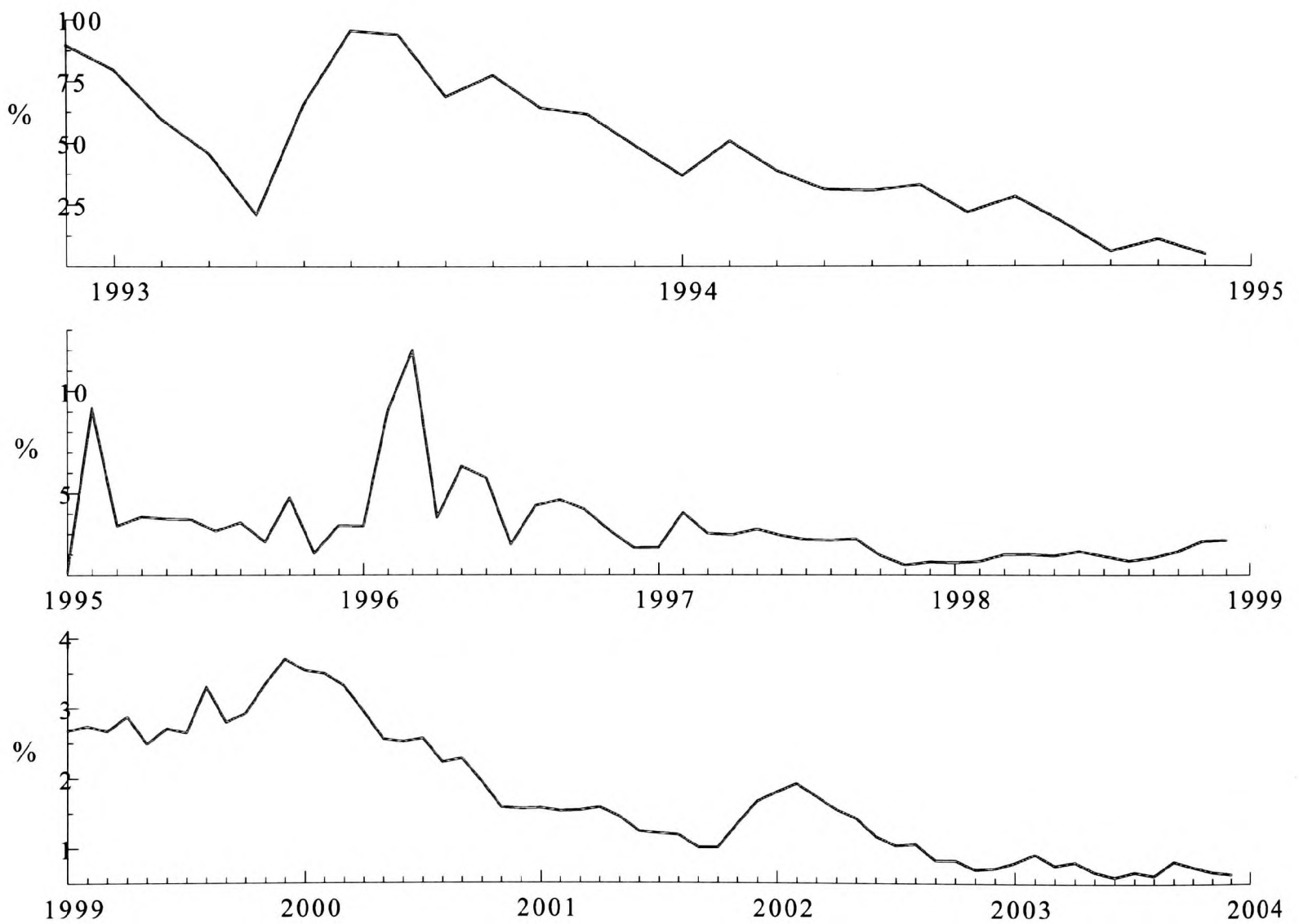
**Figure 4-2:** Lithuania: Deposit Dollarisation Ratio ( $DDR = \frac{FCD}{DCD+FCD}$ )



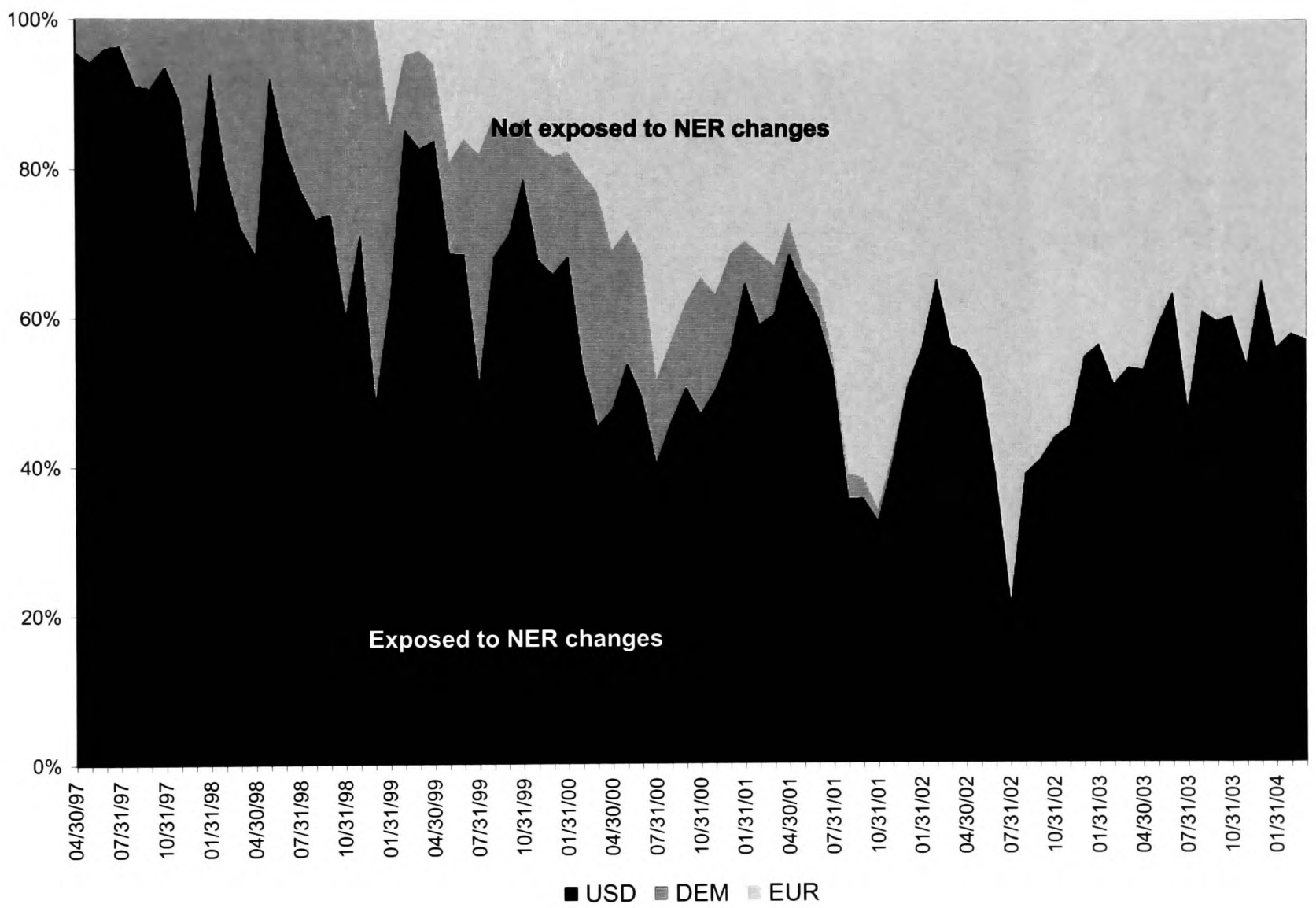
**Figure 4-3:** Estonia: spread between average nominal interest rates on DC and FC 3-6 month term deposits



