South African Reserve Bank Working Paper Series WP/23/08

The inherent uncertainties in output gap estimation: a South African perspective

Cobus Vermeulen

Authorised for distribution by Konstantin Makrelov

23 August 2023

South African Reserve Bank

South African Reserve Bank Working Papers are written by staff members of the South African Reserve Bank and, on occasion, by consultants under the auspices of the South African Reserve Bank. The papers deal with topical issues and describe preliminary research findings, and develop new analytical or empirical approaches in their analyses. They are solely intended to elicit comments and stimulate debate.

South African Reserve Bank Working Papers are externally refereed.

Information on South African Reserve Bank Working Papers can be found at
https://www.resbank.co.za/en/home/publications/Papers/working-papers

Enquiries relating to the Working Paper Series can be addressed to:
Head: Economic Research Department
South African Reserve Bank
P O Box 427
Pretoria 0001

Tel. +27 12 313 3911

The inherent uncertainties in output gap estimation: a South African perspective

Cobus Vermeulen[1]

Abstract

Monetary policy actions require accurate real-time estimates of potential output and the output gap. Such estimates are, however, vulnerable to definitional uncertainty, choice of methodology and data revision. It is also not always clear whether changes in economic conditions reflect changes in actual or potential output; flawed output gap estimates could in turn lead to an inappropriate policy stance. This paper compares the SARB's current approach to estimating potential output and the output gap to the empirical literature. New estimates of the output gap and potential growth are also presented, concentrating on the impact of the COVID-19 pandemic on economic activity. Consistent with existing literature, these estimates detect a chronically negative annual output gap since 2009, which only closed briefly in 2018, and a steady decline in the potential growth rate since 2010. The COVID-19 shock has likely exacerbated existing structural economic weaknesses and led to massive negative output gaps, notably during 2020, and continuing into 2021.

Key words
Potential output, output gap, monetary policy, multivariate filter

JEL codes
C32, E32, E37, E52

[1] Department of Economics, University of South Africa, and Research Fellow, SARB Chair in Monetary Policy Studies, University of Pretoria. vermejc@unisa.ac.za. I thank Rowan Walter and Rudi Steinbach for valuable conceptual and technical input. I also thank seminar participants at the South African Reserve Bank and two anonymous reviewers for their comments and suggestions.

1. Introduction

Potential output and the output gap are important concepts in modern monetary policy-making. They inform policymakers of prevailing economic conditions and directly guide decisions on the monetary policy stance, which is "set with reference to expected inflation and the actual cyclical position of the economy" (Melolinna and Tóth 2016:1). Actual output below potential (a negative output gap) indicates slack in the economy, suggesting that some productive resources are sitting idle. Conversely, actual output beyond potential (a positive output gap) could suggest an overheating economy. It follows that the output gap can be used to "indicate whether demand in an economy is excessive or deficient relative to the available resources used in production" (Botha, Ruch and Steinbach 2018:2). Accurate and timeous identification of the current state of the economy is therefore of critical importance to ensure an appropriate monetary policy stance.

However, while actual output can be readily measured, potential output, and by extension the output gap, are not directly observable, and therefore need to be estimated. Popular estimation paradigms include statistical methods, such as the Hodrick-Prescott (HP) filter, structural models such as the production function approach, and semi-structural models such as multivariate filters or unobserved components models. However, the wide range of techniques employed introduces uncertainty in the resultant estimates, with results "sensitive to the specific method employed" (Kemp 2015:549). Moreover, data and sampling choices can also influence estimates. Sample periods could include structural breaks, shocks, and periods of extreme volatility, while data is subject to revision over time.

The matter is further complicated by the fact that potential output is defined in various ways. Okun (1962:98) first posited that potential output is associated with some concept of 'full' employment. He qualified this to mean the "maximum production without inflationary pressure". However, numerous definitions of potential output have been proposed and utilised since then, ranging from the level of output associated with either full employment or price stability, to definitions related to 'steady state', 'efficient', 'perfect competition' and 'flexible price' levels of output. The variety of definitions employed resulted in the term 'potential output' being used "to describe related, but logically distinct, concepts" (Basu and Fernald 2009:188).

To reliably obtain an estimate of potential output and the output gap, there are thus at least three dimensions of uncertainty the researcher has to navigate – definition, methodology and data – which will each be addressed through the discussion below. Given the far-reaching consequences of policy decisions made on the basis of estimates of potential output – flawed estimates of potential output can easily lead to an incorrect policy stance (Kuttner 1994) – it is critical that such estimates meet the highest levels of economic and scientific rigour.

This paper reviews the empirical paradigms employed to date to estimate potential output, output gaps and potential growth in the South African context, and compares the evolution of the South African Reserve Bank's (SARB's) approach to the contemporary literature on the subject. It considers the above mentioned uncertainties against the rich theoretical and empirical literature, with the goal of identifying the most suitable approach for the South African economy, that is, the approach which minimises these uncertainties. It then presents a multivariate HP filter model, chosen as the superior estimation paradigm for the South African economy, to estimate new measures of potential output, potential growth and the output gap for the South African economy up to 2022. Of particular interest is the severe economic shocks caused by the COVID-19 pandemic, which caused massive economic contractions around the world in early 2020. The extreme volatility in economic data during this time poses significant challenges to estimation and filtering techniques, further increasing uncertainty in the data dimension.

This paper is structured as follows. Potential output and the output gap are defined in Section 2. This section also considers uncertainties arising from alternative definitions of potential output. Section 3 evaluates the mainstream estimation paradigms and discusses uncertainties arising from choices around methodology and data. Section 4 provides a brief overview of the evolution of the SARB's output gap estimates as well as the comparable South African academic literature. The choice of the multivariate HP filter as the superior estimation paradigm for the South African economy is substantiated in Section 5. A semi-structural multivariate HP filter model is then presented to complement and update existing estimates of potential output for the South African economy. Estimation results are presented in Section 6, and are consistent with the South African empirical literature: potential growth has been falling since 2010, temporarily collapsed during the COVID-19 shock in 2020, and then recovered in 2021. The annual output gap has been persistently negative since 2009, with the exception of 2018 and 2019 when it only just closed. This section also considers the monetary policy implications of potential output uncertainty. Section 7 concludes.

2. Defining potential output

Okun's (1962) seminal work spawned a substantial literature on defining and quantifying the concept of 'potential output'. Potential output is first and foremost a measure of an economy's long-term productive capacity, and is determined by supply-side considerations such as an economy's resource endowment, human and physical capital, and productivity. It can be viewed as the "steady-state level of output associated with the long-run aggregate supply curve" (Kuttner 1994:361), or the "level of output that involves the full utilization of factor inputs" (Borio, Disyatat and Juselius 2017:657).

Early definitions of potential output included some notion of 'full' or 'maximum' employment, that is the level of output at which the labour supply is fully utilised, or literally the "maximum possible output of an economy if all resources were fully employed" (Kemp 2015:550). However, this Keynesian "Okun-originated" concept of potential output was gradually supplanted by a Monetarist "natural-rate-based concept" (Congdon 2008:164), following the breakdown of the Phillips curve in the late 1960s and the stagflation of the 1970s. Potential output was increasingly viewed as the level of output where unemployment was at some 'natural' level, where it does not cause an acceleration or deceleration in inflation – the non-accelerating inflation rate of unemployment (NAIRU). This shift was also accompanied by a greater role for monetary policy relative to fiscal policy with respect to macroeconomic stability, as policy considerations shifted away from 'full' employment to price stability (Congdon 2008).

Additional definitions of potential output are found in general equilibrium theory. These can include the 'efficient', 'natural' and 'trend' levels of output (Vetlov et al. 2011), and all have different interpretations and implications for the design of optimal monetary policy.[2]

In the context of monetary policy, an operational definition of potential output is that "level of output that may be sustained indefinitely without creating a tendency for inflation to rise or fall" (Beneš et al. 2010:5). This definition is based on Okun's (1970:132) view that potential output represents the "maximum production without inflationary pressure". Attempting to exceed this level of production will put upward pressure on factor costs and ultimately on inflation (Bodnár et al. 2020). Because this paper is concerned with a monetary authority tasked with maintaining price stability,[3] potential output will therefore be defined here as Okun's (1970) 'maximum non-inflationary' level of production. This serves to anchor the definition of potential output and avoid ambiguity relative to alternative

[2] While a full-fledged dynamic stochastic general equilibrium (DSGE) model falls beyond the scope of this paper, it could be valuable to investigate output gaps arising from these different definitions. This is left for future research.

[3] The SARB is mandated by the Constitution of the Republic of South Africa (1996:S224(1)) to "protect the value of the currency in the interest of balanced and sustainable growth". This is embodied in a formal inflation target of 3–6%.

definitions.

Finally, potential growth, equivalent to the growth rate of potential output, is then defined as the "rate of growth possible without accelerating inflation" (SARB 2017:22), or, in other words, "the rate beyond which such growth will face production capacity constraints and hence deliver price increases" (Anvari, Ehlers and Steinbach 2014:3).

2.1 Supply and demand

The output gap compares the "*realised* output of an economy to its potential level of output" (Botha, Ruch and Steinbach 2018:2). That is, $\hat{Y}_t = Y_t - \bar{Y}_t$, where $\hat{Y}_t$, Y_t and $\bar{Y}_t$ represent, respectively, the output gap, actual output and potential output. Conceptually, economic slack, for example, an unemployment rate higher than the 'natural' rate, or inflation below its target,[4] will manifest in a negative output gap ($Y_t < \bar{Y}_t$), and vice versa.

Bodnár et al. (2020:42) suggest that "potential output typically reflects supply conditions", while "fluctuations around potential output are related to demand factors." The output gap could therefore be interpreted as a mismatch between supply (potential) and demand (actual) which gives rise to the business cycle, with the business cycle resulting "primarily from movements in aggregate demand in relation to a slow moving level of aggregate supply" (Scacciavillani and Swagel 1999:5). The output gap then provides "a measure of aggregate demand pressure relative to potential in an economy at a particular time" (Kemp 2015:549). The inflationary pressures which could result from this mismatch are important for monetary policymaking; monetary policy's key focus is therefore "the business cycle, with the aim of stabilising inflation" (Lienert and Gillmore 2015:7). Since a positive output gap is usually thought to be inflationary,[5] it normally induces monetary tightening in accordance with some monetary policy reaction function (MPRF)[6] to slow down aggregate demand.

