UNSW Business School

Working Paper

Never Stand Still | Business School | Economics

UNSW Business School Research Paper No. 2016 ECON 12

Is Business Cycle Asymmetry Intrinsic in Industrialized Economies?

James Morley
Irina B Panovska

Is Business Cycle Asymmetry Intrinsic in Industrialized Economies?

James Morley[1]
University of New South Wales
Sydney NSW 2052
Australia
Email: james.morley@unsw.edu.au

Irina B. Panovska
Lehigh University
Bethlehem, PA 18015
USA
Email: irp213@lehigh.edu

This Draft: September 21st 2016

Abstract

We consider a model-averaged forecast-based estimate of the output gap to measure economic slack for ten industrialized economies. Our measure takes changes in the long-run growth rate into account and, by accounting for model uncertainty using equal weights on different forecast-based estimates, is robust to different assumptions about the underlying structure of the economy. For each country, we find that the estimated output gap is highly asymmetric, with much larger negative movements during recessions than positive movements in expansions, suggesting that this particular form of business cycle asymmetry is an intrinsic characteristic of industrialized economies. Furthermore, the estimated output gap is strongly negatively correlated with future output growth and unemployment and positively correlated with capacity utilization in each case. It also implies a convex Phillips Curve in many cases.

JEL Codes: E32; E37
Keywords: output gap; model averaging; business cycle asymmetry; convex Phillips Curve

[1] Corresponding author. An earlier version of this study that focused on Asia-Pacific economies circulated under the title of "Measuring Economic Slack: A Forecast-Based Approach with Applications to Economies in Asia and the Pacific". We thank Stephane Dees, Jun Il Kim, Aaron Mehrotra, Tim Robinson, and James Yetman, as well as conference and seminar participants at the People's Bank of China-BIS Conference on "Globalisation and Inflation Dynamics in Asia and the Pacific", the "Continuing Education in Macroeconometrics workshop at the University of New South Wales, the Sydney Macroeconomics Readings Group, the European Central Bank, and the University of Technology Sydney for helpful questions and comments. The usual disclaimers apply.

1. Introduction

There is relatively little consensus in macroeconomics about how best to measure economic slack. Even settling on the output gap (i.e., the difference between actual and potential log real GDP for an economy) as the preferred measure, there remains the challenge of defining and calculating "potential". Widely-used decomposition methods that assume a linear structure for the economy, such as the Hodrick-Prescott (1997) filter, an unobserved components (UC) model with uncorrelated components (Clark, 1987), and a UC model with correlated components (Morley, Nelson, and Zivot, 2003), can lead to very different estimates of the output gap, as shown by, for example, Morley, Nelson, and Zivot (2003) or Perron and Wada (2015). Furthermore, there is a vast literature that documents a possible nonlinear structure for the economy (see, for example, Hamilton, 1989, Kim, 1994, Kim and Nelson, 1999, Kim, Morley and Piger, 2005, and Sinclair, 2010). However, as discussed in detail in Section 3, formal hypothesis tests provide only mixed evidence that nonlinear models of aggregate output are preferable to linear models. Given a lack of strong evidence for a single empirical specification of the economy that outperforms all other models, we argue in this paper for a model-averaged forecast-based estimate of the output gap as the appropriate measure of economic slack.

In terms of the forecast-based approach that we adopt in this paper, it is based on the idea that the presence or absence of economic slack directly implies whether an economy can or cannot grow faster than its long-run average growth rate without necessarily leading to subpar growth in the future. In particular, if the optimal forecast of future output growth is above/below average, then output will be estimated to be below/above potential. This approach implicitly defines "potential" as the stochastic trend of log real GDP and has its origins in the influential study by Beveridge and Nelson (1981, BN hereafter), with this particular interpretation of the BN decomposition discussed in Morley (2011).

Given a forecast-based approach to estimating the output gap, we need to confront the question of how to construct an optimal forecast of future output growth. BN consider low-order ARMA models, which result in small output gaps, often with counterintuitive sign (e.g., the estimated gap is often positive during recessions). Motivated by the different results and mixed evidence for different models discussed above, as well as the forecasting literature and recent studies on estimating the output gap by Garratt, Mitchell, and Vahey (2014) and Morley and Piger (2012), we consider model-averaged forecasts instead of relying on one particular time series model or class of models. Importantly, we follow Morley and Piger (2012) by including nonlinear time series models in the model set under consideration. Notably, this approach does not necessarily result in output gap estimates of counterintuitive sign as long as the model-averaged forecasts imply negative serial correlation in economic growth at longer horizons.

For our analysis, we investigate economic slack for a group of ten industrialized economies.[2] Importantly, our analysis takes into account structural breaks in long-run growth. The resulting measure of economic slack is a modified version of the model-averaged estimate of the output gap used by Morley and Piger (2012) for US real GDP. In particular, we consider the same broad set of both linear and nonlinear models from Morley and Piger (2012), but we place equal weights on all models considered and we incorporate some prior beliefs from previous analysis in Bayesian estimation of some model parameters. Given the diverse set of linear and nonlinear models, the simpler approach of using equal weights produces similar results to estimating optimal weights, while equal weights and Bayesian estimation are much easier to implement for a broad range of economies than the approach to model averaging and maximum likelihood estimation of the nonlinear models taken in Morley and Piger (2012).[3]

Our main finding is that model-averaged estimates of the output gap are highly asymmetric for all economies, regardless of the underlying statistical evidence supporting nonlinearity. This is notable because it suggests this form of business cycle asymmetry is not just a characteristic of the US economy, as previously established by Morley and Piger (2012), but is intrinsic in industrialized economies more generally. Furthermore, the estimated output gaps have strong negative forecasting relationships with future output growth in all cases and are closely related to narrower measures of slack given by the unemployment rate and capacity utilization. These results support the accuracy of the model-averaged estimates in comparison with model-specific estimates of the output gap. The results for a Phillips curve relationship with inflation are more mixed, but there is evidence in favor of a convex relationship for a number of economies,

---

[2] We are motivated to consider industrialized economies to determine whether there are any intrinsic characteristics of their output gaps, much like Levin and Piger (2006) investigated intrinsic characteristics of inflation rates for industrialized economies.

[3] GAUSS code for calculation of the model-averaged estimate of the output gap is available at https://sites.google.com/site/jamescmorley/research/code.

arguing against the imposition of a linear relationship when estimating output gaps, such as is done by Kuttner (1994) and in many other studies. Finally, we find very strong correlations between the measures of slack across economies, including dynamic linkages based on pairwise Granger Causality tests, and some evidence of an "English-speaking" business cycle or "Euro-only" business cycle.

The rest of this paper is organized as follows. Section 2 discusses the data, including the possible presence of structural breaks in long-run output growth for each economy. Section 3 motivates the model-averaging approach by demonstrating the sensitivity of the estimate of the output gap to the time series model under consideration. Section 4 presents the empirical models and methods used in the analysis. Section 5 reports the results first for the benchmark US case and then for a group of other industrialized economies. Section 6 concludes. The technical details are relegated to an appendix.

2. Data

We consider macroeconomic data for the United States (US) and nine other industrialized economies: Australia (AU), Canada (CA), France (FRA), Germany (DEU), Italy (IT), Japan (JP), Korea (KR), New Zealand (NZ), and the United Kingdom (UK). Our sample was selected with the intention of examining a representative set of industrialized economies. In particular, we include the large to medium-sized G7 economies, an additional medium-sized economy with many similar characteristics to the G7 economies (i.e., Australia), a somewhat smaller economy that also has many similar characteristics to the G7 economies (i.e., New Zealand), and an emergent medium-sized industrialized economy that has undergone several structural changes, but has reliable data (i.e., Korea). Data series for real GDP, the price level, the unemployment rate, and capacity utilization were sourced from OECD databases and from relevant national data sources. See Table A.1 in the appendix for full details.

For quarterly real GDP, we use the available seasonally-adjusted series and construct quarterly growth rates by taking first differences of 100 times the natural logs of the levels. The available sample periods for quarterly growth rates of real GDP are listed in Table 1.

For the price level, we use the core PCE deflator for the United States, core CPI for Canada, Germany, France, and the United Kingdom, and headline CPI for the remaining economies. These choices were determined by a general preference for core measures, but only when they are available for a relatively long sample period in comparison to real GDP. We calculate inflation as the year-on-year percentage change in the price level and then construct 4-quarter-ahead changes in inflation. The relevant sample periods based on common availability of both real GDP and price level data are listed in Table 3 in the next section.

