PART III

Identification and indicator of monetary policy: Operating procedures in Switzerland

Abstract

We analyze different identification frameworks based on the Swiss National Bank's operating procedures in order to measure monetary policy since the breakdown of the Bretton Woods system. We use a two-stage VAR methodology to identify monetary shocks built either on monetary and nonmonetary residuals (without extraction) or only on the orthogonal portion of monetary residuals relative to nonmonetary residuals (with extraction). Based on these models, we construct various monetary policy indicators. We report them as weighted sums of monetary policy variables. Our main indicator reveals that the exchange rate was the dominant variable at the end of the seventies. During the eighties, aggregates had their golden age, while in the nineties, the call rate showed up as an indicator. In addition to the comparison of different models and their dynamics, we focus on econometric problems arising with such a VAR methodology.

JEL Codes: E50, E52.

Keywords: Identification, Indicator, Monetary policy, Operating procedures, Shocks, Switzerland, VAR.

---

1 Introduction and Overview

This introduction begins with a brief description of three observations that motivate our approach. Empirical search for overall indicators of monetary policy in Switzerland based on operating procedures guides our study. We go on to discuss what we do and then summarize the obtained results.

1.1 Search for Indicators

One of the most challenging goals of monetary economics is the identification of monetary policy (Grossman, 1991). Zha (1997) describes this concept as the process of sorting out the central bank's behavior from that of the many other interacting agents in the economy. We can regard these interactions like a loop between a central bank and other players as financial market participants, producers, and consumers. Thus, the central bank influences economic conditions and at the same time responds to events emanating from the economic activity. Hence, to identify monetary policy consists in separating out the exogenous actions from the systematic reactions of the central bank, or put in other words, isolating exogenous monetary shocks generated by the central bank.

When isolated, these shocks help us focus on the dynamic effects of monetary policy on the economy through its transmission mechanism. Only responses to an exogenous variable can measure the effects of policy-induced changes in that variable (Cochrane, 1994). Furthermore, following this identification, a composite construction, based on these exogenous shocks and an implicit endogenous reaction function, can be used as an indicator of monetary policy. Thus, such a measure reveals the direction and shape of monetary policy and particularly its relative restrictiveness during specific periods of time.

The situation in Switzerland about indicators of monetary policy is similar to the one in other European countries. Different indicators produce different assessments about exogenous and overall policy stance with different terms of validity. When reporting on monetary policy in its main forum, the quarterly publication 'Money, Currency, and Business Cycle' (now called 'Quarterly Bulletin'), the Swiss National Bank (SNB) refers to 'monetary conditions' as indicators. It describes on one hand the evolution of several monetary aggregates, following its official position, and on the other hand also short-term interest rates on the money and financial markets. However, we think it is worth building up new indicators and analyzing whether they perform better in appraising Swiss monetary policy than the traditional view based on sometimes unstable monetary aggregates and noisy short-term interest rates, summarized under 'monetary conditions'[1].

1.2 Use of Operating Procedures

A modern technique to look for an indicator, initiated by Bernanke and Mihov (1997, 1998), is to focus on the behavior of a central bank using its operating procedures. Models of operating procedures directly shape the exogenous implementation of monetary policy. They also provide an economic interpretation for diverse econometric methods criticized by their lack of economic foundations.

Still according to Zha (1997), the search for monetary shocks is de facto an empirical issue. We need sophisticated econometric tools, in particular vector autoregressive (VAR) econometrics, in order to isolate the central bank's behavior from all other behaviors observed in the data. Operating procedures are a 'bridge' between conceptual and empirical identifications. Bernanke and Mihov (1997, 1998) suitably recommend to apply methods allowing for structural changes in the economy, or more precisely, changes in operating procedures. It offers a substantial advantage over the focus on a unique sample or over models that do not consider these changes over time at all.

Using operating procedures to structure reduced-form (RF) models is new and attractive in Switzerland. First, this is new because our model nests the approaches of Clarida and Gertler (1997) and Bernanke and Mihov (1997, 1998), that both use operating procedures following different VAR methods. We focus on operating procedures to select among various scenarios the best one describing the operating choices of the SNB. Second, this setup is particularly appealing, because since the breakdown of Bretton Woods, we suspect several evolutions in Swiss operating procedures due to instabilities on the money and financial markets, legal changes, and electronic improvements in the payments system[2]. They all challenge the role of monetary aggregates as an indicator. Furthermore, Switzerland is characterized by unique features of its central banking economics, as

[1] Searching for an alternative indicator, Lengwiler (1997) applied an index-based indicator using a Monetary Conditions Index (MCI) to Switzerland following a model used by the Bank of Canada. He discovered that this MCI could not outperform the monetary base as a policy indicator.

[2] See Birchler (1988) for more details about legal changes and Vital (1998) concerning the improved payments system introduced at the end of the eighties.

the important role given to exchange rate movements that we have to take into account. This is part of the so-called 'disciplined discretion' (Laubach and Posen, 1997). The SNB has always emphasized the state contingent nature of its monetary targets and rules. In the event of unexpected disturbances (especially an appreciation of the Swiss franc (CHF)), the SNB is prepared to deviate from its monetary targets (Rich, 1997). Thus, we estimate and test, for various samples, models based on different assumptions concerning targeting strategies at the operative level.

1.3 Econometrics

Our last observation concerns VAR econometrics[3]. The extraction of shocks within this class of models may display econometric flaws that we analyze more explicitly. Given that only innovations can be interpreted as monetary shocks, we use a structural VAR (SVAR) methodology. Even if they have been evolving a lot for twenty years, VAR are still severely criticized because they often lack a structural economic model. All our SVAR are thus backed up by our structural model which nests all tested setups. Moreover, SVAR are not always robust with respect to changes in VAR identification assumptions[4]. Worse, even changes in estimation procedures may play a role. We shed light on these criticisms in performing VAR regressions under different VAR identification schemes and with different estimation methods.

1.4 Results

Based on these three observations, our goal is to apply an econometric framework to identify Swiss monetary policy shocks using operating procedures. This application aims at producing new indicators

[3]VAR is a seemingly unrelated regressions model (SURE) where all equations have identical explanatory variables. They have got revival with the seminal Sims' paper (1980) that criticized and challenged the research based on large-scale macroeconomic models.

[4]Identification of VAR and identification of monetary policy are two separate concepts. VAR or econometric identification consists in recovering a structural system from a RF expression. We cannot directly estimate SVAR, we only estimate RF. The difficulty is that there exists an infinity of SVAR for a single RF. In order to recover a particular SVAR, we identify it in putting restrictions on the coefficient matrices. After this econometric identification, SVAR can be used to identify monetary policy shocks, so to isolate exogenous monetary policy.

for Swiss monetary policy and looking at the dynamics of monetary policy shocks. This framework is not completely new for large closed economies, but we perform this analysis for Switzerland, a small open economy, introducing an exchange rate element in the analysis. At the same time, we concentrate on econometric problems not treated in the traditional empirical literature about monetary VAR.

The contribution of this paper is then threefold. First, we show that our model nests different approaches to model VAR residuals. It also reveals the economic (and not technical) ambiguities linked to the comparison of nested models founded on different residuals calculations. These two different calculations are the so-called setups 'without extraction' by Clarida and Gertler (1997) and 'with extraction' by Bernanke and Mihov (1997, 1998). Second, we find that the model using orthogonal residuals (with extraction) can overcome a few flaws found in the nonorthogonal model (without extraction), in particularly avoiding using instrumental variables. Third, albeit the mixed evidence about our results, in particular concerning the puzzling dynamics, we provide a new indicator for the monetary policy stance in Switzerland during the period 1976-1997. All tested indicators are weighted sums of policy variables including the call rate, a monetary aggregate, and the Deutschmark (DM) exchange rate. Our main indicator reveals that the exchange rate was the dominant variable at the end of the seventies. During the eighties, aggregates had their golden age, while in the nineties, the call rate showed up as an indicator. This indicator offers the advantage to directly beam the overall stance of monetary policy actions.

The structure of this study is as follows. Section 2 presents the nesting model and their extrapolated versions. We also describe the data. The third section is devoted to the estimation, the obtained results, and their interpretation. Section 4 concludes.

2 Methodology and Data

2.1 Model

Our model is a two-stage SVAR nesting the approaches of Clarida and Gertler (1997) and Bernanke and Mihov (1997, 1998). After having first estimated a SVAR(k) and stored its RF residuals, we model them in a second nonrecursive SVAR(0).

VAR methodology is the natural approach for the following reasons: i) all variables are endogenous; ii) VAR supplant simultaneous equations systems (SES) due to the endogeneity of all variables and no a priori restriction in the estimation process[5]; iii) possibility to isolate innovations that should receive the name of monetary policy shocks; iv) flexibility in structuralization, particularly imposition of a minimal set of restrictions on the first VAR and possibility to test imposed restrictions on the second VAR.

We go on to present the first SVAR that is common to all our models. However, the second SVAR is different with respect to the treatment of the first VAR residuals. This second SVAR is either called without or with extraction depending on this treatment.

2.1.1 First Step SVAR(k)

Let $\mathbf{z}_t$ be a $[(m+n)\times 1]$ vector of macroeconomic variables and ε_t a $[(m+n)\times 1]$ vector of structural disturbances affecting the economy. The elements of ε_t are mutually orthonormal-iid shocks with a diagonal variance-covariance matrix $E\left(\varepsilon_t\varepsilon_t'\right)=\mathbf{\Sigma}$. Let $\mathbf{A}_0, \mathbf{A}_1 \ldots \mathbf{A}_k$, and $\mathbf{B}$ be square coefficient matrices. Matrix $\mathbf{A}_0$ has diagonal elements normalized to zero and matrix $\mathbf{B}$ has diagonal elements normalized to one, for convenience. For matter of matrix algebra, the vector size is the same for $\mathbf{z}_t$ and ε_t[6]. This first SVAR(k) is a general representation of a macroeconomic framework that determines $\mathbf{z}_t$. All macroeconomic variables are endogenous and, in addition, depend on k lags of all variables in the vector $\mathbf{z}_t$. The true economy is summarized by the structural vector equation (1).

$$\mathbf{z}_t=\sum_{i=0}^{k}\mathbf{A}_i\mathbf{z}_{t-i}+\mathbf{B}\varepsilon_t \tag{1}$$

In order to isolate monetary shocks, an element ε_t^s of ε_t, it is important to make a distinction between variables that the SNB can directly influence and other variables that the SNB cannot directly

[5] See Amisano and Giannini (1997) for a description of advantages and disadvantages of SVAR over SES.

[6] We impose as a priori restrictions that matrix $\mathbf{B}$ is square. There are no economic reason to impose a square $\mathbf{B}$ matrix, implying that the number of independent original disturbances is equal to the number of elements in the vector $\mathbf{z}_t$, here $m+n$. Even if a rectangular $\mathbf{B}$ matrix of size $[(m+n)\times d]$, where $d>m+n$ is possible, we do not follow this way, because we need a non-singular matrix $\mathbf{B}$ to compute indicators of monetary policy, that are presented in the next section.

influence. Because this definition is quite loose, we use a timing assumption (Bernanke and Blinder, 1992) to sort out variables within $\mathbf{z}$. We split this vector into two separate categories according to Clarida and Gertler (1997) and Bernanke and Mihov (1998). We divide elements of $\mathbf{z}$ into m nonpolicy ($\overline{\mathbf{z}}$) and n policy variables ($\underline{\mathbf{z}}$). Thus, we define policy variables as variables that the SNB influences within the current considered period, generally a month. This timing assumption guides us to use data at the monthly frequency. Monthly data offers the advantage to partially solve the degrees of freedom problem faced by VAR. However, from an economic point of view, we could also apply this timing assumption to quarterly data without prejudice. Because of rigidities, we know that the SNB begins to influence nonpolicy variables with a lag, but there is no evidence whether it is a month or a quarter[7].

The nonpolicy variables in the vector $\mathbf{z}$ include a commodity price index $\overline{z}^{com}$ as an indicator of external price shocks, gross domestic product $\overline{z}^{gdp}$, retail sales $\overline{z}^{rs}$, price level $\overline{z}^{pl}$, and the German call rate $\overline{z}^{fcr}$. The data is precisely described in the next section. For the policy variables, we consider the Swiss overnight rate $\underline{z}^{cr}$, a real monetary aggregate $\underline{z}^{mon}$, and the real DM exchange rate $\underline{z}^{exr}$. Concerning money stocks, one can expect the single choice of a narrow defined aggregate due to the central bank's direct control when modeling operating procedures. We nevertheless think that this restriction is not necessary, as we model the exogenous part of monetary policy only with VAR residuals. Thus, the use of broader defined aggregates does not violate the assumption of SNB direct control over monetary aggregates. Henceforth, we pick the monetary base $M0$ $\underline{z}^{mon_0}$ and the money stock $M1$ $\underline{z}^{mon_1}$ [8].

We now rewrite model (1) of the true economy with $\overline{\mathbf{z}}_t$ and $\underline{\mathbf{z}}_t$:

$$\begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} = \sum_{i=0}^{k} \begin{pmatrix} \mathbf{A}_i^{\overline{z}\overline{z}} & \mathbf{A}_i^{\overline{z}\underline{z}} \\ \mathbf{A}_i^{\underline{z}\overline{z}} & \mathbf{A}_i^{\underline{z}\underline{z}} \end{pmatrix} \begin{pmatrix} \overline{\mathbf{z}}_{t-i} \\ \underline{\mathbf{z}}_{t-i} \end{pmatrix} + \begin{pmatrix} \mathbf{B}^{\overline{z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{\underline{z}} \end{pmatrix} \begin{pmatrix} \varepsilon_t^{\overline{z}} \\ \varepsilon_t^{\underline{z}} \end{pmatrix}. \quad (2)$$

The different matrices are now written using partitioned matrix al-

[7] We estimated our model and looked at its dynamics with monthly and quarterly data and noticed that this VAR is robust with respect to the assumed data frequency.

[8] We also performed our estimations with the sight deposits of commercial banks at the SNB (*giro* deposits). Due to poor estimations, results using sight deposits are not reported.

gebra with corresponding sizes[9]. Matrix $\mathbf{B}$ allows the various structural shocks, also split into nonpolicy and policy shocks, to enter each equation with the single restriction that we do not allow the monetary world shocks to independently enter the nonpolicy sphere. They certainly affect the economy but only through the effects on policy variables[10]. This assumption is not too restrictive, because we can imagine processes generating these shocks as totally independent of each other (e.g. with an independent central bank, we can assume such a disconnection). Composite residuals for each variable, or more precisely, for each equation in the system, are then a mix of the different individual structural shocks[11].

System (2) is not econometrically identified. Without restrictions imposed on this true structure, it is not possible to retrieve its coefficients after its RF estimation. Although it is too early to mention in detail the econometric identification problem, a first step towards this identification is to break the loop of contemporaneous influences between nonpolicy and policy variables in this dynamic setup. In order to solve this problem, we use the mentioned timing assumption again, based on the fact that the SNB cannot directly influence in a timing dimension the nonpolicy variables. After the introduction of this timing assumption in the system, implying $\mathbf{A}_0^{\overline{z}\underline{z}} = \mathbf{0}$, system (2) becomes system (3)[12].

$$\begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} = \begin{pmatrix} \mathbf{A}_0^{\overline{z}\overline{z}} & \mathbf{0} \\ \mathbf{A}_0^{\underline{z}\overline{z}} & \mathbf{A}_0^{\underline{z}\underline{z}} \end{pmatrix} \begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} + \sum_{i=1}^{k} \begin{pmatrix} \mathbf{A}_i^{\overline{z}\overline{z}} & \mathbf{A}_i^{\overline{z}\underline{z}} \\ \mathbf{A}_i^{\underline{z}\overline{z}} & \mathbf{A}_i^{\underline{z}\underline{z}} \end{pmatrix} \begin{pmatrix} \overline{\mathbf{z}}_{t-i} \\ \underline{\mathbf{z}}_{t-i} \end{pmatrix} + \begin{pmatrix} \mathbf{B}^{\overline{z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{\underline{z}} \end{pmatrix} \begin{pmatrix} \varepsilon_t^{\overline{z}} \\ \varepsilon_t^{\underline{z}} \end{pmatrix} \tag{3}$$

This system can be reduced for each group of variables and esti-

[9] The partitioned matrix $\mathbf{A}_j^{ab}$ is a coefficient matrix linking explained vector $\mathbf{a}_t$ to explanatory vector $\mathbf{b}_{t-j}$.

[10] For example ε_t^s, an element of vector $\varepsilon_t^{\underline{z}}$, representing expansionary monetary shocks, directly influences all the other elements of vector $\underline{\mathbf{z}}_t$, but has no direct influence on the elements of vector $\overline{\mathbf{z}}_t$.

[11] Equation (1) can be written as $\mathbf{z}_t = \sum_{i=0}^{k} \mathbf{A}_i \mathbf{z}_{t-i} + \boldsymbol{\eta}_t$ where $\boldsymbol{\eta}_t$ is a $[(m+n) \times 1]$ vector of composite residuals. Each element of $\boldsymbol{\eta}_t$ corresponds to an equation in the system. Variance-covariance matrix of $\boldsymbol{\eta}_t$ is a non-symmetric block diagonal matrix, because m is generally different from n.

[12] See Appendixes A-H for all technical details.

mated with ordinary least squares (OLS) equation by equation[13]:

$$\begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} = \sum_{i=1}^{k} \mathbf{\Pi}_i \begin{pmatrix} \overline{\mathbf{z}}_{t-i} \\ \underline{\mathbf{z}}_{t-i} \end{pmatrix} + \begin{pmatrix} \mathbf{r}_t^{\overline{z}} \\ \mathbf{r}_t^{\underline{z}} \end{pmatrix}, \tag{4}$$

where $\begin{pmatrix} \mathbf{r}_t^{\overline{z}} & \mathbf{r}_t^{\underline{z}} \end{pmatrix}'$ are RF residuals after the first estimation. By reduction, we know that vector $\mathbf{r}_t$ is defined as equation (5)

$$\begin{pmatrix} \mathbf{r}_t^{\overline{z}} \\ \mathbf{r}_t^{\underline{z}} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{B}^{\overline{z}} & \mathbf{0} \\ \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{B}^{\overline{z}} & \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} \end{pmatrix} \begin{pmatrix} \varepsilon_t^{\overline{z}} \\ \varepsilon_t^{\underline{z}} \end{pmatrix} \tag{5}$$

representing at the same time the connection between VAR residuals and structural shocks $\begin{pmatrix} \boldsymbol{\varepsilon}_t^{\overline{z}} & \boldsymbol{\varepsilon}_t^{\underline{z}} \end{pmatrix}'$. We define $\boldsymbol{\varepsilon}_t^{\underline{z}} = \begin{pmatrix} \varepsilon_t^s & \varepsilon_t^d & \varepsilon_t^x \end{pmatrix}'$ as a vector including the mentioned monetary policy shock ε^s, a money demand shock ε^d, and an exchange rate shock ε^x.

Equation (5) is the core of this paper and constitutes the base for the second step. We use it to make the distinction between the different models of SNB behavior. The major difference between the two approaches we apply to Switzerland is to decide whether we model policy shocks directly from the vector $\begin{pmatrix} \mathbf{r}_t^{\overline{z}} & \mathbf{r}_t^{\underline{z}} \end{pmatrix}'$, or whether we first extract from $\mathbf{r}_t^{\underline{z}}$ an intermediate vector $\mathbf{u}_t^{\underline{z}}$. We then look for policy shocks from this new vector[14].

2.1.2 Second Step SVAR(0)

For the second estimation, we use an economic model to econometrically identify this nonrecursive SVAR. We have to decide whether we directly use the VAR residuals $\mathbf{r}_t^{\underline{z}}$ from the first regression and try to express them in terms of true structural disturbances $\boldsymbol{\varepsilon}_t$. In this

[13] Where $\mathbf{A}_{0,-1}^{**} = (\mathbf{I} - \mathbf{A}_0^{**})^{-1}$ and $\mathbf{\Pi}_i = \begin{pmatrix} \mathbf{\Pi}_i^{\overline{zz}} & \mathbf{\Pi}_i^{\overline{z}\underline{z}} \\ \mathbf{\Pi}_i^{\underline{z}\overline{z}} & \mathbf{\Pi}_i^{\underline{zz}} \end{pmatrix}$, or more precisely

$$\mathbf{\Pi}_i = \begin{pmatrix} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{zz}} & \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{z}\underline{z}} \\ \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{zz}} + \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_i^{\underline{z}\overline{z}} & \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{z}\underline{z}} + \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_i^{\underline{zz}} \end{pmatrix}.$$

[14] For all models we assume a minimum set of restrictions on the policy part. Concerning the nonpolicy variables, we may use a Cholesky decomposition, implying a specific ordering of the nonpolicy variables, or a more complicated model. This would completely identify the SVAR and allow analyzing its IRF for all types of shocks, in particular nonpolicy shocks. This complete identification is however not necessary to discover policy shocks, to look at their dynamics, and to construct policy indicators.

case, this is the method without extraction. Alternatively, we can extract from $\mathbf{r}_t^{\underline{z}}$ new series $\mathbf{u}_t^{\underline{z}}$ that are the portion of VAR residuals in the policy block that is orthogonal to the VAR residuals in the nonpolicy block. This is the way with extraction. The extraction, to get the new generated residuals $\mathbf{u}_t^{\underline{z}}$, consists in regressing $\mathbf{r}_t^{\underline{z}}$ on $\mathbf{r}_t^{\overline{z}}$:

$$\mathbf{r}_t^{\underline{z}} = \mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{r}_t^{\overline{z}} + \mathbf{u}_t^{\underline{z}} \tag{6}$$

where $\mathbf{u}_t^{\underline{z}} = \mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \boldsymbol{\varepsilon}_t^{\underline{z}}$. We thus model $\mathbf{u}_t^{\underline{z}}$ with help of $\boldsymbol{\varepsilon}_t^{\underline{z}}$.

