BANK OF GREECE

EUROSYSTEM

Working Paper

Quantifying Federal Reserve credibility

Stephen Hall
George Tavlas

340

APRIL 2025

BANK OF GREECE
Economic Analysis and Research Department – Special Studies Division
21, E. Venizelos Avenue
GR-102 50 Athens
Tel: +30210-320 3610
Fax: +30210-320 2432

www.bankofgreece.gr

Published by the Bank of Greece, Athens, Greece

ISSN: 2654-1912 (online)
DOI: https://doi.org/10.52903/wp2025340

QUANTIFYING FEDERAL RESERVE CREDIBILITY

Stephen G. Hall

Department of Economics, Leicester University, Bank of Greece, and University of Pretoria

George S. Tavlas

Bank of Greece and the Hoover Institution, Stanford University

ABSTRACT

We derive an index that quantifies the Federal Reserve's credibility from 1965 until 2024. The credibility measure is derived by using the Kalman filter to extract an unobserved component from data, the movements of which are affected by central-bank credibility. We extend previous work using the Kalman filter in that we standardize the variables thought to affect credibility so that they have zero mean and unit variance. Consequently, there is no need to estimate parameters in the Kalman filter measurement equations. The credibility index is used to identify seven historical episodes during which the degree of credibility differed.

Keywords: central bank credibility, Federal Reserve, Kalman filter, unobserved component

JEL Classification: C32, E52, E58

Acknowledgments: We thank Elisavet Bosdelekidou and Maria Monopoli for research assistance.

Correspondence:
George Tavlas
Bank of Greece,
21 E. Venizelos Ave., Athens, 102 50, Greece
Tel. no. +30 210 320 2370
email: gtavlas@bankofgreece.gr

1. Introduction

Beginning with the taming of inflation in the first half of the 1980s by the Federal Reserve under the chairmanship of Paul Volcker, credibility – typically defined as the confidence that the private sector has in a central bank's ability to deliver its inflation objective -- has assumed a paramount role as a determinant of the effectiveness of central-bank policies (*e.g.*, Yellen, 2006; Svensson, 2008, 321; Bordo and Siklos, 2015, 35).[1] Volcker became Fed chair in October 1979. At that time, CPI inflation was over 12 percent and the 10-year Treasury note was 10.7 percent. Five years later, in October 1984, the 10-year note stood at 12.4 percent although inflation had fallen to 4.3 percent. The Fed had not yet established that it was a credible central bank and, thus, inflation expectations (embedded in the 10-year note) remained at high levels. With the fall in inflation to low single digit levels from the mid-1980s until 2020, credibility came to be seen as important to policymakers because, in contrast to the situation during the first half of the 1980s, the acquisition of credibility was seen as reducing the costs of disinflation. Correspondingly, the acquisition of credibility was viewed as important in reducing the cost of keeping inflation down once it was low, while it helped garner support for central bank independence (Blinder, 2000; Ehrmann, 2024).[2]

Although the ideas of what credibility is and why it is important have gained widespread acceptance, the ability to measure credibility has been a different matter. Because credibility is unobserved, there is no accepted way of measuring credibility, and only a few studies have attempted to quantify the concept. As we discuss below, those studies share a common underpinning: they rely on the basic idea that credibility depends on the distance between a central bank's inflation objectives, whether explicit or implied, and either (1) a measure of the expectation that the central bank will achieve that objective or (2) the actual inflation outcome.

In this paper, we develop a different way of measuring credibility. We begin with the observation that there are many macroeconomic variables – not just inflation expectations and/or actual inflation – that are affected by a central bank's credibility.

[1] The above definition of credibility pertains to a regime under which the central bank targets the inflation rate. Under different regimes, the definition of credibility would also be different. For example, under the gold standard, credibility depended on a commitment to a fixed exchange rate parity (Bordo and Kydland, 1995, 425).

[2] As Ermann (2024, 2) points out, credibility is often used interchangeably with trust and reputation.

These variables include short-term and long-term interest rates, real output growth, and money growth. For example, if an identical inflationary shock (as in the case of the Covid shock) affects two regions, one with a credible central bank and the other with a central bank that lacks credibility, the former central bank would likely have to raise interest rates less, and undergo reduced unemployment, to achieve a given rate of inflation than the other central bank. The point of departure of this paper is that, if we could use a group of variables affected by central bank policies to extract an unobserved common component that affects each of the variables, the unobserved component can be identified as credibility. To be sure, that unobserved component could be called sunspots. But given that we are concerned with a group of variables directly affected by a central bank's policies – including the private sector's trust in those policies – we believe that the unobserved component is more likely to represent credibility.

How can the unobserved component be extracted from the data? We extend a methodology used for constructing measures of underlying – but unobserved – economic activity developed by Stock and Watson (1991) and Garratt and Hall (1996). Those authors developed a framework under which they constructed an optimal estimate of economic activity from a set of observable variables that were taken to measure underlying economic activity. Stock and Watson (1991) and Garrett and Hall (1996) noted that, while there are many macroeconomic series that measure the level of economic activity, they are all subject to distortion and measurement error. Those studies used the Kalman filter to extract a measure of economic activity which represented all the co-movements of the observable series. We follow a similar procedure, but in our case, that unobserved component is credibility because our data are directly affected by the central bank. Our extension to previous studies using the Kalman filter lies in standardizing all the observed variables so that they have a zero mean and unit variance. The standardization removes the need to estimate any parameters in the measurement equations in terms of both coefficients and variances. The more variables that are impacted by credibility, the better our extracted measure of credibility. For example, we know that energy prices affect inflation. If we were to use consumer price inflation as the only variable in our information set, we would have difficulty in determining that it was credibility, and not oil prices themselves, that affects inflation. If we add other variables to our information set, such as short-term and long-term interest rates, direct measures of inflation expectations, and M2 growth, which are less likely to be affected by oil prices than is inflation, we will be in a better position to claim that the underlying factor affecting all the variables is credibility.

The remainder of the paper consists of four sections. Section 2 presents a review of previous studies that have attempted to quantify credibility. Section 3 describes our empirical approach of extracting a measure of the Fed's credibility from a group of variables affected by that central bank's policies. That section also describes the variables used. Section 4 presents the results of our credibility measure. Section 5 concludes.

2. Literature review

A substantial literature has emerged in recent years that attempts to quantify central bank credibility. By-and-large, this literature aims to operationalize Blinder's definition of credibility: "A central bank is credible if people believe it will do what it says" (2000, 1422). Initial studies used a measure of central bank independence as a proxy for credibility – see, for example, Cukierman and Meltzer (1986) and Faust and Svensson (2001). Blinder (2000) initiated an approach under which surveys are used to determine key characteristics which the respondents believe determine credibility; those characteristics are then quantified to construct a credibility index.