2.2 Supply and supply?

The usual interpretation of a negative output gap is that "demand is too weak" (SARB 2017:24). However, a negative output gap may also be caused by an increase in potential output with actual output (demand) remaining temporarily unchanged or slow to adjust, or

[4] However, inflation below its target does not always imply economic slack. It may just be driven by positive supply shocks, which could mask demand-driven inflationary pressures in an overheating economy.

[5] Specifically, this is demand-pull inflation, where, in an overheating economy, demand runs ahead of supply.

[6] The SARB utilises a typical Taylor rule MPRF, where the policy rate is a function of, among others, inflation and the output gap (SARB 2017).

a situation in which potential output grows by more than actual output. Similarly, a positive output gap could be the result of a fall in potential.[7] If productive resources are idle, the economy will necessarily produce less than its potential. This implies that there could also be a wedge between actual and potential production, irrespective of the level of aggregate demand. Potential output can therefore be "driven by exogenous productivity shocks to aggregate supply" (Scacciavillani and Swagel 1999:6), which influence both the long-term growth trend and short-term fluctuations in output.[8]

It follows that the output gap may change *even if nothing changes on the demand side*. Negative supply shocks may well, while not directly or immediately influencing demand, constrain an economy's productive sectors to produce less than they are potentially able to. Alichi (2015:4, own emphasis) suggests that "macroeconomic shocks are not exclusive to the *actual* output; *potential* can also be hit by shocks." Supply shocks can therefore be thought to "affect the economy's productive potential rather than the output gap" (Botha, Ruch and Steinbach 2018:5). If, for example, a drought is considered to be a temporary decrease in potential output, an existing positive output gap would widen further. However, if the drought is not recognised as a supply shock, the only other way to explain a widening output gap and accompanying inflationary pressures would be to assume an overheating economy, and excess demand generating a larger positive output gap. The latter assumption would, however, misrepresent the true nature of demand pressures in the economy.

It is therefore of vital importance to assign economic information to the correct component of the output series. Not accounting for supply shocks or other transitory phenomena (e.g. the financial cycle or commodity windfalls, discussed below) could distort estimates of the output gap, and subsequently the extent of demand pressures in the economy, which could lead to an inappropriate monetary policy response. Similarly, positive demand shocks which are not correctly identified may be "confused for evidence of robust potential growth" (Janse van Rensburg, Fowkes and Visser 2019:1). The first criteria in choosing the estimation paradigm is therefore that it has to enable the researcher to disentangle transitory from more permanent effects on actual and potential output.

[7] From the output gap equation $\hat{Y}_t = Y_t - \bar{Y}_t$, $\downarrow \bar{Y}_t \implies \uparrow \hat{Y}_t$ while Y_t remains unchanged in the case of a positive output gap. The converse is true in the case of a negative output gap.

[8] Blanchard and Quah (1989) argue that supply shocks permanently influence output, while the influence of demand shocks is only temporary.

3. Measuring potential output

3.1 Estimation paradigms

As a starting point, any economic time series can be presented as the sum of a trend and cyclical component (Burns and Mitchell 1946). That is, $y_t = y_t^T + y_t^C$, where y_t^T and y_t^C represent, respectively, the trend and the cycle.[9] In the context of the output gap, the trend is equated to potential, while the cycle represents the output gap (equivalent to the difference between actual and trend, i.e. $y_t^C = y_t - y_t^T$). Approaches to estimating potential output therefore involve decomposing real gross domestic product (GDP) into its trend and cyclical components, where data are filtered to "extract the unobservable underlying potential output level from cyclical variations" (Alichi 2015:4).

From the academic literature, three main techniques[10] of estimating potential output can be identified. These are the classic HP filter, the production function (PF) approach, and the multivariate HP (MVHP) filter. The HP filter is a pure univariate statistical technique that utilises no more information than that contained in the output (GDP) series. It extracts estimates of potential output by filtering out trend and cyclical components at a particular frequency band (Borio, Disyatat and Juselius 2017). At the other extreme, structural models are anchored on theorised economic relationships by making "assumptions about the structure of the economy" (Álvarez and Gómez-Loscos 2018:828). The PF approach is such a full structural model that incorporates a host of other macroeconomic information beyond GDP. The MVHP filter is a hybrid of the two approaches, and is therefore classified as a semi-structural model.

These approaches "vary substantially in terms of the economic information they incorporate" (Borio, Disyatat and Juselius 2017:660). Univariate techniques are computationally easy and rely on only one time series, while the advantage of structural approaches is that rich macroeconomic information is incorporated in the estimation. However, the model set-up is much more onerous and sensitive to the researcher's choice of priors. Semi-structural approaches try to find a middle ground between computational ease and a richer estimation by incorporating some economic information, but with looser restrictions than under a full structural model.

3.1.1 HP filter

Hodrick and Prescott (1997) disaggregate a time series into a long-term trend and short-

[9] See also Nelson and Plosser's (1982) seminal work.

[10] See Álvarez and Gómez-Loscos (2018) for an exhaustive recent discussion on potential output estimation paradigms, including some fringe approaches.

term cycle. It is assumed that the "trend is stochastic and varies smoothly over time" (Álvarez and Gómez-Loscos 2018:828). Real GDP can be disaggregated into these two components by minimising the following loss function:

$$L = \sum_{t=1}^{T}(y_t - y_t^T)^2 + \lambda \sum_{t=1}^{T}\left[(y_t^T - y_{t-1}^T) - (y_{t-1}^T - y_{t-2}^T)\right]^2 \tag{1}$$

where y_t^T represents potential output and $y_t - y_t^T$ is the output gap. The smoothing parameter λ is a "positive number which penalises variability in the growth component series" (Hodrick and Prescott 1997:3).[11]

The HP filter is a simple technique to extract estimates of potential and the output gap using basic statistical software. Since it requires only one data series, it is "simple, transparent, and can be applied to any country where GDP data exist" (Blagrave et al. 2015:4). It is therefore particularly popular in some emerging economies where data scarcities preclude more data-intensive approaches. However, it suffers from three major shortcomings, which make the HP filter at best an imperfect tool for real-time output gap estimation:

1. It does not contain any structural economic information. For this reason, some shocks may mechanically be incorrectly assigned to potential (trend) output instead of actual output, and vice versa. Moreover, in the monetary policy context of this paper, no information on inflation is incorporated, so it is impossible to know whether the extracted trend truly represents a *non-inflationary* path of potential output.

2. The "choice of the degree of cycle versus trend to include. . . is chosen subjectively and is often prone to debate" (Ehlers, Mboji and Smal 2013:1–2). Álvarez and Gómez-Loscos (2018:828) argue that "theoretically, there exist an infinite number of possibilities of breaking down an economic series into a trend and a cyclical component" (by varying the value of the λ parameter), while "neither economic theory nor econometrics suggest a unique definition of trend." The convention of setting $\lambda = 1600$ is also subject to criticism: Du Toit (2008:22), for example, argues that "different optimal values of lambda arise due to different censoring rules regarding the duration of business cycles".

3. It suffers from the endpoint problem prevalent in statistical filters; that is, the "perva-

[11] The larger the value of λ, the smoother the estimated trend series, while a smaller λ will yield a better fit of the series. If $\lambda = 0$, the trend will equal the original series. For quarterly data, $\lambda = 1600$ is typically chosen (Hodrick and Prescott 1997).

sive unreliability of end-of-sample estimates of the trend in output" (Orphanides and Van Norden 2002:569). Because a statistical filter assumes that the "average deviation of actual output from its potential level should be zero" (Chen and Górnicka 2020:6), it usually "yields a path for the trend that tends to converge towards the actual data at the end points of the estimation sample" (Anvari, Ehlers and Steinbach 2014:8). This implies that the HP filter is often "biased toward indicating a *closed* output gap at the end of the sample" (2014:8).[12] Furthermore, because GDP data are often revised over time, output gap estimations are also revised as more data become available; therefore, "HP filter estimates of the output gap can only be considered reliable once a data point is a few years old" (2014:8).

3.1.2 PF approach

The PF approach "disaggregates real GDP in terms of the factors of production" (Anvari, Ehlers and Steinbach 2014:9). This approach incorporates structural information such as the capital stock, labour force and participation rates, and total factor productivity (TFP) in a production function specification. Estimates of individual production factors' potential levels can then be aggregated to determine total potential output.[13] This requires assumptions for 'normal' capacity utilisation, the 'natural' rate of unemployment, and trend TFP (Álvarez and Gómez-Loscos 2018).

However, this approach "may bring back through the backdoor some of the problems that plague univariate statistical approaches" (Borio, Disyatat and Juselius 2017:661). For example, estimates of trend factor utilisation are often calculated using univariate statistical filters; these estimates are then in themselves vulnerable to the endpoint problem. Moreover, variables such as the natural rate of unemployment are sensitive to the researcher's assumptions, and are therefore subject to heavy criticism. Overestimation of, for example, trend employment (tantamount to an underestimation of the NAIRU) may yield a structurally higher estimate of potential labour and consequently potential output (Arsov and Watson 2019). Doubtful estimates or diverging definitions of 'potential' capital and labour then cast doubt on the reliability of the resultant potential output estimates. Assumptions of perfect competition and constant returns to scale may also be critiqued, while the choice of functional form (e.g. Cobb-Douglas or constant elasticity of substitution) would also

[12] The endpoint problem can be alleviated to some degree by adding forecasted observations to the end of the sample in order to artificially remove the 'endpoint' of the most recent actual observations. However, such forecasts are naturally also prone to error and uncertainty.

