Discussion Papers No. 562, October 2008
Statistics Norway, Research Department

Roger Hammersland

Classical identification: A viable road for data to inform structural modeling

Abstract:
This paper addresses how to enhance the role of data in structural model design by utilizing structural breaks and superfluous information as auxiliary tools of exact identification. To illustrate the procedure and to study the simultaneous interplay between financial variables and the real side of the economy a simultaneous equation model is constructed on Norwegian aggregate data. In this model, while innovations to stock prices and credit do cause short run movements in real activity, such innovations do not precede real economy movements in the long run.

Keywords: Structural vector Error Correction modeling, Identification, Cointegration, Financial variables and the real economy.

JEL classification: C30, C32, C50, C51,C53, C53, E44

Acknowledgement: I am grateful for comments by Terje Skjerpen, participants at the ESEM 2007 meeting in Budapest. I would also like to thank Karsten Gerdrup with whom I wrote the forerunner of this paper.

Address: Roger Hammersland, Statistics Norway, Research Department. E-mail: rhs@ssb.no

1 Introduction

These days a priori information has more or less completely got the upper hand on data in the process of structural model identification and design. For instance, in the structural vector autoregressive (SVAR) and simultaneous equation (SEM) model literature it has been, and still is, common to exactly identify the system by combining the imposition of a diagonal structural form covariance matrix of the errors with either (non-testable) a priori restrictions on the contemporaneous feedback matrix or analogous restrictions on the matrix of parameters that characterizes the long run solution of the system.[1] Often these kind of restrictions imply a lower or upper block triangular contemporaneous feedback matrix which gives importance to the ordering of the variables in the block diagonal part of the system in that the short run responses implied by the lower or upper triangularity should be in accordance with some perceived a priori view of "delayed" reaction.[2] In general little attention is paid in this literature to the issue of model design beyond what is implied by this process of identification.

The inherent problem of structural model design is that there is no way to test for the exact identifying restrictions of a structural or simultaneous equation model. As long as the exact identifying restrictions reflect subjective a priori information of substantial interest and consequence for the properties of the model this introduces necessarily a significant trace of arbitrariness in model design and specification. In fact, in some cases one might even speak of design where the outcome is more or less fully manipulated and determined by the researcher's a priori subjective belief or wishful thinking. Furthermore, the fact that most of the structural literature resorts to a diagonal structural error covariance matrix as a design criterion further impairs the possibility of developing a data congruent model as it contributes to make the model less elastic when confronted with data. In particular the price paid for securing a structural interpretation of shocks ex ante in this respect, i.e. as a part of the exact identifying set of restrictions ,could be unduly high in terms of miss-specification and lack of congruency.

To reduce the degree of arbitrariness inherent in structural modelling this paper strikes a blow for classical identification techniques aimed at giving more emphasis to data in the process of structural model specification and design. The strategy is based on the idea of making the models "more elastic" when confronted with data and thus to avoid laying the exact identifying restrictions on information laden parts of the model and on parts that would make it harder to come up with an admissible and congruent deterministic structure, i.e the covariance matrix. The advantage of such a strategy should be obvious as after the system is exactly identified tests for over identifying restrictions are at ones disposal and one can enter into a design process where the data are allowed to speak, i.e. a process where both the ordering of the variables and the contemporaneous structure of

[1] There is a huge and growing literature in this area and to render justice to all of its contributors is clearly outside the scope of this paper. However, not to mention Sims (1980) seminal paper where he introduces the idea of exact identification through recursive identification would indisputably have to be characterized as an oblivion. Papers that deserve mention for the introduction of restrictions on the systems long run properties are, respectively, Blanchard and Quah (1989), Shapiro and Watson (1988) and Gali (1992).

[2] Notably there are authors that have tried to avoid the recursive identification scheme, see e.g. Bernanke (1986) and Blanchard and Watson (1986) among others who introduced non-recursive restrictions on the contemporaneous interactions among variables for identification.

the model is the outcome of a testable dialog with the data and not divine information. As regards the covariance matrix, this advocates a strategy where the structural shock restrictions are tested for and potentially imposed ex post, i.e. after the deterministic part of the model has got its final structural form.

Ruling out the use of the contemporaneous feedback matrix and the covariance matrix as sources of exact identification limits the set of ways to exactly identify the system. However, it is important to point out that several alternatives still remain at our disposal. A classical approach to the problem would for instance imply that one puts to use exogenous information, information about structural breaks together with the model's lag structure. In addition this paper suggests to utilize what I have chosen to call *superfluous* information as an auxiliary tool to help with the exact identification part of the problem. In this context *superfluous* information refers to information that originally entered in a more general version of the model, with richer dynamics and longer lags, but that was later discarded in the process of model reduction. Typically, this information constitutes lagged dynamic terms and though often discarded due to poor explanatory power, will serve the purpose as an auxiliary identification tool as long as the parameter estimates are numerically different from zero. The idea is then to get rid of these *superfluous* terms in the early process of structural model design, i.e. the process which final scope is to end up with an overidentified parsimonious structural representation of an exactly identified point of departure.

To illustrate the procedure and to study the mutual interplay between financial variables and the real economy a simultaneous equation model is constructed using Norwegian aggregate data. In the case of Norway, it turns out that to illustrate the leading indicator properties of financial variables in the setting of a fully simultaneous equation model that adequately and congruently portrays the evolvement of the real economy one can do with a surprisingly small information set. In fact in addition to real GDP the information set that forms the basis of our preferred structural vector error correction model comprises only stock prices and an indicator for domestic credit. As regards the results the model do substantiate the leading indicator properties of the financial variables. However, what is interesting and new in this respect compared to earlier research is the finding that this property has a substantially different origin for the variables considered. That is, while the deterministic part of the model identifies a direct and contemporaneously link from domestic credit to real GDP, the stock price variable, on the other hand, only affects the real economy indirectly via the error correction term of the credit equation. The error correction term on the other hand implies that there is a one- to -one relationship between credit and stock prices in the long run while no such relationship relationship is found to tie together the real and financial side of the economy. These findings have two implications: i) the effect of a structural shock to stock prices is only gradually transmitted to the real sphere over the short to medium run and ii) while domestic credit and stock prices are closely linked in the long-run, in general shocks to asset prices only have transitory effects on the real economy. Noteworthy, this stands in contrast to what is found in Beaudry and Portier (2006, 2005) where shocks to stock prices have a lasting long run effect on the US and Japanese real economy.

The remaining sections of the paper are structured as follows. In Section 2, in addition to give some background information I present some stylized facts related to the potential leading indicator properties of financial variables for the development of real mainland

DGP of Norway. Section 3 then set up the empirical model framework and runs through a modelling exercise with the aim of illustrating the potential of a data based structural design procedure and to show how it can be used to shed light on the sources of economic fluctuation. Finally, Section 4 offers some concluding comments.

