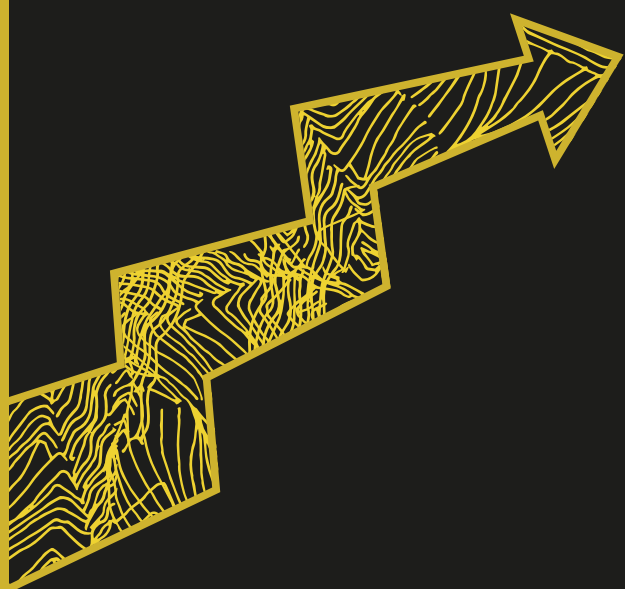


# Documentos de trabajo



**N. 2**

Abril, 2024

---

**Variation Index of the Output Gap  
(VIOG): A  
New Way of Testing Potential GDP  
Estimations**

# **Variation Index of the Output Gap (VIOG): A New Way of Testing Potential GDP Estimations**

Serie Documentos de Trabajo 2024

Edición No. 2

Edición digital

Abril del 2024

© 2024 Valor Público, centro de estudios e incidencia.

Universidad EAFIT

[Valorpublico@eafit.edu.co](mailto:Valorpublico@eafit.edu.co)

## **Autores**

Alejandro Pinilla Barrera

Álvaro Hurtado Rendón

Hermilson Velásquez Ceballos

## **Serie documentos de trabajo de Valor Público EAFIT**

El centro de estudios e incidencia Valor Público, de la Universidad EAFIT, es un escenario para la comprensión y la transformación de problemas que requieren de la intervención colectiva y la decisión compartida: los asuntos públicos. Sus documentos de trabajo dan a conocer los resultados de los proyectos de investigación. Esta serie reúne trabajos de sus cuatro iniciativas de investigación y trabajo: gobierno y democracia, seguridad y justicia, inclusión y diversidad y desarrollo económico.

Los documentos de trabajo están disponibles en:

<https://www.eafit.edu.co/escuelas/economiayfinanzas/valorpublico/Paginas/documentos-de-trabajo.aspx>

# *Variation Index of the Output Gap (VIOG): A New Way of Testing Potential GDP Estimations*

Alejandro Pinilla Barrera\*  
Álvaro Hurtado Rendón†  
Hermilson Velásquez Ceballos‡

## Abstract

The potential GDP plays a fundamental role in macroeconomic models used by policymakers to make decisions. Its nature as an unobservable variable has led to the proposition of various estimation techniques based on different assumptions about its generating process. This work proposes a methodology to evaluate the different techniques for estimating potential GDP: the *Variation Index of Output Gap (VIOG)*. Inspired by the statistic mean absolute deviation and the work of [Darvas, Vadas, et al. \(2003\)](#), the *VIOG* measures the average absolute gap derived from a potential GDP estimation technique. Unlike other methodologies that assess the statistical performance of estimation techniques, the *VIOG* aims to provide a practical response to the subjective dilemma of potential GDP estimation depending on what policymakers seek: more aggressive economic policies, where a greater output gap is permitted, or more conservative ones, which show a smaller output gap. To illustrate this, an application using the Taylor rule is presented. The results suggest that the differences between using different methodologies to estimate potential GDP can be significant, especially in times of crisis, affecting policymakers' decision-making, implying that depending on the technique used, more or less restrictive economic policy decisions may be taken.

**Keywords:** potential GDP, output gap, business cycle, Taylor's rule, statistical filters.

**JEL Codes:** C22, C32, E23, E32, E37, E61, E63.

---

\*Msc in Economics from Universidad EAFIT, Colombia. Email: apinill2@eafit.edu.co

†Phd in Economics from National University of the South, Argentina. Professor and researcher at Universidad EAFIT, Colombia. Email: ahurtad1@eafit.edu.co

‡Phd in Mathematical Sciences from Polytechnic University of Valencia, Spain. Email: evelas@eafit.edu.co

# 1 Introduction

Within the central discussions of economic policy lies a concept widely elusive due to being an unobservable variable, yet fundamentally important—this is the potential or trend product. The former is defined as the maximum level of production that an economy can achieve without generating an acceleration in inflation, referring to the current level of productive capacity, which includes various transient productivity shocks and resource allocation issues that describe the economy at any given moment. On the other hand, there is the notion of trend product, a concept that refers to the productive capacity in the absence of transient productivity shocks, meaning when productive inputs are used at their normal capacity. In this case, the accumulation of positive and negative transient shocks tends to cancel out in the long run. However, empirically, both concepts tend to overlap because the difference is clear in theory but not in practice. It assumes identifying fluctuations as a result of demand and supply shocks and then establishing the transient and permanent nature of these fluctuations ([Fontanari, Palumbo, & Salvatori, 2020](#)).

The discussion can initially be traced back to [Okun \(1962\)](#) under a Keynesian approach and then extended to the neoclassical framework. In the former, reference is made to potential GDP, an element that leads to the presence of involuntary unemployment and the absence of inflationary pressures. Generally, the most commonly used method is to resort to an aggregate production function, where the concepts of the Non-Accelerating Inflation Rate of Unemployment (NAIRU) and the Non-Accelerating Inflation Capacity Utilization Rate (NAICU) become predominant in defining demand economic policy. In the second approach, the trend GDP, the neoclassical approach is found, where equilibrium is determined by supply factors and institutional variables in the absence of price and wage rigidities. According to [US Congressional Budget Office \(2004\)](#), this difference establishes the use of different methods for estimation, which can be classified into: methods supported by production functions; and statistical methods, such as statistical filtering techniques ([Baxter & King, 1999](#); [Butterworth, 1930](#); [Christiano & Fitzgerald, 2003](#); [Hodrick & Prescott, 1997](#); [Kalman, 1960](#)), multivariate filters, and multivariate time series models, such as VAR and SVAR models.

The discussions about the aforementioned methods focus on their advantages and disadvantages. In the first case, these methods are attributed the advantage of explicitly showing the supply side. It cannot be denied that potential GDP is a measure of productive capacity; these methods allow for a transparent accounting of growth and, being disaggregated methods, they can reveal more knowledge of aggregate activity. However, they present disadvantages in terms of the simplicity imposed on certain parameters, such as the weighting of factors that sometimes do not correspond to empirical evidence, the use of deterministic temporal trends to adjust model variables, and they are also supported

by other unobservable variables such as NAIRU and NAICU, whose measurement reflects the slack in different markets. In the second case, a key advantage is flexibility, as these methods are more flexible in estimating variables and parameter values. Additionally, they closely follow the data, allowing for the identification of trend changes as the economy evolves. Their deficiencies lie in the inability to compare the result with a measure of the economy's capacity, showing that trend GDP accelerates and decelerates according to the economic cycle. Subjective judgments about the value of parameters are necessary. They are also subject to the end-of-sample problem. Furthermore, time series models are highly aggregated and consequently hide a series of economic relationships that may be of utmost importance for decision-makers. These discussions have been addressed by different authors (Darvas et al., 2003; Giorno, Richardson, Roseveare, & Van den Noord, 1995), where problems are generally addressed, alluding to statistical characteristics without resorting to the consequences derived from estimation.

While acknowledging the utmost relevance of these discussions, the primary contribution of this paper is to illustrate the discussion about statistical characteristics without delving deeply into them and to propose that the focus should be on the consequences of measurement, which ultimately serve as signals for decision-making by economic policymakers. With this objective in mind, the *Variation Index of the Output GAP (VIOG)* is constructed based on the work of Darvas et al. (2003). This index allows us, firstly, to clearly observe the consequences since it will be an indicator of more restrictive or more expansive economic policies. In other words, the normative characteristic of economic policymakers is decisive in the results of the index, and furthermore, the construction of this index is a more robust method against sample changes, a characteristic implicit in the discussions of different methods. Despite recognizing the importance of this discussion, this paper will not consider such a distinction. Instead, our concern will be the effects that different measurements of trend and potential GDP have on the restrictive or expansive nature of monetary or fiscal policy. In essence, we acknowledge the theoretical difference derived from this, but our focus will be on the implications it may have for decision-making by economic policymakers.

As a second contribution of this study, it is proposed that the discussion between potential and trend GDP does not exist in the long term, in other words, the statistical and economic discussion is unimportant because this issue pertains to the long term. This hypothesis is tested through a cointegration exercise, demonstrating that the different methods cointegrate in the long term. This means that while they may differ in the short term, in the long term, potential GDP and trend GDP are two series that converge.

With these objectives, the article is divided in 5 sections, where the first is this introduction. The second section presents a literature review of the different discussions about the methods used to estimate potential or trend GDP and

their implications. The third section introduces the proposed methodology for the *VIOG* and provides an application of it through the Taylor rule. This section is complemented by the proposal of cointegration analysis established for the different methods with the aim of demonstrating the long-term analysis. Section 4 presents the results, and Section 5 offers the conclusions.

## 2 Estimation of potential GDP: a current discussion

In 1962, Okun introduced the concept of potential production as a measure of an economy's production capacity under conditions of full employment, understanding the latter as the state that allows for maximum production without causing inflationary pressures. However, as mentioned by Carnazza, Fontanari, Liberati, and Palumbo (2021), over time, this interpretation of potential production has shifted from a Keynesian theoretical framework to an interpretation more aligned with the neoclassical school. This view considers that the current level of production fluctuates around potential production, understood as an equilibrium determined by supply factors and institutional variables, in the absence of price or wage rigidities (Romer, 2012; Anzoategui and Kim, 2021).