In a world in which inflation targeting, whether explicit or implicit targeting of inflation, has predominated as the primary objective of central banks, much of the empirical literature on central bank credibility has focused on the degree to which economic agents expect that central banks will attain their inflation targets (whether explicit or implicit). Consequently, studies have used various measures of inflation expectations (also using alternative measures of inflation) to assess their degree of compatibility with central banks' inflation objectives. Initial studies in this genre include Cukierman and Meltzer (1986), Cukierman (1992), Bomfim and Rudebusch (2000) and Cecchetti and Krause (2002). In what follows, we review more recent studies that have not only sought to determine how closely central banks have achieved their inflation targets but have constructed quantitative measures of credibility.

Several studies by Bordo and Siklos (2014, 2015, 2017) have estimated credibility indices for various groups of countries and for individual countries. In Bordo and Siklos (2014), those authors quantified credibility during the period 2005 through

2013 (annual data) for three groups of countries: advanced (including the United States), the euro zone, and emerging markets. To derive a measure of credibility for each of these groups, the authors defined credibility as the squared difference between observed inflation in a particular year and the mean of inflation forecasts (for each group) derived from AR(1) and IMA (1,1) models (2014, 75). A central bank was deemed credible when it delivered its inflation objective conditional on the monetary regime in place. For the group of advanced economies, the following results (under which high credibility registered a low score) were obtained: (1) credibility was higher using post-2000 data compared with data using a longer sample period;[3] and (2) credibility was high during most of the sample (based on their measure, credibility was near zero) but deteriorated in 2008 (to about 7) and 2009 and 2010 (to near 2). Credibility measures were not derived for individual central banks.

Bordo and Siklos (2015) constructed individual measures of credibility for ten central banks, including the Fed. Using panel (annual) data, the authors measured credibility as the squared difference between a 3-year average of actual inflation and an implicit, time-varying inflation target. As in their 2014 paper, the authors estimated expected inflation using both an AR(1) model and an IMA(1,1) model. For the Fed, the estimation period was 1913 (the year that the Fed established) to 2010. For the post-1970 period, Bordo and Siklos found sharp deteriorations in the Fed's credibility in the early-1970s and the late-1970s. Credibility improved sharply in the early-1980s and remained at high levels throughout the remainder of the estimation period.

Bordo and Siklos (2017) investigated the drivers of credibility for a group of both advanced and emerging economies over the period 1980 to 2014 using annual data. Two measures of credibility are used: (1) the difference between the expected inflation rate one-year ahead and a moving average of past inflation rates. Bordo and Siklos (2017) related those two measures of credibility to sets of economic (*e.g.,* GDP growth), financial stability (*e.g.,* non-performing loans), and institutional (*e.g.,* rule of law) indicators. For the United States, estimates of the Fed's credibility were presented with the credibility measure ranging from zero (high credibility) to 3 (low credibility). Bordo and Siklos found that the Fed's credibility fell sharply during the 2007-09

3 The longer sample period covered three monetary regimes: the gold standard (1880-1917 and 1922-1933), Bretton Woods (1959-1972), and flexible exchange rates (1980-2014).

financial crisis; otherwise, the Fed's credibility measure was between 0 and 1 (2017, 34).

Park (2023) constructed numerical measures of credibility for both the Fed and the ECB using a New Keynesian model to generate forecasts of inflation and then comparing those forecasts with the inflation rate estimated with a time-varying-parameter model. That comparison provided the basis for the construction of a credibility measure that ranged from zero (low credibility) to unity (high credibility). For the Fed, the credibility measure was computed during the quarterly interval 1968:4 to 2014:4. The Fed credibility was found to be highest (between 0.8 and 1.0) in 1975 and 1977-78; credibility dropped to .5 in 2004 and to .4 in 2008. Overall, mean credibility was found to be above 0.85 for the entire estimation period (2023, 159).

Jabbar *et al.* (2023) used Big Data Analytics to construct credibility indices for Bank Indonesia. The indices were based on text mining of public perceptions toward the central bank's credibility that are reported in the news media – namely, text mining of economic and financial news. Ranking credibility from 0 percent to 100 percent during the period from 2012:1 to 2020:1 (semi annual data), the authors found an average credibility score of 63.4 percent, with the credibility score demonstrating a slight upward trend during the sample record (2023, 14, Figure 4).

Issler and Soares (2023) constructed a credibility index for the Brazilian central bank, based on the difference between survey data on expected consumer price inflation 12-months ahead and the central bank's explicit inflation target. The data cover the period November 2001 until April 2017. The difference between the survey data and the target inflation rate was used to construct confidence intervals; whenever the target fell within the interval, the authors considered the bank to have been credible. Using GMM estimation, the authors constructed a credibility index – 0 to 1, with 1 being highly credible – and found that the Brazilian central bank "was credible 65 percent of the time, with the exception of a few months in the beginning of 2007 and during the interval between mid-2013 and throughout mid-2016" (2023, 1268 and 1283, Figure 7).

A broadly similar procedure to estimate credibility of the Brazilian central bank was provided by Hecq, Issler, and Voisin (2024). However, instead of using survey data on inflation as the basis of comparison, those authors used the predicted densities obtained from a mixed autoregressive model (MAR) model. Specifically, the authors computed the probability that annual inflation would remain "within target bounds in the future as an indicator of whether the central bank in credible and to construct a credibility index" (2024, 11), with values ranging from 0 (not credible) to 1 (highly credible). The authors found that the central bank was highly credible (credibility index near 1) during most of the period from 2007 to 2017, although there were several periods of very low credibility (credibility index near 0).[4]

3. Empirical approach

3.1 Background

In what follows, we develop a statistical approach to quantify central bank credibility based on a dynamic factor and estimated using the Kalman filter. There are two basic forms of factor models: static and dynamic. The best known of the static approaches is principal components. This technique essentially averages a set of observed variables together with a set of weights that maximize the ability of each factor to explain a set of observed variables. The principal components technique, however, is static because it focuses on each point in time in isolation, forming a weighted average of the variables of interest at a particular point in time. This means that the factors, or principal components, can vary erratically over time. The first stage of principal components is to normalize all the variables so that they have a zero mean and a unit variance such that a final weighted component will have an equal weight in its construction from each of the observed variables. The advantage of principal components is that it provides a number of factors, all orthogonal to each other, that explain any amount of variation in the observed data that is desired. The disadvantage of this technique is that its static nature can yield implausible results if the underlying factor moves smoothly and slowly through time.

[4] The paper by Hecq, Issler, and Voisin (2024) contained a figure that compared their credibility index with the one constructed by Issler and Soares (2023). The two indices give very different results during several time periods. Hecq, Issler, and Voisin (2024, 11) stated: "The discrepancies between the two methods stem from the horizon and perspective of each index" – namely, the Issler and Soares index was based on inflation expectations one year ahead whereas the Hecq, Issler, and Voisin index was based on predicted inflation 1, 3, and 6 months ahead.

