Investigating weak identification in estimates of the Taylor rule for the Norwegian economy

Anders Merckoll Helseth

Master's Thesis, Department of Economics

University of Oslo

May 2015

Summary

This thesis investigates weak identification when a forward looking Taylor rule is estimated with GMM. Lagged values of the nominal interest rate, inflation and the output gap are used as instruments. In the New Keynesian sticky price model, the strength of identification is determined by the dynamic structure. Without any backward looking terms in the inflation and output gap equations, serial correlation in inflation and the output gap shocks is necessary for identification. The order of the AR process in the shock terms determines the number of lags that are relevant instruments. Simulation of a DSGE model shows that higher persistence in the shocks increases the strength of identification. Increasing the degree of interest rate smoothing weakens identification. Monetary policy that leads to an indeterminate equilibrium increases identification, compared to a determinate equilibrium.

Identification robust inference methods show that the forward looking Taylor rule is weakly identified for Norwegian data. Even though GMM estimates produce tight confidence intervals, we observe large areas of disagreement between identification robust confidence sets and confidence sets based on standard GMM inference. Possible explanations are suggested. The high degree of policy intertia observed for monetary policy will likely lead to weaker identification. Inflation expectations seem to be well anchored. This reduces the explained variation in the endogenous variables from the instruments, and hence weakens identification. Inconsistency, or discretion, in Norges Bank reaction pattern will also make identification more difficult.

Preface

I would like to thank my supervisor Tord Krogh for all his work in helping me write this thesis. His academic and dispositional input has greatly improved the quality of my work. Most of all, I appreciate his willingnes to help and make time in a busy schedule. I truly believe I couldn't have been more lucky in the assignment of supervisor.

I would like to thank the Department of Economics for providing me with a scholarship and office space for the past year, and Steinar Holden for the "student scholarship in macro and monetary policy issues".

Finally, I would like to thank my fellow students for creating a fun atmosphere at Blindern, both academically and socially.

Contents

List of Figures

List of Tables

1 Introduction

In the last 15 years or so, the generalized method of moments (GMM) has become a popular tool for estimating forward looking models with rational expectations. In the linear IV regression, problems can arise when instruments are weak, i.e., they are weakly correlated with the endogenous variables they serve as instruments for. The conceptually same problem can arise in GMM estimation. In the context of GMM, weak instruments correspond to weak identification of some or all of the unknown parameters. Weak identification leads to non-normal distributions, leaving coventional GMM inference invalid. For the linear IV regression many methods for handling weak instruments are proposed and widely applied. However, as pointed out in Stock, Wright and Yogo (2002), weak instruments is a much more difficult problem in general nonlinear GMM than in linear IV regression. In the context of GMM, the usual rank condition for identification is not sufficient to guarantee reliable inference using GMM in finite samples. Some methods are however available. Papers such as Mavroeidis (2010) and Krogh (2015) explore weak identifcation through obtaining confidence sets that are fully robust to weak identification. The identification robust confidence sets are then compared with conventional GMM inference.

A large number of papers argue that there might be problems with weak identification when estimating forward looking models with rational expectations. Mavroeidis (2004) shows that the usual "weak instruments" problem can arise naturally, when the predictable variation in expected future values is small relative to unpredictable future shocks.

In this thesis I will explore the possibility of weak identification when estimating simple interest rate rules for Norway. This issue has received attention in the recent literature, especially following the study by Mavroeidis (2010) on US data. In section 2 I will present the most famous interest rate rule, the Taylor rule. Section 3 provides a short overview of monetary policy in Norway. In section 4 the Generalized method of moments (GMM) is introduced. The moment conditions for GMM estimation of the Taylor rule are presented in the end of this section. Section 5 tries to give an intuitive explanation of the necessary dynamics for the Taylor rule to be identified. Section 6 illustrates the consequenses of strength of identifcation for GMM estimate distributions on simulated data. Section 7 is divided into two parts. In the first part, identification robust inference methods are presented. The second part reinvestigates the simulated data by estimating an identification robust confidence set. In section 8 I turn to real data for the Norwegian economy. In the first subsection, the measures of nominal interest rate, inflation and the output gap are presented. The second subsection explores weak identification in the standard GMM estimation of the Taylor rule. By the standard GMM I refer to the case where lagged values of the interest rate, inflation and output gap are used as instruments. This was the setup used by Clarida, Gali and Gertler (2000), which arguably serves as a benchmark for GMM estimation of the Taylor rule. Subsection 3 explores adding additional instruments to strengthen identification.

2 Interest rate rules

When conducting monetary policy, simple interest rate rules are used as a check of the robustness of the monetary policy. To my knowledge, no central bank claims to strictly follow an interest rate rule when setting interest rates. Instead most central banks minimize a loss function. A functional form of the loss function for Norges Bank will be presented in section 3. Estimating simple interest rate rules is however a popular exercise in monetary economics. This serves a dual purpose. On the one hand, they often approximate optimal policy. This provides a benchmark for central banks to compare their desired nominal interest rate against. On the other hand, they have proved to be good empirical representations of monetary policy. The rules estimated can be viewed as what rule would the central bank follow if we transform the loss function into a simple interest rate rule. The most famous interest rate rule is the Taylor rule proposed by John Taylor. The structure of the rule will be outlined below.

2.1 The original Taylor Rule

Taylor (1993) proposes the following rule for interest rate setting in the US.

$$r = p + 0.5y + 0.5(p - 2) + 2 \tag{1}$$

where r is the federal funds rate, p is the rate of inflation over the previous four quarters and y is the percent deviation of real GDP from a target. This rule was proposed for the US economy and an equivalent rule for the Norwegian Economy would likely have different coefficients. When interpreting the original Taylor rule it is useful to start at the point where both inflation and output is at its respective target levels. With an inflation target of 2, the interest rate that will prevail in equilibrium is 4. This can be seen as the equilibrium nominal interest rate. Equilibrium real interest rate will hence be 2. Along with the Taylor rule came what has been known as the Taylor principle. The Taylor principle states that the nominal interest rate should move more than one-for-one with respect to changes in inflation. The specification above suggests a 1.5-to-one ratio. The intuition is that the real interest rate drives output, which in turn affects inflation. To counteract the rise in inflation the Taylor principle must be active to help decrease inflation. If the Taylor principle was not active, an increase in inflation would be further strengthened because of the unability to increase the real interest rate.

2.2 Forward-looking Taylor rule with interest rate smoothing

The original Taylor rule was backward looking. The interest rate was set based on previous inflation and output gap. Many argued that central banks were looking at expected inflation and output gap when setting interest rates. This led to the introduction of a forward looking Taylor rule. The principle is still the same, but instead of looking at previous inflation and output, the rule now uses expected future values of these variables. The forward looking target rate seen in among others Clarida, Gali and Gertler (2000) is defined as follows.

$$i_t^* = i^* + \psi_\pi (E[\pi_{t+k}|\Omega_t] - \pi^*) + \psi_x E[x_{t+q}|\Omega_t] \tag{2}$$

where

i_t^* is the nominal interest according to the forward looking Taylor rule. I will call this the target interest rate for reasons that will be clear when introducing interest rate smoothing,

i^* is the equilibrium nominal interest rate that will prevail when both inflation and output is at its respective targets,

π_{t+k} is the annual inflation rate between period t and period t+k,

π^* is the inflation target,

x_{t+q} is the average output gap between period t and period t+q,

Ω_t is the information set containing all known information at time t,

ψ_π and ψ_x are the weight coefficients for respectively inflation and the output gap

In addition to basing the interest rate on expected future outcomes of inflation and output, it is widely recognised that central banks tend to smooth interest rates. One reason for this could be that the central bank wants to avoid large fluctuations in asset prices. Introducing interest rate smoothing the central bank's policy rule can be written as

$$i_t = (1 - \rho)i_t^* + \rho i_{t-1} + \varepsilon_{i,t} \tag{3}$$

where ρ is the smoothing coefficient,

i_t is the nominal interest rate prevailing at time t

$\varepsilon_{i,t}$ is an exogenous interest rate shock at time t.

By including a shock term we recognize the fact that changes to other variables than inflation and the output gap could affect nominal interest rates directly. We now observe that the nominal interest rate is set as a weighted average of the nominal interest rate last period and the target interest rate according to the forward looking Taylor rule. Say the target interest rate is 100 basis points higher than the actual nominal interest rate the previous period. Due to interest rate smoothing, the central bank will not choose the target rate as their nominal interest rate. In fact, the gap between the previous interest rate and the target rate will only be closed by the proportion $(1 - \rho)$. Inserting for the target interest rate in (4) gives us the following policy rule

$$i_t = (1 - \rho)(\alpha + \psi_\pi E[\pi_{t+k}|\Omega_t] + \psi_x E[x_{t+q}|\Omega_t]) + \rho i_{t-1} + \varepsilon_{i,t} \tag{4}$$

where $\alpha = i^* - \psi_\pi \pi^*$ for ease of notation.

Clarida, Gali and Gertler (2000) estimate the Taylor rule in (4) on US time series data. They use GMM as their estimation method, with lagged values of inflation, the output gap and the nominal interest rate as instruments. They find inflation coefficients below one up to 1980, and coefficients above one since then. They argue that this is evidence of monetary policy leading to an indeterminate equilibrium up until 1980, refered to as the pre-Volcker period. From 1980-1997, named the Volcker-Greenspan period, they argue that monetary policy lead to a determinate equilibrium.

Mavroeidis (2010) reinspects the results of Clarida, Gali and Gertler using identification robust inference. He finds that identification is weaker when monetary policy leads to a determinate equilibrium.

Skumsnes (2013) estimates the Taylor rule in (4) for the Norwegian economy with GMM. He finds that the inflation coefficient lies in the determinancy region for most of his regression specifications. The findings of Mavroeidis (2010) suggest that the estimated parameters from Skumsnes (2013) might suffer from weak identification. Based on this, I will conduct an analysis similar to that of Mavroeidis (2010) for Norwegian data later in this thesis.

3 Monetary Policy in Norway

Norges Bank is responsible for conducting monetary policy in Norway. The central banking law of 1985 regulates Norges Banks operations. The law states that Norges Bank shall be an executive as well as advisory organ for monetary, credit and currency policy. The bank's core objectives are to assist in keeping prices stabile through monetary policy, encourage fincancial stability and create robust and effective financial structures and payment systems. In addition, they are to manage the portfolio of The Government Pension Fund Global and handle the banks own currency reserves. Since 2001 Norges Bank has been regulated by the monetary policy regulation issued by the Ministry of Finance. The regulation states that the monetary policy objective should be low and stable inflation, which over time should be near 2.5%. Further, the long term objective of the monetary policy should be to provide a nominal anchor in the economy. A nominal anchor is a necessary prerequisite for stability in financial and housing markets. Norges Bank should also aim at stabilizing the Norwegian currency and expectations in development of the currency. The existence of multiple objectives is referred to as a flexible inflation targeting regime.

As stated in the monetary policy report (Norges Bank, 2012a), Norges Bank conducts monetary policy in order to minimize a loss function of the form

$$L = (\pi_t - \pi^*)^2 + \lambda(y_t - y^*)^2 + \gamma(i_t - i_{t-1})^2 + \tau(i_t - i^*)^2 \tag{5}$$

where π is inflation, y is output and i is the nominal interest rate. Subscripts represent the time period of the variable, and * denotes target values. The loss function is a weighted sum of inflation deviation from its target, output deviation from its target, the change in the interest rate and deviation from the interest target. The quadratic form of the loss function suggests that Norges Bank puts equal weight on deviations above and beyond target. The size of the coefficents indicates how much weight Norges Bank puts on the different deviations. Strict inflation targeting would suggest $\lambda = \gamma = \tau = 0$. Higher coefficients would suggest more flexible inflation targeting.

The main instrument for Norges Bank is the key policy rate. This is the rate commercial banks get from deposits in the central bank. Hence the key policy rate can be viewed as the floor of the interest rate corridor. No banks would be willing to lend out money at a lower interest rate than what it would get from a deposit in the central bank. Accordingly, the ceiling is the interest rate on banks overnight loans from the central bank.

4 GMM-estimation of the forward looking Taylor Rule

The original Taylor rule was based on predetermined values of inflation and the output gap and didn't include interest rate smoothing. When the Taylor rule is without interest rate smoothing, it will be linear in the parameters. This suggests the original Taylor rule could be estimated with linear methods such as ordinary least squares or two-stage least-squares (TSLS), depending on our assumptions between the explanatory variables and the error term. When including a smoothing parameter in the Taylor rule, it becomes non-linear in the parameters. This is a violation of the OLS- and TSLS assumptions that ensure unbiased and consistent estimators. Hence non-linear estimation methods should be used when facing interest rate smoothing.