**Figure 4-4:** Lithuania: spread between average nominal interest rates on DC and FC 3-6 month term deposits



**Figure 4-5:** Estonia: currency composition of gross monthly inflows of FC deposits



**Figure 4-6:** Lithuania: currency composition of gross monthly inflows of FC deposits

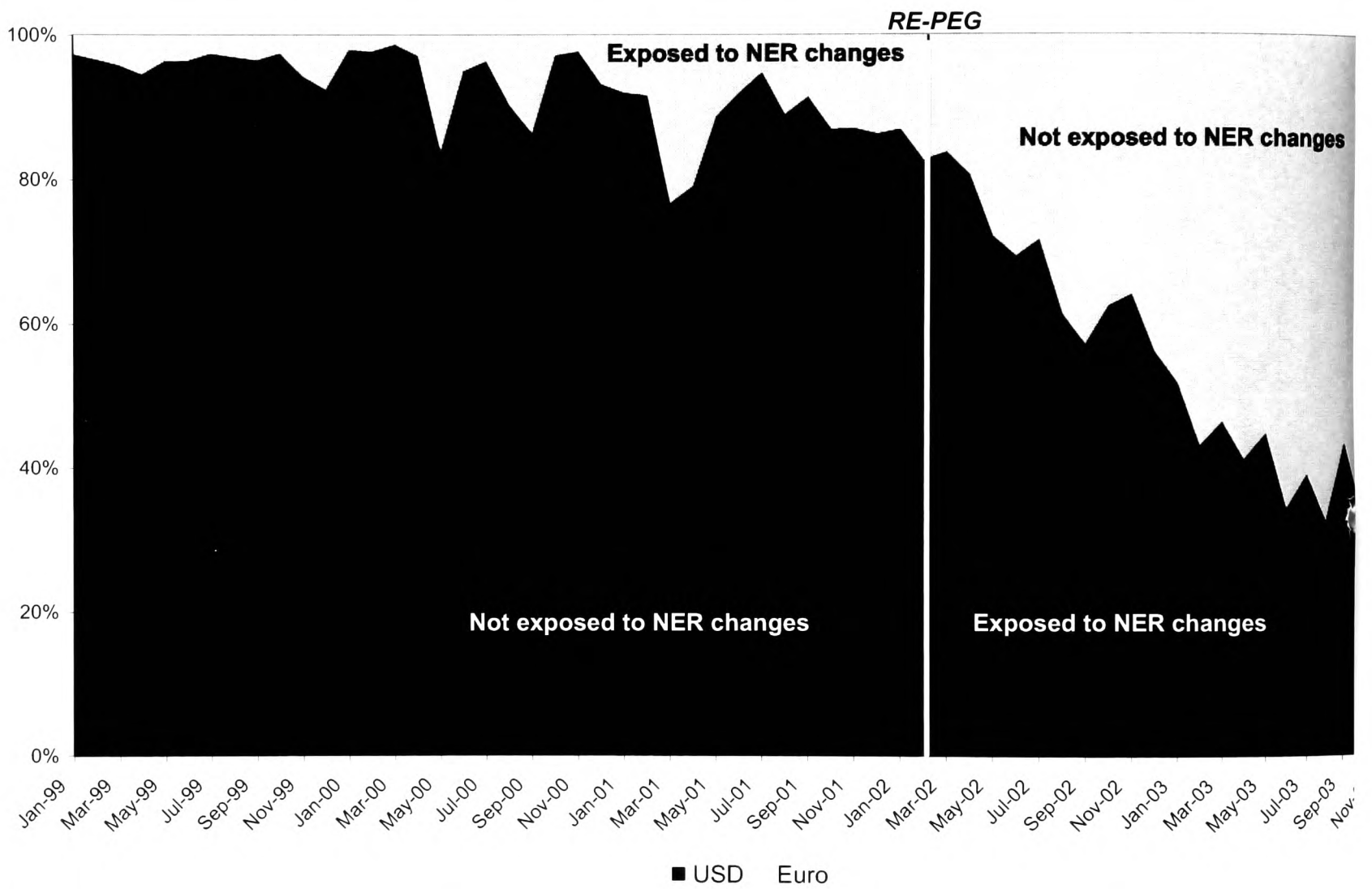
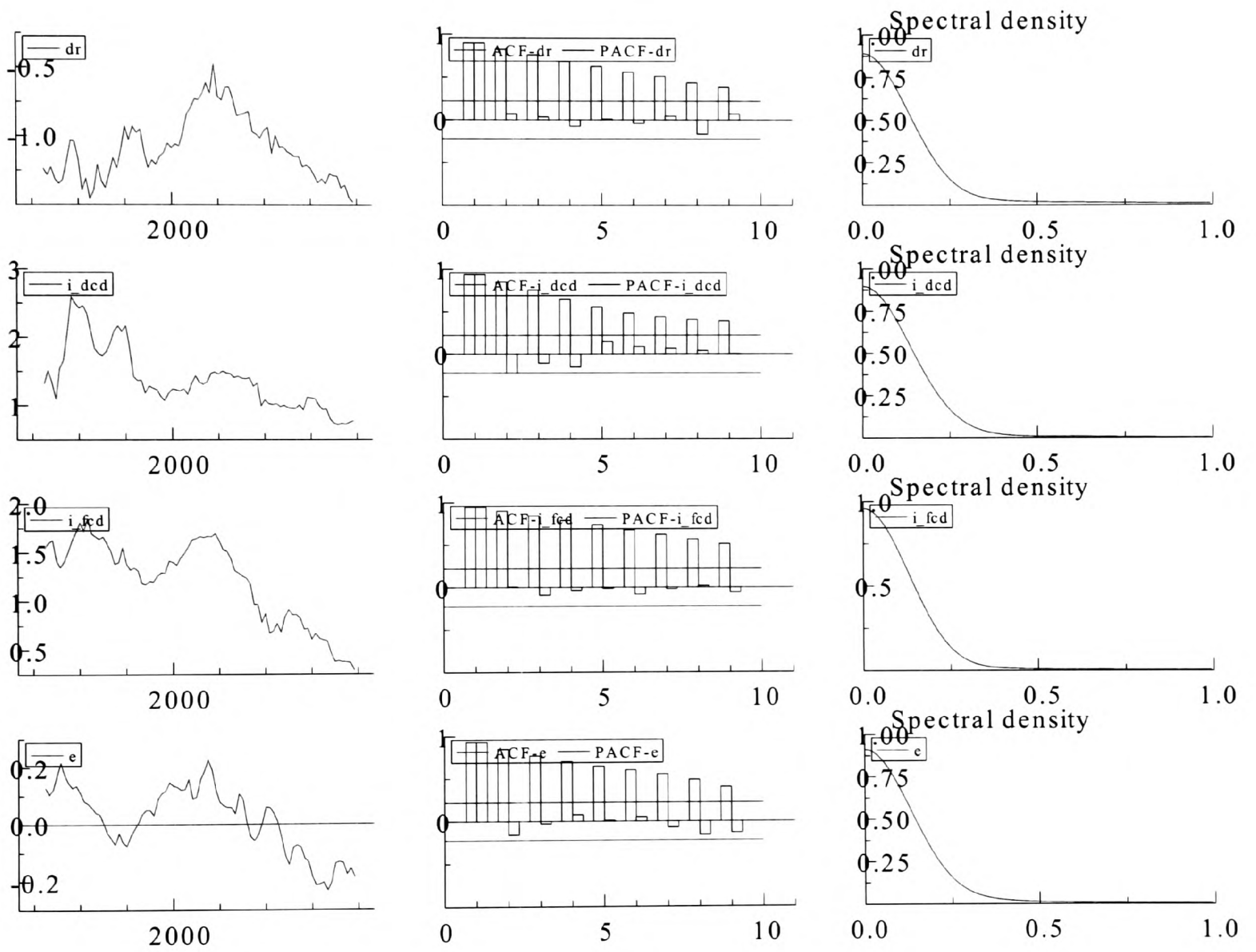
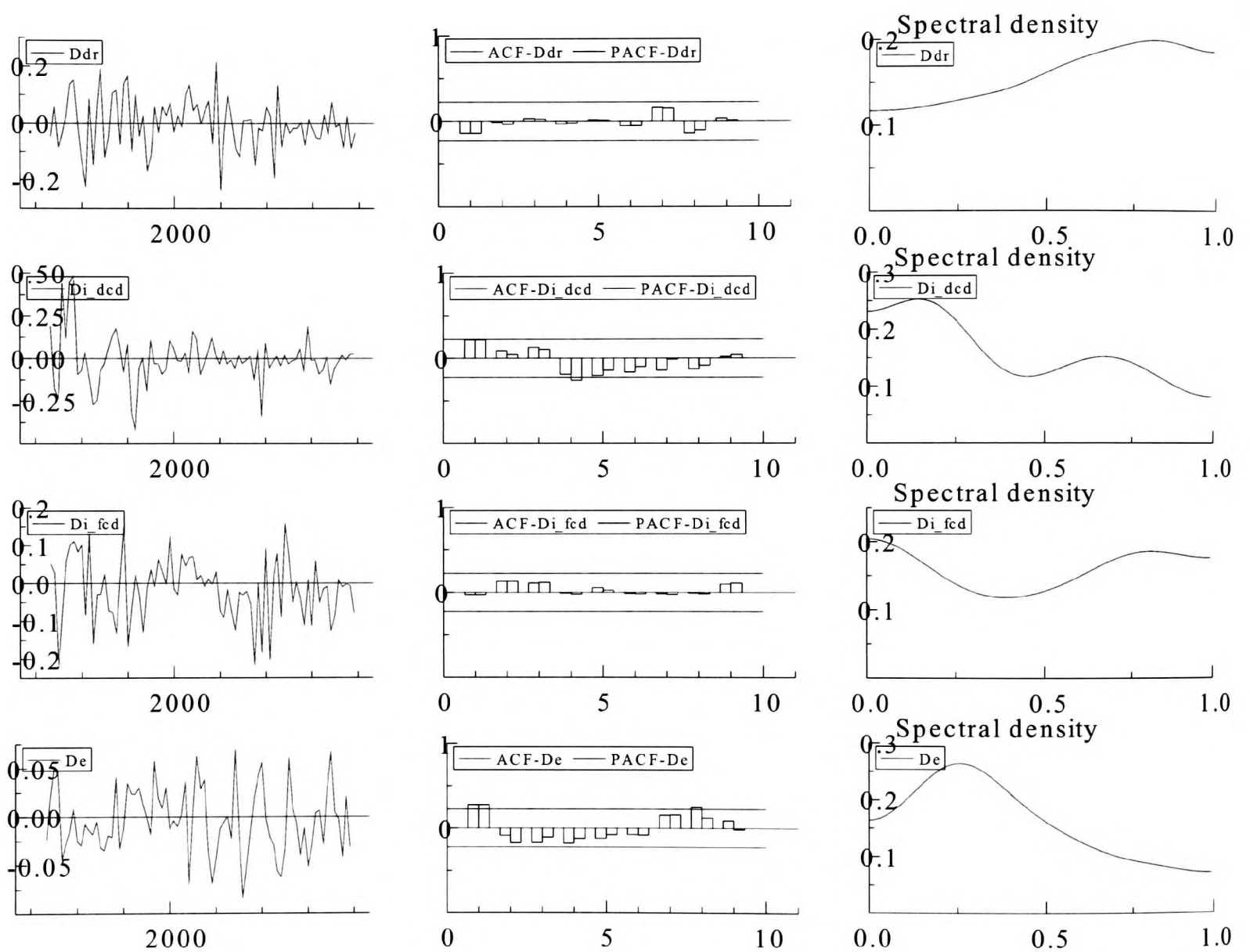


Figure 4-7: Estonia: Time Series Characteristics of Variables in Levels



**Figure 4-8: Estonia: Time Series Characteristics of Variables in First Differences**



**Figure 4-9:** Lithuania: Time Series Characteristics of Variables in Levels

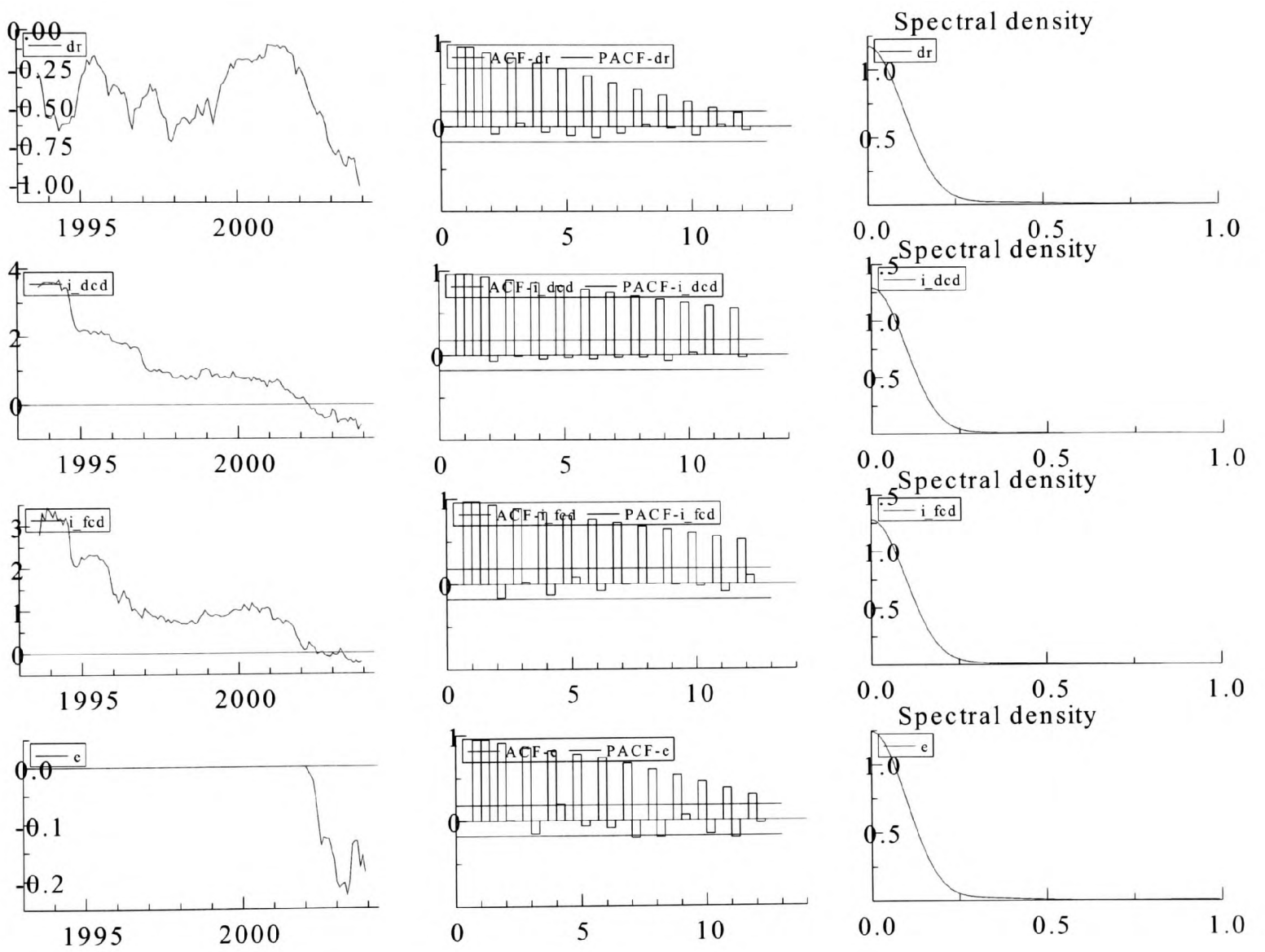


Figure 4-10: Lithuania: Time Series Characteristics of Variables in First Differences