This theoretical debate has implications when estimating potential production because the theoretical perspective adopted will determine the technique to be used (Carnazza et al., 2021). Methods with a more statistical focus are more in line with the neoclassical theory of potential production, where potential production is seen as a long-term trend. Meanwhile, methods more focused on economic fundamentals are more related to Keynesian theory, which interprets potential production as the maximum productive capacity of the economy according to its characteristics.

Within statistical methods, the initial works of Beveridge and Nelson (1981) and Clark (1987) stand out, as they extract potential output by decomposing production into two components: one permanent (trend) and the other transitory (cyclical), using ARIMA models (Box & Jenkins, 1970) and unobservable components models (Harvey, 1989), respectively. More recently, and following the idea of decomposing output, various time series filtering techniques have been proposed to estimate potential output, such as the filters by Hodrick and Prescott (1997), Baxter and King (1999), Christiano and Fitzgerald (2003), and Kalman (1960) (Harvey & Jaeger, 1993).

However, despite the fact that decomposition or filtering techniques allow for a direct and simple estimation, they suffer from several drawbacks: Mishkin (2007) points out that depending on the technique, different statistical assumptions are required, leading to different estimations of potential output. Moreover, these methods do not provide information about their association with

other macroeconomic variables such as inflation, which is essential for the development of monetary policy. On the other hand, [Anzoategui and Kim \(2021\)](#) adds that they are subject to criticism from [Lucas \(1976\)](#) and therefore may be biased. Additionally, [Altăr, Necula, and Bobeică \(2010\)](#) and [Carnazza et al. \(2021\)](#) highlight their lack of economic foundation and interpretation. Furthermore, these criticisms persist against multivariate approaches that link production with other macroeconomic variables like inflation ([Laxton and Tetlow, 1992](#); [Kuttner, 1994](#)). Due to these issues, approaches grounded in economic theory have been preferred in recent years, with the production function and DSGE models being the most prominent.

The Keynesian approach (also known as the production function method) estimates potential output based on its theoretical determinants: the potential levels of labor, capital, and total factor productivity, combined in an aggregate production function ([Burns, Van Rensburg, Dybczak, & Bui, 2014](#); [Carnazza et al., 2021](#); [Gerali, Locarno, Notarpietro, & Pisani, 2018](#)). Several authors have pointed out the advantages of using this approach to estimate potential output: [Mishkin \(2007\)](#) acknowledges that this method focuses on the drivers of potential growth rather than simply considering the historical behavior of production or its historical relationship with other variables. Furthermore, its disaggregated nature allows for the use of more data, which is valuable during periods of structural changes. On the other hand, [Havik et al. \(2014\)](#) notes that another advantage of the method is its consistency between economic theory and practical estimation. Finally, [Altăr et al. \(2010\)](#) states that the approach enables establishing significant links between policy reform measures and actual outcomes. Additionally, it supports forecasting and scenario building based on growth prospects, making explicit assumptions about technological, demographic, and institutional trends.

In spite of its advantages, the significant uncertainty surrounding the estimation of its components, the large amount of data required for estimation—some of which may not be reliable—and the process of decomposing total factor productivity by applying filtering techniques to the Solow residual, inheriting the drawbacks of statistical approaches, render this method not entirely accepted ([Mishkin, 2007](#); [Altăr et al., 2010](#)).

On the other hand, DSGE models have been developed since the publication of the work by [Kydland and Prescott \(1982\)](#). These models provide more realistic and theoretically elegant representations of the economy, allowing for a different perspective on potential output as production in the absence of inefficiencies caused by price and wage rigidities. This theoretical advantage makes them more informative than other techniques when designing economic policies. However, they also present several disadvantages. Firstly, DSGE models require statistical filtering techniques to obtain the series' cycles used as input in their development. Moreover, [Anzoategui and Kim \(2021\)](#) notes that these models require the estimation of a large set of parameters and shocks, thus raising

concerns about parameter identification. Finally, the results are model-specific, as evidenced by the contradictory conclusions between the studies of [Andrés, Lopez-Salido, and Nelson \(2005\)](#), [Neiss and Nelson \(2005\)](#), and [Edge, Kiley, and Laforte \(2008\)](#), and the studies of [Sala, Söderström, and Trigari \(2008\)](#), [Galí, Smets, and Wouters \(2012\)](#), and [Justiniano, Primiceri, and Tambalotti \(2013\)](#) regarding the improvement of potential output estimations compared to other techniques.

The wide variety of techniques for estimating output has led to the ongoing debate, and authors like [Mishkin \(2007\)](#) and [Altar et al. \(2010\)](#) suggest estimating several of these methods and then using researchers' or policymakers' criteria to choose the best among them. However, in practice, the production function method is the most commonly used for estimating potential output, as supported by various organizations such as the U.S. Congressional Budget Office ([Shackleton, 2018](#)) or the European Commission ([D'Auria et al., 2010](#); [Havik et al., 2014](#)).

### 3 Methodology

The unobservable nature of potential GDP has hindered the existence of a consensus on which methodology or technique is most suitable for its estimation. In fact, authors like [Darvas et al. \(2003\)](#) argue that there are no solid grounds for selecting one method over others because all of them face criticism from various aspects. Therefore, following an economic theoretical approach, a methodology is proposed to select and evaluate the relevance of the main techniques used to estimate potential GDP, considering their short and long-term implications. For short-term analysis, based on the works of [Darvas et al. \(2003\)](#) and [Giorno et al. \(1995\)](#), an indicator based on the output gap is proposed to measure the relevance of the various techniques for estimating potential GDP. On the other hand, for long-term analysis, the cointegration technique will be used as a basis to measure the differences between different methods.

In this article, the main techniques for estimating potential GDP were considered. From a statistical approach, filters by [Hodrick and Prescott \(1997\)](#), [Christiano and Fitzgerald \(2003\)](#), [Baxter and King \(1999\)](#), [Butterworth \(1930\)](#), and [Kalman \(1960\)](#) were estimated. The description of the estimation methodology is briefly presented at the end of the article, in the appendix. From economically grounded filters, the production function method was considered. The estimation of potential GDP using this methodology was obtained from FRED (Federal Reserve Bank of St. Louis) and is detailed in [Shackleton \(2018\)](#), therefore it will not be presented in this article.

### 3.1 Short-term analysis: *Variation Index of the Output Gap*

The importance of potential GDP derives from its use in macroeconomic models by policymakers to make decisions. An example of this is the Taylor rule by Taylor (1993) or the methodology to estimate government structural budget balances (Giorno et al., 1995). It is natural to consider using such models to construct a methodology that evaluates the performance of different estimations of potential GDP, allowing the development of a theoretical criterion to select the most suitable estimation technique.

Antecedents of this can be found in the work of Giorno et al. (1995), who use the output gap to compare the differences between estimating potential GDP using the production function and the filter by Hodrick and Prescott (1997), in order to choose the most suitable technique for estimating the structural budget balances of the OECD governments (Organisation for Economic Co-Operation and Development). Although their analysis allows for the discovery of several significant differences both qualitatively and quantitatively between estimating potential GDP using the production function or the filter by Hodrick and Prescott (1997), it lacks a clear frame of reference, and many of its results refer to statistical characteristics rather than the consequences in estimating structural budget balances.

A different approach, based on the discrepancy generated when estimating potential GDP using different sample sizes, is proposed by Darvas et al. (2003). This approach suggests the construction of an index, termed the average revision by the authors, which allows not only the evaluation of various techniques for estimating potential GDP but also the construction of a new estimation technique: the consensus potential GDP. This consensus GDP is nothing more than a weighted average of all potential GDP estimation techniques, using weights constructed from the authors' proposed average revision indicator. Although this methodology proposes a clearer and more concise criterion than that proposed by Giorno et al. (1995), it reduces the criterion to a statistical property of the techniques: the best technique for estimating potential GDP is the one that is more robust to changes in sample size.

To overcome the limitations of both methodologies, a new approach is proposed that combines the best of both worlds. It aims to clearly and concisely determine the best estimation of potential GDP, not solely relying on statistical arguments but also economic theory. Inspired by the statistical metric, mean absolute deviation, and the work of Darvas et al. (2003), the proposal involves constructing an index based on the total magnitude of the output gap, denoted as the *Variation Index of the Output Gap* (VIOG). Specifically, for the

potential GDP estimation technique  $j$ , the  $VIOG$  is defined as:

$$VIOG_j = \frac{1}{T} \sum_{t=1}^T \left| q_t - \bar{q}_t^j \right| \quad (1)$$

where  $T$  is the sample size,  $q_t$  is the GDP at time  $t$ , and  $\bar{q}_t^j$  is the estimated potential GDP using technique  $j$  at time  $t$ . Note that the  $VIOG$  is greater when there is a larger discrepancy between the GDP and the estimated potential GDP. Thus, techniques that better fit the GDP series will have a lower  $VIOG$ , while techniques that do not fit the GDP series as well will have a higher  $VIOG$ .

Under this indicator, *a priori*, the “best” technique is not the one with a higher or lower  $VIOG$ , but rather it depends on the researcher’s or policy maker’s objective or the model being used. To illustrate this, consider the rule by Taylor (1993)<sup>1</sup>. This rule calculates the federal funds rate based on the gaps between current inflation and the target inflation, as well as between the current GDP and the potential GDP. According to this rule, larger output gaps suggest a more aggressive monetary policy compared to when the output gap is smaller. If the policymaker seeks a more aggressive monetary policy, the best estimation of potential GDP will be the one with the higher  $VIOG$ . Conversely, if their goal is a less aggressive monetary policy, then the best estimation of potential GDP will be the one with a lower  $VIOG$ . This analysis can be easily applied to other rules or models in which the output gap plays a significant role or can be used to determine the preferable scenario.