In addition to the issue of interest rate smoothing, the Taylor rule in (4) is forward looking. The central bank uses expected future inflation to determine interest rates rather than current inflation. When using a forward looking-specification, the explanatory variables will be endogenous. The generalized method of moments (GMM) can handle non-linear equations with endogenous explanatory variables. Hence GMM would be a natural method for estimating a forward-looking Taylor rule.

When estimating the Taylor rule by GMM we make use of the rational expectations assumption. By replacing expectations with realizations and including the forecast errors, $(E_t x_{t+q} - x_{t+q})$ and $(E_t \pi_{t+k} - \pi_{t+k})$, with the correct weights in the disturbance, the Taylor rule can be rewritten as.

$$i_t = (1-\rho)(\alpha + \psi_\pi \pi_{t+k} + \psi_x x_{t+q}) + \rho i_{t-1} + \epsilon_t \tag{6}$$

where the error term consists of the original disturbance in addition to the forecast errors

$$\epsilon_t = -(1-\rho)(\psi_\pi(\pi_{t+k} - E[\pi_{t+k}|\Omega_t]) + \psi_x(x_{t+q} - E[x_{t+q}|\Omega_t])) + \varepsilon_{i,t} \tag{7}$$

GMM makes use of a set of moment conditions. The moment conditions are then used to solve for the parameters of our model. To be able to solve the system we need at least as many moment conditions as parameters we want to estimate. When the vector of moment conditions equals the number of parameters, we have exact identification. If we have more moment conditions than parameters, we have over-identification.

The moment conditions can be written in the form (Drukker, 2010)

$$E\left[\mathbf{m}(\mathrm{y_t}, \mathbf{x_t}, \mathbf{z_t}, \theta) = \mathbf{0}\right] \tag{8}$$

where $\mathbf{m}$ is a $q \times 1$ vector of functions whose expected values are zero in the population, y_t is the left hand side variable, $\mathbf{x_t}$ is the explanatory variable vector, $\mathbf{z_t}$ is the the instrument vector with dimension $q \times 1$ and θ is the parameter vector with dimension $k \times 1$, where $k \leq q$. The sample moments that correspond to the population moments are

$$\bar{\mathbf{m}}(\theta) = \frac{\mathbf{1}}{\mathbf{T}} \Sigma_{t=1}^{T} \mathbf{m}(\mathrm{y_t}, \mathbf{x_t}, \mathbf{z_t}, \theta) \tag{9}$$

When $k < q$, GMM minimizes the following objective function with respect to the vector, θ, of parameters

$$\hat{\theta}_{\mathbf{GMM}} \equiv \Sigma_{t=1}^{T} argmin_{\theta} \bar{\mathbf{m}}(\theta)' \mathbf{W} \bar{\mathbf{m}}(\theta) \tag{10}$$

where $\mathbf{W}$ is an estimation weighting matrix. The estimation weighting matrix can take on different forms, depending on the type of GMM estimator we want. In general, it's not possible to derive an explicit formula for the GMM estimator when the system is overidentified. Hence, minimizing the objective function makes us of numerical optimization methods.

When $k = q$ the GMM estimator solves $\bar{m}(\theta)$ exactly so that $\bar{\mathbf{m}}(\theta)' \mathbf{W} \bar{\mathbf{m}}(\theta) = \mathbf{0}$

The moment conditions in our Taylor regression will be of the form

$$E[\epsilon_t | \mathbf{Z_t}] = E[i_t - (1 - \rho)(\alpha + \psi_\pi \pi_{t+k} + \psi_x x_{t+q}) - \rho i_{t-1} | \mathbf{Z_t}] = \mathbf{0} \tag{11}$$

where $[\epsilon_t|\mathbf{Z_t}]$ is $l \times q$ matrix of the l moment conditions and $\mathbf{Z_t}$ is the vector containing the q variables the moments are conditioned on. We are estimating three (four including the constant) parameters and hence need at least three moment conditions in addition to $E[\epsilon_t] = 0$. Because of rational expectations and $E_{t-1}\varepsilon_{r,t} = 0$, any predetermined variable will be orthogonal to the disturbance term. In addition, current variables that are uncorrelated with the error term could be used as instruments. We suspect current values of inflation and the output gap to be correlated with the error term. Thus the list of possible instruments includes, but is not restricted to, all lags of inflation, interest rate and the output gap.

5 Identification

In the last section I showed that GMM fits parameter values that brings the moment conditions, often called orthogonality conditions, as close to their assumed value of 0 as possible. In addition to the orthogonality conditions, we need the instruments to be relevant for the parameters to be identified. An instrument will be relevant when the instrument can explain at least some of the variation in the endogenous variables. i_{t-1} is in the instrument set, and is clearly a relevant instrument with respect to i_{t-1}. Lagged values of π_t and x_t are however not necessarily relevant instruments for π_{t+k} and x_{t+q} without imposing additional assumptions. Cochrane (2011) proposes two assumptions that are needed for identification in a three equation New Keynesian sticky price model. The FED's reaction to deviations from the target values for inflation and output has to be the same as the relation between equilibrium interest rates and equilibrium inflation and output. Furthermore, the FED's choice of equilibrium must be such that the equlibrium quantities can be expressed in the autoregressive representation, with "equilibrium" parameters in the zone of determinancy. Below I try to give an intuitive explanation of the dynamics needed in a New Keynesian Sticky price model for the parameters to be identified.

5.1 Dynamics needed for the Taylor rule to be theoretically identified

For the purpose of simplicity, I will in this part set the smoothing parameter to 0. As long as the smoothing coefficient is within the unit circle it will not affect whether inflation or output is theoretically identified. Without smoothing the Taylor rule will look the following. Inflation horizon is set to 1 and the output gap considered is the current.

$$i_t = \psi_\pi E_t \pi_{t+1} + \psi_x x_t + \varepsilon_{i,t} \tag{12}$$

To discuss identification we need a closing model for the endogenous regressors. To close out the model, I will use a New Keynesian sticky price model in the line of Clarida, Gali and Gertler (2000). The model consists of a forward-looking Phillips curve, allowing for an exogenous inflation shock, and an output gap relation, allowing for an exogenous demand shock. The equilibrium conditions are log linearized around a steady state.

$$\pi_t = \beta E_t \pi_{t+1} + \lambda x_t + z_t \tag{13}$$

$$x_t = E_t x_{t+1} - (i_t - E_t \pi_{t+1}) + g_t \tag{14}$$

where g_t and z_t are the exogenous shocks and they could follow an AR(1) process

$$z_t = \rho_\pi z_{t-1} + \varepsilon_{\pi,t} \tag{15}$$

$$g_t = \rho_x g_{t-1} + \varepsilon_{x,t} \tag{16}$$

where ρ_π and ρ_x are the AR-coefficients of the inflation shock and output gap shock respectively. $\varepsilon_{i,t}$, $\varepsilon_{\pi,t}$ and $\varepsilon_{x,t}$ are assumed to be innovations with respect to time $t-1$.

Furthermore I will assume that there is a determinate equilibrium. Formally, the following condition ensures a determinate equilibrium in the above model specification.

$$\psi_\pi + \frac{1-\beta}{\lambda}\psi_x - 1 \geq 0 \tag{17}$$

Clarida, Gali and Gertler (2000) set $\beta = 0.99$ and $\lambda = 0.3$. Accepting those parameter values and looking at condition (17), we can observe that the condition for a determinate equilibrium largely coincides with the Taylor principle ($\psi_\pi > 1$). If the Taylor principle is satisfied and the output gap coefficient has the expected sign, there will always be a determinate equilibrium.

In a determinate equilibrium we know that the solution for inflation and the output gap can be written in the form:

$$\pi_t = D_1(z_t, v_t) \tag{18}$$

and

$$x_t = D_2(z_t, v_t) \tag{19}$$

where $v_t = g_t - \varepsilon_{i,t}$, and will be linear in the error terms.

$$\pi_t = a_{11}z_t + a_{12}g_t + a_{13}\varepsilon_{i,t} \tag{20}$$

$$x_t = a_{21}z_t + a_{22}g_t + a_{23}\varepsilon_{i,t} \tag{21}$$

By using the method of undetermined coefficients, the closed form solution becomes. See appendix B for calculations.

$$\begin{pmatrix} \pi_t \\ x_t \end{pmatrix} = \begin{pmatrix} \rho_\pi & \rho_x \\ \rho_\pi & \rho_x \end{pmatrix} \frac{D}{D_{det}} \begin{pmatrix} d_{22}\pi_{t-1} - d_{12}x_{t-1} \\ -d_{21}\pi_{t-1} + d_{11}x_{t-1} \end{pmatrix} + D \begin{pmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{x,t} - \varepsilon_{i,t} \end{pmatrix} \tag{22}$$

Looking at closed form solution we need $\rho_\pi \neq 0$ and $\rho_x \neq 0$ to make both 1-lagged values of inflation and the output gap possible instruments. If only one of shock processes were autoregressive of order 1, we would only have one instrument for two parameters, and hence the parameters would be underidentified. Adding additional dynamics will make more lagged values relevant instruments. In the case of the inflation and output gap shocks being AR(1), adding inflation smoothing and output smoothing would make the two-lagged values of inflation and output gap possible instruments. In general the sum of q and k will be the number of relevant instruments when the shock processes are AR(q) and inflation and output gap is an AR(k) process. It is important to note that these conclusion rely heavily on the interest rate shock having no autocorrelation. Krogh (2015) notes that the GMM estimator in the New Keynesian Phillips curve will be biased for any instrument in Z_{t-1} when the cost shock is autocorrelated. This leaves the parameters unidentified no matter the dynamics of the rest of the system. This also applies to the Taylor rule estimation in the sense that the interest rate shock can not be autocorrelated. This point is also made in Cochrane (2011).

6 Illustrating identification

Above we have seen that if the exogenous shocks to inflation and output gap are AR(1)-processes, π_{t-1} and x_{t-1} can be used as instruments for π_{t+1} and x_t. In theory this is of course true as long as both the lag coefficients are within the unit circle. In practice however, we often encounter the problem of weak identification. Weak identification arises when the predictable variation in the endogenous variables are weakly correlated with the variation in the instruments, relative to variation in the endogenous variables the instruments can not predict.

6.1 Investigating strength of identification through simulation

6.1.1 Higher AR-coefficients in the inflation and output gap shocks make identification stronger

In a determinate equilibrium, the expected future path of all endogenous variables are unique at time $t-2$. At time $t-1$ the i.i.d. portion of our exogenous shocks are realized and alters the expected future path of the endogenous variables. In all periods going forwards this happens. Because of no lagged endogenous variables in our model specification (still no smoothing), all expected variation in the endogenous variables comes from the AR(1)-process in the shock terms. π_{t-1} and x_{t-1} are instruments for π_{t+1}. The realizations of the i.i.d. disturbances at time $t-1$ will impact the realization of π_{t+1}. By the time π_{t+1} is realized, i.i.d disturbances from both time t and $t+1$ are realized, which also affects the realization of π_{t+1}. The strength of identification will in this case be how much the realization of i.i.d disturbances from time $t-1$ changes π_{t+1} relative to the realization of i.i.d disturbances from time t and $t+1$. Higher values of ρ_π and ρ_x should make the variation coming from time $t-1$ higher relative to that coming from t and $t+1$. I will investigate this by simulating the model from part 3 in Matlab with an extension called Dynare. Dynare handles simulation of DSGE-models with rational expectations. I will simulate three cases, one with low values on the AR(1) coefficients, one with medium valued AR-coefficients and a case with high valued AR-coefficients.