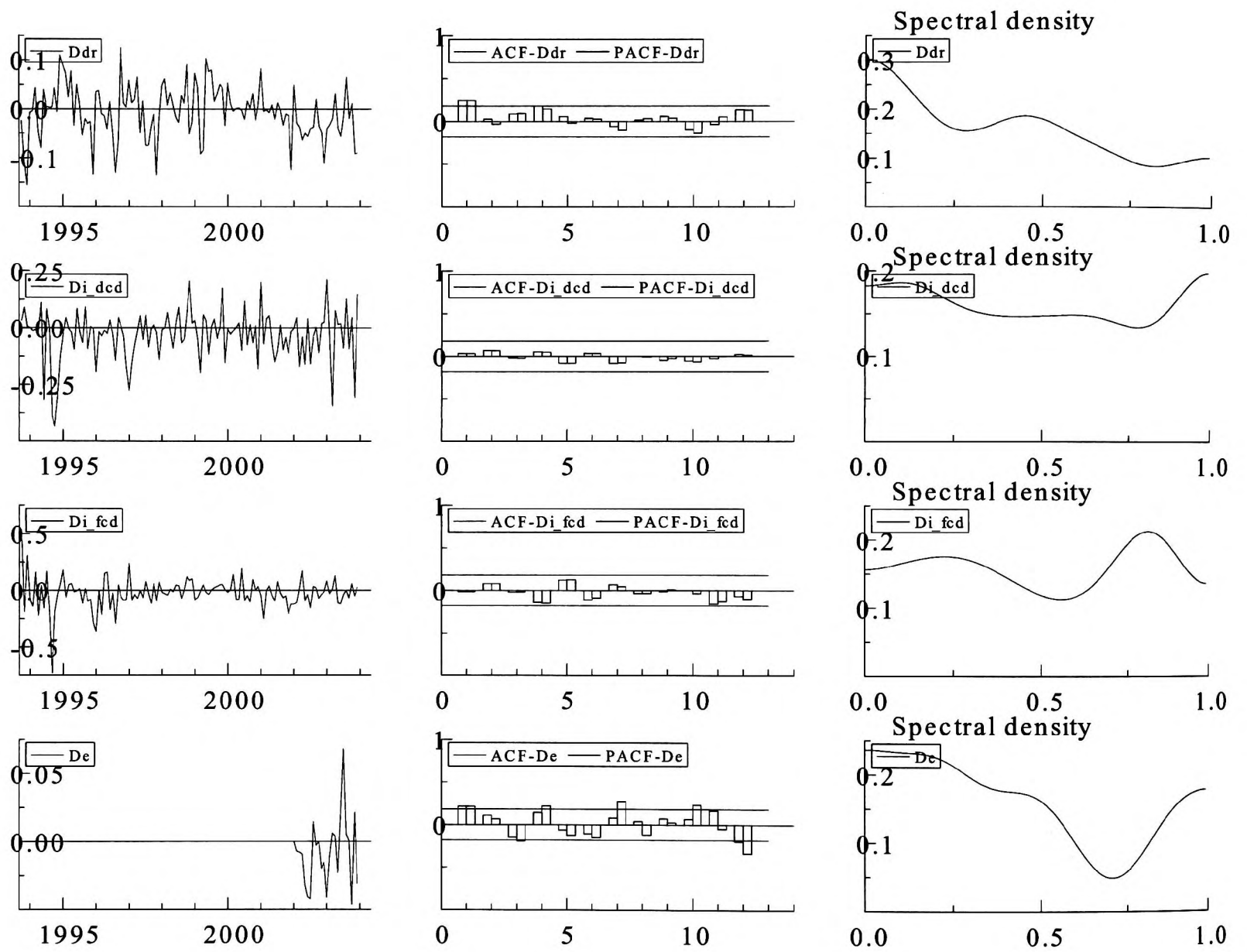
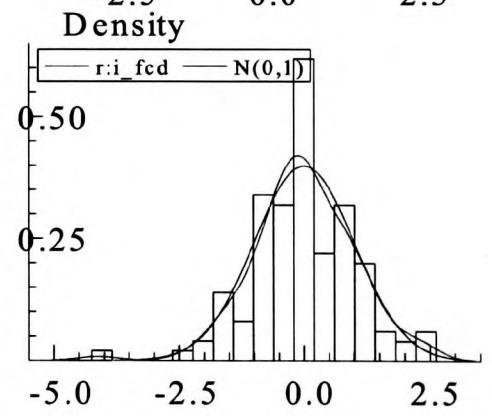
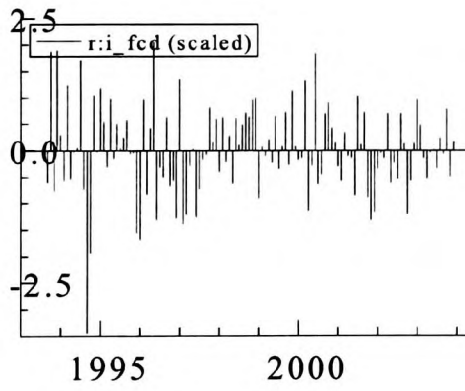
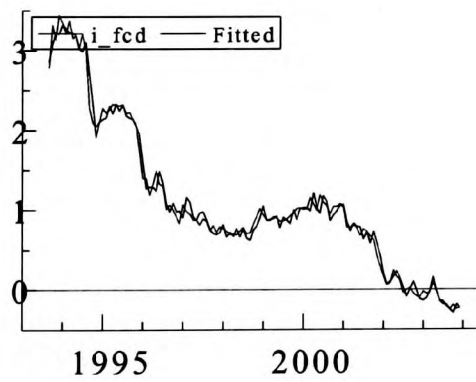
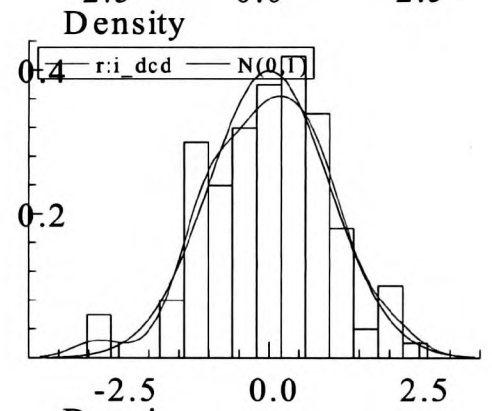
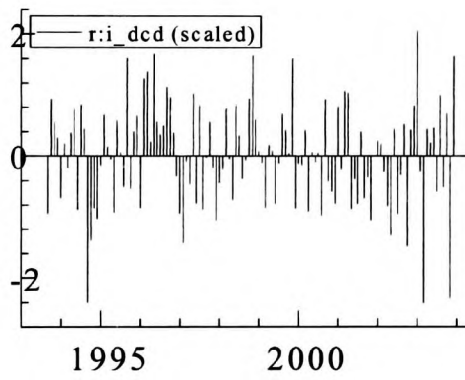
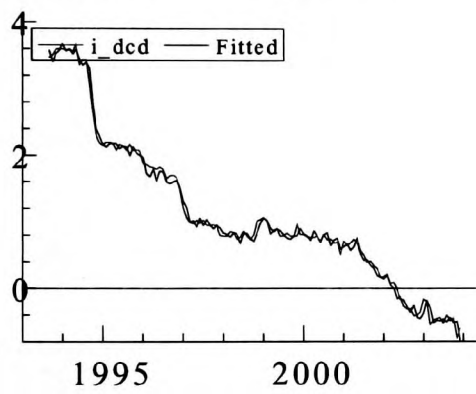
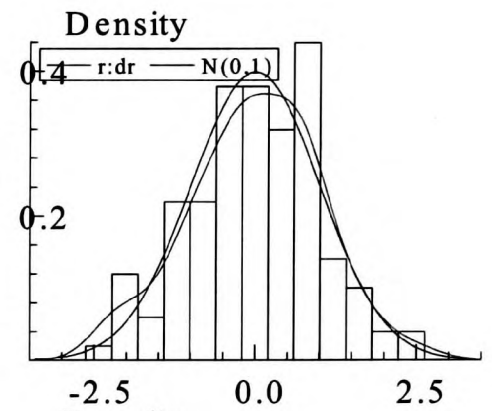
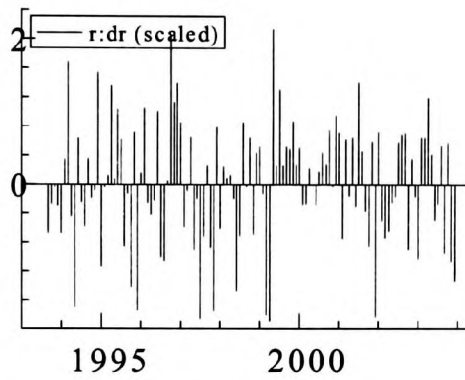
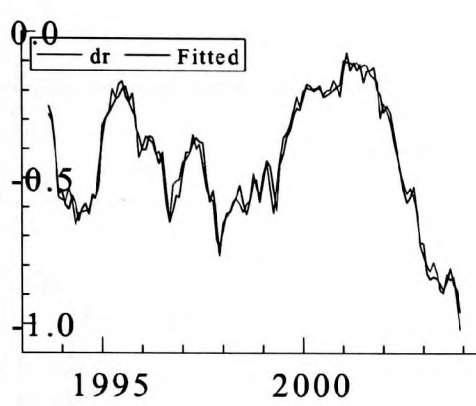


Figure 4-11: Lithuania: UVAR(8)



**Figure 4-12:** Lithuania: Recursive Estimates of Elements of  $\beta$

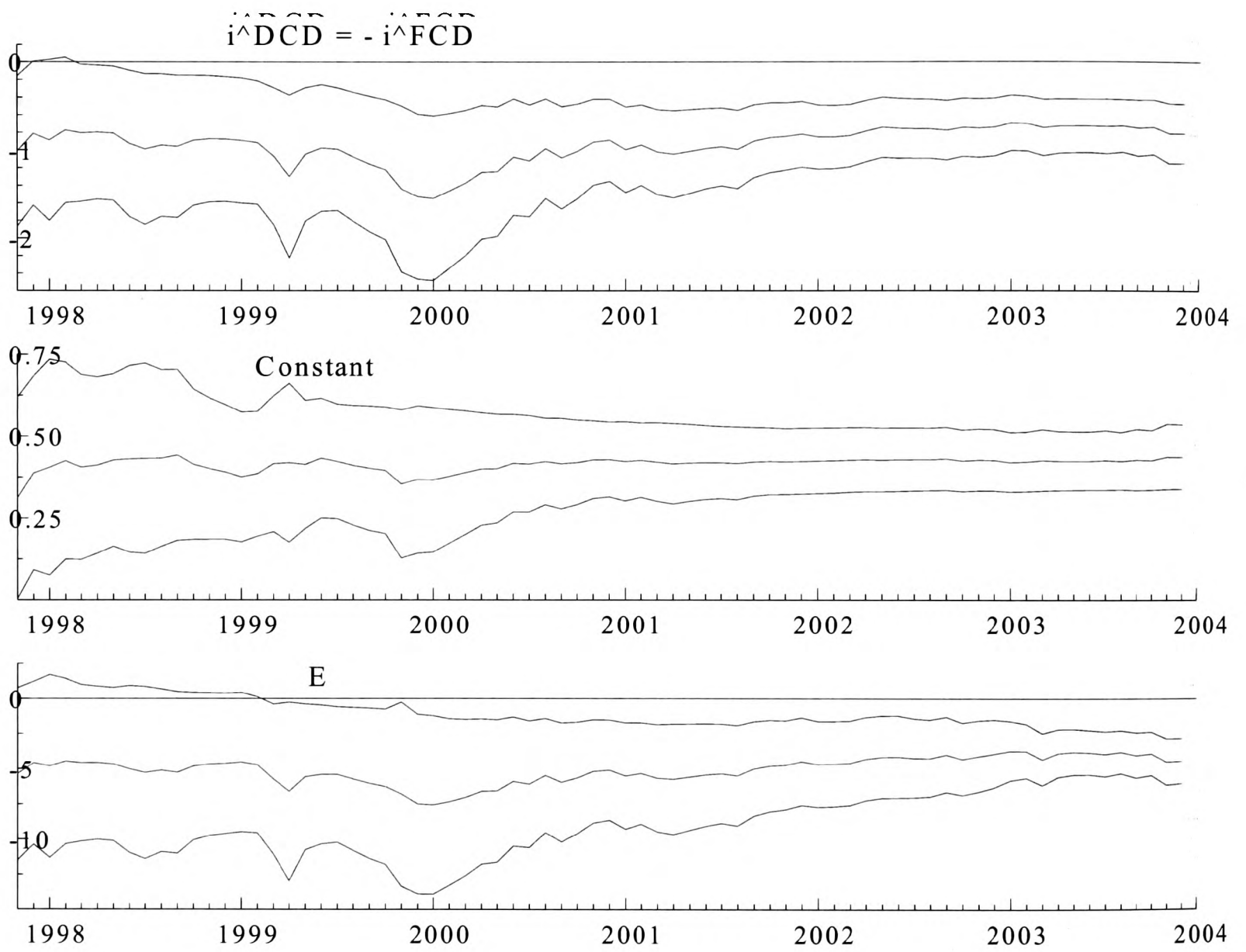
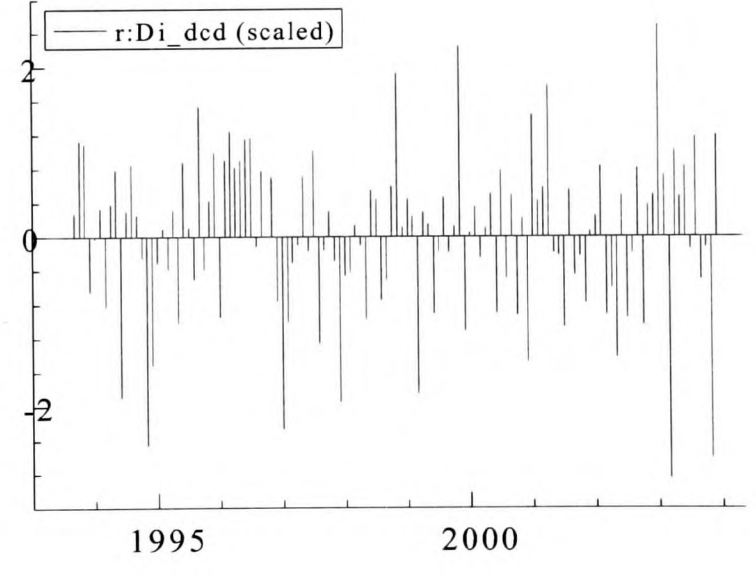
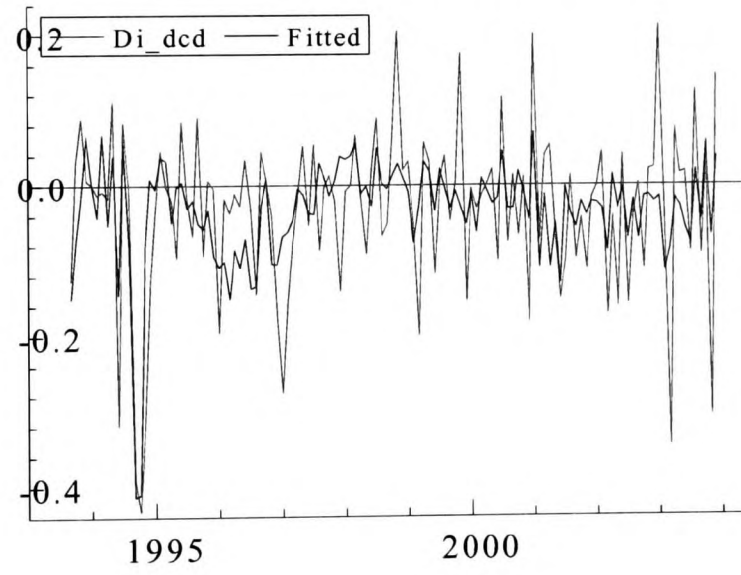
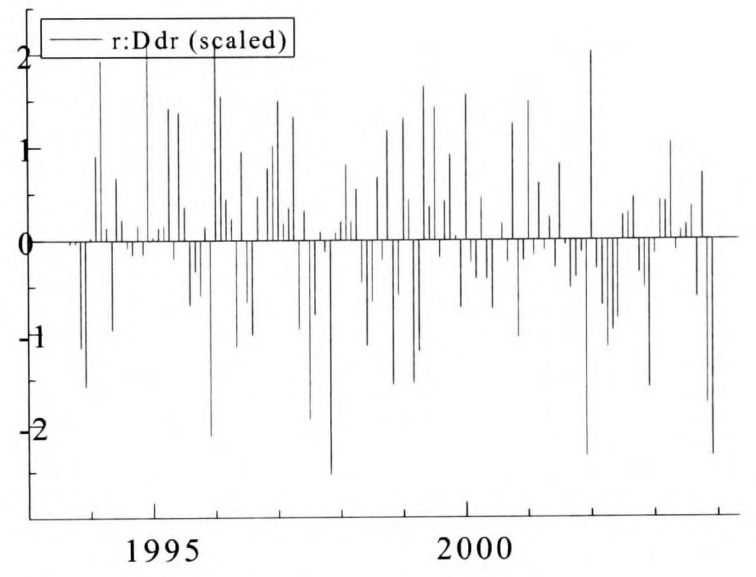
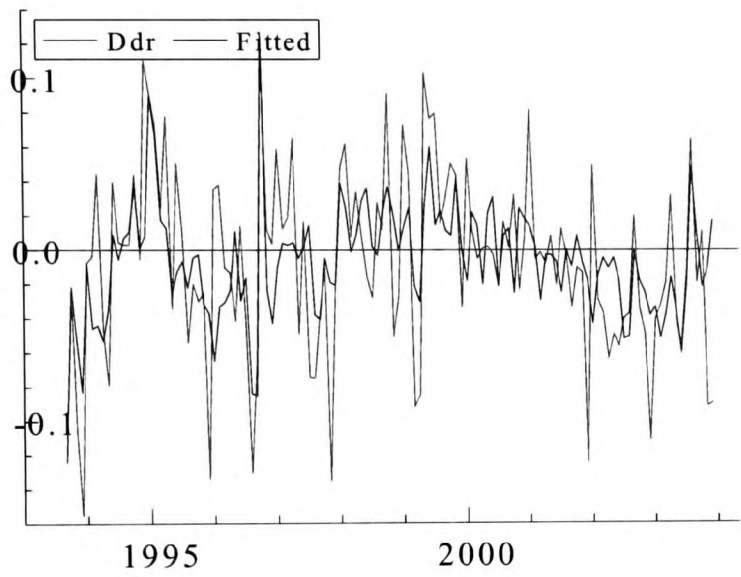