Econometric difficulties may appear with this second VAR. On one hand, we use generated regressors ($\mathbf{r}$ or $\mathbf{u}$) and on the other hand we are going to use generated instruments ($\mathbf{r}$ and $\boldsymbol{\varepsilon}$) in the setup without extraction. We focus on these two specific problems, when it is crucial to present them.

2.1.2.1 Without Extraction This approach, applied to the German Bundesbank by Clarida and Gertler (1997) and to the Bank of Japan by Chinn and Dooley (1997), directly models the residuals from the first VAR regression. The true model is still the general equation (1) now with matrix $\mathbf{B} = \mathbf{I}_{m+n}$ restricting the model interpretation. This simplification follows the decision that the monetary policy indicator is assumed before estimating the model. For the Bundesbank, Clarida and Gertler (1997) chose the call rate as a policy indicator, generating henceforth a direct relationship between $\underline{z}_t^{cr}$ and ε_t^s . We replicate their research and use operating procedures to model three equations representing the behavior of the SNB in innovation form. We thus configure the policy section of system (5) between $\mathbf{r}_t^{\underline{z}}$ and $\varepsilon_t^{\underline{z}}$ without giving any interpretation to $\mathbf{u}$ because we do not compute them[15]. This application of operating procedures is natural when we know that each policy equation in a VAR can be interpreted as the sum of an endogenous part, a so-called implicit rule, and an exogenous part, representing deviations from the rule or monetary shocks. The central bank's behavior behind these exogenous shocks is definitely linked to its operating actions.

In this framework, the goal is to analyze the central bank's behavior and to have a benchmark for alternative indicators and dynamics

[15] Information about Swiss operating procedures can be found in the large literature on this issue (Bisignano (1996), Borio (1997a, 1997b), Landmann and Jerger (1997), Laubach and Posen (1997), Spörndli and Moser (1997), and Swank and Velden (1997)).

of monetary policy. As already explained, in this setup the monetary policy indicator is not calculated. It is an assumption that we defend or denounce based on econometric and dynamic results of estimated equations. Three equations, one for each policy variable, a money supply function (7), a money demand function (8), and an explanation of real exchange rate (9) are used to configure the policy section.

$$r_t^{cr} = \theta_1 r_t^{com} + \theta_2 r_t^{mon} + \theta_3 r_t^{exr} + \varepsilon_t^s \quad (7)$$

$$r_t^{mon} = \theta_4 r_t^{gdp} + \theta_5 r_t^{cr} + \varepsilon_t^d \quad (8)$$

$$r_t^{exr} = \theta_6 r_t^{com} + \theta_7 r_t^{gdp} + \theta_8 r_t^{rs} + \theta_9 r_t^{pl} + \theta_{10} r_t^{fcr} + \theta_{11} r_t^{cr} + \theta_{12} r_t^{mon} + \varepsilon_t^x \quad (9)$$

This specification is very appealing for its econometric and economic simplicity. These equations are estimated with a generalized method of moments (GMM) or with a two-stage least squares (2SLS) estimator both using instrumental variables (IV)[16]. IV are nonpolicy $\mathbf{r}_t^{\overline{\underline{z}}}$, policy $\mathbf{r}_t^{\underline{z}}$, ε_t^s and ε_t^d.

In equation (7), we assume that the call rate $\underline{z}_t^{cr}$ can be interpreted as an indicator of overall monetary policy and that ε_t^s represents the monetary shock we look for. This equation is thus a reaction function in innovation form against inflation pressure stemming from extern supply shocks r_t^{com}, increases in money demand r_t^{mon}, and exchange rate appreciations r_t^{exr}. All other things being equal, we expect positive coefficients. The second equation (8) is a money demand equation in innovation form with a scale variable r_t^{gdp} and an opportunity cost for keeping wealth in cash form r_t^{cr}. We except a positive sign for the output coefficient and a negative one with the interest rate coefficient. The third equation (9) is an unrestricted representation of the exchange rate explained by a commodity price index, GDP, price level, retail sales, the German call rate, the Swiss call rate, a monetary aggregate, and a real exchange rate shock.

The first two equations are restricted while the exchange rate one is not. In this model, we reproduce the Clarida and Gertler (1997) setup and assume that the SNB reacts to inflation threats.

[16] Intuitively, the use of IV should 'correct' the setup without extraction for circularity problem between regressors and structural shocks. It is however difficult to isolate this correction effect as we see in the next section 'Comparison without and with Extraction'.

The money demand equation is standard. When looking at the effects of the monetary sector ($\varepsilon^{\underline{z}}$) on nonpolicy and policy variables, the coefficients of nonpolicy variables ($\theta_2, \theta_4, \theta_6$-$\theta_{10}$) do not enter the impulse response functions (IRF) calculation, but their integration in the equations influences the estimated coefficients (the θ's premultiplying policy variables in the three equations) that are involved in monetary IRF. So, even if not always significant, they can play an important role.

Despite these interesting features, we have to be careful with such an approach. First, we decide that the call rate should represent the monetary policy indicator and interpret residuals of this equation as monetary shocks. It also means that the indicator of overall stance is per se the call rate. We correct this first characteristic in the model with extraction, where among different restrictions, we try to pick up the most appropriate one implying a suitable indicator.

Second, we use residuals from a first SVAR as regressors ($\mathbf{r}$) and residuals of a second univariate regression ($\boldsymbol{\varepsilon}$) as instruments implying a 'double Pagan problem' (Pagan, 1984). The use of generated regressors, produced by the first VAR(k), in the money supply, money demand, and exchange rate equations, do not bias the coefficients of our three equations. However, the use of generated instruments in the money demand (ε_t^s) and in the exchange rate equation (ε_t^s and ε_t^d) may unfortunately bias the regressions, such that the coefficient inference and IRF could be misleading.

Third, we have a serious robustness problem because the used instruments (for equations (8) and (9)) come from two regressions where we make identification assumptions (Sarte, 1997). With other assumptions, we would get different results and therefore, probably different instruments for the subsequent regressions. Selected IV can even loose their instrumental power. It is worth noticing that Clarida and Gertler (1997) never mention this robustness problem that remains a main disadvantage of this model. This non-robustness is also exacerbated by the estimation of different samples. We see, in the section about models with extraction, that we do not have this robustness problem any more, because we use other estimation methods.

Finally, our results without extraction are also different when we use a 2SLS or an overidentified GMM estimator. Probable heteroscedasticity of residuals guides us to use the overidentified GMM.

As a benchmark, we keep the same specification as Clarida and

Gertler (1997) for the sake of comparison with the Bundesbank. Even if this approach is not the core of Clarida and Gertler (1997), we think it is interesting to have a benchmark to compare our different extensions.

2.1.2.2 With Extraction This approach, initiated by Bernanke and Mihov (1997, 1998), is our second method to model residuals $\mathbf{r}_t$. First of all, we extract residuals $\mathbf{u}_t^z$ from RF residuals. A priori, without empirical results, it is difficult to point out the advantages of this extraction over the framework chosen by Clarida and Gertler (1997). From a theoretical point of view, this is merely another model that avoids the flaws connected to the framework without extraction. However, we also see the limit of the comparison between our two setups. Series in each setup do not represent the same part of monetary policy. In the previous model, it is clear that in addition to the econometric problems, we use 'polluted' residuals to model SNB behavior. The approach with extraction has nevertheless the advantage to analyze, relatively to the general model, series that only symbolize exogenous monetary policy.

After the extraction (6), we store the new residuals $\mathbf{u}_t^z$ that now represent the link to structural policy shocks: $\mathbf{u}_t^z = \mathbf{A}_{0,-1}^{zz}\mathbf{B}^z\boldsymbol{\varepsilon}_t^z$. These new series in innovation form are $\mathbf{u}_t^z = \begin{pmatrix} u_t^{cr} & u_t^{mon} & u_t^{exr} \end{pmatrix}'$. Representing the autonomous policy of the central bank, it is intuitive to use operating procedures again in this framework as explained in Bernanke and Mihov (1997, 1998). Moreover, we include in the policy variables an exchange rate element typical for a small open economy. A difficulty with the quoted papers is that they only consider closed economies like the US or Germany. In Switzerland, we cannot construct models without considering the exchange rate. It is a strategic variable for the SNB and for various lobbies in Switzerland. Aggregate demand very often increases first through its external components, and then through its absorption. This decision enables to keep the same variables as Clarida and Gertler (1997) to facilitate the comparison between these two angles.

The setup, to configure the policy section, is the same as the structure without extraction, a money supply function (11), a money demand function (12), and an expression for the exchange rate in monetary innovations (13). As before, three shocks $\boldsymbol{\varepsilon}_t^z$ influence this system. This market for bank reserves in innovation form must be in equilibrium (10). We do not write time subscripts.

$$u_s^{mon} = u_d^{mon} \tag{10}$$

$$u_s^{mon} = \lambda\varepsilon^d + \phi\varepsilon^x + \varepsilon^s \tag{11}$$

$$u_d^{mon} = \rho u^{cr} + \varepsilon^d \tag{12}$$

$$u^{exr} = \delta u^{cr} + \varepsilon^x \tag{13}$$

Money supply equation (11) allows the SNB to react on the reserves market affected by money demand shocks and exchange rate shocks. The monetary authority can thus accommodate or not money demand shocks and shocks of the CHF external value. The SNB has also the opportunity to unilaterally implement monetary policy shocks ε^s. Money demand equation (12) is a standard money demand equation without a scale variable, because we are in the policy sphere only. All things being equal, a negative ρ means that an increase in opportunity costs of holding money reduces the money demand. Finally, last equation (13) pictures an explanation of the exchange rate in innovation form, where we do not expect a specific sign for δ.

System (10)-(13) can be reduced and expressed in matrix notation corresponding to $\mathbf{u}^{\underline{z}} = \mathbf{A}^{\underline{zz}}_{0,-1}\mathbf{B}^{\underline{z}}\varepsilon^{\underline{z}}$:

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & \frac{(\lambda-1)}{\rho} & \frac{\phi}{\rho} \\ 1 & \lambda & \phi \\ \frac{\delta}{\rho} & \frac{\delta(\lambda-1)}{\rho} & \left(1+\frac{\delta\phi}{\rho}\right) \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix}. \tag{14}$$

We estimate this system (14) following Bernanke and Mihov (1997, 1998). We use a GMM estimator for stationary variables and a variance-covariance structure as moment conditions. System (14) is underidentified (6 conditions $V[u^{cr}], Cov[u^{mon}, u^{cr}], Cov[u^{exr}, u^{cr}]$, $V[u^{mon}]$, $Cov[u^{exr}, u^{mon}]$, and $V[u^{exr}]$ for 4 coefficients λ, ϕ, ρ, δ and 3 variances $V[\varepsilon^s]$, $V\left[\varepsilon^d\right]$, and $V[\varepsilon^x]$). Henceforth, we just-identify and overidentify this system in order to find out the best models carried by the data. We then perform various tests, in particular Hansen (1982) tests, on these constrained setups in order to discover whether some restrictions imposed on the system best catch the variable dynamics. Because a specific assumption may be too restrictive for the whole sample, we split the sample into various subperiods. We thus hope to find for each subsample the appropriate restrictions to apply to the model. This will signal for each subsample different operating procedures changing over time.

We present now five different restricted setups. Three setups use two restrictions and similarly correspond to three strategies to operate monetary policy. The central bank can, on a day-to-day basis, target total bank reserves, the call rate, or the exchange rate. In this framework, we cannot regard inflation targeting as a potential strategy because we only focus on operational targets[17].

We also look at two schemes using only a single restriction: $\delta = 0$ implying that structural exchange rate shocks are orthogonal and $\lambda = 0$ meaning that the central bank does not react to money demand shocks. Hereafter, we briefly sketch the model $\mathbf{u}^{\underline{z}} = \mathbf{A}_{0,-1}^{\underline{zz}}\mathbf{B}^{\underline{z}}\varepsilon^{\underline{z}}$ for each strategy.

Bank Reserves Targeting (BR) Coefficients λ are ϕ equal zero. It implies that the SNB does not react to money demand shocks and to exchange rate shocks. It results $\varepsilon^s = u^{mon}$.

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & -\frac{1}{\rho} & 0 \\ 1 & 0 & 0 \\ \frac{\delta}{\rho} & -\frac{\delta}{\rho} & 1 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix} \tag{15}$$

Call Rate Targeting (CR) Coefficient λ takes on 1 and ϕ on 0. The SNB does not react to exchange rate shocks, but does accommodate money demand shocks to target the call rate. It results $\varepsilon^s = \rho u^{cr}$.

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & 0 & 0 \\ 1 & 1 & 0 \\ \frac{\delta}{\rho} & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix} \tag{16}$$

Exchange Rate Targeting (ER) Coefficients $\lambda = 1$ and $\phi = -\frac{\rho}{\delta}$ constrain the system (14). The SNB accommodates money demand shocks and partially offsets exchange rate shocks. It results $\varepsilon^s = \frac{\rho}{\delta} u^{exr}$.

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & 0 & -\frac{1}{\delta} \\ 1 & 1 & -\frac{\rho}{\delta} \\ \frac{\delta}{\rho} & 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix} \tag{17}$$

[17] A plausible alternative would be to add an expected inflation series (as a policy variable) in $\underline{\mathbf{z}}$ and to introduce expected inflation in the representation of the market for bank reserves (10)-(13).

Setup $\delta = 0$ It removes from this system the structure of the last equation. It results $\varepsilon^s = \lambda\rho u^{cr} + (1-\lambda) u^{mon} - \phi u^{exr}$.

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & \frac{(\lambda-1)}{\rho} & \frac{\phi}{\rho} \\ 1 & \lambda & \phi \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix} \tag{18}$$

Setup $\lambda = 0$ It implies that the central bank does not react to money demand shocks. It results $\varepsilon^s = u^{mon} - \phi (u^{exr} - \delta u^{cr})$.

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & \frac{-1}{\rho} & \frac{\phi}{\rho} \\ 1 & 0 & \phi \\ \frac{\delta}{\rho} & \frac{-\delta}{\rho} & \left(1 + \frac{\delta\phi}{\rho}\right) \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix} \tag{19}$$

2.1.3 Indicator Construction

Each presented setup implies an indicator of monetary shocks and of the overall stance of monetary policy[18]. They are based on assumptions or on econometric key figures. In the case without extraction, it is by assumption the call rate that measures the overall stance of monetary policy. ε^s from equation (7) reveals only exogenous policy. In the case with extraction, there is a specific way to construct a composite indicator proposed by Bernanke and Mihov (1997, 1998). Their method calculates a weighted sum of policy variables. A $[3 \times 3]$ matrix premultiplying vector $\underline{\mathbf{z}}_t$ produces a $[3 \times 1]$ vector of indicators.

The following equation, the policy vector from the true model (1), can shed light on this construction:

$$\begin{aligned} \underline{\mathbf{z}}_t &= \sum_{i=1}^{k} \mathbf{\Pi}_i^{\underline{z}\overline{z}} \overline{\mathbf{z}}_{t-i} + \sum_{i=1}^{k} \mathbf{\Pi}_i^{\underline{z}\underline{z}} \underline{\mathbf{z}}_{t-i} + \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{z}\overline{z}} \mathbf{B}^{\overline{z}} \right) \varepsilon_t^{\overline{z}} \\ &\quad + \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right) \varepsilon_t^{\underline{z}}. \end{aligned} \tag{20}$$

We premultiply $\underline{\mathbf{z}}_t$ by the inverse of the multiplicand of structural shocks. We take then the element in the vector $\left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right)^{-1} \underline{\mathbf{z}}_t$ that corresponds to the line having the element ε_t^s. With our ordering, it is

[18] Six models: one without extraction and five constrained models with extraction.

always the first element. This ordering does not play any role because all the matrices are computed according to the chosen structure. This first element indicates the overall monetary policy, while ε^s only stands for exogenous monetary policy.

2.1.4 Comparison without and with Extraction

We have seen that the comparison between both approaches is obvious concerning the construction of the two models. Moreover, it is technically possible to compare both approaches regarding the produced indicators. It is more difficult to derive the economic intuition behind the disparities of those two types of indicators.

Major differences are on one hand the distribution of structural shocks among the different equations ($\mathbf{B}$), and on the other hand the construction of matrix $\mathbf{A}_0$, put together $\mathbf{A}^{\underline{z}\underline{z}}_{0,-1}\mathbf{B}^{\underline{z}}$. This matrix represents also the link to add to RF VAR IRF in order to gain SVAR IRF.

Table 1: Potential Indicators

Weights for Policy Variables	$\underline{z}^{cr}$	$\underline{z}^{mon}$	$\underline{z}^{exr}$
Without	τ	$\tau\left(\theta_3\theta_{12}+\theta_2\right)$	$\tau\theta_3$
BR	–	1	–
CR	ρ	–	–
ER	–	–	$\frac{\rho}{\delta}$
$\delta=0$	$\lambda\rho$	$1-\lambda$	$-\phi$
$\lambda=0$	$\phi\delta$	1	$-\phi$

Note: Without = Model without extraction; BR = Bank reserves targeting; CR = Call rate targeting; ER = Exchange rate targeting; $\delta, \lambda = 0$ = One-restriction model with extraction. $\tau = \frac{1}{1-\theta_3\theta_5\theta_{12}-\theta_3\theta_{11}-\theta_2\theta_5}$.

We report in table 1 the first line of matrix $\left(\mathbf{A}^{\underline{z}\underline{z}}_{0,-1}\mathbf{B}^{\underline{z}}\right)^{-1}$ creating therefore with $\underline{\mathbf{z}}$ the desired indicator in the setup with extraction. The setup without extraction also produces such a weighting matrix and so an implicit indicator. It contains only θ's that premultiply policy variables in equations (7)-(9). We thus ignore the restricted part linking $\mathbf{r}^{\underline{z}}_t$ to $\mathbf{r}^{\overline{\underline{z}}}_t$ and the assumption that the indicator of this model without extraction is the call rate. On the other hand, in the framework with extraction, this part is unrestricted and estimated during the extraction. Table 1 summarizes all potential indicators and shows for each indicator the weights to apply to each policy variable in $\underline{\mathbf{z}}$.

2.2 Data

We use almost the same nonpolicy variables as Clarida and Gertler (1997), namely a commodity price index, GDP, retail sales, price level, and the German call rate[19]. Concerning the policy variables, we still follow Clarida and Gertler (1997) and Bernanke and Mihov (1998) with respect to the call rate and monetary aggregates. We report data for real $M1$ and the real monetary base, but we do not highlight the difference between borrowed and non-borrowed reserves. We think that this difference is too marginal to have a significant impact on our results. However, we introduce an exchange rate element in order to consider the open economy. We use monthly data as explained in the previous section and detail them in table 2 and figure 1.

Table 2: Data Description

75:10-97:12				
	μ	σ	JB	ADF
$\overline{z}^{com}$	96.22	13.26	6.22**	-2.99
$\overline{z}^{gdp}$	23589.80	2490.18	23.74*	-0.85
$\overline{z}^{rs}$	85.99	16.23	22.36*	-2.08
$\overline{z}^{pl}$	80.15	15.94	18.84*	-1.35
$\overline{z}^{fcr}$	5.94	2.38	24.92*	-2.94
$\underline{z}^{cr}$	3.36	2.36	33.63*	-1.94
$\underline{z}^{mon_0}$	39485.21	7086.36	5.04	-0.87
$\underline{z}^{mon_1}$	90697.09	8761.23	46.04*	-2.39
$\underline{z}^{exr}$	0.99	0.07	82.43*	-2.43

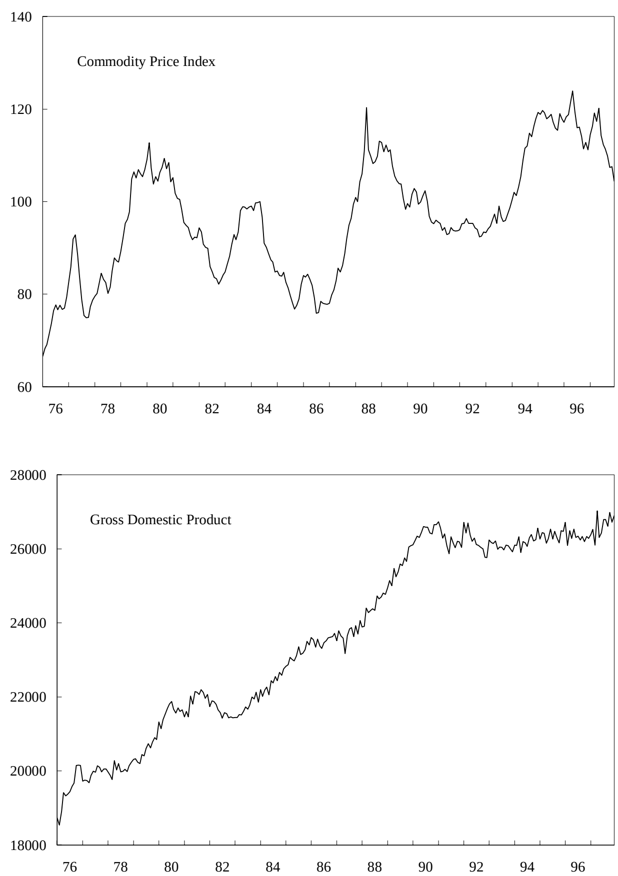
Note: $\overline{z}^{com}$ = Commodity price index; $\overline{z}^{gdp}$ = Gross domestic product (mio CHF); $\overline{z}^{rs}$ = Value of retail sales (index); $\overline{z}^{pl}$ = Price level index; $\overline{z}^{fcr}$ = German call rate; $\underline{z}^{cr}$ = Call rate; $\underline{z}^{mon_0}$ = Real monetary base (mio CHF); $\underline{z}^{mon_1}$ = Real $M1$ (mio CHF); $\underline{z}^{exr}$ = Real exchange rate (CHF/DM); μ = Mean; σ = Standard deviation; JB = Jarque-Bera test; ADF = Augmented Dickey-Fuller test. Null hypotheses: i) JB test, H_0: normal distribution; ii) ADF test, H_0: unit root. Rejection of the null hypothesis at the 1% significance level (*) and at the 5% significance level (**). Source: Datastream, SNB, and Cuche and Hess (1999a, 1999b).