Model simulated:

$$i_t = \rho i_{t-1} + (1-\rho)(\psi_\pi E_t \pi_{t+1} + \psi_x x_t) + \varepsilon_{i,t} \tag{23}$$

$$\pi_t = \beta E_t \pi_{t+1} + \lambda x_t + z_t \tag{24}$$

$$x_t = E_t x_{t+1} - (i_t - E_t \pi_{t+1}) + g_t \tag{25}$$

$$z_t = \rho_\pi z_{t-1} + \varepsilon_{\pi,t} \tag{26}$$

$$g_t = \rho_x g_{t-1} + \varepsilon_{x,t} \tag{27}$$

Table 1: Simulation. Parameters being equal in all specifications

	ψ_x	β	λ
Parameter values	0.4	0.99	0.3

I start the simulation from the steady state for the model. For simplicity, the above model is constructed so that all steady state values are 0. In each period values for the i.i.d shock terms are drawn from a standard normal distribution. The values of the endogenous variables are then determined. Expectations are calculated using an end condition and applying rational expectations. Each DGP is simulated 500 times producing sample sizes of 1000 or 50. The parameters are then estimated using GMM in Stata. The distribution of the parameter estimates is “plotted” using a gaussian kernel.

a) (1)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.2	0.2

b) (2)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.5	0.5

c) (3)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.9	0.9

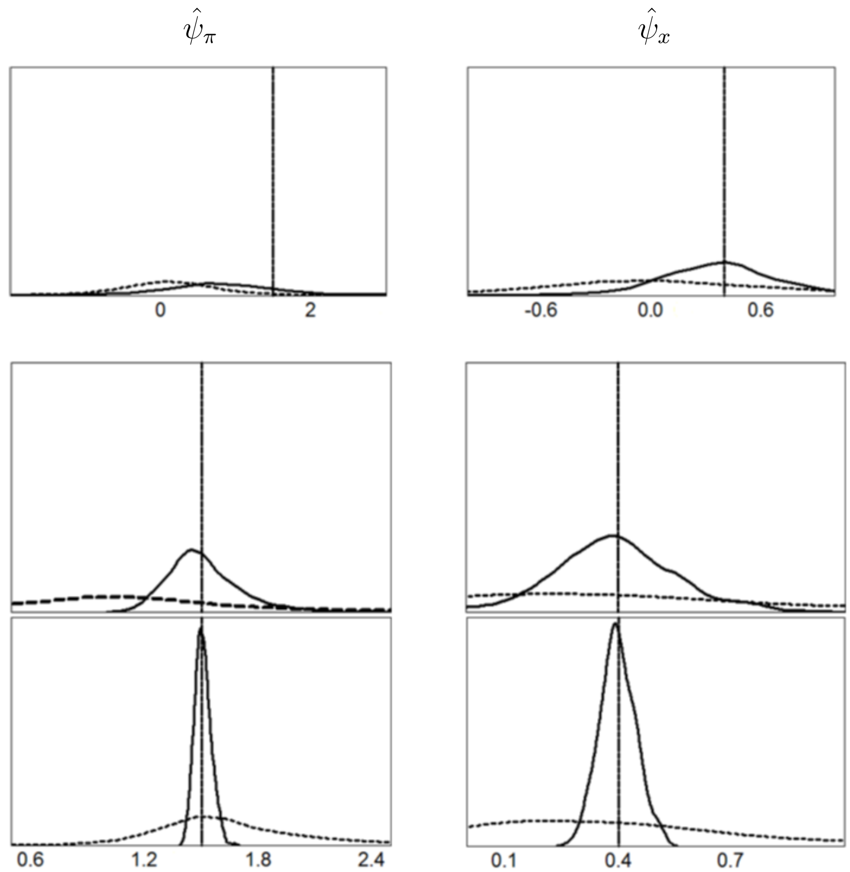


Figure 1: GMM parameter distributions on simulated data. Persistence in shocks and strength of identification.
a) Distribution of GMM estimates for ψ_π and ψ_x based on 500 simulations of specification 1.
b) Distribution of GMM estimates for ψ_π and ψ_x based on 500 simulations of specification 2.
c) Distribution of GMM estimates for ψ_π and ψ_x based on 500 simulations of specification 3.
Solid line represents sample size of 1000. Dotted line represents sample size of 50. Vertical line represents true parameter value.

Figure (1) gives the distribution of GMM estimates for different AR(1) coefficients in the inflation and output gap shocks. Simulation shows that the distribution gets narrower and more dense the higher the lag coefficients are. For the case of very weak identification, the inflation coefficient is very poorly estimated. The variation in our endogenous variables are in this case weakly correlated with instruments, and hence leads to weak identification. Our estimates get gradually better as identification becomes stronger, represented through the lag coefficients in the shock terms.

Worth noting is that the distribution does not get "dense" regardless of strength of identification for the sample size of 50. It gets better as identification strengthens, but even in specification (3) the variance of the GMM estimates are high.

6.1.2 Indeterminancy. Sunspot fluctuations improve identification.

I will now investigate what happens to identification when the model equilibrium is indeterminate. The model will be indeterminate when the Taylor principle is not satsified ($\psi_\pi < 1$). In theory an indeterminate equilibrium will increase strength of identifcation of the Taylor rule. The sunspot shock will add additional exogenous dynamics to the path of inflation and the output gap. This will strengthen the covariance between lagged and future realisations of inflation and the output gap. Mavroeidis (2010) finds that the parameter estimates for the pre-Volcker period are well identified. The parameter estimates for the Volcker-Greenspan period, however, do not seem to be well identified. According to Mavroeidis one possible explanation for this is that a determinate monetary policy removes the possibility of sunspot fluctuations, and mitigates the effect of shocks on future inflation and output. As a result, the expectations of these variables becomes less variable than they would be under an indeterminate monetary policy.

To be able to simulate an indeterminate equilibrium I have to modify the model to allow for sunspot fluctuations. I follow the modifications suggested by Farmer and Khramov (2013). $E_t\pi_{t+1}$ is replaced with pi_t in equations (23), (24) and (25). The realization of inflation at time t is then modeled as a fundemental shock.

$$\pi_t = pi_{t-1} + \sigma_s * sunspot_t \tag{28}$$

where σ_s is a parameter $sunspot_t$ is an iid. disturbance term and pi_{t-1} should be interpreted as $E_{t-1}\pi_t$. The dynamics of the model can now seem somewhat strange. The added equation (28) determines current inflation as the expectation from last period, plus a sunspot shock. Equations (23)-(27) now determine pi_t instead of π_t, in addition to the other endogenous variables. Expected inflation, pi_t, will thus be driven by the sunspot shock in addition to the persistence in the inflation and output gap shocks.

a) (2)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.5	0.5

b) (4)

ψ_π	ρ	ρ_π	ρ_x
0.5	0	0.5	0.5

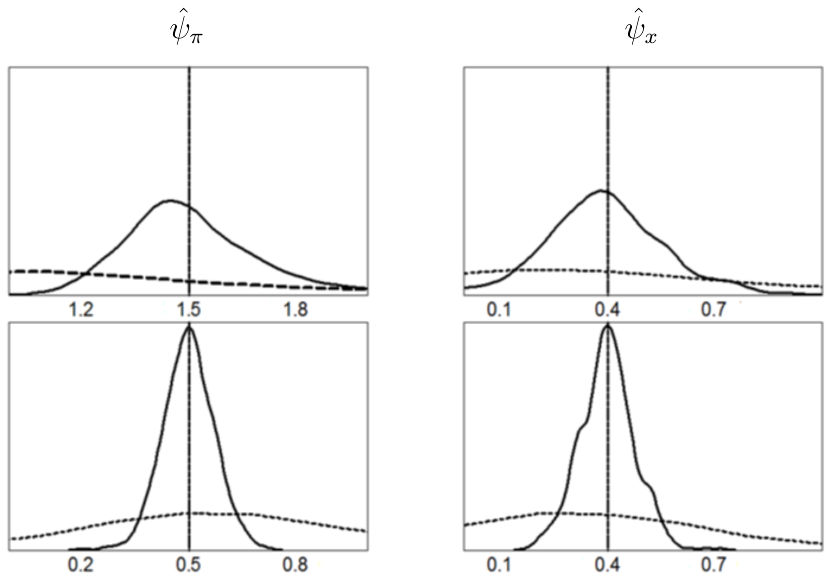


Figure 2: GMM parameter distributions on simulated data. Indeterminacy and strength of identification.
a) Distribution of GMM estimates for ψ_π and ψ_x based on 500 simulations of specification 2.
b) Distribution of GMM estimates for ψ_π and ψ_x based on 500 simulations of specification 4.
Solid line represents sample size of 1000. Dotted line represents sample size of 50. Vertical line represents true parameter value.

We see that the distribution of the parameter estimates are narrower and more dense for the indeterminate equilibrium. The key difference from the determinate equilibrium is that the endogenous variables π_t and x_t now also depend on the additional state variable pi_t. Since pi_t is autocorrelated, pi_{t-1} provides an additional relevant forcing variable for lagged values of inflation and the output gap. The sunspot shock associated with the indeterminate equilibrium produces more predictable variation on the endogenous variables from our instruments, and hence strengthens identification.

6.1.3 A high smoothing parameter makes it harder to identify the Taylor rule

By looking at the interest rate equation the potential problem of high smoothing becomes quite clear. If there is a high smoothing parameter, the interest rate will move less to changes in expected inflation and the output gap. The data could therefore support the claim of lower parameters for inflation and output gap even though this is not the case. Mavroeidis (2010) finds a higher smoothing parameter for the weakly identified period (1979-1997) than the well identified period (1961-1979). This suggests that policy has become more gradual over time, and provides an additional explanation for why the parameters have become weakly identified in the period 1979-1997.

a) (5)

ψ_π	ρ	ρ_π	ρ_x
1.5	0.8	0.9	0.9

b) (2)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.9	0.9

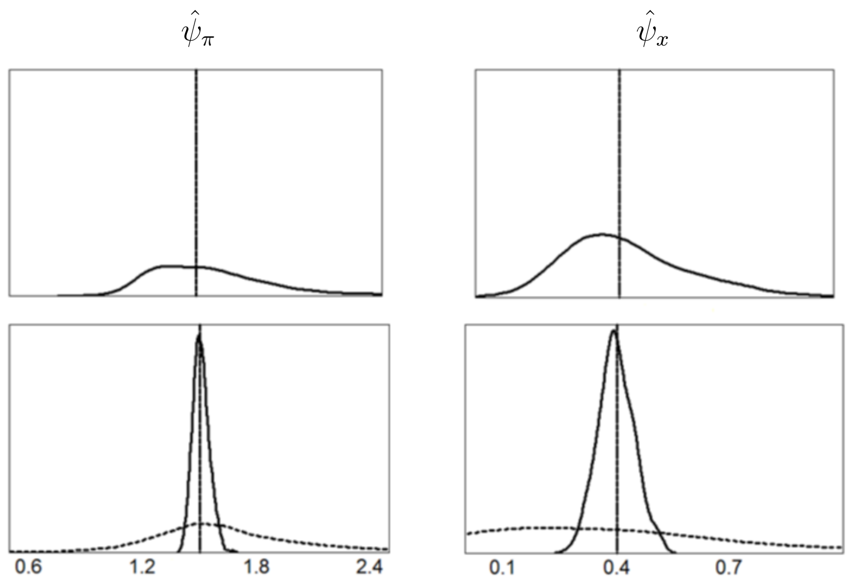


Figure 3: GMM parameter distributions on simulated data. Interest rate smoothing and strength of identification.
a) Distribution of GMM estimates for ψ_π and ψ_x based on 500 simulations of specification 5 for a sample size of 1000.
b) Distribution of GMM estimates for ψ_π and ψ_x based on 500 simulations of specification 2.
Solid line represents sample size of 1000. Dotted line represents sample size of 50. Vertical line represents true parameter value.

Introducing interest rate smoothing makes the distribution of the estimated parameters less dense. The smoothing parameter reduces the variation in the interest rate from changes in our endogenous variables. The result is that identification becomes weaker.

7 Investigating strength of identification with identification-robust inference

To investigate identification we first have to make the assumption that the true parameter vector is unique at θ_o. In GMM, the parameter vector θ is identified by the conditional mean restrictions $E[h(Y_t,\theta_o)|Z_t] = 0$, where θ_o is the true value of θ and Z_t is the vector of instruments. This implies that the numerator components of the GMM objective function has expectation zero under the true parameter value, $E[\phi_t(\theta_0)] = 0$. A necessary condition for θ to be identified is $E[\phi_t(\theta)] \neq 0$ for all $\theta \neq \theta_0$. The conditional moment restrictions are only satisfied when θ is at its true parameter value θ_0. Weak instruments can be seen as the case where $E[\phi_t(\theta)]$ is nearly 0 for some or many values of $\theta \neq \theta_0$.

Stock, Wright and Yogo (2002) provide an overview on the theory of weak instruments in linear instrumental variables (IV) regression, and weak identification in generalized method of moments (GMM) estimation. Weak instruments in IV estimation arise when the instruments are weakly correlated with included endogenous variables. In GMM estimation weak instruments correspond to weak identification of some or all of the unknown parameters. If instruments are weak, the sampling distribution of GMM statistics are in general nonnormal, and standard GMM point estimates, hypothesis tests, and confidence intervals are unreliable. The authors suggest one formal test for weak identification in nonlinear GMM proposed by Wright (2003). The conventional asymptotic theory of GMM requires the gradient of the moment conditions $\phi_t(\theta_o)$ to have full column rank. The test for weak identification formulates a null hypothesis of complete failure of this rank condition. Formally, this makes the test proposed by Wright (2003), a test for nonidentification or underidentification, not for weak identification. In simulations the proposed test has very little power in a sample size of 100. Hence, it will not be very useful to use this test on our empirical data, consisting of sample sizes of 100 or less.