In contrast, under dynamic factor analysis, while the objective is essentially the same as under principal components -- to derive a set of factors that explain the observed data -- but in this case to also be relatively smooth. In other words, instead of averaging the data at each point in time separately, dynamic factor analysis averages the data over a number of periods (with a set of geometrically declining weights either side of the estimation point). To do this, the dynamic factor models use the state space form and the Kalman filter. Initial work in this area by Stock and Watson (1991) and Garratt and Hall (1996) was only able to derive a single dynamic factor. Recently, Gibson, Hall, and Tavlas (2022) demonstrated how multiple factor models can be generated using the Kalman filter and also how standard principal components can be generated in the Kalman filter approach, thus giving the two frameworks a unified form. For purposes of this study, the intuition is that we have many macro series which are affected by central bank credibility and they are all also subject to other influences and measurement error. Consequently, no single series can fully reflect central bank credibility. By finding the unobserved component which underlies all the series, we are able to identify numerically the degree to which the central bank is credible over time.

3.2. The Kalman filter: an overview

The Kalman filter is typically used in two main ways. The first, and probably the most common, way is to generate time-varying parameters in what otherwise appears to be a standard econometric equation. The second approach is to provide estimates of some unobserved component. The idea here is that, given a range of variables that are affected by this unobserved component, the filter works backwards to provide an estimate of the underling unobserved component. The intuition underlying the studies by Stock and Watson (1991) and Garratt and Hall (1996) was that various measures of GDP (income, output, expenditure) provide different estimates of the state of the economy, but they share a common true unobserved level of underlying economic activity. Using the Kalman filter, those studies took the actual observations of a set of variables and produce an optimal maximum likelihood estimate of the unobserved state of the economy.[5] Our point of departure is the following: we assume that there is a number of observed variables that are affected by central bank credibility. These variables include interest rates, the money supply, expectations of inflation and the

[5] For a detailed account of the Kalman filter, see Cuthbertson, Hall, and Taylor (1992, chap. 7).

outcome for inflation itself relative to a target rate. We argue that these observed variables can be used to filter-out a measure of central bank credibility.

The Kalman filter requires a model to be set up in state space form, so that there are two sets of equations. The first are the measurement equations, which take the following general form:

$$Y_t = S_t X_t + \varepsilon_t \tag{1}$$

where Y_t is an nx1 vector of measured variables, S_t is an mx1 vector of unobserved state variables, X_t is an 1xm vector of exogenous variables (in our case these will be unity so they can be dropped from the formulation) and ε_t is a nx1 vector of error terms which are distributed as $N(0, \Omega_t)$, where Ω_t is usually assumed to be diagonal (the fact that this matrix has a t subscript is important, as discussed below).

Given our use of the unobserved component version of the state space form, we can simplify the measurement equations as follows:

$$Y_t = S_t + \varepsilon_t \tag{2}$$

In line with the first stage of principal components, the Y_t measured variables are normalized to have a zero mean and unit variance. Consequently, all the observed variables have an equal weight in the construction of the underlying credibility series. The second part of the state space form is given by the state equations:

$$S_t = S_{t-1} + \zeta_t \tag{3}$$

where ζ_t is a set of error terms distributed as $N(0, \Psi)$. Under the normality assumption for the two sets of errors, the Kalman filter gives maximum likelihood estimates of the underlying state variables. (If this assumption is relaxed, it gives least squares estimates.) This model will filter any common information from the *X* variables into *S* and all remaining variation in *X* will be relegated to the idiosyncratic effect. The state equations (3) can take a number of alternative forms; if the lagged variable is dropped (so that $S_t = \xi$), then the dynamic factor is no longer dynamic and the Kalman filter gives the same result as principal components. If a constant is added (so that

$S_t = \alpha + S_{t-1} + \zeta$ (that is, a random walk with drift), then the state variable would trend upwards or downwards depending on the sign of α, the drift parameter. The random walk without drift formulation (3) is generally preferred because it allows the state variable to move freely without any prior restriction.

The two ways of using the Kalman filter can then be made explicit. In a time-varying parameter model, Y_t is the standard dependent variable, X_t (in (1)) are the independent variables and S_t are the time-varying parameters. In the unobserved component model, Y_t is a set of observed variables (in our case, all the things that may be affected by credibility), X_t effectively disappears and becomes a set of ones, and S_t is a single variable -- our unobserved measure of credibility. In typical applications, the measurement covariance matrix Ω_t is simplified to be time invariant. But in our case using a time-varying covariance matrix is important, as discussed below.

The Fed began explicit inflation targeting in 2012; consequently, before that year, there was no announced inflation target and the Fed's success in containing inflation provides different information than it does after 2012. One approach for dealing with this issue would be to start estimation after 2012, but then we could say nothing about credibility during the earlier period. However, by setting the diagonal element of Ω_t, which corresponds to hitting the target, to a very large number before the initiation of inflation targeting and then resetting it to a small number after inflation targeting was announced, we are able to use a much longer data period and define the earlier period as one in which the announced target does not affect our measure of credibility. Moreover, recent articles by Ireland (2007) and Milani (2020) provided evidence that, although the Fed did not announce an explicit inflation target before 2012, it operated as if it had an implicit target for a longer period. In this regard, Ireland (2007) estimated a New Keynesian model over the period 1959 to 2004 for which the results indicated that the Fed's implied inflation target rose from 1¼ percent in 1959 to hit twin peaks at or above 8 percent in the mid- to late-1970s, before falling back to below 2¼ percent in 2004. Milani (2020) estimated a monetary model under both subjective and national expectations in which he found that under the former, the Fed's implied target was between 3 percent between 1960 to 2005; under national expectations, the implied target was 2 percent in the early-1960s, and rose to 8 percent in the 1970s, before falling to 2 percent after the mid-1980s. Clearly, it is not possible to confirm the accuracy of these estimated implied targets. We assume that the Fed had an implicit target of 2 percent during the part of our sample period prior to 2012.[6] The resulted report below would not be affected by assuming differing implied targets. To allow for this possibility the 2 percent implied targets, we provide one measure of deviations in inflation from target (target1) for the announced period with a variance which captures the announcement period through the variance (announcement), and one variable for deviations of inflation from the implicit target for our entire period (target2).

Our explicit model is as follows. The measurement equations -- collectively we refer to the measurement equations as equation (4) -- are:

$rfedfund_t$	$= +S_t + \varepsilon_{1t}$	the real federal funds rate
$rm2_t$	$= -S_t + \varepsilon_{2t}$	the real growth rate of M2
$target1$	$= -S_t + \varepsilon_{3t}$	the difference between actual inflation and the Fed's explicit inflation target (after 2012)
$target2$	$= -S_t + \varepsilon_{4t}$	the difference between actual inflation and the Fed's implicit inflation target (4)
$target3$	$= -S_t + \varepsilon_{5t}$	the difference between expected inflation 5 years ahead and Fed's explicit inflation target
$target4$	$= -S_t + \varepsilon_{6t}$	the difference between expected inflation 5 years ahead the Fed's implicit target
$rlbond_t$	$= +S_t + \varepsilon_{7t}$	the real yield on 10-year Treasury bonds
gdp_t	$= +S_t + \varepsilon_{8t}$	real GDP growth
$unemp_t$	$= -S_t + \varepsilon_{9t}$	unemployment rate
$rexch_t$	$= S_t + \varepsilon_{10t}$	real effective dollar exchange rate

where S_t is the state variable which is our measure of central bank credibility.