2 Background, Stylized Facts and Data

There is a huge literature suggesting that asset price movements reflect the markets expectation of future developments in the economy.[3] For instance, as a test of standard valuation models on US data, Fama (1990) shows that monthly, quarterly and annual stock returns on the NYSE are highly correlated with future production growth rates for 1953-1987. Two recent contributions aimed at demonstrating this are the studies by Beaudry and Portier (2006, 2005), who using US and Japanese data, find that innovations to stock prices precede most of the long run movements in total factor productivity. These findings are cited in support for future production growth reflecting information about future cash flows that is impounded in stock prices.

Moreover, enterprises often finance a share of their total purchases with loans from credit institutions or by issuing bonds when making new investments. An increase in corporate credit as registered in monthly credit statistics may thus provide information about the development in business fixed investments before the national accounts are published. This suggests that not only asset prices but also credit figures might possess leading indicator properties for the development in the real economy.

Looking at the stylized facts there are much to indicate that what is said in the two above paragraphs has much to recommend it. For instance Figure 1 shows that there has been a positive correlation between the real equity price gap, the enterprise investment gap and the consumption gap since the beginning of the 1990s. Furthermore, real equity prices seem to function as a leading indicator of investment, while this variable is more like a coincident indicator of private consumption.

Likewise, according to Figure 2 there seem to have been a close relationship between cyclical developments in real house prices and housing investment since the beginning of the 1990s. Real house prices seem moreover to be related to the consumption gap.

Finally, Figure 3 shows there has been a positive relationship between growth in domestic real credit to enterprises and cyclical developments in mainland business fixed investment since the beginning of the 1990s. For example, growth in real domestic credit to enterprises picked up sharply in 1992 and was fairly instantly followed by a marked increase in mainland business investment. Growth in both credit and investment was sluggish in 2003-2004 during the downturn in the Norwegian economy, but since 2004 they have both simultaneously gathered headway.

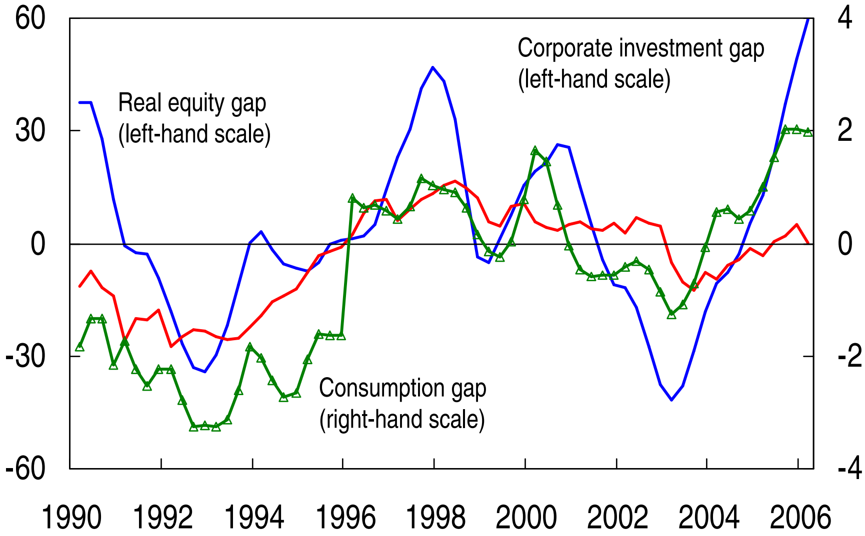
Based on the stylized facts and the outcome of a preliminary correlation analysis [4] the primary system of variables that interests us is one composed of an index of real equity prices, S, real house prices, real domestic credit to enterprises, C, and real mainland GDP. [5] However, based on the results of a number of preliminary modelling attempts, house

[3] For a comprehensive review of the literature see Stock and Watson (2003)

[4] For a detailed account see Gerdrup et al. (2006)

[5] All nominal to real transformations have been accomplished by deflating the nominal quantities by

Figure 1: Real equity gap, corporate investment gap and consumption gap.[1)] Per cent. 1990 Q1 2006 Q1



[1)] Real equitiy prices, fixed investment in non-financial enterprises in Mainland Norway and private consumption as a percentage of trend. The trends are estimated with a HP-filter ($\lambda = 40000$) also using data from the 1980s. The series are adjusted for seasonality and irregular components. Sources: EcoWin, Norges Bank and Statistics Norway

prices were at the end excluded from the information set due to a lack of explanatory power.[6]

Despite referring to deviations from HP filtered trends and seasonally adjusted data in the stylized facts above, in the econometric analysis all variables in the information set consist of unadjusted variables, i.e. variables that have not been de-trended or adjusted for noise or seasonality. The reason for this is that trends and seasonality can be taken explicitly into account in an econometric analysis by specifying seasonal dummies and deterministic trends. Furthermore, the econometric analysis in this paper is confined to the period from the first quarter of 1990 to the last quarter of 2005. This is because it is likely that the relationships between the real economy and financial variables have changed over time, making information from earlier periods less relevant for the purpose of model design. In this respect, it suffices to mention that figures for the 1980s are heavily influenced by the liberalisation of money, credit and capital markets and other economic policy changes.[7]

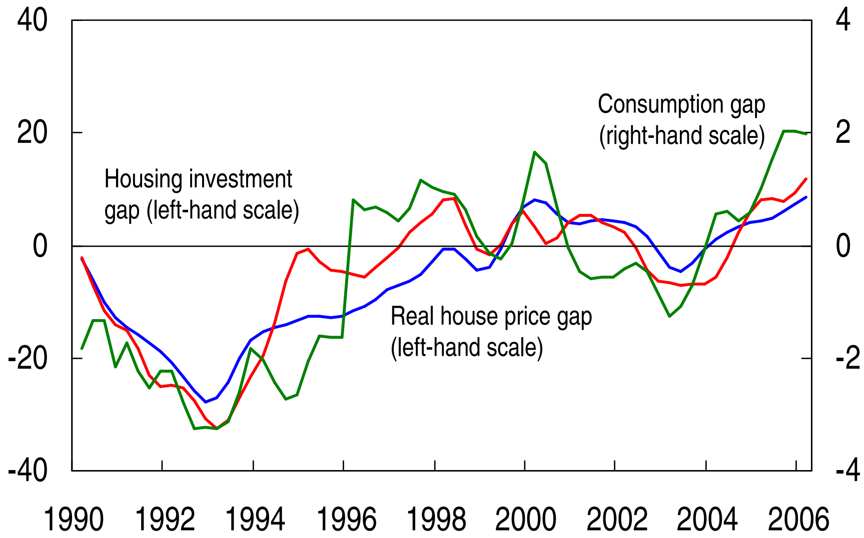
---

the Norwegian consumer price index adjusted for taxes and energy, CPIATE.

