Alternative Methods of Estimating Potential Output and the Output Gap: An Application to Sweden

Valerie Cerra and Sweta Chaman Saxena
IMF Working Paper

European I Department

Alternative Methods of Estimating Potential Output and the Output Gap: An Application to Sweden

Prepared by Valerie Cerra and Sweta Chaman Saxena

Authorized for distribution by Sharmini Coorey

March 2000

Abstract

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

This paper reviews a number of different methods that can be used to estimate potential output and the output gap. Measures of potential output and the output gap are useful to help identify the scope for sustainable noninflationary growth and to allow an assessment of the stance of macroeconomic policies. The paper then compares results from some of these methods to the case of Sweden, showing the range of estimates.

JEL Classification Numbers: E32, C22, C32

Keywords: Sweden, Business Cycles, Output Gap, Potential Output, Detrending, Production Function, Vector Autoregression, Unobserved Components Models

Authors’ E-Mail Addresses: vcerra@imf.org; ssaxena+@pitt.edu

1 European I Department, IMF; and University of Pittsburgh, respectively. The authors would like to thank Claes Berg, Jim Boughton and Alun Thomas for helpful comments and suggestions. The usual disclaimers apply.
Contents

I. Introduction ..................................................................................................................... 3

II. Approaches to Estimating Potential Output and the Output Gap ........................................ 4
    A. The Rodrick-Prescott filter ................................................................................... 4
    B. The unobserved components methods .................................................................. 5
        Beveridge-Nelson decomposition ........................................................................ 5
        Univariate unobserved components model ....................................................... 6
        Bivariate unobserved components model .......................................................... 7
        Common permanent and temporary components ............................................... 8
        Common components with asymmetric growth rates ........................................... 9
    C. The structural VAR approach by Blanchard and Quah ......................................... 10
    D. The production function approach ....................................................................... 13
    E. Demand-side model ............................................................................................. 14
    F. System estimates of potential output and the NAIRU .......................................... 14

III. Empirical Estimates of Potential Output and the Output Gap ......................................... 16
    A. HP filter ............................................................................................................ 16
    B. Simple unobserved components methods ............................................................ 16
    C. Common permanent and cyclical components ..................................................... 20
    D. Structural VAR ................................................................................................... 20
    E. Production function ........................................................................................... 23
    F. Systems unobserved components ......................................................................... 26

IV. Conclusions ................................................................................................................... 30

References ................................................................................................................................ 31

Tables

Figures
1. HP filter with smoothing parameters 100, 200, and 1000 ............................................ 18
2. Bivariate unobserved components models with unemployment or inflation ............... 19
3. Common permanent and cyclical components .......................................................... 21
4. Probabilities of permanent and temporary output losses ........................................... 22
5. Structural VAR with output, the real exchange rate, and relative prices ................... 24
6. Structural VAR with output, unemployment, and inflation ................................... 25
7. Production function approach ................................................................................. 27
8. TFP and labor contributions to the levels of GDP and potential GDP ...................... 28
9. System unobserved components method .................................................................. 29
I. INTRODUCTION

This paper describes a variety of methodologies for estimating a country's potential output level and presents empirical estimates for Sweden. The paper explains why these methods produce a variety of results, some of which are more plausible than others. The usefulness of estimating the output gap is to help identify the scope for sustainable noninflationary growth and allow an assessment of the stance of macroeconomic policies.\footnote{The output gap is defined as actual output minus potential output relative to potential output, \((y-y*)/y^*\), in percent.}

The level of actual output relative to the potential level of output determines whether economic policy should be directed toward raising aggregate demand or whether structural issues should be given more prominence.

Very little empirical research has been done to estimate Sweden's output gap. Exceptions are papers by Apel, Hanssen, and Lindberg (1996) and Apel and Jansson (1997) which estimated output gaps for Sweden through 1996. The former use an HP filter, production function, and unobserved components approach and find output gaps ranging from about \(-\frac{1}{2}\) percent to \(-2\frac{1}{2}\) percent in 1996. The latter use a system unobserved components approach and find output gaps ranging from roughly \(-3\) to \(-8\) in 1996 depending on the measure of inflation and the equation specification used in the estimation. (These methodologies are described below.) In addition to these in-country analyses, a recent issues paper by the European Commission (May 1999) compares output gaps for European countries through 1998 using an HP filter and other trend estimation methods (band pass filter, linear time trend, and production function). Results from these trend estimation methods indicate that the output gap in Sweden was nearly closed in 1998, except that the production function method shows an output gap of about \(-1\) percent (+1 percent when including a trend break).\footnote{Thomas (IMF 99) reveals that the strong improvement in public finances since the early 1990s has added about 0.3 percentage points per annum to pre-existing estimates of potential output growth.}

The definition and estimation of the trend and cyclical components of output raise a number of theoretical and empirical questions, which reflect the ongoing controversy over the origins of economic fluctuations. As potential output is an unobserved variable, a number of statistical and economic approaches have been developed to estimate it and the corresponding output gap. Since such measures are known to be fairly uncertain, this paper presents estimates derived from different techniques, highlighting the sensitivity of the results to alternative methodologies.
The paper is organized as follows. Section II briefly describes alternative methods for estimating potential output and the output gap and related determinants of potential output, such as the non-accelerating inflation rate of unemployment (NAIRU). Section III presents the estimates obtained for Sweden. Section IV summarizes the main findings.

II. APPROACHES TO ESTIMATING POTENTIAL OUTPUT AND THE OUTPUT GAP

There are two basic methodologies for estimating potential output: statistical detrending and estimation of structural relationships. The former attempt to separate a time series into permanent and cyclical components; the latter attempt to isolate the effects of structural and cyclical influences on output, using economic theory. The set of statistical methods discussed here include the HP filter, Beveridge-Nelson, and various unobserved components methods (univariate, bivariate, and common permanent and cyclical components). Methods that use economic theory to identify the output gap include the structural VAR, production function, demand-side model, and multivariate system models.