The relevant sample periods based on common availability of the unemployment rate data with real GDP are listed in Table 6 in Section 4, and the relevant sample periods based on common availability of capacity utilization data with real GDP are listed in Table 7 in Section 4.

In addition to sample periods for the real GDP growth rate data, Table 1 reports estimated structural break dates for long-run growth rates—i.e., expected growth in the absence of shocks. Perron and Wada (2009) argue that it is crucial to account for a structural break in the long-run growth rate of US real GDP when measuring economic slack for the US economy using unobserved components models. They impose a break date of 1973Q1 based on the notion of a productivity growth slowdown at that time. Similarly, Perron and Wada (2015) show that that the popular Hodrick-Prescott (HP) filter is sensitive to the treatment of structural breaks. In particular, they show that that accounting for structural breaks can lead to very different inference about the output cycle in G7 economies. Thus, we follow Perron and Wada (2009, 2015) and allow for structural breaks in long-run growth rates. The full structural break test results are presented in Table A.2 in the Appendix.

Applying Bai and Perron's (1998, 2003) sequential testing procedure for structural breaks in the mean growth rate of US real GDP, we do not detect any break in the early 1970s. Instead, we find the estimated break date is 2000Q3. This break is significant at the 1% level and corresponds to a reduction in the mean growth rate. There is only weak evidence in favor of a second structural break in 1973Q1 (p-value is 0.13). However, following much of the literature, including Perron and Wada (2009, 2015), and allowing for the possibility of power and size distortions in finite samples, we also allow for a second structural break in 1973Q1.[4] We discuss the consequences of imposing different break dates when measuring economic slack for the US economy in Section 5 below. It also turns out also to be important to account for structural breaks in long-run expected growth for the other economies as well. With the exception of Australia and New Zealand, we find structural breaks in the expected long-run growth rates for all other economies. The estimated break dates and the corresponding sequence of mean growth regimes are reported in Table 1. We find evidence of one structural break for Canada, France, Italy, Korea, and the UK and evidence in favor of two structural breaks for Germany and Japan.[5] To account for structural breaks in subsequent analysis, the output growth series are mean-adjusted based on the estimated average growth rate in each regime until there is no remaining evidence of additional breaks.

3. Motivation

We motivate the model-averaging approach to measuring economic slack described in the next section by first considering forecast-based estimates of the output gap based on two commonly used models: an AR(1) model and Harvey and Jaeger's (1993) unobserved components (UC)

[4] Following much of the applied literature, we consider trimming of 15% of the sample from its end points and between breaks for admissible break dates. But even when using 5% trimming, we find no evidence of an additional structural break for the US in the mid-1970s at even the 10% level. As discussed in more detail in Section 5, not allowing for a second break in 1973 leads to estimates of output slack that are very strongly at odds with measures of slack from the previous literature and with more narrowly defined measures of slack, such as the unemployment rate. Given the broad evidence in favor of a break in 1973 from the previous literature, we impose a second break in 1973Q1. In general, we find that it is more problematic to underestimate than to overestimate the number of structural breaks when calculating forecast-based output gaps. Specifically, forecast-based output gaps can display permanent movements that proxy for large structural breaks in growth rates when these are not directly accounted for in the data, while accounting for smaller or possibly misspecified structural breaks tends to have little impact on forecast-based output gaps.

[5] The regression model for testing structural breaks includes only a constant. The evidence for structural breaks is generally weaker when allowing for serial correlation. In addition, the p-value for the test statistics for the second structural break in Germany in 1991Q2 was only significant at the 0.11 level. Similarly, the test statistics for the structural break in the UK in 1973Q1 was only significant at the 0.15 level. The OEDC series for German GDP is adjusted for the reunification level shift, but there is still evidence, albeit somewhat weak, in favor of a slope shift. However, previous studies for Germany that use a different set of empirical models (see, inter alia, Klinger and Weber, 2016, and Perron and Wada, 2015) find evidence of a break in the early 1990s following the reunification. In addition, when using year-on-year growth rates, we find stronger evidence in favor of a structural break in the UK and of second structural break in Germany. For the UK, when the 1973Q2 break is not taken into account, almost all measures of slack considered here imply that the UK output gap was below trend from 1973Q1 throughout 2016Q1. We therefore impose a structural break in the UK in 1973Q1 and a second structural break in 1991Q2 for Germany. All other breaks reported in Table 1 were significant at the 10% level. Allowing for additional structural breaks led to model-averaged estimates of the output gap that are very similar to those reported in the paper.

model that corresponds to the commonly used Hodrick-Prescott (HP) filter with a smoothing parameter of 1,600 (denoted UC-HP hereafter). The AR(1) model is estimated for quarterly real GDP growth and the output gap is estimated using the BN decomposition for an AR(1) model (see Morley, 2002, for details of this calculation). The UC-HP model is estimated for 100 times the natural logs of quarterly real GDP and the output gap is estimated using the Kalman filter. Although it is specified in terms of log levels, the UC-HP model provides an implicit forecast of future output growth, with the Kalman filter calculating the long-horizon conditional forecast of future output at each point of time.

Figure 1 plots the estimated output gaps based on the AR(1) and UC-HP models for US real GDP. As discussed in Morley and Piger (2012), these estimates are very different from each other, with the output gap based on the AR(1) model being of small amplitude and positive during NBER-dated recessions, while the output gap based on the UC-HP being of much larger amplitude and negative during NBER-dated recessions.

At first sight, it might seem obvious that the UC-HP output gap is preferable, especially given its more intuitive relationship with recessions. However, the AR(1) model fits the data much better than the UC-HP model by any standard metric used for model comparison, including AIC and SIC, a result that was highlighted in Morley and Piger (2012).[6]

Furthermore, as pointed out by Nelson (2008), the notion of an output gap as a measure economic slack directly implies that it should have a negative forecasting relationship with future output growth. Specifically, when the economy is above trend and the output gap is positive, future growth should be below average as the economy returns to trend and vice versa. Motivated by the analysis in Nelson (2008), we calculate the correlation between a given estimate of the output gap and the subsequent 4-quarter output growth.[7] Table 2 reports these

[6] We follow the approach in Morley and Piger (2012) to ensure the adjusted sample periods are equivalent for all models under consideration. For the linear and nonlinear AR models discussed below, this involves backcasting sufficient observations based on the long-run growth rate to condition on in estimation. For the UC models discussed below, it involves placing a highly diffuse prior on the initial level of the stochastic trend and evaluating the likelihood for the same observations as for the models of growth rates. In the case of the US, the AIC for the AR(1) model is -357.207 and the AIC for the UC model is -599.478, where the AIC is rescaled as in Morley and Piger (2012), with larger values being preferred. See the original study for details on the rescaling.

[7] Nelson (2008) considers regressions that capture the correlation between a given estimate of the output gap and 1-quarter-ahead US output growth. Our results for the US data are qualitatively similar to his even though we consider 4-quarter-ahead output growth, which arguably provides a better sense of forecasting ability at a policy-relevant

correlations and, consistent with the findings in Nelson (2008), the correlation for the US output gap based on the AR(1) model is negative, while the correlation for the UC-HP model is positive. This result directly suggests that the output gap based on the AR(1) model provides a more accurate measure of economic slack than the UC-HP model, even if its relationship with recessions seems counterintuitive.

Figure 2 plots the estimated output gaps based on the AR(1) and UC-HP models for real GDP data for the other nine industrialized economies. The figure makes it clear that the very different implications of the two models for the estimated output gap are not just a quirk of the US data. As in Figure 1, the output gap based on the AR(1) model is always smaller in amplitude than the output gap based on the UC-HP model and often of the opposite sign. The correlation results for these other economies in Table 2 are a bit more mixed, but the correlation with future output growth is still negative for more of the AR(1) model output gaps than the UC-HP model output gaps. Finally, any formal model comparison, including based on AIC or SIC, strongly favors the AR(1) model in every case.

More favorable to the UC-HP model is the forecasting relationship between the competing model-based output gaps and future inflation. Table 3 reports correlations between output gap estimates and subsequent 4-quarter changes in inflation. Consistent with most conceptions of the Phillips curve, the correlation is always positive for the UC-HP model output gap. By contrast, it is negative for 8 out of 10 economies when considering the AR (1) model output gap.