The data are monthly from 1965m1 to 2024m10. Inflation is the annual change in the CPI; rm2 is the real annual growth rate of M2; rfedfund is the real federal funds rate; rlbond is the real 10 year government bond yield; target3 is the difference between expected inflation 5 years ahead and the explicit target of 2 percent -- this series begins in 2012 with the announcement of an explicit target; target 4 is the difference between

[6] As noted below, our sample period extends backward to 1965. King and Lu (2022, 2) reported that during 1996 the Federal Open Market Committee "coalesced on an internal long-run goal of 2 percent inflation, but they chose not to make it public."

expected inflation 5 years ahead and the Fed's implicit target of 2 percent. The inflation expectations series begins in 2003m2, so before this date it does not exist and we again allow for this by setting ε_{6t} to a large number before 2003m2; gdp is the growth rate of gross domestic product, which is available quarterly: we interpolated it to a monthly frequency using a cubic interpolation technique; unemp is the unemployment rate. Rexch is the real effective exchange rate, defined as the ratio of a trade-weighted basket of currencies to the dollar; thus, an increase in the ratio corresponds to an appreciation of the dollar. All data are from the St. Louis Fed database FRED.

In contrast to Stock and Watson (1991) and Garratt and Hall (1996) – and, to our knowledge, all previous studies using the Kalman filter – the observed variables are standardized so that they have a unit variance and zero mean. In Stock and Watson (1991) and Garratt and Hall (1996) the variables differed in their scaling (for example, the difference between M2 growth and a real interest rate) so that the estimated coefficients in the measurement equations compensated for the different scaling. However, standardization removes the need for scaling and, thus, there is no need to estimate the parameters. This circumstance also applies to the variances of the measurement equations: these variances would typically be estimated because the measurement of the variances could differ substantially from each other. Again, by standardizing the variables, they all have the same variance so that the variances can all be set to unity, simplifying the estimation procedure.

Regarding the set of measurement equations in (4), the real federal funds rate, the real long bond rate, and GDP have positive signs because increases in these variables are taken to mean the Fed is gaining credibility (*e.g.*, raising real interest rates to reduce inflation); the other variables have negative signs under the presumption that rises in these variables – inflation, the unemployment rate, and money growth – reduce credibility. The smoothed state variable, which is produced by the Kalman filter, has an arbitrary scaling. Therefore, to construct a credibility index we have rescaled it so that its maximum is 10. At the end of each equation is an error term ε_{it}; our assumption about the error variance is that they are all set to unity except for ε_{3t}, ε_{5t} and ε_{6t} which are the variables involving the Fed's announced inflation target or inflation expectations. In these three cases (ε_{3t}, ε_{5t}, and ε_{6t}), the variance is set equal to a variable which has the value 1,000 before the data started and 1 after the start of inflation targeting. This implies that these three variables do not play a part in the construction of the credibility variable until 2012, the year in which the target was announced (in the case of ε_{3t}) or after the beginning of the inflation expectations data (in the cases of ε_{5t} and ε_{6t}). The single state equation is then equation (3), and the variance ξ_t was chosen to be 0.1, so that the state variable would be relatively smooth.

The Kalman filter works in three stages: prediction, updating and smoothing. It initially works through the sample from the first period to the last, at each period predicting what the state variable will be based on equation (3). Then, when we are able to actually observe the measured variables, it updates its initial estimate in light of the outcomes for these variables. Thus, the initial one step ahead prediction for the state variable is

$$S_{t|t-1} = S_{t-1} \tag{5}$$

In addition, we require a matrix, P_t, which is the state error covariance matrix and is predicted at each point in time, to be

$$P_{t|t-1} = X_t P_t X_t' + \Psi_t \tag{6}$$

We then observe the measurement variables at time t and are then able to update the initial estimates of $S_{t|t-1}$ *and* $P_{t|t-1}$ with the information at time t using the updating equations

$$S_{t|t} = S_{t|t-1} + K_t(Y_t - S_{t|t-1}) \tag{7}$$

and

$$P_{t|t} = (I - K_t X_t)P_{t|t-1} \tag{8}$$

where K_t is the Kalman gain.

These equations are used recursively from the first period to the last to produce the updated state variable $S_{t|t}$. The final stage of the filter is to work backwards through time producing the smoothed estimates, that is $S_{t|T}$ where T is the last observation. This procedure gives the optimal estimate of S_t based on the full sample. If we are interested in forecasting, we would normally focus on the one step ahead predictions, but in this case the smoothed estimates give us the best overall measure of the state variable, and this is what we report below.

4. Results

We begin by showing the underlying measurement variables. As mentioned, we have two groups of variables: those that we believe have a positives effect on credibility (such as higher interest rates) and those that have a negative effect on credibility (such as inflation exceeding its target). We show these two groups separately in what follows. All the variables are standardised so that they have a unit variance and zero mean.

Figure 1 displays the time profiles of the standardised interest rates, that is, the real 10-year bond rate, and the real federal funds rate -- variables that have a positive effect on credibility. There is clear co-movement in the two series. Figure 2 shows the standardised value of the real GDP growth rate and the real effective U.S. dollara rate. There were several periods during which the behavior of real GDP diverged from movements in the two interest rate series – notably in the early-1980s, during the global financial crisis of 2007-08, and during the outbreak of the Covid shock in 2020-21; during those periods the Fed initially hiked interest rates as real output fell; subsequently, interest rates fell sharply.

Figure 3 displays the standardized values of the four target variables, each of which is assumed to have a negative effect on credibility. These variables clearly move together although, because of the starting dates for explicit targeting and the beginning of the expectations data, they start at different points of time. Figure 4 shows the remining two variables, unemployment and M2 growth, that are assumed to have a negative effect on credibility. In sum, the idiosyncratic component of each series appears to be significant such that additional information is gained by using them together.

Figure 5 reports our measure of credibility. As mentioned, we have taken credibility to be an unobserved factor. Because that factor is normalized in the setting up of the measurement equations, we can give it any absolute value we choose. We have constructed it so that it corresponds to an index that ranges from zero to 10, where zero is the lowest possible credibility and 10 is the maximum amount of credibility.

To facilitate the interpretation and performance of our credibility measure, we have identified seven separate regions (R1-R7) in Figure 5. We briefly comment on monetary policies in each of these regions.