In addition to traditional data characteristics, we focus on their stationarity properties. We include in table 2 augmented Dickey-Fuller (ADF) tests for the series in level. Because of the low power

[19] We use a monthly GDP based on the methodology of Cuche and Hess (1999a). Due to structural breaks in the GDP series during our sample 1975-1997, we interpolate our GDP with their optimized parameters used for the period 1981-1997 in order to match their interpolated series.

of ADF-tests and because the reported series almost succeed in passing the ADF-test (many passed at the 10% significance level), we decide to use the variables in log-level. This decision is also motivated by the purpose of the first regression, i.e. to form residuals for the second VAR. This decision furthermore facilitates the economic interpretation of the dynamics[20]. These choices are partially confirmed by the results of cointegration tests following the Johansen procedure (1991) reported in table 3. In table 3, we see that it is not possible to find a cointegrating vector which authorizes a plausible economic interpretation.

Figure 1: Data 75:10-97:12



[20] All IRF produce stationary results.

Figure 1 *Continued*

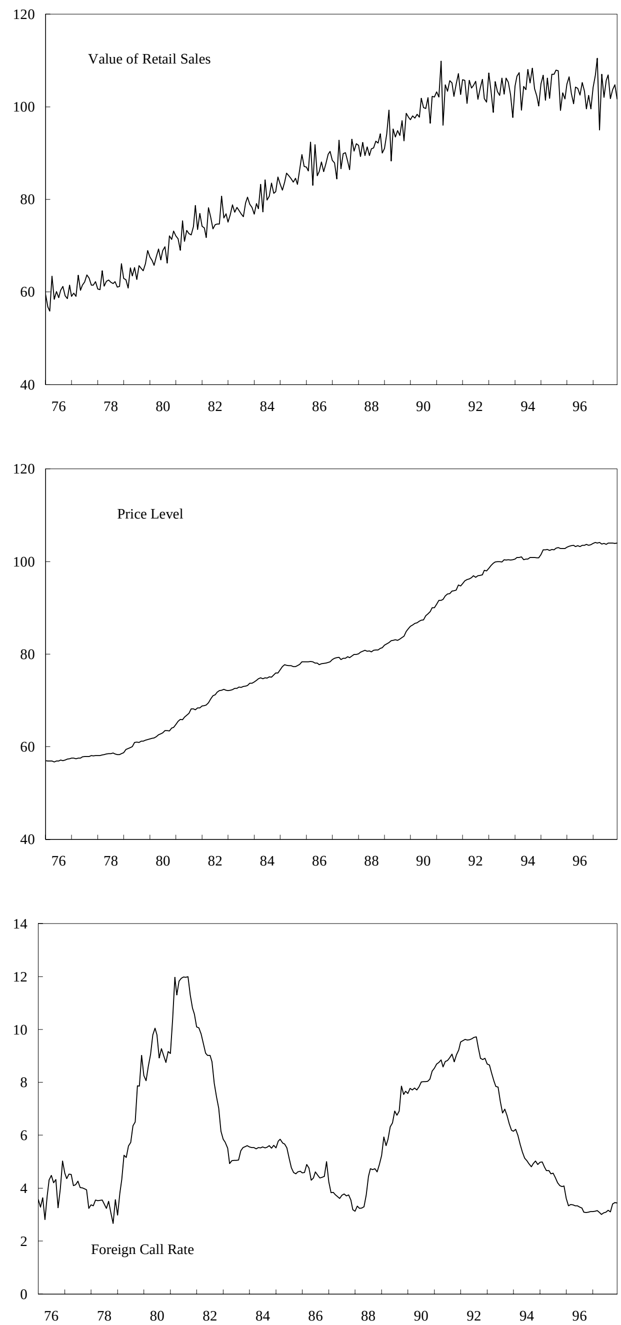


Figure 1 *Continued*

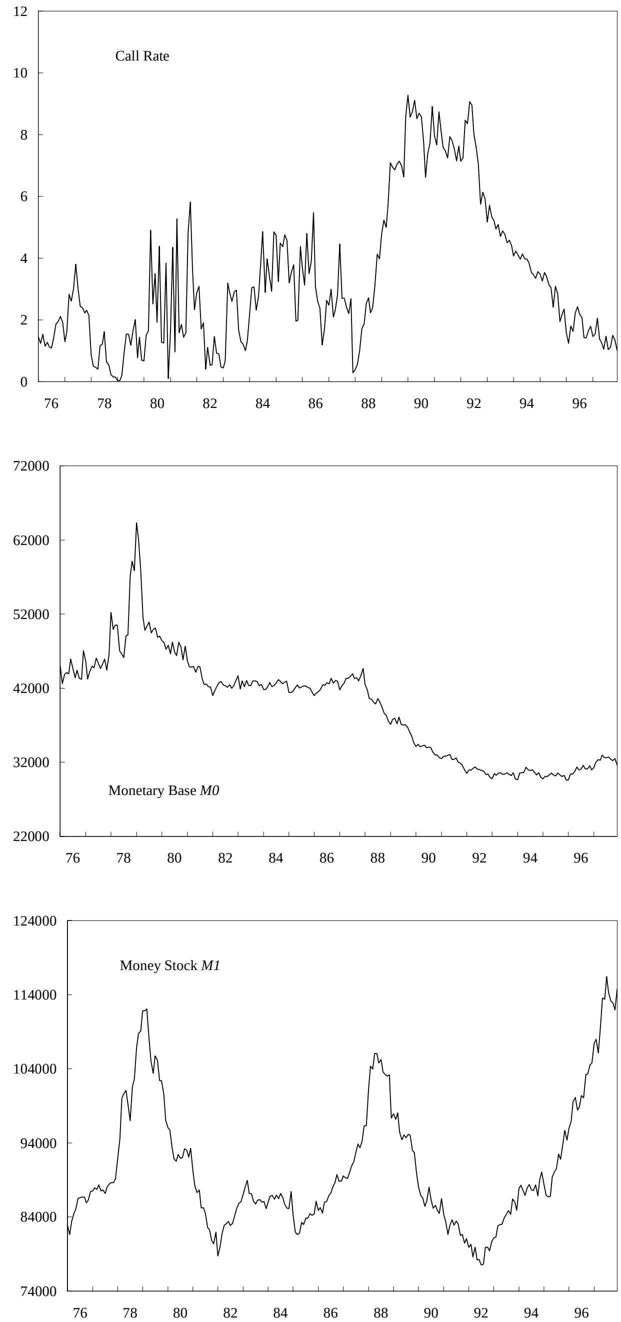
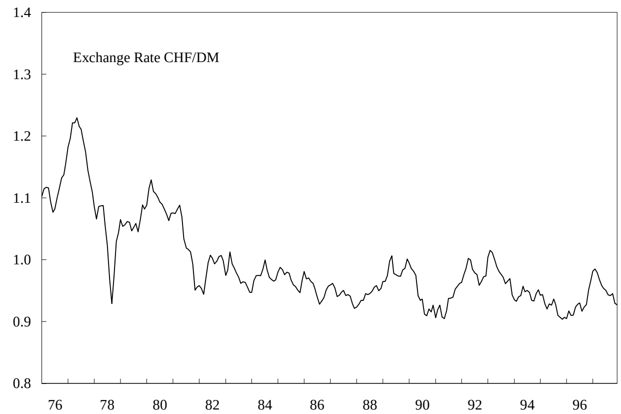


Figure 1 *Continued*



Note: Graphs represent series given in table 2. Source: Datastream, SNB, and Cuche and Hess (1999a, 1999b).

Hence, we decide not to use a vector error-correction mechanism (VECM) approach in this paper. We emphasize two special cointegration vectors that Clarida and Gertler (1997) found for Germany. They concern on one hand cointegration between money and industrial production (velocity of money) and on the other hand between retail sales and industrial production. We do not find the velocity vector cointegrated, but retail sales and GDP are cointegrated. Cuche and Hess (1999b) also found multiplicative cointegration at the quarterly level between retail sales and GDP. For the sake of simplicity, we do not use it here. We nevertheless estimate the setup without extraction within a VECM and notice that it does not perform better than the reported results in the next section.

We report data only for the whole period 1975-1997. This is worth mentioning that we do not re-estimate the first VAR in order to compute residuals for different subsamples. When we split the sample, we do it with the residuals calculated from the first VAR covering the whole considered range. Thus, we avoid considering the degrees of freedom problem known with short samples[21]. The estimation of the first VAR is performed with twelve lags ($k = 12$) to take into account the possible seasonality pattern of certain variables.

[21] An alternative would be to perform Bayesian VAR to improve our results with respect to the problem of degrees of freedom.

Table 3: Cointegration Test

$M0$ 75:10-97:12

A	H_0	H_a	LR	B	H_0	H_a	LR
A1	0	8	242.55*	B1	0	1	60.46*
A2	1	8	182.09*	B2	1	2	55.81*
A3	2	8	126.28*	B3	2	3	51.23*
A4	3	8	75.05**	B4	3	4	27.26**
A5	4	8	47.79**	B5	4	5	24.69**
A6	5	8	23.10	B6	5	6	15.04
A7	6	8	8.06	B7	6	7	5.20
A8	7	8	2.86	B8	7	8	2.86

$M1$ 75:10-97:12

A	H_0	H_a	LR	B	H_0	H_a	LR
A1	0	8	204.95*	B1	0	1	63.45*
A2	1	8	141.50*	B2	1	2	50.79*
A3	2	8	90.71	B3	2	3	29.54
A4	3	8	61.17	B4	3	4	26.03
A5	4	8	35.14	B5	4	5	17.53
A6	5	8	17.61	B6	5	6	12.43
A7	6	8	5.18	B7	6	7	5.17
A8	7	8	0.01	B8	7	8	0.01

Note: $M0$ = Test performed with monetary base; $M1$ = Test performed with monetary aggregate $M1$. H_0 = null hypothesis; H_a = alternative hypothesis; for each hypothesis, given figure is number of cointegration relations; LR = likelihood ratio statistic. Tests are run assuming linear trend in data and an intercept in the cointegrating equation and in the vector autoregression. Twelve lags are included. Test A, null hypothesis of h cointegrating relations against the alternative of no cointegration. LR is the weighted sum of the $(9-h)$-smallest eigenvalues. Test B, null hypothesis of h cointegrating relations against the alternative of $h+1$ relations. LR is the weighted h^{th} largest eigenvalue. Rejection of the null hypothesis at the 1% significance level (*) and at the 5% significance level (**).

3 Results

All our estimations are performed for five samples. We split the whole sample into four sections where we presume a changing behavior in SNB operating procedures. There is a first subsample before 1980. During this period, a strong appreciation of the CHF forced the SNB to massively intervene and to temporarily abandon monetary targeting. The second subsample goes from 1980 to 1987 when the SNB introduced a new payments system (SIC), and commercial banks faced new liquidity restrictions. The next subsample encloses the period 1988-1992 to see what happened after these legal improvements and the 1987 crash. Finally, the period since 1993 onwards is our last subsample. We expect to detect a diminishing role for aggregates during this period. All these sample cuts are based on presumed changes in the SNB's behavior raised using its official publications[22].

We present the results of various estimations and then look at plausible indicators carried by the data. All setups are estimated for both aggregates $M0$ and $M1$. We split them into two sections considering or not the extraction.

3.1 Without Extraction

Results for the model without extraction and its three regressions are given in table 4 for both considered monetary aggregates. Table 4 reports for each θ the estimated value and its t-statistic. All regressions are estimated by overidentified GMM using IV.

In general, reported results are poor, because it is not possible to find a setup, among the five different subsamples, that displays the expected coefficient signs. The use of two different aggregates does not change this deplorable image. It was only possible to achieve similar results as Clarida and Gertler (1997) obtained for Germany using a business cycle indicator, based on survey data, instead of the interpolated monthly Swiss GDP (results not reported)[23]. However, even with these better results, the performance of the regression still remains low. The German evidence, measured by the model without extraction, does not apply to Switzerland.

Furthermore, we show that the reported coefficients, in amplitude and direction, are not robust when we split the sample. This implies two contrary conclusions. On one hand, it means that the setup without extraction is not robust over sample changes. On the other hand, this lack of robustness strengthens the presence of changing operating procedures during the considered period.

[22] Quarterly bulletins and annual reports.

[23] Clarida and Gertler (1997) found orthodox signs for Germany, but many were not significant. For Switzerland with this business cycle indicator, we get for the twelve coefficients θ the following values: 0.05, 0.01, 3.48, 1008.16, -419.01, -0.46, -11.83, 0.64, -0.61, -3.12, 12.21, and 0.01.

Table 4: Estimation without Extraction

Whole Sample 75:10-97:12

$M0$ and $M1$

$$
\begin{aligned}
r_t^{cr} &= -\underset{(-1.8617)}{3.5949}\ r_t^{com} + \underset{(0.2268)}{2.7038}\ r_t^{mon_0} \\
&\quad + \underset{(0.7855)}{21.5214}\ r_t^{exr} + \varepsilon_t^{s} \\
r_t^{mon_0} &= \underset{(1.1610)}{0.1585}\ r_t^{gdp} - \underset{(-4.6653)}{0.0194}\ r_t^{cr} + \varepsilon_t^{d} \\
r_t^{exr} &= -\underset{(-0.8253)}{0.0281}\ r_t^{com} + \underset{(0.9342)}{0.0942}\ r_t^{gdp} + \underset{(1.6952)}{0.0464}\ r_t^{rs} \\
&\quad -\underset{(-1.1800)}{0.2934}\ r_t^{pl} + \underset{(1.0038)}{0.0142}\ r_t^{fcr} - \underset{(-5.0351)}{0.0139}\ r_t^{cr} \\
&\quad + \underset{(0.6666)}{0.0276}\ r_t^{mon_0} + \varepsilon_t^{x}
\end{aligned}
$$

$$
\begin{aligned}
r_t^{cr} &= -\underset{(-2.3859)}{2.4637}\ r_t^{com} - \underset{(-0.7756)}{8.1218}\ r_t^{mon_1} \\
&\quad + \underset{(0.2787)}{6.4518}\ r_t^{exr} + \varepsilon_t^{s} \\
r_t^{mon_1} &= -\underset{(-0.6735)}{0.0851}\ r_t^{gdp} + \underset{(0.1521)}{0.0004}\ r_t^{cr} + \varepsilon_t^{d} \\
r_t^{exr} &= -\underset{(-0.2853)}{0.0096}\ r_t^{com} + \underset{(0.1077)}{0.0109}\ r_t^{gdp} + \underset{(0.4113)}{0.0111}\ r_t^{rs} \\
&\quad -\underset{(-1.9501)}{0.4196}\ r_t^{pl} - \underset{(-0.6708)}{0.0087}\ r_t^{fcr} - \underset{(-1.2560)}{0.0024}\ r_t^{cr} \\
&\quad - \underset{(0.7098)}{0.0364}\ r_t^{mon_1} + \varepsilon_t^{x}
\end{aligned}
$$

Before 1980 75:10-79:12

$M0$ and $M1$

$$
\begin{aligned}
r_t^{cr} &= -\underset{(-2.0297)}{14.9174}\ r_t^{com} + \underset{(0.9041)}{16.9717}\ r_t^{mon_0} \\
&\quad + \underset{(0.0934)}{7.1309}\ r_t^{exr} + \varepsilon_t^{s} \\
r_t^{mon_0} &= \underset{(1.6547)}{1.2901}\ r_t^{gdp} - \underset{(-5.5868)}{0.1049}\ r_t^{cr} + \varepsilon_t^{d} \\
r_t^{exr} &= -\underset{(-0.1961)}{0.0201}\ r_t^{com} + \underset{(0.4619)}{0.1233}\ r_t^{gdp} + \underset{(0.1610)}{0.0136}\ r_t^{rs} \\
&\quad + \underset{(0.6142)}{0.4430}\ r_t^{pl} - \underset{(-0.3473)}{0.0093}\ r_t^{fcr} - \underset{(-0.3045)}{0.0029}\ r_t^{cr} \\
&\quad + \underset{(0.1457)}{0.0109}\ r_t^{mon_0} + \varepsilon_t^{x}
\end{aligned}
$$

$$
\begin{aligned}
r_t^{cr} &= -\underset{(-1.2182)}{3.1534}\ r_t^{com} - \underset{(-1.3293)}{15.8855}\ r_t^{mon_1} \\
&\quad + \underset{(0.3057)}{5.0338}\ r_t^{exr} + \varepsilon_t^{s} \\
r_t^{mon_1} &= -\underset{(-1.0514)}{0.2715}\ r_t^{gdp} - \underset{(-1.0072)}{0.0040}\ r_t^{cr} + \varepsilon_t^{d} \\
r_t^{exr} &= -\underset{(-0.6484)}{0.0756}\ r_t^{com} + \underset{(1.2264)}{0.4011}\ r_t^{gdp} + \underset{(0.7770)}{0.0667}\ r_t^{rs} \\
&\quad -\underset{(-0.3760)}{0.2686}\ r_t^{pl} - \underset{(-1.9876)}{0.0592}\ r_t^{fcr} - \underset{(-1.0736)}{0.0070}\ r_t^{cr} \\
&\quad + \underset{(1.2666)}{0.2840}\ r_t^{mon_1} + \varepsilon_t^{x}
\end{aligned}
$$

Table 4 *Continued*

Beginning Eighties 80:01-87:12

$M0$ and $M1$

$$
\begin{aligned}
r_t^{cr} &= -\underset{(-2.2957)}{5.5333}\, r_t^{com} + \underset{(0.1105)}{0.7154}\, r_t^{mon_0} \\
&\quad + \underset{(0.5413)}{9.4753}\, r_t^{exr} + \varepsilon_t^s \\
r_t^{mon_0} &= \underset{(1.7262)}{0.4146}\, r_t^{gdp} - \underset{(-2.6163)}{0.0090}\, r_t^{cr} + \varepsilon_t^d \\
r_t^{exr} &= \underset{(0.4640)}{0.0241}\, r_t^{com} - \underset{(-1.3820)}{0.1939}\, r_t^{gdp} + \underset{(2.8536)}{0.1130}\, r_t^{rs} \\
&\quad - \underset{(0.6651)}{0.1750}\, r_t^{pl} + \underset{(0.8129)}{0.0149}\, r_t^{fcr} - \underset{(-1.4879)}{0.0025}\, r_t^{cr} \\
&\quad - \underset{(-0.1731)}{0.0088}\, r_t^{mon_0} + \varepsilon_t^x
\end{aligned}
$$

$$
\begin{aligned}
r_t^{cr} &= -\underset{(-2.5798)}{6.2070}\, r_t^{com} - \underset{(-0.8020)}{7.6208}\, r_t^{mon_1} \\
&\quad + \underset{(0.3188)}{6.9093}\, r_t^{exr} + \varepsilon_t^s \\
r_t^{mon_1} &= -\underset{(-0.7087)}{0.1386}\, r_t^{gdp} + \underset{(0.8639)}{0.0030}\, r_t^{cr} + \varepsilon_t^d \\
r_t^{exr} &= \underset{(0.8761)}{0.0517}\, r_t^{com} - \underset{(-1.1889)}{0.1850}\, r_t^{gdp} + \underset{(2.0813)}{0.0877}\, r_t^{rs} \\
&\quad + \underset{(0.2532)}{0.0739}\, r_t^{pl} - \underset{(-0.3483)}{0.0065}\, r_t^{fcr} + \underset{(0.1886)}{0.0004}\, r_t^{cr} \\
&\quad + \underset{(0.7058)}{0.0570}\, r_t^{mon_1} + \varepsilon_t^x
\end{aligned}
$$

End Eighties 88:01-92:12

$M0$ and $M1$

$$
\begin{aligned}
r_t^{cr} &= -\underset{(-0.4705)}{0.6901}\, r_t^{com} - \underset{(-1.3811)}{13.7521}\, r_t^{mon_0} \\
&\quad + \underset{(2.2261)}{19.0009}\, r_t^{exr} + \varepsilon_t^s \\
r_t^{mon_0} &= -\underset{(-2.5990)}{0.7549}\, r_t^{gdp} + \underset{(2.4754)}{0.0258}\, r_t^{cr} + \varepsilon_t^d \\
r_t^{exr} &= \underset{(0.5101)}{0.0224}\, r_t^{com} + \underset{(2.0208)}{0.3300}\, r_t^{gdp} - \underset{(-0.5865)}{0.0281}\, r_t^{rs} \\
&\quad - \underset{(-2.9044)}{1.4588}\, r_t^{pl} + \underset{(2.3744)}{0.0620}\, r_t^{fcr} - \underset{(-2.6815)}{0.0159}\, r_t^{cr} \\
&\quad - \underset{(-3.0261)}{0.3917}\, r_t^{mon_0} + \varepsilon_t^x
\end{aligned}
$$

$$
\begin{aligned}
r_t^{cr} &= \underset{(0.0348)}{0.0711}\, r_t^{com} - \underset{(-2.1503)}{24.3981}\, r_t^{mon_1} \\
&\quad + \underset{(0.8152)}{10.4623}\, r_t^{exr} + \varepsilon_t^s \\
r_t^{mon_1} &= -\underset{(-1.5162)}{2.0646}\, r_t^{gdp} + \underset{(1.5653)}{0.1118}\, r_t^{cr} + \varepsilon_t^d \\
r_t^{exr} &= -\underset{(-0.2997)}{0.0191}\, r_t^{com} + \underset{(1.5163)}{0.2588}\, r_t^{gdp} - \underset{(-2.7252)}{0.0975}\, r_t^{rs} \\
&\quad - \underset{(-2.5703)}{1.2248}\, r_t^{pl} + \underset{(0.9379)}{0.0233}\, r_t^{fcr} - \underset{(-0.6252)}{0.0040}\, r_t^{cr} \\
&\quad - \underset{(-1.9024)}{0.1978}\, r_t^{mon_1} + \varepsilon_t^x
\end{aligned}
$$