Procedures that are fully robust to weak identification includes, the Nonlinear Anderson-Rubin statistic and Kleibergen's Statistic. Both procedures test $\theta = \theta_0$ in the nonlinear GMM setting. The tests are based on the continous-updating GMM objective function (Hansen et. al 1996), in which the weight matrix is evaluated at the same parameter value as the numerator.

Stock and Wright point out some symptoms of weak identification in empirical data. These include the objective function being clearly nonquadratic and has plateaus or ridges that in terms of LR statistics are not far from its minimum value. Another symptom is that the confidence sets computed by one of the two methods mentioned above will differ substantially from the conventional GMM confidence sets. In addition, one could look for asymptotically equivalent GMM estimators producing substantially different results.

The simplest method for obtaining identification-robust confidence sets is based on the S statistic proposed by Stock and Wright(2000). This is a generalization to GMM for the identification robust method due to Anderson and Rubin (1949). The Anderson-Rubin approach fixes the values of the parameters that may suffer from weak identification. In the Taylor rule setting, again without smoothing for ease of notation, this involves fixing the parameters ψ_π and ψ_x , to some values ψ_π^0 and ψ_x^0. Then running the regression

$$i_t - \psi_\pi^0 \pi_{t+1} - \psi_x^0 x_t = QV_t + u_t \tag{29}$$

where V_t is the vector of instruments, and Q is the vector of associated parameters. Testing the null hypothesis $(\psi_\pi, \psi_x) = (\psi_\pi^0, \psi_x^0)$ is equivalent to test the null hypothesis of $Q = 0$ because the instruments should be irrelevant when (ψ_π^0, ψ_x^0) coincides with the true parameter values. The 95% confidence set is obtained by keeping all parameter pairs (ψ_π^0, ψ_x^0) for which the null hypothesis is not rejected at the 5% level. Outlined above is the Anderson Rubin approach in the linear IV setting. The GMM version of the test proposed by Stock and Wright does not include regressing the residuals from the restricted model on the instruments. First ψ_π and ψ_x are fixed to some values ψ_π^0 and ψ_x^0. Thereafter the restricted parameter estimates ($\hat{\alpha}(\psi_\pi^0, \psi_x^0)$ and $\hat{\rho}(\psi_\pi^0, \psi_x^0)$ in case of smoothing) are obtained through GMM estimation on the restricted model. The S-statistic is then constructed by evaluating the CUE objective function at the fixed and restricted parameter estimates.

$$S_T^{CU}(\hat{\theta}) = \left[\sqrt{\frac{1}{T}}\sum \phi_t(\hat{\theta})\right]' \hat{W}\left(\hat{\theta}\right)^{-1}\left[\sqrt{\frac{1}{T}}\sum \phi_t(\hat{\theta})\right] \tag{30}$$

where $\hat{\theta} = \psi_\pi^0, \psi_x^0, \hat{\alpha}(\psi_\pi^0, \psi_x^0), \hat{\rho}(\psi_\pi^0, \psi_x^0)$. Since $\phi_t(\hat{\theta})$ is serially correlated $\hat{W}\left(\hat{\theta}\right)$ is replaced by an estimator of the spectral density of $\phi_t(\hat{\theta})$ at frequency 0. In my tests this estimator will be the HAC Bartlett with number of lags decided by the Newey West method. The number of lags will be decided by the formula, $lags = int\left(4(\frac{T}{100})^{\frac{2}{9}}\right)$. At this point I feel it is important to note that the number of lags included in the HAC Bartlett estimator can impact the value of the S-statistic significantly. In my tests, increasing the number of lags lead to substantial reduction in the S statistic for a given set of fixed and restricted parameters. To avoid "cheating" from adjusting the number of lags, I will in all constructed statistics use the number of lags decided by the above formula. The statistic obtained in (31) is evaluated against its critical value from the chi-square distribution. $S_T^{CU}(\theta)$ will be χ^2 distributed with degrees of freedom equal the number of instruments minus the number restricted estimated parameters. If $S_T^{CU}(\hat{\theta})$ is beneath the critical value, (ψ_π^0, ψ_x^0) is included in the identification robust inference confidence set.

Mavroeidis(2010) refers to the S-statistic based on the Anderson and Rubin approach as the AR-S statistic. Instead of using the AR-S statistic, he uses the conditional score statistic proposed by Kleibergen(2005), denoted K-LM. He argues that AR-S statistic is less powerful than the Wald test when identification is strong. In his words, "K-LM, which can be thought of as a version of the AR-S statistic that uses only an optimal combination of instruments". The KL-M confidence sets are based on the K statistic which is computed using the formula

$$K\left(\beta_0\right) = \frac{1}{4T}\left(\frac{\partial Q(\theta)}{\partial \theta'} \mid \hat{\theta}_0\right)\left[\hat{D}_T(\hat{\theta}_0, r)'\hat{V}_{ff}(\hat{\theta}_0)^{-1}\hat{D}_T(\hat{\theta}_0, r)\right]\frac{1}{4T}\left(\frac{\partial Q(\theta)}{\partial \theta'} \mid \hat{\theta}_0\right) \quad (31)$$

I now want to revisit the simulation data from the previous part, using identification robust inference. Based on the GMM results I will construct a confidence ellipse for the possibly weakly identified parameters. I will use the K-LM statistic to construct an identification robust confidence set for the same parameters. In addition I will plot all the actual GMM estimates to check whether the confidence set based on the GMM-estimates or KL-M statistic does the best job of reflecting the true distribution of the estimated parameters. To construct the expected confidence ellipse of the GMM-estimators, I will use the median values of the GMM estimates (ψ_π, ψ_x) and median value of the elements of the covariance matrix. By the same line of argument I will use the median values for the KL-M statistic associated with each parameter pair. The median values are chosen to represent a typical realization of the simulated data.

7.1 The KL-M confidence set converges towards the GMM confidence ellipse when identification is stronger

Since the AR-S and KL-M confidence sets take into consideration the non normality of the distribution when identification is weak, we expect these sets to be bigger than those from conventional GMM when identification is weak. In my model simulation strength of identification is represented by the size of the lag coefficient in the inflation shock and output gap shock. First, I want to show that the identification robust inference confidence set closes in on the GMM Wald ellipse when identification is stronger. As identification gets stronger the distribution will become more similar to the distributions assumed by conventional GMM inference. Hence the rejection regions, and when inverted the acceptance regions, will become more similar under stronger identificaiton.

a) (2)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.5	0.5

b) (3)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.9	0.9

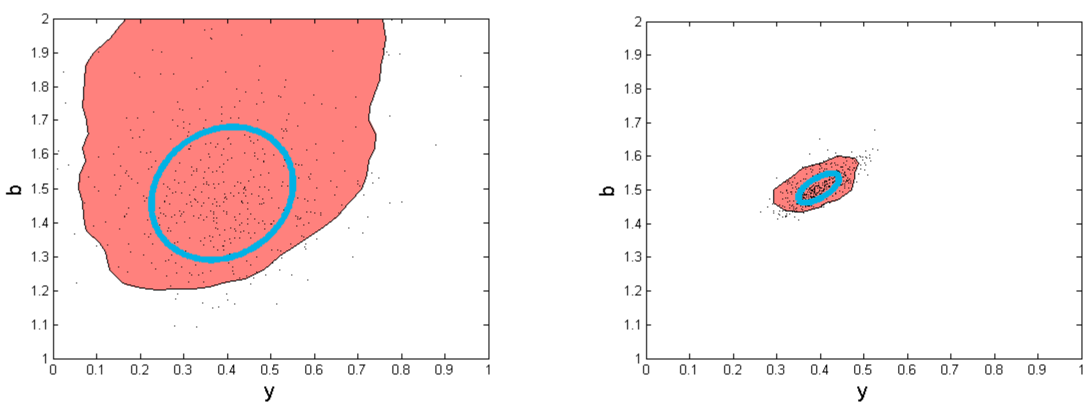


Figure 4: Simulation results. Strength of identification and the GMM Wald ellipse vs. KL-M confidence set.
a) Wald ellipse from GMM estimation and the KL-M confidence set for specification 2.
b) Wald ellipse from GMM estimation and the KL-M confidence set for specification 3.
In both figures actual GMM estimates are plotted with small black dots from 500 simulations. Sample size is set to 1000. Both ellipses and confidence sets are at the 95% level.

Figure 4a shows that there is quite a large region of disagreement between the Wald ellipse and the KL-M confidence set when the lag coefficients are set to 0.5 for both the inflation shock and the output gap shock. In figure 4b, the region of disagreement has become much smaller in absolute terms when the lag coefficients are set to 0.9. In absolute terms, the increase in the KL-M confidence set region is much bigger than the increased region of the Wald ellipse when identification becomes weaker, as we move from figure 4b to 4a. This is well in line with what theory on weak identifcation suggests. Large areas of disagreement between the Wald ellipse and identification robust inference confidence sets is a symptom of weak identification.

7.2 Indeterminancy makes identification stronger

As shown in Mavroeidis (2010) an indeterminate equilibrium leads to the endogenous variables π_t and x_t also depending on an additional state variable u_t that is driven by the sunspots. Since u_t is autocorrelated, u_{t-1} strengthens the correlation between the lagged values of π_t and x_t with future realizations. Identification robust confidence sets for the determinate and indeterminate equilibrium is reported below, when all other parameters are kept the same.

a) (2)

ψ_π	ρ	ρ_π	ρ_x
1.5	0	0.5	0.5

b) (4)

ψ_π	ρ	ρ_π	ρ_x
0.5	0	0.5	0.5

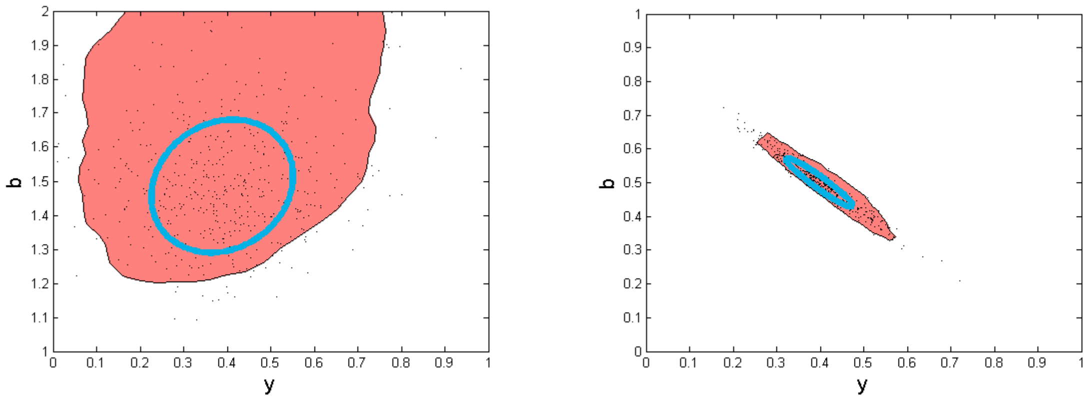


Figure 5: Simulation results. Indeterminacy and strength of identification. GMM Wald ellipse vs KL-M confidence set.
a) Wald ellipse from GMM estimation and the KL-M confidence set for specification 2.
b) Wald ellipse from GMM estimation and the KL-M confidence set for specification 4.
In both figures actual GMM estimates are plotted with small black dots from 500 simulations. Sample size is set to 1000. Both ellipses and confidence sets are at the 95% level.

When the equlilibrium is indeterminate, represented by $\psi_\pi = 0.5$, the KL-M confidence set becomes more similar to the Wald ellipse. This is the same as happened when identification got stronger in the last section. Hence, the simulation shows that sunspot fluctuation increases identification when all other parameters are kept equal. As noted above this is due to the fact that the sunspot fluctuations are represented by an additional state variable that produces increased correlation between the instruments and the endogenous variables.

8 Exploring weak identification in GMM estimation on Norwegian data

In this section I will explore whether GMM estimation of a Taylor rule for the Norwegian economy suffers from weak identification. To my knowledge this has not been done before.