Now, following the ideas of Darvas et al. (2003), from the proposed indicator, it is possible to construct a new estimator for potential GDP as the weighted average of potential GDP from all available techniques, with weights constructed based on  $VIOG$ . Once again, the researcher’s or policy maker’s criterion determines the construction of the weights. If one wishes to value more the techniques that allow for a smaller average output gap, then the weight for technique  $j$  would be:

$$w_j = \frac{\frac{1}{VIOG_j}}{\sum_{i=1}^I \frac{1}{VIOG_i}}$$

where  $I$  is the total number of potential GDP estimation techniques available. On the other hand, if techniques that, on average, allow for a larger output gap are valued more, then the weight for technique  $j$  is:

$$w_j = \frac{VIOG_j}{\sum_{i=1}^I VIOG_i}$$

thus, the new estimated potential GDP would be:

$$\bar{q}_t = \sum_{i=1}^I w_i \bar{q}_t^i$$

---

<sup>1</sup>A detailed explanation of this rule is provided at the end of the article in the Appendix.

### 3.2 Long-term analysis: *Cointegration*

The methodology introduced earlier based on *VIOG* compares potential GDP estimation techniques by their differences in estimating the output gap, focusing its analysis on the GDP cycle or short-term variations. However, the significance of potential GDP lies not only in short-term analysis but also in long-term assessments, reflecting the trend. For this purpose, the most appropriate methodology is cointegration.

According to [Engle and Granger \(1987\)](#), a vector of  $n$  variables  $x_t = (x_{1,t}, \dots, x_{n,t})$  is said to be cointegrated of order  $d - b$  ( $CI(d, b)$ ) if:

- All components of  $x_t$  are integrated of order  $d$ .
- There is at least one vector  $\beta = (\beta_1, \dots, \beta_n)$  such that  $\beta x_t = \beta_1 x_{1,t} + \dots + \beta_n x_{n,t}$  is integrated of order  $d - b$ , where  $b > 0$ . The vector  $\beta$  is called the cointegration vector.

If the vector  $x_t$  is cointegrated, then the variables that constitute it share common stochastic trends ([Stock & Watson, 1988](#)). This allows interpreting cointegration as long-term equilibriums, implying that the degree of deviation from the long-term equilibrium determines the temporary paths of the cointegrated variables. Essentially, if the system is away from equilibrium, some variables must adjust to restore the system to equilibrium. This interpretation allows considering cointegration models as error correction models. Formally, a vector of variables  $x_t = (x_{1,t}, \dots, x_{n,t})$  integrated of order 1 has an error correction representation if it can be expressed as:

$$\Delta x_t = \pi_0 + \pi x_{t-1} + \pi_1 \Delta x_{t-1} + \dots + \pi_p \Delta x_{t-p} + \varepsilon_t$$

here,  $\pi_0$  is a vector of intercepts of size  $n$ ;  $\pi_i$  are coefficient matrices of dimension  $n \times n$ ;  $\pi$  is a matrix with at least one non-zero element; and  $\varepsilon_t$  is a vector of possibly correlated errors. If there is an error correction representation of the variables that constitute  $x_t$ , then necessarily there is a linear combination of  $I(1)$  variables that is stationary. By rearranging  $x_t$  from the error correction representation:

$$\pi x_{t-1} = \Delta x_t - \pi_0 - \pi_1 \Delta x_{t-1} - \dots - \pi_p \Delta x_{t-p} - \varepsilon_t$$

Since the right-hand side of the above expression is stationary, then  $\pi x_{t-1}$  must be stationary, i.e., it must be cointegrated because, as shown by [Engle and Granger \(1987\)](#), if  $x_t$  is a vector of  $I(1)$  variables, the rank of the matrix  $\pi$  is  $0 \leq r < n$ , where  $r$  is the number of linearly independent cointegration vectors. If the variables in  $x_t$  cointegrate ( $r \neq 0$ ), due to Granger's representation theorem, the error correction model and cointegration are equivalent representations ([Engle & Granger, 1987](#); [Johansen, 1991](#)). Hence, it is possible to use the rank of  $\pi$  to determine if a vector of variables is cointegrated ([Johansen, 1988](#); [Stock & Watson, 1988](#)).

There are various ways to assess cointegration; however, the most commonly used methods are the Engle-Granger procedure (Engle & Granger, 1987), which aims to determine if the residuals of the equilibrium relationship are stationary, and the methodology proposed by Johansen (1988), which employs the rank of the matrix  $\pi$ . While both options are viable, the Johansen (1988) methodology has been more widely accepted in the literature because it offers a more comprehensive evaluation of cointegration by considering multiple cointegrating vectors and providing more robustness in its methodology. This is because the Engle-Granger procedure has two main difficulties: estimating the long-term equilibrium regression requires designating one variable as endogenous and the others as exogenous. Depending on the variable chosen, it's possible to obtain different results in the cointegration test, which is not desirable; and the procedure is conducted in two stages: the first stage estimates the residuals, and the second uses these estimated residuals to estimate the regression used to evaluate cointegration, therefore, any error in the first stage can carry over to the second, potentially affecting the results.

Now, as mentioned earlier, assessing cointegration using the methodology of Johansen (1988) boils down to evaluating the rank of the matrix  $\pi$ . To do this, the matrix  $\pi$  is estimated using the error correction model. Then, the rank of  $\pi$  is assessed using the significance of the estimated characteristic roots since the rank of a matrix is equal to the number of characteristic roots that are different from zero. This test can be carried out using the following statistics:

$$\begin{aligned}\lambda_{traza}(r) &= -T \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i) \\ \lambda_{max}(r, r+1) &= -T \ln(1 - \hat{\lambda}_{r+1})\end{aligned}$$

where  $\hat{\lambda}_i$  are the estimated values of the characteristic roots obtained from the estimation of  $\pi$ ;  $T$  is the number of usable observations;  $n$  is the total number of potentially cointegrated variables;  $r$  is the number of possible cointegration relations. The first statistic evaluates the null hypothesis that the number of different cointegration relationships is less than or equal to  $r$ , against the general alternative hypothesis, which is that it is greater than  $r$ . On the other hand, the second statistic tests the null hypothesis that the number of different cointegration relationships is  $r$ , against the alternative hypothesis that it is  $r+1$ . In both cases, the aim is to assess whether the statistics are statistically different from zero for cointegration to exist.

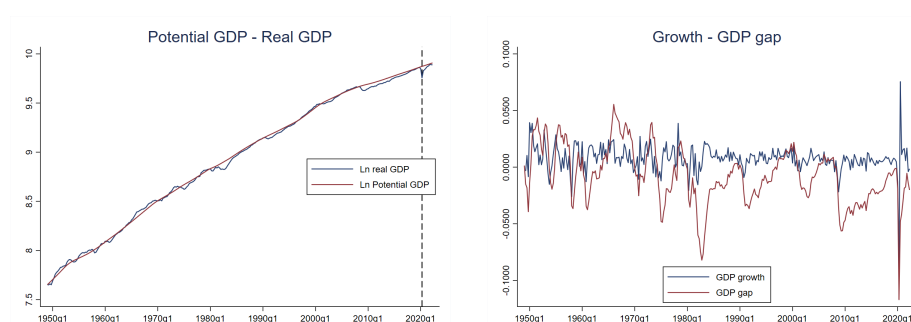
## 4 Results

### 4.1 Data

The data retrieved from FRED (Federal Reserve Bank of St. Louis) on real GDP and potential GDP was used, with the latter estimated by the Congressional

Budget Office (CBO) using the production function method (see [Shackleton, 2018](#)). The series are in quarterly frequency, spanning from the first quarter of 1949 to the second quarter of 2022, totaling 294 observations. As usual, the series were considered in logarithms. In [Figure 1](#), the graphs of the logarithm of real GDP and potential GDP, as well as real growth and the output gap or cycle, can be observed. The joint analysis of the graphs allows us to identify periods of crisis and booms that the American economy has experienced, among which the crisis of 2008 stands out, as well as the most recent one generated by the COVID-19 pandemic.

**Figure 1:** Potential GDP vs Real GDP and GDP growth vs GDP gap.



*Notes:* Graphs of the GDP and the potential GDP in logarithms, as well as the GDP growth and the output gap.

*Source:* Authors' construction with data retrieved from the FRED and the CBO.

The descriptive statistics for growth and the output gap are presented in [table 1](#). As expected from the information in the graphs, both series are centered around zero; moreover, the output gap shows more volatility than the growth.

**Table 1:** Descriptive statistics of the GDP growth and the output gap.

Series	Media	Standard Deviation	Coefficient of Variation	Min	Max
Growth	0,7725%	1,1610%	1,5028	-8,9373%	7,5475%
Gap	-0,7352%	2,5076%	3,4106	-11,6967%	5,5402%

*Source:* Authors' construction with data retrieved from the FRED and the CBO.

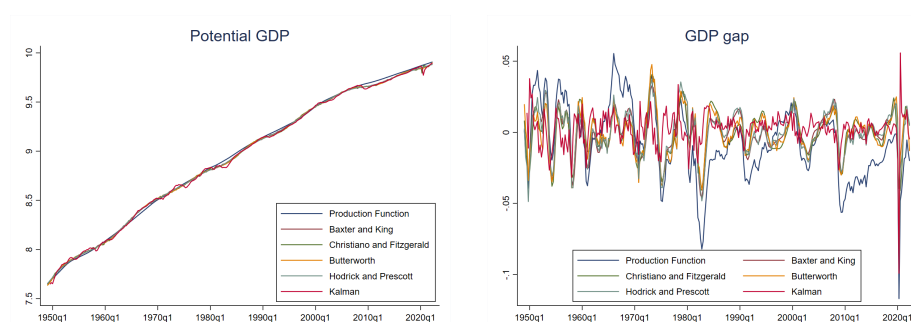
On the other hand, analyzing the economic growth oscillations in more detail, [tables 12](#) and [13](#) are presented (see the appendix), detailing the most pronounced increases and declines in GDP. These refer to quarterly growth or decreases that are more than one standard deviation away from the mean. The most significant periods so far have been the quarters of the first year of the COVID-19 pandemic. In the second quarter of 2022, there was a GDP decline of more than 3 standard deviations (approximately  $-9\%$ ), followed by an equally dramatic production increase or recovery (equivalent to  $7.5\%$ ). Additionally, the years 1950, 1952, 1958, 1978, 1980, 1982, and 2008 stand out, showing quarters with growth/decline between two and three standard deviations from the quarterly

mean. These periods correspond to the most significant crises in the American country: the crisis of 1949, the late '90s energy crisis, and the 2008 subprime crisis.

