International Great Inflation and Common Monetary Policy∗

Jacek Suda† Anastasia S. Zervou‡

January 2014

Abstract

The Great Inflation of the 1970s was an international phenomenon. We study whether monetary authorities in the G7 countries were changing their responses to inflation in a similar manner during and following the Great Inflation era. Our results suggest that the common to the G7 countries inflation pattern during the Great Inflation period is associated with a common pattern in the monetary policy response to inflation. Specifically, first, we find that until the early 1980s monetary authorities in the G7 countries responded mildly to inflation and they systematically fought it throughout the 1980s. Second, we find that the estimated Taylor-rule coefficients on inflation are cointegrated, implying the existence of a long run relationship in the responses to inflation, during and right after the Great Inflation period. Third, we conduct a principal component analysis on the residuals of the estimated Taylor rules and conclude that the shocks’ structure cannot account enough for the monetary policies’ comovements. Finally, we find that the response to inflation weakens during the 2000s.

JEL classification: E52; E58; C22.

Keywords: International Monetary Policy, International Great Inflation, Time Varying Parameter Model.

∗We are very grateful to James Morley for his direction and support. We are very grateful to Chang-Jin Kim for sharing computer routines. We also thank James Bullard, Anindya Banerjee, Siddhartha Chib, Yunjong Eo, Anna Mikusheva, Athanasios Orphanides, B. Ravikumar, Barbara Rossi, Aarti Singh, Thanasis Stengos, participants of seminars at the University of Houston, Texas A&M University, Banque de France, the 2012 Midwest Macroeconomics Conference, 16th Conference on Computing in Economics and Finance, the 15th Texas Camp Econometrics, the 8th Conference on Research on Economic Theory & Econometrics and the Midwest Econometric Group for helpful comments.

†Corresponding author. Banque de France, jacek.suda@gmail.com, http://www.jaceksuda.com.

‡Texas A&M University, 4228 TAMU, College Station, TX 77843, USA, azervou@econmail.tamu.edu, http://econweb.tamu.edu/azervou/, tel. +1 979-8457309, fax. +1 979-8478757.
1 Introduction

The Great Inflation was an international phenomenon. As Figure 1 shows all of the G7 countries experienced higher than usually inflation at approximately the same period, from the mid-1960s until the beginning of the 1990s. We explore Friedman and Allen (1970)’s view that "Inflation is always and everywhere a monetary phenomenon [...]" using the international setting. Specifically, we approximate monetary policy with a forward-looking, time-varying parameter Taylor rule and we estimate the G7 monetary authorities’ responses to inflation. Our results suggest that the G7 countries had been responding in a similar manner to inflation during the Great Inflation period; they were responding weakly to inflation during the 1970s and strongly from the mid-1980s until the mid-1990s. We find that the series of responses are cointegrated, revealing a long run relationship across the monetary policy of the G7 countries. These results suggest that the origins of the common inflation patterns in the G7 countries during the Great Inflation period, lay on the conduct of common monetary policy.\footnote{Previous work, using a different model that estimates Philips Curves in the G7 countries, finds that trend inflation is at its highest during the 1970s in the G7 countries (Morley et al., 2014). Note that trend inflation is related to long run expectations that the central banks may influence.}

Most of the previous research searches for an explanation for the Great Inflation in the United States.\footnote{With Nelson (2005a,b, 2007, 2008) as notable exceptions.} Figure 2 illustrates that after 1965 the CPI inflation in the US rose, reached 6% in 1969 and peaked at 15.7% in 1980. While some of the explanations suggested are based on the oil price shock or fiscal policy, it is the monetary policy that has been at the center of attention.\footnote{Blinder (1982) argues that the oil price shocks are responsible for the inflation pattern observed during the 1970s. Meltzer (2005) suggests an explanation based on the coordination of Federal Reserve System with the Treasury. He argues that although the monetary authorities were aware of the too high inflation they could not react due to political pressure. He points out that the Great Inflation started when the coordination of monetary policy become necessary for the government and ended when the Federal Reserve became independent.} DeLong (1997) argues that during the 1970s, the Federal Reserve was biased against sustaining low inflation due to the recent at that time experience of deflation and high unemployment of the Great Depression. Clarida et al. (2000) suggests that during the pre–Volker era, the nominal interest rate’s response to inflation was less than one to one creating a self-fulfilling expectations trap and leading to further inflation increase.\footnote{Lubik and Schorfheide (2004) estimate a DSGE New-Keynesian model and find that monetary policy...} Romer and Romer (2002, 2004) claim that during the 1960s and 1970s, monetary
authorities treated inflation as a cost-push phenomenon with the long-run tradeoff between inflation and unemployment discouraging them from aiming in controlling inflation.\footnote{5}{Similarly, Sargent (1999), Cogley and Sargent (2005) and Sargent et al. (2006) consider the argument that the Great Inflation was promoted by the monetary authorities changing views about the natural rate hypothesis. Nelson (2005a,b, 2007, 2008) propose the monetary neglect hypothesis according to which the Great Inflation was the result of monetary authorities attributing inflation to non-monetary reasons during the 1970’s.}

However, as mentioned earlier, looking outside the United States, we see that the inflation pattern across developed countries seems strikingly similar. Figure 1 shows that inflation was higher than usually at approximately the same period in the G7 countries, a feature that not all explanations of Great Inflation in the United States are able to account for. Specifically, the timing of the Great Inflation in other countries does not correlate with their Central Banks independence, and thus monetary authorities’ independence does not seem to be the reason behind the international Great Inflation.\footnote{6}{For example, Bundesbank was an independent bank from very early on, and the Banque de France, Bank of England and Bank of Japan became independent during the 1990s, when inflation had already leveled off.} In addition, observing Figure 1 we see that while the general inflation pattern is shared among the countries con-
considered, in many countries (e.g. France, UK and Japan) inflation had started increasing before the first oil shock. Thus, the oil shock does not seem to be the most important reason behind the common patterns of inflation across the G7 countries. Finally, the self-fulfilling expectations situation requires the trigger of the oil price shock, and thus the former critique applies also here.

This paper considers the Great Inflation as an international phenomenon and attempts to gather information from the cross-sectional dimension that the international experience offers, in order to contribute towards identifying its cause.

We study the evolution of monetary policy decisions in order to explore monetary authorities’ responsibility in sustaining high inflation in the G7 countries during the 1970s. In our empirical model we approximate these decisions using a forward-looking Taylor rule for the nominal interest rate and estimate monetary authorities’ response to inflation, output gap and smoothing parameter. We explore the evolution of monetary policy decisions over time using the time varying parameter framework (Kim and Nelson, 1989, 2006). This framework allows monetary authorities to change their policies in terms of their response

\[ \text{Equation} \]

\[ \text{Figure 2: CPI Inflation rate for the USA, quarterly, seasonally adjusted data.} \]
to inflation, output gap and in terms of smoothing behavior.

In addition, we consider the possibility that these changes have similar pattern in the G7 countries during the Great Inflation period. We explore the possibility of a long run, cointegrated relationship in monetary policy responses across countries at that period. The multi-country dimension of our analysis permits to examine whether there is commonality in the learning processes of the monetary authorities. Furthermore, we apply principal component analysis in the residuals of the Taylor rules in order to examine if there are strong common components in the shocks structure of the countries considered.

We find that the developed countries in our sample (Canada, France, Germany, Italy, Japan, UK and US) have not only similar inflation patterns but also similar monetary policy patterns. Specifically, we show that during the 1970s the G7’s monetary authorities were reluctant to increase interest rates targets in response to high inflation, while they were responding aggressively to inflation later and until the mid 1990s. We observe that after the mid 1990s the reaction to inflation diminishes for many of the G7 countries. Possibly, as inflation decreases and stabilizes central banks assign lower importance to it. Earlier work (Kim et al., 2006) has documented such a result, and we are able to identify it for other countries as well.

Furthermore, we find that the monetary authorities in the G7 countries exhibit a cointegrated relationship in their inflation responses during the Great Inflation period. The existence of common stochastic trend in the policy parameters indicates that there is a long-run relationship in policy conduct concerning inflation. In particular, if we exclude Japan we cannot reject the hypothesis that there is only one stochastic trend that is responsible for the long-run changes in monetary policy in all remaining countries. Thus, the common inflation patterns observed during the Great Inflation era are indeed related to common monetary policy patterns.

Lastly, we perform principal component analysis in the residuals of the G7 Taylor rules and we do not find strong common components. This result indicates that it was mostly common policies rather than common shocks that contributed towards the international phenomenon of the Great Inflation.

Previous work (DeLong (1997), Sargent (1999), Romer and Romer (2002, 2004), Cog-
ley and Sargent (2005), Nelson (2005a,b, 2007, 2008) Romer (2005), Sargent et al. (2006), DiCecio and Nelson (2009)) proposes that the Great Inflation of the 1970s was the result of perceptions of what monetary policy can and ought to do; these perceptions encouraged accommodative policy during the Great Inflation era.\footnote{Examples of such perceptions are the permanent trade-off between inflation and unemployment and monetary policy’s ineffectiveness towards cost-push inflation.} Romer (2005) argues that these perceptions were shared in a number of countries as policy makers read the same academic papers and form similar beliefs about the role of monetary policy at any point in time. Furthermore, Nelson (2005a,b, 2007, 2008) suggest a unified framework of beliefs prevailing among various developed countries. If these arguments are correct we should find similar cross-country patterns not only in inflation, as Figure 1 indicates, but also in the monetary authorities’ response to inflation. Our empirical results support this argument and empirically identify the missing link in the literature on the ideas-driven common inflation pattern by finding a common pattern in monetary policies response to inflation.

The rest of the paper is organized as follows. Section 2 introduces the model and Section 3 presents the estimation technique. Section 4 describes the data and Section 5 presents the results. Finally, Section 6 concludes.

\section{The Model}

Following Kim and Nelson (2006) and Kim (2008) we use a time-varying parameter model to estimate monetary policy rules for the G7 countries: Canada, France, Germany, Italy, Japan, the UK and the US.\footnote{Kim and Nelson (2006) used this framework for estimating a monetary policy rule for the United States. We use the same rule for all the countries considered in order to have a unified framework.}

We consider a forward-looking Taylor-type rule where the short term interest rate $r_{i,t}^*$ and the actual short term interest rate $r_{i,t}$ at time $t$ for country $i$ are specified as

$$r_{i,t}^* = \beta_{0,t}^* + \beta_{1,i,t}(E_t(\pi_{i,t,J}) - \pi_{i,t}^*) + \beta_{2,i,t}E_t(y_{i,t,J})$$

and

$$r_{i,t} = (1 - \theta_i,t)r_{i,t}^* + \theta_i,tr_{i,t-1} + m_{i,t}, \quad 0 < \theta_i < 1,$$

with
where for country $i$, $\pi_{i,t}^{*}$ is the target inflation rate, $\pi_{i,t,J}$ is the inflation rate from period $t$ to period $t + J$, $y_{i,t,J}$ is the average output gap from period $t$ to period $t + J$ and $\beta_{0,i,t}^{*}$ is the target short term interest rate when both inflation and output gap are equal to their target values. $\theta_{i,t}$ denotes the smoothing parameter (also time–varying) while $m_{i,t}$ denotes the random disturbance term.