[6]In this context its also worth mentioning that in the run-up to this paper a number of models were estimated on information sets that, in addition to the variables mentioned in the text, included short- and long-term interest rates and their differentials. Non of these modelling attempts turned out to be successful, however, in the sense of producing a well specified interpretable simultaneous model with good forecasting properties.

[7]Moreover, there was a banking crisis in Norway in the period 1988–1993. Since 1993, the economic situation has been more stable. It is therefore likely that the relationships between the real economy and

Figure 2: Real house price gap, consumption gap and housing investment gap. [1)] Per cent. 1990 Q1 2006 Q1



[1)] Real house prices, private consumption and housing investment as a percentage of trend. The trends are estimated with a HP-filter (λ =40000) also using data from the 1980s. We have used house price data from the RIMINI (a macro model earlier used in Norges Bank) data base for the 1980s. The series are adjusted for seasonality and irregular components. Sources: ECON/NEF, Norges Bank and Statistics Norway.

3 Model, Identification and Results

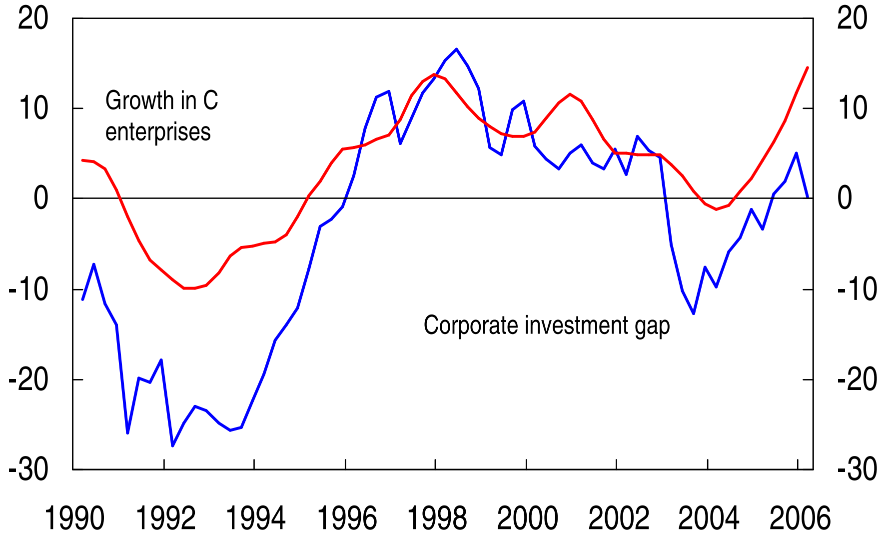
The point of departure is the error correction version of the vector autoregressive reduced form model. In the general case this can be given the following representation:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \epsilon_t \qquad (1)$$

where X_t represents a $p \times 1$ variable vector, k, the order of the VAR, D_t, a vector composed of deterministic variables like dummies, trend and a constant, and ϵ_t a Gaussian white noise term with covariance matrix Ω. The rank of the Π matrix gives us information about the cointegration properties of the model, and in the case the rank, r, is less than full, i.e. less than p, the Π matrix may be written as the product of two $p \times r$ matrices, α and β, with full column rank equal to $r < p$. The level term in equation (1) can then be written as $\Pi X_{t-1} = \alpha\beta' X_{t-1}$ where $\beta' X_{t-1}$ represents the r cointegrating linear combinations of

financial variables have been more stable since 1993 than over a longer period. Nevertheless, we use data from 1990 when we undertake our analysis. The background for this is that we use a model with several variables and lags, and therefore need somewhat longer data series (i.e. several degrees of freedom) to estimate fairly precisely the coefficients of the model. This may be justified by the fact that we can take structural breaks into account in an econometric study, thus benefiting from data for a somewhat longer period.

Figure 3: Corporate investment gap and growth in real domestic credit to enterprises, C. [1)] Per cent. 1990 Q1 2006 Q1



[1)] 1) Fixed investment in non-financial enterprises in mainland Norway as a percentage of trend. The trend is estimated with a HP-filter (λ=40000) also using data from the 1980s. The series are adjusted for seasonality and irregular components. Sources: Norges Bank and Statistics Norway.

the variables while the α matrix has got the interpretation of a coefficient matrix with error correction coefficients or loadings. The cointegration analysis in connection with the preparation of the SEM[8] is based on a three dimensional VAR of order 2,[9] where all the variables where specified as logarithms of the original level series and both a constant and a trend were restricted to lie in the space spanned by the α matrix. Since we are utilizing unadjusted data concentrated seasonal dummies were specified to enter unrestrictedly

[8]To distinguish the type of structural model developed in this paper from the SVAR model type we have chosen to use the term Simultaneous Equation Model interchangeably with the statistical concept of a structural form. A more precise connotation would perhaps have been a Structural Vector error correction Model or SVecM.

[9]The VAR of order 2 amounts to a valid reduction of a data congruent VAR of order 6 where in fact none of the individual equation hypotheses for absence of autocorrelation and heteroscedasticity plus normality are rejected at a level below 10 per cent. The system diagnostics of the VAR(6), given below and where the figures in parentheses are the respective tests'significance probabilities, do neither give rise to any concern.

Vector AR 1-4 test:	$F(36,77)$	=	1.5061[0.0676]
Vector Normality test:	$\chi^2(6)$	=	10.164[0.1179]
Vector Heterosc. test:	$\chi^2(228)$	=	246.99[0.1849]

The F-test statistic for the elimination of all lags greater than 3 from the model is F(36,113)=1.2712[0.1716], where the figure in parenthesis is the test's significance probability. Nor where any of the partial reductions implied by the model reduction rejected.

Table 1: **Johansen's test for the number of cointegrating vectors**

VAR order: 2, constant and trend restricted to lie in the α space, unrestricted centered seasonal dummies. Estimation period: 1990 Q1 to 2005 Q4.

Trace Eigenvalue test:		
H_0	H_1	Values of test statistics
r=0	r $\leq$ 3	82.844[0.000]**
r $\leq$ 1	r $\leq$ 3	42.825[0.000]**
r $\leq$ 2	r $\leq$ 3	11.093[0.086]

1) The values in parentheses are the respective tests' significance probabilities.

2) * and ** signify that the test is significant at a level of 5 and 1%, respectively.

together with a dummy for a structural break in the credit series in 1994Q2.[10] The VAR was then estimated by Full Maximum Likelihood. In this context it is, as pointed out by Johansen (2006), worth noting that there is a price paid by using maximum likelihood in estimating VARs. Namely that the model must fit the data in the sense of constituting a congruent representation of the data generating process(DGP). Footnote 9 shows, however, that this requirement does not pose any problems in our case.