## 4.2 Estimations of potential GDP

As it was mentioned in the methodology, univariate filters such as [Hodrick and Prescott \(1997\)](#), [Christiano and Fitzgerald \(2003\)](#), [Baxter and King \(1999\)](#), [Butterworth \(1930\)](#), and [Kalman \(1960\)](#) were considered. The details of the estimations for each technique are provided in the appendix. Figure 2 displays the trends and gaps of the various estimated filters, including the production function:

**Figure 2:** Potential GDP and output gap.



*Notes:* Potential GDP and output gap estimated by different methodologies. GDP and potential GDP series were considered in logarithms.

*Source:* Authors' construction with data retrieved from the FRED and the CBO.

Note that depending on the methodology, both the estimates of potential GDP and the output gap change significantly. The potential GDP estimated by the production function is a methodology more akin to the Keynesian conception of potential GDP, based on the idea that potential GDP is the maximum level of production that can be achieved. Therefore, it is not surprising that the estimation by this technique tends to be higher than the rest of the estimates, that the output gap is much more volatile, and of greater magnitude, as observed in Table 2, which displays production gap estimates as a percentage of potential GDP for each of the methods. On the other hand, statistical filters, being consistent with the neoclassical conception that potential GDP is the long-term trend of actual GDP, present less volatile output gaps, and their estimated levels of potential GDP are below the potential GDP estimated by the production function. It is worth noting that the [Kalman \(1960\)](#) filter, unlike the other filters, appears to generate less volatility in the output gap. This evidence is natural because this filter is based on a recursive algorithm that allows updating the estimation as new information becomes available (see [Harvey, 1989](#)).

**Table 2:** Output gap based on different methodologies.

Method/Year	1950	1952	1958	1978	1980	1982	2008	2020
Production Function	0,1601%	0,2722%	-0,3197%	0,1355%	-0,2496%	-0,7866%	-0,1147%	-0,5707%
Baxter and King		0,1577%	-0,3649%	0,1950%	-0,0039%	-0,4128%	0,1064%	
Christiano and Fitzgerald	-0,1179%	0,0687%	-0,2854%	0,1775%	-0,0135%	-0,3673%	0,0713%	-0,3229%
Butterworth	0,0039%	-0,0284%	-0,3311%	0,1684%	0,0042%	-0,3462%	0,0610%	-0,2602%
Hodrick and Prescott	-0,0402%	0,1144%	-0,3381%	0,2486%	0,0217%	-0,4173%	0,1055%	-0,3429%
Kalman	0,3066%	-0,0687%	-0,0328%	0,1366%	-0,0735%	-0,1157%	-0,1043%	-0,1141%

*Notes:* Output gap as a percentage of potential GDP by each of the considered estimation methods for the years that presented any quarterly growth deviation of more than two standard deviations from the mean. For the Baxter and King filter, the values for the years 1950 and 2020 are not presented due to the truncation necessary to estimate the filter.

*Source:* Authors' construction using potential GDP estimates from the production function method and the filters of Baxter and King, Christiano and Fitzgerald, Butterworth, Hodrick and Prescott, and Kalman.

### 4.3 Variation Index of the Output Gap

In the previous subsection, the estimations of each of the potential GDP estimation techniques considered for this work were presented. Next, we will describe what was outlined in the methodology as short-term analysis: the estimation of the *Variation Index of the Output Gap (VIOG)*. The results of the *VIOG* estimation as well as the two types of weights are presented in Table 3.

**Table 3:** *VIOG* estimation.

Method	VIOG	1/VIOG	VIOP weights	1/VIOG weights
Production Function	0,0225	44,4388	27,4734%	9,3529%
Baxter and King	0,0113	88,5541	14,3092%	18,6378%
Christiano and Fitzgerald	0,0128	78,0122	15,6499%	16,4191%
Butterworth	0,0130	76,6717	15,9235%	16,1369%
Hodrick and Prescott	0,0133	74,7414	16,3348%	15,7307%
Kalman	0,0089	112,7137	10,8317%	23,7226%

*Notes:* Results of the *VIOG* estimation and the weights computed based on the latter.

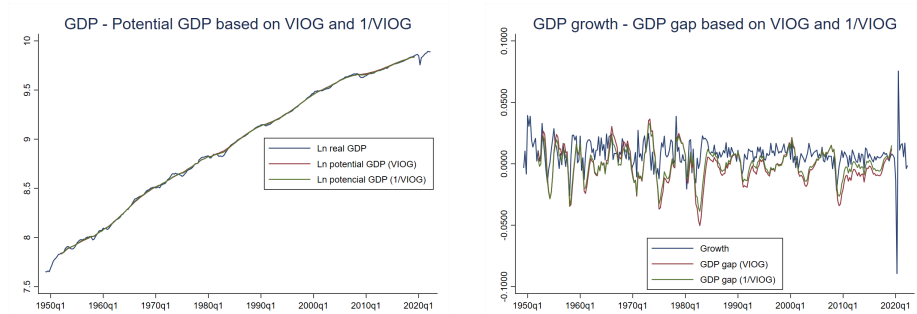
*Source:* Authors' construction based on the estimations of the *VIOG*.

The estimations suggest that if the weights constructed from *VIOG* are considered, the best method would be the one that allows for a greater output gap, in this case, the production function. However, as indicated in the methodology, choosing the method that maximizes the permissiveness in the output gap is not the only alternative. It is possible to change the criterion for consideration to the opposite, the method that allows for a smaller gap, that is, the one that is more stable, in which case the best method would be the [Kalman \(1960\)](#) filter. Regarding the rest of the techniques, the [Hodrick and Prescott \(1997\)](#), [Christiano and Fitzgerald \(2003\)](#), and [Butterworth \(1930\)](#) filters stand out, maintaining a moderate performance according to either of the two selection criteria, emerging as good candidates if, instead of choosing between allowing the largest or smallest output gap, an intermediate point is sought. This result could be considered as an empirical argument in favor of using the [Hodrick and Prescott \(1997\)](#) filter, in spite of the various criticisms it has received from a

theoretical perspective (see [Hamilton, 2018](#)), a fact also noted by [Darvas et al. \(2003\)](#) based on the results they obtained using their proposed index.

Two new versions of potential GDP based on  $VIOG$  and its reciprocal  $1/VIOG$  were estimated from their corresponding constructed weights. Figure 3 compares these two new versions of the potential GDP and the output gap with the real GDP and the GDP growth, respectively. There are no significant differences between the two versions in both potential GDP and the output gap. This outcome is not entirely surprising, as averaging tends to neutralize the differences observed due to the various techniques used to estimate potential GDP and its gap (see Figure 2).

**Figure 3:** GDP vs potential GDP and GDP growth vs output gap based on  $VIOG$  and  $1/VIOG$  estimations.



*Notes:* Potential GDP and output gap estimated from the  $VIOG$  and  $1/VIOG$  weights against GDP and GDP growth respectively.

*Source:* Authors' construction based on the estimations of the  $VIOG$ .

**Table 4:** Output gap based on  $VIOG$  estimations.

Method/Year	1952	1958	1978	1980	1982	2008
$VIOG$	0,1139%	-0,2944%	0,1741%	-0,0751%	-0,4668%	0,0099%
$1/VIOG$	0,0631%	-0,2593%	0,1768%	-0,0397%	-0,3601%	0,0224%

*Notes:* Output gap as a percentage of potential GDP for each of the  $VIOG$  weighted averages for the years presented between two and three quarterly standard deviations from the average growth. The years 1950 and 2020 were not taken into account due to the truncation of the Baxter and King filter.

*Source:* Authors' construction based on the estimations of the  $VIOG$ .

Likewise, as was done with each of the filters, the annual average gaps estimated as the average of the estimated potential GDP are displayed in table 4. The estimations do not present significant differences between the two weighted averages of potential GDP, confirming what was previously mentioned: they are estimations without remarkable distinctions.

#### 4.4 Testing cointegration

Before proceeding with policy analysis, a long-term analysis is presented: cointegration. To carry this out, the order of integration of the real GDP and potential GDP from each of the techniques is first evaluated using the most well-known tests: the Dickey-Fuller test (Dickey & Fuller, 1979) and the Phillips-Perron test (Phillips & Perron, 1988)<sup>2</sup>. Table 5 presents the results for the series in levels, while Table 6 shows the results for the first differences. From the tests, it can be concluded that the series are integrated of order 1.

**Table 5:** Unit root tests.

Variable	Phillips-Perron	Augmented Dickey-Fuller
GDP	0,8653	0,8986
Production Function	0,9968	1,0000
Baxter and King	1,0000	0,9970
Christiano and Fitzgerald	0,9653	0,9414
Butterworth	0,8770	0,8123
Hodrick and Prescott	0,9882	0,9965
Kalman	0,7018	0,7892

*Notes:* Unit root tests for variables in logarithms. The p-values of both tests are presented in this table.

*Source:* Authors' construction based on the estimations of the Phillips-Perron and Augmented Dickey-Fuller unit root tests.

**Table 6:** Unit root test for first differences.

Variable	Phillips-Perron	Augmented Dickey-Fuller
GDP	0,0000	0,0000
Production Function	0,1873	0,0023
Baxter and King	0,0000	0,0001
Christiano and Fitzgerald	0,0000	0,0000
Butterworth	0,1065	0,0009
Hodrick and Prescott	0,0365	0,0031
Kalman	0,0000	0,0000

*Notes:* Unit root tests for variables in differences. The p-values of both tests are presented in this table.

*Source:* Authors' construction based on the estimations of the Phillips-Perron and Augmented Dickey-Fuller unit root tests.

Once the order of integration has been determined, the optimal lags to include in the cointegration test by Johansen (1988) are estimated. As the series are quarterly, up to 4 possible lags are evaluated. Table 7 shows the results of various tests to determine the optimal number of lags, which turns out to be 4 according to all methods considered<sup>3</sup>.