A. The Hodrick-Prescott filter

The Hodrick-Prescott (HP) filter is a simple smoothing procedure that has become increasingly popular because of its flexibility in tracking the characteristics of the fluctuations in trend output. Trend output (denoted by $y^*$) derived using the HP filter is obtained by minimizing a combination of the gap between actual output ($y$) and trend output and the rate of change in trend output for the whole sample of observations ($T$):

$$\text{Min} \sum_{i=0}^{T} (y_i - y_i^*)^2 + \lambda \sum_{i=2}^{T-1} (y_{i+1}^* - y_i^*) - (y_i^* - y_{i-1}^*),$$

where $\lambda$ determines the degree of smoothness of the trend.

The properties and shortcomings of the HP filter have been well documented (Harvey and Jaeger, 1993). A major drawback comes from the difficulty in identifying the appropriate detrending parameter $\lambda$—which is generally overlooked by using arbitrary values popularized by the real business cycle literature. Mechanical detrending based on the HP filter can lead to spurious cyclicality with integrated or nearly integrated time series and an excessive smoothing of structural breaks. A second important flaw of the HP filter arises from its high end-sample biases, which reflect the symmetric trending objective of the method across the whole sample and the different constraints that apply within the sample and at its edges. This flaw is particularly severe when the focus of attention is directed at the most recent observations in the sample in an effort to draw conclusions for policy implementation and make projections for the immediate future.
B. The unobserved components methods

The unobserved components method is an approach to estimating unobserved variables such as potential output and the NAIRU using information from observed variables. This approach has the advantage that explicit relationships can be specified between output, unemployment, and inflation. The relationships are first written in state space form. State-space form is a general way of representing dynamic systems such that the observed variables are specified as a function of the unobserved state variables in the measurement (or observation) equation and a separate transition equation specifies the autoregressive process for the state variables. Once a dynamic time series model is written in a state-space form, the unobserved state vector can be estimated using a recursive algorithm known as the Kalman filter that uses guesses for the unobserved variables to create predictions for the observed variables and then updates the guesses based on the prediction errors. The approach has the disadvantage of requiring considerable programming. In addition, results are often sensitive to the initial guesses for the parameters.

Beveridge-Nelson decomposition

A Beveridge-Nelson decomposition is a detrending method using unobserved components. Output is assumed to contain unobserved permanent and temporary components consisting of a random walk with drift and a stationary autoregressive process, respectively.

Consider an ARMA(p,q) model for the changes in output:

\[ \phi(L)\Delta y_t = c + \theta(L)e_t, \quad e_t \sim iid(0, \sigma^2), \] \hspace{1cm} (2)

\[ \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p \] \hspace{1cm} (3)

\[ \theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q \] \hspace{1cm} (4)

and where \(|\phi| < 1\) and \(|\theta| < 1\). The ARMA model can be written in its moving average representation (Wold form) as:

\[ \Delta y_t = \mu + \Psi(L)e_t \quad \text{where} \quad \Psi(L) = \phi(L)^{-1}\theta(L) = \sum_{j=0}^{\infty} \psi_j L^j. \] \hspace{1cm} (5)

The Beveridge-Nelson decomposition is given by:

---

\(^4\) Refer to Kim and Nelson (1999) for the technical details of the Kalman filter.
\[ y_t = y_0 + \delta t + \Psi(1) \sum_{j=1}^{t} \varepsilon_j + \tilde{\varepsilon}_t \quad (6) \]

where \( \tilde{\varepsilon}_t = \Psi(L)\varepsilon_t \), \( \Psi(L) = \sum_{k=0}^{\infty} \psi_k L^k \), \( \psi_k = -\sum_{j=0}^{\infty} \psi_j \),

\[ TD_t = y_0 + \delta t = \text{deterministic trend}, \quad (7) \]

\[ TS_t = \sum_{j=1}^{t} \varepsilon_j = \text{stochastic trend}, \]

and \( C_t = \tilde{\varepsilon}_t = \text{temporary or cyclical component} \quad (8) \)

To proceed with the decomposition, an ARMA(p,q) is estimated on the changes in output. Various ARMA models are estimated up to an ARMA (2,2) and the Schwarz criterion is used to select the best model. Then the series is decomposed into stationary and trend components using the BN decomposition technique described above.

**Univariate unobserved components model**

Unlike the Beveridge-Nelson decomposition, the UC model decomposes the series \( y_t \) into two independent components: a stochastic trend component, \( y_{1t} \) and a cyclical component, \( y_{2t} \). That is, whereas shocks to the two components are perfectly correlated in the BN decomposition, the shocks are uncorrelated in the univariate unobserved model.

\[ y_t = y_{1t} + y_{2t} \quad (9) \]

\[ y_{1t} = \delta + y_{1,t-1} + e_{1t} \quad (10) \]

\[ y_{2t} = \phi_1 y_{2,t-1} + \phi_2 y_{2,t-2} + e_{2t} \quad (11) \]

\[ e_{it} \sim i.i.d. N(0, \sigma_i^2), \quad i = 1, 2, \quad E[e_{1t} e_{2s}] = 0 \text{ for all } t \text{ and } s \]

where the roots of \( (1 - \phi_1 L - \phi_2 L^2) = 0 \) lie outside the unit circle. Taking both \( y_{1t} \) and \( y_{2t} \) as unobserved state variables, this model could be written in the state-space form as follows:

\[ y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{2,t-1} \end{bmatrix} \quad (12) \]
The model is then estimated using the Kalman filter.