The results of the reduced form cointegration analysis is given in Table 1 and Table 2 below and gives unambiguous support for the existence of two cointegrating vectors. Moreover, the F-test for the number of overidentifying restrictions, in Table 2, shows that the identified system, consisting of two cointegrating relationships, constitutes a valid restriction of a corresponding exactly identified long run structure. The first of the structural long-run relationships implies that GDP mainland Norway is a trend stationary variable with a yearly growth rate of approximately 2.9%. In this respect it is worth noting that the output gap, as estimated in this way, is fairly similar to that presented in Norges Bank (2006) from 1996 and onwards (See Figure 4).[11] The second cointegrating relationship on the other hand implies that the ratio of domestic credit of enterprises to equity prices is constant over time, which due to the logarithmic specification and a small abuse of terminology amounts to saying that a percentage increase in the equity price feeds into an equivalent increase in domestic credit of enterprises in the long run. However, to substantiate what was here hinted at, namely that the long-run causal link between credit and equity prices goes from equity prices to credit, an analysis of the error correction coefficient matrix,α,is necessary.

Looking at the loadings we do, however, observe that the second cointegrating vector while being non-significant in the equity price equation, is strongly significant in the re-

[10]The inclusion of the structural break dummy is due to a redefinition of the credit variable in 1994 2 where some households were reclassified as enterprises and vice versa. The dummy is equal to one in the second quarter of 1994 and zero otherwise and thus represents a level shift in the marginal processes for the levels of the variables.

[11]See Inflation Report 1/06 on http://www.norges-bank.no/.

Table 2: **The identified system of cointegrating linear combinations given r=2, the loading matrix and a test of overidentifying restrictions** [1)]

The identified long run structure given 2 cointegrating relations:

$$\begin{pmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} & \beta_{51} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} & \beta_{52} \end{pmatrix} \begin{pmatrix} gdp_t \\ c_t \\ s_t \\ 1 \\ \text{TREND} \end{pmatrix} =$$

$$\begin{pmatrix} gdp_t & - & 0.0073\,\text{TREND} & - & 3.90 \\ & & (0.00016) & & (0.013) \\ c_t & - & s_t & - & 0.50 \\ & & & & (0.055) \end{pmatrix}$$

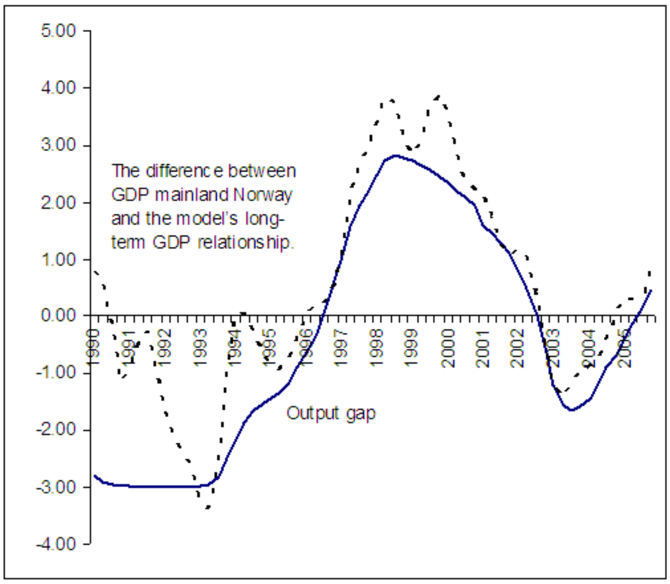
Error correction coefficient matrix:

$$\begin{matrix} \Delta gdp \\ \Delta c \\ \Delta s \end{matrix} \; : \; \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix} = \begin{pmatrix} - & 0.634 & - & 0.026 \\ & (0.111) & & (0.001) \\ & 0.127 & - & 0.032 \\ & (0.122) & & (0.0075) \\ - & 1.532 & & 0.03 \\ & (0.584) & & (0.03) \end{pmatrix}$$

LR-test of overidentifying restrictions: $\chi^2(3) = 0.603[0.896]$

[1)] The value in parenthesis under each coefficient is the coefficient's standard error while the value in parenthesis following the test of over-identifying restrictions is the test's significance probability. The variables gdp_t ,c_t and s_t stands for, respectively, real mainland GDP, real domestic credit to enterprises and real equity prices, lower4case letters indicating that all the quantities are logarithmic transformations of the original variables referred to in the text.

Figure 4: Output gap from Inflation Report 1/06 and the difference between GDP mainland Norway[1)] and the model's long-term GDP relationship. Per cent. 1990 Q1 2005 Q4



[1)] GDP mainland Norway is adjusted for seasonality and irregular components. Sources: Norges Bank and Statistics Norway

duced form equation for domestic credit. This backs up our former suspicion of a one-way causal link between credit and equity prices, going from equity prices to credit, though to fully substantiate such a claim would necessitate the construction of a fully simultaneous equation model. Furthermore, the loading matrix reveals that both cointegrating relationships contribute to error correction in the reduced form equation of Norwegian mainland GDP. While the first of these error correction terms represents ordinary error correction, in the sense of representing a mechanism that gradually seeks to eliminate output gaps, the second term is less trivial in that there is little to suggest independent error correction from credit gaps in the structural GDP equation. However, the loadings of Table 2 relates to the reduced form of a underlying structural form suggesting that the second significant error correction term might be due to a contemporaneous dynamic link between output and domestic credit. If so, the second reduced form error correction coefficient does nothing else than reflecting the error correction properties of the credit equation. To unveil whether this is the case or whether the second error correction term in the reduced form GDP equation is due to an independent structural effect and/or both is one of the aims in the structural analysis below. Otherwise, one can note that the first cointegrating relationship does not contribute to error correction in the reduced form equation for domestic credit. This indicates absence of a contemporary dynamic effect of GDP in the behavioral domestic credit equation of the simultaneous equation model and that the contemporary causal direction thus is from credit growth to GDP growth and not the other way around.

The model that so far has been analyzed is a reduced form representation of the variables in our information set. To be able to explicitly address the topic of dynamic contemporary causality and to construct a model that is more in accordance with the idea of economic systems by nature being simultaneous we will now move on and on the basis of the reduced form analysis develop a simultaneous equation model for our three variables. However, before presenting the results of this modelling exercise we will first turn to a brief discussion of the technique being used to exactly identify the behavioral system.