Table 4 *Continued*

After 1993 93:01-97:12

$M0$ and $M1$

$$r_t^{cr} = -\underset{(-0.1028)}{0.3218}\, r_t^{com} + \underset{(0.6580)}{10.3367}\, r_t^{mon_0} + \underset{(1.1665)}{23.8230}\, r_t^{exr} + \varepsilon_t^s$$

$$r_t^{mon_0} = \underset{(0.8779)}{0.2189}\, r_t^{gdp} - \underset{(-4.4833)}{0.0496}\, r_t^{cr} + \varepsilon_t^d$$

$$r_t^{exr} = -\underset{(-1.0907)}{0.1270}\, r_t^{com} + \underset{(0.6777)}{0.1521}\, r_t^{gdp} + \underset{(0.5632)}{0.0357}\, r_t^{rs} - \underset{(-1.8906)}{1.6102}\, r_t^{pl} + \underset{(0.1983)}{0.0100}\, r_t^{fcr} - \underset{(-1.7978)}{0.0204}\, r_t^{cr} + \underset{(2.6954)}{0.5471}\, r_t^{mon_0} + \varepsilon_t^x$$

$$r_t^{cr} = -\underset{(-0.0161)}{0.0272}\, r_t^{com} - \underset{(-1.2092)}{10.1301}\, r_t^{mon_1} - \underset{(-0.6194)}{8.2851}\, r_t^{exr} + \varepsilon_t^s$$

$$r_t^{mon_1} = \underset{(1.0010)}{0.2626}\, r_t^{gdp} + \underset{(1.7735)}{0.0190}\, r_t^{cr} + \varepsilon_t^d$$

$$r_t^{exr} = -\underset{(-0.5661)}{0.0401}\, r_t^{com} - \underset{(-0.9105)}{0.1916}\, r_t^{gdp} + \underset{(0.0610)}{0.0035}\, r_t^{rs} - \underset{(-1.5496)}{1.0038}\, r_t^{pl} + \underset{(0.6530)}{0.0218}\, r_t^{fcr} + \underset{(3.5733)}{0.0187}\, r_t^{cr} + \underset{(2.9855)}{0.3074}\, r_t^{mon_1} + \varepsilon_t^x$$

Note: $M0$ = Estimated with monetary base; $M1$ = Estimated with monetary aggregate $M1$; r_t^{com} = Commodity price index; r_t^{gdp} = Gross domestic product; r_t^{rs} = Value of retail sales; r_t^{pl} = Price level index; r_t^{fcr} = German call rate; r_t^{cr} = Call rate; $r_t^{mon_0}$ = Real monetary base; $r_t^{mon_1}$ = Real $M1$; r_t^{exr} = Real exchange rate (Deutschmark). t-values are given in parentheses. All the coefficients are estimated by GMM with IV. IV are in the first equation: r_t^{com}, r_t^{gdp}, r_t^{rs}, r_t^{pl}, and r_t^{fcr}; in the second equation: r_t^{com}, r_t^{gdp}, r_t^{rs}, r_t^{pl}, r_t^{fcr}, and ε_t^s; in the third equation: r_t^{com}, r_t^{gdp}, r_t^{rs}, r_t^{pl}, r_t^{fcr}, ε_t^s , and ε_t^d.

We further do not recognize the assumed fight against inflation pressures in the money supply equation. Similarly, the opportunity cost in the money demand equation sometimes appears with a positive coefficient which is inconceivable with our economic intuition. All these coefficients are not robust for alternative specifications of the first VAR as well. We estimated the same setup as Clarida and Gertler (1997) with lags 1-6, 9, and 12, in order to best solve the degrees of freedom problem, but still without succeeding in finding plausible results (not reported). In addition, we were not more successful with other lag specifications.

A potential explanation for these poor results is the econometric flaws described in the section concerning the generated instruments.

Moreover, when the reaction function of the bank in innovation form (7) is misspecified, the subsequent use of its residuals as instruments can only worsen the next estimations. A second explanation are IV themselves. They remove the endogeneity problem faced by the nonorthogonal residuals $\mathbf{r}$ used in the different regressions, but they do not clean the residuals of nonpolicy influences. Only the next approach with extraction does. Finally, and this is the most important reason, the assumption about the call rate as a unique gauge of overall monetary policy is probably too strong for Swiss data.

Henceforth, we suggest that this model with the call rate as a measure of monetary policy cannot portray a suitable overall indicator for Swiss monetary policy. Based on these regressions, we reject the call rate as a single indicator of monetary policy for the period 1975-1997. We also note that this specification without extraction cannot outperform two Cholesky scenarios (not reported) where a plausible interpretation is also impossible[24].

3.2 With Extraction

Before commenting the results with extraction, we briefly look at statistical properties of the two sets of series $\mathbf{r}^{\underline{z}}$ and $\mathbf{u}^{\underline{z}}$. While their theoretical origins are clear, it is not the case about their statistical features. A plain statistical analysis of series $\mathbf{r}^{\underline{z}}$ and $\mathbf{u}^{\underline{z}}$ is not able to show striking differences. We could indeed almost assume that their generating processes are the same. Henceforth, we think that only an economic interpretation about these two vectors makes sense - cleaned or not from nonpolicy influences - and that a pure statistical focus is aimless.

The same unsatisfactory feeling appears when we want to give an economic interpretation to the comparison of the implied indicator without extraction with respect to the different indicators with

[24] In addition to the frameworks without and with extraction and for the sake of comparison, we also identify our system (1) with two Cholesky decompositions using vector $\mathbf{z}$ ordered as presented for the first VAR. The goal is to have a simple dynamics to qualify the improvements of our more complicated models. This comparison was initiated in Switzerland by Jordan (1998). The first Cholesky identification (Cholesky diagonal) is a three-matrix triangular decomposition of $(\mathbf{I}_{m+n} - \mathbf{A}_0)^{-1}\mathbf{B}$. It imposes a diagonal normalization of one on the triangular matrices and a nonidentity diagonal variance-covariance matrix $\mathbf{\Sigma}$. The second one (Cholesky identity) is slightly different because we assume that $\mathbf{\Sigma} = \mathbf{I}_{m+n}$, removing the normalization of one on the diagonals of triangular matrices. We do not report the results due to poor estimations.

extraction presented in table 1. This is virtually insurmountable. It clearly illustrates that similar equations, e.g. money demand in innovation form in both frameworks, can denote different dynamics and indicator constructions that we cannot differentiate any more with simple economic thinking. A comforting decision is to reject the model without extraction, thus to stop further investigating this comparison.

3.2.1 Just-Identified and Overidentified Estimations

We report first our results for the two just-identified setups and then for the three overidentified cases. Table 5 shows the results using the just-identified framework. We report our results for the five considered samples.

Table 5 contains the different coefficients and the corresponding matrices $\mathbf{A}^{zz}_{0,-1}\mathbf{B}^{z} = \mathbf{H}$ linking extracted VAR residuals to structural shocks. We see that these results are not robust relative to the used aggregates, $M0$ or $M1$. Still searching for robustness, splitting the sample disturbs the image given by the whole sample and thus shows that these results are not robust over time with heavy corrections for the opportunity costs of money demand, the exchange rate coefficients, and the money supply parameters.

Table 5: Just-Identified Estimation

Whole Sample 75:10-97:12, $\delta = 0$ (top) and $\lambda = 0$ (bottom)

	$M0$			$M1$	
λ	ϕ	ρ	λ	ϕ	ρ
0.5873	−0.0928	−0.0431	0.6167	−0.0861	0.0269
	H			**H**	
−23.2019	9.5754	2.1531	37.1747	−14.2491	−3.2007
1.0000	0.5873	−0.0928	1.0000	0.6167	−0.0861
0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
δ	ϕ	ρ	δ	ϕ	ρ
0.0013	−0.0711	−0.1982	0.0034	0.1255	−0.1399
	H			**H**	
−5.0454	5.0454	0.3587	−7.1480	7.1480	−0.8971
1.0000	0.0000	−0.0711	1.0000	0.0000	0.1255
−0.0066	0.0066	1.0005	−0.0243	0.0243	0.9969

Table 5 *Continued*

Before 1980 75:10-79:12, $\delta = 0$ (top) and $\lambda = 0$ (bottom)

	$M0$			$M1$	
λ	ϕ	ρ	λ	ϕ	ρ
0.0793	−0.1279	−0.1658	0.3479	0.1488	−0.0430
	H			**H**	
−6.0314	5.5531	0.7714	−23.2558	15.1651	−3.4605
1.0000	0.0793	−0.1279	1.0000	0.3479	0.1488
0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
δ	ϕ	ρ	δ	ϕ	ρ
0.0002	−0.1194	−0.2024	−0.0027	0.1169	−0.0798
	H			**H**	
−4.9407	4.9407	0.5899	−12.5313	12.5313	−1.4649
1.0000	0.0000	−0.1194	1.0000	0.0000	0.1169
−0.0010	0.0010	1.0001	0.0338	−0.0338	1.0040

Beginning Eighties 80:01-87:12, $\delta = 0$ (top) and $\lambda = 0$ (bottom)

	$M0$				$M1$
λ	ϕ	ρ	λ	ϕ	ρ
0.3283	0.1091	−0.0481	0.8602	0.0598	0.0091
	H			**H**	
−20.7900	13.9647	−2.2682	109.8901	−15.3626	6.5714
1.0000	0.3283	0.1091	1.0000	0.8602	0.0598
0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
δ	ϕ	ρ	δ	ϕ	ρ
0.0006	−0.0986	−0.2162	0.0029	0.0866	−0.1891
	H			**H**	
−4.6253	4.6253	0.4561	−5.2882	5.2882	−0.4580
1.0000	0.0000	−0.0986	1.0000	0.0000	0.0866
−0.0028	0.0028	1.0003	−0.0153	0.0153	0.9987

End Eighties 88:01-92:12, $\delta = 0$ (top) and $\lambda = 0$ (bottom)

	$M0$			$M1$	
λ	ϕ	ρ	λ	ϕ	ρ
0.1857	−0.2604	−0.0783	1.0226	−0.1081	−0.0231
	H			**H**	
−12.7714	10.3997	3.3257	−43.2900	−0.9784	4.6797
1.0000	0.1857	−0.2604	1.0000	1.0226	−0.1081
0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
δ	ϕ	ρ	δ	ϕ	ρ
0.0021	0.2364	−0.1582	0.0057	0.0116	−0.0985
	H			**H**	
−6.3211	6.3211	−1.4943	−10.1523	10.1523	−0.1178
1.0000	0.0000	0.2364	1.0000	0.0000	0.0116
−0.0133	0.0133	0.9969	−0.0579	0.0579	0.9993

Table 5 *Continued*

After 1993 93:01-97:12, $\delta = 0$ (top) and $\lambda = 0$ (bottom)

$M0$			$M1$		
λ	ϕ	ρ	λ	ϕ	ρ
0.7771	0.0450	0.0249	0.1939	0.2281	0.1321
	H			**H**	
40.1606	−8.9518	1.8072	7.5700	−6.1022	1.7267
1.0000	0.7771	0.0450	1.0000	0.1939	0.2281
0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
δ	ϕ	ρ	δ	ϕ	ρ
0.0041	0.0704	−0.2436	0.0049	0.2568	−0.2937
	H			**H**	
−4.1051	4.1051	−0.2890	−3.4048	3.4048	−0.8744
1.0000	0.0000	0.0704	1.0000	0.0000	0.2568
−0.0168	0.0168	0.9988	−0.0167	0.0167	0.9957

Note: $M0$ = Estimated with monetary base. $M1$ = Estimated with monetary aggregate $M1$. Equations with extraction: $u_s^{mon} = u_d^{mon}$, $u_s^{mon} = \lambda\varepsilon^d + \phi\varepsilon^x + \varepsilon^s$, $u_d^{mon} = \rho u^{cr} + \varepsilon^d$, $u^{exr} = \delta u^{cr} + \varepsilon^x$.

This non-robustness leads to the same conclusion as in the case without extraction. When we split the sample, changing results confirm the need to allow the model to catch different setups for different samples. Moreover, we are not able to discriminate between the two just-identified setups. This failure is then corrected with the overidentified setup where we calculate an overidentification statistic, the J-statistic based on Hansen (1982), giving thus a way of sorting out different scenarios.

We report our results for overidentified cases in table 6. For each scenario we report the estimated coefficients, the assumptions, matrix **H** linking extracted residuals to structural shocks, and the Hansen (1982)-J-statistic[25]. Despite some shortcomings of Hansen (1982)-J-statistic[26], it remains an important selection mechanism among overidentified cases. We report the value of the minimized function and multiply the J-statistic by the size of the considered sample. This new statistic is χ_1^2-distributed due to a first-order overidentification. We see that for the samples 1980-1987 and 1988-1992, we cannot reject the null hypothesis that the overidentifying restrictions are satisfied at the 5% significance level. For other samples, overidentifications are not accepted at the 5% significance level. On

[25] Our overidentified system satisfies rank and order conditions.

[26] Hansen (1982)-J-test can easily fail to detect a misspecified model. See Hamilton (1994) for more details.

the other hand, when we consider a less rigorous significance level, J-statistics keep going to be comparable and useful. We thus decide to also use them as a selection mechanism for the periods before 1980 and after 1993.

Table 6: Overidentified Estimation

Whole Sample 75:10-97:12, BR, CR, ER (top to bottom)

$M0$			$M1$		
δ	ρ	J	δ	ρ	J
−1.0500	−7.1250	0.2667	−2.3650	−0.0215	0.4821
	H			**H**	
−0.1404	0.1404	0.0000	−46.5116	46.5116	0.0000
1.0000	0.0000	0.0000	1.0000	0.0000	0.0000
0.1474	−0.1474	1.0000	110.0000	−110.0000	1.0000
δ	ρ	J	δ	ρ	J
−3.4123	−5.7840	0.3966	−2.1249	−4.1265	0.3307
	H			**H**	
−0.1729	0.0000	0.0000	−0.2423	0.0000	0.0000
1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
0.5900	0.0000	1.0000	0.5149	0.0000	1.0000
δ	ρ	J	δ	ρ	J
8.2350	−2.8950	0.2672	−7.5470	−2.4555	0.3492
	H			**H**	
−0.3454	0.0000	−0.1214	−0.4072	0.0000	0.1325
1.0000	1.0000	0.3515	1.0000	1.0000	−0.3254
−2.8446	0.0000	0.0000	3.0735	0.0000	0.0000

Before 1980 75:10-79:12, BR, CR, ER (top to bottom)

$M0$			$M1$		
δ	ρ	J	δ	ρ	J
−0.7456	−14.5200	1.0039	4.2560	−7.1250	0.8822
	H			**H**	
−0.0689	0.0689	0.0000	−0.1404	0.1404	0.0000
1.0000	0.0000	0.0000	1.0000	0.0000	0.0000
0.0513	−0.0513	1.0000	−0.5973	0.5973	1.0000
δ	ρ	J	δ	ρ	J
−1.0500	−8.1250	1.0284	−1.0230	−3.2560	0.6620
	H			**H**	
−0.1231	0.0000	0.0000	−0.3071	0.0000	0.0000
1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
0.1292	0.0000	1.0000	0.3142	0.0000	1.0000
δ	ρ	J	δ	ρ	J
7.8920	−0.2120	0.4611	7.2145	−0.6570	0.4607
	H			**H**	
−4.7170	0.0000	−0.1267	−1.5221	0.0000	−0.1386
1.0000	1.0000	0.0269	1.0000	1.0000	0.0911
−37.2264	0.0000	0.0000	−10.9810	0.0000	0.0000

Table 6 *Continued*

Beginning Eighties 80:01-87:12, BR, CR, ER (top to bottom)

$M0$			$M1$		
δ	ρ	J	δ	ρ	J
0.0010	−4.9100	0.0109	0.0030	−4.8125	0.0196
	H			**H**	
−0.2037	0.2037	0.0000	−0.2078	0.2078	0.0000
1.0000	0.0000	0.0000	1.0000	0.0000	0.0000
−0.0002	0.0002	1.0000	−0.0006	0.0006	1.0000
δ	ρ	J	δ	ρ	J
−0.4125	−4.5550	0.3292	−1.2756	−4.3129	0.4132
	H			**H**	
−0.2195	0.0000	0.0000	−0.2319	0.0000	0.0000
1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
0.0906	0.0000	1.0000	0.2958	0.0000	1.0000
δ	ρ	J	δ	ρ	J
6.2545	−2.9545	0.3329	−6.9878	−2.9245	0.4734
	H			**H**	
−0.3385	0.0000	−0.1599	−0.3419	0.0000	0.1431
1.0000	1.0000	0.4724	1.0000	1.0000	−0.4185
−2.1169	0.0000	0.0000	2.3894	0.0000	0.0000

End Eighties 88:01-92:12, BR, CR, ER (top to bottom)

$M0$			$M1$		
δ	ρ	J	δ	ρ	J
2.8500	−0.0350	0.5045	0.0040	−0.0800	0.0138
	H			**H**	
−28.5714	28.5714	0.0000	−12.5000	12.5000	0.0000
1.0000	0.0000	0.0000	1.0000	0.0000	0.0000
−81.4286	81.4286	1.0000	−0.0500	0.0500	1.0000
δ	ρ	J	δ	ρ	J
3.2540	−6.9458	0.6953	−0.0900	−8.5680	0.7280
	H			**H**	
−0.1440	0.0000	0.0000	−0.1167	0.0000	0.0000
1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
−0.4685	0.0000	1.0000	0.0105	0.0000	1.0000
δ	ρ	J	δ	ρ	J
−6.9988	−3.2500	0.7086	−7.1524	−3.2002	0.6898
	H			**H**	
−0.3077	0.0000	0.1429	−0.3125	0.0000	0.1398
1.0000	1.0000	−0.4644	1.0000	1.0000	−0.4474
2.1535	0.0000	0.0000	2.2350	0.0000	0.0000

Table 6 *Continued*

After 1993 93:01-97:12, BR, CR, ER (top to bottom)

$M0$			$M1$		
δ	ρ	J	δ	ρ	J
−2.5550	−12.0500	0.6953	−1.5588	−6.9100	0.5646
	H			**H**	
−0.0830	0.0830	0.0000	−0.1447	0.1447	0.0000
1.0000	0.0000	0.0000	1.0000	0.0000	0.0000
0.2120	−0.2120	1.0000	0.2256	−0.2256	1.0000
δ	ρ	J	δ	ρ	J
0.0300	−0.9458	0.3707	5.2360	−1.1354	0.3745
	H			**H**	
−1.0573	0.0000	0.0000	−0.8807	0.0000	0.0000
1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
−0.0317	0.0000	1.0000	−4.6116	0.0000	1.0000
δ	ρ	J	δ	ρ	J
−2.1540	−0.8157	0.3752	2.5890	−2.100	0.3769
	H			**H**	
−1.2259	0.0000	0.4643	−0.4762	0.0000	−0.3862
1.0000	1.0000	−0.3787	1.0000	1.0000	0.8111
2.6407	0.0000	0.0000	−1.2329	0.0000	0.0000

Note: $M0$ = Estimated with monetary base. $M1$ = Estimated with monetary aggregate $M1$. BR = Bank reserves targeting; CR = Call rate targeting; ER = Exchange rate targeting. Equations with extraction: $u_s^{mon} = u_d^{mon}$, $u_s^{mon} = \lambda\varepsilon^d + \phi\varepsilon^x + \varepsilon^s$, $u_d^{mon} = \rho u^{CR} + \varepsilon^d$, $u^{exr} = \delta u^{cr} + \varepsilon^x$.

First of all, results for the whole sample confirm the need to allow for more flexibility in the model in order to catch changing procedures over time. Moreover, J-statistics are quite similar, showing the difficulty to focus on a single sample. Splitting our sample, we have the opportunity to select the setup catching best what residuals represent for each subsample. We then use the J-test to do it. For the sample before 1980, this implies that we select the model with exchange rate targeting constructed with $M1$. For the two subsequent subsamples, we unambiguously select the bank reserves targeting model. For the period 1980-1987, it does not matter whether we use $M0$ or $M1$. This indicates that this period was the golden age of monetary targeting with a relationship between aggregates and price level stable over time. For the period 1988-1992, we still choose the bank reserves targeting setup, but now only produced with the model using $M1$. The rejection of the null hypothesis for the model using $M0$ confirms the changing environment at the end of the eighties. The new electronic payments system caused a radical change in the base demand and implied a strategic re-orientation towards $M1$. Finally, for the last sample after 1993, we select the call rate model even if the J-statistic produced by the model assuming exchange rate targeting is quite similar.

3.2.2 Dynamics

Before computing an indicator based on these results and interpreting it, we turn to the model dynamics. We notice that the first column of the matrix linking $\mathbf{u}$ to ε, influencing the dynamics of the economy after an exogenous monetary shock, is theoretically similar for all setups and empirically quite near due to similar coefficients. This is purposely a feature of our nesting model, because exogenous monetary shocks should affect the economy independently of the assumed scenarios about operating procedures. However, the second and third column of this same matrix depend on our different hypotheses and do not display the same dynamics after demand and exchange rate shocks for each operating scenario. It is thus tempting to use IRF after such shocks to strengthen our choices based on J-statistics only.