Rewriting equations (1) and (2) we obtain the equation to be estimated:

\[ r_{i,t} = (1 - \theta_{i,t})(\beta_{0,i,t} + \beta_{1,i,t}\pi_{i,t} + \beta_{2,i,t}y_{i,t}) + \theta_{i,t}r_{i,t-1} + e_{i,t}, \]
\[ e_{i,t} \sim i.i.d.N(0, \sigma_{e,i}^{2}), \]

where $\pi_{i,t}$ denotes the current inflation rate, $y_{i,t}$ is the average output gap at the current period for country $i$, 

\[ \beta_{0,i,t} = \beta_{0,i,t}^{*} - \beta_{1,i,t}\pi_{i,t}, \]

and

\[ e_{i,t} = -(1 - \theta_{i,t})[\beta_{1,i,t}(\pi_{i,t} - E_{t}(\pi_{i,t+1})) + \beta_{2,i,t}(y_{i,t} - E_{t}(y_{i,t+1}))] + m_{i,t}, \]

where we let $J = 1$.

We assume that the time varying coefficients follow random walk dynamics:

\[ \beta_{k,i,t} = \beta_{k,i,t-1} + \epsilon_{k,i,t}, \]
\[ \epsilon_{k,i,t} \sim i.i.d.N(0, \sigma_{\epsilon,k,i}^{2}), \quad k = 0, 1, 2, 3. \]

In addition, the smoothing parameter $\theta_{i,t}$ is constrained to take values between 0 and 1 by setting it as follows:

\[ \theta_{i,t} = \frac{1}{1 + \exp(-\beta_{3,i,t})}. \]

Our system of equations consists of the interest rate equation (3) and equations (5) and (6) that describe the transition of the time–varying coefficients.

The above model does not account for the possibility that the variance of the shocks could be changing over time. In addition, the above specification is subject to endogeneity problem as regressors are correlated with the error term. Finally, this is a non-linear model which we are going to linearize before we use the Kalman filter to estimate it.
We address all these issues below. In Section 2.1 we allow for a GARCH process for the variance of the disturbance term of the monetary policy rule. In Section 2.2 we follow Kim (2006) in implementing a two-step estimation technique, in order to account for the endogeneity. Finally, in Section 2.3 we address the non-linearities in the estimation of coefficients of the monetary policy rule.

2.1 Heteroscedastic Disturbances

Sims and Zha (2006) point out that the US interest rate instability could be due to the time–varying variance of the disturbances and not due to the time–varying parameters of the monetary policy reaction equation. As a result, a simple time–varying parameter model that does not account for a changing variance of the error term could produce spurious variation in the time–varying parameters. To address this issue we allow for a GARCH(1,1) process for the variance of the error term in the interest rate and replace equation (3) with equation (7):

\[
\begin{align*}
    r_{i,t} &= (1 - \theta_{i,t})(\beta_{0,i,t} + \beta_{1,i,t}\pi_{i,t} + \beta_{2,i,t}y_{i,t}) + \theta_{i,t}r_{i,t-1} + e_{i,t}, \\
    e_{i,t} | I_{t-1} &\sim i.i.d. N(0, \sigma_{e,i,t}^2),
\end{align*}
\]

where

\[
\sigma_{e,i,t}^2 = a_0 + a_1 e_{i,t-1}^2 + a_2 \sigma_{e,i,t-1}^2,
\]

and \( I_{t-1} \) summarizes information up to time \( t - 1 \).

2.2 Endogeneity

In equation (7), the regressors are correlated with the error term \( e_{i,t} \) given by equation (4). In this case, the standard Kalman filter cannot be applied. To address the endogeneity we employ instrumental variables and the augmented Kalman filter framework introduced by Kim (2006, 2008) and Kim and Nelson (2006). This approach, based on Limited Information Maximum Likelihood, generalizes Pagan’s (1984) work on generated regressors to the time varying parameters framework.

Let \( z_{i,t} \) be a \( L \times 1 \) vector of instruments not correlated with \( e_{i,t} \). Then, we can write the
instrumental variables equations:

$$\pi_{i,t} = z_{i,t}' \delta_{\pi,i,t} + v_{\pi,i,t}, \quad v_{\pi,i,t} \sim i.i.d. N(0, \sigma_{\pi,i,t}^2),$$

(9)

and

$$y_{i,t} = z_{i,t}' \delta_{y,i,t} + v_{y,i,t}, \quad v_{y,i,t} \sim i.i.d. N(0, \sigma_{y,i,t}^2),$$

(10)

where $\delta_{\pi,i,t}$ is a $L \times 1$ vector of time-varying coefficients, $v_{\pi,i,t} = \sigma_{\pi,i,t} v_{\pi,i,t}^*$, $v_{y,i,t} = \sigma_{y,i,t} v_{y,i,t}^*$, and $v_{\pi,i,t}^* \sim i.i.d. N(0, 1)$, and $v_{y,i,t}^* \sim i.i.d. N(0, 1)$ are the standard errors. Since there could be shocks that alter the relationship between the endogenous variables in the current period and the instruments, we allow for a time-varying relationship between the endogenous variables and the instruments. Such specification has a desirable economic interpretation: the time-varying model is employed when inflation and output gap are estimated and forecasted by the central bank. Parameters in equations (9) and (10) are assumed to follow a random walk process:

$$\delta_{\nu,i,t} = \delta_{\nu,i,t-1} + \zeta_{\nu,i,t}, \quad \zeta_{\nu,i,t} \sim i.i.d. N(0_L, \Sigma_{\zeta,\nu,i}), \quad \nu = \pi, y,$$

(11)

where $\zeta_{\nu,i,t}$ is a $L \times 1$ vector of disturbances for the transition of the IV coefficients, $0_L$ is a $L \times 1$ vector of zeros and $\Sigma_{\zeta,\nu,i}$ is a $L \times L$ diagonal variance-covariance matrix of the IV coefficients.

In addition, we allow the error terms in the instrumental variable equations to be heteroscedastic and to follow a GARCH(1,1) process:

$$\sigma_{\nu,i,t}^2 = a_{\nu,0,i} + a_{\nu,1,i} v_{\nu,i,t-1}^2 + a_{\nu,2,i} \sigma_{\nu,i,t-1}^2, \quad \nu = \pi, y.$$

(12)

Given the oil price shocks of 1973 and 1979, it is important to allow for changing variance of the shocks affecting inflation and output gap. Moreover, our specification allows us to take into account changing uncertainty about future expected inflation and output gap estimates.\(^\text{10}\)

In the standard instrumental variable approach it is enough to model endogeneity by

\(^\text{10}\)This is true even for i.i.d. disturbances $v_{\nu,i,t}$, given the time-varying parameter framework. See Kim and Nelson (2006) for further discussion.
having \( \text{Cov}(v^*_{\nu,i,t}, e_{i,t}) \neq 0, \nu = \pi, y \). However, the time-variation in the IV equations, through \( \zeta_{\nu,i,t} \), offers an additional source of endogeneity. To address the issue we allow for \( e_{i,t} \) to be contemporaneously correlated with both \( v^*_{\nu,i,t} \) and \( \zeta_{\nu,i,t} \); that is, \( \text{Cov}(v^*_{\nu,i,t}, e_{i,t}) \neq 0 \) and \( \text{Cov}(\zeta_{\nu,i,t}, e_{i,t}) \neq 0 \) for \( \nu = \pi, y \). Since \( v^*_{\nu,i,t} \) and \( \zeta_{\nu,i,t} \) are not separately identified, the decomposition of \( e_{i,t} \) into a part that depends on \( v^*_{\nu,i,t} \), a part that depends on \( \zeta_{\nu,i,t} \), and an exogenous part is not feasible. Instead, using equation (11), we can decompose the instrumental variable equations (9) and (10) into predicted values and prediction errors, as follows:

\[
\pi_{i,t} = z_{i,t}' \delta_{\pi,i,t} - 1 + \eta_{\pi,i,t}, \eta_{\pi,i,t} | z_{i,t} \sim i.i.d. N(0, f_{\pi,i,t}), \quad (13)
\]

and

\[
y_{i,t} = z_{i,t}' \delta_{y,i,t} - 1 + \eta_{y,i,t}, \eta_{y,i,t} | z_{i,t} \sim i.i.d. N(0, f_{y,i,t}), \quad (14)
\]

where \( \eta_{\nu,i,t} = z_{i,t}' \zeta_{\nu,i,t} + \sigma_{\nu,i,t} v^*_{\nu,i,t} \) for \( \nu = \pi, y \), and we define

\[
f_{\chi,i,t} = \begin{bmatrix} f_{\pi,i,t} & \text{Cov}(v^*_{\pi,i,t}, v^*_{y,i,t}) \\ \text{Cov}(v^*_{\pi,i,t}, v^*_{y,i,t}) & f_{y,i,t} \end{bmatrix} = \begin{bmatrix} \sigma^2_{\pi,i} z_{i,t} + \sigma^2_{\pi,i,t} & \text{Cov}(v^*_{\pi,i,t}, v^*_{y,i,t}) \\ \text{Cov}(v^*_{\pi,i,t}, v^*_{y,i,t}) & \sigma^2_{\pi,i} z_{i,t} + \sigma^2_{y,i,t} \end{bmatrix},
\]

where \( v^*_{\nu,i,t}, \zeta_{\nu,i,t} \) do not correlate with each other, for \( \nu = \pi, y \). In such a formulation of instrumental variable equations (13) and (14), there is a part of the instrument that is correlated with \( e_{i,t} \) (that is \( \eta_{\pi,i,t}, \eta_{y,i,t} \)) and a part that is uncorrelated with \( e_{i,t} \). We define the standardized prediction error \( \eta^*_{\chi,i,t} \equiv f_{\chi,i,t}^{-1/2} \eta_{\chi,i,t} \), with \( \eta_{\chi,i,t} \equiv [ \eta_{\pi,i,t} \eta_{y,i,t} ]' \) and describe the correlation between \( e_{i,t} \) and \( \eta^*_{\chi,i,t} \) as

\[
\begin{bmatrix} \eta^*_{\chi,i,t} \\ e_{i,t} \end{bmatrix} \sim N \left( \begin{bmatrix} 0_2 \\ 0 \end{bmatrix}, \begin{bmatrix} I_2 & \rho \sigma_{e_{i,t}} \\ \rho \sigma_{e_{i,t}} & \sigma^2_{e_{i,t}} \end{bmatrix} \right), \quad (15)
\]

with \( \rho \equiv [ \rho_{\pi,i} \rho_{y,i} ]' \). In the above formulation, a non-zero correlation between \( e_{i,t} \) and \( \eta^*_{\chi,i,t} \), i.e., \( \rho \neq 0_2 \), corresponds to non-zero correlations \( \text{Cov}(v^*_{\nu,i,t}, e_{i,t}) \neq 0 \) and \( \text{Cov}(\zeta_{\nu,i,t}, e_{i,t}) \neq 0 \) for \( \nu = \pi, y \).