Taken together, the results in Tables 2 and 3 suggest that neither forecast-based estimate of the output gap provides a particularly accurate measure of economic slack. Put another way, even if we restrict ourselves only to two widely-used linear models, there is considerable uncertainty about the appropriate measure of economic slack. The AR(1) model fits the data better and its corresponding output gap generally provides a better forecast of future real GDP growth. But the

---

horizon. Also, Nelson (2008) conducts a pseudo out-of-sample forecasting analysis by estimating models and output gaps using data only up to when the forecast is made (it is a pseudo out-of-sample forecast because the data are revised, although Orphanides and van Norden, 2002, find that using revised or real-time data matters much less than incorporating future data in estimation of the output gap at any point in time). However, even though we use the whole sample to estimate models, we are implicitly using data only up to when the forecast is made to estimate output gaps. This is straightforward for the Harvey and Jaeger (1993) UC-HP model, which directly allows for filtered inferences, as opposed to the traditional HP filter, which is a two-sided filter, explaining why Nelson (2008) considers the out-of-sample forecasting analysis when evaluating the forecasting properties the output gap based on the traditional HP filter.

UC-HP model output gap is more consistent with widely-held beliefs about the relationship between economic slack and recessions, as well as generally providing a better forecast of future changes in inflation.

Given the fact that both the AR(1) and the UC-HP models are linear, a natural question that arises is whether accounting for any potential nonlinearities would provide a better measure of the business cycle and output slack. While nonlinear are more highly parametrized, there is some evidence that nonlinear models fit US output growth better than the corresponding linear AR(p) models (see, for example, Hamilton, 1989, or Kim, Morley, and Piger, 2005). Table 4 presents the results of the Carrasco, Hu, and Ploberger (2014) test for the Hamilton and bounceback Markov-switching models with normal and t-distributed errors versus a linear AR(2) model and a Monte-Carlo based likelihood ratio (LR) test for a depth-based bounceback model versus an AR(2) model (discussed in more detail in the next subsection). Again, the results are inconclusive in many cases, with the test statistics being right around the threshold critical values in many cases and the results sensitive to the assumptions about the distribution of the disturbances. For example, we can reject the null of linearity against a Markov-switching model for Australia, but the critical values are right around the 10% threshold. Similarly, we can reject the null of linearity against a depth-based bounceback model for the US, but the critical values are right around the 5% critical value.

These mixed results motivate the methods outlined in the next section. In particular, drawing from an insight going back at least to Bates and Granger (1969) that combined forecasts can outperform even the best individual forecast, we follow Morley and Piger (2012) and construct a model-averaged estimate of the output gap, averaging over a range of linear and nonlinear forecasting models.

4. Methods

The analysis used here closely follows the approach to estimating a model-averaged output gap (MAOG) developed in Morley and Piger (2012) for US real GDP. However, we consider a few modifications that make the approach more easily applicable to data from other economies. The approach, including the modifications, is outlined in this section. The full details of the approach are in the original study and are also set out in the appendix.

Morley and Piger (2012) focus on univariate models of real GDP, which includes the AR(1) and UC-HP models discussed in the previous section. As is evident from Figures 1 and 2, the univariate models capture a range of possibilities about the nature of the output gap. Also, univariate analysis allows us to test multivariate relationships rather than assume the answer *a priori*. The benefits of this approach for the relationship with inflation in particular will become evident when the results are presented below.

All of the models allow for a stochastic trend in real GDP, which is motivated by standard unit root and stationarity tests, even when allowing for structural breaks in long-run growth. The results for all countries for the standard unit root tests (Augmented Dickey-Fuller and Elliott-Rothenberg-Stock point-optimal Dickey Fuller), the standard stationarity tests (Leybourne and McCabe, 1992, and the KPSS test proposed by Kwiatkowski et al., 1992), and the unobserved-components based stationarity test based on Morley, Panovska, and Sinclair (2016) are presented in Table 5.[8] This is important because many off-the-shelf methods such as linear detrending, traditional HP filtering, and Bandpass filtering produce large spurious cycles when applied to time series with stochastic trends (see Nelson and Kang, 1981, Cogley and Nason, 1995, and Murray, 2003). By contrast, as long as the models under consideration avoid overfitting the data, the forecast-based approach will not produce large spurious cycles.

Following Morley and Piger (2012), we consider linear AR(p) models of orders p = 1, 2, 4, 8, and 12 with Gaussian errors or Student t errors, the linear UC-HP model due to Harvey and Jaeger (1993), the linear UC0 and UCUR models with AR(2) cycles from Morley, Nelson, and Zivot (2003), the nonlinear bounceback (BB) models from Kim, Morley, and Piger (2005) with BBU, BBV, and BBD specifications and AR(0) or AR(2) dynamics, the nonlinear UC0-FP model with an AR(2) cycle from Kim and Nelson (1999), and the nonlinear UCUR-FP model with an AR(2) cycle from Sinclair (2010).[9] Again, see the appendix and the original studies for more details of these models.

---

[8] Based on the Monte Carlo analysis in Morley, Panovska, and Sinclair (2016), we consider the bootstrapped p-values for all stationarity tests to correct for potential size distortions in finite samples.

[9] As a minor modification to Morley and Piger (2012), we drop the linear AR(0) models and nonlinear Markov-switching model from Hamilton (1989) with AR(0) and AR(2) dynamics. In the former case, the output gap is always zero by construction, so its inclusion merely serves to shrink the model-averaged output gaps towards zero. In the latter case, the output gap is linear by construction, so its inclusion as a nonlinear model puts additional prior

The first major modification from Morley and Piger (2012) is that models are estimated using Bayesian methods instead of maximum likelihood estimation (MLE). This allows incorporation of informative priors in the estimation. The priors we used here are not particularly strong, with estimates based on the posterior mode virtually identical to MLE for many of the models.[10] However, for economies with relatively short samples for real GDP or other quirks in the data such as large outliers, there is some tendency for MLE of the UC models and the nonlinear models to overfit the data. By incorporating more informative priors about the persistence of the autoregressive dynamics or the persistence of Markov-switching regimes based on US estimates from Morley and Piger (2012), we are able to avoid problems associated with shorter samples and outliers, while obviating the need to undertake a long, protracted search for the best model specifications for each economy.[11] The full details of the priors are presented in the appendix.

The second major modification from Morley and Piger (2012) is that we consider equal-weights on the models when constructing MAOGs rather than weights based on Bayesian model averaging (BMA). Although a number of models receive nontrivial weight based on the SIC approximation of BMA when considering the US data in Morley and Piger (2012), this is not always the case for other economies. For example, a simple AR(0) model would receive all weight for Australian real GDP based on SIC if it were included in the model set. However, such

---

weight on a linear output gap. As demonstrated below, dropping these models has very little practical impact on the model-averaged estimate of the output gap for US real GDP. If the Hamilton model is included in the set of models, the correlation between the MAOG computed using equal weights that includes the Hamilton Model and the MAOG that does not include the Hamilton model is 0.99995. Furthermore, as shown in Table 5, the Carrasco et al. (2014) bootstrap test for Markov-Switching parameters cannot reject the null of no switching for all economies except New Zealand, Italy, and Australia, with p-values higher than 10% in all cases except for Italy. However, the null of linearity can be strongly rejected in favor of the BBD model for those three economies. The null of linearity can also be rejected in favor of the BBU model for Germany, Japan, Korea, New Zealand, and the UK, and in favor of the BBD model for all economies except Italy and New Zealand. Therefore, our set of models does not lose empirical relevance by excluding the Hamilton models.

[10] The AR(1) and UC-HP models discussed in previous section were estimated using the posterior mode. But the estimated output gaps for these models are indistinguishable from those based on MLE. For example, for the US data, the correlation between the Bayesian and MLE output gaps is >0.999999.

[11] In principle, this setup would also make it possible to apply the approach outlined in this paper even given severe data limitations or a desire to impose tighter priors based on strongly held beliefs. For example, in an earlier version of this study, Morley (2014) estimated the output gap for a set of 13 economies in the Asia and Pacific, many with very short sample periods and extreme outliers. In terms of imposing tighter priors on characteristics such as the smoothness of trend, see the approaches outlined in Harvey, Trimbur, and van Dijk (2007) for UC models and Kamber, Morley, and Wong (2016) for AR models. However, given the strong evidence for a volatile stochastic trend in Morley, Panovska, and Sinclair (2016) and in Table 5 of this paper, we avoid imposing smoothness priors as it could potentially lead to spurious cycles.

a model implies the output gap is always exactly zero by construction (not just zero on average), which clearly runs contrary to widely and strongly held beliefs and, as will be seen below, would produce inferior forecasts of future output growth and changes in inflation in comparison to the Australian MAOG.