Region 1, 1965 to 1975. This was a period during which (CPI) inflation rose from 1.6 percent in 1965 to double digit levels in 1974 and 1975; the unemployment rate rose from 4.0 percent in 1965 to 8.2 percent in 1975. William McChesney Martin was Fed chair during the early part of the period, until January 1970, when he was replaced by Arthur Burns. Figure 5 shows that the credibility measure mostly declined during the years from 1965 to 1975, from above 8 at the beginning of the period to about 2 at the end of the period. There was a temporary rebound in the early-1970s from the declining trend. This rebound corresponds to President Richard M. Nixon's announcement on August 15, 1971 of his New Economic Policy, which included a 90-day freeze on wages and prices to control inflation, the closure of the gold window, and a 10 percent surcharge of all dutiable imports. The package of measures was initially well received by the markets; inflation fell slightly after the imposition of the price and wage controls, but inflation surged after the controls became voluntary or ended in 1973, and the New Economic Policy was judged to be a failure (Meltzer, 2003, 759-60).

Region 2, 1976-80. There were three Fed chairs during this period: Arthur Burns, G. William Miller, who became Fed chair in April 1978, and Paul Volcker, who was Fed chair from August 1979 until August 1987. Inflation fell from double digit levels in 1974 and 1975, to between 5.8 percent (1976) and 7.6 percent (1978) the following three years, before again surging to double digit levels in 1979 and 1980. The unemployment rate ranged from 6 percent to 8 percent throughout 1976 to 1980. Burns' tenure at the Fed is considered to have been a failure (Meltzer, 2003, chap. 7; Hetzel, 2022, chap. 22). He did not believe in the effectiveness of monetary policy: he held a cost-push view of inflation (Nelson, 2024, chap. 12, 147). Miller had little experience with monetary policy: his appointment to the Fed chair owed to his political connections (Meltzer, 2003, 848). In these circumstances, credibility would have been expected to be low, and this is what our credibility measure shows. It remained between near 0 to 4 throughout the period, reaching its lowest level in 1980.

Regions 3, 1981-85. Volcker's actions in the early-1980s were consistent with a central bank that lacks credibility and tries to acquire it. To bring down inflation expectations, Volcker believed, monetary policy had to tighten abruptly and needed to remain tight, even if it meant bringing the economy into recession (Silber, 2012, chap. 10).[7] Under Volcker, the Fed raised policy rates to near 20 percent in the early-1980s and the economy went through two recessions. Between September 1980 and March 1985 the trade-weighted value of the dollar rose by 54 percent (Blinder, 2022, 145). Inflation fell, from 10.1 percent in 1981 to 3.5 percent in 1985. Our credibility measure shows a sharp rise from about 1 at the beginning of the period to about 7 at the end of the period.

Region 4, 1986-2000 (The Great Moderation). Inflation moved in a narrow range – mainly between 2 percent to 4 percent -- during the period. Reflecting a drop in inflation expectations, the yield on the 10-year Treasury note fell steadily from over 10 percent in 1985 to near 5 percent at the end of the period. The unemployment rate fell from about 7 percent at the beginning of the period to 3.9 percent in 2000. Alan Greenspan succeeded Volcker as Fed chair in August 1987 and remained in that position throughout the period. Greenspan's policies until the early-2000s have been characterized as having implicitly followed a Taylor type rule until the early-2000s (Taylor, 2012). Our credibility measure registers consistently high numbers – between 5 and 9 – during the Great Moderation.

Region 5, 2001-2010. Although credibility measure was high during the early part of this period, it subsequently declined in the ten years 2001 to 2010, especially during the Great Financial Crisis (GFC) of 2008-10. The unemployment rate rose from about 6 percent at the beginning of the period to over 9 percent in both 2009 and 2010. Inflation fluctuated in a range of 1.5 percent to 3.8 percent from 2001 to 2009, before declining by 0.3 percent in 2010. Alan Greenspan was Fed Chair until 2006, when Ben Bernanke took over. Under Bernanke, the Fed undertook quantitative easing, expanding its balance sheet during 2007 to 2010. The economy was in a recession from December 2007 to June 2009, during which period it contracted by 5.1 percent. The credibility measure started the period near 8; it ended the period near 4.

[7] This policy differed from that being recommended by leading monetarists, including Milton Friedman, and Karl Brunner and Allan Meltzer, during the late-1970s and early-1980s. The monetarists believed that monetary policy needed to be tightened gradually. See Tavlas (2025).

Region 6, 2011-2018. Inflation was near 2 percent throughout the period. The unemployment rate dropped steadily from 8.5 percent in 2011 to 3.9 percent in 2018. The yield on the 10-year Treasury note fluctuated in a very narrow range – from 1.8 percent to 2.8 percent. Janet Yellen was Fed Chair until February 2018, when Jerome Powell took over. Credibility rose sharply from 3 about 9 at the beginning of the period and remained high after that.

Region 7, 2019-2024. With the outbreak of Covid in 2020, the economy entered a brief but very sharp recession during which real GDP declined by 19 percent (at the beginning of 2020). Inflation rose strongly in 2021 and 2022: due to the Fed's delayed response to the increase in inflation, credibility fell sharply. Our credibility registers a decline from about 9 in 2019, to near 4 in 2021 and 2022. However, with the sharp tightening in the Fed's policy stance in 2022, the decline in credibility proved to be short-lived – credibility rose to about 9 at the end of the period.

5. Conclusions

We applied the Kalman filter to derive a measure of credibility for the Federal Reserve by deriving an unobserved component from a group of variables that are affected by central bank credibility. We extend previous studies that use the Kalman filter in that we standardize all the variables to that they have a zero mean and unit variance, removing the need to estimate parameters in the measurement equations. There is inevitably an issue in identifying and interpreting any unobserved component in terms of a real-world counterpart, such as credibility. However, we believe that our resulting index conforms very well with our understanding of historical events and that it provides a useful formalisation of our understanding of the evolution of the credibility of the Fed.