**Bivariate unobserved components model**

Through a bivariate unobserved components model, definitions of potential output, the NAIRU, and the inflation rate can be explicitly incorporated in the decomposition and can be simultaneously estimated. Following Clarke (1989), the cyclical movement in output is measured using a bivariate unobserved components model, where output and unemployment (or alternatively inflation) each have their own trend components, but the cyclical component is common to the two series. Assume that output, \( y_t \), contains a stochastic trend, \( n_t \), and a stationary cyclical component, \( x_t \). The unemployment rate, \( z_t \), has a trend component, \( L_t \), and a stationary component, \( C_t \). The model is:

\[
y_t = n_t + x_t
\]

(14)

\[
n_t = \delta + n_{t-1} + v_t, \quad v_t \sim i.i.d. N(0, \sigma_v^2)
\]

(15)

\[
x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + e_t, \quad e_t \sim i.i.d. N(0, \sigma_e^2)
\]

(16)

\[
z_t = L_t + C_t
\]

(17)

\[
L_t = L_{t-1} + v_{lt}, \quad v_{lt} \sim i.i.d. N(0, \sigma_{Lt}^2)
\]

(18)

\[
C_t = \alpha_0 x_t + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + e_{ct}, \quad e_{ct} \sim i.i.d. N(0, \sigma_{et}^2)
\]

(19)

All errors are white noise. The cyclical component of unemployment, \( C_t \), is assumed to be a function of the current and past transitory components of output.

Treating all the variables in the above model as unobserved, the model could be represented in the state-space form as follows:

\[
\begin{bmatrix}
y_t \\
z_t
\end{bmatrix} =
\begin{bmatrix}
1 & 1 & 0 & 0 & 0 \\
0 & \alpha_0 & \alpha_1 & \alpha_2 & 1
\end{bmatrix}
\begin{bmatrix}
n_t \\
x_t \\
x_{t-1} \\
x_{t-2} \\
L_t
\end{bmatrix}
+ \begin{bmatrix}
e_t \\
e_{ct}
\end{bmatrix}
\]

(20)
One advantage of using the inflation rate is the closer linkage of the constructed output gap to its definition based on stable inflation. Cyclically high rates of inflation would be correlated with the cyclical component of output through the component $x_t$.

Common permanent and temporary components

The idea that economic times series have common components dates back to the landmark study of Burns and Mitchell (1946) which found that the business cycle is characterized by simultaneous comovement in many economic activities. A generalization of the above bivariate unobserved (cyclical) component model would include both common permanent and common cyclical factors. Examples of specifications using common permanent factors abound in the business cycle literature following the econometric formalization of the idea by Stock and Watson (1989, 1991, 1993). The focus on permanent components of output followed from the strand of empirical literature which provided evidence that output follows a stochastic trend. Such evidence implies that shocks to growth persist and therefore a recession permanently reduces output. An additional implication of the stochastic trend view is that since there is no temporary component to recessions and expansions, the output gap cannot be defined. However, a series of recent papers challenges the view that changes in output are permanent (Wynne and Balke (1992), Beaudry and Koop (1993), and Sichel (1994)). This recent evidence points to a strong recovery phase following a recession, an idea explicit in Friedman’s (1964, 1993) “plucking” model of the business cycle, whereby output is plucked down from its trend growth during recessions (as in the path of a string attached to the underside of a board) and then springs back to the upper limit set by resource constraints. This section takes an agnostic approach to the debate by permitting both temporary and permanent factors.

Formally, each individual time series $Y_i$ (in logs), for $i=1, \ldots, N$, consists of a deterministic time trend $DT_i$, a stochastic permanent component with a unit root $P_i$, and a transitory component $T_i$. Each series can be written as:

$$Y_i = DT_i + P_i + T_i$$

(22)
where \(N_t\) and \(x_t\) are the common permanent and common transitory components, respectively; \(\zeta_{it}\) and \(\omega_{it}\) are the idiosyncratic permanent and transitory components, respectively. The \(\gamma_i\) terms are permanent factor loadings, and indicate the extent to which each series is affected by the common permanent component, \(N_t\). Similarly, the transitory factor loadings, \(\lambda_i\), indicate the extent to which each series is affected by the common transitory component, \(x_t\).

Taking first differences, the model can be written in deviations from means:

\[
\Delta y_{it} = \gamma_i \Delta n_t + \lambda_i \Delta x_t + \zeta_{it} + \omega_{it}
\]

where \(\Delta y_{it} = \Delta Y_{it} - \Delta \bar{Y}_i\), \(\Delta n_t = \Delta N_t - \phi(1)^{-1} \delta\), and \(\zeta_{it} = \Delta \zeta_{it} + \Delta \omega_{it}\).

\[
\phi(L) \Delta n_t = v_t, v_t \sim iid \ N(0,1)
\]

\[
\phi^*(L) x_t = u_t, u_t \sim iid \ N(0,1)
\]

Equations (27) and (28) describe the dynamics of the permanent and transitory components, respectively.

**Common components with asymmetric growth rates**

The other regularity of the business cycle documented by Mitchell (1927) is asymmetry—the idea that expansions are fundamentally different from recessions in their duration and in the abruptness of changes in growth: “Business contraction seems to be a briefer and more violent process than business expansion.” Therefore, an extension of the model above permits asymmetry in the growth rates. Asymmetry is obtained by specifying regime switching in the common permanent component, as proposed by Hamilton (1989), or regime switching in the common temporary component, as advocated by Kim and Nelson (1999).