The problem of BMA putting too much weight (from a forecasting perspective) on one model has been highlighted by Geweke and Amisano (2011). They find that linear pooling of models produces better density forecasts than BMA and discuss the calculation of optimal weights for linear pooling of models. However, as long as the model set is relatively diverse, applying equal weights to models works almost as well as optimal weights and is much easier to implement in practice. Thus, we take this simple approach of using equal weights for the reasonably diverse set of linear and nonlinear models discussed above.[12] Again, see the appendix for more details of the model averaging.

5. Results

We first consider the United States as a benchmark case in order to provide perspective on the impact of the modifications to Morley and Piger (2012) described in the previous section, as well as providing context for the other results.

To begin, we compare the updated MAOG based on the US real GDP data described in Section 2, equal weights, and Bayesian estimation to the original MAOG reported in Morley and Piger (2012) based on a shorter sample period, a different vintage of data, BMA weights, and MLE. We also consider the updated MAOG based on the Morley and Piger (2012) BMA weights and MLE for the full sample. Figure 3 plots these three MOAGs together. The most noticeable thing is their similarity, with the major finding in Morley and Piger (2012) of a highly asymmetric shape holding for the updated MAOGs. The correlation between the updated MAOG based on BMA weights and MLE and the updated MAOG based on equal weights and Bayesian estimation is 0.96.

[12] To be specific, we place equal weights on linear and nonlinear classes of models and divide those equal weights up evenly amongst the models within the classes. Because the nonlinear models nest linear dynamics in their parameter space, there is still more implicit prior weight on linear than nonlinear dynamics, although this is addressed somewhat by the somewhat informative priors for parameters in the nonlinear models.

The impact of incorporating prior information about parameters may be obscured in Figure 3 given that the priors were calibrated based on the estimates for US data in Morley and Piger (2012). However, it is important to emphasize that the asymmetric shape of the output gap is in no way driven by the priors on the nonlinear models. The priors for the Markov-switching parameters favor regime shifts in the mean growth rate corresponding to business cycle phases, along the lines of Hamilton (1989), but there is no prior that shocks have more temporary effects in recessions than in expansions. Figure 4 makes this clear by applying the modified approach to data simulated from a simple random walk with drift.[13] For this data, the true output gap is always zero. The estimated MAOG is not always zero, but, unlike what would be the case for the HP filter given a random walk, the spurious cycle is quite small in magnitude relative to the US MAOG and, importantly, it fluctuates symmetrically around zero. Thus, any finding of asymmetry for the MAOGs reflects the data, not the incorporation of prior information in estimating model parameters.[14]

As displayed in Figures 3 and 4, our results indicate that there is little remaining economic slack for the US economy at the end of the sample in 2016Q1. This result is consistent with the Federal Reserve's views (see, for example, Yellen, 2015). These results, however, turn out to be sensitive to allowing for a structural break in long-run growth in 2000Q3. Figure 5 plots the updated US MAOG (allowing for a structural break in 1973Q1 and 2000Q3) against a version under the assumption of no structural break and a version that imposes only one break in 1973Q1 based on previous literature. Assuming no change in the long-run growth, the US economy appears to still be below trend at the end of the sample. Given uncertainty about the structural break, it might make sense to average across these two scenarios, which would still imply the economy remains slightly below trend at the end of the sample, although not by as much as in the no break case. If we assume that the US economy was at trend at the end of the sample, this would clearly imply that recessions can permanently shift the trend path of output downwards, which is the implication of many forecasting models for US real GDP, including low-order

[13] The drift and standard deviation of shocks are both set to 1, which is a surprisingly reasonable calibration for 100 times the natural logs of quarterly US real GDP.

[14] Indeed, model averaging would tend to understate asymmetry if it were present in the data generating process (e.g., suppose Kim and Nelson's, 1999, UC0-FP model was the true model) by shrinking the mean of the estimated MAOG towards zero given that the mean of the estimated output gap is zero by construction for all of the linear models under consideration.

AR(p) models, Hamilton's (1989) Markov-switching model, and, to some extent, the bounceback models of Kim, Morley, and Piger (2005). In a recent paper, Huang, Luo, and Startz (2016) find that recessions prior to 1984 can be described as U-shaped, but recessions after 1984 can be better described using Hamilton's (1989) L-shaped model. Figure 6 plots the estimated trend in US real GDP based on the model-averaged output gap. A permanent negative effect of the Great Recession of the trend path is quite evident for this estimate of trend and is much larger than for previous recessions.[15]

One way to judge the plausibility of the US economy being at trend at the end of the sample is to compare the US MAOG to other narrower measures of slack. Figure 7 plots the US MAOG against the US unemployment rate and US capacity utilization. Similar to the findings in Morley and Piger (2012), there is a clear relationship between the MAOG and these variables. More supportive of relatively little remaining slack at the end of the sample is the simple fact that the MAOG in the no break case would imply relatively fast growth and downward pressure on inflation in the period immediately after the Great Recession. In particular, returning to Tables 2 and 3, the US MAOG has a negative correlation of -0.38 with future output growth and positive correlation of 0.46 with future changes in inflation. These results are much stronger than those for the output gaps based on the AR(1) and UC-HP models and support the MAOG as a highly relevant measure of economic slack. But, given lacklustre growth and stable inflation after the Great Recession, these results also support the MAOG allowing for a structural break and the idea that the US economy is actually close to trend at the end of the sample, noting that the trend path is lower than before the recession, as suggested in Figure 6.

Having demonstrated how the modified approach works in the benchmark US case, at least when liberally allowing for structural breaks in long-run growth, we now apply the approach to the remaining G7 economies, Australia, New Zealand, and Korea.

Figure 8 plots MAOGs for the various economies. For all cases considered, the MAOGs are highly asymmetric, similar to the US results. Specifically, the output gaps take on much larger

[15] Allowing for one structural break in 1973Q1 leads to similar results. Similarly, allowing for a structural break in 2000Q3 but not in 1973Q1 leads to an estimated MAOG that is large and negative during the 1990-1991 recession and very deep during 2001 recession, which is at odds with previous estimates of output slack, and with more narrow measures of slack, such as unemployment and capacity utilization, where both the 1990 and 2001 recession were relatively shallow. This further motivates our inclusion of a second structural break in 1973Q1.

negative values than positive ones.[16] The only possible exception is Italy, where the output fluctuations are relatively more symmetric, but there is still strong evidence that a couple of the contractions in 1969 and 2008-2009 caused highly asymmetric movements. The ubiquity of this form of business cycle asymmetry across the ten economies under consideration suggest that it is an intrinsic characteristic in industrialized economies, not just a feature of the US economy in particular. This is a potentially important result for theory-based modeling of the business cycle, which tends to focus on linear dynamics for convenience, although there are many exceptions.[17]

How plausible are the MAOGs as measures of economic slack? As with the US benchmark, we compare the MAOGs to other narrower measures of slack. Table 6 reports the correlation of each MAOG with the corresponding unemployment rate. For comparison, we also report correlations for output gaps based on AR(1) and UC-HP models. Corresponding to an Okun's Law relationship, the MAOG has the most negative correlation with the unemployment rate in all 10 cases (including the US benchmark), with many of the correlations being quite large in magnitude.

Table 7 reports the corresponding correlations with capacity utilization. The MAOG has the most positive correlation with capacity utilization in 7 out of 10 cases (including the US benchmark) and has very high correlations in two of the other cases (Germany and the United Kingdom).

Overall, the strong coherence with other measures of slack lends credence to the MAOGs. The coherence is particularly notable given that the MAOGs are estimated using only univariate

[16] Perron and Wada (2015) emphasize the importance of large outlier shocks in their paper, and use the 1968 strike as one example that causes distortions and large discrepancies between the HP estimate and their UC estimate of the output gap. The MAOG identifies the 1968Q2 strike in France as a temporary large deviation, with output returning to trend the subsequent quarter. Treating the 1968Q2 strike in France as a one-time event using a dummy variable before estimating the MAOG does not change the results, with the only difference being that the estimated MAOG in 1968Q2 (during the strike). By including models with t-errors in our set of models, we indirectly allow for the possibility of occasional large shocks such as strikes. Furthermore, it is not implausible to assume that output would be below trend during a strike.