Figure 1: Interest Rates

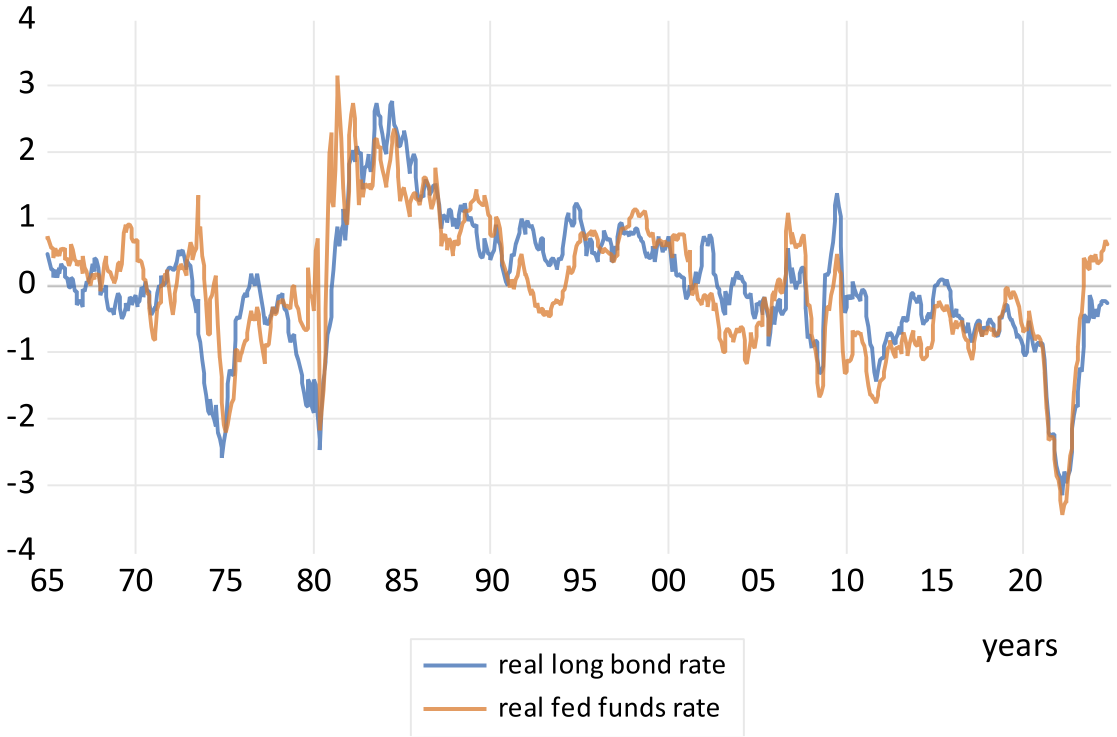


Figure 2: Real GDP Growth and the Real Effective Exchange Rate

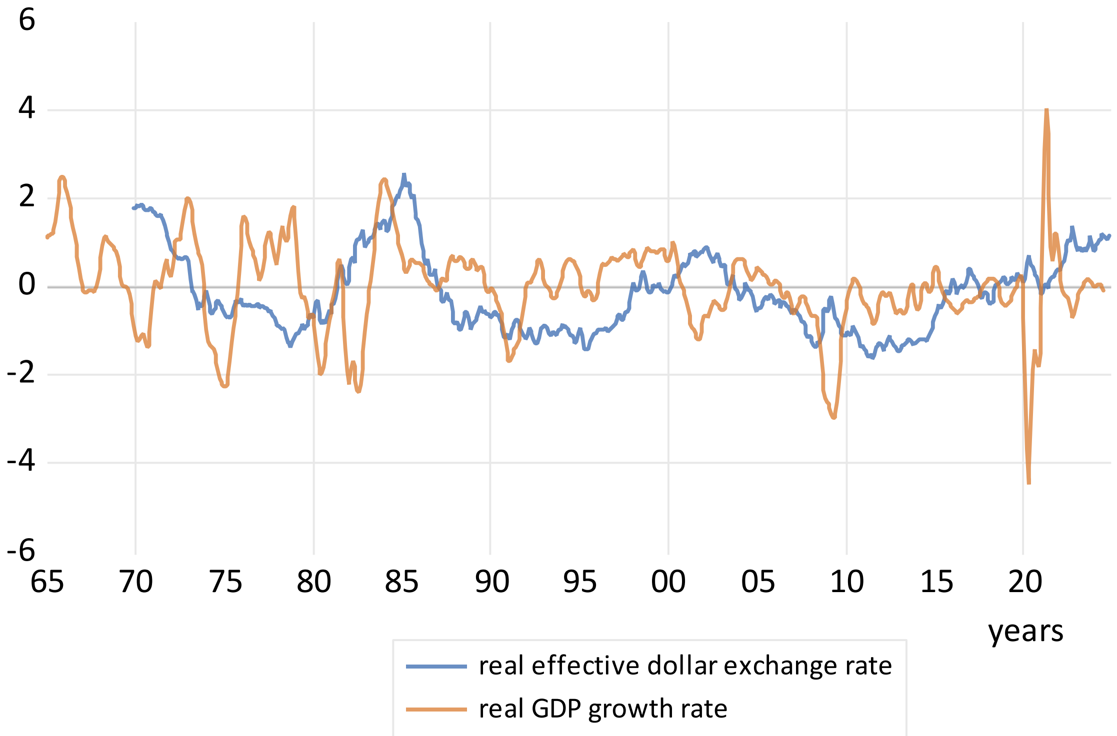


Figure 3: The Four Variables Involving the Fed's Target

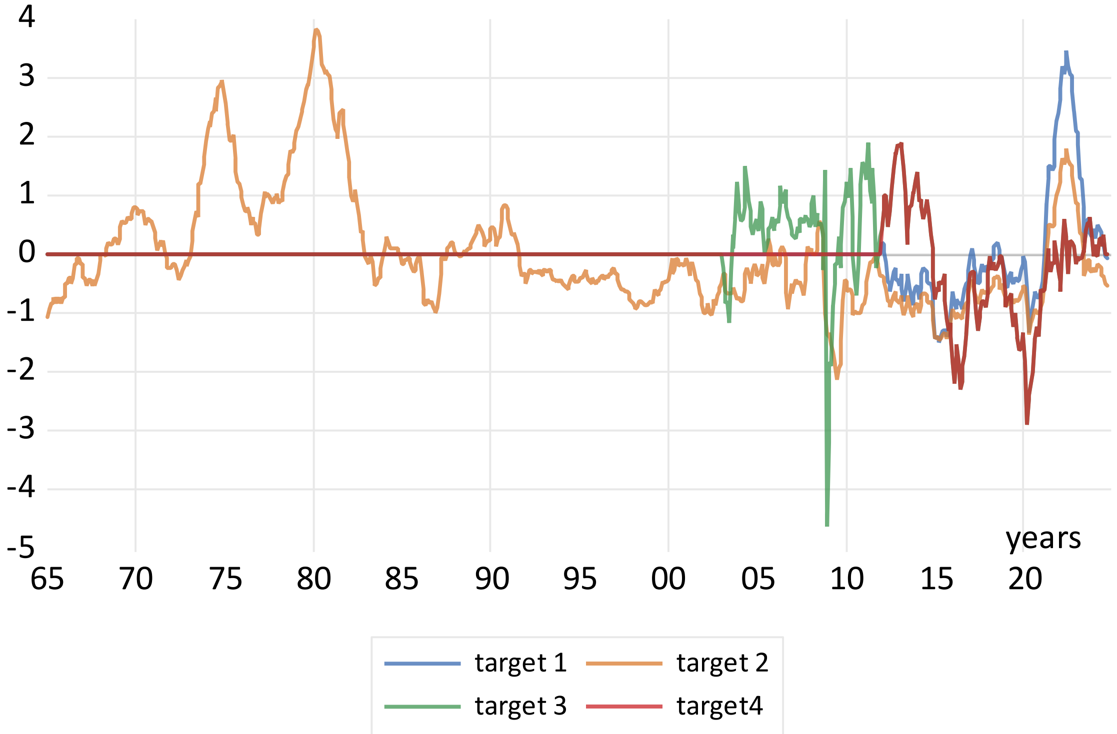


Figure 4: M2 Growth and the Unemployment Rate

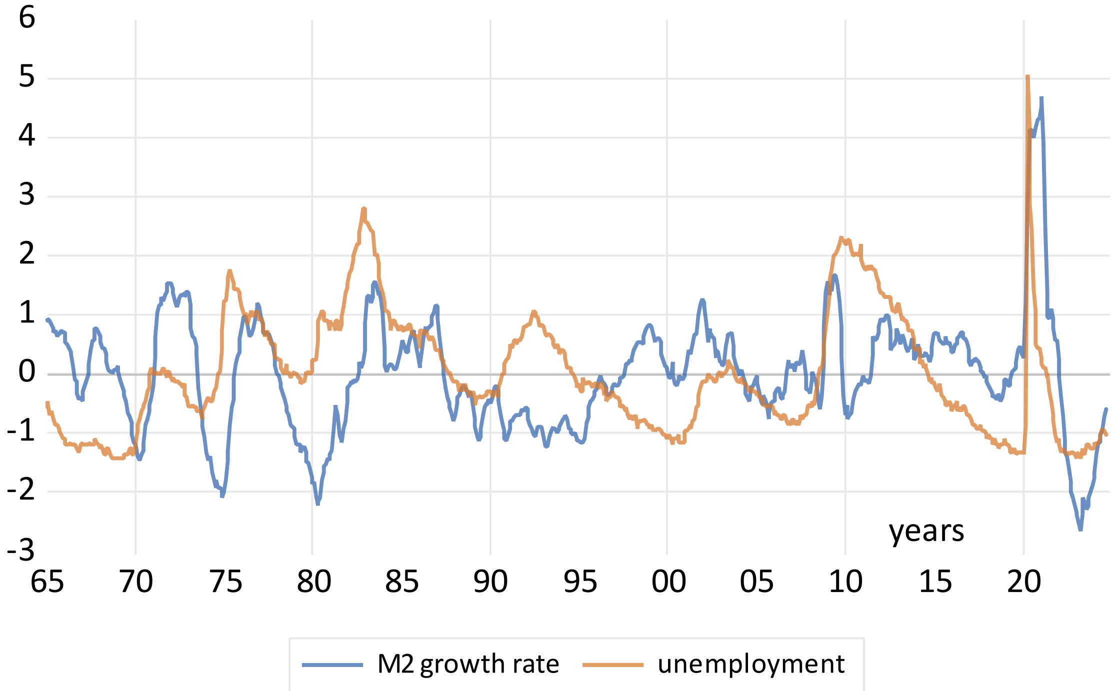
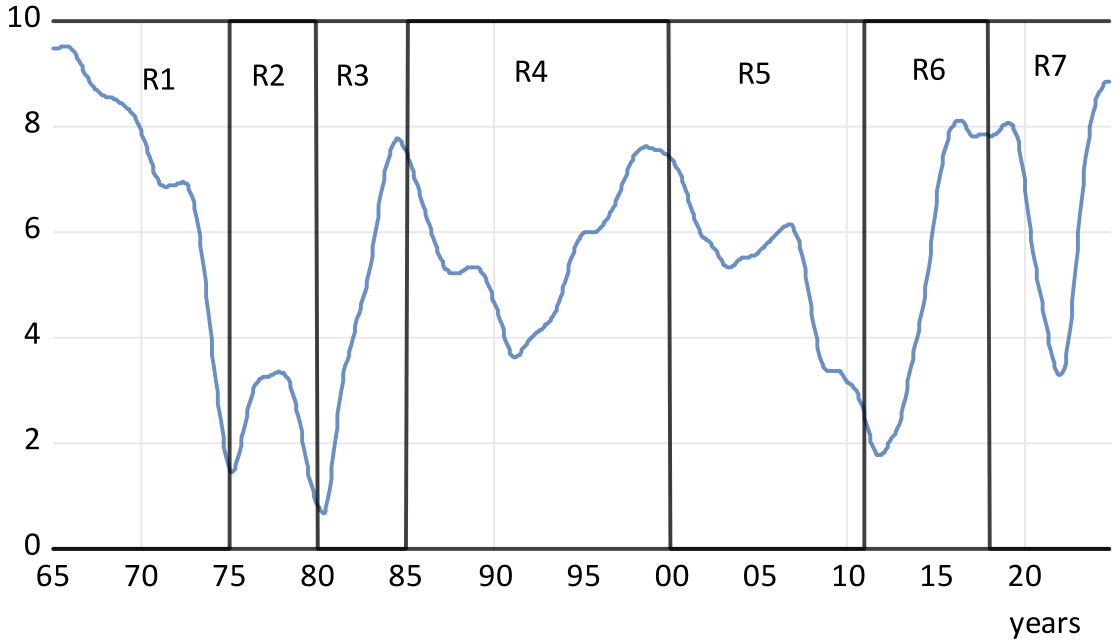


Figure 5: Credibility Measure



References

Blinder, Alan S. (2000) Central bank credibility: why do we care? How do we build it? *American Economic Review* 90(5), 1421-1431.

Blinder, Alan S. (2022) *A Monetary and Fiscal History of the United States, 1961-2021*. Princeton: Princeton University Press.

Bomfim, Antulio N. and Glenn D. Rudebusch. (2000) Opportunistic and deliberate disinflation under imperfect credibility. *Journal of Money, Credit and Banking* 32(4) Part 1, 707-721.

Bordo, Michael D. and Finn E. Kydland. (1995) The gold standard as a rule: an essay in exploration. *Explorations in Economic History* 32, 423-464.

Bordo, Michael D. and Pierre L. Siklos. (2014) Central bank credibility, reputation and inflation targeting in historical perspective. National Bureau of Economic Research *Working Paper* 20693.

Bordo, Michael D. and Pierre L. Siklos. (2015) Central bank credibility: A historical and quantitative exploration. National Bureau of Economic Research *Working Paper* 20824. In *Central Banks at a Crossroads: What Can We Learn from History?* edited by M.D. Bordo, Ø. Eitrheim and M. Flandreau. Cambridge, Mass.: Cambridge University Press, 62-144.

Bordo, Michael D. and Pierre L. Siklos. (2017) Central bank credibility before and after the crisis. *Open Economies Review* 28(1), 19-45.

Cecchetti, Stephen and Stefan Krause Montalbert. (2002) Central bank structure, policy efficiency, and macroeconomic performance: exploring empirical relationships. *Review, Federal Reserve Bank of St. Louis* 84(4), 47-60.

Cukierman, Alex. (1992) Central bank strategy, credibility, and independence: theory and evidence. MIT Press Books, vol. 1, number 0262031981 (December).