Formally, the above model can be altered as follows.
\[ \phi(L) \Delta n_t = \mu_{S_{i,t}} + v_t, \ v_t \sim iid \ N(0,1) \quad (29) \]
\[ \mu_{S_{i,t}} = \mu_0 + \mu_1 S_{i,t}, \ S_{i,t} = \{0,1\} \quad (30) \]
\[ \Pr[S_{i,t} = 0 | S_{i,t-1} = 0] = q_1, \ \Pr[S_{i,t} = 1 | S_{i,t-1} = 1] = p_1. \quad (31) \]

\( S_{i,t} \) is a latent Markov-switching state variable that switches between 0 and 1 with transition probabilities given by equation (31). The common permanent component, \( n_t \), grows at rate \( \phi(1)^{-1}(\mu_0) \) when \( S_{i,t} = 0 \), and at rate \( \phi(1)^{-1}(\mu_0 + \mu_1) \) when \( S_{i,t} = 1 \).

\[ \phi^*(L)x_t = \tau_{S_{2,t}} + u_t, \ u_t \sim iid \ N(0,1) \quad (32) \]
\[ \tau_{S_{2,t}} = \tau_{S_{2,t}} S_{2,t}, \ S_{2,t} = \{0,1\} \quad (33) \]
\[ \Pr[S_{2,t} = 0 | S_{2,t-1} = 0] = q_2, \ \Pr[S_{2,t} = 1 | S_{2,t-1} = 1] = p_2. \quad (34) \]

\( S_{2,t} \) is a latent Markov-switching state variable, independent of \( S_{i,t} \), whose transitions are governed by the probabilities in equation (34). The term, \( \tau \), is the size of the 'pluck'. If \( \tau < 0 \), then the transitory component is plucked down during a recession. Following the pluck then there is a tendency for output to revert to its previous peak.

The idiosyncratic components of each series have the following autoregressive structure:
\[ \psi_i(L)z_t = e_{it} + \epsilon_{it} \sim iid \ N(0,\sigma_i^2). \quad (35) \]

The innovation variances of the common components have been normalized to unity to identify the model; all innovations are assumed to be mutually and serially uncorrelated at all leads and lags; and the roots of \( \phi(L) = 0 \), \( \phi^*(L) = 0 \), and \( \psi_i(L) = 0 \) lie outside the unit circle.

The common growth component, \( \Delta n_t \), and the common transitory component, \( x_t \), can be governed by one state variable \( S_{i,t} \) or two different state variables, \( S_{i,t} \) and \( S_{2,t} \). The latter relaxes the assumption that all recessions are alike by permitting different episodes to be based on temporary or permanent causes.

C. The structural VAR approach by Blanchard and Quah

This method stems from the traditional Keynesian and neoclassical synthesis, which identifies potential output with the aggregate supply capacity of the economy and cyclical
fluctuations with changes in aggregate demand. Based on a vector autoregression (VAR) for output growth and unemployment, Blanchard and Quah (1989) identify structural supply and demand disturbances by assuming that the former have a permanent impact on output, while the latter can have only temporary effects on it.\(^6\) The analysis can be extended to include temporary nominal shocks by including a price variable that is affected by nominal shocks in the short and long runs. A similar approach has been used by Clarida and Gali (1994) to estimate the effects of supply, demand, and nominal shocks on relative output, the real exchange rate, and the price level of the home country relative to its trading partners.

Formally, the method, as applied to the Clarida and Gali model, expresses the log of relative output, the log of the real exchange rate and the log of relative CPIs as a vector in first differences (assuming that both variables are difference stationary):

\[
\Delta x_t = \begin{bmatrix} \Delta y_t & -\Delta r_t, \Delta R_{er} & \Delta p_t - \Delta p_t^* \end{bmatrix}
\]

The vector has a moving-average structural representation given by:

\[
\Delta X_t = C(L)e_t
\]

where \(L\) is the lag operator and \(e_t = [e_n, e_d, e_a]\) is a vector of exogenous, unobserved structural shocks; \(e_n\) is the aggregate supply shock, \(e_d\) is the aggregate demand shock and \(e_a\) is the aggregate nominal shock. The errors are serially uncorrelated and have a variance-covariance matrix normalized to the identity matrix. Since the vector of structural shocks is not observed directly, the trick is to recover \(e_t\) by estimating an unrestricted VAR, which can be inverted to yield the moving-average representation:

\[
\Delta X_t = A(L)u_t
\]

The first matrix in the polynomial \(A(L)\) is the identity matrix and \(u_t\) is a vector of reduced form residuals with the covariance matrix \(\Sigma\). Equations 37 and 38 imply a linear relationship between the reduced form residuals and the shocks of the structural model:

\[
u_t = C_0 e_t
\]

It is necessary to identify the 3x3 matrix \(C_0\) to be able to recover the vector of structural shocks \(e_t\) from the estimated disturbance vector \(u_t\). The symmetric matrix \(\Sigma = C_0 C_0'\) imposes six of the nine restrictions that are required, and therefore we need only three more identifying restrictions. Blanchard and Quah (1989) suggest that we can use economic theory to impose these restrictions. Economic theory has a number of implications regarding the

\(^6\)Extensions of the method and other types of identification can be found in King et al. (1991) and Bayoumi and Eichengreen (1992).
long-run behavior of variables in response to shocks and therefore imposing these long-run restrictions allows us to properly identify the shocks.

The long-run representation of equation (37) can be written as:

\[
\begin{bmatrix}
\Delta Y \\
\Delta RER \\
\Delta CPI
\end{bmatrix} =
\begin{bmatrix}
C_{11}(1) & C_{12}(1) & C_{13}(1) \\
C_{21}(1) & C_{22}(1) & C_{23}(1) \\
C_{31}(1) & C_{32}(1) & C_{33}(1)
\end{bmatrix}
\begin{bmatrix}
\varepsilon_x \\
\varepsilon_d \\
\varepsilon_n
\end{bmatrix}
\]

(40)

where \( C(1) = C_0 + C_1 + \ldots \) is the long-run effect of \( \varepsilon_t \) on \( \Delta X \).