Surprisingly, with respect of the positive results of the estimation, we discover mixed evidence regarding IRF after a monetary shock ε^s. We display these IRF after an expansionary monetary shock for both aggregates in figures 2 and 3. Figure 2 concerns IRF using M0 and the best overidentified model for the whole sample, namely the bank reserves targeting model.

Figure 2: IRF with M0 and Bank Reserves Targeting

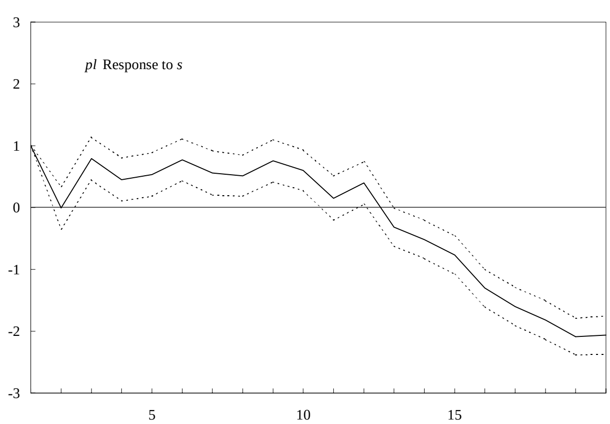


Figure 2 *Continued*

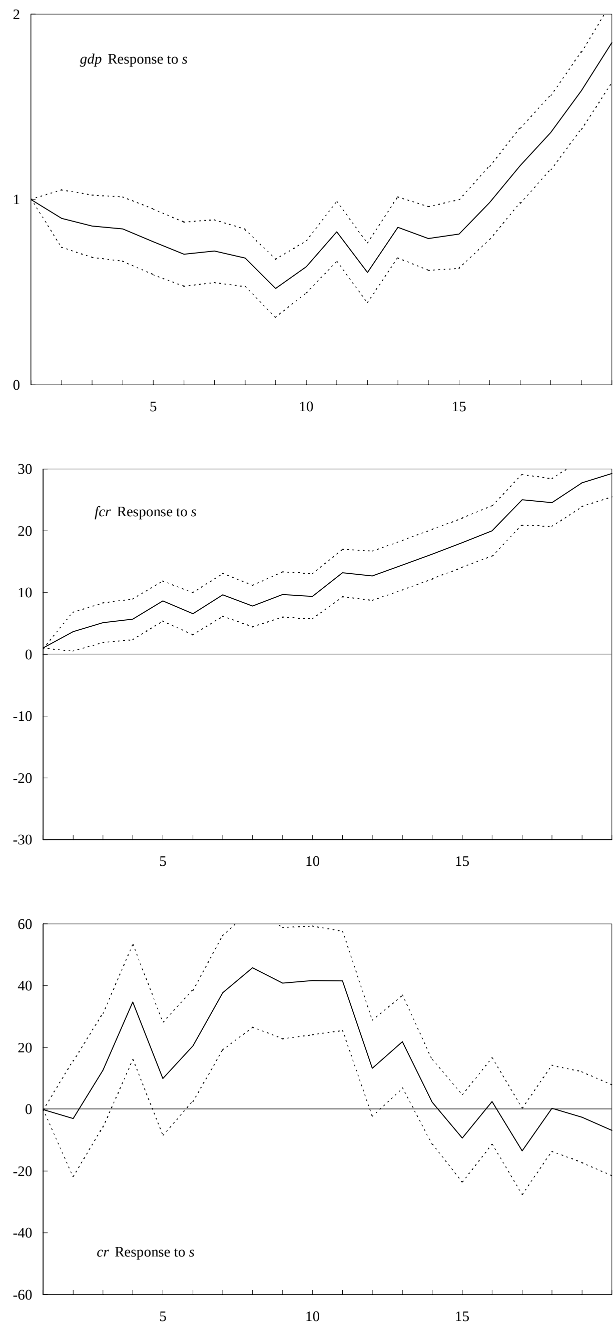
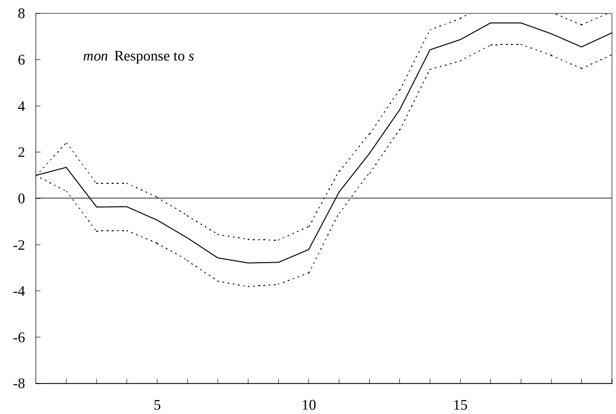
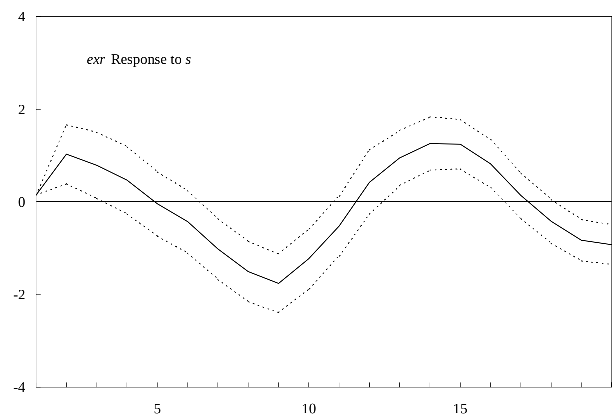


Figure 2 *Continued*





Note: IRF = Impulse response function. IRF are plotted with a 95 % confidence interval.

Figure 3 plots IRF using M1 and the best overidentified model for the whole sample, namely the call rate targeting model. Reported dynamics is quite poor according to three aspects[27].

First, the reaction of the foreign call rate is too high. This is not plausible to assume such an influence of Switzerland on Germany. Second, the IRF are particularly puzzling about the huge reaction of the interest rate, where we would expect a liquidity effect, implying

[27] It is worth mentioning that the number of variables could play a role, while well-behaved VAR models generally have less variables than ours. We also note that the presence of a commodity price index does not solve the puzzles.

a decrease in the interest rate. We disappointingly observe that the liquidity puzzle is linked to a marginal increase in money. Third, when we allow the model to best catch a particular scenario, we open the way to different dynamics for each setup. This is a feature and disadvantage of this approach confirmed by both examples reported in figures 2 and 3.

Based on this mixed evidence about the dynamics after a monetary shock, we give up using IRF after demand and exchange rate shocks as a selection mechanism. This puzzling dynamics does not shade the results based on the overidentified mechanism. It however reveals that the model as a whole, and not only the exogenous analysis, is not correctly specified to analyze phenomena as the transmission mechanism.

Figure 3: IRF with M1 and Call Rate Targeting

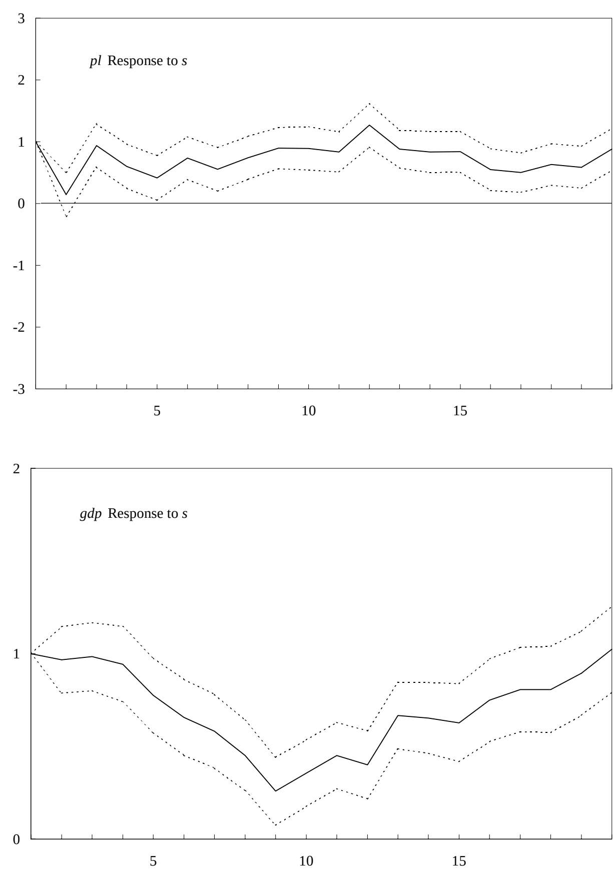


Figure 3 *Continued*

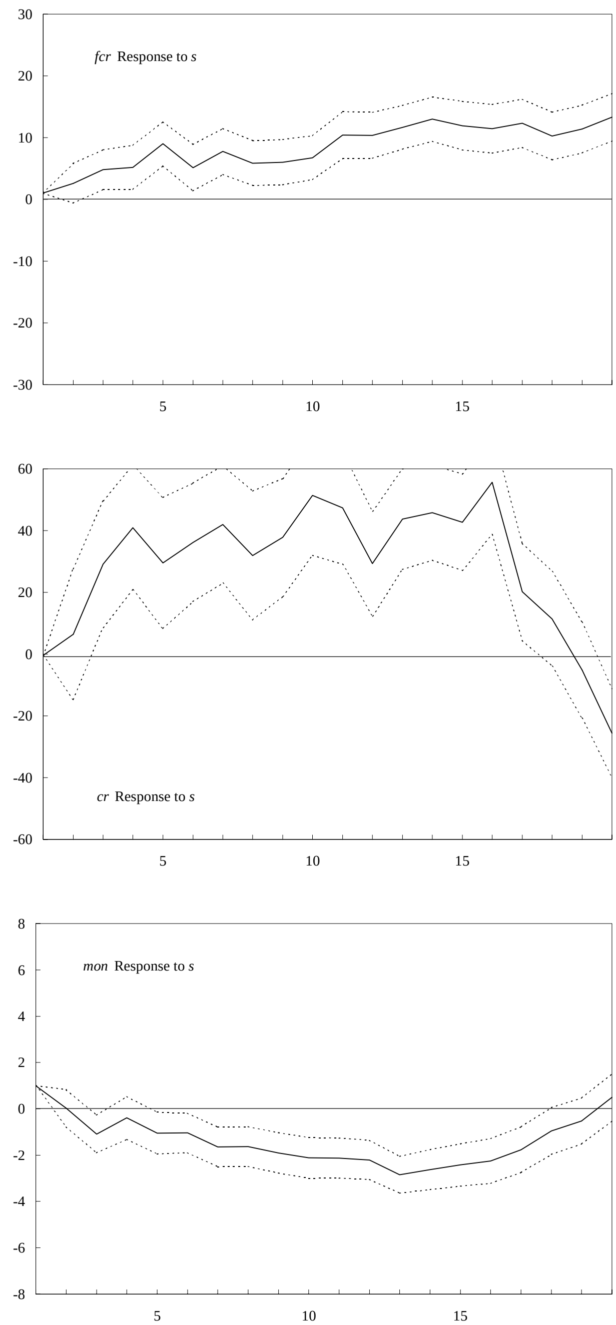
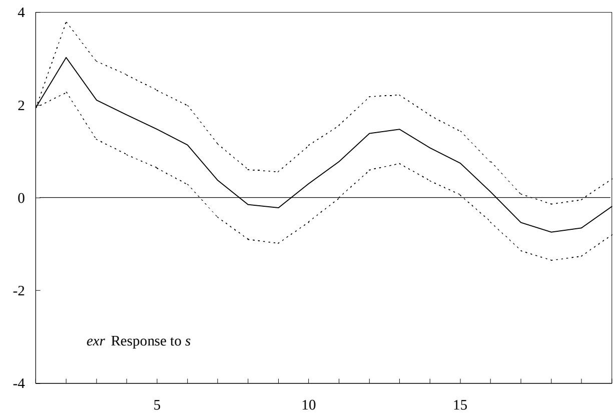


Figure 3 *Continued*



Note: IRF = Impulse response function. IRF are plotted with a 95 % confidence interval.

3.2.3 A Main Indicator

We compute an indicator for our subsamples based on the results of the overidentified setup and report it in figure 4. Figure 4 shows the indicator, both smoothed and normalized to be represented in a single figure[28]. The normalization allows the indicator to be comparable over the whole period. Thus, we have to interpret the indicator, its size and its direction, with respect to the average stance that is, per definition, also normalized to zero. There is a trade-off between clarity and accuracy in this indicator construction. Our goal is to display the indicator on a single plot for the whole considered sample. This is only possible after some normalization, and it comes at a cost of accuracy with respect to raw indicator figures. In figure 4, vertical lines mark the subsamples where we have different models arising from overidentified estimations. The first sample is based on exchange rate targeting, the next two periods are based on reserves targeting, and finally the last one is based on call rate targeting.

Compared to the traditional aggregate $M0$ as an indicator, we see that statistical methods clearly confirm the use of aggregates for the eighties. On the other hand, these same methods reveal for the end

[28] We subtracted a moving average of the last 12 months from the original values of our indicator. We moreover normalized the indicator in order to have a variance of one and a mean of zero for the whole sample.

of the seventies and for the nineties other indicators. These methods thus show the periods where the SNB officially explained that it deviated from its monetary targeting. It is true for the exchange rate targeting strategy before 1980 when the SNB temporarily stopped fixing objectives in terms of aggregate due to turbulences on the financial markets. The SNB had to massively react to a CHF appreciation in particular with respect to the DM. This is however less clear for the period after 1992 when the SNB was reluctant to admit that it focused on other variables, in particular the call rate. This revelation is probably explained by the changing announcement policy in the early nineties[29], a changing relationship between aggregates and price level, and finally once again a CHF appreciation. This is worth mentioning that the model using an exchange rate strategy in the period after 1992 almost succeeded in passing the selection test.

Figure 4: Indicator 1977-1997

Note: Vertical dashed lines signal changing periods.

Before 1980, our indicator captures the expansionary policy due to the CHF appreciation. However, because the indicator is function

[29] We do not discover a changing behavior right after the implementation of the new announcement policy. We catch this change from the beginning of 1992 onwards.

of the external value of the CHF, it suffers from a short delay compared to the stance announcement made at that time by the SNB. This problem is less significant for other periods. Concerning the golden age of monetary targeting, we see that the beginning of the eighties was quite restrictive following a CHF probably too weak. At the same time, the elimination of restrictions on capital imports also resulted in a diminishing demand for money, that was partially accommodated. During the eighties, movements of the indicator correspond more or less to the official announcements. Moreover, the indicator is also able to catch the turbulences in the mid-eighties, the CHF appreciation, and the effects of the crash, all met by an expansionary policy. For the period after 1990, still with the reserves targeting model up until 1992, and then with the call rate model, we have some difficulty to interpret the movements in form of high swings in the indicator. This is particularly true for the period after 1995 where we have the impression, if we trust the indicator, that the SNB did not follow a consequent and constant policy. For this last section of the path, we think however that our produced indicator is not very accurate. The method to calculate the indicator, based on the call rate targeting strategy, unfortunately exacerbates the indicator swings.

4 Conclusion

Our framework nests models that use VAR residuals in order to identify monetary policy and produce overall stance indicators for Swiss monetary policy. The contribution of this paper is threefold.

First, our model nests two approaches similar in economic terms and different in the treatment of VAR residuals. Differences proceed rather from econometric considerations than economic ones. With the results provided by the estimations, we then realize that the method with extraction performs better than the one without.

Second, the setup without extraction cannot produce good results due to econometric and economic flaws in its specification. The use of 'polluted' monetary residuals in modeling operating procedures cannot generate elaborate conclusions about the stance of Swiss monetary policy. We show that the shortcomings of this approach are not function of the Swiss data but are more general.

Third, we produce a new indicator for overall Swiss monetary policy with help of the identification based on overidentified models with extraction. While these models are not able to accurately analyze the mechanism of transmission, statistical methods allow confirming SNB strategic decisions during these last fifteen years. In particular, our indicator catches changing operating procedures over time at the end of the seventies and during the eighties. Our results state that the period before 1980 was conducted following an exchange rate targeting strategy. During the eighties, bank reserves targeting was the leading strategy. We call this period the golden age of monetary targeting. Finally, the last period, since 1993 onwards, was guided by a call rate targeting strategy.

References

Amisano, G. and Giannini, C. (1997). *Topics in Structural VAR Econometrics*, Springer.

Bernanke, B. S. and Blinder, A. S. (1992). The federal fund rate and the channels of monetary transmission, *American Economic Review* 82(4): 901–921.

Bernanke, B. S. and Mihov, I. (1997). What does the Bundesbank target?, *European Economic Review* 41(6): 1025–1053.

Bernanke, B. S. and Mihov, I. (1998). Measuring monetary policy, *Quarterly Journal of Economics* 113(3): 869–902.

Birchler, U. W. (1988). Neue Liquiditätsvorschriften und Geldpolitik, *Quarterly Bulletin, Swiss National Bank* 6(1): 75–81.

Bisignano, J. (1996). Varieties of monetary policy operating procedures: Balancing monetary objectives with market efficiency, *Working paper 35*, BIS.

Borio, C. E. V. (1997a). The implementation of monetary policy in industrial countries: A survey, *Economic paper 47*, BIS.

Borio, C. E. V. (1997b). Monetary policy operating procedures in industrial countries, *Conference papers 3, Implementation and Tactics of Monetary Policy*, BIS.

Chinn, M. D. and Dooley, M. P. (1997). Monetary policy in Japan, Germany and the United States: Does one size fit all?, *Working paper 6092*, NBER.

Clarida, R. and Gertler, M. (1997). How the Bundesbank conducts monetary policy, *Reducing Inflation, Motivation and Strategy ed. by C. D. Romer and D. H. Romer*, NBER Studies in Business Cycles 30, University of Chicago and NBER.

Cochrane, J. H. (1994). Shocks, *Carnegie-Rochester Conference Series on Public Policy* 41: 295–364.

Cuche, N. A. and Hess, M. K. (1999a). Eine monatliche Datenreihe für das schweizerische Bruttoinlandprodukt, Konjunkturindikator und Hilfsgrösse für die Forschung, *Die Volkswirtschaft, Magazin für Wirtschaftspolitik* 72(9): 32–33.

Cuche, N. A. and Hess, M. K. (1999b). Estimating monthly GDP in a general Kalman filter framework: Evidence from Switzerland, *Working paper 9902*, Study Center Gerzensee.

Grossman, H. I. (1991). Monetary economics: A review essay, *Journal of Monetary Economics* 28(2): 323–345.

Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica* 50: 1029–1054.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* 59(6): 1551–80.

Jordan, T. J. (1998). Monetary policy shocks in Switzerland: Is there a liquidity effect? mimeo.

King, R. G. and Watson, M. W. (1997). Testing long-run neutrality, *Economic Quarterly, Federal Reserve Bank of Richmond* 83(3): 69–99.

Landmann, O. and Jerger, J. (1997). 20 Jahre Geldmengensteuerung in der Schweiz: Erfahrungen und Alternativen. mimeo.

Laubach, T. and Posen, A. S. (1997). Disciplined discretion: The German and Swiss monetary targeting frameworks in operation, *Research paper 9707*, Federal Reserve Bank of New York.

Lengwiler, Y. (1997). Der 'Monetary Conditions Index' für die Schweiz, *Quarterly Bulletin, Swiss National Bank* 15(1): 61–72.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors, *International Economic Review* 25(1): 221–247.

Rich, G. (1997). Monetary targets as a policy rule: Lessons from Swiss experience, *Journal of Monetary Economics* 39(1): 113–141.

Sarte, P.-D. G. (1997). On the identification of structural VAR, *Economic Quarterly, Federal Reserve Bank of Richmond* 83(3): 45–67.

Sims, C. A. (1980). Macroeconomics and reality, *Econometrica* 48(1): 1–48.

Spörndli, E. and Moser, D. (1997). Monetary policy operating procedures in Switzerland, *Conference papers 3, Implementation and Tactics of Monetary Policy*, BIS.

Swank, J. and Velden, L. V. (1997). Instruments, procedures and strategies of monetary policy: An assessment of possible relationships for 21 OECD countries, *Conference papers 3, Implementation and Tactics of Monetary Policy*, BIS.

Vital, C. (1998). Swiss Interbank Clearing (SIC): Erfahrungen und Einschätzungen, *Banktopologie ed. by H. Geiger and K. Spremann, Bank- und finanzwirtschaftliche Forschungen*, Haupt.

Zha, T. (1997). Identifying monetary policy: A primer, *Economic Review, Federal Reserve Bank of Atlanta* 82(2): 26–43.