Figure 4-13: Lithuania: PVECM(7)



**Figure 4-14:** Lithuania: Stability of Estimated Coefficients of PVECM(7)

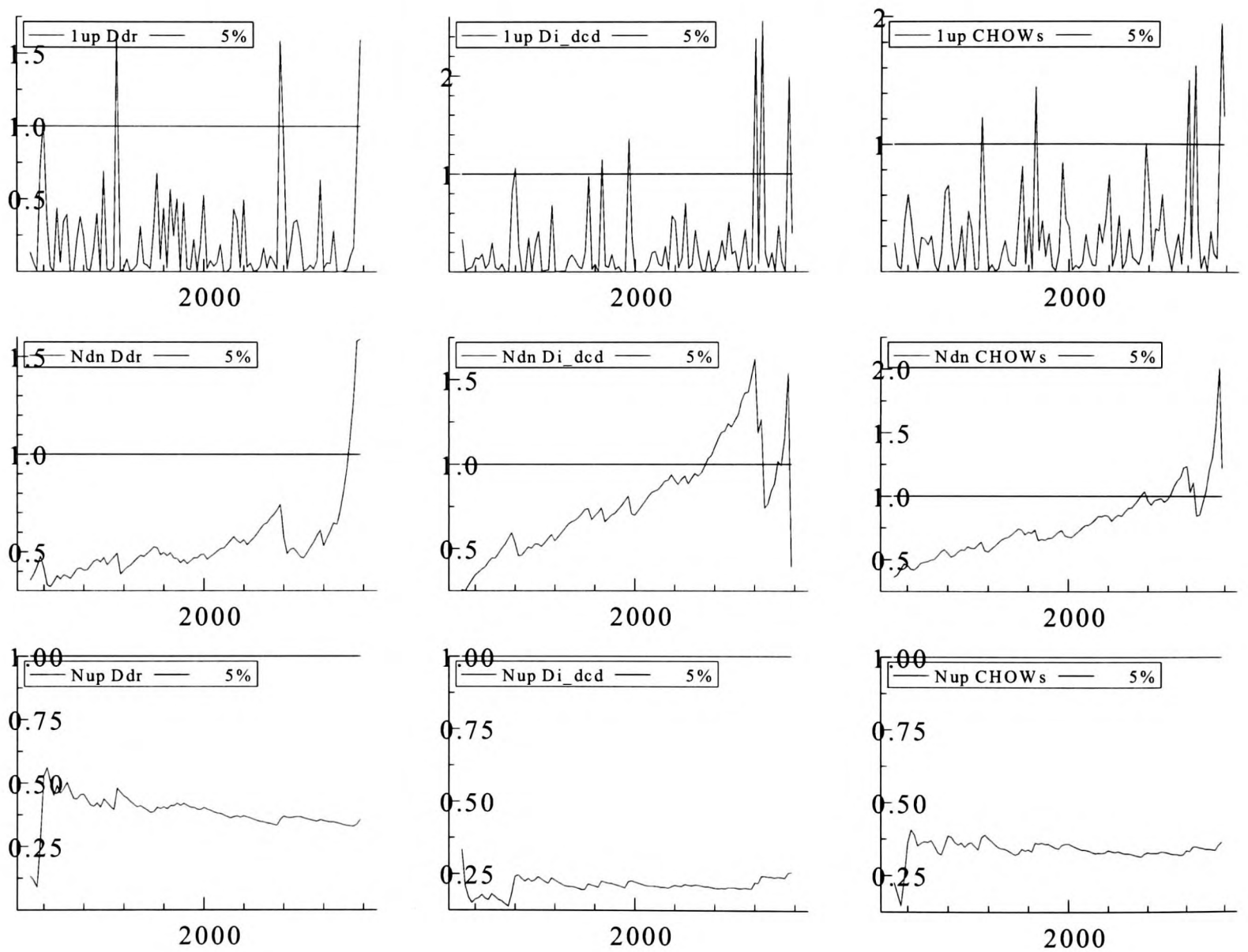
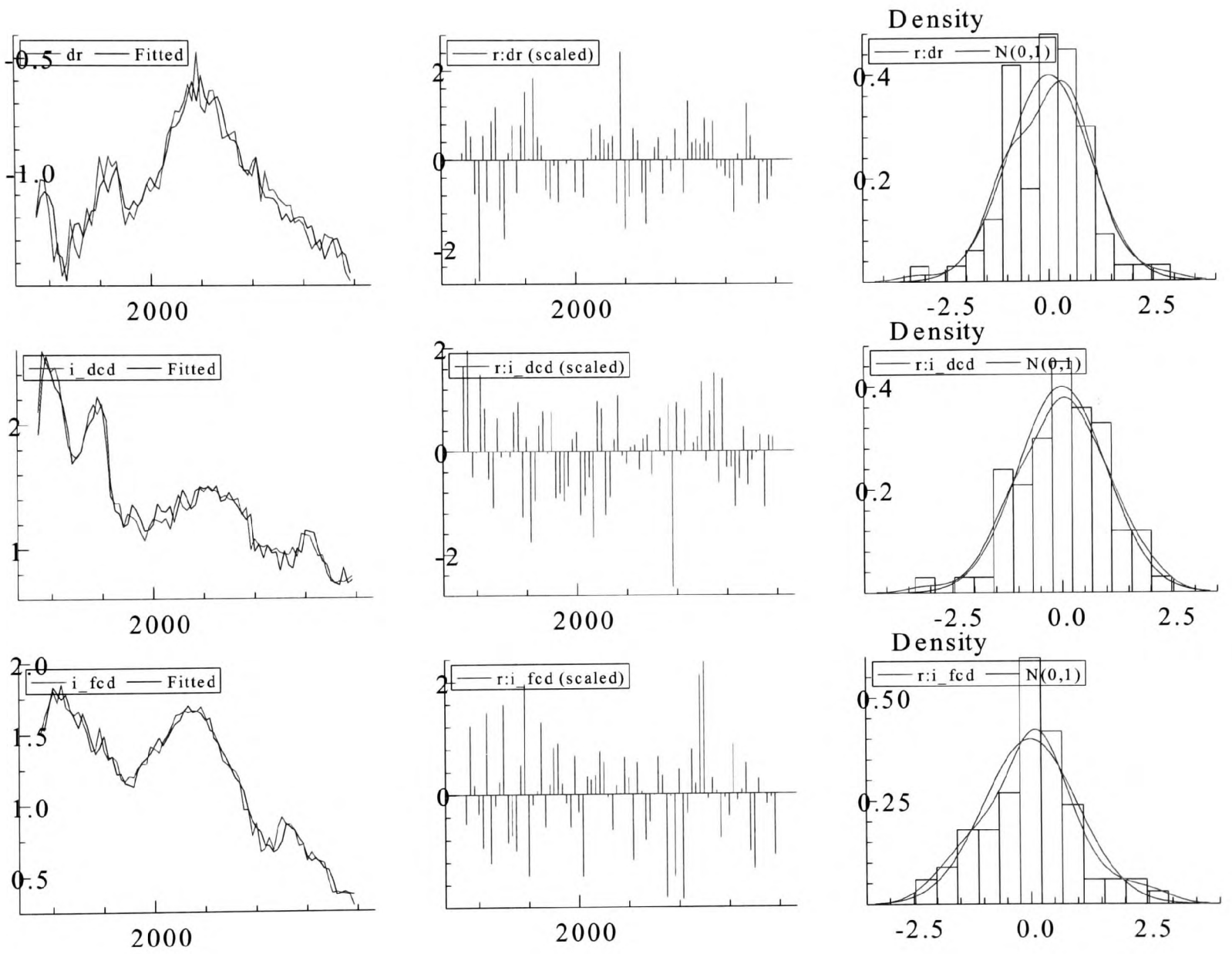


Figure 4-15: Estonia: UVAR(6)