[13] For example, if the production function is defined as $Y_t = A_t K_t^{\alpha} N_t^{\beta}$, potential output can be determined from $\bar{Y}_t = \bar{A}_t \bar{K}_t^{\alpha} \bar{N}_t^{\beta}$.

yield different results (Steenkamp 2018).

3.1.3 MVHP filter

Semi-structural models combine the efficiency of purely statistical methods with economic theory by incorporating useful information contained in other variables (Anvari, Ehlers and Steinbach 2014, Álvarez and Gómez-Loscos 2018). The general approach is to model potential output as an unobserved stochastic trend, and then give deviations of output from this trend (i.e. the output gap) some economic definition or interpretation. This can be done by linking output to, for example, unemployment or inflation, based on a priori theorised or empirical relationships between output and these observed variables, and utilising actual observations in these data to help identify the output gap. These relationships can therefore condition or restrict movements in the output gap to provide more economic interpretation and intuition than under a univariate approach. Various structural restrictions can then be employed to "expand the information set used in the estimation of the output gap" (Álvarez and Gómez-Loscos 2018:840).

Borio, Disyatat and Juselius's (2017) influential paper[14] proposes a parsimonious multivariate filter approach based on, among others, Laxton and Tetlow (1992) and Beneš et al. (2010). It involves augmenting the standard HP filter with structural economic relationships and estimating the resulting system using a Kalman filter and Bayesian techniques. This approach leverages the HP filter's ease-of-calculation advantage, while allowing structural economic relationships to contribute to a richer and more robust estimation of potential output. Borio, Disyatat and Juselius (2017:656) argue that "financial developments contain information about the *cyclical* component of output" (own emphasis). Because financial developments are inherently cyclical, failure to account for these factors would lead to the impact of financial developments on output to be erroneously ascribed to *potential* output, which would lead to an overestimation of potential and underestimation of the gap.[15] This approach captures "the information content that financial factors have for the cyclical, potentially highly persistent, variations in output" (Borio, Disyatat and Juselius 2017:659), which yields 'finance-neutral' measures of the output gap.

Generally speaking, this strand of research contends that estimates of *potential* output should filter out transitory shocks to *actual* output, including, for example, the financial cycle (Borio, Disyatat and Juselius 2017) or commodity prices (Botha and Schaling 2020).

[14] This was first published as a Bank for International Settlements (BIS) working paper in 2013.

[15] This is a key shortcoming of univariate filters, which do not incorporate economic information beyond the filtered series itself.

This approach can be generalised to exploit the information content of observable variables which co-move with the business cycle (Melolinna and Tóth 2016); this helps to "disentangle the trend from the cycle" (Andersson et al. 2018:51) and allows the researcher to "isolate the effects of structural vs. cyclical influences on output" (Kemp 2015:550).[16] These models can (i) be estimated more precisely and (ii) are more robust in real time than alternative approaches.

3.2 Data uncertainty

The final requirement from the estimation paradigm is that it has to be robust to data revisions. As was mentioned above, GDP data are often revised over time. This also applies to potential output, an unobserved variable of which "historical estimates . . . have been revised substantially" (SARB 2017:23), and subsequently output gaps.[17] According to Kemp (2015:554), "policy actions that may seem perfectly reasonable and appropriate might prove to be wholly inappropriate as new data become available."

In addition, statistical estimation techniques can be vulnerable to structural breaks, shocks, and data volatility. Output gaps are even harder to measure in the vicinity of a large shock like the global financial crisis or COVID-19 (Melolinna and Tóth 2016).

3.3 Summary: estimation and data

The advantages and disadvantages of these three mainstream approaches are summarised in Table 1. While other approaches are theoretically interesting and may be useful under certain conditions, they are currently not that popular.

Table 1: Summary of international literature

Approach	Advantages	Disadvantages
HP filter	Computationally easy Suitable where other data are inadequate	Endpoint problem Lack of structural information Sensitive to GDP revisions
PF	Rich economic interpretation	Estimating underlying trends Data requirements Vulnerable to specification error
MVHP	Less sensitive to specification errors Various relationships can be tested Real-time robustness	Data requirements

[16] Indeed, a feature of the empirical literature is that estimates of potential output are often revised downward when controlling for more temporary factors.
[17] This critique applies around the world. See Orphanides and Van Norden (2002).

Several leading central banks, including the Fed (Mishkin 2007), the Reserve Bank of New Zealand (Lienert and Gillmore 2015), the Bank of Canada and the Bank of Japan (Arsov and Watson 2019) endorse the PF approach. These central banks do, however, have the advantage of rich data and relatively low and stable unemployment rates; this could ameliorate some of the challenges of the PF approach highlighted above. The multivariate filter is highly regarded by international financial institutions such as the International Monetary Fund (IMF) (Beneš et al. 2010, Alichi 2015 and Blagrave et al. 2015) and the BIS (Borio, Disyatat and Juselius 2013, Borio, Disyatat and Juselius 2014, Borio, Disyatat and Juselius 2017 and Alberola et al. 2016), and is employed by central banks such as the Bank of England (Melolinna and Tóth 2016) and the SARB (Botha, Ruch and Steinbach 2018).[18]

4. The South African context

4.1 Evolution of the SARB's output gap estimates

The SARB's approach to estimating potential output has evolved along with the empirical literature. While the output gap was a feature in the SARB's monetary policy models from the early 2000s, its first output gap estimates were published in 2013,[19] whereafter updates to the approach were published in 2014 and 2018. Over this period, the SARB's estimations evolved from the PF approach, which was the dominant paradigm in the early 2000s, to an aggregation approach (incorporating a range of statistical and structural estimates), to a multivariate semi-structural approach. The SARB adopted a finance-neutral model in 2014, which was augmented in 2018 to more accurately reflect short-lived supply shocks.

4.1.1 Aggregation approach (2013)

Ehlers, Mboji and Smal (2013) employ four estimation paradigms – a standard HP filter, a structural PF model, a semi-structural MVHP filter, and a potential general equilibrium model – which are then aggregated to create a more robust, equally weighted average estimate of potential output, with the aim of "moderating biases associated with each estimation technique" (2013:6).[20]

[18] Fedderke and Mengisteab (2017) also provide a recent summary of methodologies employed by international institutions.

[19] It was only with the publication of Ehlers, Mboji and Smal's (2013) paper that the SARB started making its approach explicit in the public domain. This marked a notable shift in its communication strategy, and contributed to improving the transparency of monetary policy.

[20] A similar approach was followed by the Fed around the same time (see Edge and Rudd (2016)).

4.1.2 Semi-structural finance-neutral model (2014)

Anvari, Ehlers and Steinbach (2014) propose a move away from the aggregation approach of Ehlers, Mboji and Smal (2013) in favour of a single semi-structural model. It was initially thought that "an average of several methods, each with its own shortcomings, is likely to be more accurate than any individual method" (Anvari, Ehlers and Steinbach 2014:10). However, they criticise the PF approach, citing challenges in estimating potential labour and capital, which contributes to "uncertainty surrounding the eventual estimate of potential output" (Anvari, Ehlers and Steinbach 2014:9), while re-emphasising the end-of-sample problems plaguing the standard HP filter. The PF approach can also often capture temporary developments (e.g. commodity or financial cycles alluded to earlier), which are erroneously ascribed to potential – instead of actual – output. This new model is enhanced with, among others, features of the financial cycle, and is able to "replicate the fundamental properties of purely statistical filters without any of the inherent shortcomings thereof" (Anvari, Ehlers and Steinbach 2014:i).

4.1.3 Short-lived supply shocks (2018)

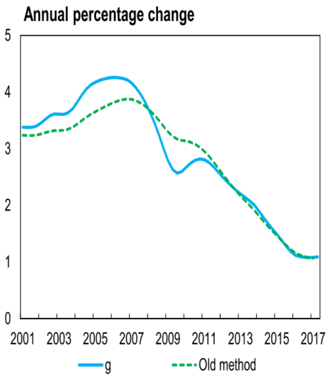
One shortcoming of the Anvari, Ehlers and Steinbach (2014) model was that it did not "account for short-term fluctuations in growth caused by supply shocks" (SARB 2017:23), which subsequently "misdiagnoses short-term supply shocks as demand phenomena" (2017:23). Botha, Ruch and Steinbach (2018) therefore attempt to isolate the effects of short-lived supply shocks on estimates of potential output. By controlling for supply-side conditions, they find that the "resulting output gap more accurately reflects a measure of demand pressures in the economy" (Botha, Ruch and Steinbach 2018:1).

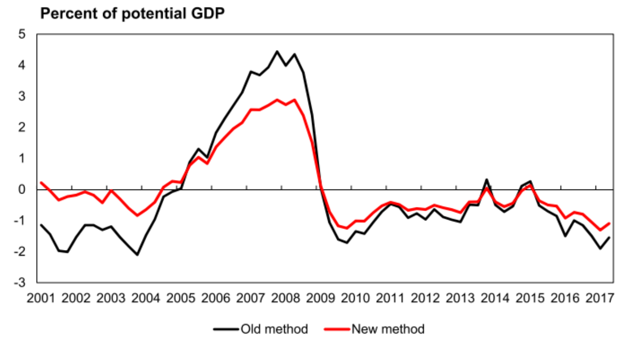
The advantage of this approach is that potential output is now much more responsive to current developments affecting supply. Temporary supply-side events are now correctly captured in potential output, and no longer "reflected in the output gap as shocks to demand" (Botha, Ruch and Steinbach 2018:14). This leads to a more dynamic underlying potential growth series g as compared to the 'old' method, as illustrated in Figure 1, and ultimately a generally narrower output gap.

4.2 South African academic literature

In addition to the international literature, local research may also have informed the SARB's output gap estimates. This literature is summarised in Table 2. While the PF approach was popular in the earlier years, later studies (2014 onwards) have relied almost exclusively on the MVHP framework. Over comparable samples, the PF approach often yielded higher

Figure 1: New estimates of underlying growth and the output gap





Source: Botha, Ruch and Steinbach (2018)

estimates of potential output,[21] which is perhaps attributable to the difficulty in pinning down an accurate, stable measure of the NAIRU or some benchmark unemployment rate (Du Toit, Ground and Van Eyden 2006). Irrespective of the method(s) employed, however, the main takeaway is a concerning pattern of a continual reduction in the pace of potential growth.