Appendix

Appendix A First Step SVAR(k)

Our model is based on a two-step SVAR. In this appendix, we estimate a SVAR(k) and store its RF residuals. In a second step (Appendixes B and C), they become explained and explanatory variables in a SVAR($\mathbf{0}$). We construct the economy RF using either matrix algebra with vector $\mathbf{z}$ (containing all the variables) or partitioned matrix algebra with two vectors $\overline{\mathbf{z}}$ (m nonpolicy variables) and $\underline{\mathbf{z}}$ (n policy variables), such as $\mathbf{z}_t = \begin{pmatrix} \overline{\mathbf{z}}_t & \underline{\mathbf{z}}_t \end{pmatrix}'$:

$$
\begin{aligned}
\text{SVAR}(k) \quad &: \quad \mathbf{z}_t = \sum_{i=0}^{k} \mathbf{A}_i \mathbf{z}_{t-i} + \boldsymbol{\eta}_t, \\
\text{RF} \quad &: \quad \mathbf{z}_t = \sum_{i=1}^{k} \left(\mathbf{I}_{m+n} - \mathbf{A}_0\right)^{-1} \mathbf{A}_i \mathbf{z}_{t-i} + \mathbf{r}_t,
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{\eta}_t &= \mathbf{B}\boldsymbol{\varepsilon}_t, \\
E\left(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'\right) &= \boldsymbol{\Sigma}, \\
\mathbf{z}_t &= \begin{pmatrix} \overline{z}_t^{com} & \overline{z}_t^{gdp} & \overline{z}_t^{rs} & \overline{z}_t^{pl} & \overline{z}_t^{fcr} & \underline{z}_t^{cr} & \underline{z}_t^{mon} & \underline{z}_t^{exr} \end{pmatrix}', \\
\mathbf{r}_t &= \left(\mathbf{I}_{m+n} - \mathbf{A}_0\right)^{-1} \mathbf{B}\boldsymbol{\varepsilon}_t, \text{ or more precisely} \\
\mathbf{r}_t &= \mathbf{A}_0 \mathbf{r}_t + \mathbf{B}\boldsymbol{\varepsilon}_t.
\end{aligned}
$$

Setting restrictions on contemporaneous relationship $\mathbf{A}_0$ and on $\mathbf{B}$, we partly identify this SVAR(k) in order to write down the RF estimated by OLS. The model becomes

$$
\begin{aligned}
\begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} = & \begin{pmatrix} \mathbf{A}_0^{\overline{z}\overline{z}} & \boxed{\mathbf{0}} \\ \mathbf{A}_0^{\underline{z}\overline{z}} & \mathbf{A}_0^{\underline{z}\underline{z}} \end{pmatrix} \begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} + \sum_{i=1}^{k} \begin{pmatrix} \mathbf{A}_i^{\overline{z}\overline{z}} & \mathbf{A}_i^{\overline{z}\underline{z}} \\ \mathbf{A}_i^{\underline{z}\overline{z}} & \mathbf{A}_i^{\underline{z}\underline{z}} \end{pmatrix} \begin{pmatrix} \overline{\mathbf{z}}_{t-i} \\ \underline{\mathbf{z}}_{t-i} \end{pmatrix} \\
& + \begin{pmatrix} \mathbf{B}^{\overline{z}} & \boxed{\mathbf{0}} \\ \boxed{\mathbf{0}} & \mathbf{B}^{\underline{z}} \end{pmatrix} \begin{pmatrix} \varepsilon_t^{\overline{z}} \\ \varepsilon_t^{\underline{z}} \end{pmatrix},
\end{aligned}
$$

expressed in RF

$$
\begin{aligned}
\begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} = & \begin{pmatrix} \mathbf{I}_m - \mathbf{A}_0^{\overline{z}\overline{z}} & \mathbf{0} \\ -\mathbf{A}_0^{\underline{z}\overline{z}} & \mathbf{I}_n - \mathbf{A}_0^{\underline{z}\underline{z}} \end{pmatrix}^{-1} \sum_{i=1}^{k} \begin{pmatrix} \mathbf{A}_i^{\overline{z}\overline{z}} & \mathbf{A}_i^{\overline{z}\underline{z}} \\ \mathbf{A}_i^{\underline{z}\overline{z}} & \mathbf{A}_i^{\underline{z}\underline{z}} \end{pmatrix} \begin{pmatrix} \overline{\mathbf{z}}_{t-i} \\ \underline{\mathbf{z}}_{t-i} \end{pmatrix} \\
& + \begin{pmatrix} \mathbf{I}_m - \mathbf{A}_0^{\overline{z}\overline{z}} & \mathbf{0} \\ -\mathbf{A}_0^{\underline{z}\overline{z}} & \mathbf{I}_n - \mathbf{A}_0^{\underline{z}\underline{z}} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{B}^{\overline{z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{\underline{z}} \end{pmatrix} \begin{pmatrix} \varepsilon_t^{\overline{z}} \\ \varepsilon_t^{\underline{z}} \end{pmatrix},
\end{aligned}
$$

or rewritten using $\mathbf{A}_{0,-1}^{**} = (\mathbf{I} - \mathbf{A}_0^{**})^{-1}$

$$\begin{aligned}\begin{pmatrix} \overline{\mathbf{z}}_t \\ \underline{\mathbf{z}}_t \end{pmatrix} &= \sum_{i=1}^{k} \begin{pmatrix} \mathbf{\Pi}_i^{\overline{zz}} & \mathbf{\Pi}_i^{\overline{z}\underline{z}} \\ \mathbf{\Pi}_i^{\underline{z}\overline{z}} & \mathbf{\Pi}_i^{\underline{zz}} \end{pmatrix} \begin{pmatrix} \overline{\mathbf{z}}_{t-i} \\ \underline{\mathbf{z}}_{t-i} \end{pmatrix} \\ &\quad + \underbrace{\begin{pmatrix} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{B}^{\overline{z}} & \mathbf{0} \\ \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{B}^{\overline{z}} & \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} \end{pmatrix} \begin{pmatrix} \varepsilon_t^{\overline{z}} \\ \varepsilon_t^{\underline{z}} \end{pmatrix}}_{\begin{pmatrix} \mathbf{r}_t^{\overline{z}} & \mathbf{r}_t^{\underline{z}} \end{pmatrix}'},\end{aligned}$$

with

$$\mathbf{\Pi}_i = \begin{pmatrix} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{zz}} & \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{z}\underline{z}} \\ \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{zz}} + \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_i^{\underline{z}\overline{z}} & \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{A}_i^{\overline{z}\underline{z}} + \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_i^{\underline{zz}} \end{pmatrix}.$$

We store residuals $\begin{pmatrix} \mathbf{r}_t^{\overline{z}} & \mathbf{r}_t^{\underline{z}} \end{pmatrix}'$ for the second step.

Appendix B Second Step without Extraction SVAR(0)

We present here the SVAR(0) that links RF residuals $\mathbf{r}^{\underline{z}}$ with structural shocks $\boldsymbol{\varepsilon}^{\underline{z}}$ in the second step. Without extraction, this link is direct.

Model in Innovation Form

We assume $\mathbf{B}^{\overline{z}} = \mathbf{I}_m$ and $\mathbf{B}^{\underline{z}} = \mathbf{I}_n$ giving thus a new equation linking $\mathbf{r}^{\underline{z}}$ and $\boldsymbol{\varepsilon}^{\underline{z}}$, that we write following three different ways.

$$\begin{aligned}
\begin{pmatrix} \mathbf{r}_t^{\overline{z}} \\ \mathbf{r}_t^{\underline{z}} \end{pmatrix} &= \begin{pmatrix} \mathbf{I}_m - \mathbf{A}_0^{\overline{zz}} & \mathbf{0} \\ -\mathbf{A}_0^{\underline{z}\overline{z}} & \mathbf{I}_n - \mathbf{A}_0^{\underline{zz}} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\varepsilon}_t^{\overline{z}} \\ \boldsymbol{\varepsilon}_t^{\underline{z}} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{A}_{0,-1}^{\overline{zz}} & \mathbf{0} \\ \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} & \mathbf{A}_{0,-1}^{\underline{zz}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\varepsilon}_t^{\overline{z}} \\ \boldsymbol{\varepsilon}_t^{\underline{z}} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{A}_0^{\overline{zz}} & \mathbf{0} \\ \mathbf{A}_0^{\underline{z}\overline{z}} & \mathbf{A}_0^{\underline{zz}} \end{pmatrix} \begin{pmatrix} \mathbf{r}_t^{\overline{z}} \\ \mathbf{r}_t^{\underline{z}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_t^{\overline{z}} \\ \boldsymbol{\varepsilon}_t^{\underline{z}} \end{pmatrix}
\end{aligned}$$

Three equations are used as restrictions and estimated by overidentified GMM[30] to get bottom partitioned matrices $\mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}}$ and $\mathbf{A}_{0,-1}^{\underline{zz}}$ (12 coefficients):

$$\begin{aligned}
r_t^{cr} &= \theta_1 r_t^{com} + \theta_2 r_t^{mon} + \theta_3 r_t^{exr} + \varepsilon_t^s, \\
r_t^{mon} &= \theta_4 r_t^{gdp} + \theta_5 r_t^{cr} + \varepsilon_t^d, \\
r_t^{exr} &= \theta_6 r_t^{com} + \theta_7 r_t^{gdp} + \theta_8 r_t^{rs} + \theta_9 r_t^{pl} + \theta_{10} r_t^{fcr} + \theta_{11} r_t^{cr} + \theta_{12} r_t^{mon} + \varepsilon_t^x,
\end{aligned}$$

or in matrix notation where $\mathbf{A}_0^{\overline{zz}}$ is not identified $(\times)$, and $\boldsymbol{\varepsilon}_t^{\underline{z}} = \begin{pmatrix} \varepsilon_t^s & \varepsilon_t^d & \varepsilon_t^x \end{pmatrix}'$:

$$\begin{pmatrix} \mathbf{A}_0^{\overline{zz}} & \mathbf{0} \\ \mathbf{A}_0^{\underline{z}\overline{z}} & \mathbf{A}_0^{\underline{zz}} \end{pmatrix} = \left(\begin{array}{ccccc|ccc} \times & \dots & & & & 0 & \dots & \\ \vdots & \ddots & & & & \vdots & \ddots & \\ \hline \theta_1 & 0 & 0 & 0 & 0 & 0 & \theta_2 & \theta_3 \\ 0 & \theta_4 & 0 & 0 & 0 & \theta_5 & 0 & 0 \\ \theta_6 & \theta_7 & \theta_8 & \theta_9 & \theta_{10} & \theta_{11} & \theta_{12} & 0 \end{array} \right).$$

Dynamics

Dynamics implied by the policy sector shocks $\boldsymbol{\varepsilon}_t^{\underline{z}}$, besides RF IRF, is function of the above matrix. RF IRF are given by a VMA(∞) representation of $\mathbf{z}_t = \sum_{i=1}^k \boldsymbol{\Pi}_i \mathbf{z}_{t-i} + \mathbf{r}_t = \boldsymbol{\Psi}(L)\,\mathbf{r}_t = \mathbf{I}_{m+n}\mathbf{r}_t + \boldsymbol{\Psi}_1 \mathbf{r}_{t-1} + \boldsymbol{\Psi}_2 \mathbf{r}_{t-2} + \dots$, implying RF IRF $\frac{\partial \mathbf{z}_{t+l}}{\partial \mathbf{r}_t}$ that are equal to $\frac{\partial \mathbf{z}_t}{\partial \mathbf{r}_{t-l}} = \boldsymbol{\Psi}_l = \mathbf{K}\mathbf{M}^l\mathbf{K}'$, where $[8k \times 8k]$ $\mathbf{M}$ and $[8 \times 8k]$ $\mathbf{K}$ are respectively $\begin{pmatrix} \boldsymbol{\Pi}_1 & \boldsymbol{\Pi}_2 & \dots & \boldsymbol{\Pi}_{k-1} & \boldsymbol{\Pi}_k \\ \mathbf{I}_{m+n} & \mathbf{0} & & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_{m+n} & \mathbf{0} \end{pmatrix}$ and $\begin{pmatrix} \mathbf{I}_{m+n} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}$.

[30] IV for r^{cr} equation: r_t^{com}, r_t^{gdp}, r_t^{rs}, r_t^{pl}, and r_t^{fcr}; for r^{mon} equation: same IV and ε_t^s; for r^{exr} equation: same IV as for r^{mon} equation and ε_t^d.

Hence, SVAR dynamics $\frac{\partial \mathbf{z}_t}{\partial \boldsymbol{\varepsilon}_{t-l}}$ is equal to $\frac{\partial \mathbf{z}_t}{\partial \mathbf{r}_{t-l}} \frac{\partial \mathbf{r}_{t-l}}{\partial \boldsymbol{\varepsilon}_{t-l}} = \boldsymbol{\Psi}_l \left(\mathbf{I}_{m+n} - \mathbf{A}_0\right)^{-1}$ where only the last three columns are definite. For standard error bands around IRF, see Hamilton (1994).

Extra: Variance Decomposition

To perform a variance decomposition, we have to just-identify the upper left part $\mathbf{A}_0$ ($\mathbf{A}_0^{\overline{zz}}$ with 25 coefficients) and its corresponding $\boldsymbol{\varepsilon}$ variances (15 elements). In order to estimate these 40 coefficients, 25 restrictions are needed (15 $\mathbf{r}$ moments are known): i) variance-covariance matrix of $\boldsymbol{\varepsilon}$ shocks is diagonal (10 restrictions); ii) $\mathbf{A}_0^{\overline{zz}}$ has a Cholesky structure (15 restrictions).

$$\mathbf{A}_0 = \left(\begin{array}{ccccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_2 & \alpha_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_4 & \alpha_5 & \alpha_6 & 0 & 0 & 0 & 0 & 0 \\ \alpha_7 & \alpha_8 & \alpha_9 & \alpha_{10} & 0 & 0 & 0 & 0 \\ \hline \theta_1 & 0 & 0 & 0 & 0 & 0 & \theta_2 & \theta_3 \\ 0 & \theta_4 & 0 & 0 & 0 & \theta_5 & 0 & 0 \\ \theta_6 & \theta_7 & \theta_8 & \theta_9 & \theta_{10} & \theta_{11} & \theta_{12} & 0 \end{array}\right)$$

Cholesky decomposition of the second moments is given by the following equation:

$$V \begin{bmatrix} r_t^{com} \\ r_t^{gdp} \\ r_t^{rs} \\ r_t^{pl} \\ r_t^{fcr} \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\alpha_1 & 1 & 0 & 0 & 0 \\ -\alpha_2 & -\alpha_3 & 1 & 0 & 0 \\ -\alpha_4 & -\alpha_5 & -\alpha_6 & 1 & 0 \\ -\alpha_7 & -\alpha_8 & -\alpha_9 & -\alpha_{10} & 1 \end{pmatrix}^{-1} V \begin{bmatrix} \varepsilon_t^{com} \\ \varepsilon_t^{gdp} \\ \varepsilon_t^{rs} \\ \varepsilon_t^{pl} \\ \varepsilon_t^{fcr} \end{bmatrix}$$
$$\cdot \begin{pmatrix} 1 & -\alpha_1 & -\alpha_2 & -\alpha_4 & -\alpha_7 \\ 0 & 1 & -\alpha_3 & -\alpha_5 & -\alpha_8 \\ 0 & 0 & 1 & -\alpha_6 & -\alpha_9 \\ 0 & 0 & 0 & 1 & -\alpha_{10} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} .$$

The variance decomposition measures the contribution of each innovation to the mean-squared error (MSE) of the j-period-ahead forecast: $\mathbf{MSE}(\hat{\mathbf{z}}_{t+j|t})$

$$\begin{aligned}
&= V\left[\sum_{i=0}^{j-1} \boldsymbol{\Psi}_i \left(\mathbf{I}_{m+n} - \mathbf{A}_0\right)^{-1} \boldsymbol{\varepsilon}_{t+j-i}\right] = \sum_{i=0}^{j-1} \boldsymbol{\Psi}_i \mathbf{A}_{0,-1} \boldsymbol{\Sigma} \mathbf{A}_{\mathbf{0},-\mathbf{1}}' \boldsymbol{\Psi}_i', \\
&= V\left[\varepsilon_t^{com}\right] \left(\sum_{i=0}^{j-1} \boldsymbol{\Psi}_i \mathbf{A}_{0,-1} \begin{pmatrix} 1 & 0 & \cdots \\ 0 & 0 & \\ \vdots & & \ddots \end{pmatrix} \mathbf{A}_{0,-1}' \boldsymbol{\Psi}_i'\right) + \\
&\quad V\left[\varepsilon_t^{gdp}\right] \left(\sum_{i=0}^{j-1} \boldsymbol{\Psi}_i \mathbf{A}_{0,-1} \begin{pmatrix} 0 & 0 & \cdots \\ 0 & 1 & \\ \vdots & & \ddots \end{pmatrix} \mathbf{A}_{0,-1}' \boldsymbol{\Psi}_i'\right) + \ldots, \\
&= V\left[\varepsilon_t^{com}\right] \left(\sum_{i=0}^{j-1} \boldsymbol{\Psi}_i \mathbf{C}_1 \mathbf{C}_1' \boldsymbol{\Psi}_i'\right) + V\left[\varepsilon_t^{gdp}\right] \left(\sum_{i=0}^{j-1} \boldsymbol{\Psi}_i \mathbf{C}_2 \mathbf{C}_2' \boldsymbol{\Psi}_i'\right) + \ldots,
\end{aligned}$$

where $\mathbf{C}_1$ is the first column of the matrix $(\mathbf{I}_{m+n} - \mathbf{A}_0)^{-1}$, $\mathbf{C}_2$ the second column, etc. The percentage contribution can be calculated for each j-period-ahead forecast.

Appendix C Second Step with Extraction SVAR(0)

We show here the SVAR(0) that links RF residuals $\mathbf{r}^{\underline{z}}$ with structural shocks $\boldsymbol{\varepsilon}^{\underline{z}}$ in the second step with extraction.

Extraction, Model in Innovation Form, and Dynamics

The link between $\mathbf{r}^{\underline{z}}$ and $\boldsymbol{\varepsilon}^{\underline{z}}$ is defined with help of three new series $\mathbf{u}_t^{\underline{z}} = \begin{pmatrix} u_t^{cr} & u_t^{mon} & u_t^{exr} \end{pmatrix}'$ extracted from $\mathbf{r}^{\underline{z}}$. We extract this vector from the RF using OLS:

$$\begin{pmatrix} \mathbf{r}_t^{\overline{z}} \\ \mathbf{r}_t^{\underline{z}} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{B}^{\overline{z}} \boldsymbol{\varepsilon}_t^{\overline{z}} \\ \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{zz}} \mathbf{B}^{\overline{z}} \boldsymbol{\varepsilon}_t^{\overline{z}} + \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} \boldsymbol{\varepsilon}_t^{\underline{z}} \end{pmatrix},$$

where $\mathbf{r}_t^{\underline{z}} = \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{r}_t^{\overline{z}} + \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} \boldsymbol{\varepsilon}_t^{\underline{z}}$ and $\mathbf{u}_t^{\underline{z}} = \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} \boldsymbol{\varepsilon}_t^{\underline{z}}$. Theoretical dynamics of the structural model $\frac{\partial \mathbf{z}_t}{\partial \boldsymbol{\varepsilon}_{t-l}}$ is similar to the one without extraction:

$$\boldsymbol{\Psi}_l \begin{pmatrix} \times & \mathbf{0} \\ \times & \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} \end{pmatrix} = \begin{pmatrix} \times & \frac{\partial \overline{z}_{t+l}^{com}}{\partial \varepsilon_t^s} & \frac{\partial \overline{z}_{t+l}^{com}}{\partial \varepsilon_t^d} & \frac{\partial \overline{z}_{t+l}^{com}}{\partial \varepsilon_t^x} \\ \vdots & \vdots & \vdots & \vdots \\ \times & \frac{\partial \underline{z}_{t+l}^{exr}}{\partial \varepsilon_t^s} & \frac{\partial \underline{z}_{t+l}^{exr}}{\partial \varepsilon_t^d} & \frac{\partial \underline{z}_{t+l}^{exr}}{\partial \varepsilon_t^x} \end{pmatrix}.$$

Specification for $\mathbf{u}_t^{\underline{z}} = \mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} \boldsymbol{\varepsilon}_t^{\underline{z}}$

Based on operating procedures, we construct different models to restrict the relationship between $\mathbf{u}^{\underline{z}}$ and $\boldsymbol{\varepsilon}^{\underline{z}}$. Four equations link these two vectors. We do not write time subscripts.

$$\begin{aligned} u_s^{mon} &= u_d^{mon} \\ u_s^{mon} &= \lambda \varepsilon^d + \phi \varepsilon^x + \varepsilon^s \\ u_d^{mon} &= \rho u^{cr} + \varepsilon^d \\ u^{exr} &= \delta u^{cr} + \varepsilon^x \end{aligned}$$

We write them in matrix form linking $\mathbf{u}^{\underline{z}}$ and $\boldsymbol{\varepsilon}^{\underline{z}}$:

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{\rho} & \frac{(\lambda-1)}{\rho} & \frac{\phi}{\rho} \\ 1 & \lambda & \phi \\ \frac{\delta}{\rho} & \frac{\delta(\lambda-1)}{\rho} & \left(1 + \frac{\delta\phi}{\rho}\right) \end{pmatrix}}_{\mathbf{A}_{0,-1}^{\underline{zz}} \mathbf{B}^{\underline{z}} = \mathbf{H}(\boldsymbol{\theta}) \text{ and } \boldsymbol{\theta} = \lambda, \phi, \rho, \delta} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix}.$$

GMM Estimation

We estimate our system by overidentified GMM. Goal is to estimate the coefficient vector $\boldsymbol{\theta}$ and the diagonal matrix $V[\boldsymbol{\varepsilon}] = \boldsymbol{\Sigma}\ (V[\varepsilon^s], V\left[\varepsilon^d\right]$, and $V[\varepsilon^x]$

to estimate) with the variance-covariance matrix of $\mathbf{u} = \mathbf{H}(\boldsymbol{\theta})\boldsymbol{\varepsilon}$ that is known. This implies $V[\mathbf{u}] = \mathbf{H}(\boldsymbol{\theta})V[\boldsymbol{\varepsilon}]\,\mathbf{H}(\boldsymbol{\theta})'$, or more precisely $E(\mathbf{u}_t\mathbf{u}_t') = \mathbf{H}(\boldsymbol{\theta})\boldsymbol{\Sigma}\mathbf{H}(\boldsymbol{\theta})'$:

$$\begin{pmatrix} V[u^{cr}] & Cov[u^{cr}, u^{mon}] & \cdots \\ Cov[u^{mon}, u^{cr}] & V[u^{mon}] & \\ \vdots & & \ddots \end{pmatrix} = \underbrace{\begin{pmatrix} \omega^{11} & \omega^{12} & \cdots \\ \omega^{21} & \omega^{22} & \\ \vdots & & \ddots \end{pmatrix}}_{\boldsymbol{\Omega}\left(\boldsymbol{\theta}', V[\varepsilon^s], V\left[\varepsilon^d\right], V[\varepsilon^x]\right) = \boldsymbol{\Omega}(\boldsymbol{\theta}^{*\prime})}$$

where $\boldsymbol{\Omega} = \begin{pmatrix} h^{11}(\boldsymbol{\theta}) & \cdots \\ h^{21}(\boldsymbol{\theta}) & \\ \vdots & \ddots \end{pmatrix} \begin{pmatrix} V[\varepsilon^s] & 0 & \cdots \\ 0 & V\left[\varepsilon^d\right] & \\ \vdots & & \ddots \end{pmatrix} \begin{pmatrix} h^{11}(\boldsymbol{\theta}) & \cdots \\ h^{12}(\boldsymbol{\theta}) & \\ \vdots & \ddots \end{pmatrix}$.