<sup>2</sup>The null hypothesis for both unit root tests is non-stationarity.

<sup>3</sup>The methods considered are: likelihood ratio (LR), final prediction error (FPE), Akaike in-

**Table 7:** Optimal lags tests.

Lag	LL	LR test statistic	P-value	FPE	AIC	HQIC	SBIC
0	5.540,5200			0,0000	-41,6054	-41,5675	-41,5111
1	9.346,7300	7.612,4000	0,0000	0,0000	-69,8551	-69,5520	-69,1007
2	11.786,3000	4.879,1000	0,0000	0,0000	-87,8293	-87,2610	-86,4148
3	12.652,6000	1.732,5000	0,0000	0,0000	-94,0268	-93,2312	-92,0465
4	13.381,3000	1.457,5000	0,0000*	0,0000*	-99,1903*	-98,1674*	-96,6441*

*Notes:* Tests to determine the optimal lags for the cointegration test. The log likelihood (LL), the test statistic and the p-value are presented for the LR test. The lags marked with \* are the optimal lags according to each method.

*Source:* Authors' construction based on the estimations of the optimal lags tests: likelihood ratio, final prediction error, Akaike information criterion, Schwarz Bayesian information criterion, and Hannan-Quinn information criterion.

With the determined number of lags, the cointegration test by [Johansen \(1988\)](#) was conducted, considering both the trace and the maximum eigenvalue tests, as well as information criteria. [Table 8](#) presents the test results. According to the [Johansen \(1988\)](#) tests and the Schwarz Bayesian information criterion (SBIC), there are 5 cointegration relationships, while according to the Hannan-Quinn information criterion (HQIC), there could be up to 6 relationships.

**Table 8:** Cointegration tests.

Max lag	Eigen value	Trace statistic	Critic value		Max statistic	Critic value		SBIC	HQIC	AIC
			5%	1%		5%	1%			
0		1.132,7777	124,2400	133,5700	486,0843	45,2800	51,5700	-100,7815	-102,0227	-102,8562
1	0,8392	646,6933	94,1500	103,1800	359,5100	39,3700	45,1000	-102,3360	-103,6820	-104,5858
2	0,7412	287,1834	68,5200	76,0700	183,6155	33,4600	38,7700	-103,4567	-104,8913	-105,8546
3	0,4986	103,5679	47,2100	54,4600	45,9991	27,0700	32,2400	-103,9580	-105,4652	-106,4773
4	0,1588	57,5688	29,6800	35,6500	37,7518	20,9700	25,5200	-103,9840	-105,5476	-106,5976
5	0,1323	19,8170*	15,4100	20,0400	14,4583*	14,0700	18,6300	-104,0210*	-105,6249	-106,7019
6	0,0529	5,3587	3,7600	6,6500	5,3587	3,7600	6,6500	-104,0124	-105,6404*	-106,7337
7	0,0199							-104,0115	-105,6476	-106,7463

*Notes:* Optimal values are marked with \*.

*Source:* Authors' construction based on the estimations of the cointegration tests: trace test, max statistic test, Akaike information criterion, Schwarz Bayesian information criterion, and Hannan-Quinn information criterion.

Regardless of the number of cointegration relationships that exist, the existence of cointegration implies convergence to common stochastic trends ([Stock & Watson, 1988](#)). This implies that in the long term, the potential GDP estimated by any of the estimation techniques is statistically the same, an argument that shifts the discussion about the optimal methodology for estimating national production towards the short term.

## 4.5 Policy implications

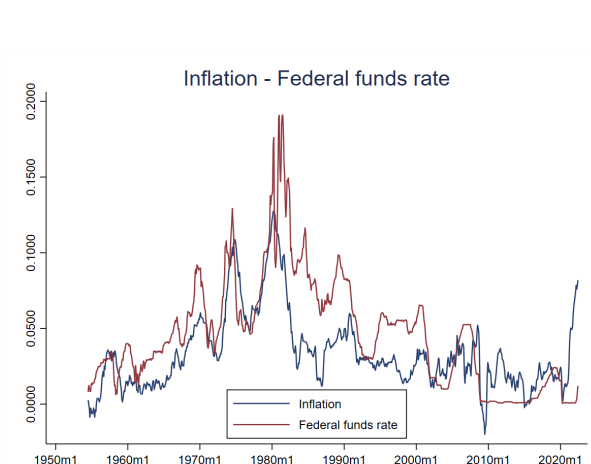
So far, the performance of each technique has been evaluated from a highly statistical standpoint. However, the goal of economic policy is applicability, the potential of these techniques to aid in the decision-making process for policy

formation criterion (AIC), Schwarz Bayesian information criterion (SBIC), and Hannan-Quinn information criterion (HQIC). For further details on each methodology, refer to [Lütkepohl \(2005\)](#).

makers, who in many cases focus more on the usefulness and practical consequences of these techniques than their statistical properties. For this reason, and as mentioned earlier, the short-term analysis will be extended using the Taylor (1993) rule.

To conduct the analysis based on Taylor (1993), data on inflation was retrieved from the US Bureau of Labor Statistics, calculated from the CPIAUCSL (Consumer Price Index for All Urban Consumers: All Items).<sup>4</sup> Additionally, data of the federal funds rate was collected from the FRED.<sup>5</sup> Figure 4 illustrates the evolution of the series from 1954 to 2022. Within this time frame, significant periods stand out: high inflation during the 1970s due to the oil crisis, the decline and stabilization of inflation throughout the 1980s and 1990s, a significant rise and fall in the early 2000s due to the subprime crisis, and a high inflation period following the COVID-19 crisis.

**Figure 4:** Inflation vs Federal funds rate.



Source: Authors' construction based on the data retrieved from the US Bureau of Labor Statistics and the FRED.

It is necessary to highlight the way the federal funds rate moves as inflation evolves: periods of high inflation are characterized by periods with very high federal funds rates, while periods of falling inflation are marked by declines in the federal funds rate. This reflects the outcomes of monetary policy in the United States. Table 9 presents the descriptive statistics of both series. As expected *a priori*, the federal funds rate is much more volatile than inflation. Additionally, it's noteworthy that the average inflation has been approximately around 3%, while the federal funds rate is 100 basis points higher.

<sup>4</sup>For more details on the methodology used for the index calculation, please visit <https://fred.stlouisfed.org/series/CPIAUCSL>.

<sup>5</sup>For further information, please visit <https://fred.stlouisfed.org/series/FEDFUNDS>.

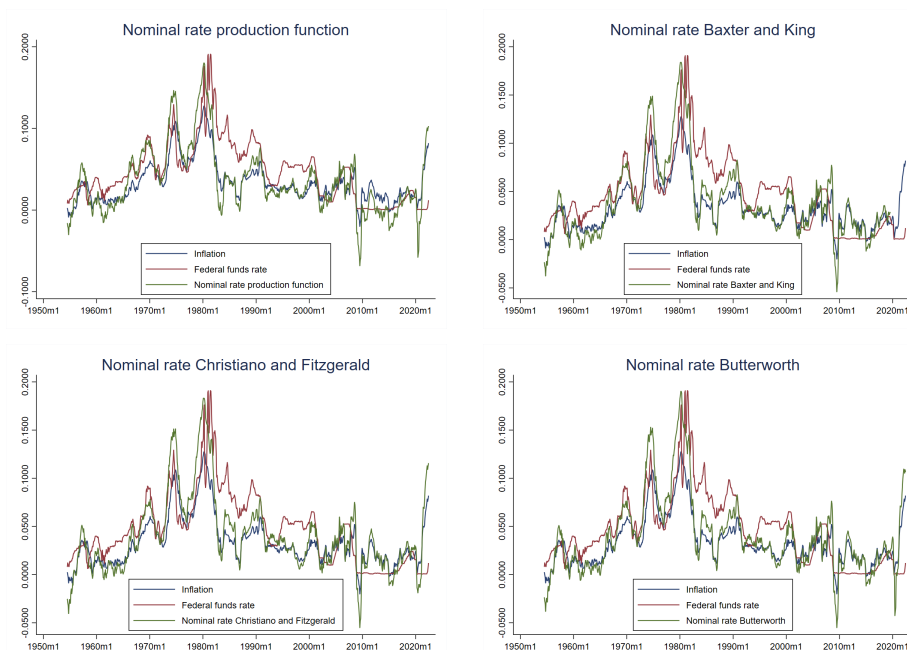
**Table 9:** Descriptive statistics of the inflation and the federal funds rate.

Variable	Mean	Standard Deviation	Coefficient of Variation	Min	Max
Federal funds rate	4,6078%	3,6277%	0,7873	0,0500%	19,1000%
Inflation	3,3623%	2,5285%	0,7520	-1,9979%	12,7341%

*Source:* Authors' construction based on the data retrieved from the US Bureau of Labor Statistics and the FRED.

The calculation of the nominal interest rate using Taylor's rule (Taylor, 1993) by different methodologies to estimate the output gap is summarized in figures 5 and 6, and table 10. Graphically, the results are quite similar regardless of the estimated output gaps. In general, the federal funds rate tends to cover not only inflation but also each of the nominal rates, leading to the conclusion that Taylor's rule tends to underestimate the nominal rate concerning the observed data. Additionally, the evolution of the nominal rate calculated using the output gap estimated by the production function differs slightly from the other nominal rates, showing more discrepancy from the inflation behavior. This confirms the previously obtained results that the production function method allows for more gap, translating into less adjustment concerning inflation. Finally, it is worth noting that unlike the federal funds rate, which considers practical criteria such as the non-negativity of the rate, the estimates by Taylor's rule have no restrictions and can take negative values.