[17] For example, Diebold, Schorfheide, and Shin (2016) find that incorporating nonlinearities in the exogenous driving processes and allowing for stochastic volatility in a DSGE model markedly improves the density forecast performance of the model. Auroba, Bocola and Schorfheide (2013) highlight the fact that asymmetric wage and price adjustments lead to inherent nonlinearity in DSGE models, and argue in favor of using a nonlinear time-series model to evaluate the performance and predictive ability of DSGE models. Guerrieri and Iacoviello (2016) find that collateral constraints in a DSGE model lead to macroeconomic asymmetries—in particular, when constraints are slack, expanding wealth makes small contribution to consumption growth, but tightened constraints can sharply exacerbate recessions.

models of real GDP. At the same time, the MAOGs provide a broad and useful measure of slack, even when unemployment rate or capacity utilization data are distorted as measures of slack by long-run structural factors.

Revisiting Table 2, the MAOGs provide a stronger signal about future economic growth than the two other output gap estimates for all economies in our sample. This result provides the most direct support of the MAOGs as measures of economic slack based on the definition considered in this paper. It also confirms the possibility that output growth can be somewhat predictable even when standard model comparison metrics would select a random walk model, as the SIC would in the case of Australia.

Looking back at Table 3, the results for the MAOGs in terms of correlation with future changes in inflation are more mixed. The MAOGs provide a stronger signal than the UC-HP model output gap in only 4 of the 10 cases (including the US benchmark). However, a correlation coefficient may be too simplistic as a measure of the relationship between the output gap and inflation. Figure 9 displays a scatterplot of the US MAOG (x-axis) against the subsequent 4-quarter change in US inflation (y-axis) and there is a clear nonlinear, convex Phillips Curve relationship between the output gap and future changes in inflation that would only be partially captured by a correlation coefficient.

Figure 10 displays the corresponding scatterplots for the nine other industrialized economies. The same convex relationship as for the US data is evident for Australia, France, Japan, and Korea. For some of the other cases, such as Canada and New Zealand, the Phillips Curve relationships look more linear. However, a clear implication of Figures 9 and 10 is that it is important not to impose a linear (or any other) specification for the Phillips Curve relationship *a priori*, as is done in some other approaches to estimating output gaps (e.g., Kuttner, 1994). In particular, if the imposed relationship is incorrectly specified, then the resulting output gap estimate will necessarily be distorted and cannot be used to determine a better specification of a Phillips Curve relationship. The convexity of the Phillips Curve in some cases argues against a linear specification. Also, there is some evidence that the relationship between the output gap and inflation has evolved over time, with many of the observations of stable inflation following large negative output gaps corresponding to the recent Global Financial Crisis. Consistent with the Lucas's (1976) famous critique that reduced-form Phillips Curve relationships should change with policy regimes, this apparent breakdown in the previous pattern near the end of the sample could be due to an anchoring of inflation expectations (see IMF, 2013) and argues strongly against imposing a fixed relationship with inflation when estimating the output gap.

Given general support for the MAOGs as measures of economic slack, especially in terms of the crucial definitional sense of forecasting future economic growth, the last question considered in this paper is whether the MAOGs are related across economies.[18] For example, Mitra and Sinclair (2012) find that cycles for G-7 economies obtained using a correlated UC model are highly correlated across economies. Using a time-varying framework, Del Negro and Otrok (2008) find that Japan has been decoupling from other G-7 countries since the 1970s. Similarly, several other studies, including Canova et al. (2007) and Stock and Watson (2005) find evidence of co-movement across cycles, such as "English-speaking business cycle" or "Euro business cycle"—that is, high correlation across groups of economies that has increased over time (in particular, after the mid-1980s and early 1990s). While there are a lot of studies that attempt to explain movements in business cycles in a multi-country setting and quantify the importance of world and regional or group shocks, our goal here is to simply evaluate whether cross-country links between MAOGs are in line with the previous literature. There is strong within-country support in favor of the MAOG as a measure of slack from Tables 1, 2, 3, 6, and 7, and we are interested in evaluating if the MAOGs are also supported by cross-country comparisons. To answer this question, we consider the correlations between the MAOGs for all 10 economies, and conduct pairwise Granger Causality tests.

Table 8 reports correlations for output gaps across countries. The correlations are positive in most cases, indicating that the MAOGs tend to move together. The correlations also become stronger post 1984 for the US, Australia, Canada, and the UK, and France, Italy, and Germany, thus lending informal support to the "English-speaking business cycle" and "Euro-cycle". However, there is no evidence of decoupling between Japan and the rest of the world.

Table 9 presents the results for the Granger Causality tests. At the 10% level, the output gaps appear to be related across many of the economies, with 24 rejections of no Granger Causality.

[18] In principle, we could also consider whether the estimated trend shocks are correlated, as in Blonigen et al. (2014), but our focus is on output gaps and leave analysis of trends for future research.

The patterns are generally sensible, although the 10% may include some rejections of the null merely due to random sampling given that a total of 90 tests were conducted. At the same time, the fact that the number of rejections holds up at 19 for the tests at the 5% level and 10 for tests at the 1% level suggests that many of rejections are simply because the null hypothesis is false.

In terms of general patterns for cross-economy spillovers and focusing on the results at the 5% level, the output gaps for Canada and the US appear to Granger-cause the largest number of other economies, while the output gaps for France and the United Kingdom are Granger-caused by the largest number of other economies. Furthermore, there is evidence of an "English-speaking" cycle both before and after 1984, and some evidence of an "Euro-cycle" after 1984.

6. Conclusions

There is much more uncertainty about the degree of economic slack than is commonly acknowledged in academic and policy discussions, which often treat the output gap as if is directly observed. Canova (1998) argues that this uncertainty has huge implications in terms of "stylized facts" about the business cycle used to motivate theoretical analysis. Also, Orphanides (2002) argues that this uncertainty is responsible for huge policy mistakes in the past, especially in terms of the high inflation in the 1970s.

In light of this uncertainty about the degree of economic slack prevailing in an economy at any given point of time and its importance for policy, we argue for a model-averaged forecast-based estimate of the output gap. For all industrialized economies considered here, the model-averaged estimate is closely related to the narrower measures of slack given by the unemployment rate and, consistent with the notion of an output gap as a measure economic slack, has a strong negative forecasting relationship with future output growth. Most importantly, the model-averaged output gap estimates are all highly asymmetric, as was found for US real GDP in Morley and Piger (2012). This directly suggests that this particular form of business cycle asymmetry is intrinsic in industrialized economies and should be addressed in theoretical models of the economy.[19]

[19] As emphasized in Kiley (2013) and noted by many others, theory-oriented DSGE models imply reduced-form VAR, VECM, or VARMA models. Thus, forecast-based output gap estimates provide robust measures of economic slack across a wide range of different economic assumptions used to identify a structural model, at least as long as

Evidence for a Phillips Curve relationship between the model-averaged output gap and inflation is more mixed. But the overall results strongly argue against imposing a linear relationship in estimating output gaps. As an example of why imposing a fixed relationship is so problematic, consider Stock and Watson (2009, 2010). Their analysis suggests that inflation is difficult to forecast using standard measures of economic slack, except when the estimated output gap (or unemployment gap) is large in magnitude. This directly suggests possible mismeasurement due to imposition of symmetry and/or a nonlinear Phillips Curve relationship (see Dupasquier and Ricketts, 1998, and Meier, 2010). Our measure of economic slack allows for a full investigation of the nature of the relationship between the output gap and inflation, including the possibility of a nonlinear relationship.

Finally, there are notable dynamic linkages between the model-averaged output gaps across many economies and some evidence of an "English-speaking" business cycle or "Euro-only" business cycle. In general, as with inflation, our estimates allow us to explore possible cross-economy relationships without imposing them *a priori*.

---

the reduced-form model or models used to calculate the optimal forecast capture the dynamics in the data (this point relates back to Sims, 1980—also see Fernandez-Villaverde et al., 2007).

Appendix 1: Empirical Models

Following Morley and Piger (2012), we define the output gap, c_t, as the deviation of log real GDP, y_t, from its stochastic trend, τ_t, as implied by the following trend/cycle process:

$$y_t = \tau_t + c_t, \tag{1}$$

$$\tau_t = \tau_{t-1} + \eta_t^*, \tag{2}$$

$$c_t = \sum_{j=0}^{\infty} \psi_j \omega_{t-j}^*, \tag{3}$$

where $\psi_0 = 1$, $\eta_t^* = \mu + \eta_t$, and $\omega_t^* = \overline{\omega} + \omega_t$, with η_t and ω_t following martingale difference sequences. The trend, τ_t, is the permanent component of y_t in the sense that the effects of the realized trend innovations, η_t^*, on the level of the time series are not expected to be reversed. By contrast, the cycle, c_t, which captures the output gap, is the transitory component of y_t in the sense that the Wold coefficients, ψ_j, are assumed to be absolutely summable such that the realized cycle innovations, ω_t^*, have finite memory. The parameter μ allows for non-zero drift in the trend, while the parameter $\overline{\omega}$ allows for a non-zero mean in the cycle, although the mean of the cycle is not identified from the behaviour of the time series alone, as different values for $\overline{\omega}$ all imply the same reduced-form dynamics for Δy_t, with the standard identification assumption being that $\overline{\omega} = 0$.