**Figure 4-16:** Estonia: Recursive Estimates of Elements of  $\beta$

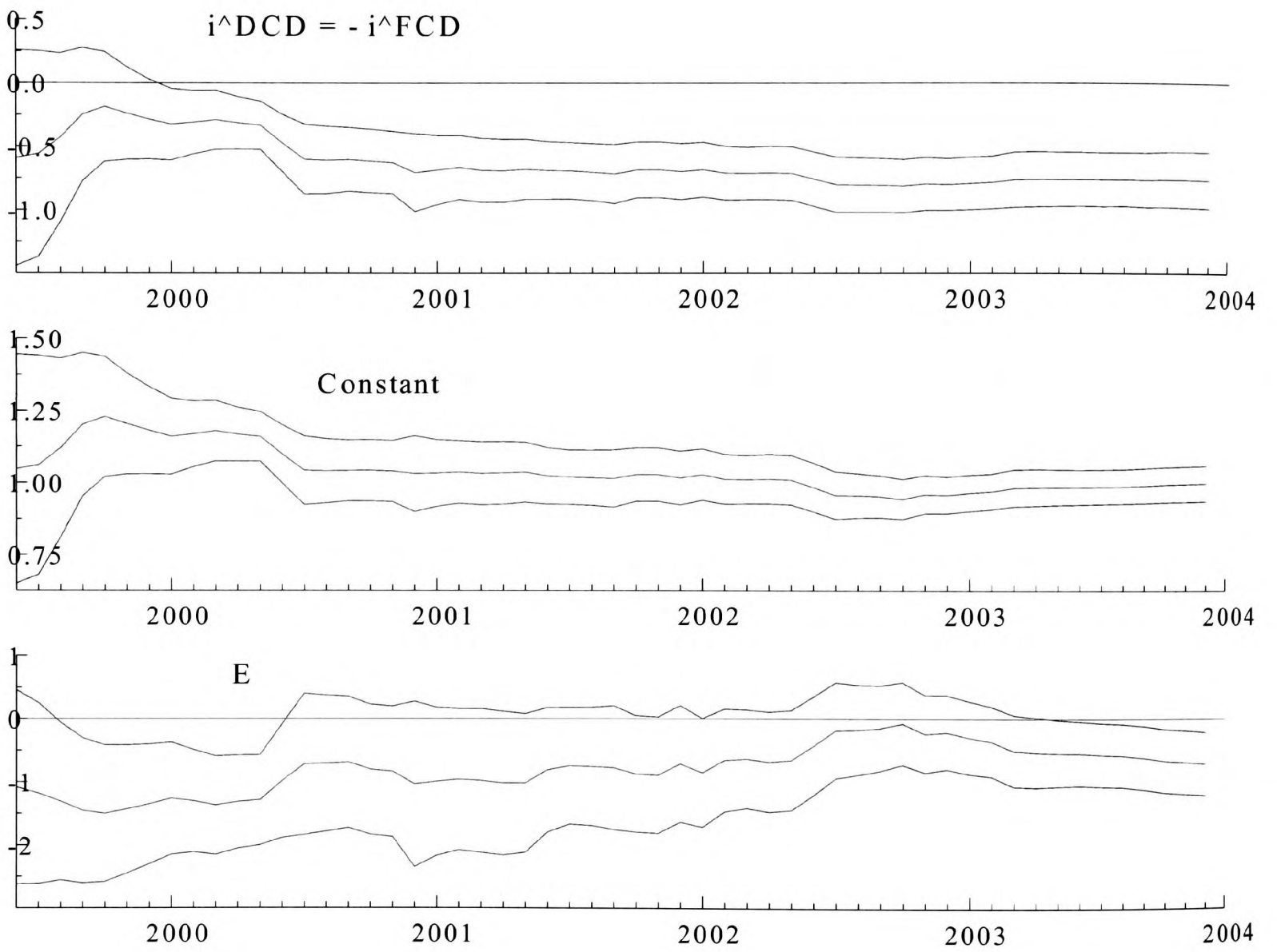
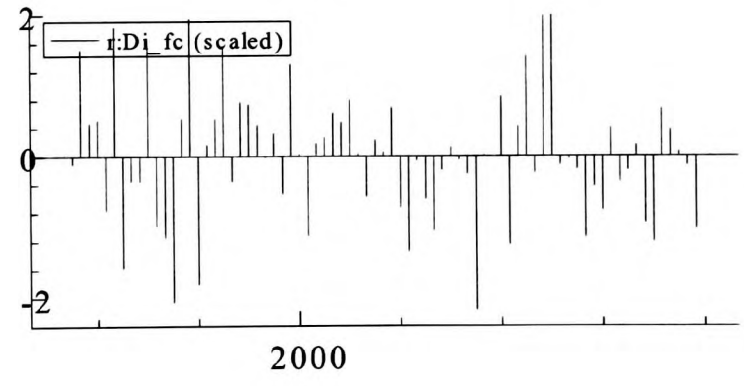
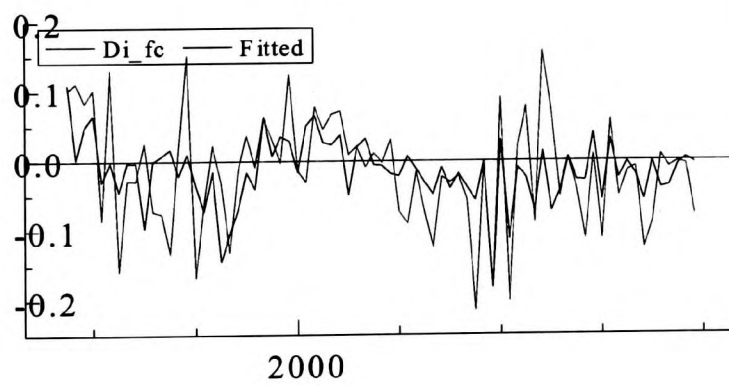
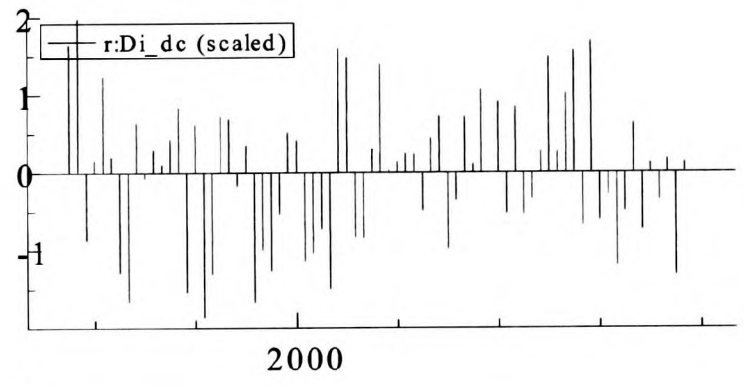
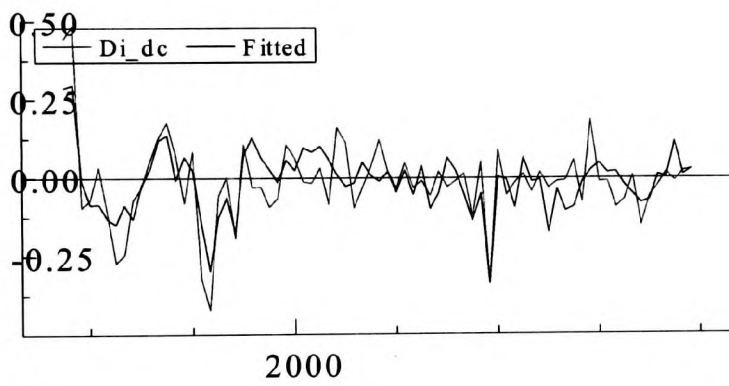
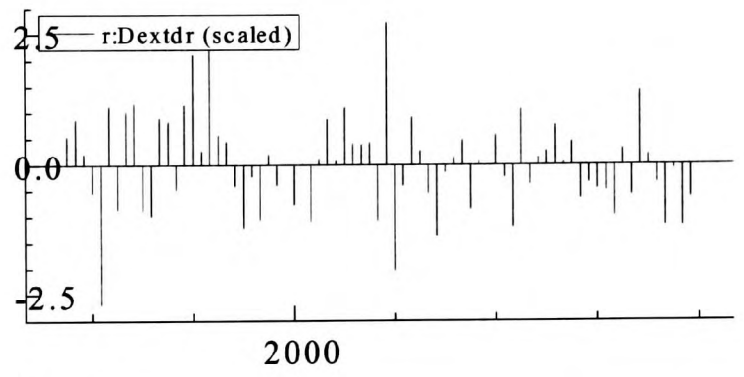
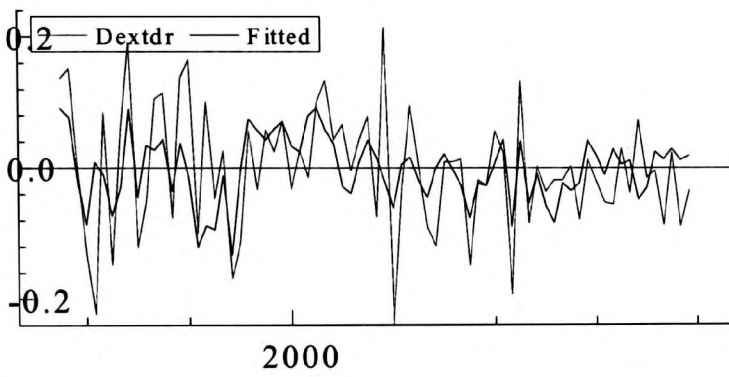
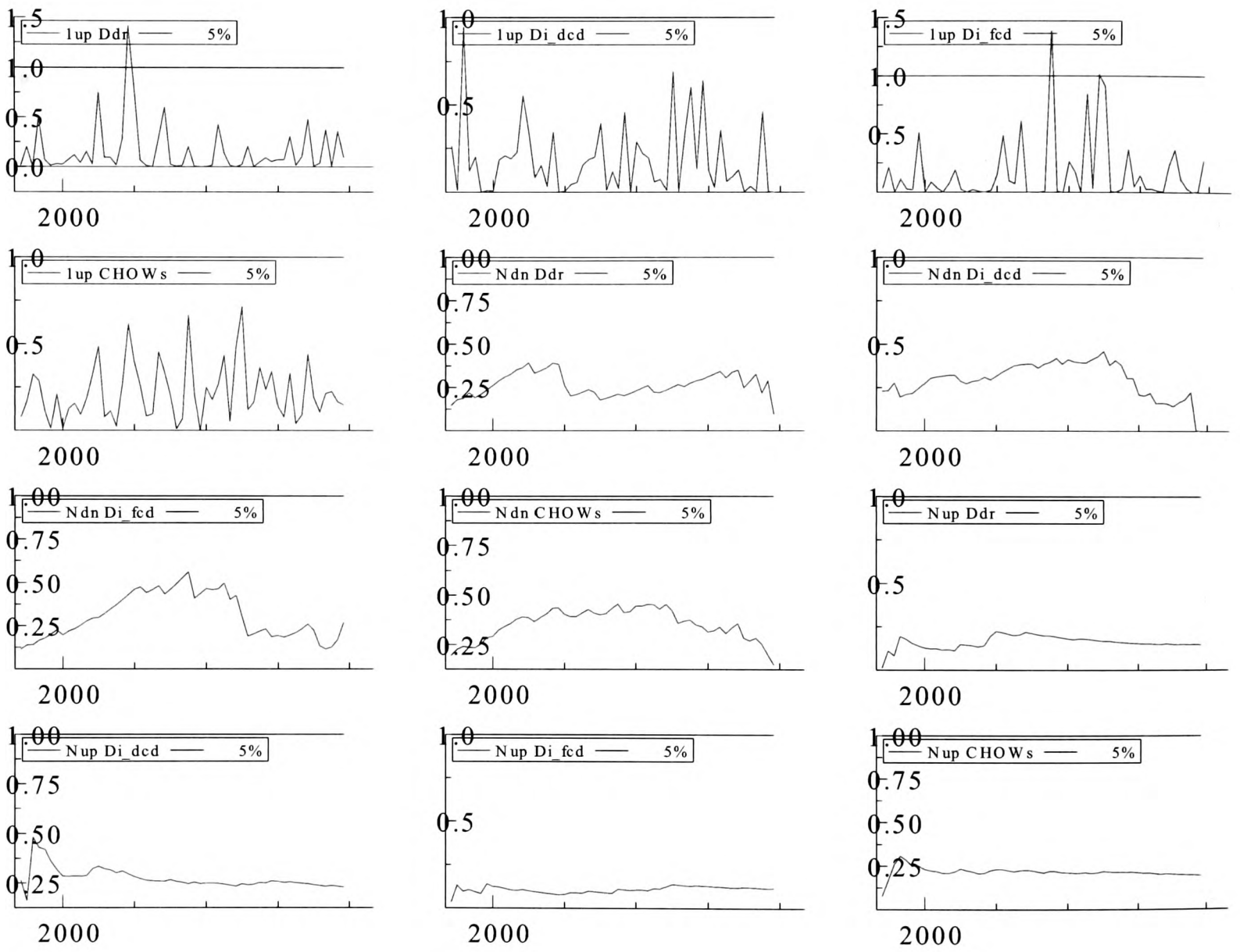


Figure 4-17: Estonia: PVECM(5)