Similarly to Kim (2006), we decompose \( [\eta^*_{\chi,i,t} e_{i,t}]' \) into two independent normal shocks with unit variance using a Cholesky decomposition of the variance-covariance matrix in
The above equation helps us decompose \( e_{i,t} \) as

\[
e_{i,t} = \rho_{\pi,i} \sigma_{\pi,i,t} \eta^*_{\pi,i,t} + \rho_{y,i} \sigma_{\pi,i,t} \eta^*_{y,i,t} + \sqrt{1 - \rho_{\pi,i}^2 - \rho_{y,i}^2} \sigma_{\pi,i,t} \tau_{i,t},
\]

where \( \omega_{i,t} \equiv \sqrt{1 - \rho_{\pi,i}^2 - \rho_{y,i}^2} \sigma_{\pi,i,t} \tau_{i,t} \), and rewrite equation (7) as

\[
\tau_{i,t} = (1 - \theta_{i,t})(\beta_{0,i,t} + \beta_{1,i,t} \pi_{i,t} + \beta_{2,i,t} y_{i,t}) + \theta_{i,t} \tau_{i,t-1} + \rho_{\pi,i} \sigma_{\pi,i,t} \eta^*_{\pi,i,t} + \rho_{y,i} \sigma_{\pi,i,t} \eta^*_{y,i,t} + \omega_{i,t},
\]

\[
\omega_{i,t} \sim i.i.d. N(0, (1 - \rho_{\pi,i}^2 - \rho_{y,i}^2) \sigma^2_{\pi,i,t}),
\]

where \( \sigma^2_{\pi,i,t} \) is given by equation (8) and the disturbance term \( \omega_{i,t} \) is not correlated with the regressors.

The above model is estimated in two steps. Step 1: Estimate the system of IV’s, equations (9) - (12), and produce estimates for \( \eta^*_{\pi,i,t}, \eta^*_{y,i,t} \), i.e., \( \hat{\eta}^*_{\pi,i,t}, \hat{\eta}^*_{y,i,t} \). Step 2: Estimate the system of equations (5), (6) and (8) together with the following one:

\[
\tau_{i,t} = (1 - \theta_{i,t})(\beta_{0,i,t} + \beta_{1,i,t} \pi_{i,t} + \beta_{2,i,t} y_{i,t}) + \theta_{i,t} \tau_{i,t-1} + \gamma_{\pi} \hat{\eta}^*_{\pi,i,t} + \gamma_{y} \hat{\eta}^*_{y,i,t} + u_{i,t},
\]

\[
u_{i,t} = \omega_{i,t} + \gamma_{\pi}(\eta_{\pi,i,t} - \hat{\eta}^*_{\pi,i,t}) + \gamma_{y}(\eta_{y,i,t} - \hat{\eta}^*_{y,i,t}).
\]

For the estimation in Step 1 the correlation between \( \nu_{\pi,i,t} \) and \( \nu_{y,i,t} \) is assumed to be constant and \( \zeta_{\nu,i,t} \) is not correlated with \( \nu_{\nu,i,t} \), for \( \nu = \pi, y \). Also, for running the Kalman filter in the second step, it has to be that \( \epsilon_{k,i,t} \) is not correlated with \( \nu_{k,i,t}, \zeta_{\nu,i,t} \) or \( e_{i,t} \), for

\[\text{(15)}\]

\[
\begin{bmatrix}
\eta^*_{X,i,t} \\
e_{i,t}
\end{bmatrix}
= 
\begin{bmatrix}
I_2 & 0_2 \\
\rho' \sigma_{e,i,t} & \sqrt{1 - \rho' \rho \sigma_{e,i,t}}
\end{bmatrix}
\begin{bmatrix}
\eta^*_{X,i,t} \\
\tau_{i,t}
\end{bmatrix}
\sim i.i.d. N(0_3, I_3).
\]

The Cholesky decomposition decomposes the Variance-Covariance matrix in a lower diagonal matrix and its transpose:

\[
\begin{bmatrix}
I_2 & \rho' \sigma_{e,i,t} \\
0_2 & \sqrt{1 - \rho' \rho \sigma_{e,i,t}}
\end{bmatrix}
= 
\begin{bmatrix}
I_2 & 0_2 \\
\rho' \sigma_{e,i,t} & \sqrt{1 - \rho' \rho \sigma_{e,i,t}}
\end{bmatrix}.
\]

Then we can decompose \( \eta^*_{X,i,t} e_{i,t} \) into two independent shocks with unit variance (\( \eta^*_{X,i,t} \) and \( \tau_{i,t} \)), multiplied by the Cholesky factor, so the product will have the same variance as the the Variance-Covariance matrix of \( \eta^*_{X,i,t} e_{i,t} \).
\( k = 0, 1, 2, 3 \) and \( \nu = \pi, y \).

The main rational for this two-step approach is the following. The Kalman filter cannot be immediately applied because the regressors in equation (7) correlate with the error term. But if we decompose \( e_{i,t} \) in two components, one that correlates with the regressors and one that does not, i.e., \( e_{i,t} = E[e_{i,t} | \pi_{i,t}, y_{i,t}] + \omega_{i,t} \), then we can run the Kalman filter in a model that explicitly models the endogenous part of \( e_{i,t} \), and the exogenous error term, here \( \omega_{i,t} \). To this end, we first compute \( E[e_{i,t} | \pi_{i,t}, y_{i,t}] = \rho_{\pi,i} \sigma_{e_{i,t}} \eta_{\pi,i,t}^* + \rho_{y,i} \sigma_{e_{i,t}} \eta_{y,i,t}^* \), which correlates with the regressors. Next, we can compute \( \omega_{i,t} = e_{i,t} - E[e_{i,t} | \pi_{i,t}, y_{i,t}] \) that does not correlate with the regressors. Given that \( \eta_{\pi,i,t}^*, \eta_{y,i,t}^* \) are not observed, we use estimates \( \hat{\eta}_{\pi,i,t}^*, \hat{\eta}_{y,i,t}^* \) in equation (18). Now, given that in equation (17) the error term is exogenous (\( \omega_{i,t} \) does not correlate with \( \eta_{\pi,i,t}^* \) or \( \eta_{y,i,t}^* \)), we can apply the Kalman filter.

Note that the standard errors of the estimated coefficients \( \beta_{k,i,t}, k = 0, 1, 2, 3 \) must be adjusted for the generated regressors at the second step. The true variance of \( \beta_{k,i,t}, k = 0, 1, 2, 3 \) is calculated using the variance of \( e_{i,t} \) (given by equation (16)). The problem arises because when estimating equation (18) we compute variance for \( \beta_{k,i,t}, k = 0, 1, 2, 3 \) using the variance of \( u_{i,t} \), which actually converges to the variance of \( \omega_{i,t} \) rather than that of \( e_{i,t} \). In Section 3 we adjust for this issue.

### 2.3 Non-linear Estimation

Our equation of interest (18) is non-linear with respect to the coefficients that we want to estimate. Following Kim and Nelson (2006) we rewrite it as

\[
\begin{align*}
  r_{i,t} &= f(x_{i,t}; \beta_{i,t}) + \gamma_{\pi} \hat{\eta}_{\pi,i,t}^* + \gamma_{y} \hat{\eta}_{y,i,t}^* + u_{i,t}, \\
  u_{i,t} &= \omega_{i,t} + \gamma_{\pi} (\eta_{\pi,i,t}^* - \hat{\eta}_{\pi,i,t}^*) + \gamma_{y} (\eta_{y,i,t}^* - \hat{\eta}_{y,i,t}^*), \\
  \omega_{i,t} &\sim \text{i.i.d.} N(0, (1 - \rho_{\pi,i}^2 - \rho_{y,i}^2) \sigma_{e_{i,t}}^2),
\end{align*}
\]

where \( \beta_{i,t} = [\beta_{0,i,t} \beta_{1,i,t} \beta_{2,i,t} \beta_{3,i,t}]', x_{i,t} = [1 \ \pi_{i,t} \ y_{i,t} \ r_{i,t-1}]' \), and

\[
f(x_{i,t}; \beta_{i,t}) = (1 - \frac{1}{1 + \exp(-\beta_{3,i,t})})(\beta_{0,i,t} + \beta_{1,i,t} \pi_{i,t} + \beta_{2,i,t} y_{i,t}) + \frac{1}{1 + \exp(-\beta_{3,i,t})} r_{i,t-1}.
\]
The model is linearized around $\beta_{i,t} = \beta_{i,t|t-1} = E(\beta_{i,t}|I_{t-1})$ using first order Taylor expansion. Therefore,

$$f(x_{i,t}; \beta_{i,t}) \simeq f(x_{i,t}; \beta_{i,t|t-1}) + \frac{\partial f(x_{i,t}; \beta_{i,t|t-1})}{\partial \beta_{i,t}}(\beta_{i,t} - \beta_{i,t|t-1}),$$

and the linearization of equation (19) yields:

\begin{align*}
    r_{i,t} &= C_{i,t} + X_{i,t}' \beta_{i,t} + \gamma_{\pi} \hat{\eta}_{\pi,i,t} + \gamma_{y} \hat{\eta}_{y,i,t} + u_{i,t}, \quad (20) \\
    u_{i,t} &= \omega_{i,t} + \gamma_{\pi} (\eta_{\pi,i,t} - \hat{\eta}_{\pi,i,t}) + \gamma_{y} (\eta_{y,i,t} - \hat{\eta}_{y,i,t}), \\
    \omega_{i,t} &\sim i.i.d. N(0, (1 - \rho_{\pi,i}^2 - \rho_{y,i}^2)\sigma_{\omega,i,t}^2),
\end{align*}

where

\begin{equation}
    C_{i,t} = \frac{r_{i,t-1}}{1 + e^{\exp(-\beta_{3,i,t|t-1})}} - \frac{e^{\exp(-\beta_{3,i,t|t-1})} \beta_{3,i,t|t-1} (r_{i,t-1} - \beta_{0,i,t|t-1} \pi_{i,t} - \beta_{2,i,t|t-1} y_{i,t})}{(1 + \exp(\beta_{3,i,t|t-1}))^2}, \quad (21)
\end{equation}

and

\begin{equation}
    X_{i,t} = \begin{bmatrix}
        1 - \frac{1}{1 + \exp(-\beta_{3,i,t|t-1})} \\
        \pi_{i,t} - \frac{\pi_{i,t}}{1 + \exp(-\beta_{3,i,t|t-1})} \\
        y_{i,t} - \frac{y_{i,t}}{1 + \exp(-\beta_{3,i,t|t-1})} \\
        \frac{\exp(-\beta_{3,i,t|t-1}) (r_{i,t-1} - \beta_{0,i,t|t-1} \pi_{i,t} - \beta_{2,i,t|t-1} y_{i,t})}{(1 + \exp(\beta_{3,i,t|t-1}))^2}
    \end{bmatrix}. \quad (22)
\end{equation}

After the linearization, the interest rate is linear with respect to the $\beta$s in (20).

3 Empirical Strategy

We look for a long-run relationship between monetary policy changes in the G7 countries. To this end we employ a two-stage procedure. In the first stage of our empirical analysis, we estimate the Taylor rule for each country. We recover the estimated time-varying policy parameters and use them in the second stage, where we look for a cointegrating relationship between the estimated series. In the second stage we also perform principal component analysis on the Taylor rules’ residuals, generated at the first stage of our empirical analysis. Below we describe both these stages in detail.
3.1 Stage 1 Estimation: Country by Country Taylor Rule

Given the regressors’ endogeneity, we employ a two-step estimation procedure introduced by Kim (2006) and Kim and Nelson (2006), and briefly outlined in Section 2.2. At the first step we estimate the system of IVs equations (9) - (12), and generate estimates $\hat{\eta}^{*}_{\pi,i,t}$ and $\hat{\eta}^{*}_{y,i,t}$. Then we estimate the linearized system of equation (20)-(22) together with equations (5), (6) and (8), taking into consideration the non-linearity of the model and the heteroscedastic disturbances. Below we explain these two steps in detail. Note that the timing of the model assumes that the right hand side variables are available at the beginning of period $t$ and this knowledge is used to make inferences about the left hand side variables. New observations for the left hand side variables become available at the end of each period $t$. The notation $t|t-1$ denotes usage of knowledge of period $t$ for the right hand variables and of period $t-1$ for the left hand side variables.