5. Methodology and estimation

5.1 Choice of estimation paradigm

Based on the empirical literature, the MVHP filter is deemed the superior approach for the South African economy. This approach meets all three criteria for reliably estimating potential output set out above, thereby minimising the various dimensions of uncertainty. The classic univariate HP filter may have been the best approach in a previous era of relative data scarcity; however, advances in computing power and the quality and abundance of data now allow for much richer multivariate techniques. The PF approach utilises much more structural economic information, but it is extremely vulnerable to misspecification. Steenkamp (2018) shows how alternative model specifications could provide implausible output gap estimates, while high and volatile South African unemployment rates make it difficult to pin down a stable NAIRU (Du Toit, Ground and Van Eyden 2006). Moreover, if the production factors' underlying trends are extracted using an HP or other univariate filter, the model's real-time properties will not be reliable.

The MVHP filter's first advantage is that it is not as restrictive as the PF approach: instead of strictly imposing specific relationships, it lets the data 'speak', thus being *less vulnerable*

[21] See also Table 1 in Klein (2011).

Table 2: Summary of South African literature

Author(s)	Approach	Pot. growth	Sample
Smit and Burrows (2002)	HP, PF, AMV	nc/p	n/a
Du Toit, Ground and Van Eyden (2006)	PF	3%	1971–2003
Du Plessis, Smit and Sturzenegger (2008)	SVAR	2.9%	2004–2008
Ehlers, Mboji and Smal (2013)†	HP, PF, MVHP	3.9% 2.9%	2000–2007 2009–2011
Anvari, Ehlers and Steinbach (2014)†	MVHP	3.5% 3.0%	2000–2008 2009–2014
Kemp (2015)	MVHP	3.6% 2.6%	2001–2007 2009–2014
Kemp and Smit (2016)	MVHP	3.2% 2.2%	1994–2007 2011–2014
Fedderke and Mengisteab (2017)‡	various	3.5–3.8% 2.0–2.4%	2005–2010 2010–2015
Botha, Ruch and Steinbach (2018)†	MVHP	2.1%	2008–2016
Steenkamp (2018)	PF	2.6%	2001–2017
Botha and Schaling (2020)	MVHP	nc/p	1972–2019

† represents methodologies officially followed by the SARB. AMV = adapted multivariate filter. nc/p = not calculated or published.
‡: Fedderke and Mengisteab (2017) estimate a narrow range of potential growth rates using various statistical filters.

to misspecification and allowing a wide range of conditioning restrictions to be evaluated. In particular, by building in a conditioning role for inflation, it enables output gap estimates to be *anchored in a non-inflationary definition*.

Secondly, the MVHP filter has been empirically shown to yield more precise and more robust real-time output gap estimates than other methods. While all estimation paradigms are vulnerable to data revision, particularly in GDP figures, the MVHP approach is the *least vulnerable to data uncertainty*. This is because this approach is anchored in structural economic relationships and incorporates data on a range of economic variables; these additional variables are less susceptible to revision than GDP figures are, thus minimising ex post output gap revisions even as GDP data are revised. Therefore, even though the MVHP filter remains vulnerable to end-of-sample problems, albeit less so than pure univariate approaches by taking advantage of information from other economic variables, it provides arguably the most robust real-time estimates of potential output. This result holds across numerous economies, samples and datasets.[22]

[22] MVHP models usually have lower mean absolute output gap revisions than competing models. See

Finally, a key result from the literature review is that researchers must be able to "distinguish between *permanent movements in* potential output and *transitory movements around* potential" (Scacciavillani and Swagel 1999:6, own emphasis). Neither the HP filter nor the PF approach lends itself to this type of decomposition, but by controlling for these relationships in the MVHP framework more reliable estimates of potential output and the output gap can be obtained.

5.2 Model specification

In this section, an MVHP model is estimated, building on the methodology described in Botha, Ruch and Steinbach (2018), Borio, Disyatat and Juselius (2017) and Blagrave et al. (2015), to obtain new estimates of potential output for South Africa.

5.2.1 State-space form

The standard HP filter optimisation problem can be cast in state-space form as

$$\Delta\bar{y}_t = \Delta\bar{y}_{t-1} + \varepsilon_{0,t} \tag{2}$$

$$y_t = \bar{y}_t + \varepsilon_{1,t} \tag{3}$$

The measurement equation 3 can then be augmented to embed economic information. Specifically, it can be rewritten as

$$y_t = \bar{y}_t + \beta(y_{t-1} - \bar{y}_{t-1}) + \gamma' X_t + \varepsilon_{2,t} \tag{4}$$

where y_t is the natural logarithm of real GDP, $\bar{y}_t$ is potential output, and $\hat{y}_t = y_t - \bar{y}_t$ is the output gap. X_t is a vector of economic variables with parameters γ' upon which the path of potential output can be conditioned.[23] The state equation 2 produces an estimate of unobservable potential output, which is informed by an augmented measurement equation 4 incorporating observable data.

Observed data on actual output y_t as well as variables in the vector X_t can help to pin down the output gap,[24] and ultimately determine the evolution of potential output. In principle, X_t can contain any variable that "embeds useful information on the slack/tightness of the

Beneš et al. (2010), Anvari, Ehlers and Steinbach (2014), Blagrave et al. (2015), Kemp (2015), Melolinna and Tóth (2016), Borio, Disyatat and Juselius (2017) and Botha and Schaling (2020).

[23] In the special case where $\gamma_i = 0 \;\forall\, i$, this specification would yield an estimate for potential identical to the standard HP filter.

[24] Equation 4 can also be expressed in terms of the output gap only, that is $\hat{y}_t = \beta\hat{y}_{t-1} + \gamma' X_t + \varepsilon_{2,t}$.

economic landscape" (Alberola et al. 2016:10). Additional variables are informative only insofar "movements in potential output affect them differently than the cyclical movements in actual output" (Chen and Górnicka 2020:6). Output gap estimates "are only improved relative to a simple statistical filtration if the structural relationships specified in the filter are valid ones" (Alichi 2015:5).

The error terms $\varepsilon_{0,t}$ and $\varepsilon_{2,t}$ are assumed to be normally and independently distributed with mean zero and variances σ_0^2 and σ_2^2, respectively. The corresponding noise-to-signal ratio is given by the parameter $\lambda_2 = \sigma_2^2/\sigma_0^2$. To "generate measures that are of comparable cyclicality to the standard HP filter" (Borio, Disyatat and Juselius 2017:662), λ_2 is set so that $\frac{\text{var}(y_t - \bar{y}_{(3),t})}{\text{var}(\Delta^2 \bar{y}_{(3),t})} = \frac{\text{var}(y_t - \bar{y}_{(4),t})}{\text{var}(\Delta^2 \bar{y}_{(4),t})}$, where $\bar{y}_{(3),t}$ and $\bar{y}_{(4),t}$ are estimates of potential output obtained from, respectively, equations 3 and 4.

5.2.2 Short-lived supply shocks

This model also embeds Botha, Ruch and Steinbach's (2018) decomposition of the underlying potential growth process:

$$g_t = \rho_g g_{t-1} + (1 - \rho_g) g^{ss} + \varepsilon_t^g \tag{5}$$

where g_t and g^{ss} represent, respectively, the underlying trend growth rate and the long-run steady-state growth rate. $\rho_g < 1$ is an autoregressive term and ε_t^g is a residual term. Potential output (eq. 2) is augmented with this underlying trend growth rate as well as an error-correction term $\theta(\bullet)$ to return potential output to its steady-state path:

$$\bar{y}_t = \bar{y}_{t-1} + g_t - \theta(\bar{y}_{t-1} - \bar{y}_{t-2} - g_{t-1}) + \varepsilon_t^{\bar{y}} \tag{6}$$

where $\varepsilon_t^{\bar{y}}$ acts as a shock to the level of potential output.[25] Botha, Ruch and Steinbach (2018:10) argue that this approach allows for a potential output growth process with the following features:

- a constant long-run potential growth rate equal to the steady-state growth rate g^{ss}, which represents the long-term equilibrium;
- medium-term equilibrium dynamics in the form of persistent deviations from the long-run equilibrium which are brought about by shocks to the underlying trend growth rate

[25] If this shock term is excluded and ρ_g is set equal to 1 in 5, the model collapses to the earlier specification of Anvari, Ehlers and Steinbach (2014) where temporary supply shocks are not accounted for.

g_t (eq. 5); and

- short-lived deviations of potential growth from these medium-term dynamics driven by transitory shocks to the level of potential output (eq. 6).

5.3 Structural restrictions

The MVHP paradigm allows structural equations to link the evolution of the unobservable output gap to observable macroeconomic data. Numerous theoretical relationships are tested here to supplement the information contained in output data with economic theory, including inflation, the financial cycle, capacity utilisation and commodity prices.

5.3.1 Phillips curve

The core notion of potential output is some level of output that is *not inflationary*. A forward-looking open-economy Phillips curve, similar to the specification in Botha et al. (2017), is added:

$$\pi_t = \alpha^{\pi}\pi_{t-1} + (1-\alpha^{\pi})E_t\pi_{t+1} + \alpha^{\hat{y}}\hat{y}_t + \alpha^{q}\hat{q}_t + \varepsilon_t^{\pi} \tag{7}$$

Inflation π_t is influenced by past inflation π_{t-1}, expected inflation $E_t\pi_{t+1}$, the output gap $\hat{y}_t$ and the real exchange rate gap $\hat{q}_t$. $\alpha^{\pi} < 1$ is an autoregressive term and ε_t^{π} is a residual term.

5.3.2 Financial cycle

The financial cycle has formed an integral part of the SARB's potential output models since Anvari, Ehlers and Steinbach (2014), and has been ubiquitous in the South African literature thereafter (Kemp 2015, Kemp and Smit 2016, Botha, Ruch and Steinbach 2018, and Botha and Schaling 2020). The same specification is therefore employed here, with growth in private sector credit extension $\hat{c}_t$ included in the vector X_t, and $\hat{c}_t$ evolving following a first-order autoregressive (AR(1)) process of the form $\hat{c}_t = \upsilon_c\hat{c}_{t-1} + \varepsilon_t^{\hat{c}}$.