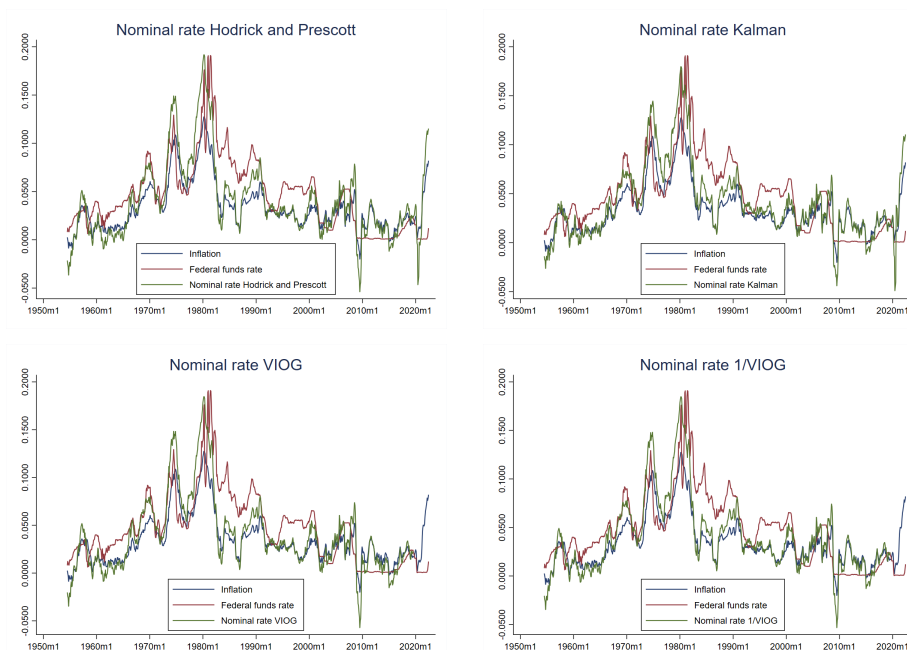
**Figure 5:** Nominal rate using different potential GDP methodologies.



*Notes:* Nominal rate calculated using the Taylor rule and different estimates of the output gap, including: production function, Baxter and King, Christiano and Fitzgerald, and Butterworth. *Source:* Authors' construction based on the data retrieved from the US Bureau of Labor Statistics and the FRED, and the potential GDP estimations.

The results from Table 10 suggest that the method that best fits the observed data is the Kalman (1960) filter. In all the descriptive statistics shown, the Kalman (1960) filter is the one that closely matches the statistics of the federal funds rate. Considering the production function method, its higher variability is notable, with the estimated nominal rate averaging half a percentage point less than most of the other nominal rates. This difference is quite significant and demonstrates the potential difference that can arise in the policy rate according to the Taylor's rule and the technique chosen to calculate the output gap. These discrepancies in the nominal rates can lead to a policy being more or less aggressive, resulting in significantly different outcomes in economies, as it was highlighted previously.

**Figure 6:** Nominal rate using different potential GDP methodologies.



*Notes:* Nominal rate calculated using the Taylor rule and different estimates of the output gap, including: Hodrick and Prescott, Kalman, *VIOG*, and  $1/VIOG$ .

*Source:* Authors' construction based on the data retrieved from the US Bureau of Labor Statistics and the FRED, and the potential GDP estimations.

**Table 10:** Descriptive statistics of the nominal rates calculated by the different output gap estimation techniques.

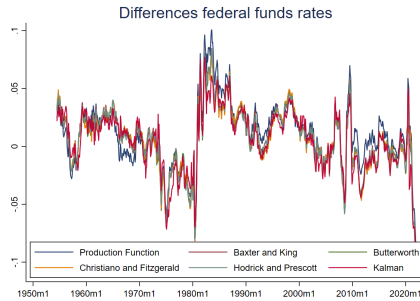
Variable	Mean	Standard Deviation	Coefficient of Variation	Min	Max
Production Function	3,5968%	4,0383%	1,1228	-6,8144%	18,0139%
Baxter and King	3,9851%	3,9678%	0,9957	-5,3933%	18,3905%
Christiano and Fitzgerald	4,0324%	3,9705%	0,9847	-5,4955%	18,3075%
Butterworth	4,0388%	3,9812%	0,9857	-5,5002%	19,0219%
Hodrick and Prescott	4,0256%	4,0195%	0,9985	-5,3806%	19,1661%
Kalman	4,1584%	3,7673%	0,9060	-4,8904%	17,9659%
<i>VIOG</i>	3,9232%	3,9193%	0,9990	-5,7061%	18,4549%
$1/VIOG$	4,0096%	3,8875%	0,9695	-5,3204%	18,4644%

*Source:* Authors' construction based on the data retrieved from the US Bureau of Labor Statistics and the FRED, and the potential GDP estimations.

Finally, the differences between the nominal rates and the federal funds rate are analyzed. In Figure 7, it can be observed that the differences do not seem to be significant among the techniques for estimating the output gap, except for the production function, which shows notable differences in some periods such as the 1960s or the second decade of the 2000s. On the other hand, Table 11 presents the descriptive statistics of these differences. The greatest variability

in this measure is exhibited by the derivative of the [Kalman \(1960\)](#) filter, which has a considerably higher coefficient of variation, whereas the lowest variability is observed in the derivative based on the production function. However, concerning the average, the roles are reversed. The estimation by the [Kalman \(1960\)](#) filter presents the smallest average difference with the federal funds rate, while the production function displays the largest discrepancy with an average difference of half a percentage point.

**Figure 7:** Differences between the federal funds rate and the various nominal rates estimations.



*Source:* Authors' construction based on the data retrieved from the US Bureau of Labor Statistics and the FRED, and the potential GDP estimations.

**Table 11:** Descriptive statistics of the difference between the federal funds rate and the various nominal rates estimations.

Variable	Media	Standard Deviation	Coefficient of Variation	Min	Max
Production Function	1,0110%	2,8609%	2,8297	-9,8721%	10,0699%
Baxter and King	0,8038%	2,5818%	3,2121	-8,1773%	8,4582%
Christiano and Fitzgerald	0,5754%	2,8350%	4,9266	-11,0907%	8,0107%
Butterworth	0,5690%	2,8222%	4,9597	-10,8742%	7,9570%
Hodrick and Prescott	0,5822%	2,8432%	4,8833	-11,0838%	8,4232%
Kalman	0,4495%	2,7321%	6,0785	-10,7686%	7,7095%
VIOG	0,8656%	2,5323%	2,9255	-7,5966%	8,4895%
1/VIOG	0,7792%	2,5106%	3,2221	-7,6859%	7,9200%

*Source:* Authors' construction based on the data retrieved from the US Bureau of Labor Statistics and the FRED, and the potential GDP estimations.

## 5 Conclusions

In this article, the empirical implications of different potential GDP estimation techniques were studied, both in the short and long term. The long-term analysis focused on assessing the cointegration between the estimated techniques. The results of this analysis indicate that the series, integrated of order 1, are cointegrated, implying convergence towards a common stochastic trend ([Stock & Watson, 1988](#)). This suggests that in the long term, the debate over which

technique to use becomes unnecessary.

On the other hand, to evaluate the different techniques in the short term, an index measuring the total magnitude of the output gap, the *VIOG*, as well as its inverse  $1/VIOG$ , was proposed. The Taylor's rule was used to contrast the estimation results with real data. The findings indicate that the technique with the highest *VIOG* is the production function, while the one with the lowest *VIOG* is the Kalman (1960) filter. This suggests that for more aggressive approaches allowing a larger output gap, the production function is the appropriate technique, while in more conservative approaches, the Kalman (1960) filter would be the technique to use. Finally, based on the results derived from the Taylor rule analysis, it is inferred that the Kalman (1960) filter has the best fit measured through statistics of real nominal interest rate values such as mean and standard deviation.

The conclusions drawn from the long and short-term analyses allow us to affirm that beyond theoretical discussions, empirically the various potential GDP estimation techniques differ significantly in the short term. Therefore, policy makers must be aware of the consequences this might entail for economic policies where potential GDP is used as an input. For instance, for the Taylor's rule, average differences of up to 50 basis points are found in the estimated nominal interest rate among the different techniques.

This study stands out from others in three main aspects. Firstly, the combination of diverse methodologies for analyzing the different potential GDP estimation techniques. Secondly, it adopts an empirical rather than theoretical approach. It focuses not on the theoretical discussion between potential GDP and trend GDP but seeks to show the practical effects of considering one potential product estimation technique over another. Lastly, it proposes an index (*VIOG*) that helps determine which technique is more suitable depending on policy makers' objectives, allowing the proposal of a new potential product estimation as a weighted average of potential GDP techniques, using *VIOG* for constructing the weight.

In conclusion, regarding to potential GDP, unresolved aspects include empirical results for fiscal policies, analysis of the performance of more innovative techniques such as multivariate approaches and those employing DSGE models, and the development of other methodologies to broaden the empirical analyses of potential product estimations. This makes potential GDP remain an attractive area for exploration in macroeconomics.