8.1 Data

I will use the same data as in Skumsnes (2013), while also looking at a longer time span in certain situations. The longest sample period will be from 1985 to 2012. A longer time span is chosen to avoid small sample problems as far as possible. When including additional instruments, the sample period is restricted by the length of the time series for the additional instruments.

8.1.1 The interest rate

The three month Norwegian Interbank Offered rate (NIBOR) will be used as the measure for the interest rate. NIBOR is calculated by taking the trimmed mean interest rate from six panel banks that operate in Norway, leaving out the highest and lowest value. The panel banks for NIBOR are DNB Bank ASA, Danske Bank, Handelsbanken, Nordea Bank Norge ASA, SEB AB and Swedbank. NIBOR is calculated for different maturities. The time horizon of three months is chosen to match with our quarterly data. The NIBOR rate is often used as a reference for the money market rates between the different banks. Hence the NIBOR rate is supposed to reflect what rates the banks require to lend to other banks. I want to use the money market rates as a proxy for the short term nominal interest rate. Because of the close correlation between money market rates and NIBOR, NIBOR is chosen as our measure of the money market rates. Figure 6 depicts the time series of NIBOR from 1985 to 2012.

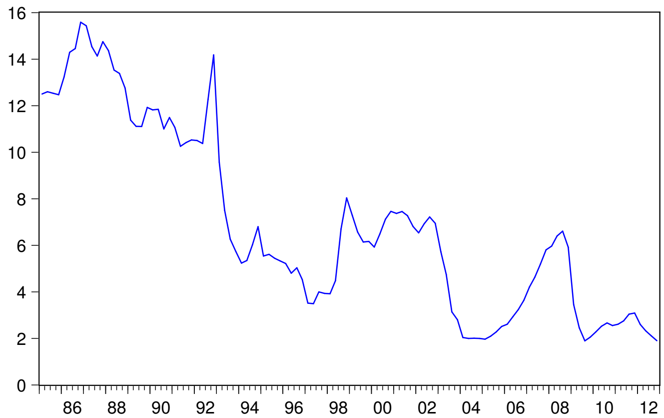


Figure 6: Time series of NIBOR from 1985 to 2012

8.1.2 Inflation

The consumer price index adjusted for tax changes and excluding energy commodities, CPIATE, will be used to construct the measure for inflation. CPIATE is often referred to as core inflation. As stated in Skumsnes (2013): "It is reasonable to exclude energy commodities for small open economies like Norway because these prices can be taken as exogenous and say little about price changes in Norway". The four quarter log differential is used as the measure for 1-year inflation. In figure 7 the time series of inflation measured by the CPIATE is drawn against the inflation target of 2.5. It is important to note that Norges Bank didn't receive the instruction of targeting an inflation rate of 2.5 before 2001. The inflation target is still drawn for the whole period, just as a comparison.

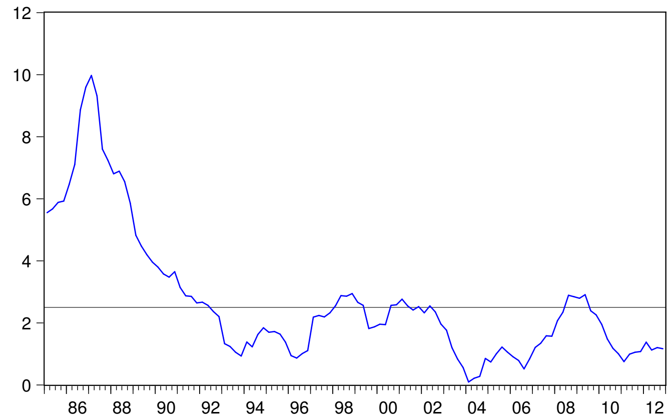


Figure 7: Time series of inflation measured by CPIATE from 1985 to 2012. Inflation target of 2.5 percent marked by the straight line for comparison.

8.1.3 The output gap

The output gap is unobservable. To estimate the output gap I use gross domestic product for Mainland-Norway. Mainland-Norway consists of all domestic production except from exploration of crude oil and natural gas, transport via pipelines and ocean transport. First a four quarter smoothing average is constructed for mainland GDP. Then a Hodrick Prescott filter is used to create a trend for mainland GDP. The Hodrick Prescott filter minimizes the following equation with respect to τ

$$Min_{\tau}\left(\sum(y_t-\tau_t)^2+\lambda\sum\left[(\tau_{t+1}-\tau_t)-(\tau_t-\tau_{t-1})\right]^2\right) \tag{32}$$

where y is actual output and τ is the trend. $\lambda = 0$, makes the trend equal to actual output. Increasing λ, increases the weight put on smoothing in the trend. When λ approaches infinity the trend approaches a straight line. Hodrick and Prescott (1997) suggested using $\lambda = 1600$ for quarterly American data. Statistics Norway uses $\lambda = 40000$, as they argue this fits the Norwegian economy better.

The output gap (OG) is simply the percent deviation in actual output from its trend estimated by the Hodrick Prescott filter

$$OG = 100\left(\frac{y-\tau}{\tau}\right) \tag{33}$$

The estimated output gap is of course sensitive to the choice of λ. Thus, choosing the "wrong" λ could produce severe problems when estimating the Taylor rule. As there is no observable counterpart to the estimated output gap, we are not able to deduce the correct λ with certainty. Skumsnes (2013) finds that changing λ from 40 000 to 1600 affects the estimated parameters for inflation and the output gap. Choosing a HP filter with $\lambda = 1600$ leads to the policy rule putting more emphasis on the output gap and less on inflation, compared to the HP filter with $\lambda = 40000$. I will choose $\lambda = 40000$ to compute the output gap in all regressions. Figure 8 displays the estimated output gap.

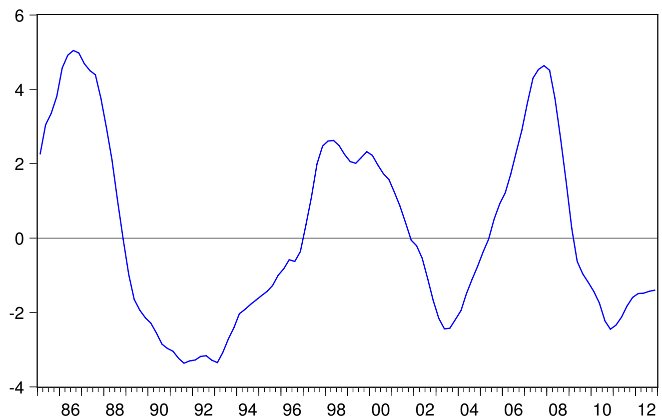


Figure 8: Time series of the output gap from 1985-2012

8.2 Estimating the Taylor rule for Norwegian data

8.2.1 Lags of interest rate, inflation and the output gap as instruments

Table 2: GMM results: Using 4 lags of interest rate, inflation and output gap as instruments.

α	ρ	ψ_π	ψ_x	*# of instruments*	J-stat (J-prob)	Se of reg
1.58	0.92*	2.06*	0.43		7.49	0.809
(1.31)	(0.04)	(0.59)	(0.79)		(0.59)	

Estimating $i_t = \rho i_{t-1} + (1-\rho)(\psi_\pi \pi_{t+4} + \psi_x x_{t+1}) + \epsilon_t$
Sample period: 1985Q1 - 2012Q4
Standard errors of the parameters are reported in paranthesis.
J-stat is the value of the objective function at the estimated parameter values with a 1-step weighting matrix.
J-prob is the p-value of the test of overidentifying restrictions.
Rejection of the null hypothesis at the 5% level is denoted by *.

Using 4 lags of the interest rate, output gap and inflation as instruments produce an estimate of ψ_π well within the determinancy region. Hence the results suggest that the Taylor principle was satisified in the period 1985-2012. Both the inflation parameter and the smoothing parameter are highly significant. The constant and output gap parameter are, however, not significant at the 5% level. Especially the output gap parameter has a low t-statistic. The output gap parameter is much lower than the inflation parameter. This suggests that Norges Bank reacts much stronger to deviations in inflation from target, than deviations in output gap from the target. The smoothing parameter is high and highly significant. This can be seen as evidence for Norges Bank having a high preference for gradualism when setting the nominal interest rate. At the same time a low weight is put on the target rate from the forward looking Taylor rule. To discuss the possible problem of weak identification I construct the Wald ellipse from the GMM estimates and the AR-S confidence set. The reported regression results use a sequential 1-step weighting matrix and coefficient iteration. To construct the AR-S statistic I have to use the CUE objective function. For my data, the CUE objective function produces lower function values than the 1-step weighting matrix and coefficient iteration regression. Hence, the p-values associated with the AR-S statistic are higher than those suggested by the reported regression results.

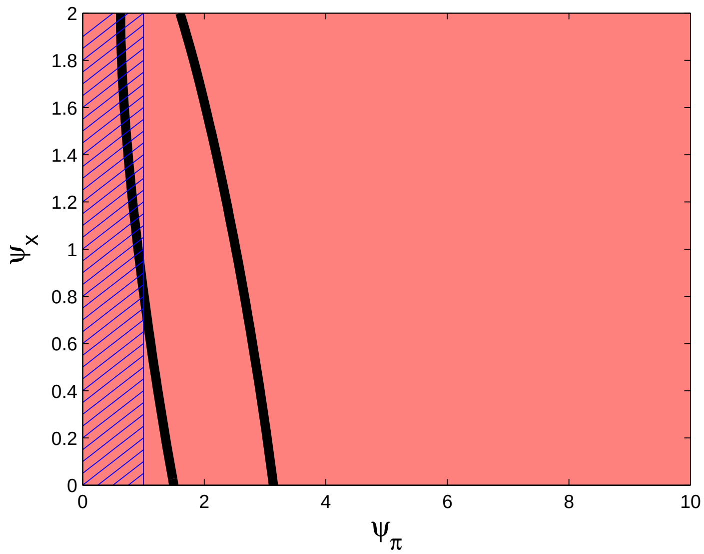


Figure 9: 95% Wald ellipse from GMM results and AR-S confidence set. Instrument set: 4 lags of nominal interest rate, inflation and the output gap.
Sample period: 1985Q1 - 2012Q4.
Plot represents GMM point estimates.

In figure 9, we observe that there are huge areas of disagreement between the Wald ellipse and the AR-S confidence set. The AR-S statistic tests a null hypothesis for the values of the pair $\psi_{\pi o}$ and ψ_{xo}. When varying our null hypothesis for the pair $\psi_{\pi o}$ and ψ_{xo}, between 0 and 10, and 0 and 2, respectively, we are unable to reject the null hypothesis for all combinations. This could be a sign of severe problems with weak identification in our GMM regression. The GMM ellipse lies largely within the determinancy region. The Wald ellipse by itself suggests that Norges Bank was conducting an acitve monetary policy in the period 1985 to 2012. The AR-S confidence set, however, lies both within the indeterminancy and determinancy region. These results warrant some further inspection of our regression results. In table (3) I have reported the regression output when $\psi_{\pi o}$ and ψ_{xo} are fixed at 10 and 2 respectively. We observe that the smoothing parameter is approaching 1, suggesting very little weight is put on the target interest rate.

Table 3: GMM results: Using 4 lags of interest rate, inflation and output gap as instruments, and fixing ψ_π and ψ_x

α	ρ	ψ_π	ψ_x	*# of instruments*	J-stat (J-prob)	Se of reg
-15.8	0.99*	10	2		9.74	0.809
(6.17)	(0.004)				(0.55)	

Estimating $i_t = \rho i_{t-1} + (1-\rho)(10\pi_{t+4} + 2x_{t+1}) + \epsilon_t$
Sample period: 1985Q1 - 2012Q4
Standard errors of the parameters are reported in paranthesis.
J-stat is the value of the objective function at the estimated parameter values with a 1-step weighting matrix.
J-prob is the p-value of the test of overidentifying restrictions.
Rejection of the null hypothesis at the 5% level is denoted by *.

When increasing the inflation parameter and output gap parameter, the restricted estimate of ρ reacts by approaching unity, and the constant parameter takes a "high" negative value. The change in the constant parameter happens almost by construction, since $i^* - \psi_\pi \pi^* = \hat{\alpha}$. The unrestricted regression in table 2 implies an equilibirum interest rate ($i*$) of 6.7, while the restricted regression in table 3 implies an equilibrium interest rate of 9.2. Both the J-stat and SE of the unrestricted and restricted regression are close. Based on this I want to investigate how much increasing $\psi_{\pi o}$ and ψ_{xo} actually affects the estimate of the nominal interest rate, and thereby the residuals. This is shown in figure (10).