Moment conditions are $E(\mathbf{J}(\boldsymbol{\theta}^*, \mathbf{u}_t)) = \mathbf{0}$, where $\mathbf{J}(\boldsymbol{\theta}^*, \mathbf{u}_t)$ equals $vech(\mathbf{u}_t\mathbf{u}_t') - vech(\mathbf{H}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\mathbf{H}')$ or $\boldsymbol{\gamma}_t - \boldsymbol{\xi}_t(\boldsymbol{\theta}^*)$ and corresponds to the following vector:

$$\begin{pmatrix} u_t^{cr}u_t^{cr} - \xi_t^{11}(\boldsymbol{\theta}^*) \\ u_t^{mon}u_t^{cr} - \xi_t^{21}(\boldsymbol{\theta}^*) \\ u_t^{exr}u_t^{cr} - \xi_t^{31}(\boldsymbol{\theta}^*) \\ u_t^{mon}u_t^{mon} - \xi_t^{22}(\boldsymbol{\theta}^*) \\ u_t^{exr}u_t^{mon} - \xi_t^{32}(\boldsymbol{\theta}^*) \\ u_t^{exr}u_t^{exr} - \xi_t^{33}(\boldsymbol{\theta}^*) \end{pmatrix}.$$

Sample counterpart of $E(\mathbf{J}(\boldsymbol{\theta}^*, \mathbf{u}_t))$ is $\mathbf{g}(\hat{\boldsymbol{\theta}}^*, \mathbf{u}_T) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{J}\left(\hat{\boldsymbol{\theta}}^*, \mathbf{u}_t\right) \neq \mathbf{0}$:

$$\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} u_t^{cr}u_t^{cr} - \xi_t^{11}\left(\hat{\boldsymbol{\theta}}^*\right) \\ u_t^{mon}u_t^{cr} - \xi_t^{21}\left(\hat{\boldsymbol{\theta}}^*\right) \\ u_t^{exr}u_t^{cr} - \xi_t^{31}\left(\hat{\boldsymbol{\theta}}^*\right) \\ u_t^{mon}u_t^{mon} - \xi_t^{22}\left(\hat{\boldsymbol{\theta}}^*\right) \\ u_t^{exr}u_t^{mon} - \xi_t^{32}\left(\hat{\boldsymbol{\theta}}^*\right) \\ u_t^{exr}u_t^{exr} - \xi_t^{33}\left(\hat{\boldsymbol{\theta}}^*\right) \end{pmatrix} = \begin{pmatrix} V[u^{cr}] - \omega^{11}\left(\hat{\boldsymbol{\theta}}^*\right) \\ Cov[u^{mon}, u^{cr}] - \omega^{21}\left(\hat{\boldsymbol{\theta}}^*\right) \\ Cov[u^{exr}, u^{cr}] - \omega^{31}\left(\hat{\boldsymbol{\theta}}^*\right) \\ V[u^{mon}] - \omega^{22}\left(\hat{\boldsymbol{\theta}}^*\right) \\ Cov[u^{exr}, u^{mon}] - \omega^{32}\left(\hat{\boldsymbol{\theta}}^*\right) \\ V[u^{exr}] - \omega^{33}\left(\hat{\boldsymbol{\theta}}^*\right) \end{pmatrix}.$$

Coefficients equal $\hat{\boldsymbol{\theta}}^*_{GMM} = \arg\min_{\mathbf{c}} \mathbf{g}(\mathbf{c}, \mathbf{u}_T)'\hat{\mathbf{S}}_T^{-1}\mathbf{g}(\mathbf{c}, \mathbf{u}_T)$ where $\hat{\mathbf{S}}_T^{-1}$ is a weighting matrix to select accordingly to the serial correlation of the sample equivalent of moment conditions. Overidentification test (OT) statistic is $\left(\sqrt{T}\mathbf{g}(\hat{\boldsymbol{\theta}}^*, \mathbf{u}_T)\right)'\hat{\mathbf{S}}_T^{-1}\left(\sqrt{T}\mathbf{g}(\hat{\boldsymbol{\theta}}^*, \mathbf{u}_T)\right)$ and is $\chi^2_{(\text{degree of over.})}$ (Hansen, 1982). OT statistic represents the J-statistic (value of the function at the optimum) times the sample size. OT tests whether sample moments represented by $\mathbf{g}(\hat{\boldsymbol{\theta}}^*, \mathbf{u}_T)$ are as close to zero as would be expected if the corresponding population moments $E(\mathbf{J}(\boldsymbol{\theta}^*, \mathbf{u}_t))$ were truly zero. We use it as a selection mechanism where the smallest statistic among different overidentification schemes is the better setup. However, a statistic bigger than a critical value would mean that the overidentification scheme is not as good as expected. We would reject H_0: $\mathbf{g}(\hat{\boldsymbol{\theta}}^*, \mathbf{u}_T) = \mathbf{0}$.

Moment Conditions Using Covariance Structure

Based on the matrix representation above, we take the variances of $\mathbf{u}$ and of our model to construct moment conditions.

$$
\begin{aligned}
& V\left[\begin{array}{ccc} u^{cr} & u^{mon} & u^{exr} \end{array}\right]' \\
= & \begin{pmatrix} \frac{1}{\rho} & \frac{(\lambda-1)}{\rho} & \frac{\phi}{\rho} \\ 1 & \lambda & \phi \\ \frac{\delta}{\rho} & \frac{\delta(\lambda-1)}{\rho} & \left(1+\frac{\delta\phi}{\rho}\right) \end{pmatrix} \mathbf{\Sigma} \begin{pmatrix} \frac{1}{\rho} & 1 & \frac{\delta}{\rho} \\ \frac{(\lambda-1)}{\rho} & \lambda & \frac{\delta(\lambda-1)}{\rho} \\ \frac{\phi}{\rho} & \phi & \left(1+\frac{\delta\phi}{\rho}\right) \end{pmatrix}
\end{aligned}
$$

$$
= \begin{pmatrix}
\begin{array}{c} \left(\frac{1}{\rho}\right)^2 V\left[\varepsilon^s\right] \\ +\left(\frac{\lambda-1}{\rho}\right)^2 V\left[\varepsilon^d\right] \\ +\left(\frac{\phi}{\rho}\right)^2 V\left[\varepsilon^x\right] \end{array} &
\begin{array}{c} \frac{1}{\rho} V\left[\varepsilon^s\right] \\ +\frac{\lambda(\lambda-1)}{\rho} V\left[\varepsilon^d\right] \\ +\frac{\phi^2}{\rho} V\left[\varepsilon^x\right] \end{array} &
\begin{array}{c} \frac{\delta}{\rho^2} V\left[\varepsilon^s\right] \\ +\frac{\delta(\lambda-1)^2}{\rho^2} V\left[\varepsilon^d\right] \\ +\frac{\phi(\rho+\delta\phi)}{\rho^2} V\left[\varepsilon^x\right] \end{array} \\
\begin{array}{c} \frac{1}{\rho} V\left[\varepsilon^s\right] \\ +\frac{\lambda(\lambda-1)}{\rho} V\left[\varepsilon^d\right] \\ +\frac{\phi^2}{\rho} V\left[\varepsilon^x\right] \end{array} &
\begin{array}{c} V\left[\varepsilon^s\right] \\ +\lambda^2 V\left[\varepsilon^d\right] \\ +\phi^2 V\left[\varepsilon^x\right] \end{array} &
\begin{array}{c} \frac{\delta}{\rho} V\left[\varepsilon^s\right] \\ +\frac{\lambda\delta(\lambda-1)}{\rho} V\left[\varepsilon^d\right] \\ +\phi\left(1+\frac{\delta\phi}{\rho}\right) V\left[\varepsilon^x\right] \end{array} \\
\begin{array}{c} \frac{\delta}{\rho^2} V\left[\varepsilon^s\right] \\ +\frac{\delta(\lambda-1)^2}{\rho^2} V\left[\varepsilon^d\right] \\ +\frac{\phi(\rho+\delta\phi)}{\rho^2} V\left[\varepsilon^x\right] \end{array} &
\begin{array}{c} \frac{\delta}{\rho} V\left[\varepsilon^s\right] \\ +\frac{\lambda\delta(\lambda-1)}{\rho} V\left[\varepsilon^d\right] \\ +\phi\left(1+\frac{\delta\phi}{\rho}\right) V\left[\varepsilon^x\right] \end{array} &
\begin{array}{c} \left(\frac{\delta}{\rho}\right)^2 V\left[\varepsilon^s\right] \\ +\left(\frac{\delta(\lambda-1)}{\rho}\right)^2 V\left[\varepsilon^d\right] \\ +\left(1+\frac{\delta\phi}{\rho}\right)^2 V\left[\varepsilon^x\right] \end{array}
\end{pmatrix}.
$$

Assumptions about Identification

6 different moments and 7 parameters imply 2 restrictions for the system to be overidentified and 1 restriction in case of just-identification. For each model, moment conditions correspond to $V\left[u^{cr}\right]$, $Cov\left[u^{mon}, u^{cr}\right]$, $Cov\left[u^{exr}, u^{cr}\right]$, $V\left[u^{mon}\right]$, $Cov\left[u^{exr}, u^{mon}\right]$, and $V\left[u^{exr}\right]$. The selected constrained forms are presented in the following table.

Setups with Extraction

#	Name	# Restrictions	Restrictions
1	Bank reserves targeting (BR)	2	$\lambda = \phi = 0$
2	Call rate targeting (CR)	2	$\lambda = 1, \phi = 0$
3	Exchange rate targeting (ER)	2	$\lambda = 1, \phi = \frac{-\rho}{\delta}$
4	$u^{exr} = \varepsilon^x$	1	$\delta = 0$
5	No reaction to ε^d	1	$\lambda = 0$

#1 $\lambda = \phi = 0$, $\varepsilon^s = u^{mon}$ meaning

$$
\begin{pmatrix}
\frac{1}{\rho^2}\left(V\left[\varepsilon^s\right] + V\left[\varepsilon^d\right]\right) & \ldots & \\
\frac{1}{\rho} V\left[\varepsilon^s\right] & V\left[\varepsilon^s\right] & \ldots \\
\frac{\delta}{\rho^2}\left(V\left[\varepsilon^s\right] + V\left[\varepsilon^d\right]\right) & \frac{\delta}{\rho} V\left[\varepsilon^s\right] & \frac{\delta^2}{\rho^2}\left(V\left[\varepsilon^s\right] + V\left[\varepsilon^d\right]\right) + V\left[\varepsilon^x\right]
\end{pmatrix},
$$

$$
\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & \frac{-1}{\rho} & 0 \\ 1 & 0 & 0 \\ \frac{\delta}{\rho} & \frac{-\delta}{\rho} & 1 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix}.
$$

#2 $\lambda = 1$ and $\phi = 0$, $\varepsilon^s = \rho u^{cr}$ meaning

$$
\begin{pmatrix}
\frac{1}{\rho^2} V\left[\varepsilon^s\right] & \ldots & \\
\frac{1}{\rho} V\left[\varepsilon^s\right] & V\left[\varepsilon^s\right] + V\left[\varepsilon^d\right] & \ldots \\
\frac{\delta}{\rho^2} V\left[\varepsilon^s\right] & \frac{\delta}{\rho} V\left[\varepsilon^s\right] & \frac{\delta^2}{\rho^2} V\left[\varepsilon^s\right] + V\left[\varepsilon^x\right]
\end{pmatrix},
$$

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & 0 & 0 \\ 1 & 1 & 0 \\ \frac{\delta}{\rho} & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix}.$$

<u>#3</u> $\lambda = 1$ and $\phi = \frac{-\rho}{\delta}$, $\varepsilon^s = \frac{\rho}{\delta} u^{exr}$ meaning

$$\begin{pmatrix} \frac{1}{\rho^2} V[\varepsilon^s] + \frac{1}{\delta^2} V[\varepsilon^x] & \dots & \\ \frac{1}{\rho} V[\varepsilon^s] + \frac{\rho}{\delta^2} V[\varepsilon^x] & V[\varepsilon^s] + V\left[\varepsilon^d\right] + \frac{\rho^2}{\delta^2} V[\varepsilon^x] & \dots \\ \frac{\delta}{\rho^2} V[\varepsilon^s] & \frac{\delta}{\rho} V[\varepsilon^s] & \left(\frac{\delta}{\rho}\right)^2 V[\varepsilon^s] \end{pmatrix},$$

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & 0 & \frac{-1}{\delta} \\ 1 & 1 & \frac{-\rho}{\delta} \\ \frac{\delta}{\rho} & 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix}.$$

<u>#4</u> $\delta = 0$, $\varepsilon^s = (1 - \lambda) u^{mon} + \lambda \rho u^{cr} - \phi u^{exr}$ meaning

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & \frac{(\lambda-1)}{\rho} & \frac{\phi}{\rho} \\ 1 & \lambda & \phi \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix}.$$

<u>#5</u> $\lambda = 0$, $\varepsilon^s = u^{mon} - \phi (u^{exr} - \delta u^{cr})$ meaning

$$\begin{pmatrix} u^{cr} \\ u^{mon} \\ u^{exr} \end{pmatrix} = \begin{pmatrix} \frac{1}{\rho} & \frac{-1}{\rho} & \frac{\phi}{\rho} \\ 1 & 0 & \phi \\ \frac{\delta}{\rho} & \frac{-\delta}{\rho} & \left(1 + \frac{\delta\phi}{\rho}\right) \end{pmatrix} \begin{pmatrix} \varepsilon^s \\ \varepsilon^d \\ \varepsilon^x \end{pmatrix}.$$

Appendix D Indicators

We show for each setup the construction of indicators. In the setup without extraction, the indicator for the overall stance of monetary policy is given according to the assumptions of the model. However, in the framework with extraction it is possible to compute a new set of indicators. We isolate the policy vector

$$\begin{aligned}\underline{\mathbf{z}}_t \quad = \quad & \sum_{i=1}^{k} \mathbf{\Pi}_i^{\underline{z}\overline{z}} \overline{\mathbf{z}}_{t-i} + \sum_{i=1}^{k} \mathbf{\Pi}_i^{\underline{z}\underline{z}} \underline{\mathbf{z}}_{t-i} + \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{z}\overline{z}} \mathbf{B}^{\overline{z}} \right) \boldsymbol{\varepsilon}_t^{\overline{z}} \\ & + \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right) \boldsymbol{\varepsilon}_t^{\underline{z}},\end{aligned}$$

and premultiply it by the inverse of the coefficient matrix of the structural shocks.

$$\begin{aligned}\left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right)^{-1} \underline{\mathbf{z}}_t \quad = \quad & \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right)^{-1} \sum_{i=1}^{k} \mathbf{\Pi}_i^{\underline{z}\overline{z}} \overline{\mathbf{z}}_{t-i} + \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right)^{-1} \sum_{i=1}^{k} \mathbf{\Pi}_i^{\underline{z}\underline{z}} \underline{\mathbf{z}}_{t-i} \\ & + \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right)^{-1} \left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{A}_0^{\underline{z}\overline{z}} \mathbf{A}_{0,-1}^{\overline{z}\overline{z}} \mathbf{B}^{\overline{z}} \right) \boldsymbol{\varepsilon}_t^{\overline{z}} + \boldsymbol{\varepsilon}_t^{\underline{z}}\end{aligned}$$

It gives then three different indicators $\left(\mathbf{A}_{0,-1}^{\underline{z}\underline{z}} \mathbf{B}^{\underline{z}} \right)^{-1} \underline{\mathbf{z}}_t$. We select the one that corresponds to the equation with the element ε_t^s. In our setup this is the first row.

Appendix E Dynamics Comparison

We compare dynamics without and with extraction. For both types of model, IRF display the form: $\boldsymbol{\Psi}_l \begin{pmatrix} \times & \mathbf{0} \\ \circ \text{ or } \times & \mathbf{A}^{zz}_{0,-1}\mathbf{B}^{\underline{z}} \end{pmatrix}$. $\times$ means 'not identified' and $\circ$ means 'restricted'. We focus on differences between our models.

#0 Case without extraction[31]

$$\begin{pmatrix} \times & \mathbf{0} \\ \circ & \frac{1}{\tau}\begin{pmatrix} 1 & (\theta_3\theta_{12}+\theta_2) & \theta_3 \\ \theta_5 & (1-\theta_3\theta_{11}) & \theta_3\theta_5 \\ (\theta_5\theta_{12}+\theta_{11}) & (\theta_2\theta_{11}+\theta_{12}) & (1-\theta_2\theta_5) \end{pmatrix}^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} \times & \mathbf{0} \\ \circ & \begin{pmatrix} 1 & -\theta_2 & -\theta_3 \\ -\theta_5 & 1 & 0 \\ -\theta_{11} & -\theta_{12} & 1 \end{pmatrix} \end{pmatrix}.$$

#1 Case with extraction and bank reserves targeting

$$\begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} 0 & 1 & 0 \\ -\rho & 1 & 0 \\ -\delta & 0 & 1 \end{pmatrix}^{-1} \end{pmatrix} = \begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \frac{1}{\rho} & \frac{-1}{\rho} & 0 \\ 1 & 0 & 0 \\ \frac{\delta}{\rho} & \frac{-\delta}{\rho} & 1 \end{pmatrix} \end{pmatrix}.$$

#2 Case with extraction and call rate targeting

$$\begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \rho & 0 & 0 \\ -\rho & 1 & 0 \\ -\delta & 0 & 1 \end{pmatrix}^{-1} \end{pmatrix} = \begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \frac{1}{\rho} & 0 & 0 \\ 1 & 1 & 0 \\ \frac{\delta}{\rho} & 0 & 1 \end{pmatrix} \end{pmatrix}.$$

#3 Case with extraction and exchange rate targeting

$$\begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} 0 & 0 & \frac{\rho}{\delta} \\ -\rho & 1 & 0 \\ -\delta & 0 & 1 \end{pmatrix}^{-1} \end{pmatrix} = \begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \frac{1}{\rho} & 0 & \frac{-1}{\delta} \\ 1 & 1 & \frac{-\rho}{\delta} \\ \frac{\delta}{\rho} & 0 & 0 \end{pmatrix} \end{pmatrix}.$$

#4 Case with extraction and $\delta = 0$

$$\begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \rho\lambda & (1-\lambda) & -\phi \\ -\rho & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \end{pmatrix} = \begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \frac{1}{\rho} & \frac{(\lambda-1)}{\rho} & \frac{\phi}{\rho} \\ 1 & \lambda & \phi \\ 0 & 0 & 1 \end{pmatrix} \end{pmatrix}.$$

[31] $\tau = (1-\theta_3\theta_5\theta_{12}-\theta_3\theta_{11}-\theta_2\theta_5)^{-1}$.

#5 Case with extraction and $\lambda = 0$

$$\begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \phi\delta & 1 & -\phi \\ -\rho & 1 & 0 \\ -\delta & 0 & 1 \end{pmatrix}^{-1} \end{pmatrix} = \begin{pmatrix} \times & \mathbf{0} \\ \times & \begin{pmatrix} \frac{1}{\rho} & \frac{-1}{\rho} & \frac{\phi}{\rho} \\ 1 & 0 & \phi \\ \frac{\delta}{\rho} & \frac{-\delta}{\rho} & \left(1+\frac{\delta\phi}{\rho}\right) \end{pmatrix} \end{pmatrix}.$$

Summary (Weights for policy variables)

#0

$$\tau \begin{pmatrix} 1 & (\theta_3\theta_{12}+\theta_2) & \theta_3 \\ \theta_5 & (1-\theta_3\theta_{11}) & \theta_3\theta_5 \\ (\theta_5\theta_{12}+\theta_{11}) & (\theta_2\theta_{11}+\theta_{12}) & (1-\theta_2\theta_5) \end{pmatrix}$$

#1

$$\begin{pmatrix} 0 & 1 & 0 \\ -\rho & 1 & 0 \\ -\delta & 0 & 1 \end{pmatrix}$$

#2

$$\begin{pmatrix} \rho & 0 & 0 \\ -\rho & 1 & 0 \\ -\delta & 0 & 0 \end{pmatrix}$$

#3

$$\begin{pmatrix} 0 & 0 & \frac{\rho}{\delta} \\ -\rho & 1 & 0 \\ -\delta & 0 & 1 \end{pmatrix}$$

#4

$$\begin{pmatrix} \rho\lambda & (1-\lambda) & -\phi \\ -\rho & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

#5

$$\begin{pmatrix} \phi\delta & 1 & -\phi \\ -\rho & 1 & 0 \\ -\delta & 0 & 1 \end{pmatrix}$$

Appendix F GMM-versus-2SLS (without Extraction)

Overidentified IV coefficients can be estimated by GMM or 2SLS. We show that they produce the same results when there is homoscedasticity of the error terms. If there is heteroscedasticity, GMM offers a better estimator. In the model $r_t = \mathbf{r^*}'_t\boldsymbol{\beta} + \varepsilon_t$ for $t = 1 \ldots n$, we assume a regressors $\mathbf{r}^*$ and q instruments $\mathbf{z}$ $(q > a)$. q moment conditions are then given by $E(\mathbf{z}_t\varepsilon_t) = \mathbf{0}$, producing a $[a \times 1]$ GMM estimator $\hat{\boldsymbol{\beta}}_{GMM} = \left(\mathbf{R}^{*\prime}\mathbf{ZWZ}'\mathbf{R}^*\right)^{-1}\left(\mathbf{R}^{*\prime}\mathbf{ZWZ}'\mathbf{r}\right)$ where $\mathbf{W} = \left(V\left[\mathbf{z}_t\varepsilon_t\right]\right)^{-1}$.