**Figure 4-18: Estonia: Stability of Estimated Coefficients of PVECM(5)**



# Bibliography

- Adam, C., Ndulu, B., and Sowa, N. K. (1996) "Liberalisation and Seignorage Revenue in Kenya, Ghana and Tanzania", *Journal of Development Studies*, 32(4), 531-53.
- Akçay, O. C., Alper, C. E., and Karasulu, M. (1997) "Currency Substitution and Exchange Rate Instability: The Turkish Case", *European Economic Review*, 41, 827-53.
- Akerlof, G. (1970) "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, 84, 488-500.
- Alami, T. (2001) "Currency Substitution versus Dollarization: A Portfolio Balance Model", *Journal of Policy Modelling*, 23, 473-79.
- Alchian, A. (1977) "Why Money?", *Journal of Money, Credit, and Banking*, 9, 133-40.
- Allen, F., and Gale, D. (2000) "Financial Contagion", *Journal of Political Economy*, 108(1), 1-33.
- Alonso-Gama, P., Fabrizio, S., Kramarenko, V., and Wang, Q. (2002) "Lithuania: History and Future of the Currency Board Arrangement". IMF Working Paper 02/127. Washington, D.C.: International Monetary Fund.
- Aristotle (1932) *Politics*. London: William Heinemann Ltd.
- Arrow, K. and Debreu, G. (1954) "Existence of an Equilibrium for a Competitive Economy", *Econometrica*, 12, 265-90.
- Arteta, C. Ó. (2003) "Are Financially Dollarized Countries More Prone to Costly Crises?" Board of Governors of the Federal Reserve System International Finance Discussion Paper No.460. Washington, D.C.: Board of Governors of the Federal Reserve System.
- Aukutsionek, S. (1998) "Industrial Barter in Russia", *Communist Economies and Economic Transformation* 10(2), 179-88.
- Babić, A. (2000) "The Monthly Transaction Demand for Money in Croatia", Croatian National Bank Working Paper W-5. Zagreb: Croatian National Bank.
- Bahmani-Oskooee, M. and Techaratenachai, A. (2001) "Currency Substitution in Thailand", *Journal of Policy Modeling*, 23, 141-5.
- Baliño, T., Bennett, A., and Borzenstein, E. (1999) Monetary Policy in Dollarized Economies. IMF Occasional Paper 171. Washington, D.C.: International Monetary Fund.
- Banerjee, A. and Maskin, E. (1996) "A Walrasian Theory of Money and Barter", *Quarterly Journal of Economics*, 111(4), 955-1005.
- Barro, R. and Gordon, D. (1983) "A Positive Theory of Monetary Policy in a Natural Rate Model", *Journal of Political Economy*, 91, 589-610.

- Baumol, W. (1952) "The Transactions Demand for Cash", *Quarterly Journal of Economics*, 67(4), 545-56.
- Beck, T., and Levine, R. (2002) "Industry Growth and Capital Allocation: Does Having a Market- or Bank-Based System Matter?", *Journal of Financial Economics*, 64(2), 147-80.
- Beck, T., Levine, R., and Loyaza, N. (2000) "Finance and the Sources of Growth", *Journal of Financial Economics*, 58(1-2), 261-300.
- Biais, B., and Gollier, C. (1997) "Trade Credit and Credit Rationing", *Review of Financial Studies* 10(4), 903-37.
- Bogetić, Ž. (2000) "Full Dollarization: Fad or Future?", *Challenge*, 43(2), 17-48.
- Bojnec, Š. (2002) "Payments, Insolvency and Finance during Economic Transformation: Slovenia on the Way to European Union Accession", *Europe-Asia Studies*, 54(2), 277-97.
- Bordo, M. (2003) "Exchange Rate Regime Choice in Historical Perspective". IMF Working Paper 03/160. Washington, D.C.: International Monetary Fund.
- Beaugrand, P. (2003) "Overshooting and Dollarization in the Democratic Republic of Congo". IMF Working Paper 03/115. Washington, D.C.: International Monetary Fund.
- Berg, A., and Borzenstein, E. (2000) "The Choice of Exchange Rate Regime and Monetary Target in Highly-Dollarized Economies", *Journal of Applied Economics*, 3(2), 285-324.
- Brittain, B. (1981) "International currency substitution and the apparent instability of velocity in some Western European economies and in the United States", *Journal of Money, Credit and Banking* 13, 135-55.
- Brana, S. and Maurel, M. (1999) "Barter in Russia: Liquidity Shortage versus Lack of Restructuring", CEPR Working Paper 2258. London: CEPR.
- Broecker, T. (1990) "Credit-Worthiness Tests and Interbank Competition", *Econometrica* 58(2), 429-52.
- Broda, C., and Levy Yeyati, E. (2003) Endogenous Deposit Dollarization. Federal Reserve Bank of New York Staff Report No. 160. New York: Federal Reserve Bank of New York.
- Buiter, W. (2005) "New Developments in Monetary Economics: Two Ghosts, Two Eccentricities, a Fallacy, a Mirage and a Mythos", *Economic Journal* 115(502), C1-C31.
- and Grafe, C. (2002) "Anchor, Float, or Abandon Ship: Exchange Rate Regimes for Accession Countries". CEPR Discussion Paper No.3184. London: Centre for Economic Policy Research.
- Burkart, M., and Ellingsen, T. (2004) "In-Kind Finance", *American Economic Review* 94(3), 569-90.

- Burnside, C., Eichenbaum, M., and Rebelo, S. (2001) "Hedging and financial fragility in fixed exchange rate regimes", *European Economic Review*, 45, 1151-93.
- Caballero, R., and Krishnamurthy, A. (2002) "A Dual Liquidity Model for Emerging Markets", *American Economic Review* 92(2), 33-37.
- Calvo, G. and Reinhart, C. (2002) "Fear of Floating", *Quarterly Journal of Economics*, 117(2), 379-408.
- Calvo, G., and Vegh, C. (1996) "From Currency Substitution to Dollarization and Beyond: Analytical and Policy Issues", pp.153-75 in Calvo, G. (1996) *Money, Exchange Rates, and Output*, Cambridge, MA: MIT Press.
- Carlin, W., Fries, S., Schaffer, M., and Seabright, P. (2000) "Barter and Non-Monetary Transactions in Transition Economies", in Seabright, P. ed. *The Vanishing Rouble: Barter Networks and Non-Monetary Transactions in Post-Soviet Societies*, Cambridge, Cambridge University Press.
- Carruth, A., and Sanchez-Fung, J. R. (2001) "Money Demand in the Dominican Republic", *Applied Economics*, 32(11), 1439-49.
- Catão, L. and Terrones, M. (2000) "Determinants of Dollarization: the Banking Side". IMF Working Paper 00/146. Washington, D.C.: International Monetary Fund.
- Catán, T. (2002) "Argentina snowed under by paper IOUs: Pesos or pacificos? A dizzying array of quasi-currencies now fill up the tills", *Financial Times*, April 11, 2002.
- Cavalcanti, R. de O. and Wallace, N. (1999) "Inside and Outside Money as Alternative Media of Exchange", *Journal of Money, Credit and Banking*, 31(3), 443-457.
- Chari, V. V., and Jagannathan, R. (1988) "Banking Panics, Information, and Rational Expectations Equilibrium", *Journal of Finance*, 43(3), 749-61.
- Chaudhri, E., and Hakura, D. (2001) "Exchange Rate Pass-Through to Domestic Prices: Does the Inflationary Environment Matter?". IMF Working Paper 01/194. Washington, D.C.: International Monetary Fund.
- Cipolla, C. M. (1956) *Money, Prices, and Civilisation in the Mediterranean World*. Princeton, NJ: Princeton University Press.
- Clarida, R., Gali, J., and Gertler, M. (1998) "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory". NBER Working Paper 6442. Cambridge, MA: National Bureau of Economic Research.
- Clower, R. W. (1967) "A Reconsideration of the Microfoundations of Monetary Theory", *Western Economic Journal*, 6(1), 1-9.
- Cohen, (1992) *Athenian Economy and Society: a Banking Perspective*. Princeton: Princeton University Press

- Commander, S., and Mummsen, C. (2000) "Non-monetary Transactions in Russia", in Paul Seabright, ed. (2000) *The Vanishing Rouble: Barter Networks and Non-monetary Transactions in Post-Soviet Societies*. Cambridge: CUP.
- Commander, S., Dolinskaya, I., and Mumssen, C. (2002) "The Determinants of Barter in Russia: an Empirical Analysis", *Journal of Development Economics* 67, 275-307.
- Coricelli, F. (2002) "Exchange Rate Policy During Transition to the European Monetary Union: The Option of Euroization", *Economics of Transition* 10(2), 405-17.
- Cowen, D., Unterberdoerster, O., Thacker, N., and Kiroshita, N. (2002) Vietnam: Selected Issues and Statistical Appendix. IMF Country Report 02/05. Washington, D.C.: International Monetary Fund.
- Cuddington, J. T. (1983) "Currency substitutability, capital mobility and money demand", *Journal of International Money and Finance*, 2, 111-33.
- Cuthbertson, K., and Bredin, D. (2001) "Money Demand in the Czech Republic Since Transition", *Journal of Policy Reform*, 4(4), 271-90.
- Dalton, G. (1982) "Barter", *Journal of Economic Issues* 16(1), 181-90.
- De Blasio, G. (2003) "Does Trade Credit Substitute for Bank Credit? Evidence from Firm-Level Data", IMF Working Paper 03/166. Washington, D.C.: IMF.
- De la Torre, A., Levy Yeyati, E., and Schmukler, S. L. (2002) "Financial Globalisation: Unequal Blessings". World Bank Policy Research Working Paper 2903. Washington, D.C.: The World Bank.
- (2003) "Living and Dying with Hard Pegs: The Rise and Fall of Argentina's Currency Board", *Economia* 5(2), 43-107.
- Demirgüç-Kunt, A. and Maksimović, V. (2002) "Firms as Financial Intermediaries: Evidence from Trade Credit Data". Mimeo, The World Bank.
- De Nicoló, G., Honohan, P., and Ize, A. (2003) "Dollarisation of the Banking Sector: Good or Bad?" World Bank Policy Research Working Paper 3116. Washington, DC: The World Bank.
- De Roover, R. (1954) "New Interpretations of the History of Banking", *Cahiers de l'Histoire du Monde*, 1954, p.38-76.
- (1974) *Business, Banking, and Economic Thought in Late Mediaeval and Early Modern Europe*. Chicago: University of Chicago Press.
- De Zamaróczy, M. and Sa, S. (2002) "Macroeconomic Adjustment in a Highly Dollarized Economy: The Case of Cambodia". IMF Working Paper 02/92. Washington, D.C.: International Monetary Fund.
- Demirgüç-Kunt, A., and Levine, R. (2001) "Bank-Based and Market-Based Financial Systems: Cross-Country Comparisons", in Demirgüç-Kunt, A., and Levine, R., edd.

*Financial structure and economic growth: a cross-country comparison of banks, markets, and development*, Cambridge and London: MIT Press.

Diamond, D. (1984) "Financial Intermediation and Delegated Monitoring", *Review of Economic Studies* 51(3), 393-414.

— and Dybvig, P. (1983) "Bank runs, deposit insurance, and liquidity", *Journal of Political Economy*, 91(3), 401-19.

Doornik, J. (2003) "Asymptotic tables for cointegration tests based on the Gamma-distribution approximation". Mimeo: Nuffield College, Oxford.

Dornbusch, R. (1976) "Expectations and Exchange Rate Dynamics", *Journal of Political Economy*, 84(December), 1161-76.

Dowd, K., and Greenaway, D. (1993) "Currency competition, network externalities and switching costs: towards an alternative view of optimum currency areas", *Economic Journal*, 103, 1180-89.

Dukić, L., and Šošić, V. (2001) "Nemonetarne transakcije u Hrvatskoj". Mimeo: Croatian National Bank.

Edwards, S. and Magendzo, I. I. (2003) "Dollarization and economic performance: what do we really know?", *International Journal of Finance and Economics*, 8, 351-63.

Eichengreen, B., and Hausman, R. (1999) "Exchange Rates and Financial Fragility", NBER Working Paper 7418. Cambridge, MA: National Bureau of Economic Research.

Einzig, P. (1966) *Primitive Money in its Ethnological, Historical, and Economic Aspects*. London: Pergamon Press.

Ellingsen, T. (2000) "Payment in Kind". Mimeo: Stockholm School of Economics.

Engineer, M. and Bernhardt, D. (1991) "Money, Barter, and the Optimality of Legal Restrictions", *Journal of Political Economy*, 99(4), 743-70.

Erjavec, N., and Cota, B. (2001) "The Stability of Money Demand in Croatia in Poststabilization Period", *Ekonomski Pregled* 52(9-10), 1149-1171.

Favero, C. (2004) *Applied Macroeconometrics*. Oxford: Oxford University Press.

Feige, E., Faulend, M., Šonje, V., and Šošić, V. (2000) "Currency Substitution, Unofficial Dollarization, and Estimates of Foreign Currency Held Abroad: The Case of Croatia". Paper presented at the Sixth Dubrovnik Economic Conference, Dubrovnik, June 28-29, 2000.

— and Dean, J. (2002) "Dollarization and Euroization in Transition Countries: Currency Substitution, Asset Substitution, Network Externalities and Irreversibility". Paper presented at the Fordham University International Conference on Euro- and Dollarization: Forms of Monetary Union in Integrating Regions, New York, April 5-6, 2002.