## 6 References

Altăr, M., Necula, C., & Bobeică, G. (2010). Estimating potential gdp for the romanian economy. an eclectic approach. *Romanian Journal of Economic*

- Forecasting*, 3, 5–25.
- Andrés, J., Lopez-Salido, J. D., & Nelson, E. (2005). Sticky-price models and the natural rate hypothesis. *Journal of Monetary Economics*, 52(5), 1025–1053.
- Anzoategui, D., & Kim, M. (2021). Re-estimating potential gdp: new evidence on output hysteresis. *Mimeo.*
- Baxter, M., & King, R. G. (1999). Measuring business cycles: approximate band-pass filters for economic time series. *Review of economics and statistics*, 81(4), 575–593.
- Beveridge, S., & Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘business cycle’. *Journal of Monetary Economics*, 7(2), 151–174.
- Box, G. E. P., & Jenkins, G. M. (1970). Time series analysis, forecasting and control.
- Burns, A., Van Rensburg, T. J., Dybczak, K., & Bui, T. (2014). Estimating potential output in developing countries. *Journal of Policy Modeling*, 36(4), 700–716.
- Butterworth, S. (1930). On the theory of filter amplifiers. *Wireless Engineer*, 7(6), 536–541.
- Carnazza, G., Fontanari, C., Liberati, P., & Palumbo, A. (2021). From potential gdp to structural balance: A theoretical reassessment and new evidence for italy. *Review of Political Economy*, 1–31.
- Chen, G.-m., Firth, M., & Rui, O. M. (2002). Stock market linkages: evidence from latin america. *Journal of Banking & Finance*, 26(6), 1113–1141.
- Christiano, L. J., & Fitzgerald, T. J. (2003). The band pass filter. *international economic review*, 44(2), 435–465.
- Clark, P. K. (1987). The cyclical component of us economic activity. *The Quarterly Journal of Economics*, 102(4), 797–814.
- Darvas, Z., Vadas, G., et al. (2003). *Univariate potential output estimations for hungary* (Tech. Rep.). Magyar Nemzeti Bank (Central Bank of Hungary).
- D’Auria, F., Havik, K., Mc Morrow, K., Planas, C., Raciborski, R., Roger, W., . . . others (2010). *The production function methodology for calculating potential growth rates and output gaps* (Tech. Rep.). Directorate General Economic and Financial Affairs (DG ECFIN), European . . .
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427–431.
- Edge, R. M., Kiley, M. T., & Laforde, J.-P. (2008). Natural rate measures in an estimated dsge model of the us economy. *Journal of Economic Dynamics and control*, 32(8), 2512–2535.
- Enders, W. (2008). *Applied econometric time series*. John Wiley & Sons.
- Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276.
- Fontanari, C., Palumbo, A., & Salvatori, C. (2020). Potential output in theory and practice: A revision and update of okun’s original method. *Structural*

- Change and Economic Dynamics*, 54, 247–266.
- Fuller, W. A. (2009). *Introduction to statistical time series*. John Wiley & Sons.
- Galí, J., Smets, F., & Wouters, R. (2012). Unemployment in an estimated new keynesian model. *NBER macroeconomics annual*, 26(1), 329–360.
- Gerali, A., Locarno, A., Notarpietro, A., & Pisani, M. (2018). The sovereign crisis and italy’s potential output. *Journal of Policy Modeling*, 40(2), 418–433.
- Giorno, C., Richardson, P., Roseveare, D., & Van den Noord, P. (1995). Estimating potential output, output gaps and structural budget balances.
- Hamilton, J. D. (2018). Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics*, 100(5), 831–843.
- Harvey, A. C. (1989). Forecasting, structural time series models and the kalman filter.
- Harvey, A. C., & Jaeger, A. (1993). Detrending, stylized facts and the business cycle. *Journal of applied econometrics*, 8(3), 231–247.
- Havik, K., Mc Morrow, K., Orlandi, F., Planas, C., Raciborski, R., Röger, W., ... others (2014). *The production function methodology for calculating potential growth rates & output gaps* (Tech. Rep.). Directorate General Economic and Financial Affairs (DG ECFIN), European ...
- Hodrick, R. J., & Prescott, E. C. (1997). Postwar us business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, 1–16.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2-3), 231–254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: journal of the Econometric Society*, 1551–1580.
- Justiniano, A., Primiceri, G. E., & Tambalotti, A. (2013, April). Is there a trade-off between inflation and output stabilization? *American Economic Journal: Macroeconomics*, 5(2), 1-31. Retrieved from <https://www.aeaweb.org/articles?id=10.1257/mac.5.2.1> doi: 10.1257/mac.5.2.1
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems.
- King, R., & Rebelo, S. (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, 17(1-2), 207-231. Retrieved from <https://EconPapers.repec.org/RePEc:eee:dyncon:v:17:y:1993:i:1-2:p:207-231>
- Kuttner, K. N. (1994). Estimating potential output as a latent variable. *Journal of business & economic statistics*, 12(3), 361–368.
- Kydland, F. E., & Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, 1345–1370.
- Laxton, D., & Tetlow, R. (1992). *A simple multivariate filter for the measurement of potential output* (Vol. 59). Bank of Canada Technical Report No. 59. Ottawa.
- Lucas, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, 1, 19-46. Re-

trieved from <https://www.sciencedirect.com/science/article/pii/S0167223176800036> doi: [https://doi.org/10.1016/S0167-2231\(76\)80003-6](https://doi.org/10.1016/S0167-2231(76)80003-6)

- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- Mills, T. C. (2003). *Modelling trends and cycles in economic time series* (Vol. 10). Springer.
- Mishkin, F. (2007). *Estimating potential output: a speech at the conference on price measurement for monetary policy, federal reserve bank of dallas, dallas, texas, may 24, 2007* (Speech No. 299). Board of Governors of the Federal Reserve System (U.S.). Retrieved from <https://EconPapers.repec.org/RePEc:fip:fedgsq:299>
- Neiss, K. S., & Nelson, E. (2005). Inflation dynamics, marginal cost, and the output gap: Evidence from three countries. *Journal of Money, Credit and Banking*, 1019–1045.
- Okun, A. M. (1962). Potential gnp: Its measurement and its significance. *Proceedings of the American Statistical Association, Business and Economic Statistical Section*, 89–104.
- Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *biometrika*, 75(2), 335–346.
- Pollock, D. (2000). Trend estimation and de-trending via rational square-wave filters. *Journal of Econometrics*, 99(2), 317–334. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0304407600000282> doi: [https://doi.org/10.1016/S0304-4076\(00\)00028-2](https://doi.org/10.1016/S0304-4076(00)00028-2)
- Romer, D. (2012). *Advanced macroeconomics, 4e*. New York: McGraw-Hill.
- Sala, L., Söderström, U., & Trigari, A. (2008). Monetary policy under uncertainty in an estimated model with labor market frictions. *Journal of Monetary Economics*, 55(5), 983–1006.
- Shackleton, R. (2018). *Estimating and projecting potential output using cbo's forecasting growth model*. Congressional Budget Office.
- Stock, J. H., & Watson, M. W. (1988). Testing for common trends. *Journal of the American statistical Association*, 83(404), 1097–1107.
- Taylor, J. B. (1993). Discretion versus policy rules in practice. In *Carnegie-rochester conference series on public policy* (Vol. 39, pp. 195–214).
- US Congressional Budget Office. (2004). A summary of alternative methods for estimating potential gdp.
- Wei, W. (2006). *Ws 2006. time series analysis: Univariate and multivariate methods*. USA: Pearson Educations, Inc.
- Whittaker, E. T. (1922). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, 41, 63–75. doi: 10.1017/S0013091500077853

## 7 Appendix

### 7.1 Tables

**Table 12:** GDP increases and decreases that are at least 1 standard deviation from the mean from 1949 to 1970.

Year	Quarter	Distance	Change (%)
1949	4	1 standard deviation	-0,84
1950	1	2 standard deviations	3,93
1950	2	1 standard deviation	3,05
1950	3	2 standard deviations	3,86
1951	3	1 standard deviation	2,06
1952	4	2 standard deviations	3,28
1953	3	1 standard deviation	-0,56
1953	4	1 standard deviation	-1,51
1954	1	1 standard deviation	-0,48
1954	4	1 standard deviation	1,96
1955	1	1 standard deviation	2,86
1957	4	1 standard deviation	-1,03
1958	1	2 standard deviations	-2,60
1958	3	1 standard deviation	2,31
1958	4	1 standard deviation	2,34
1959	2	1 standard deviation	2,26
1960	1	1 standard deviation	2,25
1960	2	1 standard deviation	-0,54
1960	4	1 standard deviation	-1,28
1961	4	1 standard deviation	1,96
1963	3	1 standard deviation	2,20
1964	1	1 standard deviation	2,11
1965	1	1 standard deviation	2,42
1965	3	1 standard deviation	2,22
1965	4	1 standard deviation	2,30
1966	1	1 standard deviation	2,43
1968	1	1 standard deviation	2,04
1969	4	1 standard deviation	-0,49
1970	4	1 standard deviation	-1,07

Source: Authors' construction based on the data retrieved from the FRED.

**Table 13:** GDP increases and decreases that are at least 1 standard deviation from the average from 1971 to 2022.

<b>Year</b>	<b>Quarter</b>	<b>Distance</b>	<b>Change (%)</b>
1971	1	1 standard deviation	2,72
1972	2	1 standard deviation	2,27
1973	1	1 standard deviation	2,47
1973	3	1 standard deviation	-0,53
1974	1	1 standard deviation	-0,86
1974	3	1 standard deviation	-0,95
1975	1	1 standard deviation	-1,22
1976	1	1 standard deviation	2,25
1977	2	1 standard deviation	1,94
1978	2	2 standard deviations	3,86
1980	2	2 standard deviations	-2,06
1981	1	1 standard deviation	1,96
1981	2	1 standard deviation	-0,74
1981	4	1 standard deviation	-1,09
1982	1	2 standard deviations	-1,55
1983	2	1 standard deviation	2,28
1983	3	1 standard deviation	2,00
1983	4	1 standard deviation	2,09
1984	1	1 standard deviation	1,95
1990	4	1 standard deviation	-0,91
1991	1	1 standard deviation	-0,47
2001	3	1 standard deviation	-0,40
2008	1	1 standard deviation	-0,41
2008	3	1 standard deviation	-0,53
2008	4	2 standard deviations	-2,18
2009	1	1 standard deviation	-1,17
2020	1	1 standard deviation	-1,30
2020	2	3 standard deviations	-8,94
2020	3	3 standard deviations	7,55
2022	1	1 standard deviation	-0,40

*Source:* Authors' construction based on the data retrieved from the FRED.

## 7.2 Taylor's rule

In 1993, Taylor presented a monetary policy rule to be used by central banks to stabilize economic activity by determining short-term interest rates. This rule calculates the federal funds rate based on the gap between current and target inflation and the gap between current and potential GDP:

$$i_t = \pi_t + r_t + a_\pi(\pi_t - \pi_t^*) + a_y(y_t - \bar{y}_t)$$

where  $i_t$  is the nominal interest rate,  $\pi_t$  is the inflation rate,  $\pi_t^*$  is the target inflation rate,  $r_t$  is the real interest rate,  $y_t$  is the natural logarithm of GDP and  $\bar{y}_t$  is the natural logarithm of potential GDP. The coefficients  $a_\pi$  and  $a_y$  must be positive according to the results obtained by Taylor (1993). This restriction suggests that the interest rate should be high when inflation is above its target or when output is above its full employment level to reduce the pressure on prices. In the opposite case, to stimulate output, the rule suggests a low interest rate. In phenomena such as stagflation where inflation is above its target and output below its full employment level the rule is not conclusive and its results will depend on the assignment of values to the weights  $a_\pi$  and  $a_y$ . Taylor (1993) suggests  $a_\pi = a_y = 0.5$ .