The optimal estimate (in a minimum mean-squared-error sense) of trend for a range of trend/cycle processes as in (1)-(3), including those with regime-switching parameters, can be calculated using the regime-dependent steady-state (RDSS) approach developed in Morley and Piger (2008). The RDSS approach involves constructing long-horizon forecasts using a given time series model to capture the dynamics of the process. Importantly, the long-horizon forecasts are conditional on sequences of regimes and then marginalized over the distribution of the unknown regimes. Specifically, the RDSS measure of trend is

$$\hat{\tau}_t^{RDSS} \equiv \sum_{\tilde{S}_t} \left\{ \hat{\tau}_t^{RDSS}\left(\tilde{S}_t\right) \cdot p^M\left(\tilde{S}_t \middle| \Omega_t\right) \right\}, \tag{4}$$

$$\hat{\tau}_t^{RDSS}\left(\widetilde{S}_t\right) \equiv \lim_{j\to\infty} \left\{ E^M\left[y_{t+j} \middle| \left\{S_{t+k} = i^*\right\}_{k=1}^{j}, \widetilde{S}_t, \Omega_t \right] - j \cdot E^M\left[\Delta y_t \middle| \left\{S_t = i^*\right\}_{-\infty}^{\infty} \right] \right\}, \tag{5}$$

where $\widetilde{S}_t = \{S_t, \ldots, S_{t-m}\}'$ is a vector of relevant current and past regimes for forecasting a time series, $p^M(\cdot)$ is the probability distribution with respect to the forecasting model, S_t is an unobserved state variable that takes on *N* discrete values according to a fixed transition matrix, and i^* is the "normal" regime in which the mean of the transitory component is assumed to be zero. The choice of "normal" regime i^* is necessary for identification. Meanwhile, for a given forecasting model with Markov-switching parameters, the probability weights in (4), $p^M\left(\widetilde{S}_t \middle| \Omega_t\right)$, can be obtained from the filter given in Hamilton (1989). Note that the RDSS trend simplifies to the Beveridge and Nelson (1981) trend in the absence of regime switching.

In practice, the correct model for the dynamics of the time series process is unknown. Thus, following Morley and Piger (2012), we consider a range of models, as listed in the main text. The linear and nonlinear AR(p) models are specified as follows:

$$\phi(L)\left(\Delta y_t - \mu_t\right) = e_t \tag{6}$$

$$\mu_t = \mu(S_t, \ldots, S_{t-m}), \tag{7}$$

where $\phi(L)$ is p^{th} order. We consider versions of the AR(p) models with Gaussian errors (i.e., $e_t \sim N(0, \sigma_e^2)$) or Student *t* errors (i.e., $e_t \sim t(\nu, 0, \sigma_e^2)$). $S_t = \{0,1\}$ is a Markov state variable with fixed continuation probabilities $\Pr[S_t = 0 \mid S_{t-1} = 0] = p_{00}$ and $\Pr[S_t = 1 \mid S_{t-1} = 1] = p_{11}$. In the linear case, $\mu_t = \mu$, while there are three different specifications of μ_t in the nonlinear case that correspond to the BB models developed by Kim, Morley, and Piger (2005):

1. "U"-Shaped Recessions (BBU)

$$\mu_t = \gamma_0 + \gamma_1 S_t + \lambda \sum_{j=1}^{m} \gamma_1 S_{t-j}, \tag{8}$$

2. "V"-Shaped Recessions (BBV)

$$\mu_t = \gamma_0 + \gamma_1 S_t + (1 - S_t)\lambda \sum_{j=1}^{m} \gamma_1 S_{t-j}\,, \tag{9}$$

3. Recovery based on "Depth" (BBD)

$$\mu_t = \gamma_0 + \gamma_1 S_t + \lambda \sum_{j=1}^{m} (\gamma_1 + \Delta y_{t-j}) S_{t-j}\,, \tag{10}$$

where the state $S_t = 1$ is labeled as the low-growth regime by assuming $\gamma_1 < 0$. Following Kim, Morley, and Piger (2005), we assume $m = 6$. See the original study for the full motivation of these specifications.

The linear and nonlinear UC models are based on (1)-(3), with the following parametric specification of the transitory component in (3):

$$\phi(L)c_t = \omega^*_{t-j}\,, \tag{11}$$

where $\overline{\omega} = 0$ for the linear UC0 and UCUR models and $\overline{\omega} = \tau S_t$ for the nonlinear UC0-FP and UCUR-FP models, with the state $S_t = 1$ labeled by assuming $\tau < 0$. The shocks to the trend and cycle are Gaussian (i.e., $\eta_t \sim N(0, \sigma^2_\eta)$, $\omega_t \sim N(0, \sigma^2_\omega)$ for the UC0 and UC0-FP models and $(\eta_t, \omega_t)' \sim N(0, \Sigma_{\eta\omega})$ for the UCUR and UCUR-FP models). Given an AR(2) cycle, the covariance for the UCUR and UCUR-FP models is identified (see Morley, Nelson, and Zivot, 2003).

As mentioned in the main text, Bayesian estimates for these models are based on the posterior mode. The priors for the various model parameters are set out in Table A.3. Note that the prior for bounceback coefficient has zero mean, implying a prior mean of zero for the output gap. The prior for the mean of the transitory shock for the UC-FP models has a negative mean, but this has very little impact on the prior mean of the model-averaged output gap given the small weights on any given model. The prior on the AR coefficients clearly places them in the stationary region. Finally, the prior for the continuation probabilities is centered at 0.95 for the expansion regime and 0.75 for the other regime. This is calibrated based on the results for US data in Morley and Piger (2012).

In practice, given parameter estimates, we use the BN decomposition or, in the case of the UC models, the Kalman filter to estimate the output gap for the linear models. Note that the filtered inferences from the Kalman filter are equivalent to the BN decomposition using the corresponding reduced-form of the UC model, while the BN decomposition is equivalent to the RDSS approach in (4)-(5) in the absence of regime-switching parameters. To estimate the output gap for the nonlinear forecasting models, we use the RDSS approach or, in the case of the nonlinear UC models, the Kim (1994) filter, which combines the Kalman filter with Hamilton's (1989) filter for Markov-switching models. For the nonlinear models, we follow Kim and Nelson (1999) and Sinclair (2010) by assuming the "normal" regime $i^* = 0$, which corresponds to an assumption that the cycle is mean zero in expansions.

Finally, the MAOG is calculated as follows:

$$\tilde{c}_t = \sum_{i=1}^{N} c_{i,t} \Pr\left(M_i\right), \tag{12}$$

where i indexes the N models under consideration, $c_{i,t}$ is the estimated output gap for model i, M_i is an indicator for model i, and $\Pr\left(M_i\right)$ denotes the weight placed on model i. As discussed in footnote 11, we place equal weights on linear and nonlinear classes of models and divide those equal weights up evenly amongst the models within the classes. Given 13 linear models (five linear AR models with two types of errors and three linear UC models) and 14 nonlinear models (two nonlinear AR models with three BB specifications and two types of errors and two nonlinear UC models), the weight on each linear model is 3.9% and the weight on each nonlinear model is 3.6%.

References

Auroba, S. Boragan, Luigi Bocola, and Frank Schorfheide, 2013, "Assessing DSGE Model Nonlinearities." Federal Reserve Bank of Philadelphia Working Paper No. 13-47.

Bates, John M., and Clive W.J. Granger, 1969, "The Combination of Forecasts," *Operations Research Quarterly* 20, 451-468.

Bai, Jushan, and Pierre Perron, 1998, "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica* 66, 47-78.

Bai, Jushan, and Pierre Perron, 2003, "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics* 18, 1-22.

Bloningen, Bruce, A., Jeremy M. Piger, and Nicholas Sly, 2014, "Comovement in GDP Trends and Cycles Among Trading Partners," *Journal of International Economics* 94, 239-247.