Using Clarida and Gali's (1994) identifying assumptions, the long-run restrictions imposed on the model are \( C_{12} = 0, C_{13} = 0, \) and \( C_{23} = 0 \). These restrictions make the matrix \( C(1) \) upper triangular and we can use this fact to recover \( C_0 \). The restrictions imply that in the long-run, output is affected only by supply shocks. The nominal shock can be distinguished from demand shocks by the fact that only the latter impacts the real exchange rate in the long run. Nominal shocks have permanent effects only on the price level.

The residuals from the unrestricted VAR and the estimated parameters of \( C_0 \) can be used to construct the vector of exogenous structural shocks. Since potential output corresponds to the permanent component of output in the system, the equation for the change in potential output can be derived using the vector of supply shocks:

\[
\Delta Y_t = \mu_y + C_{11}(L)\varepsilon_x
\]

(41)

where \( \mu_y \) is the (previously ignored) linear trend in output.

Compared with other multivariate detrending techniques, this method relies on clear theoretical foundations and does not impose undue restrictions on the short-run dynamics of the permanent component of output. In particular, the estimated potential output is allowed to differ from a strict random walk (Dupasquier et al., 1997). In addition, the output gap estimates derived by this method are not subject to any end-sample biases. One obvious drawback of this approach is that the identification chosen may not be appropriate in all circumstances. This is true when changes in the real exchange rate (in the Clarida and Gali model) or the unemployment rate (in the Blanchard-Quah model) do not provide good

---

\(^7\) Potential output can be described as following a random walk if production inputs evolve as a stochastic trend, for example, if productivity growth depends on the stochastic arrival of new technologies.
indications of cyclical developments in output. Standard deviations of the output gap estimates also suggest that these measures are particularly uncertain. 8

This approach is also limited by its ability to identify at most only as many types of shocks as there are variables. Moreover, the model of Clarida and Gali assumes that the orthogonally constructed exogenous innovations correspond to pure uncorrelated supply, demand and nominal shocks. However, economic theory could identify many types of shocks with varying supply, demand or nominal characteristics. For example, a technological advancement identified as a supply shock may simultaneously increase demand owing to wealth or demonstration effects. An increase in government spending on productive infrastructure is associated with a demand shock, but would likely also have long-run supply-side effects. As a consequence, it is often difficult to relate the composite pure shocks to specific economic variables. Economic theory can also have divergent implications for the effects of shocks on economic variables. The theoretical foundation for the identification of shocks in the Clarida Gali analysis stems from the Mundell-Fleming model. According to this model, a positive demand shock leads to an appreciated exchange rate in the long run. However, in models with traded and nontraded sectors, the long-run effect of a demand shock (from say higher government spending) on the real exchange rate depends on which sector spending is directed toward. As a result of these identification problems, the structural VAR may not produce results that appropriately correspond to the assumptions.

D. The production function approach

In its simplest form, this approach postulates a simple two-factor Cobb-Douglas production function for the business sector (Giorno et al., 1995):

$$\ln(Y^*) = c + \alpha \ln(L^*) + (1-\alpha) \ln(K^*) + tfp + e$$  \hspace{1cm} (42)

where $Y$, $L$, and $K$ are the value added, employment, and capital stock of the business sector, respectively; $tfp$, the trend total factor productivity (in log form); $c$, a constant; and $e$, the residual.

With parameter $\alpha$ approximated by labor's share in value added, the contributions of labor and capital to output can be computed and subtracted from the value added of the business sector (in log form). The trend total factor productivity is then derived by smoothing the residuals of the equation.

Potential output for the business sector is then computed as:

$$\ln(Y^*) = c + \alpha \ln(L^*) + (1-\alpha) \ln(K^*) + tfp$$  \hspace{1cm} (43)

8See Staiger et al. (1997) for a discussion of estimates of the NAIRU.
where $L^*$ is the trend labor input of the business sector calculated as:

$$L^* = P_{wa} \cdot Part^* \cdot (1 - NAIRU) - EG,$$  \hspace{1cm} (44)

with the trend labor force constructed as the product of $P_{wa}$ (the working age population) and $Part^*$ (the trend participation rate). $EG$ is employment in the government sector. Potential output for the whole economy is then computed by assuming that output of the government sector—measured by the government wage bill—is always at its potential.

Compared with other methods, the production function approach can provide useful information on the determinants of potential growth. This approach relies, however, on an overly simplistic representation of the production technology, and the estimates of potential output and the output gap are crucially dependent on the NAIRU estimates and sensitive to the detrending techniques used for smoothing the components of the factor inputs. Problems of trend elimination for GDP is shifted to the trend estimates of the inputs. For example, the estimates from the production function approach share the end-sample biases that affect the underlying detrending techniques that are used for labor, capital, and productivity. These estimates may also be affected by measurement errors in factor inputs, particularly in the capital stock. Moreover, the production function approach may suffer from omitted variable bias due to improper use of value-added data and imperfectly competitive output markets. In addition, the so-called labor-hoarding hypothesis emphasizes transaction costs of adjustments in the labor force: firms may find it profitable to substitute labor utilization rates for measured labor input when the labor force cannot be modified without costs, and as a result effort levels may change over the business cycle instead of measured inputs.

E. Demand-side model

Bayoumi (1999) proposes estimating the output gap directly from measures of slack in the economy such as the unemployment rate, the ratio of job seekers to job offers, and capacity utilization. The output gap is the fitted value from these measures in a regression that includes a polynomial time trend. This construction assumes that these measures of slack are contemporaneous with the cycle. However, since unemployment, for example, consists of both structural and cyclical components, this method overestimates the output gap unless the structural components of the slack variables are first removed from the totals.