Homoscedasticity

With homoscedasticity, GMM estimator equals 2SLS estimator. Homoscedasticity implies $V[\mathbf{z}_t\varepsilon_t] = V[\varepsilon_t]\,V[\mathbf{z}_t] = \frac{1}{T}\sum\varepsilon_t^2 \cdot \frac{1}{T}\sum\mathbf{z}_t\mathbf{z}_t'$ or $V[\mathbf{z}_t\varepsilon_t] = \frac{s^2}{T}\mathbf{Z}'\mathbf{Z}$. This implies

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_{GMM} &= \left(\mathbf{R}^{*\prime}\mathbf{Z}\left(V\left[\mathbf{z}_t\varepsilon_t\right]\right)^{-1}\mathbf{Z}'\mathbf{R}^*\right)^{-1}\left(\mathbf{R}^{*\prime}\mathbf{Z}\left(V\left[\mathbf{z}_t\varepsilon_t\right]\right)^{-1}\mathbf{Z}'\mathbf{r}\right), \\
&= \left(\mathbf{R}^{*\prime}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{R}^*\right)^{-1}\left(\mathbf{R}^{*\prime}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{r}\right) = \hat{\boldsymbol{\beta}}_{2SLS}.
\end{aligned}
$$

Proof

For the first stage of 2SLS, we regress $\mathbf{r}^*{}_t$ on $\mathbf{z}_t$ producing a coefficient vectors of size q ($\hat{\boldsymbol{\beta}}new_1 = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{r}^{*1}, \ldots, \hat{\boldsymbol{\beta}}new_a = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{r}^{*a}$). In matrix notation

$$
\underset{[a\times q]}{\hat{\boldsymbol{\beta}}new_{tot}{}'} = \begin{pmatrix} \hat{\boldsymbol{\beta}}new_1' \\ \vdots \\ \hat{\boldsymbol{\beta}}new_a' \end{pmatrix} \text{ or } \hat{\boldsymbol{\beta}}new_{tot} = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{R}^*.
$$

This gives fitted values $\hat{\mathbf{R}}^*$:

$$
\underset{[n\times a]}{\hat{\mathbf{R}}^*} = \mathbf{Z}\hat{\boldsymbol{\beta}}new_{tot}.
$$

For the second stage of 2SLS, we regress r_t on $\hat{\mathbf{r}}_{\mathbf{t}}^*$.

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_{2SLS} &= (\hat{\mathbf{R}}^{*\prime}\hat{\mathbf{R}}^*)^{-1}\hat{\mathbf{R}}^{*\prime}\mathbf{r} \\
&= (\hat{\boldsymbol{\beta}}new_{tot}'\mathbf{Z}'\mathbf{Z}\hat{\boldsymbol{\beta}}new_{tot})^{-1}\hat{\boldsymbol{\beta}}new_{tot}'\mathbf{Z}'\mathbf{r} \\
&= \left(\mathbf{R}^{*\prime}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{R}^*\right)^{-1}\left(\mathbf{R}^{*\prime}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{r}\right) = \hat{\boldsymbol{\beta}}_{GMM}
\end{aligned}
$$

Extra: Just-identification

In case of just-identification $\hat{\boldsymbol{\beta}}_{2SLS} = \hat{\boldsymbol{\beta}}_{IV}$.

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_{2SLS} &= \left(\mathbf{R}^{*\prime}\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{R}^*\right)^{-1}\left(\mathbf{R}^{*\prime}\mathbf{Z}\cdot\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\cdot\mathbf{Z}'\mathbf{r}\right) \\
&= \left(\mathbf{Z}'\mathbf{R}^*\right)^{-1}\left(\mathbf{Z}'\mathbf{Z}\right)\left(\mathbf{R}^{*\prime}\mathbf{Z}\right)^{-1}\left(\mathbf{R}^{*\prime}\mathbf{Z}\right)\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\left(\mathbf{Z}'\mathbf{r}\right) \\
&= \left(\mathbf{Z}'\mathbf{R}^*\right)^{-1}\left(\mathbf{Z}'\mathbf{r}\right) = \hat{\boldsymbol{\beta}}_{IV}
\end{aligned}
$$

Appendix G Generated Instruments (without Extraction)

We show that the methodology used to estimate our three equations in the setup without extraction could produce misleading results (e.g. King and Watson (1997)). To estimate the second and third equation, we use residuals of the first equation as instruments. For example, we use $\hat{\varepsilon}^1$ as an instrument in the estimation of a second equation explaining r^2. The problem is that these residuals $\hat{\varepsilon}^1_t$ are first estimated and second dependent on the structural form of the first equation explaining r^1. In this appendix, we only consider the consequences of using $\hat{\varepsilon}^1$ instead of ε^1 as an instrument[32]. We represent this problem in a two-equation setup:

$$\begin{aligned} r^1_t &= \mathbf{r}^{*1\prime}_t \boldsymbol{\theta}_1 + \varepsilon^1_t, \\ r^2_t &= \mathbf{r}^{*2\prime}_t \boldsymbol{\theta}_2 + \varepsilon^2_t. \end{aligned}$$

The first equation is estimated with IV ($\mathbf{z}_t$ or in matrix form $\mathbf{Z}$). The second equation is estimated with the same instruments plus the estimated residuals of the first equation. The standard error (SE) of $\hat{\boldsymbol{\theta}}_2$ would be 'standard' with instruments $\mathbf{U} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 & \mathbf{Z} \end{pmatrix}$, but we use the following $\hat{\mathbf{U}} = \begin{pmatrix} \hat{\boldsymbol{\varepsilon}}_1 & \mathbf{Z} \end{pmatrix}$. Hence, we recalculate the SE of $\hat{\boldsymbol{\theta}}_2$. After having stacked the observations

$$\mathbf{r}_1 = \mathbf{R}^*{}_1 \boldsymbol{\theta}_1 + \boldsymbol{\varepsilon}_1 \qquad \text{and} \qquad \mathbf{r}_2 = \mathbf{R}^*{}_2 \boldsymbol{\theta}_2 + \boldsymbol{\varepsilon}_2,$$

we rewrite

$$\begin{aligned} \hat{\boldsymbol{\varepsilon}}_1 &= \mathbf{r}_1 - \mathbf{R}^*{}_1 \hat{\boldsymbol{\theta}}_1, \\ &= \boldsymbol{\varepsilon}_1 - \mathbf{R}^*{}_1 \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1 \right), \end{aligned}$$

and[33]

$$\begin{aligned} \hat{\mathbf{U}} &= \begin{pmatrix} \hat{\boldsymbol{\varepsilon}}_1 & \mathbf{Z} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\varepsilon}}_1 & \mathbf{Z} \end{pmatrix} - \mathbf{U} + \mathbf{U}, \\ &= \mathbf{U} - \begin{pmatrix} \mathbf{R}^*{}_1 \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1 \right) & \mathbf{0} \end{pmatrix}, \end{aligned}$$

in order to construct the asymptotic distribution of $\hat{\boldsymbol{\theta}}_2 = \left(\hat{\mathbf{U}}' \mathbf{R}^*{}_2 \right)^{-1} \left(\hat{\mathbf{U}}' \mathbf{r}_2 \right)$.

$$\begin{aligned} \sqrt{T} \left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2 \right) &= \left(T^{-1} \hat{\mathbf{U}}' \mathbf{R}^*{}_2 \right)^{-1} \left(T^{-\frac{1}{2}} \hat{\mathbf{U}}' \boldsymbol{\varepsilon}_2 \right) \\ &= \left(T^{-1} \hat{\mathbf{U}}' \mathbf{R}^*{}_2 \right)^{-1} \left(T^{-\frac{1}{2}} \mathbf{U}' \boldsymbol{\varepsilon}_2 \right) \\ &\quad - \left(T^{-1} \hat{\mathbf{U}}' \mathbf{R}^*{}_2 \right)^{-1} \begin{pmatrix} T^{\frac{1}{2}} \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1 \right)' \left(T^{-1} \mathbf{R}^{*\prime}{}_1 \boldsymbol{\varepsilon}_2 \right) \\ \mathbf{0} \end{pmatrix} \end{aligned}$$

[32] We get a different series of residuals for each structural specification of the first equation. This structural problem is not treated here.

[33] $V \left[\hat{\boldsymbol{\theta}}_1 \right]$ is the 'standard' asymptotic IV variance and equals $\sigma^2_{\varepsilon_1} \left(\mathbf{Z}' \mathbf{R}^*{}_1 \right)^{-1} \left(\mathbf{Z}' \mathbf{Z} \right) \left(\mathbf{R}^{*\prime}{}_1 \mathbf{Z} \right)^{-1}$.

Proof of asymptotic distribution

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2\right) = \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right)$$

$$\begin{aligned}
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) \\
& + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} T^{-\frac{1}{2}}\left(\hat{\mathbf{U}}' - \hat{\mathbf{U}}'\right)\boldsymbol{\varepsilon}_2 \\
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) \\
& + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} T^{-\frac{1}{2}} \cdot \\
& \left(\begin{pmatrix} \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1 \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1 \\ \mathbf{0} \end{pmatrix}\right)\boldsymbol{\varepsilon}_2 \\
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) \\
& + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} T^{-\frac{1}{2}} \cdot \\
& \begin{pmatrix} \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1 - \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1 \\ \mathbf{0} \end{pmatrix}\boldsymbol{\varepsilon}_2 \\
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) \\
& + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} T^{-\frac{1}{2}} \cdot \\
& \begin{pmatrix} \left(\boldsymbol{\varepsilon}_1 - \hat{\boldsymbol{\varepsilon}}_1\right)' - \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1 \\ \mathbf{0} \end{pmatrix}\boldsymbol{\varepsilon}_2 \\
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) \\
& + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} T^{-\frac{1}{2}} \cdot \\
& \left(\begin{pmatrix} \boldsymbol{\varepsilon}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \hat{\boldsymbol{\varepsilon}}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{0} \end{pmatrix}\right) \\
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) \\
& + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} T^{-\frac{1}{2}} \cdot \\
& \left(\begin{pmatrix} \boldsymbol{\varepsilon}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{Z}'\boldsymbol{\varepsilon}_2 \end{pmatrix} - \begin{pmatrix} \hat{\boldsymbol{\varepsilon}}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{Z}'\boldsymbol{\varepsilon}_2 \end{pmatrix} - \begin{pmatrix} \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{0} \end{pmatrix}\right) \\
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} \cdot \\
& \left(T^{-\frac{1}{2}}\begin{pmatrix} \boldsymbol{\varepsilon}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{Z}'\boldsymbol{\varepsilon}_2 \end{pmatrix} - T^{-\frac{1}{2}}\begin{pmatrix} \hat{\boldsymbol{\varepsilon}}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{Z}'\boldsymbol{\varepsilon}_2 \end{pmatrix} - \begin{pmatrix} T^{-\frac{1}{2}}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\mathbf{R}^*{}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{0} \end{pmatrix}\right) \\
= \; & \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1} \cdot \\
& \left(T^{-\frac{1}{2}}\mathbf{U}'\boldsymbol{\varepsilon}_2 - T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2 - \begin{pmatrix} T^{\frac{1}{2}}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'T^{-1}\mathbf{R}^*{}'_1\boldsymbol{\varepsilon}_2 \\ \mathbf{0} \end{pmatrix}\right)
\end{aligned}$$

$$\begin{aligned}
&= \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) + \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\mathbf{U}'\boldsymbol{\varepsilon}_2\right) \\
&\quad - \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\hat{\mathbf{U}}'\boldsymbol{\varepsilon}_2\right) \\
&\quad - \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\begin{pmatrix} T^{\frac{1}{2}}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\left(T^{-1}\mathbf{R}^{*\prime}{}_1\boldsymbol{\varepsilon}_2\right) \\ \mathbf{0} \end{pmatrix} \\
&= \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\mathbf{U}'\boldsymbol{\varepsilon}_2\right) \\
&\quad - \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\begin{pmatrix} T^{\frac{1}{2}}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right)'\left(T^{-1}\mathbf{R}^{*\prime}{}_1\boldsymbol{\varepsilon}_2\right) \\ \mathbf{0} \end{pmatrix}
\end{aligned}$$

When the second term equals zero, we can write

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2\right) = \left(T^{-1}\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\mathbf{U}'\boldsymbol{\varepsilon}_2\right)$$

and this expression converges in distribution to a random variable

$$\left(T^{-1}\mathbf{U}'\mathbf{R}^*{}_2\right)^{-1}\left(T^{-\frac{1}{2}}\mathbf{U}'\boldsymbol{\varepsilon}_2\right)$$

that follows an asymptotic distribution which corresponds to the 'standard' IV estimator distribution:

$$N\left(\mathbf{0}, \sigma^2_{\varepsilon_2} \cdot \text{plim}\ \left(T\left(\hat{\mathbf{U}}'\mathbf{R}^*{}_2\right)^{-1}\left(\hat{\mathbf{U}}'\hat{\mathbf{U}}\right)\left(\mathbf{R}^{*\prime}_2\hat{\mathbf{U}}\right)^{-1}\right)\right).$$

For this second element to be zero, plim $T^{-1}\mathbf{R}^{*\prime}{}_1\boldsymbol{\varepsilon}_2 = \mathbf{0}$ is needed. It means that $\frac{1}{T}\sum \mathbf{r}^{*1}_t\varepsilon^2_t = E\left(\mathbf{r}^{*1}_t\varepsilon^2_t\right) = Cov\left[\mathbf{r}^{*1}_t, \varepsilon^2_t\right] = 0, \forall t$. This assumption is violated in the setup without extraction.

Appendix H Generated Regressors (without and with Extraction)

We show that the use of regressors generated by VAR residuals does not bias the coefficient inference. We assume a bivariate SVAR(0) representing u^1 and u^2 where $\begin{pmatrix} \hat{u}_t^1 & \hat{u}_t^2 \end{pmatrix}'$ are generated residuals from a previous regression. To keep things simple, we only focus on a coefficient μ to see the consequences of using estimated $\hat{\mathbf{u}}$ instead of true $\mathbf{u}$ in the model $\mathbf{u}_t = \mathbf{A}_{0,-1}^{zz} \mathbf{B}^{z} \boldsymbol{\varepsilon}_t$, that we express as $\begin{pmatrix} u_t^1 \\ u_t^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mu & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}$ or as $u_t^1 = \varepsilon_t^1$ and $u_t^2 = \mu \varepsilon_t^1 + \varepsilon_t^2$.

We generate $\hat{u}_t^1$ and $\hat{u}_t^2$ from a first regression explaining $\mathbf{z} = \begin{pmatrix} z^1 & z^2 \end{pmatrix}'$:

$$\begin{aligned} z_t^1 &= \mathbf{x}_t' \boldsymbol{\beta} + u_t^1 \quad \text{implying} \quad \hat{u}_t^1 = u_t^1 - \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right), \\ z_t^2 &= \mathbf{x}_t' \boldsymbol{\gamma} + u_t^2 \quad \text{implying} \quad \hat{u}_t^2 = u_t^2 - \mathbf{x}_t' \left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \right). \end{aligned}$$

We rewrite $\hat{u}_t^2$:

$$\begin{aligned} \hat{u}_t^2 &= \mu \hat{u}_t^1 + \varepsilon_t^2 + \left(\hat{u}_t^2 - u_t^2 \right) - \mu \left(\hat{u}_t^1 - u_t^1 \right), \\ &= \mu \hat{u}_t^1 + \varepsilon_t^2 + \left(-\mathbf{x}_t' \left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \right) \right) - \mu \left(-\mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \right), \\ &= \mu \hat{u}_t^1 + a_t, \end{aligned}$$

and estimate $\hat{\mu} = \frac{\sum \hat{u}_t^1 \hat{u}_t^2}{\sum \left(\hat{u}_t^1 \right)^2} = \mu + \frac{\sum \hat{u}_t^1 a_t}{\sum \left(\hat{u}_t^1 \right)^2}$. We analyze its asymptotic (*asy.*) distribution:

$$\sqrt{T} \left(\hat{\mu} - \mu \right) = \frac{\frac{1}{\sqrt{T}} \sum \hat{u}_t^1 a_t}{\frac{1}{T} \sum \left(\hat{u}_t^1 \right)^2}.$$

We develop the denominator of $\sqrt{T} \left(\hat{\mu} - \mu \right)$.

$$\begin{aligned} \frac{1}{T} \sum \left(\hat{u}_t^1 \right)^2 &= \frac{1}{T} \sum \left(u_t^1 - \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \right)^2 \\ &= \frac{1}{T} \sum \begin{pmatrix} {u_t^1}^2 + \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right)' \mathbf{x}_t \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \\ -u_t^1 \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right)' \mathbf{x}_t u_t^1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{T} \sum {u_t^1}^2 + \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right)' \left(\frac{1}{T} \sum \mathbf{x}_t \mathbf{x}_t' \right) \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \\ -\frac{1}{T} \sum u_t^1 \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right)' \left(\frac{1}{T} \sum \mathbf{x}_t u_t^1 \right) \end{pmatrix} \\ &\Downarrow \; asy. \\ & \boxed{\sigma_{u^1}^2} \end{aligned}$$

We then calculate $\hat{u}_t^1 a_t$ and develop the numerator of $\sqrt{T} \left(\hat{\mu} - \mu \right)$.

$$\begin{aligned} \hat{u}_t^1 a_t &= \left(u_t^1 - \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \right) \left(\varepsilon_t^2 - \mathbf{x}_t' \left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \right) + \mu \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \right) \\ &= u_t^1 \varepsilon_t^2 - u_t^1 \mathbf{x}_t' \left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \right) + \mu u_t^1 \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \\ &\quad - \varepsilon_t^2 \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) + \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \mathbf{x}_t \mathbf{x}_t' \left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \right) \\ &\quad - \mu \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \mathbf{x}_t \mathbf{x}_t' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \end{aligned}$$

$$\begin{aligned}
\frac{1}{\sqrt{T}}\sum \hat{u}_t^1 a_t &= \frac{1}{\sqrt{T}}\sum u_t^1\varepsilon_t^2 - \frac{1}{T}\sum u_t^1\mathbf{x}_t'\sqrt{T}\left(\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}\right) \\
&+\mu\frac{1}{T}\sum u_t^1\mathbf{x}_t'\sqrt{T}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) - \frac{1}{T}\sum \varepsilon_t^2\mathbf{x}_t'\sqrt{T}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) \\
&+\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\frac{1}{T}\sum \mathbf{x}_t\mathbf{x}_t'\right)\sqrt{T}\left(\hat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}\right) \\
&-\mu\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\frac{1}{T}\sum \mathbf{x}_t\mathbf{x}_t'\right)\sqrt{T}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) \\
&\Downarrow \quad asy.
\end{aligned}$$

$$\boxed{N(0,\sigma_{u^1}^2\sigma_{\varepsilon^2}^2) - \boldsymbol{\Sigma}_{u^1x}N(0,\sigma_{\hat{\gamma}}^2) + \mu\boldsymbol{\Sigma}_{u^1x}N(0,\sigma_{\hat{\beta}}^2) - \boldsymbol{\Sigma}_{\varepsilon^2x}N(0,\sigma_{\hat{\beta}}^2)}$$

In the VAR(p) generating $\hat{\mathbf{u}}$, $\mathbf{x}_t' = (z_{t-1}^1, z_{t-2}^1, \ldots, z_{t-p}^1, z_{t-1}^2, z_{t-2}^2, \ldots, z_{t-p}^2)$ implies $\boldsymbol{\Sigma}_{\varepsilon^2x} = \boldsymbol{\Sigma}_{u^1x} = \mathbf{0}$.

$$\begin{aligned}
\sqrt{T}\left(\hat{\mu}-\mu\right) &= \frac{\frac{1}{\sqrt{T}}\sum \hat{u}_t^1 a_t}{\frac{1}{T}\sum\left(\hat{u}_t^1\right)^2} \\
&\Downarrow \quad asy. \\
&\boxed{\frac{N(0,\sigma_{u^1}^2\sigma_{\varepsilon^2}^2)}{\sigma_{u^1}^2}}
\end{aligned}$$

This distribution is 'standard' and converges towards $N(0,\sigma_{\varepsilon^2}^2\boldsymbol{\Sigma}_{u^1u^1}^{-1})$ or $N(0,\sigma_{\varepsilon^2}^2\left(\mathbf{u}^{1\prime}\mathbf{u}^1\right)^{-1})$. Note that the asymptotic $\hat{\mu}$ distribution is function of true u^1 and not of generated $\hat{u}^1$ that we use to estimate μ.