3.1.1 Step 1: Generating $\hat{\eta}^{*}_{\pi,i,t}$ and $\hat{\eta}^{*}_{y,i,t}$

In the first step we estimate the instrumental variable equations and generate $\hat{\eta}^{*}_{\pi,i,t}$ and $\hat{\eta}^{*}_{y,i,t}$. Our state space is as follows:

$$
\nu_{i,t} = \begin{bmatrix} z_{i,t}' & 1 \end{bmatrix} \begin{bmatrix} \delta_{\nu,i,t} \\ \sigma_{\nu,i,t} v_{\nu,i,t}^{*} \end{bmatrix}, \quad \nu = \pi, y,
$$

(23)

and

$$
\begin{bmatrix} \delta_{\nu,i,t} \\ \sigma_{\nu,i,t} v_{\nu,i,t}^{*} \end{bmatrix} = \begin{bmatrix} I_L \\ 0_L' \end{bmatrix} \begin{bmatrix} \delta_{\nu,i,t-1} \\ \sigma_{\nu,i,t-1} v_{\nu,i,t-1}^{*} \end{bmatrix} + \begin{bmatrix} \zeta_{\nu,i,t} \\ \sigma_{\nu,i,t} v_{\nu,i,t}^{*} \end{bmatrix},
$$

(24)

$$
(\delta_{\nu,i,t} = A \delta_{\nu,i,t-1} + \zeta_{\nu,i,t}, \quad \zeta_{\nu,i,t} \sim i.i.d. N(0_L, \Sigma_{\zeta,\nu,i}), \quad \nu = \pi, y)
$$

Following Kim (2006) and Kim and Nelson (2006), we generate $\eta^{*}_{\nu,i,t|t-1} = f_{\nu,i,t|t-1}(\nu_{i,t} - \nu_{i,t|t-1})$ and use it as an estimate of $\eta^{*}_{\nu,i,t}$, for $\nu = \pi, y$. 

12
where $\Sigma_{\zeta,\nu,i}$ is a diagonal variance-covariance matrix and $L$ is the number of the time varying coefficients, including a time varying constant. Also, $\sigma_{\nu,i,t}^2$ is time varying and is given by equation (12).

We first run the Kalman filter and maximize the likelihood function, in order to generate estimates for the hyperparameters ($\Sigma_{\zeta,\nu,i}$, $a_{\nu,i,0}$, $a_{\nu,i,1}$, $a_{\nu,i,2}$) for the first step. The Kalman filter iteration is given below:

$$\tilde{\delta}_{\nu,i,t|t-1} = A\tilde{\delta}_{\nu,i,t-1|t-1},$$

$$P_{\nu,i,t|t-1} = AP_{\nu,i,t-1|t-1}A^\prime + \Sigma_{\zeta,\nu,i},$$

$$\tilde{\eta}_{\nu,i,t|t-1} = \nu_{i,t} - \tilde{z}_{i,t}^\prime \tilde{\delta}_{\nu,i,t|t-1},$$

$$f_{\nu,i,t|t-1} = \tilde{z}_{i,t}^\prime P_{\nu,i,t|t-1} \tilde{z}_{i,t},$$

$$\tilde{\delta}_{\nu,i,t} = \tilde{\delta}_{\nu,i,t|t-1} + P_{\nu,i,t|t-1} \tilde{z}_{i,t} f_{\nu,i,t|t-1}^{-1} \tilde{\eta}_{\nu,i,t|t-1},$$

$$P_{\nu,i,t} = P_{\nu,i,t|t-1} - P_{\nu,i,t|t-1} \tilde{z}_{i,t} f_{\nu,i,t|t-1}^{-1} \tilde{z}_{i,t}^\prime P_{\nu,i,t|t-1},$$

where $\tilde{\delta}_{\nu,i,t|t-1} = E(\tilde{\delta}_{\nu,i,t} | I_{t-1}), P_{\nu,i,t|t-1} = E(\tilde{\delta}_{\nu,i,t} - \tilde{\delta}_{\nu,i,t|t})^2$ and $\tilde{\eta}_{\nu,i,t|t-1} = \nu_{i,t} - \nu_{i,t|t-1}, f_{\nu,i,t|t-1} = E(\tilde{\eta}_{\nu,i,t|t-1}^2), \nu = \pi, y.$

Note that part of $\tilde{\delta}_{\nu,i,t}$ is the $\sigma_{\nu,i,t}$, which is not known and needs to be estimated. We use the unconditional mean of $\sigma_{\nu,i,t}^2$, calculated from equation (12), as initial value for the last diagonal element of $P_{\nu,i,t-1|t-1}$. Then, we use this information to construct an estimate of $\sigma_{\nu,i,t-1|t-1}v_{\nu,i,t-1}^\prime$, which we approximate as follows: $E(\sigma_{\nu,i,t-1|t-1}v_{\nu,i,t-1}^\prime \mid I_{t-1}) = E(v_{\nu,i,t-1} \mid I_{t-1}) = E(\sigma_{\nu,i,t-1|t-1}v_{\nu,i,t-1}^\prime \mid I_{t-1})^2 + Var(\sigma_{\nu,i,t-1|t-1}v_{\nu,i,t-1}^\prime \mid I_{t-1})$, for $\nu = \pi, y$. We get $E(\sigma_{\nu,i,t-1|t-1}v_{\nu,i,t-1}^\prime \mid I_{t-1})$ from the last component of $\tilde{\delta}_{\nu,i,t-1|t-1}$ and $Var(\sigma_{\nu,i,t-1|t-1}v_{\nu,i,t-1}^\prime)$ from the last diagonal element of $P_{\nu,i,t-1|t-1}$, as given by the Kalman filter above. Finally, we use equation (12) to derive estimates for $\sigma_{\nu,i,t}$ using $E(\sigma_{\nu,i,t-1|t-1}v_{\nu,i,t-1}^\prime \mid I_{t-1}) = E(v_{\nu,i,t-1} \mid I_{t-1})$ from above in place of $v_{\nu,i,t-1}$, and the last component of $P_{\nu,i,t-1|t-1}$ in place of $\sigma_{\nu,i,t-1}^2$.

Running the Kalman filter the second time we keep $\tilde{\eta}_{\nu,i,t|t-1}$ and $f_{\nu,i,t|t-1}$ and compute

\[\text{Note again that the notation } | I_{t-1} \text{ implies available information about } \tilde{z}_{i,t} \text{ from time } t, \text{ and available information about } \nu_{i,t-1} \text{ from period } t - 1, \text{ for } \nu = \pi, y.\]
\[ \eta_{\nu,i,t|t-1}^{\ast} = \frac{\eta_{\nu,i,t|t-1}}{\sqrt{f_{\nu,i,t|t-1}}} \], which we use in the second step of the estimation, for \( \nu = \pi, y \).

### 3.1.2 Step 2: Estimate the model, given \( \hat{\eta}_{\pi,i,t}^{\ast} \) and \( \hat{\eta}_{y,i,t}^{\ast} \)

We now estimate the model given by equations (20)-(22), (5), (6) and (8), substituting in equation (20) the elements of \( \eta_{\pi,i,t|t-1}^{\ast} \) in place of \( \hat{\eta}_{\pi,i,t}^{\ast} \) and \( \hat{\eta}_{y,i,t}^{\ast} \) estimated above. We start with equation (20) and rewrite the model in the following state-space form:

\[
\begin{align*}
    r_{i,t} & = C_{i,t|t-1} + \begin{bmatrix} X'_{i,t|t-1} & 1 \end{bmatrix} \begin{bmatrix} \beta_{i,t} \\ \omega_{i,t} \end{bmatrix} + \rho_{\pi,i} \sigma_{\epsilon_{i,t}} \eta_{\pi,i,t|t-1}^{\ast} + \rho_{y,i} \sigma_{\epsilon_{i,t}} \eta_{y,i,t|t-1}^{\ast}, \\
    (r_{i,t} &= C_{i,t|t-1} + \tilde{X}'_{i,t|t-1} \tilde{\beta}_{i,t} + \rho_{\pi,i} \sigma_{\epsilon_{i,t}} \eta_{\pi,i,t|t-1}^{\ast} + \rho_{y,i} \sigma_{\epsilon_{i,t}} \eta_{y,i,t|t-1}^{\ast})
\end{align*}
\]

and

\[
\begin{bmatrix} \beta_{i,t} \\ \omega_{i,t} \end{bmatrix} = \begin{bmatrix} I_4 & 0_4 \\ 0_4' \end{bmatrix} \begin{bmatrix} \beta_{i,t-1} \\ \omega_{i,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{i,t} \\ \omega_{i,t} \end{bmatrix},
\]

\[
\begin{bmatrix} \epsilon_{i,t} \\ \omega_{i,t} \end{bmatrix} \sim i.i.d. N \left( \begin{bmatrix} 0_4 \\ 0_4' \end{bmatrix} \begin{bmatrix} \Sigma_{\epsilon_{i,t}} & 0_4 \\ 0_4' \end{bmatrix} \begin{bmatrix} 1 - \rho_{\pi,i}^2 - \rho_{y,i}^2 \sigma_{\epsilon_{i,t}}^2 \end{bmatrix} \right),
\]

\[
(\tilde{\beta}_{i,t} = B \tilde{\beta}_{i,t-1} + \tilde{\epsilon}_{i,t}, \tilde{\epsilon}_{i,t} \sim i.i.d. N(0_5, \Sigma_{\epsilon_{i,t}}))
\]

where \( \Sigma_{\epsilon_{i,t}} \) is a 4x4 diagonal matrix with \( \sigma_{\epsilon,k,i}^2 \) as diagonal elements, for \( k = 0, 1, 2, 3 \), and \( \sigma_{\epsilon_{i,t}}^2 \) is given by equation (8).

The log-likelihood function that we maximize is:

\[
\ln L_r = \sum_{t=1}^{T} \ln \left[ \frac{1}{\sqrt{2\pi f_{i,t|t-1}}} \exp \left( -\frac{(r_{i,t} - r_{i,t|t-1})^2}{2f_{i,t|t-1}} \right) \right],
\]

where \( r_{i,t|t-1} = E(r_{i,t} | \bar{C}_{i,T}, \bar{X}_{i,T}, \bar{r}_{i,T-1}) \), for \( \bar{g} = [\bar{g}_1 \ \bar{g}_2 \ \ldots \ \bar{g}_T] \) and \( f_{i,t|t-1} = E(\eta_{i,t|t-1}^2) \).