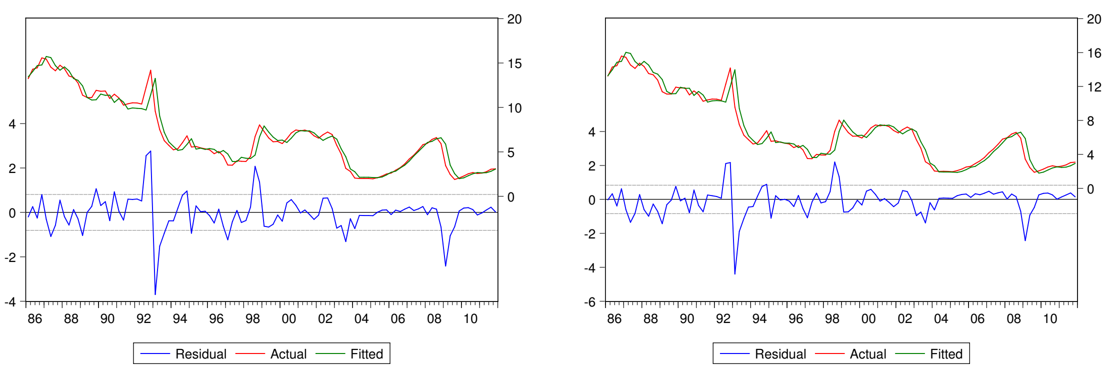


Figure 10: Actual, fitted and residuals from unrestricted and restricted regression. Instrument set: 4 lags of nominal interest rate, inflation and the output gap.
The left figure is for the unrestricted regression. The right figure is for the restricted regression $\psi_{\pi o} = 10$ and $\psi_{xo} = 2$.

When examining the residual from the unrestricted and restricted regression, we can see that the restricted estimation produces almost exactly the same residuals. The "numerator part" of the objective function consists of the sample estimates of the moment conditions on the form $\frac{1}{T}\Sigma_{t=1}^{T} Z_t \varepsilon_t$. The estimation weighting matrix used to construct the objective function for the overidentifying restriction test is also evaluated at the parameter values from the regression. Stock and Wright (2000) suggested looking for the objective function being clearly non-quadratic as a possible sign of weak identification. When concentrating out the well identified parameters α and ρ, our objective function with respect to ψ_π and ψ_π would be clearly non quadratic. In fact it would be almost flat and have no persistent rising slope when moving away from the true parameter values, assuming that the GMM estimates are the true parameter values.

To make the AR-S confidence set bounded for a "large" portion of my parameter window, I have to go as low as constructing a 25% confidence set. This is shown in figure 11.

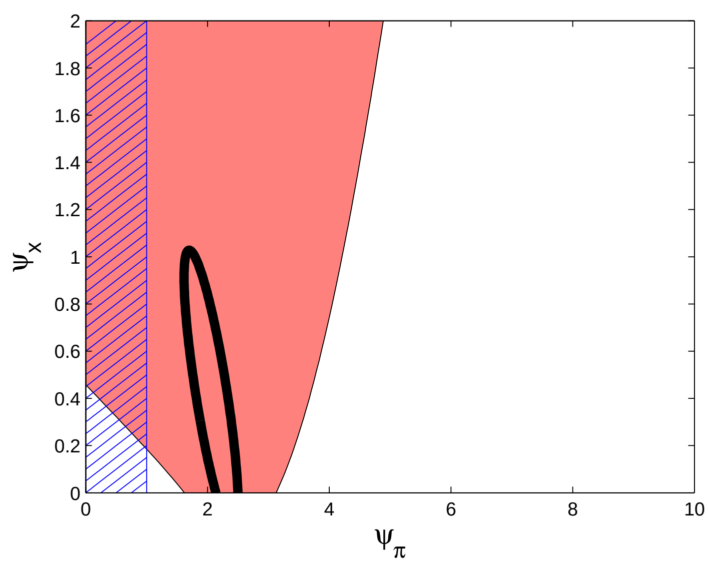


Figure 11: 25% Wald ellipse from GMM results and AR-S confidence set. Instrument set: 4 lags of nominal interest rate, inflation and the output gap.
Sample period: 1985Q1 - 2012Q4.
Plot represents GMM point estimates.

8.3 Adding instruments

Norway is a small open economy. This might suggest that additional instruments are needed to identify the parameters of our regression model. If inflation and output are partly driven from outside factors not contained in lagged values of inflation and output, including these factors should improve identification.

8.3.1 Adding commodity inflation and spread between short and long term interest rates as instruments

First of all, the spread between short and long term interest rate is not an outside factor. However, we might suspect that the spread could contain information of expectations for inflation and output gap. Hence, I try adding it as an instrument. Our inflation measure is based on the CPIATE, excluding energy and commodity prices. Nonetheless, it's not unreasonable to think that commodity prices could affect prices elsewhere in the economy. My time series of commodity inflation only stretches from 1993, so the estimation sample will be reduced. Table 4 reports the estimated parameters from the regression.

Table 4: GMM results: Adding commodity inflation and spread between long and short term interest rates as instruments

α	ρ	ψ_π	ψ_x	*# of instruments*	*J-stat (J-prob)*	*SE of reg*
0.06	0.87*	2.71*	0.27	21	11.91	0.63
(1.26)	(0.03)	(0.78)	(0.28)		(0.81)	

Estimating $i_t = \rho i_{t-1} + (1-\rho)(\psi_\pi \pi_{t+4} + \psi_x x_{t+1}) + \epsilon_t$
Sample period: 1993Q1 - 2012Q4
Standard errors of the parameters are reported in paranthesis.
J-stat is the value of the objective function at the estimated parameter values with a 1-step weighting matrix.
J-prob is the p-value of the test of overidentifying restrictions.
Rejection of the null hypothesis at the 5% level is denoted by *.

Including commodity inflation and spread as instruments don't make the inflation parameter much more significant. The t-statistic *for* ψ_π is almost unaltered. The t-statistic for ψ_x has increased, but is still insignificant. The implied long run nominal interest rate is almost unaltered at 6.8, compared to 6.7 from the regression in table 2. We saw earlier that inflation parameter was significant while the AR-S confidence set was very large. Based on the GMM estimates it's therefore not possible to conclude that identification is unchanged even though the estimates are pretty similar. I therefore again construct the AR-S confidence set and compare it with the Wald ellipse from the GMM estimates. This is shown in figure 12.

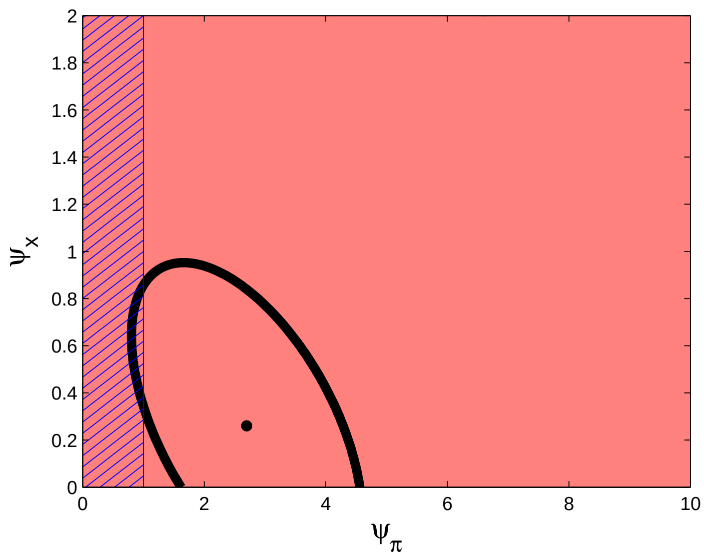


Figure 12: 95% Wald ellipse from GMM results and AR-S confidence set. Instrument set: 4 lags of nominal interest rate, inflation, the output gap, commodity inflation and spread beween short and long term bonds.
Sample period: 1993Q1 - 2012Q4.
Plot represents GMM point estimates.

Again we see huge disagreements between the GMM Wald ellipse and the AR-S confidence set. The Wald ellipse again suggests that Norges Bank was conducting an active monetary policy in the period 1993 to 2012. The identification robust confidence set is not able to rule out the points in the indeterminancy region. Again I show the regression results from the most extreme restricted estimation in table 5.

Table 5: GMM results: Adding commodity inflation and spread between long and short term interest rates as instruments, and fixing ψ_π and ψ_x

α	ρ	ψ_π	ψ_x	*# of instruments*	*J-stat (J-prob)*	*SE of reg*
-13.28*	0.97*	10	2	21	12.36	0.63
(1.63)	(0.004)				(0.87)	

Estimating $i_t = \rho i_{t-1} + (1-\rho)(10\pi_{t+4} + 2x_{t+1}) + \epsilon_t$
Sample period: 1993Q1 - 2012Q4
Standard errors of the parameters are reported in paranthesis.
J-stat is the value of the objective function at the estimated parameter values with a 1-step weighting matrix.
J-prob is the p-value of the test of overidentifying restrictions.
Rejection of the null hypothesis at the 5% level is denoted by *.

As in 8.2.1 the restricted estimate of ρ reacts by approaching unity when increasing $\psi_{\pi o}$ and ψ_{xo}. The implied long run equilibrium nominal interest rate is 11.72. Again I plot the residuals in figure 13 from the unrestricted and restricted estimation. Just as in the case with fewer instruments, the residuals are largely unchanged. This is shown in figure 13.

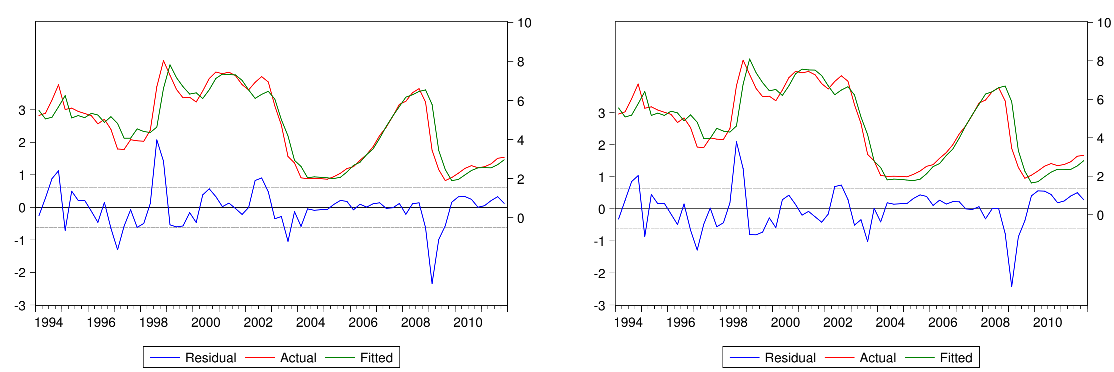


Figure 13: Actual, fitted and residuals from unrestricted and restricted regression. Instrument set: 4 lags of nominal interest rate, inflation, the output gap, commodity inflation and spread between short and long term bonds.
The left figure is for the unrestricted regression. The right figure is for the restricted regression $\psi_{\pi o} = 10$ and $\psi_{xo} = 2$.

8.3.2 Adding housing prices, equity return and foreign interest rate as instruments

Table 6: GMM results: Adding housing prices, equity return and foreign interest rate as instruments

α	ρ	ψ_π	ψ_x	*# of instruments*	*J-stat*	*SE of reg*
0.05	0.86*	2.41*	0.38*	33	12.53	0.518
(0.49)	(0.01)	(0.25)	(0.07)		(0.997)	

Estimating $i_t = \rho i_{t-1} + (1-\rho)(\psi_\pi \pi_{t+4} + \psi_x x_{t+1}) + \epsilon_t$
Sample period: 1998Q1 - 2012Q4
Standard errors of the parameters are reported in paranthesis.
J-stat is the value of the objective function at the estimated parameter values with a 1-step weighting matrix.
J-prob is the p-value of the test of overidentifying restrictions.
Rejection of the null hypothesis at the 5% level is denoted by *.

By adding more instruments, both the inflation and output gap parameter becomes highly significant.