Beveridge, Stephen, and Charles R. Nelson, 1981, "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics* 7, 151-174.

Canova, Fabio, 1998, "Detrending and Business Cycle Facts," *Journal of Monetary Economics* 41, 475-512.

Canova, Fabio, Matteo Ciccarelli, and Eva Ortega, 2007, "Similarities and Convergence in G-7 Cycles," *Journal of Monetary Economics* 54, 850-878.

Carrasco, Marine, Liang Hu, and Werner Ploberger, 2014, "Optimal Test for Markov Switching Parameters," *Econometrica* 82, 765-784.

Clark, Peter K, 1987, "The Cycle Component of the U.S. Economic Activity," *Quarterly Journal of Economics* 102,797-814.

Cogley, Timothy, and James M. Nason, 1995, "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control* 19, 253-278.

Del Negro, Marco, and Christopher Otrok, 2008, "Dynamic Factor Models with Time-Varying Parameters: Measuring Changes in International Business Cycles," Staff Report 326, Federal Reserve Bank of New York.

Dickey, David. A. and Wayne A. Fuller,1979,. "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association* 74, 427-431.

Diebold, Francis X., Frank Schorfheide, and Minchul Shin, 2016, "Real-Time Forecast Evaluation of DSGE Models with Stochastic Volatility," The

Dupasquier, Chantel, and Nicholas Ricketts, 1998, "Non-Linearities in the Output-Inflation Relationship: Some Empirical Results for Canada," Bank of Canada Working Paper 98-14.

Elliott, Graham, Thomas. J. Rothenberg, and James H. Stock, 1996. "Efficient tests for an autoregressive unit root," *Econometrica* 64,813-836.

Fernandez-Villaverde, Jesus, Juan F. Rubio-Ramirez, Thomas J. Sargent and Mark W. Watson, 2007, "ABCs (and Ds) of Understanding VARs," *American Economic Review* 97(3), 1021-1026.

Garratt, Anthony, James Mitchell, and Shaun P. Vahey, 2014, "Measuring Output Gap Nowcast Uncertainty," *International Journal of Forecasting* 30, 268-279.

Geweke, John, and Gianni Amisano, 2011, "Optimal Prediction Pools," *Journal of Econometrics* 164, 130–141.

Guerrieri, Luca, and Matteo Iacoviello, 2016, "Collateral Constraints and Macroeconomic Asymmetries," Federal Reserve Board Working Paper.

Hamilton, James D., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57, 357-384.

Harvey, Andrew C. and Albert Jaeger, 1993, "Detrending, Stylized Facts and the Business Cycle," *Journal of Applied Econometrics* 8, 231-247.

Harvey, Andrew C., Thomas M. Trimbur, and Herman K. Van Dijk, 2007, "Trends and Cycles in Economic Time Series: A Bayesian Approach," *Journal of Econometrics* 140, 618-649.

Hodrick, Robert J. and Edward C. Prescott, 1997, "Postwar US Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking* 29,1-16.

IMF, 2013, World Economic Outlook "Hopes, Realities, Risks" Chapter 3, International Monetary Fund.

Huang, Yu-Fan, Sui Luo and Richard Startz, 2016, "Are Recoveries All the Same: GDP and TFP?" University of California Santa Barbara Working Paper.

Kamber, Gunes, James Morley, and Benjamin Wong, 2016, "Intuitive and Reliable Estimates of the Output Gap from a Beveridge-Nelson Filter." BIS Working Paper, No. 584.

Kiley, Michael T., 2013, "Output Gaps," *Journal of Macroeconomics*, 37, 1-18.

King, Thomas B., and James Morley, 2007, "In Search of the Natural Rate of Unemployment," *Journal of Monetary Economics* 54, 550-564.

Kim, Chang-Jin, 1994, "Dynamic Linear Models with Markov Switching," *Journal of Econometrics* 60, 1-22.

Kim, Chang-Jin, James Morley, and Jeremy Piger, 2005, "Nonlinearity and the Permanent Effects of Recessions," *Journal of Applied Econometrics* 20, 291-309.

Kim, Chang-Jin, and Charles R. Nelson, 1999, "Friedman's Plucking Model of Business Fluctuations: Tests and Estimates of Permanent and Transitory Components," *Journal of Money, Credit and Banking* 31, 317-334.

Klinger, Sabine and Enzo Weber, 2016, "Detecting Unemployment Hysteresis: A Simultaneous Unobserved Components Model with Markov Switching," *Economics Letters* 114(c), 115-118.

Kuttner, Kenneth N, 1994, "Estimating Potential Output as a Latent Variable," *Journal of Business & Economic Statistics* 12, 361-68.

Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt and Yoncheol Shin , 1992, "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics* 54, 159-178.

Levin, Andrew, T. and Jeremy M. Piger, 2006, "Is Inflation Persistence Intrinsic in Industrial Economies?" University of Oregon Working Paper.

Leybourne, Stephen, J. and Brendan P. M. McCabe, 1992, "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics* 64, 159-178.

Lucas, Robert E., 1976, "Economic Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1, 19-46.

Meier, André, 2010, "Still Minding the Gap - Inflation Dynamics during Episodes of Persistent Large Output Gaps," IMF Working Papers 10/189, International Monetary Fund.

Mitra, Sinchan and Tara Sinclair, 2012, "Output Fluctuations in the G-7: An Unobserved Components Approach," *Macroeconomic Dynamics* 16(3), 396-422.

Morley, James, 2002, "A State-Space Approach to Calculating the Beveridge-Nelson Decomposition," *Economics Letters* 75, 123-127.

Morley, James C., 2011, "The Two Interpretations Of The Beveridge–Nelson Decomposition," *Macroeconomic Dynamics* 15, 419-439.

Morley, James, 2014, "Measuring Economic Slack: A Forecast-Based Approach with Applications to Economies in Asia and the Pacific." BIS Working Paper, No. 451.

Morley, James, Charles R. Nelson, and Eric Zivot, 2003, "Why Are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP So Different?" *Review of Economics and Statistics* 85, 235-243.

Morley, James, Irina B. Panovska, and Tara M. Sinclair (2016), "Testing Stationarity with Unobserved Components Models," Macroeconomic Dynamics (forthcoming).

Morley, James, and Jeremy Piger, 2008, "Trend/Cycle Decomposition of Regime-Switching Processes," *Journal of Econometrics* 146, 220-226.

Morley, James, and Jeremy Piger, 2012, "The Asymmetric Business Cycle," *Review of Economics and Statistics* 94, 208-221.

Murray, Christian J., 2003, "Cyclical Properties of Baxter-King Filtered Time Series," *Review of Economics and Statistics* 85, 472-476.

Nelson, Charles R., 2008, "The Beveridge–Nelson Decomposition in Retrospect and Prospect," *Journal of Econometrics* 146, 202-206.

Nelson, Charles R., and Heejoon Kang, 1981, "Spurious Periodicity in Inappropriately Detrended Time Series," *Econometrica* 49, 741-751.

Nelson, Charles R. and Plosser, Charles, 1982, "Trends and random walks in macroeconmic time series: Some evidence and implications," *Journal of Monetary Economics* 10, 139-162.

Orphanides, Athanasios, 2002, "Monetary-Policy Rules and the Great Inflation," *American Economic Review* 92, 115-120.

Orphanides, Athanasios, and Simon van Norden, 2002, "The Unreliability of Output-Gap Estimates in Real Time," *Review of Economics and Statistics* 84, 569-583.

Perron, Pierre, and Tatsuma Wada, 2009, "Let's Take a Break: Trends and Cycles in U.S. Real GDP," *Journal of Monetary Economics* 56, 749-765.

Perron, Pierre and Tatsuma Wada, 2015, "Measuring Business Cycles with Structural Breaks and Outliers: Applications to International Data," *Research in Economics* (forthcoming).

Sims, Christopher A., 1980, "Macroeconomics and Reality," *Econometrica* 48, 1-48.

Sinclair, Tara M., 2010, "Asymmetry in the Business Cycle: Friedman's Plucking Model with Correlated Innovations," *Studies in Nonlinear Dynamics and Econometrics* 14, Issue 1, Article 3.

Stock, James H. and Mark W. Watson, 2005, "Understanding Changes in International Business Cycle Dynamics," *Journal of the European Economic Association* 3(5), 968-1006.

Stock, James H., and Mark W. Watson, 2009, "Phillips Curve Inflation Forecasts," in *Understanding Inflation and the Implications for Monetary Policy, a Phillips Curve Retrospective*, Federal Reserve Bank of Boston.