The structural form or SEM representation of the reduced form is obtained by multiplying (1) by a contemporary response matrix B. This results in the simultaneous equation system:

$$B\Delta X_t = B\Pi X_{t-1} + \sum_{i=1}^{k-1} B\Gamma_i \Delta X_{t-i} + B\Phi D_t + B\epsilon_t,$$

or after having set $B\Pi = B\alpha\beta' = \alpha^*\beta'$, $B\Gamma_i = \Gamma_i^*$, $B\Phi = \Phi^*$ and $B\epsilon_t = u_t$

$$B\Delta X_t = \alpha^*\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i^* \Delta X_{t-i} + \Phi^* D_t + u_t, \tag{2}$$

Given the two previously estimated long run relationships and the fact that the cointegration analysis was undertaken on a VAR(2), (2) will have the following representation in our particular example:

$$\begin{pmatrix} 1 & b_{12} & b_{13} \\ b_{21} & 1 & b_{23} \\ b_{31} & b_{32} & 1 \end{pmatrix} \begin{pmatrix} \Delta gdp_t \\ \Delta c_t \\ \Delta s_t \end{pmatrix} = \begin{pmatrix} \alpha_{11}^* & \alpha_{12}^* \\ \alpha_{21}^* & \alpha_{22}^* \\ \alpha_{31}^* & \alpha_{32}^* \end{pmatrix} \begin{pmatrix} \text{gdp} - 0.0073\text{TREND} - 3.90 \\ c - s - 0.50 \end{pmatrix}_{t-1} \tag{3}$$

$$+ \begin{pmatrix} \gamma_{11}^* & \gamma_{12}^* & \gamma_{13}^* \\ \gamma_{21}^* & \gamma_{22}^* & \gamma_{23}^* \\ \gamma_{31}^* & \gamma_{32}^* & \gamma_{33}^* \end{pmatrix} \begin{pmatrix} \Delta gdp_{t-1} \\ \Delta c_{t-1} \\ \Delta s_{t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11}^* & \phi_{12}^* & \phi_{13}^* & \phi_{14}^* & \phi_{15}^* \\ \phi_{21}^* & \phi_{22}^* & \phi_{23}^* & \phi_{24}^* & \phi_{25}^* \\ \phi_{31}^* & \phi_{32}^* & \phi_{33}^* & \phi_{34}^* & \phi_{35}^* \end{pmatrix} \begin{pmatrix} 1 \\ S1 \\ S2 \\ S3 \\ D942 \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}$$

where we have normalized the contemporary response- or feedback matrix such that the coefficients along the main diagonal is equal to one. However, as regards estimation of (2) and (3) there evidently is a puzzle to resolve as neither of the two representations are identified – in the sense of representing a one-to-one mapping of the corresponding reduced form – without imposing further restrictions.[12] In the SVAR literature this problem is solved by assuming: i) a lower or upper triangular response matrix and ii) a diagonal empirical structural covariance matrix. However, as regards the first point, there is an inherent and insuperable problem associated with imposing exactly identifying restrictions

[12]Note that if we multiply (2) with an arbitrary non-singular F-matrix the corresponding reduced form will still be equal to (1). This illustrates that there in general does not exist a one-to one mapping between the reduced form and a SEM or structural form representation. Only in the case where the only admissible transformation matrix, F, is equal to a diagonal matrix, or in the case of (3) the identity matrix, will the simultaneous equation system be identified.

on a structural form as the exactly identifying restrictions can never be tested for. Thus one evidently runs the risk of imposing a dynamic contemporary structure that is not supported by data. This advocates a strategy where ones leaves the parts of the system perceived to be of minor importance for the purpose of exact identification and then to consider over-identifying restrictions to test for restrictions on more information laden parts of the model, like the B matrix which evidently contains important information about the causal interplay among the variables in our information set.

Moreover, though imposing the covariance matrix of the structural model's disturbances to be diagonal is theoretically substantiated, the matter presents itself quite differently when constructing empirical models on real data as there is little to suggest that the empirical covariance matrix of an estimated structural form model should inherit the stochastic properties of its theoretical equivalent. This follows both as a consequence of utilizing empirical proxies for theoretical constructs and due to the fact that empirical models in most cases are linear approximations of non-linear theoretical equivalents. Add to this the inherent problem of omitted variables and there should be no lack of reasons to substantiate why one should be careful with laying the identifying restrictions on the covariance matrix of an empirical model.

Though, ruling out the use of the contemporaneous feedback matrix and the covariance matrix as sources of exact identification represents a clear curtailment in the set of ways to exactly identify the system, it is important to point out that several alternatives still remain to our disposal. A classic approach to the problem would for instance imply that one puts to use exogenous information, information about structural breaks together with the models lag structure. Moreover, in this paper we do suggest to utilize something we have chosen to call *superfluous* information as an auxiliary tool to help with the exact identification part of the SEM. In this context *superfluous* information refers to information that originally entered in a more general version of the model, with richer dynamics and longer lags, but that was later discarded in the process of model reduction. Typically, this information constitutes lagged dynamic terms and though often discarded due to poor explanatory power, will serve the purpose as an auxiliary identification tool as long as the parameters are estimated numerically different from zero. The idea is then to get rid of these *superfluous* terms in the early process of the model design.

To relate what is said in the last paragraph to our particular example one should recall that we in modeling started out with a general VAR of order 6 for then reducing this down to a parsimonious representation of order 2 (see footnote 9). In accordance with equation (2), this means that only one lag of first differences should enter the model. However, it also means that we in the process of model reduction have discarded the total of four lags which in principle all could serve as auxiliary tools for the identification of a simultaneous equation model. Also, the fact that these lags are discarded from the parsimonious model is what makes them eligible for the characterization; carriers of *superfluous* information. Thus to be able to exactly identify system (2) an idea would be to reintroduce the first difference of the model endogenous variables lagged two periods and then to impose restrictions on their coefficients such that the SEM becomes exactly identified. In line with this an exactly identified SEM representation is given by:[13]

[13]This statement is related to the order condition. However, the claim that (4) is exactly identified hangs evidently on the rank condition also being fulfilled, which thus is tacitly and implicitly assumed in this assertion.

$$\begin{pmatrix} 1 & b_{12} & b_{13} \\ b_{21} & 1 & b_{23} \\ b_{31} & b_{32} & 1 \end{pmatrix} \begin{pmatrix} \Delta gdp_t \\ \Delta c_t \\ \Delta s_t \end{pmatrix} = \begin{pmatrix} \alpha^*_{11} & \alpha^*_{12} \\ \alpha^*_{21} & \alpha^*_{22} \\ \alpha^*_{31} & \alpha^*_{32} \end{pmatrix} \begin{pmatrix} \text{gdp} - 0.0073\,\text{TREND} - 3.90 \\ c - s - 0.50 \end{pmatrix}_{t-1}$$

$$+ \begin{pmatrix} \gamma^*_{1.11} & \gamma^*_{1.12} & \gamma^*_{1.13} \\ \gamma^*_{1.21} & \gamma^*_{1.22} & \gamma^*_{1.23} \\ \gamma^*_{1.31} & \gamma^*_{1.32} & \gamma^*_{1.33} \end{pmatrix} \begin{pmatrix} \Delta gdp_{t-1} \\ \Delta c_{t-1} \\ \Delta s_{t-1} \end{pmatrix} + \begin{pmatrix} \phi^*_{11} & \phi^*_{12} & \phi^*_{13} & \phi^*_{14} \\ \phi^*_{21} & \phi^*_{22} & \phi^*_{23} & \phi^*_{24} \\ \phi^*_{31} & \phi^*_{32} & \phi^*_{33} & \phi^*_{34} \end{pmatrix} \begin{pmatrix} 1 \\ S1 \\ S2 \\ S3 \end{pmatrix} \qquad (4)$$

$$+ \begin{pmatrix} \gamma^*_{2.11} & 0 & 0 & \phi^*_{15} \\ 0 & \gamma^*_{2.22} & 0 & \phi^*_{25} \\ 0 & 0 & \gamma^*_{2.33} & \phi^*_{35} \end{pmatrix} \begin{pmatrix} \Delta gdp_{t-2} \\ \Delta c_{t-2} \\ \Delta s_{t-2} \\ D942 \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}$$