F. System estimates of potential output and the NAIRU

System methods specify relationships between output, unemployment, and inflation on the basis of economic relationships suggested by theory. Adams and Coe (1990) jointly estimate potential output and the natural rate of unemployment based on a system of simultaneous equations for output, unemployment, inflation, and wages. They first estimate single equation preliminary specifications. Later the system is estimated using three-stage least squares where trend output and productivity used in the preliminary equations are
replaced by potential output and productivity derived from a production function. The natural rate of unemployment is incorporated in the wage equation and is consistent with the measure of potential output.

An alternative approach is to use an unobserved components model which simultaneously constructs the unobserved variables of potential output and the NAIRU. The estimated levels of these variables can be explicitly linked to their definitions of the level of output and the level of unemployment in which inflation is stable. The unobserved components models described above are examples of this approach with output, unemployment, and inflation having a common cycle and/or common stochastic trend. Kuttner (1994) specifies a bivariate model with output and inflation similar to that above linking output and inflation through a Phillips curve. The deviation of output from its potential is related to inflation through a common cycle. Apel and Jansson (1997) specify a more general model that jointly estimates potential output and the NAIRU using the Phillips curve relationship and Okun’s Law, and also includes exogenous determinants of inflation. The system can be represented by the measurement equation linking output, unemployment, and changes in inflation to potential output, NAIRU, and a cycle along with other exogenous variables:

\[
\begin{bmatrix}
Y_t \\
U_t \\
\Delta\pi_t
\end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & \eta \end{bmatrix} \begin{bmatrix} Y_t^P \\
U_t^N \\
U_t - U_t^N \\
U_{t-1} - U_{t-1}^N
\end{bmatrix} + A Z_t + \begin{bmatrix} \epsilon_t^o \\ 0 \\ \epsilon_t^i
\end{bmatrix}
\]  \hspace{1cm} (45)

The first equation in the system represents Okun’s Law. Actual GDP consists of potential GDP and a cyclical component that depends on the deviation of unemployment from the natural rate. Viewed in terms of a production function approach, this assumes that labor participation, productivity and the capital stock are at their trend levels. The third equation is the Phillips curve. Changes in inflation depend on the demand-side determinants that affect the deviation of unemployment from NAIRU and supply-side shocks (Z). The unobserved series evolve according to the transition equation:

\[
\begin{bmatrix}
Y_t^P \\
U_t^N \\
U_t - U_t^N \\
U_{t-1} - U_{t-1}^N
\end{bmatrix} = \begin{bmatrix} \delta \\ 0 \\ 0 \\ 0 
\end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & \lambda_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\
U_{t-1} \\
U_{t-1} - U_{t-1}^N \\
U_{t-2} - U_{t-2}^N
\end{bmatrix} + \begin{bmatrix} \epsilon_t^o \\ \epsilon_t^u \\ \epsilon_t^\pi \\ 0
\end{bmatrix}
\]  \hspace{1cm} (46)

Potential output and the NAIRU follow random walks, while the cycle is modeled as a second order autoregressive process.
III. Empirical Estimates of Potential Output and the Output Gap

The estimations for Sweden use seasonally adjusted quarterly data on GDP and other variables as required, including inflation and unemployment, except for the HP filter and the production function approach. The sources for data are the International Financial Statistics, OECD, and the national authorities. The HP filter was implemented on a series of annual observations in order to enlarge the sample with medium-term staff projections from the World Economic Outlook (WEO), so that end-point biases for the most recent years would be minimized. The production function approach also used annual data to maintain consistency with a companion study on Sweden (Thomas (1999)).

A. HP filter

WEO projections assume growth rates of 2.8 percent in 1999, 2.9 percent in 2000, 3 percent in 2001, 2.5 percent in 2002, leveling off to 2.2 percent in 2003–4. Using the standard value of the smoothing parameter for annual observations, 100, the output gap was estimated at −0.6 percent in 1998 (Table 1). Potential output grew by 2.1 percent from 1997 to 1998 and is projected to grow by 2.3 percent in 1999. However, as Figure 1 illustrates, estimates of the output gap vary depending on the choice of the smoothing parameter, λ. Higher values of λ attach greater weight to smoothing rates of change in the trend and therefore produce smaller (larger) fluctuations in estimates of potential growth (output gaps). This happens because the HP filter retains cycles above a certain frequency and eliminates those below that frequency (European Commission 1999). Accordingly, at lower values of λ, the cycles in potential growth for Sweden are implausibly volatile, implying a change in potential growth from around 1.2 percent in 1992 to 2.5 by the end of the projection period in 2004.

B. Simple unobserved components methods

The Beveridge-Nelson (BN) decomposition resulted in estimates for potential output which tracked actual movements in output very closely (Figure 2). This result implies that nearly all of output movements can be interpreted as permanent, structural changes. The reason for this result may stem from the long swing in output over the 1990s. Since several years passed before the depressed growth rates of output reversed, the estimated autoregressive coefficients may not be able to capture the long lag structure in output cycles over this estimation period. In the comparison of different detrending methods, Park (1996) points out that trends from an HP filter are smoother (and cycles more volatile) than from the BN filter, which includes a unit-root component. Indeed, it has been shown that the HP filter creates artificial cycles if the actual data contains a near unit root. Moreover, a series of BN decompositions shown in Park produce trend components which are similarly close to the actual data.