The first round of Kalman filter iterations estimate the model’s hyperparameters \( (\rho_{\pi,i}, \rho_{y,i}, \Sigma_{\epsilon_{i,t}}) \) maximizing the likelihood function. The Kalman filter is as follows:

\[
\tilde{\beta}_{i,t|t-1} = B \tilde{\beta}_{i,t-1|t-1},
\]
\[
P_{i,t|t-1} = BP_{i,t-1|t-1}B' + \Sigma_{\tilde{\epsilon},i},
\]
\[
\eta_{i,t|t-1} = r_{i,t} - C_{i,t|t-1} - \tilde{X}'_{i,t|t-1} \tilde{\beta}_{i,t|t-1} - \rho_{\pi,i} \sigma_{e,i} \eta_{\pi,i,t|t-1} - \rho_{y,i} \sigma_{e,i} \eta_{y,i,t|t-1};
\]
\[
f_{i,t|t-1} = \tilde{X}'_{i,t|t-1} P_{i,t|t-1} \tilde{X}_{i,t|t-1},
\]
\[
\tilde{\beta}_{i,t|t} = \tilde{\beta}_{i,t|t-1} + P_{i,t|t-1} \tilde{X}_{i,t|t-1} f_{i,t|t-1}^{-1} \eta_{i,t|t-1},
\]
\[
P_{i,t|t} = P_{i,t|t-1} - P_{i,t|t-1} \tilde{X}_{i,t|t-1} f_{i,t|t-1}^{-1} \tilde{X}'_{i,t|t-1} P_{i,t|t-1},
\]
where \(\tilde{\beta}_{i,t|t-1} = E(\tilde{\beta}_{i,t} | I_{t-1})\), \(P_{i,t|t-1} = E(\tilde{\beta}_{i,t} - \tilde{\beta}_{i,t|t})^2\) and \(\eta_{i,t|t-1} = r_{i,t} - r_{i,t|t-1}\), \(f_{i,t|t-1} = E(\eta_{i,t|t-1})\).

Note that we use the unconditional mean of \(\sigma_{e,i,t}^2\), which can be calculated from equation (8), multiplied by \((1 - \rho_{\pi,i}^2 - \rho_{y,i}^2)\), as initial value for the last diagonal element of \(P_{i,t-1|t-1}\).

Then, we use this information to construct an estimate of \(\sigma_{e,i,t}^2\), which we have assumed that follows a GARCH(1,1) process as given by equation (8). This equation requires an estimate of \(e_{i,t-1}^2\), which we approximate as follows: \(E(e_{i,t-1}^2 | I_{t-1}) = E(e_{i,t-1} | I_{t-1})^2 + \text{Var}(e_{i,t-1} | I_{t-1})\), which given the transformation in equation (17), equals: \(E(e_{i,t-1}^2 | I_{t-1}) = (\rho_{\pi,i} \sigma_{e,i,t-1} \eta_{\pi,i,t|t-1} + \rho_{y,i} \sigma_{e,i,t-1} \eta_{y,i,t|t-1} + E(\omega_{i,t-1} | I_{t-1})^2 + \text{Var}(\omega_{i,t-1} | I_{t-1})\). We get \(E(\omega_{i,t-1} | I_{t-1})\) from the last component of \(\tilde{\beta}_{i,t|t-1}\) and \(\text{Var}(\omega_{i,t-1} | I_{t-1})\) from the last diagonal element of \(P_{i,t-1|t-1}\).

After estimating the hyperparameters, we run the Kalman filter for the second time in order to get an estimate for \(\beta_{i,t}\) from the first four rows of \(\tilde{\beta}_{i,t}\). The estimate of \(\beta_{i,t}\) is given correctly by iterating the Kalman filter above. However, given the two–step approach for solving the endogeneity issue and the usage of the control function, the standard errors of the coefficients face the problem of generated regressors (see Pagan, 1984). To address this issue, we run the following Kalman filter in order to get estimates for \(\text{Var}(\beta_{i,t} | I_{t-1})\) and \(\text{Var}(\beta_{i,t} | I_{t})\) from the first 4 × 4 block of \(P_{*_{i,t}|t-1}\) and \(P_{*_{i,t}|t}\):

\[
P_{*_{i,t}|t-1} = BP_{*_{i,t-1|t-1}}B' + \Sigma_{\tilde{\epsilon},i},
\]
\[
f_{*_{i,t|t-1}} = \tilde{X}'_{i,t|t-1} P_{*_{i,t|t-1}} \tilde{X}_{i,t|t-1} + (\rho_{\pi,i}^2 + \rho_{y,i}^2) \sigma_{e,i,t}^2,
\]
\[
P_{*_{i,t|t}} = P_{*_{i,t|t-1}} - P_{*_{i,t|t-1}} \tilde{X}_{i,t|t-1} f_{*_{i,t|t-1}}^{-1} \tilde{X}'_{i,t|t-1} P_{*_{i,t|t-1}}.
\]
While the true variance of $\beta_{i,t}$ is calculated using the variance of $e_{i,t}$, given by equation (16), $P_{i,t|t}$ is calculated using the variance of $\omega_{i,t}$, which is only a part of the variance of $e_{i,t}$. The above adjustment solves this issue.

3.2 Stage 2: Cointegration

To examine whether the conduct of monetary policy and, more importantly, its changes followed similar pattern across developed countries we concentrate on the analysis of the dependence between the estimated paths of coefficients. In particular, we want to know whether we can find a long-run cross-country relationship in the responses to inflation in the Taylor rules.

Our empirical methodology is dictated by the assumption of the unit root in the time-varying parameter representation of the monetary policy. Recall that the evolution of monetary policy is given by equation (5),

$$\beta_{j,i,t} = \beta_{j,i,t-1} + \epsilon_{j,i,t}, \quad \forall j, i,$$

so in order to study the existence of cross-country relationship between responses to inflation in Taylor rule, we have to take into account the non-stationary nature of the time-series in question.

Engle and Granger (1987) suggests that even though economic variables might be non-stationary there might exists a stable well-defined linear long-run relationship between these variables. If we define $\beta_{1,t} = [\beta_{1,1,t}, \beta_{1,2,t}, \ldots, \beta_{1,k,t}]'$ as a vector of responses to inflation in the Taylor rules in all the countries then given equation (5) each element of $\beta_{1,t}$ is $I(1)$. To study the existence of the commonality in the conduct of monetary policy we ask whether we can find a cointegrated relationship in $\beta_{1,t}$, i.e. a vector $\xi$ such that $\xi'\beta_{1,t}$ is $I(0)$. If this is the case, we can say that there exists a “long-run” relationship in how monetary policy is conducted across countries. If such relation cannot be found we conclude that there does not exist a long-run common pattern in monetary policy response to inflation.
3.2.1 Testing for cointegration

To test for cointegration consider the VAR($p$) representation of a $k \times 1$ vector $\beta_{1,t}$,

$$\beta_{1,t} = \Theta_1 \beta_{1,t-1} + \ldots + \Theta_p \beta_{1,t-p} + u_{1,t}, \quad (27)$$

and its vector error-correction form

$$\Delta \beta_{1,t} = \Gamma_1 \Delta \beta_{1,t-1} + \ldots + \Gamma_{p-1} \Delta \beta_{1,t-p+1} + \Pi \beta_{1,t-1} + u_{1,t}, \quad t = 1, \ldots, T, \quad (28)$$

where $\Delta$ denotes the difference operator; $\Gamma_i = \sum_{i=1}^p \Theta_\iota$ and $\Pi = \Theta_1 + \ldots + \Theta_p - I_k$ are $p \times p$ matrices of coefficients; and $u_{1,t}$ is an error vector with i.i.d. multivariate distribution with mean 0 and covariance matrix $\Omega$.\footnote{Note that we assume there is no constant nor trend in these specifications since they are not present in our series.} In our baseline specification we set $p = 1$ so equations (27) and (28) have the following form:

$$\beta_{1,t} = \Theta_1 \beta_{1,t-1} + u_{1,t}, \quad (29)$$
$$\Delta \beta_{1,t} = \Pi \beta_{1,t-1} + u_{1,t}, \quad (30)$$

where $\Pi = \Theta_1 - I_k$ and $k = 7$.

The hypothesis of cointegration can be stated in terms of the matrix $\Pi$. This matrix, which satisfies $\Pi \beta_{1,t-1} \sim I(0)$, can be written as:

$$\Pi = \alpha \xi', \quad (31)$$

where $\xi$ is $k \times r$ matrix of cointegrating vectors such that $\xi' \beta_{1,t} \sim I(0)$; $\alpha$ is $k \times r$ matrix of associated weights; and $r = \text{rank}(\Pi)$. If $r = 0$ then $\Pi = 0$ and there does not exist a linear combination of the elements of $\beta_{1,t}$ that is stationary, while if $r = k$, $\beta_{1,t}$ is stationary. When $0 < r < k$, there exists $r$ stationary linear combinations of the elements of $\beta_{1,t}$ and $k - r$ independent stochastic trends.

To test for the number of cointegrating relationships we follow the likelihood ratio
test presented by Johansen and Juselius (1990) and Johansen (1991).\textsuperscript{15} It can be shown that (i) the maximum likelihood estimate for $\xi$ equals the matrix with the $r$ eigenvectors corresponding to the $r$ largest eigenvalues, $\lambda_j$, of $\Pi$, and (ii) if the rank($\Pi$) = $r$ and $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ then for $k - r$ smallest eigenvalues of matrix $\Pi$, $\log(1 - \hat{\lambda}_j) = 0$, $j = r + 1, \ldots, k$. Then, the Trace statistics for the null hypothesis $r \leq r_0$ against the alternative $H_1 : r_0 \leq r \leq k$ equals:

$$\lambda_{\text{trace}}(r_0) = -T \sum_{j=r_0+1}^{k} \log(1 - \hat{\lambda}_j),$$

(32)

where $\hat{\lambda}_j$, $j = 1, \ldots, k$ denotes estimated eigenvalues of $\Pi$. The maximum eigenvalue test, on the other hand, tests $H_0 : r \leq r_0$ versus $H_1 : r = r_0 + 1$ using $\lambda_{\text{max}}$ statistics:

$$\lambda_{\text{max}} = -T \log(1 - \hat{\lambda}_{t_0+1}).$$

(33)

Both tests are performed sequentially for $r = 0, 1, \ldots, k - 1$ with the testing sequence terminating when the $H_0 : r \leq r_0$ is not rejected for the first time. When this happens, we conclude that there are $r_0$ cointegrating vectors.

Note that the determination of the order, $p$, of the VAR representation (27) is important for both tests due to bias-efficiency trade-off. In general, too low $p$, i.e. too few lags in the model leads to rejection of the null hypothesis too easily, while too high $p$ decreases the power of the tests. Taking into account how $\beta_{i,j,k}$’s are generated, in our baseline specification we will consider the case with $p = 1$. Also, we will allow for an intercept in the cointegrating relationship.

The asymptotic distributions of tests statistics (32) and (33) are multivariate extensions of the Dickey-Fuller distributions and depends on a number of non-stationary components under the null hypothesis as well as on the presence of constant and/or trend. However, since $\beta_{1,t}$ are generated regressors, instead of using standard critical values, e.g. presented in MacKinnon et al. (2000), we bootstrap cointegration regression and simulate the distribution of Johansen statistics.