5.4 Full model

The full model is described below:

Output identity: $y_t = \bar{y}_t + \hat{y}_t$ (8)

Potential output: $\bar{y}_t = \bar{y}_{t-1} + g_t - \theta(\bar{y}_{t-1} - \bar{y}_{t-2} - g_{t-1}) + \varepsilon_t^{\bar{y}}$ (9)

Output gap: $\hat{y}_t = \beta^{\hat{y}}\hat{y}_{t-1} + \gamma_1\hat{c}_t + \varepsilon_t^{\hat{y}_t}$ (10)

Underlying trend growth: $g_t = \rho_g g_{t-1} + (1-\rho_g)g^{ss} + \varepsilon_t^{g}$ (11)

Potential growth: $\Delta\bar{y}_t = g_t/4 + \varepsilon_t^{\bar{y}}$ (12)

Phillips curve: $\pi_t = \alpha^{\pi}\pi_{t-1} + (1-\alpha^{\pi})E_t\pi_{t+1} + \alpha^{\hat{y}}\hat{y}_t + \alpha^{q}\hat{q}_t + \varepsilon_t^{\pi}$ (13)

Inflation exp.: $E_t\pi_t = \mu E_t\pi_{t-1} + \varepsilon_t^{E_t\pi_t}$ (14)

RER gap: $\hat{q}_t = \tau\hat{q}_{t-1} + \varepsilon_t^{\hat{q}}$ (15)

Credit growth: $\hat{c}_t = \upsilon_c\hat{c}_{t-1} + \varepsilon_t^{\hat{c}}$ (16)

An alternative specification of the financial cycle, considered by both Borio, Disyatat and Juselius (2017) and Kemp (2015), is to, instead of credit growth, use growth in real house prices as a proxy for the financial cycle. This specification is also tested below. Other specifications, including capacity utilisation and commodity prices as conditioning restrictions, are also evaluated here in an AR(1) process similar to the financial cycle restriction.

5.5 Estimation

5.5.1 Data

Data spanning 1970Q1–2022Q1 were sourced from the SARB's Quarterly Projection Model (QPM) and core model. Explanatory variables are mean adjusted: following Borio, Disyatat and Juselius (2017), Cesàro means, that is, the mean of the sequence of means obtained by successively adding one observation starting from the initial date of the sample, are subtracted from each corresponding sample observation. The Cesàro mean "converges faster to the population mean" and "generally fluctuate less as the sample size grows" (2017:664); this mitigates the problem of cyclical means.

5.5.2 Priors

Bayesian techniques are employed to estimate the coefficients. Priors are based on the empirical literature, notably Botha, Ruch and Steinbach (2018) and Anvari, Ehlers and Steinbach (2014), with selected priors described in Table 3. Parameters are only estimated up to 2020Q1, that is, right before the COVID-19 shock hit the economy. This is because attempting to estimate parameters through the extreme volatility of 2020Q2 and Q3 yields highly implausible estimates of potential output and the output gap. An alternative specification, where parameters are estimated through the COVID-19 shock, is also provided below for illustrative purposes.

Table 3: Selected list of priors

Coefficient	Prior distribution	Prior mean	Std.dev.
$\beta^{\hat{y}}$	Beta	0.88	0.05
α^{π}	Beta	0.5	0.2
$\alpha^{\hat{y}}$	Gamma	0.2	0.1
α^{q}	Gamma	0.05	0.02
τ	Beta	0.85	0.05
θ	Gamma	0.3	0.1
γ_1	Gamma	0.3	0.1
υ_c	Beta	0.34	0.2

To obtain a robust final estimate, Botha and Schaling's (2020) approach of evaluating various model specifications and testing the contribution and significance of several predictor variables is followed here. Table 4 describes various model specifications and variables' associated *t*-statistics as indications of their statistical significance. *Model 0* is the baseline model, containing the dynamic output gap and Phillips curve relationships but no additional predictor variables. *Models 1–4* have, respectively, credit growth, house prices, capacity utilisation and commodity prices included as predictor variables. Because of the very small coefficients on house prices, capacity utilisation and commodity prices, *Model 1* with credit growth is chosen as the preferred model.

Table 4: Model specifications and posterior coefficients

	Model 0	Model 1	Model 2	Model 3	Model 4
Dynamic gap	0.926 (18.522)	0.920 (18.398)	0.939 (18.772)	0.927 (18.536)	0.926 (18.510)
Credit growth		0.110 (1.096)			
House prices			0.094 (0.935)		
Capacity utilisation				0.048 (0.959)	
Commodity prices					0.015 (0.049)

Note: t-statistics in parentheses

6. Results and discussion

6.1 Output stagnation and COVID-19

The COVID-19 shock, which in 2020 wreaked havoc on an already fragile economy,[26] was a simultaneous negative supply and demand shock: capital and labour were withdrawn from economic activity due to lockdowns, while falling incomes due to firm closures and higher unemployment reduced demand for goods and services (SARB 2021). Lockdowns, imposed by governments to contain the spread of the virus, forced workers to stay home en masse and prevented them from contributing their labour to processes of production, reducing aggregate supply. Many employees subsequently lost part, or all, of their income. At the same time, many firms and stores (across the goods and services sectors) were unable to do business. Coupled with income losses, this substantially reduced demand, which manifested in massive contractions in GDP and inflation.[27]

It is, however, not obvious how such simultaneous shifts in supply and demand would influence the output gap. If factors of production remain in place while they are temporarily prevented from being utilised fully, the supply shock would be transitory. A fall in actual output while potential remains relatively unchanged will manifest in a sizeable negative output gap. On the other hand, if potential output falls to the same extent as actual output, there should be little impact on the output gap (Bodnár et al. 2020). Furthermore, the magnitude of the shock and the resultant extreme data volatility pose substantial empirical challenges to estimates of potential output and the output gap over this period, presenting an additional source of uncertainty in estimation paradigms.

6.2 Results

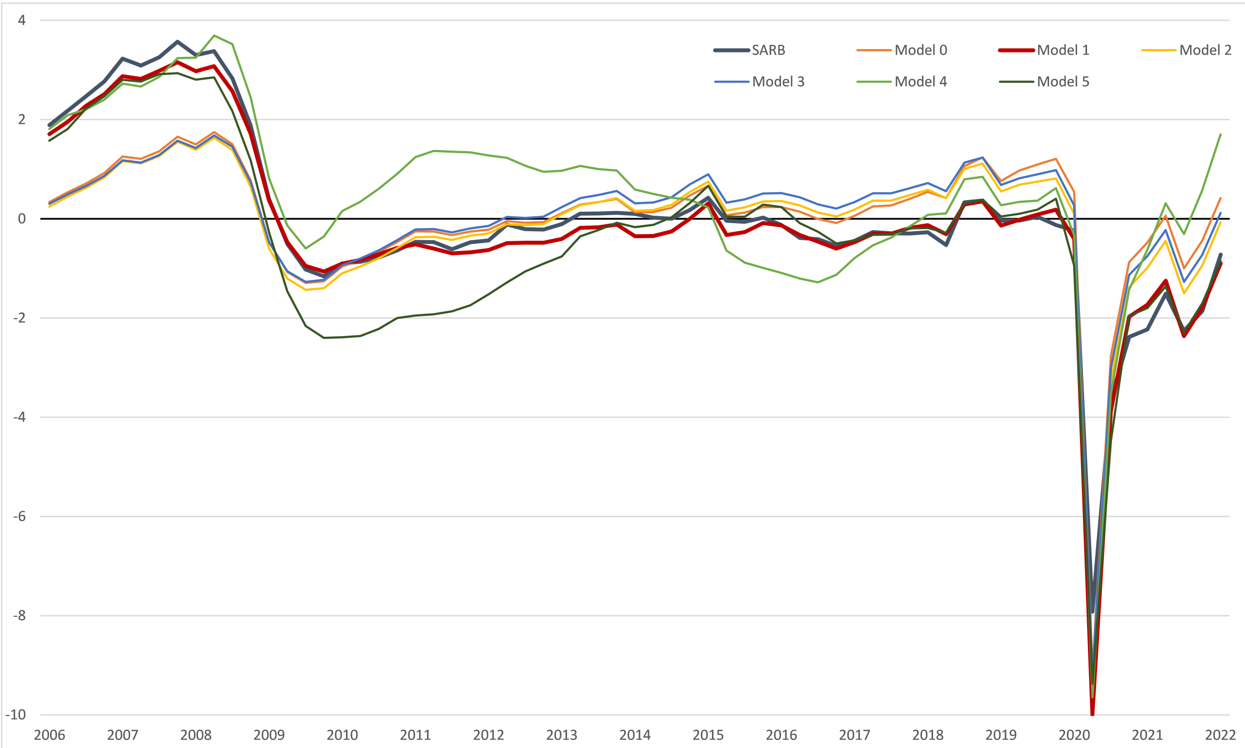
Figure 2a below provides estimates of the quarterly output gap since 2006. These include the SARB's official estimation (*SARB*) and the five specifications from Table 4 (*models 0–4*). These six models are obtained by estimating parameters only up to 2020Q1, that is, right before the COVID-19 shock hit the economy. A seventh specification (*Model 5*) estimates the preferred *Model 1* parameters through the 2020 data volatility up to the latest data point.