In (4) we have restricted the additional lags such that they only affect the behavioral equation of the variable they represent, a set of restrictions that evidently constitutes only one of many ways to exactly identify (4) by the use of redundant lags. In (4) we have also chosen to relegate the structural break dummy, $D942$, to the part of the system used for identification. The reason for this is to emphasize the dummy's important potential role in helping out with the exact identification of the system.[14]

Based on system (4) we are now ready to get down to the process of reducing our exactly identified simultaneous equation model to a parsimonious structural representation of the information contained in our data set. The fact that we by help of a clever trick have managed to avoid laying the exactly identifying restrictions on the contemporaneous response matrix implies that we by way of tests for over-identifying restrictions now are in full command when designing the contemporaneous causal structure of the model. The result of this process of design is our preferred parsimonious equation representation, given by: [15]

[14] In fact information about structural breaks in the data, like the D942 dummy in (4), is clearly advantageous to superfluous information as a tool of exact identification as the last category of variables according to Sims (1980) under certain conditions would be classified as instruments of weak identification while the first category of information always in principle leads to strong identification. When focus in this section nevertheless is on the use of superfluous information this is to demonstrate that even in the case where truly exogenous information, like structural breaks, is not at ones disposal, one might still be able to exactly identify a structural form without resorting to a priori information. However, in general, and due to their frequent existence I would like to strike a blow for the use of structural breaks in the process of exact identification.

[15] By chance the contemporaneous feedback matrix of system (5) turns out to be lower diagonal. To see this we have to reorder the equations of the simultaneous equation model such that the credit equation comes first, the GDP equation second and the equity price equation last, the reordering of the variables

$$
\begin{pmatrix} 1 & -0.36 & 0 \\ & (0.126) & \\ 0 & 1 & 0 \\ & & \\ -4.213 & 0 & 1 \\ (1.594) & & \end{pmatrix}
\begin{pmatrix} \Delta gdp_t \\ \\ \Delta c_t \\ \\ \Delta s_t \\ \\ \end{pmatrix}
=
\begin{pmatrix} -0.402 & 0.135 & 0 \\ (0.104) & (0.075) & \\ 0 & 0.34 & 0 \\ & (0.091) & \\ 3.57 & 0 & 0.272 \\ (1.133) & & (0.113) \end{pmatrix}
\begin{pmatrix} \Delta gdp_{t-1} \\ \\ \Delta c_{t-1} \\ \\ \Delta s_{t-1} \\ \\ \end{pmatrix}
\tag{5}
$$

$$
+ \begin{pmatrix} -0.535 & 0 \\ (0.12) & \\ 0 & -0.0377 \\ & (0.0078) \\ 0 & 0 \end{pmatrix}
\begin{pmatrix} \mathrm{gdp} - 0.0073\,\mathrm{TREND} - 3.90 \\ c - s - 0.50 \end{pmatrix}_{t-1}
$$

$$
+ \begin{pmatrix} 0.009 & -0.05 & -0.09 & -0.06 & 0 \\ (0.002) & (0.104) & (0.075) & (0.007) & \\ -0.0046 & 0 & 0.34 & 0 & -0.06 \\ (0.0029) & & (0.091) & & (0.013) \\ -0.041 & 3.57 & 0 & 0.271 & 0 \\ (0.023) & (1.133) & & (0.113) & \end{pmatrix}
\begin{pmatrix} 1 \\ \\ S1 \\ \\ S2 \\ \\ S3 \\ \\ D942 \\ \\ \end{pmatrix}
+ \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}
$$

System diagnostics and test of restrictions

LR-test for over-identifying restrictions:	$\chi^2(19)$	=	25.449[0.1463]
Vector test for autocorrelation of order 1-4:	$\mathrm{F}(36, 127)$	=	0.9929[0.4907]
Vector test for normality:	$\chi^2(6)$	=	9.9818[0.1254]
Vector test for heteroscedasticity:	$\mathrm{F}(120.192)$	=	1.1507[0.1925]

---

leading to a lower triangular or fully recursive contemporaneous feedback matrix given by:

$$
\begin{pmatrix} 1 & 0 & 0 \\ 0.36 & 1 & 0 \\ 0 & -4.2 & 1 \end{pmatrix}
$$

Pretending that the covariance matrix of the disturbances is diagonal this means that structural shocks to GDP and equity prices in (5) do not contemporaneously affect domestic credit while shocks to credit and GDP, on the other hand, is allowed to contemporaneously affect, respectively, GDP and equity prices. While some of these restrictions might seem reasonable others are less so, the most intriguing artifact perhaps being the absence of a contemporaneous effect of a shock to GDP on domestic credit. However, in this respect it is important to point out that the structure of (5) has been the outcome of a design process where the data have been allowed to speak, i.e. a process where both the ordering of the variables and the contemporaneous structure of the model is the outcome of a testable dialogue with the data. Compared to a procedure that more or less ignores information contained in data this should add to the reliability of the resulting identification scheme.

The first thing to notice is that the test of the over-identifying restrictions does not reject the null hypothesis that the final parsimonious simultaneous equation model, (5), constitutes a valid reduction of an exactly identified version of the model.[16] The system diagnostics given below our preferred system indicate also that the system describes data fairly well, as none of the standard vector tests indicate presence of autocorrelation, non-normality or heteroscedasticity. Moreover, the single equation and vector stability tests of Figure 5 in Appendix A do indicate that the system as such is relatively stable over the estimation period as in fact none of the recursive tests breaks a test level of 1%. Figure 6 shows furthermore that the model provides relatively good fit to GDP, real domestic credit to enterprises and real equity prices. Moreover, the model predicts GDP growth 8 quarters ahead fairly well when it is estimated using data up to and including 2003Q4 and simulated dynamically to 2005Q4 (see Figure 7). The model also predicts the development in real credit to enterprises fairly well the first six quarters of the forecast period, but does not capture the increase in real enterprice credit in the latter half of 2005. This may be because enterprises have shifted funding from foreign to domestic sources.[17] Nor was the model able to predict all of the sharp increase in real equity prices in 2004 and 2005. This may be because equity prices have been influenced by factors that are not included in the model, and perhaps among those high oil prices in particular. The forecast errors for credit and equity prices are small, however, seen in the context of the uncertainty surrounding the projections, given by 95% prediction intervals in the figures.