Our approach is similar to that considered in Psaradakis (2001) and Chang et al. (2006)

\textsuperscript{15}See Hamilton (1994) and Johansen (1995) for details.
but, as far as we know, we are the first to adapt it to VAR-based cointegration test and the Johansen’s statistics. To obtain the bootstrap samples of $\beta_{1,t}^*$ for $\beta_{1,t}$ we proceed as follows.

*Step 1:* Estimate the regression (30) under the restriction of $h$ cointegrated vectors and compute the $\hat{u}_{1,t} = \Delta \beta_{1,t} - \hat{\alpha} (\hat{\xi}' \beta_{1,t} + \hat{c})$.

*Step 2:* Construct the VAR-sieve bootstrap for $\hat{u}_{1,t}$ by estimating an AR($q$) model for $\hat{u}_{1,t}$, 

$$\hat{u}_{1,t} = \hat{\Phi}_1 \hat{u}_{1,t-1} + \ldots + \hat{\Phi}_q \hat{u}_{1,t-q} + \hat{\nu}_{qt},$$

with the choice of the lag length $q$ based on BIC, and constructing centered fitted residuals

$$\nu_t^* = \hat{\nu}_{qt} - \frac{1}{n} \sum_{t=1}^{T} \hat{\nu}_{qt}.$$ 

*Step 3:* Repeat $B$ times

(a): Draw $T$ observation from $(\nu_t^*)$ and construct

$$u_{1,t}^* = \hat{\Phi}_1 u_{1,t-1}^* + \ldots + \hat{\Phi}_q u_{1,t-q}^* + \nu_t^*.$$ 

Set $u_{1,0}^* = u_{1,-1}^* = \ldots = u_{1,-q+1}^* = 0$.

(b): Construct $\{\beta_{1,t}^*\}$

$$\beta_{1,t}^* = \beta_{1,0}^* + \sum_{k=1}^{t} u_{1,k}^*.$$ 

(c): Use the bootstrap sample $\{\beta_{1,t}^*\}$ to estimate the VECM model (30) and compute the Johansen statistics, $\lambda_{\text{max}}$ and $\lambda_{\text{trace}}$.

(d): Go to (a).

*Step 5:* Test if the null hypothesis of $h$ cointegrating vectors can be rejected.

We repeat this process for every null hypothesis of $h$ cointegrating vectors.\footnote{When we apply this procedure to the data, we use $q = 2$ and $B = 999$.}

Once we determined the number of cointegrating vectors we can recover the estimates of the cointegrating vector $\xi$ and associated matrix of weights, $\alpha$. Since different combinations of $\xi$ and $\alpha$ gives $\Pi = \alpha \xi'$, that is $\xi$ and $\alpha$ are not uniquely identified, some normalization
assumptions on $\xi$ are necessary, as discussed in Section 5 that presents the results.

3.3 Stage 2: Principal Component Analysis

In this section we study the covariance structure of the shocks that affect the monetary policy feedback rules of the G7 countries. We do that in order to examine the possibility that common shocks are important, and might shadow the importance of common monetary policy response to inflation. Finding strong relationships in the covariance structure of the shocks of the Taylor rules, is interpreted as common shocks being important for governing the G7 countries behavior.

We are using the principal component analysis in order to simplify the study of the sources of variation of our targeted covariance matrix. Specifically, we examine if we can identify a few linear combinations of the G7 Taylor rules shocks series that we generated in the first stage (the estimates of the series of $\omega_{i,t}$’s in equation (25), with $i = 1, ..., 7$), in order to examine the main components of variation of the covariance matrix that relates these series. Let $\omega_t \equiv (\omega_{1,t}, ..., \omega_{7,t})'$ be the vector with the estimated shocks for the G7 countries at time $t$, and $\omega \equiv (\omega_{1}, ..., \omega_{7})'$ the matrix of the whole sample, with covariance matrix $\Sigma_{\omega}$. Through the principal component analysis we attempt to identify a few linear combinations of $\omega_i$, $i = 1, ..., 7$, in order to understand the structure of $\Sigma_{\omega}$.

We construct linear combinations $\varepsilon_{\phi,t}$ of the vector $\omega_t$ by defining the real vector $\kappa_{\phi} \equiv (\kappa_{\phi,1}, ..., \kappa_{\phi,7})'$ with $\phi = 1, ..., 7$, as follows:

$$\varepsilon_{\phi,t} = \kappa_{\phi}' \omega_t = \sum_{i=1}^{7} \kappa_{\phi,i} \omega_{i,t}.$$

Let now $\varepsilon_{\phi}$ be a linear combination of the whole time series for our sample, which receives country weights $\kappa_{\phi} \equiv (\kappa_{\phi,1}, ..., \kappa_{\phi,7})'$ for every $t$. We can interpret $\varepsilon_{\phi}$ as the linear combination of shocks hitting each of the G7 countries, which assigns weight $\kappa_{\phi,i}$ to the shock hitting the $i$th country, $i = 1, ..., 7$.

We can derive the variance of the series of the linear combination $\varepsilon_{\phi}$ in relation to the original matrix $\Sigma_{\omega}$ as $\text{Var}(\varepsilon_{\phi}) = \kappa_{\phi}' \Sigma_{\omega} \kappa_{\phi}$ for $\phi = 1, ..., 7$. Similarly is derived the covariance between two series of linear combinations $\phi$ and $\varphi$ as $\text{Cov}(\varepsilon_{\phi}, \varepsilon_{\varphi}) = \kappa_{\phi}' \Sigma_{\omega} \kappa_{\varphi}$. 


for $\phi, \varphi = 1, \ldots, 7$.

The principal component analysis identifies vectors $\kappa_\phi$ so that the associated linear combination to have the maximum variance. Specifically, the first principal component of $\omega$ defined as $\varepsilon_1 = \kappa_1' \omega$, maximizes $\text{Var}(\varepsilon_1)$ given that $\kappa_1' \kappa_1 = 1$.\footnote{We standardize for convenience the vector.} Similarly, the second principal component of $\omega$ defined as $\varepsilon_2 = \kappa_2' \omega$, maximizes $\text{Var}(\varepsilon_2)$ given that $\kappa_2' \kappa_2 = 1$ and that $\text{Cov}(\varepsilon_1, \varepsilon_2) = 0$. The last principal component of $\omega$ defined as $\varepsilon_7 = \kappa_7' \omega$, maximizes $\text{Var}(\varepsilon_7)$ given that $\kappa_7' \kappa_7 = 1$ and that $\text{Cov}(\varepsilon_7, \varepsilon_o) = 0$, for $o = 1, \ldots, 6$.

It can be shown (see Tsay, 2010) that the proportion of the variance of $\omega$ explained by the $\phi$th principal component is the proportion of the value of the $\phi$th eigenvalue of $\Sigma_\omega$ to the sum of all eigenvalues of $\Sigma_\omega$. In Section 5.3 we compute the fraction of variation each principal component contributes towards. Having a small number of principal components contributing towards a large percentage of the variation of $\omega$ makes the simplification of identifying the principal components useful.

4 Data

We use quarterly data for Canada, France, Germany, Italy, Japan, UK and US.\footnote{Series that are available monthly are converted to quarterly using the value of the first month of each quarter. These series are the CPI for Germany and for US.} For the European countries the sample stops on 1998:4. Data are taken from the Datastream database. We have data for short term interest rates, inflation and real Gross Domestic Product.\footnote{For policy instruments we use: For Canada the three-month treasury bill rate (similarly to Nelson, 2005b); for France the call money rate; for Germany the interbank money rate (similarly to Bullard and Singh, 2008); for Italy the money market rate (which is the three-month interbank rate before February 1990, after which becomes the daily rate); for Japan the money market rate for overnight loans (similarly to Bullard and Singh, 2008); for UK the treasury bill rate (similarly to DiCecio and Nelson, 2009; for the US the Federal Funds Rate (similarly to Bullard and Singh, 2008).} We construct measures of interest rate spread and money growth to use as instrumental variables.

We use annualized inflation rate constructed from the consumer price index.\footnote{For France and Germany we construct real GDP using the nominal one and CPI. For Japan, we used the nominal GDP and the GDP deflator.} The consumer price indexes that were not originally seasonally adjusted (Canada, France, Italy, Japan, UK), were converted to seasonally adjusted series using the year-to-year percentage change in the consumer price index.

\footnote{Annualized inflation rate at time $t = \left(\frac{\text{CPI}_t - \text{CPI}_{t-1}}{\text{CPI}_{t-1}}\right) \times 400\%$.}

\[ \text{Annualized inflation rate at time } t = \left(\frac{\text{CPI}_t - \text{CPI}_{t-1}}{\text{CPI}_{t-1}}\right) \times 400\%. \]
difference adjustment.

We also use annualized money growth, constructed from a money measure.\textsuperscript{22} We seasonally adjusted the money measures that were not originally adjusted (Italy, UK) using the year-to-year percentage difference adjustment. For most of the countries we used as measure of money the M2. For the UK we use Money and Quasi-money, which is the closest available series to the definition of M2. For Canada the M2 series is too short (starts in 1968) and we used the M1 instead. Also, the M2 series for Canada ends in 2006, making the sample slightly shorter. For France and Italy we did not use a money aggregate because the M2 starts in 1977 and end of 1975, respectively. For Italy we did not use interest rate spread either. The M2 series for Japan finishes at the beginning of 2008.

We construct output gap as the deviation of seasonally adjusted log real GDP from the fitted quadratic trend.\textsuperscript{23} We finally construct interest rate spreads using the difference of a short term interest rate from a long term one.\textsuperscript{24}


\textsuperscript{22}Annualized money growth rate at time $t = \left(\frac{M_t - M_{t-1}}{M_{t-1}}\right)400\%$.

\textsuperscript{23}For log real GDP being $gdp_t$, then we fit $gdp_t = j_0 + j_1 t + j_2 t^2 + s_t$, and construct the gap as $y_t = gdp - \bar{gdp}$, for $gdp = j_0 + j_1 t + j_2 t^2$, for the estimates of $j_1, j_2$ computed above, and $t$ being the time trend.

\textsuperscript{24}For constructing the spread we use: For Canada the more than 10-year government bond yield and the 3 to 5-year government bond yield (the Treasury bill serves as policy instrument); for France the more than 5-year government bond yield, and the Treasury bill rate (the 12-month treasury bills until June 1989 and the 3-month Treasury bills after that); for Germany the long term government bond yield and the call money market rate (the Treasury bill rate becomes available only after the third quarter of 1975); for Italy we use the the long term government bond yield and the discount rate (the Treasury bill rate starts much later, in 1977); for Japan the long term government bond yield and the Japan Treasury bill rate; for UK we use the 20-year government bond yield and the 5-year government bond yield (the Treasury bill serves as policy instrument); for the US the 10-year US government bond yield and 3-month US Treasury bill rate.
5 Empirical results

We first report results from the estimations of individual countries’ Taylor rules. Next we present results of tests that measure the commonality between coefficients across countries and tests that analyze the covariance structure of the shocks. Finally, we return to the individual countries’ estimations, taking a closer look on the results.