Similar to the empirical literature, estimated output gaps are highly persistent, with the coefficient on the lagged output gap ranging between 0.920 and 0.939. *Model 1* is largely

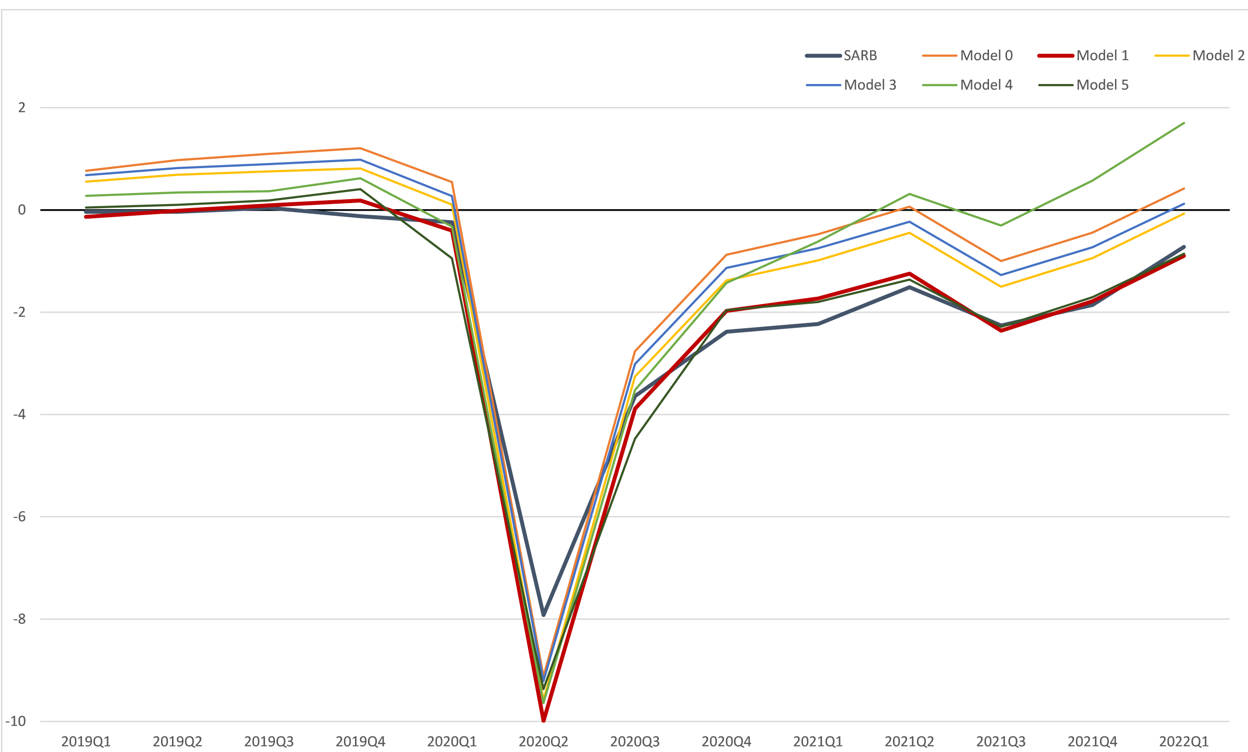
[26] Potential growth in the South African economy had been steadily falling since 2008 due to "lower capital formation and weaker productivity growth" (SARB 2017:23). Persistent electricity shortages also continue to drag down potential output.
[27] During 2020, real GDP contracted by 7.0%, while inflation slowed to 2.1% in May 2020.

Figure 2: Output gaps and the COVID-19 shock (2006–2022)

(a) Output gaps (2006–2022)



(b) COVID-19 shock (2018–2022)



in line with the SARB's current estimates. The two models show that the magnitude of the output gap is comparable in the period leading up to the financial crisis (2006-2008) and in the period thereafter (2008-2010). Both models suggest that the quarterly output gap briefly closed in 2015 and again in 2018. *Model 1*, however, estimates a more negative output gap between 2011 and 2016, resulting from a slightly higher estimate of potential output and weak actual output over this period. This also yields a small negative output gap estimated for 2013–2014, as compared to a small positive output gap according to the SARB's estimates for the same period. In fact, relative to the *SARB* model, *Model 1* generally finds a slightly higher level of potential output, which also results in a narrower output gap for the years leading up to the financial crisis.

Models 0, 2 and 3 trend together very closely, but tend to deviate from the SARB's official output gap. These models find a smaller build-up of financial imbalances during 2006–2008, manifesting in a narrower output gap, as well as a lower estimate of potential output during 2012–2019, as evidenced by a wider, more positive output gap. This illustrates the importance of credit growth in capturing the build-up of financial imbalances. Similar to Botha and Schaling's (2020) results, including commodity prices (*Model 4*) yields a sustained wider output gap between 2009 and 2014. Disentangling the effect of the commodity boom results in a lower estimate of potential output and a subsequent wider output gap for this period.

The COVID-19 shock is captured by all models (Figure 2b). *Model 1* identifies a massive negative output gap of -9.9%[28] in the second quarter of 2020, even bigger than the SARB's estimate of -7.9%, and a negative annual output gap of -4.2% for 2020 (Figure 3a). Figure 2b shows a negative output gap, ranging between -7.9% and -10%, depending on the model specification, for the second quarter of 2020.[29] Potential output contracted by 2.7% in 2020, and then grew at 2.8% in 2021 (Figure 3b). Irrespective of the model specification, the COVID-19 shock did not cause a permanent slowdown in potential output, as evidenced by the strong recovery in potential growth in 2021. However, despite the rebound in potential output, the actual output recovery was relatively weak, manifesting in a protracted negative output gap. This may be ascribed to the COVID-19 shock exacerbating pre-COVID structural economic weaknesses which had existed since 2008-09.

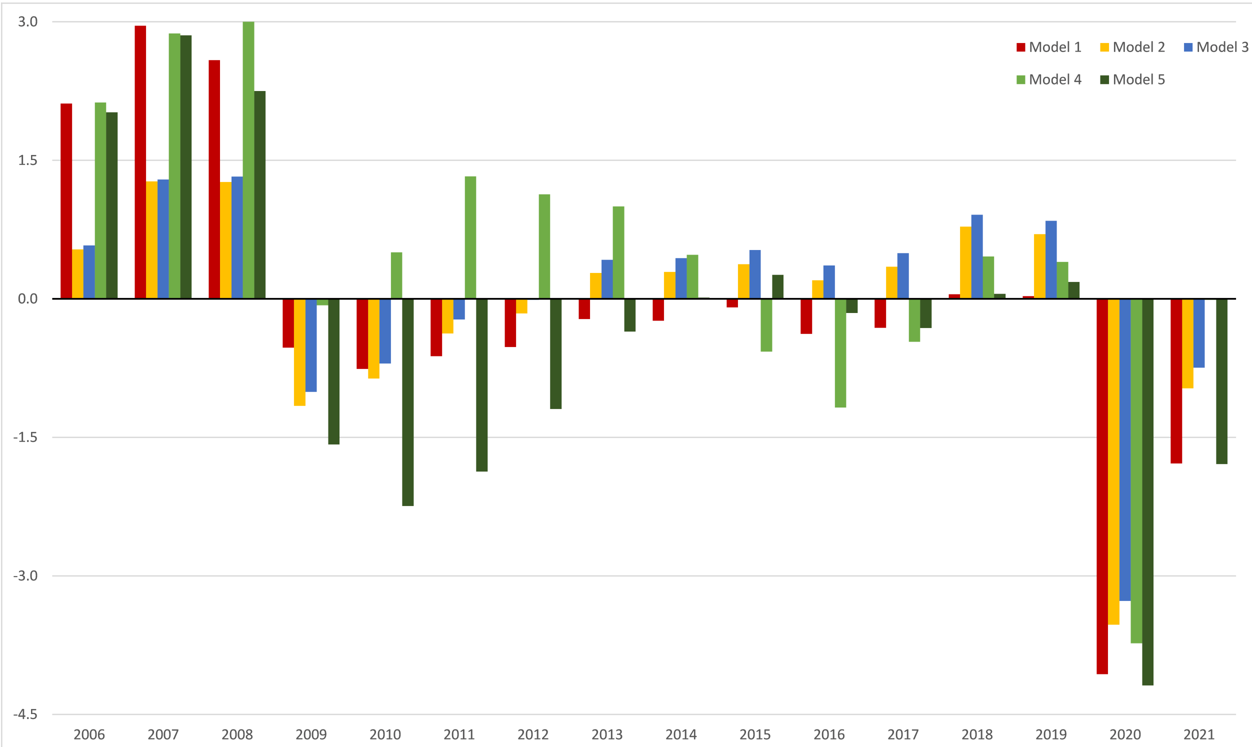
To prevent the filtering procedure from pre-empting the COVID-19 shock, *models 0–4* fix

[28] This is almost as big as the estimated output gap of about -11% in the euro area for the same quarter (Bodnár et al. 2020).

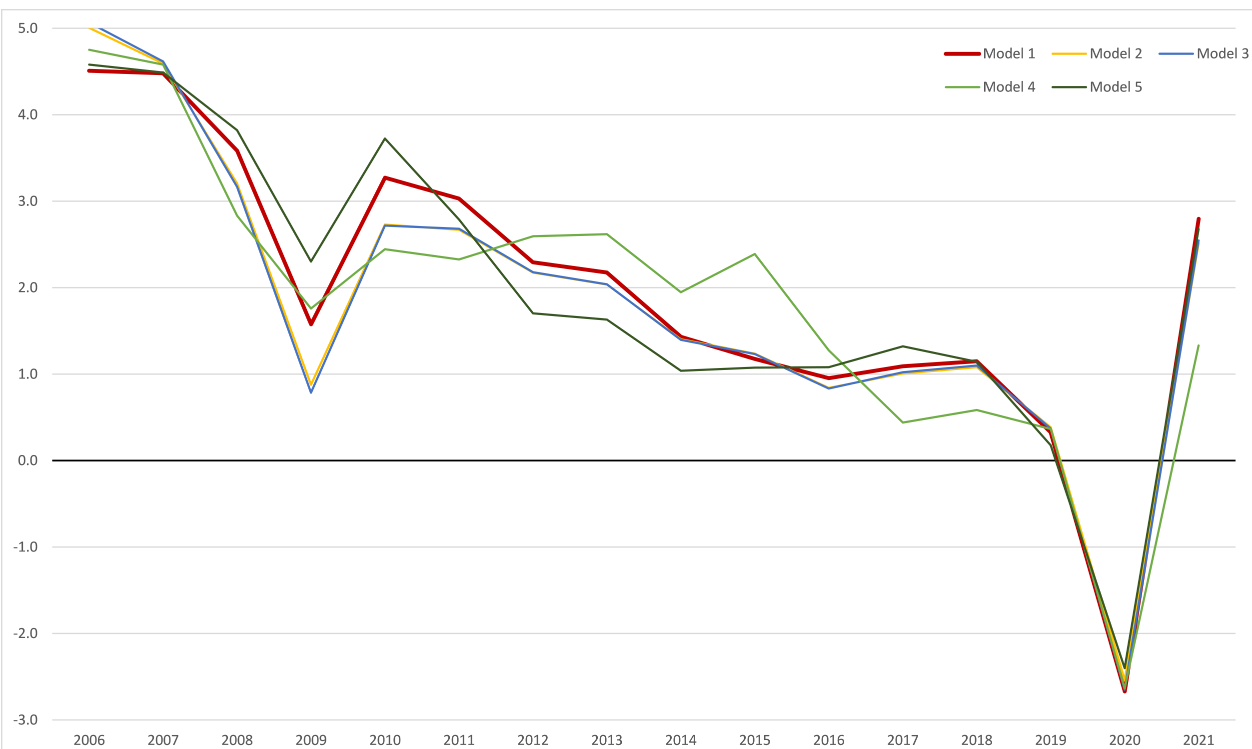
[29] In contrast, the vanilla HP filter (not shown) suggests a negative output gap of -16.4% for this quarter.