In the model, growth in real domestic credit to enterprises is influenced by credit growth in the previous quarter and by a long-term relationship that posits that the ratio of real enterprise credit to real equity prices is constant over time. This implies that real credit of enterprises will increase by 1 per cent in the long term if real equity prices increase by 1 per cent. As there is no effect of credit in the equity price equation, neither contemporaneously nor lagged or via an error correction mechanism, this long-run relationship thus represents a unidirectional causal relationship. Noteworthy there is no effect of equity prices in the GDP equation. Domestic credit on the other hand is estimated to affect real output both contemporaneously and with a lag. In the model this means that real equity prices do not affect output directly but work through channels that are captured in the real credit to enterprises. These channels can be the consumption channel, the credit channel, the investment channel and the expectations channel.

All in all, the results give some support to the leading indicator property of financial variables for predicting business cycles. However, while this is due to a direct and contemporaneously causal link for a variable like domestic credit, a structural shock to stock prices is identified to affect the real economy only indirectly via the error correction term of the credit equation. Accordingly, the effect of a structural shock to stock prices, in addition to being transient, is small and protracted.

[16]Since there is a one-to-one mapping between an exactly identified simultaneous equation model and a reduced form, an equivalent statement would be that the parsimonious simultaneous equation model does not represent a significant loss of information compared to the reduced form VAR of order 2 used in the cointegration analysis.

[17]Total real credit growth for enterprises (C3 mainland enterprises) was lower in 2005.

4 Conclusions

This paper addresses how to enhance the role of data in structural model design by utilizing structural breaks and *superfluous* information as auxiliary tools of exact identification. To illustrate the procedure and to study the simultaneous interplay between financial variables and the real side of the economy a simultaneous equation model is constructed on Norwegian aggregate quarterly data.

In the case of Norway, it turns out that to illustrate the leading indicator properties of financial variables, in the setting of a fully simultaneous equation model that adequately and congruently portrays the evolvement of the real economy, one can do with a surprisingly small information set. In fact in addition to real GDP the information set that forms the basis of our preferred SEM comprises stock prices and an indicator of domestic credit only. Though one evidently must exercise caution in drawing too strong conclusions based on such a simple and stylized description of the causal interplay between the real and financial economic spheres it is nevertheless our firm belief that the model developed herein serves to illustrating some interesting real society traits or features. Not least due to the reasonableness of results and the modelś good statistic properties.

As regards the results the model substantiates the leading indicator properties of the financial variables. However, what is interesting and new in this respect compared to former research is the finding that this property has a substantially different causal origin for the two variables. That is, while the deterministic part of the SEM identifies a direct and contemporaneously causal link from domestic credit to real GDP, the stock price variable on the other hand does only affect the real economy indirectly via the error correction term of the credit equation. The error correction term on the other hand implies that there is a one-to-one relationship between credit and stock prices in the long run while no such long-run relationship is found to tie together the real and financial side of the economy.

These findings have two implications: i) the effect of a structural shock to stock prices is only gradually transmitted to the real sphere over the short to medium run and ii) while domestic credit and stock prices are closely linked in the long-run, shocks to asset prices in general have only transitory effects on the real economy. Noteworthy, this stands in contrast to what is found in Beaudry and Portier (2004, 2005)who find that shocks to stock prices have a lasting long run effect on the real economy.

References

Beaudry, P. and F. Portier (2005), 'The news view of economic fluctuations: Evidence from aggregate Japanese data and sectoral U.S. data', *Journal of the Japanese and International Economies* **19(4)**, 635–652.

Beaudry, P. and F. Portier (2006), 'Stock prices, news and economic fluctuations', *American Economic Review* **96(4)**, 1293–1307.

Bernanke, B. (1986), Alternative explanations of the money-income correlation, *in* 'Carnegie-Rochester Conference Series on Public Policy', Vol. 25, pp. 49–99.

Blanchard, O.J. and D. Quah (1989), 'The dynamic effects of aggregate demand and supply disturbances', *American Economic Review* **79**, 655–673.

Blanchard, O.J. and M.W. Watson (1986), *Are Business Cycles all alike?*, Chicago, University of Chicago Press, pp. 123–156.

Fama, E. F. (1990), 'Stock returns, expected returns and real activity', *Journal of Finance* **45(4)**, 1089–1108.

Gali, J. (1992), 'How well does the IS-LM model fit postwar U.S. data?', *Quarterly Journal of Economics* **107(2)**, 709–738.

Gerdrup, K R, R Hammersland and B. E. Naug (2006), 'Financial variables and developments in the real economy'. Norges Bank Economic Bulletin.

Johansen, S. (2006), Confronting the economic model with the data, *in* D.Colander, ed., 'Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model', Cambridge University Press, pp. 287–300.

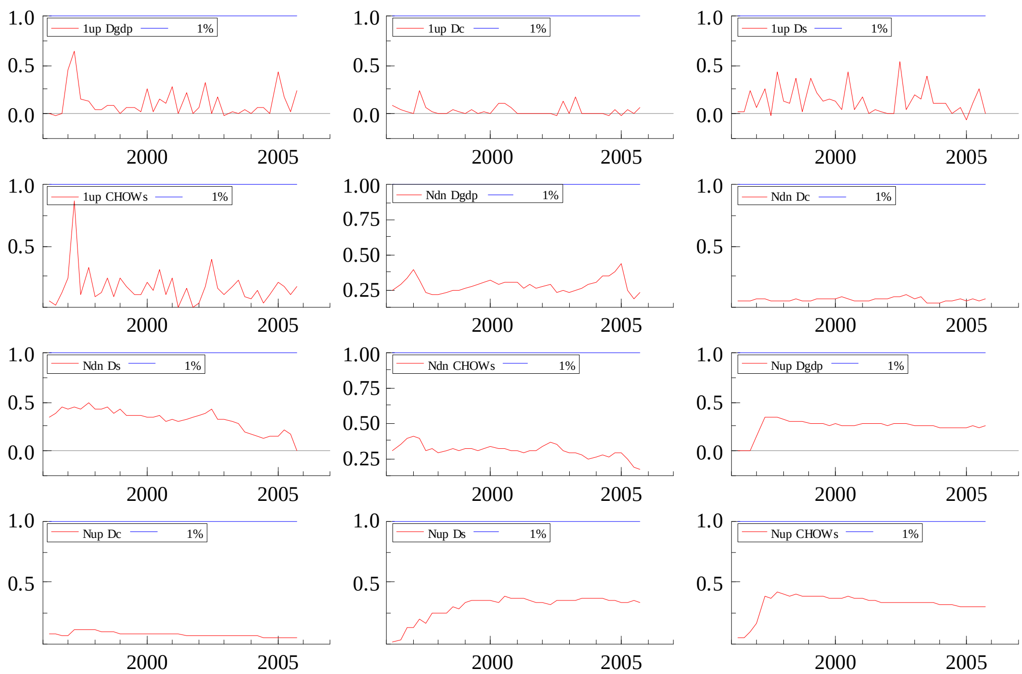
Shapiro, M.D. and M.W. Watson (1988), *Sources of Business Cycle Fluctuations*, Cambridge, MA: MIT: Press.