5.1 Taylor rule estimation results

We report the estimated series of the the time varying coefficients for each country considered, with 90% confidence bands. Figures 3 through 9 depict that in all of the G7 countries’ interest rate response to inflation is low during the 1970s (often at its lowest level), increases in the 1980s and then, for the countries with longer sample we see that decreases again after the mid 1990s. Our results suggest that responding to inflation becomes less important during the 1990s, when inflation decreases and gets under control.25

France’s and Italy’s response to inflation as shown in Figures 4 and 6, is low during the 1970s. The coefficients of interest increase and become higher than 1 only in the early 1980s. For France, this coefficient becomes statistically larger than 1 for some time during the 1980s, although for Italy it is never statistically larger than 1. Given the introduction of the Euro in 1999 we do not have the sample to fully isolate the importance of inflation in the Taylor rule during the later period. However, the change in importance assigned to fighting inflation, from weak during the 1970s to stronger during the 1980s, is apparent.

In Figure 8 we see that during the 1970s the coefficient of inflation for the UK is almost always significantly lower than 1. Similarly to the experience of France, the estimated coefficient increases in the 1980s, and becomes significantly higher than 1. Having available longer sample than what we have for Italy and France, we observe that the response becomes less ”aggressive” during the 1990s.

For the US we have the longest sample available compared to the rest of the countries. Thus, we can see more variation in the interest rate response to inflation. In Figure 9 we see strong response to inflation in the mid-1960s, which however becomes low during the

25 At the same time though, as inflation becomes more stable (see Figure 1), is harder to identify the response to inflation. This is why often the confidence bands become larger too.
1970s. The response becomes again stronger and greater than 1 during the 1980s. Although not statistically different than 1, it decreases during the 1990s and it even becomes less than 1 after the mid-1990s. The decreased response to inflation during the 1990s has been previously reported by Kim et al. (2006) who similarly to us, find three periods in monetary policy conduct, the mid 1960s to late 1970s, the 1980s, and the 1990s and afterwards.

In Figure 5 we see that the response to inflation for Germany is low during the 1970s and becomes statistically significantly higher than 1 at the beginning of the 1980s. It also stays above 1 afterwards, although it is not statistically different than 1 in the later period. Contrary to the view that Germany had traditionally responded strongly to inflation and this is why did not get high inflation during the 1970s, our estimation points out that the response to inflation was at its lowest during the 1970s. Earlier work has also documented that Germany accommodated inflation similarly to the US for the Great Inflation period (Clarida and Gertler, 1997). We find that the difference between Germany and the US refers mostly to the response to inflation during the 1990s, rather to during the 1970s. During the 1990s Germany remained aggressive in its response to inflation contrary to other developed countries.

Similarly to the US, a long sample is available for Canada. The response to inflation, as seen in Figure 3, is above 1 in the late 1960s and becomes lower than 1 during the 1970s. It increases to levels above 1 and remains high and statistically higher than 1 until the mid 1990s. This description matches the narrative evidence of Nelson (2005b). Canada’s response to inflation starts decreasing after the mid 1990s, similarly to the US. This indicates that central banks might be placing lower importance responding to inflation in recent years.

Finally, Figure 7 shows Japan’s time-varying response to inflation in the Taylor-rule. Our estimated series show that Japan’s response to inflation is low during the 1970s; however, it started to respond strongly to inflation slightly earlier than other countries, consistently with the narrative evidence of Nelson (2007). It remained above 1, although not statistically significant, until the beginning of the 1990s after which it decreased to values below 1. Our findings for Japan during the 1990s are not surprising given that during this period Japan hits the zero lower bound for the nominal interest rate.
Figures 10 to 13 show the evolution of monetary policy’s response to inflation, output gap, degree of interest rate smoothing and the intercept, represented by the time varying coefficients of the Taylor rules for all the countries considered. Figure 10 depicts the response to inflation. From there we see the similar way that monetary authorities are changing their response to inflation over time, and especially in the period right before, during, and right after the Great Inflation era. We see that all countries respond weakly to inflation during the 1970s and much stronger during the 1980s and mid 1990s. This commonality motivates us to investigate the possibility of a long run relationship among the series during and right after the Great Inflation period, presenting a cointegration test in the the next section.

Figure 12 shows that most of the countries realized the importance of smoothing after the mid 1980s when they started taking seriously into consideration the interest rate of the previous periods, in deciding policy in the current period. However, we do not observe the strong commonality that we observe for the response to inflation. The same is true for the response to output gap in Figure 11.

5.2 Cointegration analysis results

In this subsection we present the results on the existence of cointegrating relationship between variables. Using estimated series of the time-varying Taylor rule we conduct a cointegration analysis as described in Section 3.2.1. We look for a long-run relationship between the monetary policy responses to inflation in various countries, i.e., among the series of $\beta_{1,i,t}$ for the various countries $i$.

First, to verify that all series are I(1) we apply a a Dickey-Fuller unit root test to each individual component of that vector. For each of the series we cannot reject the hypothesis of unit root at the 10% significance level. In case of Japan we only marginally accept null hypothesis of unit root at 10% for ADF regression with 2 lags.

Next, we employ the Johansen procedure to find the number of cointegrating vectors. In our baseline specification the test clearly indicates the existence of long-run relationship.

\[\text{While we postulated a unit root in the } \beta \text{ specifying the nature of the time variation of the coefficients, such specification is general enough that if true } \beta \text{ was stationary, the estimated coefficients would be also stationary. We apply Augmented Dickey-Fuller tests to confirm that there are, indeed, unit roots in the estimated series and our subsequent cointegration analysis is valid.}\]

\[\text{In case of Japan we only marginally accept null hypothesis of unit root at 10\% for ADF regression with 2 lags.}\]
In Table 2 we present the results for both $\lambda_{trace}$ and $\lambda_{max}$ tests for all seven countries.

For both tests, under the $H_0$ the rank of $\xi$ is at least $r$ whereas under the alternative hypothesis, $H_1$, the rank is either strictly higher or equals $r + 1$, respectively. At the 5% level, both tests reject the null hypothesis of no cointegration. Both tests, however, do not reject the hypothesis of $r$ being less than or equal to 1. We stress that since we estimate cointegration vectors using generated regressors, the asymptotic distributions of the Johansen statistics may not be valid. In particular, using the $p-values$ based on MacKinon et al (2000) we can reject the null hypothesis of 3 cointegrating vectors at 5% level and 4 cointegrating vectors at 10% level. To correct for that we bootstrap the cointegrating regression and compute the relevant statistics as described in Section 3.2.1. We repeat it for each possible null hypothesis.

The results are stronger when we consider cointegrating regression for six countries only. In Table 3 we show the results for $\lambda_{trace}$ and $\lambda_{max}$ tests for the case in which exclude Japan. Given that the recession in Japan started in the early 1990s lasted past 1998:4 and brought both zero lower bound and unconventional monetary policy, the hypothesis that the monetary policy in Japan evolved differently from the other G7 countries is justified and confirmed by our estimates. Once we exclude Japan, we find that we can reject the null hypothesis of 4 cointegrating vectors at 10% level and 3 cointegrating vectors at 5% level. These results strongly indicate the existence of 1, or at most 2, stochastic trends that describe the behavior of responses of monetary policy to inflation in G7 countries.

We use this result to determine whether there exist a common pattern in the conduct of monetary policy. Recall that the implied vector error correction model is

$$\Delta \beta_{1,t} = \alpha \xi' \beta_{1,t-1} + u_{1,t}. \quad (34)$$

The above equation describes how policy parameters behave in the short-run, consistent with the long-run cointegrating relationship. If there is no such relationship then current changes in the stance of monetary policy would be unrelated. We interpret the existence of that many cointegrating vectors as support for the existence of a long-run pattern in how the monetary policy is conducted in developed countries. While the changes in monetary policy may vary over time in different countries there exists, nevertheless, the commonality
in the monetary policy implementation across countries in our sample.

5.3 Principal Component Analysis results

The cointegration test results above, show us that there is a long run relationship in monetary authorities’ response to inflation. We now perform principal component analysis which allows us to study the covariance structure of the Taylor rules’ innovations of the G7 countries. We interpret the cointegrated relationship of the coefficients as common policy, and the existence of strong principal components in the residuals, as common shocks.

Specifically, finding a small number of components driving a large percentage of variation would indicate that there are strong commonalities in the shocks structure of the G7 countries. On the contrary, finding that a large number of components is required in order to explain variation, would indicate that there is no strong commonality in the shocks’ structure of the G7 countries. We perform the principal component analysis as summarized in Section 3.3. Table 4 reports the results for applying this analysis in the series of $\omega_{i,t}$’s of equation (25), for $i$ being Canada, France, Germany, Italy, Japan, the UK, and the US. We see that the first component explains only 25.7%, although we need four components in order to explain 73.4% of the residuals’ variability.\footnote{As a comparison, Tsay (2010) page 487 finds in a different application, that the first component explains 53.5% of total volatility which is much larger than the volatility our first principal component explains. In addition, Tsay (2010) uses two components out of five series (which is 40% of available components) for explaining cumulatively 74% of total volatility. In our case, we need almost 60% of available components in order to explain almost the same fraction of volatility.} The results indicate that the shocks structure does not have a small number of principle components in order to approximate in a satisfactory extent the variability of the Taylor rules’ residuals in various countries, with information of common components. Then, the variability of the G7 shocks seem to originate from multiple sources, with each having limited common effect. This observation decreases the importance of common shocks in our analysis.

Related work (Chatterjee, 2010) finds that individual countries’ variation in the shocks of the Taylor rules is explained in a large extent by a common component, the G7 component, for the period 1980-2009, but also for the sub-period 1980-1987. These results do not take into account that the parameters of the the Taylor rule might be changing over time, as we do in our analysis. Also, they do not focus on the Great Inflation period, which is our
work’s period of interest. Our analysis allows to study both the commonality in the shocks structure, but also the commonality in the monetary authorities’ responses to inflation, focusing on the Great Inflation era. And we show, using the Taylor rule estimations and the cointegration tests, that there is an important common component in monetary policies response to inflation. Yet, the importance of common shocks is limited.

5.4 Inflation Adjusted Responses

In this section we get a closer look in the countries’ inflation responses. We examine whether countries started responding stronger to inflation within a few quarters of experiencing their highest inflation rates. If this was so, then it could be true that the G7 countries experienced very high inflation at similar to each other period, and each one, independently, might have decided to combat inflation. This situation weakens our hypothesis about common monetary policy responses.

To study this hypothesis, we stack the series of the countries’ responses to inflation (the $\beta_1$ coefficients), with reference point the quarter in which they exhibit maximum inflation.\(^{29}\) The results can be seen in Figure 14. The US had its maximum inflation of 15.7% in 1980:1. This date corresponds to zero in the horizontal axis. Similarly, at zero, the UK had its maximum inflation of 26.9% in 1975:3; Japan had its maximum inflation of 23.6% in 1974:3; Canada had its maximum inflation of 12.7% in 1981:3; Italy had its maximum inflation of 22.3% in 1977:1; France had its maximum inflation of 14.3% in 1981:4; Germany had its maximum inflation of 8.2% in 1981:1. This calculation helps us compare the countries’ responses to inflation, right after they had their maximum inflation incident, which is the point zero in the horizontal axis of Figure 14.

As we see, most of the countries change their behavior from less to more aggressive towards inflation, within a few quarters from the maximum inflation date. Specifically, the US, Canada, Germany and France, switch within a few quarters from their maximum inflation incidence, from responding less than one to one, to responding more than one to one to inflation. After doing so, inflation starts decreasing. This result shows that monetary policy had been important in sustaining the Great Inflation of the 1970s, and

\(^{29}\)We use the maximum inflation rate date, within the sample of inflation response for each country.
also had been important in reducing it.\textsuperscript{30} Note that these four countries experienced their maximum inflation in close proximity of time.