Figure 3: Annual output gaps and potential growth (2006–2022)

(a) Output gap



(b) Potential growth



the 2019Q1 estimate of underlying growth to the estimate from the 2019 data vintage. This prevents the filtering procedure from anticipating the 2020 shock, which would manifest in a pre-emptive slowdown in potential output, and allows the shock to 'surprise' the model. In contrast, the SARB's latest model fixes the underlying growth rate to $g_t = 0$ for 2020Q2, which might give rise to the subtle differences between *Model 1* and the *SARB* model.

Model 5 demonstrates how the volatility in the 2020 data introduces parameter instability. This leads to implausible results such as the outsized negative output gap estimated for 2009–2013. As a temporary workaround, the first six specifications' parameters were estimated on data up to 2020Q1;[30] however, this is clearly not a long-term solution. Bayesian parameter estimation in the presence of massive data volatility is therefore an important area for future research.

These results are consistent with the South African literature. All models confirm that South African potential growth has been steadily declining since 2010 (Figure 3b). *Model 1* estimates that potential growth fell to 1.6% in 2009 after peaking at 4.5% in 2006. Potential growth briefly recovered to 3.3% in 2010, but steadily fell thereafter to 0.3% in 2019. The 2020 COVID-19 shock then saw potential growth fall sharply to -2.7%, before correcting to 2.8% in 2021.

Potential growth averaged 3.7% for the decade 2000–2009, but only 1.7% for the decade 2010–2019. These results are close to Ehlers, Mboji and Smal's (2013) potential growth estimates of 3.9% for 2000–2007[31] and identical to their estimate of 2.8% for 2008–2010. Similarly, Anvari, Ehlers and Steinbach (2014) estimate an average potential growth rate of 3.5% over 2000–2009. Kemp (2015) finds that potential growth declined from an average of 3.6% from 2000–2007 to an average of 2.6% from 2009 to 2014.

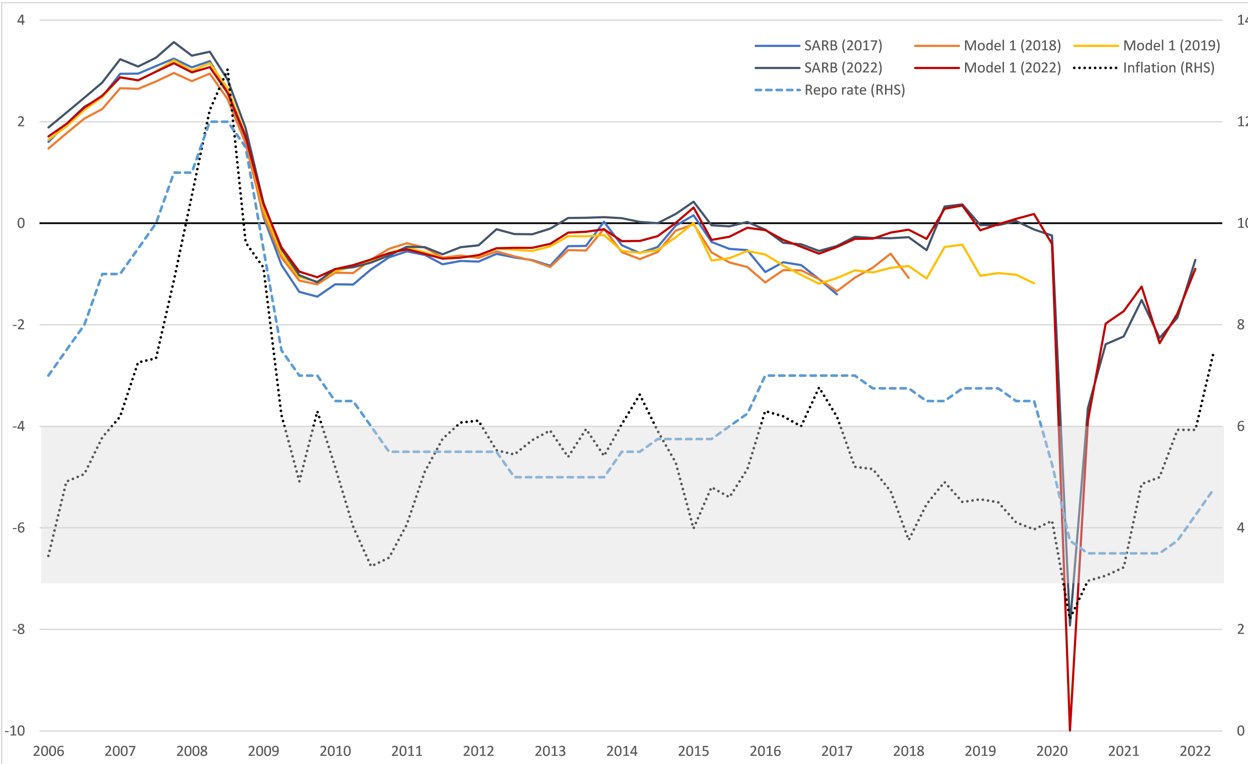
6.3 Output gap uncertainty and monetary policy

Real-time output gap estimations are exceedingly challenging. Figure 4 below illustrates the SARB's real-time estimates from 2017 and *Model 1* estimated on 2018 and 2019 vintage data, as compared to the SARB's current benchmark output gap and *Model 1* estimated on 2022 data from earlier, against the headline consumer price index (CPI) inflation rate and the repo rate.

The build-up to the global financial crisis (2006–2008) is associated with a strong positive

[30] The SARB employed the same strategy.
[31] Over the same sample this model estimates a potential growth rate of 4.0%.

Figure 4: Output gap revisions, inflation and the policy stance



Note: The shaded area represents the SARB's 3–6% inflation targeting band.

output gap[32] and rising inflation, which duly resulted in monetary policy tightening. The policy (repo) rate climbed from 7% in 2006Q1 to a high of 12% in 2008Q2. The post-crisis crash saw a sharply falling repo rate, returning to 7% by 2009Q3. Inflation returned to the target band, while the output gap turned negative. Weak economic growth and benign inflation enabled the SARB to maintain a relatively accommodative policy stance from 2010 to 2015.

In real time, a relatively wide negative output gap was estimated for 2016–2018, which, ceteris paribus, could have induced an even more accommodative monetary policy stance. At the same time, however, inflation accelerated, necessitating a tighter policy stance. However, the fact that inflation breached the upper band of its target in 2016–2017, coupled with the fact that the ex post output gap for the same period was, in fact, less negative than estimated in real time, suggests that monetary policy perhaps did not tighten enough, that is, the rate hiking cycle between 2014–2016 was not quite aggressive enough. It may be argued that the SARB's attempts to support a weak economy – which was in hindsight

[32] All these estimates are 'finance-neutral' specifications.

not as weak as was thought at the time – contributed to inflation briefly breaching its target in 2014 and 2016.

Similarly, the wider negative output gap estimated by *Model 1* as compared to the *SARB* model, coupled with inflation undershooting its target in 2020, suggests that the SARB could perhaps have been even more aggressive in lowering rates in response to the COVID-19 shock.

It should be noted, however, that the output gap is just one input in the MPRF, and with a much lower weight than the inflation forecast.[33] Moreover, while the Taylor rule serves as a guide for the policy stance, "actual monetary policy settings are not based on a mechanical Taylor rule" (Botha and Schaling 2020:20); other factors, including risks surrounding the forecast, are also considered.

7. Conclusion

This paper provided an overview of the empirical literature on potential output, and discussed various methodologies and empirical techniques by which potential output can be estimated. From the literature, three main paradigms for estimating potential output were identified: the HP filter, the PF approach, and the MVHP filter. The HP filter is an easy technique to employ, considering its relative simplicity and trivial data requirements. However, its well-documented vulnerabilities, including its end-of-sample bias and inability to incorporate economic information, should relegate it to be used as a last resort only. The notable advantage of the MVHP filter is that it simultaneously has the ability to incorporate structural economic information, while remaining less vulnerable to misspecification. Due to its coefficients being freely determined by the data, variables that do not have strong predictive power will be assigned a very low weight in the final estimation. Conversely, the strict imposition of coefficients in the PF approach makes this paradigm exceedingly vulnerable to misspecification of the underlying structural relationships and unreliable extraction of production factors' underlying trends. Newer techniques are also developing, following advances in computing power and statistical techniques. Some of these techniques, including nowcasting and DSGE modeling, may prove to be fruitful areas for future research.