Sims, Christopher A. (1980), 'Macroeconomics and reality', *Econometrica* **48**, 1–48.

Stock, H.J. and M.W. Watson (2003), 'Forecasting output and inflation: The role of asset prices', *Journal of Economic Litterature* **41**, 788–829.

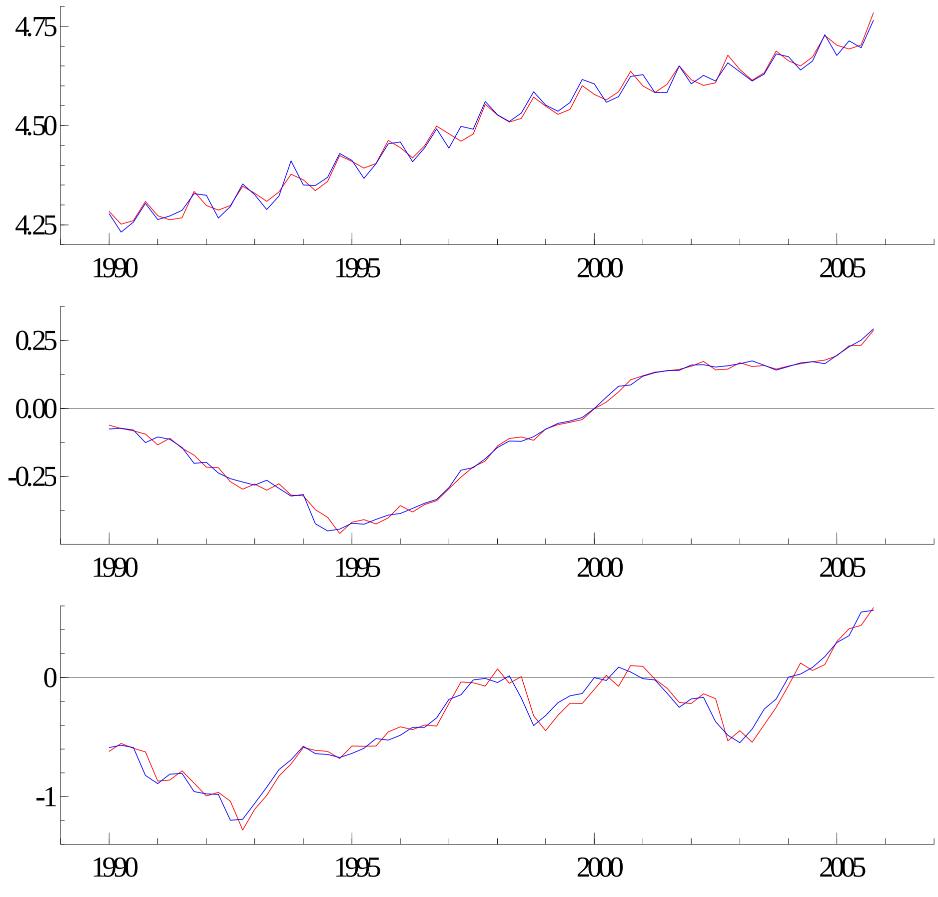
Appendix A

Figure 5: Stability tests



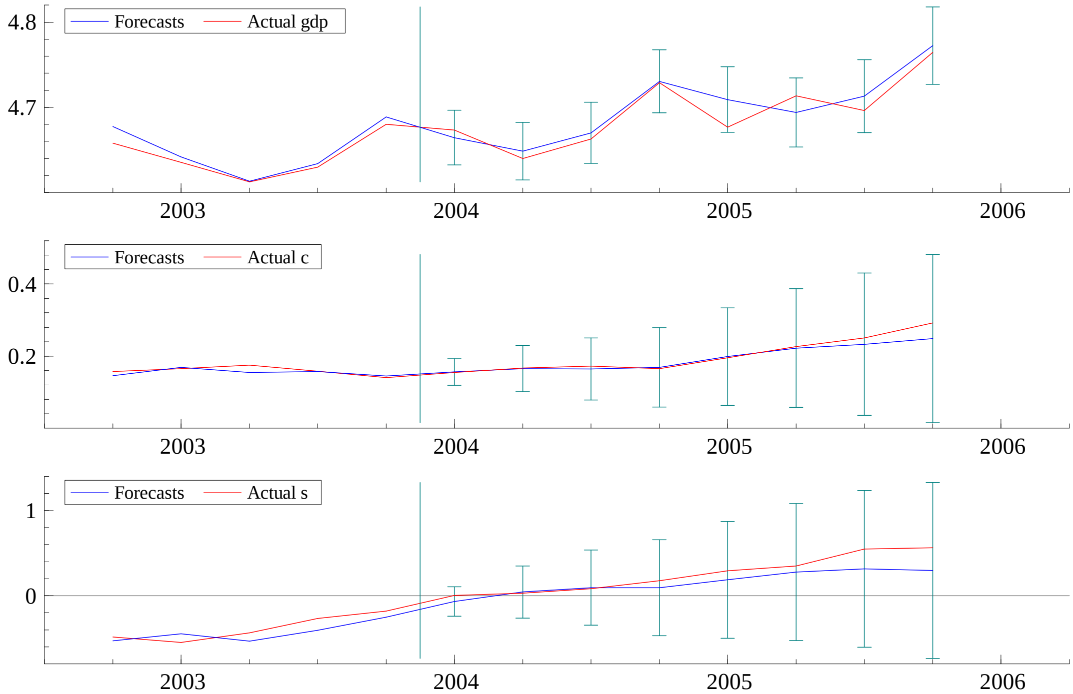
[1] The single equation stability tests are respectively represented by the 1-step forecast tests (1up), Break-point F-tests (N down) and the Forecast F-tests (Nup) tests while the vector tests are represented by the corresponding Chow tests.

Figure 6: Actual and fitted values of GDP Mainland Norway, real domestic credit to enterprises, C, and real equity prices, S. Logarithmic scale. 1990Q1 to 2005Q4



[1)] The red line symbolizes the fitted values while the blue line represents the actual ones.

Figure 7: Forecasted values of GDP Mainland Norway, real domestic credit to enterprises, C, and real equity prices, S, eight quarters ahead from 2004Q1. Logarithmic scale. 2003Q1 to 2005Q4



[1)] The forecasts are represented by the red line while the the blue lines represents the actual values.