9 An ARMA(2,2) was found to be the best model based on the Schwarz criterion.
Table 1. Sweden: Output Gaps and Potential Growth Rates in 1998
Calculated From Different Estimation Methods

<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>Output Gap</th>
<th>Potential Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hodrick-Prescott Filter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-0.6</td>
<td>2.1</td>
</tr>
<tr>
<td>200</td>
<td>-1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>1000</td>
<td>-1.7</td>
<td>1.9</td>
</tr>
<tr>
<td>Beveridge-Nelson</td>
<td>-1.7</td>
<td>3.0</td>
</tr>
<tr>
<td>Unobserved Components (Univariate)</td>
<td>0.2</td>
<td>2.8</td>
</tr>
<tr>
<td>Bivariate (with inflation)</td>
<td>0.2</td>
<td>2.8</td>
</tr>
<tr>
<td>Bivariate (with unemployment)</td>
<td>-1.1</td>
<td>1.7</td>
</tr>
<tr>
<td>Common Permanent and Transitory Components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One State (estimate 1)</td>
<td>-0.9</td>
<td>3.7</td>
</tr>
<tr>
<td>One State (estimate 2)</td>
<td>-5.5</td>
<td>0.9</td>
</tr>
<tr>
<td>Two States</td>
<td>-0.6</td>
<td>4.2</td>
</tr>
<tr>
<td>Structural VAR (with the real exchange rate)</td>
<td>-3.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Structural VAR (with unemployment)</td>
<td>-0.6</td>
<td>3.0</td>
</tr>
<tr>
<td>Production Function</td>
<td>-0.8</td>
<td>2.4</td>
</tr>
<tr>
<td>Systems Unobserved Components</td>
<td>-4.6</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Source: Staff estimates.
Figure 1. Sweden: Potential Output, Growth, and Output Gaps
HP Filter with Smoothing Parameters 100, 200, and 1000

Source: Staff estimates.
Figure 2. Sweden: Potential Output and Output Gaps
Bivariate Unobserved Components Models with Unemployment or Inflation

Source: Staff estimates.
A comparable result was obtained using the univariate unobserved component method and the bivariate unobserved component method with inflation as the second series. However, the bivariate unobserved component method with unemployment generated a much smoother series for potential GDP. Accordingly to the latter method, the output gap—at a sample low in 1993—was closed in 1998.

C. Common permanent and cyclical components

Most results from the unobserved components models with common permanent and transitory factors are similar to the univariate case. Deviations of potential output from actual are small. However, the result is sensitive to the initial guesses provided for the parameters (Figure 3). For the model using common permanent and temporary components with asymmetry governed by a single unobserved state variable, two local maxima were produced with close likelihood ratios. The resulting series from one set of parameter estimates was nearly a linear trend, while the other series tracked actual output. These results underscore the uncertainty in the estimates of potential output. The model with two unobserved states, each controlling a different component, led to an estimated output gap of -0.6 percent. The latter model relies on separate permanent and temporary state dependent switching in the average growth rate of output. This two-state model permits a characterization of historical recessions as leading to permanent versus temporary output losses. According to results from this model, recessions prior to 1993 could be attributed to a reduction in the temporary component of output growth, implying a reversion to trend (Figure 4). On the other hand, the results indicate that the output loss due to the recession in 1993 was mainly permanent. This lends support to the view that structural changes occurred in the early 1990s.\footnote{Thomas (1999) finds that Sweden's fiscal consolidation program of the 1990s, by reducing the ratio of government spending to GDP, has temporarily added an estimated 0.3 percent per annum to pre-existing estimates of potential output growth.}

There is also clear evidence that the business cycle is asymmetric—expansions are more persistent than recessions. According to the estimates, the temporary component of output expands for about 10\(\frac{1}{2}\) quarters on average but contracts for only 1\(\frac{1}{4}\) quarter. Likewise, the permanent component of output expands for about 23 years on average, but contracts for about 2\(\frac{3}{4}\) years.

D. Structural VAR

The Blanchard and Quah decomposition of output was obtained from two separate trivariate VAR models including changes in real GDP. First, the model proposed by Clarida and Gali that includes relative output, the real exchange rate and the relative price level was estimated except that Swedish output was substituted for the relative measure in order to prevent the estimation of Sweden's potential growth from being entangled with changes in the potential output of its trading partners. Second, a trivariate system was estimated that
Figure 3. Sweden: Potential Output and Output Gaps
Common Permanent and Cyclical Components

Source: Staff estimates.
Figure 4. Sweden: Probabilities of Permanent and Temporary Output Losses

Source: Staff estimates.
includes unemployment, similar to the Blanchard and Quah model except for the additional use of domestic prices. Impulse response functions from the first system of variables are similar to those reported in Thomas (1997). In accordance with the theory, supply shocks lead to positive long-run effects on output, while demand and nominal shocks have only small positive short-run effects. Correspondingly, most of the forecast variance in output stems from supply shocks. The first decomposition produces an output gap of -3.6 percent in 1998 and a growth of potential output of 2.7 percent from 1997 to 1998 (Figure 5). On the other hand, the Blanchard-Quah decomposition with output, unemployment, and inflation leads to an estimate of -0.6 percent for the 1998 output gap, and a higher potential growth rate of 3.0 percent (Figure 6). Results from this latter set of variables seem implausible, however, as the output gap is positive over much of the sample, including the early 1990s.

E. Production function

The production function approach was modeled consistent with the data and estimation of Thomas (1999). Annual private and public output, and the private capital stock series are the same as discussed in the paper. However the trend labor input was estimated using various methods as opposed to using the actual data on private hours worked which are likely to be cyclical. An OECD data series on the labor force is assumed to be at its trend level each year. Public employment was also taken from OECD. Labor’s share in output was assumed to be 65 percent. Total factor productivity was constructed as the residual of output less the capital and labor inputs. Trend productivity was constructed based on the fitted values of a regression which includes changes in the ratio of government expenditure to GDP and changes in the real capital stock of the public sector and controls for the cyclical component of productivity by including an adjusted unemployment rate and the interest rate spread.