The SARB's historic estimation approaches were then evaluated against the empirical literature. The SARB currently follows a semi-structural MVHP filter approach to estimating

[33] The weights on the inflation and output gaps in the SARB's Taylor rule MPRF are, respectively, 1.57 and 0.54 (SARB 2017).

potential output. While all models remain, at best, an approximation of reality, the MVHP filter is, given the data, technology and computing techniques at our disposal, the most suitable paradigm for the SARB to employ. This is because the MVHP paradigm is the most robust against misspecification bias and data revision, while it remains anchored in structural economic relationships, thus imparting economic interpretation and intuition to the output gap.

Finally, this paper presented an updated estimate of potential output and output gaps for the South African economy in light of the impact of the COVID-19 pandemic. Consistent with the South African empirical literature, this paper finds that potential growth has declined substantially over the last 15-odd years from a peak of 4.5% in 2006 to 0.3% in 2019. The COVID-19 shock then saw a massive simultaneous contraction in both actual and potential output during 2020. This paper estimates a negative output gap of -9.9% in the second quarter of 2020 (the ‘hard’ lockdown), and an annual potential output contraction of almost 2.7% in 2020. At the same time, actual output fell by 6.6%, leading to a large negative output gap of -4.1% for the year. Actual and potential output recovered into 2021, growing by 4.8% and 2.8%, respectively. The output gap, however, remained negative throughout 2021 and the early part of 2022. This suggests that the SARB’s easy monetary policy stance throughout 2020 and 2021, could in hindsight have been even looser. However, recent sharp increases in inflation have forced the SARB’s hand in tightening policy despite the still fledgling recovery in economic activity.

Extreme data volatility around the time of the COVID-19 shock poses significant estimation challenges. For now, this problem was overcome by cutting off the estimation sample right before the COVID-19 shock hit. However, this is not a sustainable approach. Having navigated through the COVID shock, important economic information, such as on perhaps new post-COVID structural relationships, will be lost. Incorporating Bayesian parameter estimation in the presence of massive outliers is an important area for future research.

Estimating potential output cannot be a purely mechanical process. Care needs to be taken to disentangle the impact of changing economic conditions on actual and potential output. In times of extreme volatility especially, monetary policy remains as much an art as a science.

References

Alberola, E, Gondo, R, Lombardi, M J and Urbina, D. 2016. 'Output gaps and policy stabilisation in Latin America: the effect of commodity and capital flow cycles'. BIS Working Paper No. 568 (June).

Alichi, A. 2015. 'A new methodology for estimating the output gap in the United States'. IMF Working Paper WP/15/144.

Álvarez, L J and Gómez-Loscos, A. 2018. 'A menu on output gap estimation methods'. *Journal of Policy Modeling* 40(4): 827–850.

Andersson, M, Szörfi, B, Tóth, M and Zorell, N. 2018. 'Potential output in the post-crisis period'. ECB Economic Bulletin *7*.

Anvari, V, Ehlers, N and Steinbach, R. 2014. 'A semi-structural approach to estimate South Africa's potential output'. SARB Working Paper WP/14/08.

Arsov, I and Watson, B. 2019. 'Potential growth in advanced economies'. *Reserve Bank of Australia Research Bulletin*, December: 50–60. Available at `https://www.rba.gov.au/publications/bulletin/2019/dec/potential-growth-in-advanced-economies.html`.

Basu, S and Fernald, J G. 2009. 'What do we know (and not know) about potential output?'. *Federal Reserve Bank of St. Louis Review*, July/August 4(91): 187–213.

Beneš, J, Clinton, K, Garcia-Saltos, R, Johnson, M, Laxton, D, Manchev, P B and Matheson, T. 2010. 'Estimating potential output with a multivariate filter'. IMF Working Paper WP/10/285.

Blagrave, P, Garcia-Saltos, R, Laxton, D and Zhang, F. 2015. 'A simple multivariate filter for estimating potential output'. IMF Working Paper WP/15/79.

Blanchard, O J and Quah, D. 1989. 'The dynamic effects of aggregate demand and supply disturbances'. *American Economic Review* 79(4): 655–673.

Bodnár, K, Le Roux, J, Lopez-Garcia, P and Szörfi, B. 2020. 'The impact of COVID-19 on potential output in the Euro area'. *ECB Economic Bulletin* 7.

Borio, C, Disyatat, P and Juselius, M. 2013. 'Rethinking potential output: embedding information about the financial cycle'. BIS Working Paper No. 404 (February).

Borio, C, Disyatat, P and Juselius, M. 2014. ‘A parsimonious approach to incorporating economic information in measures of potential output’. BIS Working Paper No. 442 (February).

Borio, C, Disyatat, P and Juselius, M. 2017. ‘Rethinking potential output: embedding information about the financial cycle’. *Oxford Economic Papers* 69(3).

Botha, B, De Jager, S, Ruch, F and Steinbach, R. 2017. ‘The Quarterly Projection Model of the SARB’. SARB Working Paper WP/17/01.

Botha, B, Ruch, F and Steinbach, R. 2018. ‘Short-lived supply shocks to potential growth’. SARB Working Paper WP/18/02.

Botha, B and Schaling, E. 2020. ‘Commodity prices and policy stabilisation in South Africa’. SARB Working Paper WP/20/11.

Burns, A F and Mitchell, W C. 1946. *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Chen, J and Górnicka, L. 2020. ‘Measuring output gap: is it worth your time?’. IMF Working Paper WP/20/24.

Congdon, T. 2008. ‘Two concepts of the output gap’. *World Economics* 9(1): 147–175.

Du Plessis, S, Smit, B and Sturzenegger, F. 2008. ‘Identifying aggregate supply and demand shocks in South Africa’. *Journal of African Economies* 17(5): 765–793.

Du Toit, C B, Ground, M and Van Eyden, R. 2006. ‘Does South Africa have the potential and capacity to grow at 7 per cent?: a labour market perspective’. University of Pretoria Working Paper 2006-03.

Du Toit, L. 2008. ‘Optimal HP filtering for South Africa’. Stellenbosch Economic Working Paper 07/08.

Edge, R M and Rudd, J B. 2016. ‘Real-time properties of the Federal Reserve’s output gap’. *Review of Economics and Statistics* 98(4): 785–791.

Ehlers, N, Mboji, L and Smal, M. 2013. ‘The pace of potential output growth in the South African economy’. SARB Working Paper WP/13/01.

Fedderke, J W and Mengisteab, D K. 2017. ‘Estimating South Africa’s output gap and potential growth rate’. *South African Journal of Economics* 85(2): 161–177.

Hodrick, R J and Prescott, E C. 1997. ‘Postwar US business cycles: an empirical investigation’. *Journal of Money, Credit, and Banking* 29(1): 1–16.

Janse van Rensburg, T, Fowkes, D and Visser, E. 2019. ‘What happened to the cycle? – Reflections on a perennial negative output gap’. SARB Economic Note (2019-05).

Kemp, J H. 2015. ‘Measuring potential output for the South African economy: embedding information about the financial cycle’. *South African Journal of Economics* 83(4): 549–568.

Kemp, J H and Smit, B. 2016. ‘Estimating and explaining changes in potential growth in South Africa’. *Studies in Economics and Econometrics* 40(3): 21–38.

Klein, N. 2011. *Measuring the potential output of South Africa*. Washington, D.C.: IMF.

Kuttner, K N. 1994. ‘Estimating potential output as a latent variable’. *Journal of Business & Economic Statistics* 12(3): 361–368.

Laxton, D and Tetlow, R. 1992. ‘A simple multivariate filter for the measurement of potential output’. Bank of Canada Technical Report No. 59.

Lienert, A and Gillmore, D. 2015. ‘The Reserve Bank’s method of estimating “potential output”’. Reserve Bank of New Zealand Analytical Note AN2015/01.

Melolinna, M and Tóth, M. 2016. ‘Output gaps, inflation and financial cycles in the United Kingdom’. Bank of England Staff Working Paper No. 585.

Mishkin, F S. 2007. ‘Estimating potential output’. In *Conference on Price Measurement for Monetary Policy, Federal Reserve Bank of Dallas, Dallas, Texas, May 24, 2007.* Available at `https://www.federalreserve.gov/newsevents/speech/mishkin20070524a.htm`.

Nelson, C R and Plosser, C R. 1982. ‘Trends and random walks in macroeconomic time series: some evidence and implications’. *Journal of Monetary Economics* 10(2): 139–162.

Okun, A M. 1962. ‘Potential GNP: its measurement and significance’. In *Proceedings of the Business and Economic Statistics Section*, pp. 98–104. Minneapolis, Minn.: American Statistical Association.

Okun, A M. 1970. *The political economy of prosperity*. Washington, D.C.: Brookings Institution.

Orphanides, A and Van Norden, S. 2002. ‘The unreliability of output-gap estimates in real time’. *Review of Economics and Statistics* 84(4): 569–583.

Republic of South Africa. 1996. ‘The Constitution of the Republic of South Africa’. Available at https://www.justice.gov.za/legislation/constitution/saconstitution-web-eng.pdf.

SARB. 2017. *Monetary Policy Review (October 2017)*. Pretoria: SARB. Available at https://www.resbank.co.za/content/dam/sarb/publications/reviews/monetary-policy-review/2017/8017/MPROctober2017.pdf.

SARB. 2021. *Monetary Policy Review (April 2021)*. Pretoria: SARB. Available at https://www.resbank.co.za/content/dam/sarb/publications/monetary-policy-review/2021/MPRApr21Internet.pdf.

Scacciavillani, F and Swagel, P. 1999. *Measures of potential output: an application to Israel*. Washington, D.C.: IMF.

Smit, B and Burrows, L. 2002. ‘Estimating potential output and output gaps for the South African economy’. Stellenbosch Economic Working Paper No. 5/2002.

Steenkamp, D. 2018. ‘Productivity estimates for South Africa from CES production functions’. SARB Working Paper WP/18/05.

Vetlov, I, Pisani, M, Hlédik, T, Jonsson, M and Kucsera, H. 2011. ‘Potential output in DSGE models’. ECB Working Paper No. 1351 (June).