## 7.3 Statistical filters

Time series theory considers that a time series  $y_t$  is the combination of different components such as the trend  $\tau_t$ , the cycle  $c_t$  or the seasonality  $s_t$ , among others.

$$y_t = \tau_t + c_t + s_t$$

In practice, only the series are observed; however, in many cases, the components are relevant, especially for economic policy analysis. Statistical methodologies that allow the separation of each component in time series are known as time series filtering techniques. While there are many filtering techniques in the literature, this article will consider the three most commonly used filtering techniques in economic literature: band-pass filters (Baxter and King, 1999; Christiano and Fitzgerald, 2003), high-pass filters (Butterworth, 1930; Hodrick and Prescott, 1997), and the Kalman (1960) filter.

### 7.3.1 High-pass and band-pass filters

The initial techniques for filtering economic time series, such as first differences, isolation of linear and quadratic trends, or transformation by moving averages, aim to ease the calculation of unobservable components and ensure the production of a stationary cyclical component without explicitly examining the statistical characteristics of the obtained cycle (Baxter and King, 1999; Mills, 2003). These techniques are considered from a time domain perspective. However, as mentioned by Wei (2006), an alternative approach for studying economic cycles and their estimation lies in the frequency domain analysis. This domain

describes fluctuations in cycles over time in terms of sinusoidal behaviors at different frequencies, allowing the extraction of statistical information from cycles that is impossible through time domain analysis. For this reason, band-pass and high-pass filters have been developed.

From the time domain, the autocovariances  $\gamma_j$  ( $j \in \{0, 1, \dots\}$ ) of a covariance stationary series,  $y_t$ , reflect its variance and time-dependent structure. However, from the frequency domain perspective,  $y_t$  and its autocovariances are specified in terms of stochastic cycles occurring at different frequencies  $\omega \in [-\pi, \pi]$ . The contribution of these stochastic cycles at each frequency  $\omega$  relative to the variance of  $y_t$  ( $\sigma_y^2$ ), is defined by the spectral density function  $f_y(\omega)$ .

Wei (2006) showed that if the cycle is defined by a linear filter given as:

$$c_t = \sum_{j=-\infty}^{\infty} \alpha_j y_{t-j} = \alpha(L)t_t$$

where  $\alpha(L) = \sum_{j=-\infty}^{\infty} \alpha_j L^j$  is a polynomial in the lag operator  $L$ , and  $L$  is defined such that  $L^j y_t = y_{t-j}$ , then

$$f_c(\omega) = |\alpha(\omega)|^2 f_y(\omega)$$

where  $|\alpha(e^{i\omega})|$  is known as the filter gain, which allows transforming the spectral density of  $y_t$  into the spectral density of  $c_t$ , in other words, it allows observing the impact of  $c_t$  on  $y_t$  at a frequency  $\omega$ .

The objective of frequency domain filters is to transform the series  $y_t$  into a new series  $c_t$ , for which its spectral density function  $f_c(\omega)$  equals 0 for undesired frequencies and equals  $f_y(\omega)$  for desired frequencies. Thus, for band-pass filters:

$$\alpha(\omega) = \begin{cases} 1 & si \quad |\omega| \in [\omega_0, \omega_1] \\ 0 & si \quad |\omega| \notin [\omega_0, \omega_1] \end{cases} \quad (2)$$

while for high-pass filters:

$$\alpha(\omega) = \begin{cases} 1 & si \quad |\omega| \geq \omega_0 \\ 0 & si \quad |\omega| \leq \omega_0 \end{cases} \quad (3)$$

### 7.3.2 Baxter and King filter

Baxter and King (1999) showed that for an infinite time series, there is an ideal filter whose gain function is defined as in Equation 2. However, as economic series are finite, the authors approximated the ideal filter with a moving average

filter:

$$c_t = \sum_{j=-q}^{j=q} \alpha_j y_{t-j}$$

for each  $t \in \{q + 1, \dots, T - q\}$ , where  $\alpha_{-j} = \alpha_j$ , and the coefficients sum to 0. This filter is capable of eliminating deterministic and stochastic trends of order 2 or lower (Fuller, 2009; Baxter and King, 1999).

### 7.3.3 Christiano and Fitzgerald filter

Christiano and Fitzgerald (2003), unlike Baxter and King (1999), approximate the ideal filter by minimizing the mean square error between the estimated component and the ideal component, assuming that the time series of interest is a random walk. As the authors point out, this filter is much more efficient; however, it requires the estimation of an additional parameter and loses some robust properties such as symmetry, which prevents it from removing second-order deterministic and quadratic trends.

### 7.3.4 Hodrick y Prescott filter

An alternative approach to the previous ones is the method initially proposed by Whittaker (1922) and known as the filter by Hodrick and Prescott (1997). This filter estimates the trend and cycle components by solving the optimization problem:

$$\min_{\tau_t} \left\{ \sum_{t=1}^T c_t + \lambda \sum_{t=3}^{T-2} [(\tau_t - \tau_{t-1}) - (\tau_{t-1} - \tau_{t-2})]^2 \right\}$$

Here, the parameter  $\lambda$  penalizes the growth rate of the trend component. A larger  $\lambda$  results in a higher penalty.

King and Rebelo (1993) proved that decomposing a time series using the filter proposed by Hodrick and Prescott (1997) is equivalent to decomposing it using a high-pass filter. The authors derived the gain function of this high-pass filter and demonstrated that the filter was capable of removing stochastic trends up to the fourth order.

### 7.3.5 Butterworth filter

The Butterworth filter, introduced by Butterworth (1930), has been extensively used in the engineering field due to its property of maximum flatness. In other words, the gain functions of this filter are as close as possible to a straight line at 0 for undesired frequencies and a straight line at 1 for desired frequencies. However, as with several techniques, economists have employed it in the realm of economics to solve time series filtering problems.

The Butterworth filter involves two parameters: the cutoff frequency and the filter's order. The cutoff frequency sets the point from which it begins filtering low-frequency stochastic cycles, while the filter's order establishes the slope of the gain function for a given cutoff frequency. For a given cutoff frequency, the gain function's slope at the cutoff frequency increases with the filter's order. Conversely, for a given order, the slope increases with the cutoff frequency. This flexibility resulting from the inclusion of a second parameter makes this filter better suited for time series decomposition than the [Hodrick and Prescott \(1997\)](#) filter.

Moreover, this filter shares properties of symmetry and phase neutrality (which does not shift the cycle backward or forward at a specific frequency) with the [Baxter and King \(1999\)](#) filter. However, as demonstrated by [Pollock \(2000\)](#), the Butterworth filter has crucial decomposition properties that depend on the filter's parameters. For instance, the Butterworth filter can stationarize any integrated time series of any order, provided that the filter's order is greater than the series' integration order.

### 7.3.6 Kalman filter

Different from the filters described earlier, the [Kalman \(1960\)](#) filter is not developed based on the frequency domain but on the time domain. It's a recursive method developed to find the internal state of a dynamic system (expressible as state-space models) from noisy measurements. In general, following the work of [Harvey \(1989\)](#), a state-space model is in the form:

$$\begin{aligned} y_t &= Dz_t + Fw_t + G\nu_t \\ z_t &= Az_{t-1} + Bx_t + C\varepsilon_t \end{aligned}$$

where  $A, B, C, D, F, G$  are parameter matrices, and

$y_t$	is a vector $n \times 1$ of observed endogenous variables
$w_t$	is a vector $k_w \times 1$ of exogenous variables
$\nu_t$	is a vector $r \times 1$ of observation errors
$z_t$	is a vector $m \times 1$ of unobserved state variables
$x_t$	is a vector $k_x \times 1$ of exogenous variables
$\varepsilon_t$	is a vector $q \times 1$ of state equation errors

The equation for  $z_t$  is known as the state equation, while the equation for  $y_t$  is known as the observation equation. Moreover,

$$\begin{aligned} \nu_t &\sim N(0, Q) \\ \varepsilon_t &\sim N(0, R) \\ E[\varepsilon_s, \varepsilon_t] &= 0 \text{ para todo } s \neq t \\ E[\nu_s, \varepsilon_t] &= 0 \text{ for all } s \text{ y } t \end{aligned}$$

The model parameters are obtained through maximum likelihood.

The Kalman filter method, introduced by [Kalman \(1960\)](#), allows obtaining linear forecasts of  $y_t$  through least squares, conditioned on past information. For each time instant  $t$  the filter produces the conditional expected state vector  $z_{t|t}$  and the conditional covariance matrix  $\Omega_{t|t}$ ; both conditioned on the available information up to time  $t$ . Using the model and information from previous periods, for each  $t$  the process starts by computing:

$$\begin{aligned} z_{t|t-1} &= Az_{t-1|t-1} + Bx_t \\ \Omega_{t|t-1} &= A\Omega_{t-1|t-1}A' + CQC' \\ y_{t-t-1} &= Dz_{t|t-1} + Fw_t \end{aligned}$$

These equations are known as one-step-ahead prediction equations. Subsequently, the residuals and the mean squared error matrix of the forecast errors are calculated as:

$$\begin{aligned} \tilde{v}_{t|t} &= y_t - y_{t|t-1} \\ \Sigma_{t|t} &= D\Omega_{t|t-1}D' + GRG' \end{aligned}$$

Finally, the conditional expected state vector and the conditional covariance matrix are updated as follows:

$$\begin{aligned} z_{t|t} &= z_{t|t-1} + \Omega_{t|t-1}D\Sigma_{t|t}^{-1}\tilde{v}_{t|t} \\ \Omega_{t|t} &= \Omega_{t|t-1} - \Omega_{t|t-1}D\Sigma_{t|t}^{-1}D'\Omega_{t|t-1} \end{aligned}$$