Cukierman, Alex and Allan H. Meltzer. (1986) A theory of ambiguity, credibility, and inflation under discretion and asymmetric information. *Econometrica* 54(5), 1099-1128.

Cuthbertson, Keith, Hall, Stephen G, and Mark P. Taylor. (1992) *Applied Econometric Techniques*. New York: P. Allan.

Ehrmann, Michael. (2024) Trust in central banks. Centre for Economic Policy Research *Discussion Paper* DP19781.

Faust, Jon and Lars Svensson. (2001) Transparency and credibility: monetary policy with unobservable goals. *International Economic Review* 42(2), 369-397.

Garratt, Anthony and Stephen G. Hall. (1996) Measuring underlying economic activity. *Journal of Applied Econometrics* 11(2), 135-151.

Gibson, Heather D., Hall, Stephen G., and George S. Tavlas. (2022) A suggestion for a dynamic multi factor model (DMFM). *Macroeconomic Dynamics* 26(6), 1423-1443.

Hecq, Alain, Issler, João Victor, and Elisa Voisin. (2024) A short term credibility index for central banks under inflation targeting: an application to Brazil. *Journal of International Money and Finance* 143(C).

Hetzel, Robert L. (2022) *The Federal Reserve: A New History*. Chicago: University Chicago Press.

Ireland, Peter N. (2007) Changes in the Federal Reserve's inflation target: causes and consequences. *Journal of Money, Credit and Banking* 39(8), 1851-1882.

Issler, João Victor and Ana Flávia Soares. (2023) Central bank credibility and inflation expectations: a microfounded forecasting approach. *Macroeconomic Dynamics* 27(5), 1268-1288.

Jabbar, Muhammad Abdul, Widjanarti, Anggraini, Zulen, Alvin Andhika, and Arinda Dwi Okfanti. 2023 Sentiment analysis of tourist reviews from online travel forum for improving Indonesia tourism sector. In: Irving Fisher Committee on Central Bank Statistics, *Data Science in Central Banking: Applications and Tools*, IFC Bulletin No. 59. Switzerland: Bank for International Settlements.

King, Robert G. and Yang K. Lu. (2022) Credibility and Explicit Inflation Targeting. NBER Working Paper Series, National Bureau of Economic Research *Working Paper* 30012.

Meltzer, Allan H. (2003) *A History of the Federal Reserve, Volume 1: 1913–1951*. Chicago: University of Chicago Press.

Milani, Fabio. (2020) Learning and the evolution of the FED's inflation target. *Macroeconomic Dynamics* 24(8), 1904-1023.

Nelson, Edward. (2024). *Milton Friedman and Economic Debate in the United States, 1973-2006*. 2 vols. Draft.

Park, Kwangyong. (2023) Central bank credibility and monetary policy. *International Journal of Central Banking* 19(2), 145-197.

Silber, William. (2012) *Volcker: The Triumph of Persistence*. USA: Bloomsbury Publishing.

Stock, J. and M. W. Watson. (1991) A probability model of the coincident economic indicators. In: Moore, G. and K. Lahir (eds.), *The Leading Economic Indicators: New Approaches and Forecasting Records*, pp. 63-90. Cambridge: Cambridge University Press. pp. 63-90.

Svensson, Lars E. O. (2008) Inflation targeting. In: Durlauf, Steven N. and Lawrence E. Blume (eds.), *The New Palgrave: Dictionary of Economics*, Vol. 4 (Herman-Lange), 2nd ed., pp. 319-321. London: Palgrave Macmillan.

Tavlas, George S. (2025) Discussion of "The Evolution of Monetary Policy: Critical Issues That Have Shaped Monetary Policy and the SOMC Over 50 Years" by Michael Bordo and Mickey Levy. In: Bordo, Michael, Lacker, Jeffrey M., and John B. Taylor (eds.), *A 50-Year Retrospective on the Shadow Open Market Committee and Its Role in Monetary Policy*. Standford: Hoover Press, *forthcoming.*

Taylor, John B. (2012) Monetary Policy Rules Work and Discretion Doesn't: A Tale of Two Eras. *Journal of Money Credit and Banking*, 44(6), 1017-1032.

Yellen, Janet L. (2006) Enhancing Fed credibility. Speech to the Annual Washington Policy Conference, sponsored by the National Association for Business Economics (NABE), Washington, DC, March 13.

BANK OF GREECE WORKING PAPERS

321. Bragoudakis, Z. and I. Krompas “Greek GDP forecasting using Bayesian multivariate models”, June 2023.

322. Degiannakis, S. and E. Kafousaki “Forecasting VIX: The illusion of forecast evaluation criteria”, June 2023.

323. Andreou C. P., S. Anyfantaki, C. Cabolis and K. Dellis, “Exploring country characteristics that encourage emissions reduction”, July 2023.

324. Dimakopoulou, V., Economides, G., Philippopoulos, A., and V. Vassilatos, “Can central banks do the unpleasant job that governments should do?”, December 2023.

325. Chrysanthakopoulos, C. and A. Tagkalakis, “The medium-term effects of fiscal policy rules”, January 2024.

326. Manou, K. and E. Papapetrou, “Does uncertainty matter for household consumption? A mean and a two tails approach”, February 2024.

327. Kakridis, A., “War, mobilization, and fiscal capacity: testing the bellicist theory in Greece, 1833-1939”, March 2024.

328. Mavrogiannis, C. and A. Tagkalakis, “From policy to capital: assessing the impact of structural reforms on gross capital inflows”, April 2024

329. Delis, P., S. Degiannakis, G. Filis, T. Palaskas and C. Stoforos, “Determinants of regional business cycle synchronization in Greece”, May 2024.

330. Sideris, D. and G. Pavlou, “Market power and profit margins in the Euro area countries in the post-pandemic period”, June 2024.

331. Kasimati, E. and N. Veraros, “The dry-bulk shipping market: a small econometric model”, September 2024.

332. Mermelas, G. and A. Tagkalakis, “Monetary policy transmission: the role of banking sector characteristics in the euro area”, November 2024.

333. Anastasiou, D., Pasiouras, F., Rizos, A., and A. Stratopoulou, “Do macroprudential policies make SMEs more-or-less discouraged to apply for a bank loan?”, December 2024.

334. Malliaropulos, D., Passari, E., and F. Petroulakis, “Unpacking commodity price fluctuations: reading the news to understand inflation”, December 2024

335. Degiannakis, S. and E. Kafousaki, “Disaggregating VIX”, January 2025

336. Degiannakis, S., Delis, P., Filis, G., and G. Giannopoulos, “Trading VIX on volatility forecasts: another volatility puzzle?”, February 2025

337. Papadopoulos, G., Ojea-Ferreiro, J., and R. Panzica, “Climate stress test of the global supply chain network: the case of river floods”, February 2025

338. Papaoikonomou, D., “Stochastic debt sustainability analysis: a methodological note”, March 2025

339. Dellas, H. and G. Tavlas, “The great dollar shortage debate: a modern perspective”, March 2025