Several methods were investigated in order to provide a measure of the nonaccelerating-inflation rate of unemployment (NAIRU). However, the divergence of resulting estimates attested to the uncertainty in measuring the NAIRU. First, an alternative measure of labor market slack, the NAWRU (nonaccelerating wage rate of unemployment) was estimated based on the approach proposed by Elmeskov (1993). This method assumes that changes in wage inflation are proportional to the gaps between actual unemployment and the NAWRU:

$$\Delta^2 \ln(W) = -\alpha (U - NAWRU)$$

With the additional assumption that the NAWRU does not change significantly from one year to another, the NAWRU can be approximated by:

---

11 Missing data for 1996-1998 were extrapolated based on changes in the ratio of public hours worked to total hours worked.
Figure 5. Sweden: Potential Output and Output Gap
Structural VAR with Output, the Real Exchange Rate, and Relative Prices

Source: Staff estimates.
Figure 6. Sweden: Potential Output and Output Gap
Structural VAR with Output, Unemployment, and Inflation

Source: Staff estimates.
\[ NAWRU = U - (\Delta U / \Delta^3 \ln(W)) \Delta^2 \ln(W) \]

and the resulting series are then smoothed using a HP filter to eliminate erratic movements. However, since the variance of wage growth has been small, the NAWRU tracks actual unemployment fairly closely. Hence, this method did not add to an immediate application of the HP filter. Moreover, the assumption that the NAWRU does not change significantly from one year to another is inconsistent with the result and seems to negate the value of the approach.

Second, as motivated by the discussion in Thomas (1999), time dummies for 1991, 1992, and 1993 were used to net out the sharp rise in the aggregate unemployment rate. However, this transformation removed all of the increase in unemployment, including a portion that could arguably be related to the cyclical position of the economy in those years. Finally, an HP filter was employed. In order to minimize the end-point problems with the HP filter previously discussed, the sample period was extended through 2005 using staff WEO projections and the structural level of unemployment was capped at 5 percent in accordance with estimates provided by the national authorities.

On the basis of this approach, the output gap in 1998 was estimated at -0.8 percent and potential output grew by 2.4 percent over the previous year (Figure 7). Figure 8 shows the contribution to the level of potential output of the labor input and total factor productivity (the capital stock is assumed to be equal to its trend level). Both inputs to production contributed to the sharp output gap around 1993. However, while productivity returned to and eventually eclipsed (slightly) its trend level, actual labor continues to be below its potential level. This reflects an unemployment rate which is higher than estimates of NAIRU.

F. Systems unobserved components

The system unobserved components approach updates estimations presented in Apel and Jansson (1997) using output, unemployment, and changes in inflation. An import price series is used as an exogenous variable in the equation for inflation. The estimated output gaps are broadly similar to the previous work, falling to around -7 percent in 1993 and recovering by a few percentage points in the mid-1990s (Figure 9). The results also indicate that the output gap averaged around -6 percent in 1997 and -4.6 percent in 1998, but narrowed to -3.5 percent by the last quarter 1998. Potential output grew by 1.5 percent in 1998.

This method also leads to estimates for the NAIRU. The bottom panel of Figure 9 illustrates that the NAIRU increased after 1993 to a level close to 5 percent, although it declined slightly since 1997. According to the parameter estimate on the unemployment gap (significant at the 1 percent confidence level), an actual unemployment rate that is 1 percentage point higher than the NAIRU is associated with a 1.8 percent shortfall of output from its potential level. In addition, the results indicate that import prices are important
Figure 7. Sweden: Potential Output and Output Gap
Production Function Approach

Actual and Potential GDP (business sector)

GDP
Potential Output

Output Gap

Source: Staff estimates.
Figure 8. Sweden: TFP and Labor Contributions to the Levels of GDP and Potential GDP
Production Function Approach

Source: Staff estimates.
Figure 9. Sweden: Potential Output, Output Gap, and NAIRU
System Unobserved Components Method

Source: Staff estimates.
determinants of Sweden's domestic inflation rate. The parameter estimate of 0.37 on import prices is significant at the 1 percent level and import prices appear to explain nearly half of the variation in domestic prices.

IV. CONCLUSIONS

This paper has presented new estimates of trend output and the output gap for Sweden according to different methods. The estimates are shown in Table 1. Based on the present calculations, the output gap was between -5.5 and 0.2 percent in 1998, while potential growth ranged from 0.9 percent to 4.2 percent from 1997 to 1998.

Each method has advantages and disadvantages as discussed in Section II. The main disadvantage of the statistical detrending techniques is that they are mechanistic. The latter methods rely on economic theory linking potential output to other developments in the economy. If there is confidence in the potential levels of the underlying inputs, the production function approach would likely be a preferred method since potential output would be determined on the basis of the constraints of the production factors. Attempts to produce more output than capacity would generate production bottlenecks and also lead to cost push inflation through wage pressures. However, in the case of Sweden, the estimate for the potential labor force is very uncertain due to difficulty in determining the NAIRU. Confidence in this approach is thereby weakened as the measure of the NAIRU would in turn rely on either ad hoc assumptions or atheoretical detrending methods. In this case, the system unobserved component method simultaneously generates potential output and the NAIRU with an explicit link to inflation performance. Moreover, the approach allows exogenous factors to influence inflation, although it can be argued that it is difficult to determine the complete set of contributing factors.

Although the various methods produce a range of results for the output gap, the overall evidence suggests that the large negative output gap—most pronounced in 1993—has either closed in 1998 or will close in the next 1-2 years if current trends continue. However, the evidence also suggests that at least part of the large jump in unemployment occurring in conjunction with this recent recession has become permanent. A future upswing in the business cycle may not be sufficient to restore unemployment to earlier levels; instead, structural policies to encourage a flexible and well-functioning labor market will likely be required.
REFERENCES


________, and C. Murray, 1999, “Permanent and Transitory Components of Recessions,” University of Washington manuscript.


OECD, Economic Surveys—Sweden, various issues.