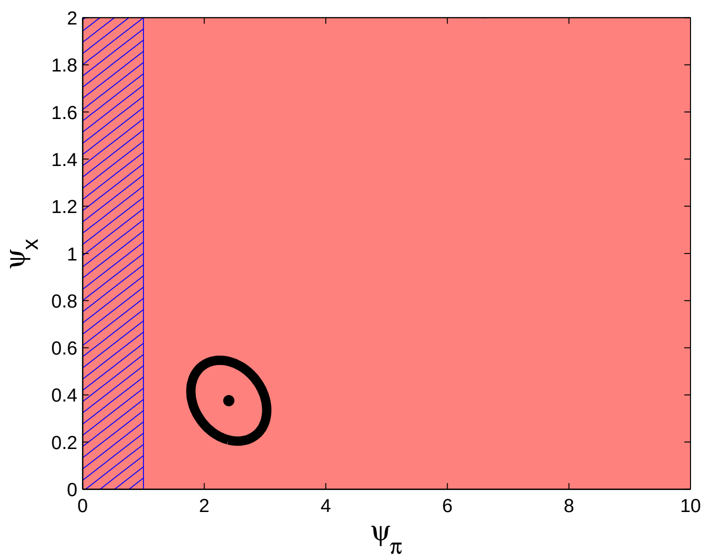


Figure 14: 95% Wald ellipse from GMM results and AR-S confidence set. Instrument set: 4 lags of nominal interest rate, inflation, the output gap, commodity inflation, spread beween short and long term bonds, housing prices, equity return and foreign interest rate.
Sample period: 1998Q1 - 2012Q4.
Plot represents GMM point estimates.

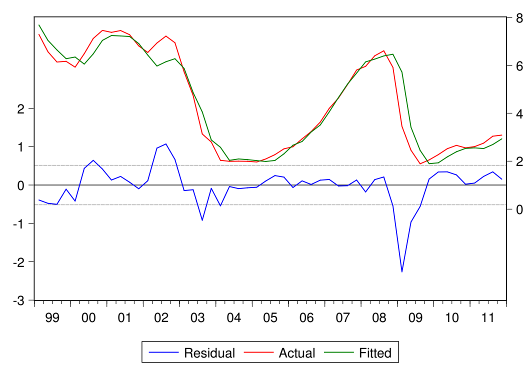
We observe that the confidene ellipse has become much smaller, and lies firmly within the determinacy region. The AR-S confidence set, however, is still unbounded within the parameter window. Hence, the additional instruments increase identification little or not at all.

Table 7: GMM results: Adding housing prices, equity return and foreign interest rate as instruments, and fixing ψ_π and ψ_x

α	ρ	ψ_π	ψ_x	*# of instruments*	*J-stat (J-prob)*	*SE of reg*
-13.28*	0.97*	10	2	33	13.10	0.549
(1.63)	(0.004)				(0.998)	

Estimating $i_t = \rho i_{t-1} + (1-\rho)(10\pi_{t+4} + 2x_{t+1}) + \epsilon_t$
Sample period: 1998Q1 - 2012Q4
Standard errors of the parameters are reported in paranthesis.
J-stat is the value of the objective function at the estimated parameter values with a 1-step weighting matrix.
J-prob is the p-value of the test of overidentifying restrictions.
Rejection of the null hypothesis at the 5% level is denoted by *.

Again ρ approaches unity when restricting ψ_π and ψ_x at 10 and 2, respectively. Actual, fitted and residuals is shown in figure 15.



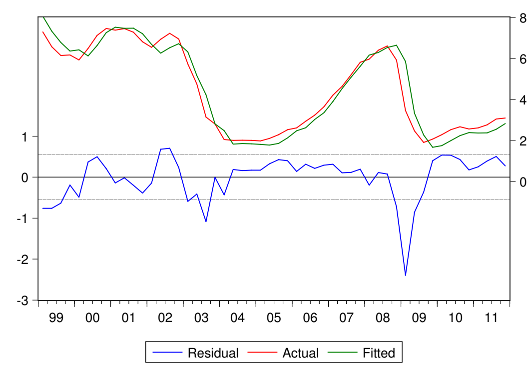


Figure 15: Actual, fitted and residuals from unrestricted and restricted regression. Instrument set: 4 lags of nominal interest rate, inflation, the output gap, commodity inflation, spread beween short and long term bonds, housing prices, equity return and foreign interest rate.
The left figure is for the unrestricted regression. The right figure is for the restricted regression $\psi_{\pi o} = 10$and $\psi_{xo} = 2$

8.3.3 Fixing the smoothing parameter

When the inflation parameter and output gap parameter are at its true value, the smoothing parameter should be well identified. From the results above there is a strong suggestion of ψ_π and ψ_x being weakly identified. Some of the unability to reject the null hypothesis seems to come from the weighting structure of the Taylor rule, combined with the strong persistence in the actual nominal interest rate. Skumsnes (2013) is not able to reject the null hypothesis of the interest rate being an AR(1) process. He argues that this could be because of the low power associated with the test. I am in no way suggesting that nominal interest rate is an AR(1) process, but it's close resemblence seems to be causing some problems in my empirical investigation. Identification is a system property, and as seen during simulations a higher smoothing parameter reduces identification. It should however not completely wipe out identification. My simulated data, although including a high smoothing parameter, produced much bigger variation in the nominal interest rate. In this case the AR-S confidence sets were able to reject null hypothesises far away from the GMM estimates. Although the simulated data arguably could be unrealistic, it perhaps provides som useful insight. With the nominal interest rate far from resembling an AR(1) process, the AR-confidence set became bounded within relatively close proximity to the GMM wald ellipse. When fixing parameters ψ_π and ψ_x the regression was not able to keep the residuals almost identical when looping through different values just by adjusting the smoothing parameter towards 1. To provide some clarity in whether this is our problem or not, I will in the following fix the smoothing parameter to its value from the GMM estimation. This will almost certainly change the residuals from the regression when looping through null hypothesis for the pair ψ_π and ψ_x. If the AR-S confidence set is still much the same, it suggests that the adjusment of the smoothing parameter to keep the residuals almost unchanged, is not the source of the weak identification.

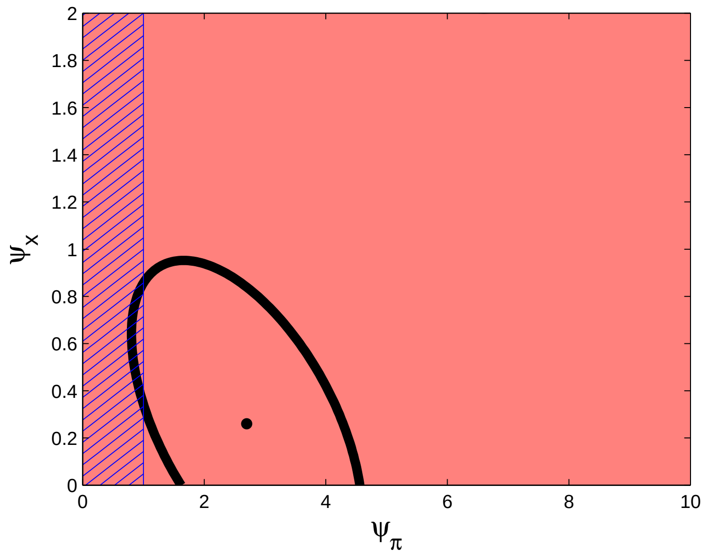


Figure 16: Fixing the smoothing parameter. 95% Wald ellipse from GMM results and AR-S confidence set. Instrument set: 4 lags of nominal interest rate, inflation, the output gap, commodity inflation and spread beween short and long term bonds. Sample period: 1993Q1 - 2012Q4.
The smoothing parameter is fixed at 0.87.
Plot represents GMM point estimates.

From figure 16 we see that fixing the smoothing parameter does not lead to any combinations of parameters being rejected in my window. This seems to suggest that the source of weak identification is not the resemblance of the interest rate series with a unit root series alone. Figure 17 shows that residuals are now considerably different, driven by the change in ψ_π and ψ_x. But this does not change the value of the CUE objective function evaluated at the estimated and fixed parameters much.

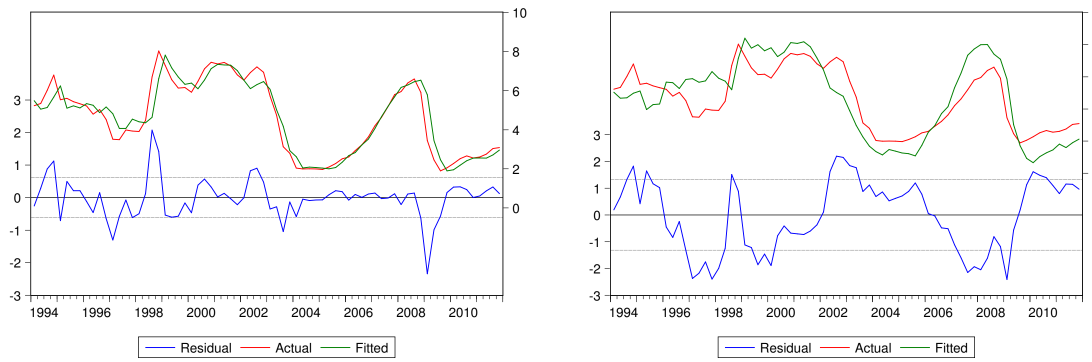


Figure 17: Fixing the smoothing parameter. Actual, fitted and residuals from unrestricted and restricted regression. Instrument set: 4 lags of nominal interest rate, inflation, the output gap, commodity inflation and spread beween short and long term bonds.
The left figure is for the unrestricted regression. The right figure is for the restricted regression $\psi_{\pi o} = 10$ and $\psi_{xo} = 2$

From the analysis above it seems reasonable to conclude that the parameters for the Taylor rule are weakly identified. The identification robust inference confidence set is unbounded within the region explored, and differs substantially from the GMM Wald ellipse.

9 Conclusion: Why is the Taylor rule weakly identified for the Norwegian economy?

This thesis has investigated identification in a forward-looking Taylor rule. Using a New Keynesian sticky price model, necessary conditions for identification is stated. In the model version used in this thesis, persistence in the inflation and output gap shocks is necessary for identification. Estimation of simulated data illustrates this point. Methods for detecting weak identification are presented and applied to both simulated data and Norwegian data. Results on Norwegian data suggest that the parameters of the Taylor rule are weakly identified.

For the last decade much emphasis has been put on transparency in monetary policy. Every three months Norges Bank publishes a monetary policy report. Since 2005 an interest rate forecast has been included in this report. One of the reasons behind publishing the forecast is to improve the general understanding of the bank's reaction pattern. To produce the forecasts, Norges Bank uses a medium sized small open economy model, named NEMO. The interest rate path is derived by minimizing a loss function representing the monetary policy mandate and the Board's policy preference. A potential source of weak identification of the Taylor rule is inconsistency, or change, in the Board's policy preferences. In monetary policy this is often referred to as discretion. If the board's preferences lead to large deviations from the interest rate suggested by simple interest rate rules, the rule obviously becomes harder to identify.

Theory on monetary economics shows that gains can be made when the central bank conducts a credible monetary policy. One example of a credible policy could simply be to strictly follow a Taylor rule. Through the monetary policy report and other communication, Norges Bank tries to make the public aware of how it bases its interest rate decisions. Evidence seems to suggest that this communication has been highly succesful. In terms of estimating the Taylor rule this could potentially be a bad thing. By clear communication, Norges bank could be inconsistent but still credible. Translating this to the Taylor rule setting, Norges Bank can be seen as following different Taylor rules for different scenarios. As long as this is well communicated in advance, it will most likely not lead to a loss of credibility.

Another possible source of weak identification is that expectations vary too little.

When monetary policy is effective this will likely be the result. As stated in Mavroeidis (2010): "Good policy removes the possibility of sunspot dynamics, and mitigates the effects of shocks on future inflation and output. As a result, the expectations of these variables become less variable than they might otherwise be". This point was illustrated when comparing confidence sets for a determinate and indeterminate equilibrium in section 7. Expectations are not directly observable, and hence cause us some problems in verifying this claim. Epinion, on behalf of Norges Bank, conducts a survey on inflation expectations. In the survey economic experts, social partners, corporate executives and households are asked about their inflation expectations for different time horizons.