- Ferguson, N. (2001) *The Cash Nexus: Money and Power in the Modern World, 1700-2000*. London : Allen Lane.
- Ferris, J. (1981) "A Transactions Theory of Trade Credit Use", *Quarterly Journal of Economics*, 94(2), 243-70.
- Firth, R. ed. (1967) *Themes in Economic Anthropology*. London: Tavistock.
- Fischer, S. (1982) "Seignorage and the Case for National Money", *Journal of Political Economy*, 90(2), 295-313.
- Fisman, R., and Love, I. (2002) "Trade Credit, Financial Intermediary Development, and Industry Growth", NBER Working Paper 8960.
- Flannery, M. (1996) "Financial Crises, Payment System Problems, and Discount Window Lending", *Journal of Money, Credit, and Banking*, 28(4), 804-24.
- Fleming, M. (1962) "Domestic Financial Policies under Fixed and Floating Exchange Rates", IMF Staff Papers, 9, 369-80.
- Frank, M., and Maksimovic, V. (1998) "Trade Credit, Collateral, and Adverse Selection". Mimeo: The World Bank.
- Frankel, H. (1977) *Money: Two Philosophies. The Conflict of Trust and Authority*. Oxford: Basil Blackwell.
- Freeman, S. (1996a) "Clearinghouse banks and banknote over-issue", *Journal of Monetary Economics*, 38, 101-115.
- (1996b) "The Payments System, Liquidity, and Rediscounting", *American Economic Review*, 86(5), 1126-1138.
- Friedman, M. (1956) "The Quantity Theory of Money: a Restatement", in Friedman, M. ed. *Essays in the Quantity Theory of Money*.
- Fritz-Krockow, B., Verreydt, E., Keller, W., and Sab, R. (2001) Haiti: Selected Issues. IMF Country Report 01/04. Washington, D.C.: International Monetary Fund.
- Furness, W. (1910) *The Island of Stone Money: Uap of the Carolines*. Philadelphia: J.B. Lippincott.
- Gaddy, C., and Ickes, B. (2002) *Russia's Virtual Economy*. Washington, D.C.: Brookings Institution Press.
- Galiani, S., Levy Yeyati, E., and Schargrotsky, E. (2003) "Financial Dollarization and Debt Deflation under a Currency Board". Mimeo: Universidad Torcuato Di Tella.
- Garvy, G. (1977) *Money, Financial Flows, and Credit in the Soviet Union*. Cambridge, MA: NBER.
- Girton, L. and Roper, D. (1981) "Theory and Implications of Currency Substitution", *Journal of Money, Credit, and Banking*, 13(1), 12-30.
- Giovannini, A. and Turtelboom, B. (1994) "Currency substitution", in van der Ploeg, ed. *The Handbook of International Macroeconomics*, Oxford: Basil Blackwell.

- Goldfeld, S. M. (1976) "The Case of the Missing Money", *Brooking Papers on Economic Activity*, 3, 683-730.
- González Anaya, J. A. (2000) "Exchange Rate Pass-through and Partial Dollarization: Is there a Link?". Stanford University Center for Research on Economic Development and Policy Reform Working Paper 81.
- Goodhart, C. (1989) *Money, Information, and Uncertainty*. Basingstoke: Macmillan.
- (1998) "The Two Concepts of Money: Implications for the Analysis of Optimal Currency Areas", *European Journal of Political Economy* 14, 407-432.
- Green, E. (1999) "Money and Debt in the Structure of Payments", *Federal Reserve Bank of Minneapolis Quarterly Review*, Spring 1999, 13-29.
- Grierson, P. (1977) *The Origins of Money*. London: The Athlone Press.
- Grossman, G. (1966) "Gold and the Sword: Money in the Soviet Command Economy", in Rostovsky, H. ed. *Industrialization in Two Systems*, New York: Jack Wiley and Sons.
- (1968) "Introduction" in Grossman, G. ed. *Money and Plan: Financial Aspects of East European Economic Reforms*, Berkeley and Los Angeles: University of California Press.
- Gulde, A.-M., Hoelscher, D., Ize, A., Leone, A., Marston, D., and Moretti, M. (2003) "Dealing with Banking Crises in Dollarized Economies", in Collyns, C. and Kincaid, G. R. edd. "Managing Financial Crises: Recent Experience and Lessons for Latin America", IMF Occasional Paper 217, Washington, D.C.: International Monetary Fund.
- Hansen, B. (1992) "Testing for Parameter Instability in Linear Models", *Journal of Policy Modeling*, 14(4), 517-33.
- Hart, O. (1995) *Firms, Contracts, and Financial Structure*. Oxford: OUP.
- Hayek, F. A. (1976) *Denationalisation of Money: An Analysis of the Theory and Practice of Concurrent Currencies*. IEA Hobart Paper No.70. London: Institute of Economic Affairs.
- Heimonen, K. (2000) "Substituting a Substitute Currency: the Case of Estonia". BOFIT Discussion Paper 11. Helsinki: Bank of Finland.
- Hicks, J. (1935) "A Suggestion for Simplifying the Theory of Money", *Economica* Feb. 1935, 1-19.
- (1989) *A Market Theory of Money*. Oxford: Oxford University Press.
- Hildebrandt, A. (2002) "Too many to fail? Inter-enterprise arrears in transition economies", BOFIT Discussion Paper 11/2002. Helsinki: Bank of Finland.
- Hiyashi, F. and Matsui, A. (1996) "A Model of Fiat Money and Barter", *Journal of Economic Theory*, 68, 111-32.
- Holmström, B. and Tirole, J. (1998) "Private and Public Supply of Liquidity", *Journal of Political Economy*, 106(1), 1-40.

- Honohan, P. and Shi, A. (2003) "Deposit Dollarization and the Financial Sector in Emerging Economies", in J. Hanson, P. Honohan, and G. Majnoni edd., *Globalization and National Financial Systems*, New York, NY: Oxford University Press.
- Honohan, P. and Stiglitz, J. E. (2001) "Robust Financial Restraint", in Caprio, G., Honohan, P. and Stiglitz, J. E. edd. (2001) *Financial liberalization: How far, how fast?* Cambridge: Cambridge University Press.
- Humphrey, C. (2000) "How is barter done? The social relations of barter in provincial Russia", in Seabright ed. (2000), p.259-97.
- Hoover, K. (1996) "Some Suggestions for Complicating the Theory of Money", in S. Pressman, ed. (1996) *Interactions in political economy: Malvern after ten years*, London and New York: Routledge.
- IMF (1998) "Croatia: Selected Issues and Statistical Appendix", IMF Country Report 98/90. Washington, D.C.: International Monetary Fund.
- (2000) "Croatia: Selected Issues and Statistical Appendix", IMF Country Report 00/22. Washington, D.C.: International Monetary Fund.
- (2001) "Croatia: 2000 Article IV Consultations and Request for a Standby Arrangement", IMF Country Report 01/50. Washington, D.C.: International Monetary Fund.
- (2002) "Introductory Remarks on the Role of the IMF Mission in Argentina by Anoop Singh, Director for Special Operations, IMF", Press Briefing, Buenos Aires, April 10, 2002. Available at <http://www.imf.org/external/np/tr/2002/tr020410.htm>.
- (2003) "Capital Markets and Financial Intermediation in the Baltic States". IMF Country Report 03/115. Washington, D.C.: International Monetary Fund.
- Imrohorglu, A. (1996) "International Currency Substitution and Seigniorage in a Simple Model of Money", *Economic Inquiry*, 34(3), 568-78.
- Ingham, G. (2004) *The Nature of Money*. Cambridge: Polity Press.
- Ize, A. and Levy Yeyati, E. (2003) "Financial Dollarization", *Journal of International Economics*, 59, 323-47.
- Ize, A. and Parrado, E. (2002) "Dollarization, Monetary Policy, and the Pass-Through". IMF Working Paper 02/188. Washington, D.C.: International Monetary Fund.
- Jácome, L. (2004), "The Late 1990s Financial Crisis in Ecuador: Institutional Weaknesses, Fiscal Rigidities, and Financial Dollarization at Work". IMF Working Paper 04/12. Washington, D.C.: International Monetary Fund.
- Jevons, W. S. (1875) *Money and the Mechanism of Exchange*. London: Appleton.
- Johansen, S. (1988) "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 231-254.
- (2005) "Interpretation of Cointegrating Coefficients in the Cointegrated Vector Autoregressive Model", *Oxford Bulletin of Economics and Statistics*, 67(1), 93-104.

- Kamin, S. and Ericsson, N. (1993), "Dollarization in Argentina", Board of Governors of the Federal Reserve System International Finance Discussion Paper No.460. New York: Board of Governors of the Federal Reserve System.
- (2003) "Dollarization in Post-Hyperinflationary Argentina", *Journal of International Money and Finance*, 22(2), 185-211.
- Keynes, J. M. (1914) "Review of A. Mitchell Innes' 'What is Money?'" *Economic Journal* 24(95), 421-425.
- Kim, H. ed. (2001) *Money and Its Use in the Ancient Greek World*. Oxford: Oxford University Press.
- King, M. (2004) "The Institutions of Monetary Policy", *American Economic Review*, 94(2), 1-13.
- King, R. G., and Levine, R. (1993) "Finance and Growth: Schumpeter Might Be Right", *Quarterly Journal of Economics*, 108(3), 717-37.
- Kiyotaki, N. and Moore, J. H. (2001) "Evil is the Root of All Money", Clarendon Lectures in Economics, University of Oxford.
- (2002a) "Evil is the Root of All Money", *American Economic Review* 92(2), 62-66.
- (2002b) "Inside Money and Liquidity". Mimeo, London School of Economics.
- Kiyotaki, N. and Wright, R. (1989) "On Money as a Medium of Exchange", *Journal of Political Economy*, 97(4), 927-954.
- Knapp, G. (1924) *The State Theory of Money*. London: The Royal Economic Society.
- Komárek, L. and Melecký, M. (2003) "Currency Substitution in a Transitional Economy with an Application to the Czech Republic", *Eastern European Economics*, 41(4), 72-99.
- Korhonen, I. (1996) "Dollarization in Lithuania", *Review of Economics in Transition* 5, 53-59.
- Kraft, E. (2001) "Foreign Banks in Croatia: Another Look", Croatian National Bank Working Paper W-10. Zagreb: Croatian National Bank.
- (2003) "Monetary Policy Under Dollarisation: the Case of Croatia", *Comparative Economic Studies* 45(3), 256-77.
- Kremers, J. J. M. and Lane, T. D. (1990) "Economic and monetary integration and the aggregate demand for money in the EMS", *IMF Staff Papers*, 37(4), 777-805.
- Krugman, P. (1996) "Are Currency Crises Self-Fulfilling?" in Bernanke, B., S. and Rotemberg, J. J., edd. (1996) *NBER Macroeconomics Annual 1996*. Cambridge, M.A.: MIT Press,
- Lahiri, A. and Vegh, C. (2001) "Living with the Fear of Floating: An Optimal Policy Perspective". NBER Working Paper No. 839. Cambridge, MA: National Bureau of Economic Research.
- Laidler, D. (1993) *The Demand for Money*. New York: Harper Collins.

- Laidler, D. and Rowe, N. (1980) "Georg Simmel's Philosophy of Money: A Review Article for Economists", *Journal of Economic Literature*, 18(1), 97-105.
- Lainela, S., and Sutela, P. (1994) *The Baltic Economies in Transition*. Helsinki: Bank of Finland.
- Leland, H. and Pyle, D. (1977) "Informational Asymmetries, Financial Structure, and Financial Intermediation", *Journal of Finance*, 32(2), 371-87.
- Levy Yeyati, E. (2003) "Financial Dollarization: Where Do We Stand?" Paper prepared for the Conference on Financial Dollarization: Policy Options at the Inter-American Development Bank, Washington, D.C., December 1-2, 2003. Mimeo.
- and Sturzenegger, F. edd. (2003) *Dollarization: Debates and Policy Alternatives*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Lizondo, S., Wolfe, A., Chua, D., Córdoba, J. P., and Richter, A. (2001) Peru: Selected Issues. IMF Country Report 01/51. Washington, D.C.: International Monetary Fund.
- Locke, J. (1980) *Second Treatise of Government*. Indianapolis: Hackett Publishing Company.
- McKinnon, R. I. (1973) *Money and Capital in Economic Development*. Washington, D.C.: Brookings Institution.
- (1982) "Currency substitution and instability in the world Dollar standard", *American Economic Review*, 72(3), 320-33.
- (1996) "Monetary and Exchange-Rate Policies for International Financial Stability: A Proposal" in McKinnon, R. I., (1996) *The Rules of the Game: International Money and Exchange Rates*, Cambridge, M.A. and London: MIT Press.
- and Oates, W. (1966) "The Implications of International Economic Integration for Monetary, Fiscal and Exchange Rate Policy". Princeton Studies in International Finance No. 16. Princeton, N.J.: Princeton University.
- Macleod, H. D. (1882) *The Principles of Economical Philosophy*. London: Longmans, Green, Reader, and Dyer.
- (1891) *The Theory of Credit*. London.
- Madox, T. (1711) *The history and antiquities of the Exchequer of the Kings of England, in two periods: to wit, from the Norman conquest, to the end of the reign of K. John; and from the end of the reign of K. John, to the end of the reign of K. Edward II*. London: printed by John Matthews and sold by Robert Knaplock.
- Makinen, G. E. and Woodward, G. (1986) "Some Anecdotal Evidence Relating to the Legal Restrictions Theory of the Demand for Money", *Journal of Political Economy*, 94(2), 260-65.
- Marin, D., Kaufmann, D., and Gorochofskij, B. (2000) "Barter in transition economies: competing explanations confront Ukrainian data" in Seabright, ed. (2000).
- Markowitz, H. (1952) "Portfolio Selection", *Journal of Finance* 7, 77-91.