On the contrary, Figure 14 shows that the UK, Japan and Italy did not become aggressive towards inflation, i.e., they did not switch to more than one to one response, in a period close to their highest inflation incidence. Specifically, the UK, Japan and Italy start fighting inflation, but their response does not switch to more than one to one until five, four and a half, and six years, respectively, after their maximum inflation date. This corresponds to switching towards aggressive policy date, on 1980:1 for the UK, 1979:1 for Japan and 1982:4 for Italy. Also note that all three countries experienced their maximum inflation before the late 1970s.

We believe that the reason why some countries delayed to switch to aggressive monetary policy is that from the mid-1960s and until the mid-1970s, there were bad ideas about what monetary policy can and ought to do in order to fight inflation (see DeLong (1997), Sargent (1999), Romer and Romer (2002, 2004), Cogley and Sargent (2005), Nelson (2005a,b, 2007, 2008), Romer (2005), Sargent et al. (2006), DiCecio and Nelson (2009)). However, during the late 1970s and at the beginning of the 1980s, monetary authorities realized that they need to become aggressive in order to fight inflation. That is why UK, Japan and Italy, although they experienced their maximum inflation much earlier, they switched to aggressive monetary policy only after the late 1970s.

\section{Conclusions}

We employ a multi-country analysis of monetary policy using a time-varying, forward-looking Taylor rule. We find that in the G7 countries the interest rate responses to inflation follow similar pattern: all countries react mildly during the 1970s and aggressively from the 1980s and up to the mid–1990s. In addition, our analysis of cointegration indicates that whereas the changes in monetary policy may vary over time in individual countries, there exists, nevertheless, commonality in the monetary policy implementation across the G7 countries during the Great Inflation period. On the contrary, the principal component

\textsuperscript{30}Note that Germany’s monetary authorities seem to have started the attempts to fight inflation about two years earlier, although after experiencing their highest inflation, their attempts became even more aggressive.
analysis reveals many sources of variation rather than one strong common component in
the shocks’s structure. We interpret the common pattern revealed by the Taylor rule
estimations and the findings of cointegrated relationship of the coefficients, as evidence of
common policy; we interpret the lack of strong principal components in the residuals as
evidence of limited common shocks effects.

Future work mandates addressing Orphanides (2001, 2002, 2004)’s critique, which em-
phasizes the effects that the potential output mis-measurement has in Taylor rule estima-
tions. This explanation could be robust with the international evidence if for example we
consider that the knowledge about output gap calculations was shared among economists
around the world, and they might have all been using methods that led to mis-measurement
errors. Kim et al. (2006) take this critique into account using a time varying parameter
model with real time data for the US. Yet, their work still concludes that monetary author-
ities responded milder to inflation during the 1970s and stronger after that. Our analysis
provides evidence in favor of common monetary policy rather than common shocks as an
explanation of the common inflation patterns in the G7 countries. However, more work on
the topic is required for addressing also the above issue.
References


Figure 3: Estimated time varying response to inflation in Canada (solid line), and 90% confidence bands (dotted lines). The dashed line indicates the value 1.
Figure 4: Estimated time varying response to inflation in France (solid line), and 90% confidence bands (doted lines). The dashed line indicates the value 1.

Figure 5: Estimated time varying response to inflation in Germany (solid line), and 90% confidence bands (doted lines). The dashed line indicates the value 1.
Figure 6: Estimated time varying response to inflation in Italy (solid line), and 90% confidence bands (doted lines). The dashed line indicates the value 1.

Figure 7: Estimated time varying response to inflation in Japan (solid line), and 90% confidence bands (doted lines). The dashed line indicates the value 1.
Figure 8: Estimated time varying response to inflation in UK (solid line), and 90% confidence bands (doted lines). The dashed line indicates the value 1.

Figure 9: Estimated time varying response to inflation in US (solid line), and 90% confidence bands (doted lines). The dashed line indicates the value 1.
Figure 10: Estimated time varying response to expected inflation in Canada, France, Germany, Italy, Japan, US and UK

Figure 11: Estimated time varying response to expected output gap in Canada, France, Germany, Italy, Japan, US and UK
Figure 12: Estimated time varying response to smoothing parameter in Canada, France, Germany, Italy, Japan, US and UK

Figure 13: Estimated time varying intercept in Canada, France, Germany, Italy, Japan, US and UK
Figure 14: Estimated time varying response to expected inflation in Canada, France, Germany, Italy, Japan, US and UK, adjusted for maximum inflation dates. The horizontal axis corresponds to US dates.
Table 1: Augmented Dickey-Fuller (unit root) test of $\beta_1$

<table>
<thead>
<tr>
<th>Country</th>
<th>$t$-statistics</th>
<th>$p$-value $AIC$</th>
<th>$p$-value $AIC$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{1,CAN}$</td>
<td>-0.423</td>
<td>0.528</td>
<td>0.622</td>
</tr>
<tr>
<td>$\beta_{1,FRA}$</td>
<td>-0.273</td>
<td>0.586</td>
<td>0.241</td>
</tr>
<tr>
<td>$\beta_{1,GER}$</td>
<td>-1.093</td>
<td>0.247</td>
<td>0.427</td>
</tr>
<tr>
<td>$\beta_{1,ITA}$</td>
<td>0.791</td>
<td>0.882</td>
<td>0.882</td>
</tr>
<tr>
<td>$\beta_{1,JAP}$</td>
<td>-1.170</td>
<td>0.219</td>
<td>0.105</td>
</tr>
<tr>
<td>$\beta_{1,UK}$</td>
<td>-0.131</td>
<td>0.637</td>
<td>0.681</td>
</tr>
<tr>
<td>$\beta_{1,US}$</td>
<td>-0.721</td>
<td>0.403</td>
<td>0.447</td>
</tr>
</tbody>
</table>

Note: $t$-statistics computed for 0 lags. $p$-value $AIC$ computed for 0 lags and $p$-value $AIC$ computed for optimal number lags according to AIC, both based on MacKinnon (1996).

Table 2: Johansen’s trace and maximum eigenvalue tests for cointegration

<table>
<thead>
<tr>
<th>$H_0$</th>
<th>$H_1^{trace}$</th>
<th>$\lambda_{trace}$</th>
<th>$p$-value $trace$</th>
<th>$H_1^{max}$</th>
<th>$\lambda_{max}$</th>
<th>$p$-value $max$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 0$</td>
<td>$r \geq 0$</td>
<td>253.38***</td>
<td>0.002</td>
<td>$r = 1$</td>
<td>96.22***</td>
<td>0.002</td>
</tr>
<tr>
<td>$r \leq 1$</td>
<td>$r \geq 2$</td>
<td>157.16</td>
<td>0.476</td>
<td>$r = 2$</td>
<td>64.46</td>
<td>0.290</td>
</tr>
<tr>
<td>$r \leq 2$</td>
<td>$r \geq 3$</td>
<td>92.71</td>
<td>0.356</td>
<td>$r = 3$</td>
<td>41.64</td>
<td>0.252</td>
</tr>
<tr>
<td>$r \leq 3$</td>
<td>$r \geq 4$</td>
<td>51.06</td>
<td>0.444</td>
<td>$r = 4$</td>
<td>29.37</td>
<td>0.144</td>
</tr>
<tr>
<td>$r \leq 4$</td>
<td>$r \geq 5$</td>
<td>21.69</td>
<td>0.834</td>
<td>$r = 5$</td>
<td>10.15</td>
<td>0.926</td>
</tr>
<tr>
<td>$r \leq 5$</td>
<td>$r \geq 6$</td>
<td>11.54</td>
<td>0.572</td>
<td>$r = 6$</td>
<td>7.46</td>
<td>0.619</td>
</tr>
<tr>
<td>$r \leq 6$</td>
<td>$r = 7$</td>
<td>4.08</td>
<td></td>
<td></td>
<td></td>
<td>0.6974</td>
</tr>
</tbody>
</table>

Note: $p$-values based on bootstrapped regression. Specification: VAR(1) in levels, no drift, no trend trend. Constant in cointegrating relationship.

Table 3: Johansen’s trace and maximum eigenvalue tests for cointegration for 6 countries, no Japan.

<table>
<thead>
<tr>
<th>$H_0$</th>
<th>$H_1^{trace}$</th>
<th>$\lambda_{trace}$</th>
<th>$p$-value $trace$</th>
<th>$H_1^{max}$</th>
<th>$\lambda_{max}$</th>
<th>$p$-value $max$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 0$</td>
<td>$r \geq 0$</td>
<td>197.21***</td>
<td>0.002</td>
<td>$r = 1$</td>
<td>83.42***</td>
<td>0.008</td>
</tr>
<tr>
<td>$r \leq 1$</td>
<td>$r \geq 2$</td>
<td>113.78</td>
<td>0.193</td>
<td>$r = 2$</td>
<td>40.95</td>
<td>0.645</td>
</tr>
<tr>
<td>$r \leq 2$</td>
<td>$r \geq 3$</td>
<td>72.83**</td>
<td>0.054</td>
<td>$r = 3$</td>
<td>32.90</td>
<td>0.25</td>
</tr>
<tr>
<td>$r \leq 3$</td>
<td>$r \geq 4$</td>
<td>39.93**</td>
<td>0.019</td>
<td>$r = 4$</td>
<td>24.56**</td>
<td>0.023</td>
</tr>
<tr>
<td>$r \leq 4$</td>
<td>$r \geq 5$</td>
<td>15.37*</td>
<td>0.052</td>
<td>$r = 5$</td>
<td>9.19</td>
<td>0.182</td>
</tr>
<tr>
<td>$r \leq 5$</td>
<td>$r \geq 6$</td>
<td>6.18</td>
<td></td>
<td></td>
<td></td>
<td>6.18</td>
</tr>
</tbody>
</table>

Note: $p$-values based on bootstrapped regression. Specification: VAR(1) in levels, no drift, no trend trend. Constant in cointegrating relationship.
Table 4: Principle Component Analysis in the series of residuals ($\omega_{i,t}$) of the Taylor rule equation (25), for $i$ being Canada, France, Germany, Italy, Japan, US and UK, for the sample 1976:2 to 1998:4. Included observations: 91 after adjustments; Balanced sample (listwise missing value deletion); Computed using ordinary correlations.

<table>
<thead>
<tr>
<th></th>
<th>Comp 1</th>
<th>Comp 2</th>
<th>Comp 3</th>
<th>Comp 4</th>
<th>Comp 5</th>
<th>Comp 6</th>
<th>Comp 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalue</td>
<td>1.803</td>
<td>1.447</td>
<td>1.040</td>
<td>0.854</td>
<td>0.741</td>
<td>0.622</td>
<td>0.494</td>
</tr>
<tr>
<td>Proportion</td>
<td>0.257</td>
<td>0.206</td>
<td>0.149</td>
<td>0.122</td>
<td>0.106</td>
<td>0.089</td>
<td>0.071</td>
</tr>
<tr>
<td>Cumulative Proportion</td>
<td>0.257</td>
<td>0.463</td>
<td>0.612</td>
<td>0.734</td>
<td>0.840</td>
<td>0.929</td>
<td>1</td>
</tr>
</tbody>
</table>