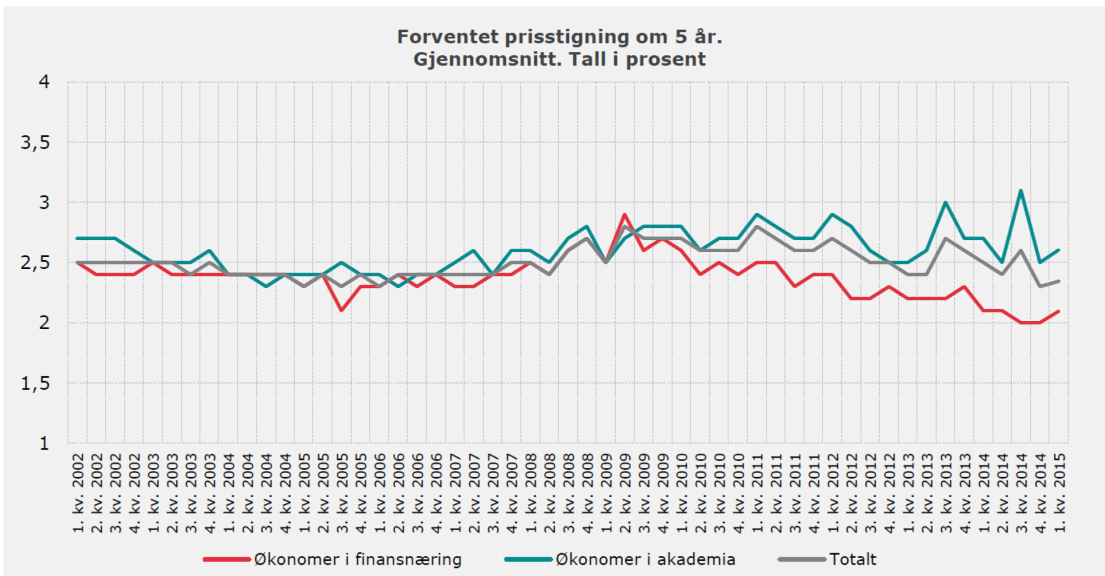


Figure 18: Five year inflation horizon expectations for economic experts. Source: Epinion, "Forventningsundersøkelse for Norges Bank 1. kvartal 2015.

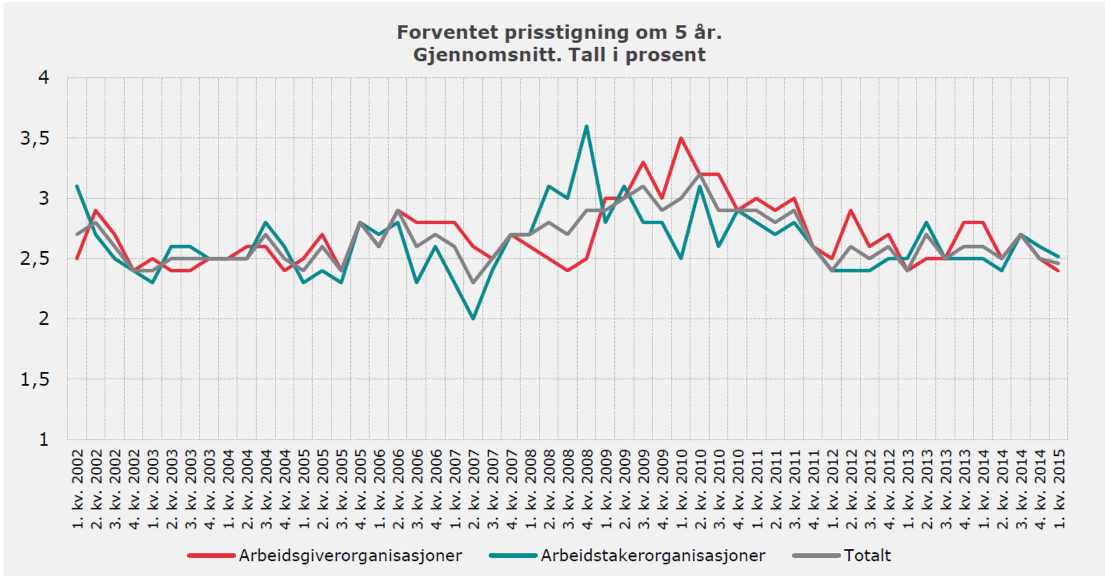


Figure 19: Five year inflation horizon expectations for social partners. Source: Epinion, "Forventningsundersøkelse for Norges Bank 1. kvartal 2015.

Figures 18 and 19 illustrates how well anchored inflation expectations are in the medium term. Expectations lie firmly anchored around the inflation target of 2.5 for most periods in the time span 2002 to 2015. Thus, the survey conducted by Epinion supports the claim of inflation expectations having low variation.

The GMM results based on Norwegian data suggest a high degree of policy intertia. As shown in section 7 and 8, this leads to problems in identifying the parameters for the target interest rates. The high degree of smoothing by Norges Bank is likely part of the reason for the observed weak identification of the forward-looking Taylor rule.

As a final note: The Taylor rule does a good job of describing monetary policy in Norway. The consequence of weak identification is simply that many different parameter values provide almost equally good fits. Thus the estimated Taylor rule still suggests that Norges Bank conducts monetary policy robust to simple interest rate rules. The estimated parameters, however, should be treated with care. They should be seen as one possible specification describing monetary policy, rather than a definitive answer to how Norges Bank sets the interest rate.

References

ALSTADHEIM, R., BACHE, I., HOLMSEN, A., MAIH, J. AND RØISLAND, Ø. (2010) (No. 11). *Staff Memo: Monetary policy analysis in practice.* Norges Bank Staff Memo.

ANDERSON, T.W. AND RUBIN, H. (1949). Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Annals of Mathematical Statistics, vol. 20, 46-63.*

BANK, NORGES. 2012a. *Norges Bank's Monetary Policy Report 1/2012.*

CLARIDA, R., GALI, J. AND GERTLER, M. (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *The Quarterly Journal of Economics, Vol. 115, 147-180.*

COCHRANE, J. (2011). Determinacy and Identification with Taylor Rules. *Journal of Political Economy, Vol. 119, 565-615.*

DRUKKER, D. M. (2010). An introduction to GMM estimation using Stata. *German Stata Users Group.*

EPINION. (2015) (February). *Forventningsundersøkelse for Norges Bank 1. kvartal 2015.*

FARMER, R. AND KHRAMOV, V. (2013). Solving and Estimating Indeterminate DSGE Models. *IMF Working Paper.*

HANSEN, L.P., HEATON, J. AND YARON, A. (1996). Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business & Economics Statistics, vol. 14(3), 262-280.*

HODRICK, R. AND PRESCOTT, E. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking, vol. 29, 1-16.*

KLEIBERGEN, F. (2005). Testing Parameters without assuming that they are identified. *Econometrica, vol. 73(4), 1103-1123.*

KROGH, T. (2015). Macro frictions and theoretical identification of the New Keynesian Phillips curve. *Journal of Macroeconomics, vol. 43, 191-204.*

Lucas, R. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy, vol. 1, 19-46.*

Mavroeidis, S. (2004). Weak identification of Forward-looking Models in Monetary Economics. *Oxford Bulletin of Economics and Statistics, vol. 66.*

Mavroeidis, S. (2010). Monetary Policy Rules and Macroeconomic Stability: Some New Evidence. *American Economic Review, vol. 100(1), 491-503.*

Skumsnes, A. H. (2013). *Estimating Taylor Rules for the Norwegian Inter Bank Offered Rate.* M.Phil. thesis, University of Oslo.

Stock, J. and Wright, J. (2000). GMM with Weak Identification. *Econometrica, Vol. 68, Issue. 5, 1055-1096.*

Stock, J., Wright, J. and Yogo, M. (2002). A survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics, vol. 20(4), 518-529.*

Taylor, J. (1993). Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy, vol. 39, 195-214.*

Wright, J. (2003). Detecting Lack Of Identification In GMM. *Econometric Theory, vol. 19(02), 322-330.*

Appendix A: Data sources

The dataset in this thesis is mainly the same dataset as used in Skumsnes (2013). For some variables the time series has been completed to allow for bigger samples in estimations.

i - Quarterly average of the 3-month Norwegian interbank offered rate, in annual percentages. Source: OECD Database.

π - Core inflation in Norway. Expressed as a 4 quarter log difference from the price index CPIJAE, in percentages. Source: Statbank, Statistics Norway.

x - Output gap for mainland Norway. Measures as a percentage deviation from a trend. The trend is a HP-trend with a smoothing parameter of 40 000. Source: Statbank, Statistics Norway.

Long-short spread in Norway - The difference between 10 year Norwegian bonds and the NIBOR. Source: OECD database.

World commodity price inflation - Expressed as a 4 quartler log difference from a price index in percentages. Source: IMF's International financial Statistics database.

Housing price gap in Norway - Measured as a percentage deviation from a trend. The trend is a HP-trend with a smoothing parameter of 40 000. Source: Statbank, Statistics Norway.

Foreign interest rates - The 4 quarter percentage difference of the euro-area 3 month interbank rate. Source: OECD database.

Equity return - The 4 quarter log difference of the OSEBX. Source: OECD database.

Appendix B

Method of undetermined coefficients calculation for equation (22)

$$\pi_t = d_{11} z_t + d_{12} v_t \tag{34}$$

$$x_t = d_{21} z_t + d_{22} v_t \tag{35}$$

$$\pi_t = \beta \left(d_{11} \rho_\pi z_t + d_{12} \rho_x v_t \right) + \lambda \left(d_{21} z_t + d_{22} v_t \right) + z_t \tag{36}$$

$$x_t = \left(d_{21} \rho_\pi z_t + d_{22} \rho_x v_t \right) - \left(\left(\psi_\pi - 1 \right) \left(d_{11} \rho_\pi z_t + d_{12} \rho_x v_t \right) + \psi_x \left(d_{21} z_t + d_{22} v_t \right) \right) + v_t \tag{37}$$

$$\pi_t = \left(\beta d_{11} \rho_\pi + \lambda d_{21} + 1 \right) z_t + \left(d_{12} \rho_x + \lambda d_{22} \right) v_t \tag{38}$$

$$x_t = \left(d_{21} \rho_\pi - \left(\psi_\pi - 1 \right) d_{11} \rho_\pi + \psi_x d_{21} \right) z_t + \left(d_{22} \rho_x - \left(\psi_\pi - 1 \right) \left(d_{12} \rho_x \right) + \psi_x d_{22} + 1 \right) v_t \tag{39}$$

$$d_{12} = d_{12} \rho_x + \lambda d_{22} \tag{40}$$

$$d_{22} = d_{22} \rho_x - \left(\psi_\pi - 1 \right) \left(d_{12} \rho_x \right) + \psi_x d + 1 \tag{41}$$

$$d_{11} = \beta d_{11} \rho_\pi + \lambda d_{21} + 1 \tag{42}$$

$$d_{21} = d_{21} \rho_\pi - \left(\psi_\pi - 1 \right) d_{11} \rho_\pi + \psi_x d_{21} \tag{43}$$

$$d_{21} = \frac{-\left(\psi_\pi - 1\right) d_{11}\rho_\pi}{1 - \psi_x - \rho_\pi} \tag{44}$$

$$d_{11} = \left(\frac{1 - \psi_x - \rho_\pi}{(1 - \beta\rho_\pi)(1 - \psi_x - \rho_\pi) + \lambda\left(\psi_\pi - 1\right)\rho_\pi}\right) \tag{45}$$

$$d_{21} = \frac{-\left(\psi_\pi - 1\right)\left(\frac{1 - \psi_x - \rho_\pi}{(1 - \beta\rho_\pi)(1 - \psi_x - \rho_\pi) + \lambda(\psi_\pi - 1)\rho_\pi}\right)\rho_\pi}{1 - \psi_x - \rho_\pi} \tag{46}$$

$$d_{22} = d_{22}\rho_x - \left(\psi_\pi - 1\right)\left(\frac{\lambda d_{22}}{1 - \rho_x}\right)\rho + \psi_x d_{22} + 1 \tag{47}$$

$$d_{12} = \frac{\lambda d_{22}}{1 - \rho_x} \tag{48}$$

$$\begin{pmatrix} \pi_t \\ x_t \end{pmatrix} = D \begin{pmatrix} z_t \\ v_t \end{pmatrix} \tag{49}$$

where $D = (D_1; D_2)$ which can be rewritten

$$\begin{pmatrix} \pi_t \\ x_t \end{pmatrix} = D \begin{pmatrix} \rho_\pi z_{t-1} \\ \rho_x v_{t-1} \end{pmatrix} + D \begin{pmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{x,t} - \varepsilon_{i,t} \end{pmatrix} \tag{50}$$

Substituting for z_{t-1} and v_{t-1} and evaluating (28) at $t-1$

$$\begin{pmatrix} \pi_t \\ x_t \end{pmatrix} = \begin{pmatrix} \rho_\pi & \rho_x \\ \rho_\pi & \rho_x \end{pmatrix} \frac{D}{D_{det}} \begin{pmatrix} d_{22}\pi_{t-1} - d_{12}x_{t-1} \\ -d_{21}\pi_{t-1} + d_{11}x_{t-1} \end{pmatrix} + D \begin{pmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{x,t} - \varepsilon_{i,t} \end{pmatrix} \tag{51}$$