- Maurel, M., and Brana, S. (1999) "Barter in Russia: Liquidity Shortage versus Lack of Restructuring", CEPR Working Paper 2258. London: CEPR.
- Mazzaferro, F., Thimann, C., and Winkler, A., (2003) "On the Sustainability of Euroisation/Dollarisation Regimes: How Important Are Fiscal Transfers, Offshore Finance and Tourism Receipts?", *Comparative Economic Studies*, 45(3), 421-36.
- Meltzer, A. (1960) "Mercantile Credit, Monetary Policy, and the Size of Firms", *Review of Economics and Statistics* 42(4), 429-436.
- Menger, C. (1892) "On the Origin of Money", *Economic Journal*, 2(6), 239-255.
- Merton, R. C. (1971) "Optimum Consumption and Portfolio Rules in a Continuous-Time Model", *Journal of Economic Theory*, 3(4), 373-413.
- Miles, M. A. (1978) "Currency substitution, flexible exchange rates and monetary independence", *American Economic Review*, 68, 428-36.
- Millet, (1991) *Lending and Borrowing in Ancient Greece*. Cambridge: Cambridge University Press.
- Mitchell Innes, A. (1913) "What is Money?", *Banking Law Journal* 30, 377-408. Also reprinted in Wray (2004).
- (1914) "The Credit Theory of Money", *Banking Law Journal* 31, 151-168.
- Mizen, P., and Pentecost, E. J. (1994) "Evaluating the Empirical Evidence for Currency Substitution: A Case Study of the Demand for Sterling in Europe", *Economic Journal*, 104(426), 1057-69.
- edd. (1996a) *The Macroeconomics of International Currencies: Theory, Policy and Evidence*. Cheltenham: Edward Elgar.
- (1996b) "Currency Substitution in Theory and Practice", in Mizen and Pentecost edd. (1996a) pp.8-43.
- Mongardini, J., and Mueller, J. (2000) "Ratchet Effects in Currency Substitution: An Application to the Kyrgyz Republic", *IMF Staff Papers*, 47(2), 218-37.
- Muldrew, C. (1998) *The Economy of Obligation: the Culture of Credit and Social Relations in Early Modern England*. London: Macmillan.
- Mundell, R. (1963) "Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates", *Canadian Journal of Economics and Political Science*, 29(4), 475-85
- Murray, J. and Powell, J. (2003) "Dollarization in Canada: Where Does the Buck Stop?" *North American Journal of Economics and Finance* 14, 145-72.
- Murphy, A. (1978) "Money in an Economy Without Banks: The Case of Ireland", *The Manchester School of Economic and Social Studies*, 46(1), 41-50.
- Ng, C. H., Smith, J. K., and Smith, R. L. (1999) "Evidence on the Determinants of Credit Terms Used in Interfirm Trade", *Journal of Finance* 54(3), 1109-29.

- Nilsen, J. H. (2002) "Trade Credit and the Bank Lending Channel", *Journal of Money, Credit, and Banking* 34(1), 226-53.
- Nuti, D. M. (2002) "Costs and Benefits of Unilateral Euroization in Central Europe", *Economics of Transition* 10(2), 419-44.
- OECD (2000) *Baltic States: a Regional Economic Assessment*. OECD Economic Surveys. Paris: OECD.
- Oomes, N. (2003) "Network Externalities and Dollarization Hysteresis in Russia". IMF Working Paper 03/96. Washington, D.C.: International Monetary Fund.
- Peiers, B. M., and Wrase, J. M. (1997) "Dollarization Hysteresis and Network Externalities: Theory and Evidence from an Informal Bolivian Credit Market". Federal Reserve Bank of Philadelphia Working Paper 97/21. Philadelphia, PA: Federal Reserve Bank of Philadelphia.
- Perotti, E. (1998) "Inertial Credit and Opportunistic Arrears in Transition", *European Economic Review* 42(9), 1703-25.
- Peterson, M. and Rajan, R. (1995) "Insiders and Outsiders: the Choice Between Informed and Arm's Length Finance", *Journal of Finance*, 47, 1367-1400.
- (1997) "Trade Credit: Theories and Evidence", *Review of Financial Studies* 10(3), 661-91.
- Pinto, B., Drebensov, V., and Morozov, A. (1999) "Give Growth and Macro Stability in Russia a Chance: Harden Budget Constraints by Dismantling Nonpayments", World Bank Working Paper 2329. Washington, D.C.: World Bank.
- (2000) Dismantling Russia's Nonpayments System: Creating Conditions for Growth, World Bank Technical Paper No.471, Washington, D.C., The World Bank.
- Poser, J. A. (1998) "Monetary Disruptions and the Emergence of Barter in FSU Economies", *Communist Economies and Economic Transformation* 10(2), 157-177.
- Postan, M. M. (1973) *Medieval Trade and Finance*. Cambridge: Cambridge University Press.
- Postlewaite, A., and Vives, X. (1987) "Bank Runs as an Equilibrium Phenomenon", *Journal of Political Economy*, 95(3), 485-91.
- Quiggin, A. H. (1949) *A Survey of Primitive Money: the Beginnings of Currency*. London: Methuen.
- Quispe, Z. (1999) "Una Aproximacion a la Demanda de los Principales Agregados Monetarios en el Peru: Junio, 1991 -- Majo, 1997". La Paz: Central Reserve Bank of Peru.
- Rajan, R. and Zingales, L. (1995) "Financial Dependence and Growth", *American Economic Review*, 88(3), 559-86.
- Reinhart, C., Rogoff, K., and Savastano, M. (2003) "Addicted to Dollars", NBER Working Paper 10015. Cambridge, MA: National Bureau of Economic Research.

- Rodríguez, M. C., and Turner, P. (2003) "Currency Substitution and the Demand for Money in Mexico", *Applied Economic Letters*, 10(1), 59-62.
- Rumbaugh, T., Ishi, K., Liang, H., and Masuda, A. (2000) "Cambodia: Selected Issues". IMF Country Report 00/135. Washington, D.C.: International Monetary Fund.
- Ryabchenko, P. (1998) "Talony vmesto deneg", *Nezavisimaya Gazeta*, ns, 13 October.
- Sahay, R., and Vegh, C. (1996) "Dollarization in Transition Economies", in Mizen, P., and Pentecost, E. J., ed., *The Macroeconomics of International Currencies: Theory, Policy and Evidence*. Cheltenham: Edward Elgar.
- Salvatore, D., Dean, J., and Willett, T. (2003) *The Dollarization Debate*. Oxford: Oxford University Press.
- Sarajevs, V. (2000) "Econometric analysis of currency substitution: a case of Latvia", BOFIT Discussion Paper 2000/4.
- Savastano, M. (1996) "Dollarization in Latin America: recent evidence and policy issues", in Mizen and Pentecost ed. (1996a).
- Samuelson, P. A. (1958) "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money", *Journal of Political Economy*, 66(6), 467-82.
- Schmukler, S., and Serven, L. (2002) "Pricing Currency Risk: Facts and Puzzles from Currency Boards". NBER Working Paper 9047. Cambridge, MA: National Bureau of Economic Research.
- Schumpeter, J. (1954) *History of Economic Analysis*. London: Allen and Unwin.
- Seabright, P. ed. (2000) *The Vanishing Rouble -- Barter Networks and Non-Monetary Transactions in Post-Soviet Societies*. Cambridge: Cambridge University Press.
- Seaford, R. (2004) *Money and the Early Greek Mind*. Cambridge: Cambridge University Press.
- Selcuk, F. (1997) "GMM Estimation of Currency Substitution in a High-Inflation Economy: Evidence from Turkey", *Applied Economic Letters*, 4(4), 225-27.
- Shaw, E. S. (1973) *Financial Deepening in Economic Development*. New York: Oxford University Press.
- Shi, S. (1995) "Money and Prices: A Model of Search and Bargaining", *Journal of Economic Theory*, 67, 467-96.
- Shubik, M. (2001) "Money and the Monetization of Credit", Cowles Foundation Working Paper 1343. Cowles Foundation for Economic Research: New Haven, CT.
- and Smith, E. (2005) "Fiat Money and the Natural Scale of Government", Cowles Foundation Working Paper 1509. Cowles Foundation for Economic Research: New Haven, CT.

- and Smith, E. (2005) "Commodity Money and the Valuation of Trade", Cowles Foundation Working Paper 1510. Cowles Foundation for Economic Research: New Haven, CT.
- Sidrauski, M. (1967) "Rational Choice and Patterns of Growth in a Monetary Economy", *American Economic Review*, 57(2), 534-544.
- Simmel, G. (1978) *The Philosophy of Money*. London: Routledge & Keegan Paul.
- Smith, A. (1981) *The Wealth of Nations*. Indianapolis: Liberty Press.
- Smith, J. (1987) "Trade Credit and Informational Asymmetry", *Journal of Finance* 42(4) 863-69.
- Smithin, J. (2000) *What is Money?* London: Routledge.
- Šonje, V., Faulend, M., and Šošić, V. (2001) "Nelikvidnost: razotkrivanje tajne" Croatian National Bank Research Paper 1-9. Zagreb: CNB.
- Stiglitz, J. and Weiss, A. (1981) "Credit Rationing with Imperfect Information", *American Economic Review*, 71, 393-410.
- Taylor, J. B. (2000) "Low Inflation, Pass-Through, and the Pricing Power of Firms", *European Economic Review*, 44(7), 1389-1408.
- Tchaidze, R. (1999) "Non-Cash Settlements of Expenditures in Baltics, Russia, and Other Former Soviet Republics". Mimeo, IMF. Available at <http://www.econ.jhu.edu/People/Tchaidze/research.html>.
- Thomas, L. R. (1985) "Portfolio Theory and Currency Substitution", *Journal of Money, Credit, and Banking*, 17(3), 347-57.
- Tobin, J. (1956) "The Interest Elasticity of the Transactions Demand for Cash", *Review of Economic Studies*, 38(3), 241-7.
- (1958) "Liquidity preference as behaviour towards risk", *Review of Economic Studies*, 25, 65-86.
- Townsend, R. (1980) "Models of Money with Spatially Separated Agents", in Kareken, J. and Wallace, N. edd. *Models of Monetary Economies*, Minneapolis, Federal Reserve Bank of Minneapolis.
- (1987) "Economic Organization with Limited Communication", *American Economic Review*, 77(6), 954-71.
- (1989) "Currency and Credit in a Private Information Economy", *Journal of Political Economy*, 97(6), 1323-44
- Uribe, M. (1997), "Hysteresis in a Simple Model of Currency Substitution", *Journal of Monetary Economics*, 40, 185-202.
- Uş, V. (2003), "Analyzing the Persistence of Currency Substitution Using a Ratchet Variable: the Case of Turkey", *Emerging Markets Finance and Trade*, 39(4), 58-81.

- Velde, F. and Veracierto, M. (2000) "Dollarization in Argentina", Federal Reserve Bank of Chicago Economic Perspectives, 24(1), 24-35.
- Vetlov, I. (2001) "Dollarization in Lithuania: an Econometric Approach", BOFIT Discussion Paper 2001/1.
- Vujčić, B. (1998) "Monetary Policy and Management of Capital Flows in a Situation of High Euroisation -- the Case of Croatia", BIS Papers 17, 79-98.
- White, B. (1921) *The Currency of the Great War*. London: Waterlow and Sons, Ltd.
- Williamson, S. (1999) "Private Money", *Journal of Money, Credit and Banking*, 31(3), 470-491.
- Williamson, S. and Wright, R. (1994) "Barter and Monetary Exchange under Private Information", *American Economic Review*, 84(1), 104-23.
- Woodruff, D. (1999) *Money Unmade: Barter and the Fate of Russian Capitalism*. London: Cornell University Press.
- Wray, R. (2004) *Credit and State Theories of Money: the Contributions of A. Mitchell Innes*. Cheltenham: Edward Elgar.
- Yildirim, J. (2003), "Currency Substitution and the Demand for Money in Five European Union Countries", *Journal of Applied Economics*, 6(2), 361-83.