Stock, James H., and Mark W. Watson, 2010, "Modeling Inflation after the Crisis," in *Macroeconomic Challenges: The Decade Ahead*, Federal Reserve Bank of Kansas City, 173-220. Kansas City, MO: Federal Reserve Bank of Kansas City.

Yellen, Janet L., 2015, The Economic Outlook and Monetary Policy. Speech at the Economic Club of Washington, D.C. December 2nd, 2015. http://www.federalreserve.gov/newsevents/speech/yellen20151202a.htm

Table 1
Structural Breaks in Long-Run Growth Rates of Real GDP

	Sample Period	Break Dates	Sequence of Growth Regimes
United States	1947Q2-2016Q1	1973Q1, 2000Q3	H, M, L
Australia	1959Q4-2015Q4	-	-
Canada	1960Q2-2015Q4	1974Q2	H, L
France	1949Q2-2016Q1	1974Q2	H,L
Germany	1960Q2-2016Q1	1973Q1,1991Q2	H,M,L
Italy	1960Q2-2016Q1	1974Q1	H, L
Japan	1955Q2-2016Q1	1973Q1,1991Q3	H,M,L
Korea	1970Q2-2016Q1	1997Q3	H,L
New Zealand	1977Q2-2016Q1	-	-
United Kingdom	1955Q2-2016Q1	1973Q2	H,L

Notes: Estimated break dates are based on Bai and Perron's (1998, 2003) sequential procedure. Breaks are significant at least at 10% level. "H", "M", "L" denote high, medium, and low mean growth regimes, respectively.

Table 2
Correlation with Subsequent 4-Quarter Output Growth

	Sample Period	AR(1) Model Output Gap	UC-HP Model Output Gap	Model-Avg. Output Gap
United States	1947Q2-2015Q1	-0.15	0.08	**-0.38**
Australia	1959Q1-2014Q4	-0.04	-0.01	**-0.27**
Canada	1960Q1-2014Q4	-0.16	-0.18	**-0.27**
Germany	1960Q1-2015Q1	-0.07	-0.001	**-0.08**
France	1949Q1-2015Q1	-0.11	0.13	**-0.18**
Italy	1960Q1-2015Q1	-0.18	0.03	**-0.33**
Japan	1955Q2-2015Q1	0.02	0.05	**-0.14**
Korea	1970Q2-2015Q1	-0.04	-0.03	**-0.20**
New Zealand	1977Q3-2015Q1	0.03	0.04	**-0.22**
United Kingdom	1955Q2-2015Q1	0.21	-0.22	**-0.35**

Note: Bold denotes the most negative correlation for each economy.

Table 3
Correlation with Subsequent 4-Quarter Change in Inflation

	Sample Period	AR(1) Model Output Gap	UC-HP Model Output Gap	Model-Avg. Output Gap
United States	1960Q1-2015Q1	-0.11	0.32	**0.46**
Australia	1959Q4-2014Q4	0.20	0.35	**0.38**
Canada	1960Q1-2014Q4	-0.25	**0.44**	0.35
Germany	1963Q1-2015Q1	-0.21	**0.49**	0.04
France	1971Q1-2015Q1	-0.17	**0.11**	-0.05
Italy	1961Q1-2015Q1	-0.26	**0.19**	-0.29
Japan	1961Q2-2015Q1	0.22	0.29	**0.33**
Korea	1970Q2-2015Q1	-0.12	0.31	**0.33**
New Zealand	1977Q3-2015Q1	-0.32	**0.39**	0.25
United Kingdom	1957Q4-2015Q1	-0.14	**0.22**	0.08

Note: Bold denotes the most positive correlation for each economy.

Table 4
Tests For Markov-Switching Models

		Alternatives		
	Null	L-shaped	U-shaped	Depth
United States	AR(2)	0.151	2.516	8.401
		(0.409)	(0.166)	(0.055)
		1.213	4.272	9.243
	AR(2)-t	0.307	1.373	9.554
		(0.164)	(0.161)	(0.035)
		0.721	2.797	9.022
Australia	AR(2)	0.880	0.197	14.826
		(0.116)	(0.688)	(0.005)
		1.833	3.270	9.218
	AR(2)-t	0.637	0.020	10.686
		(0.070)	0.999	(0.045)
		0.904	2.814	9.195
Canada	AR(2)	0.003	1.914	24.122
		(0.989)	0.221	(0.000)
		0.932	3.516	9.224
	AR(2)-t	0.003	1.914	17.825
		(0.689)	(0.221)	(0.000)
		0.932	3.516	9.575
Germany	AR(2)	0.974	3.688	59.000
		(0.210)	(0.030)	(0.000)
		1.109	3.376	8.846
	AR(2)-t	0.030	6.250	108.344
		(0.437)	(0.000)	(0.000)
		0.885	2.2886	10.756
France	AR(2)	0.001	1.220	2.803
		(1.000)	(0.432)	(0.825)
		1.223	3.915	27.829
	AR(2)-t	0.000	0.673	50.794
		(1.000)	(0.236)	(0.000)
		1.507	2.458	9.826
Italy	AR(2)	1.962	1.065	1.827
		(0.035)	(0.452)	(0.800)
		1.736	4.641	10.903
	AR(2)-t	0.057	1.171	0.473
		(0.462)	(0.201)	(0.960)
		1.356	2.255	10.732
Japan	AR(2)	0.492	2.752	36.310
		(0.146)	(0.121)	(0.081)
		1.353	4.177	54.323
	AR(2)-t	3.774	2.527	15.378
		0.000	(0.040)	(0.011)
		1.315	2.397	10.137
Korea	AR(2)	0.027	0.369	17.964
		(0.389)	(0.382)	(0.290)
		1.172	2.251	27.062
	AR(2)-t	0.026	0.0940	2.079
		(0.527)	(0.537)	(0.825)
		1.449	3.896	11.332
New Zealand	AR(2)	1.231	0.138	6.198
		(0.085)	(0.758)	(0.265)
		1.458	4.036	11.115
	AR(2)-t	1.235	0.206	2.974
		(0.030)	(0.462)	(0.570)
		0.917	2.157	10.055
United Kingdom	AR(2)	0.001	2.969	16.000
		(1.000)	(0.075)	(0.002)
		1.158	3.399	9.914
	AR(2)-t	0.001	0.065	6.592
		(1.000)	(0.708)	(0.260)
		0.993	2.440	11.464

The test statistics for the L-shaped and U-shaped Recessions are based on Carrasco et al. (2014) The test statistics for the depth-based recovery alternatives are based on estimation using a grid for the continuous probabilities. All p-values and critical values are based on parametric bootstrap experiments with 499 simulations. All tests accounted for structural breaks in the long-run growth rate.

Table 5
Unit Root and Stationarity Tests

		Test				
	Adjustment for structural breaks	ADF (asymptotic p-value)	DF ERS*	LMC (bootstrapped p-value)	KPSS (bootstrapped p-value)	MPS (bootstrapped p-value)
US	1973Q1,2000Q3	-3.201 (0.085)	8.903	0.085 (0.362)	0.163 (0.182)	1.634 (0.065)
AU	None	-1.834 (0.363)	3.209	2.088 (0.330)	0.211 (0545)	10.876 (0.015)
CA	1974Q2	-2.289 (0.438)	2.404	3.411 (0.010)	0.378 (0.116)	3.698 (0.201)
FRA	1974Q2	-1.585 (0.796)	2.575	1.897 (0.377)	0.186 (0.683)	7.835 (0.000)
DEU	1973Q1,1991Q2	-2.696 (0.239)	2.889	2.564 (0.025)	0.274 (0.055)	12.440 (0.000)
IT	1974Q1	0.525 (0.993)	2.686	1.502 (0.151)	0.318 (0.729)	3.080 (0.101)
JP	1973Q1,1991Q3	-3.147 (0.098)	2.461	0.063 (0.603)	0.152 (0.357)	0.030 (0.537)
KR	1997Q3	-3.055 (0.120)	3.078	0.071 (0.839)	0.574 (0.386)	0.430 (0.307)
NZ	None	-2.618 (0.273)	4.072	1.573 (0.261)	0.182 (0.407)	6.432 (0.100)
UK	1973Q2	-2.448 (0.353)	2.981	1.225 (0.256)	0.139 (0.708)	12.294 (0.005)

The 5% asymptotic critical value for the DF-ERS unit root tests is -1.941. We also performed unit root and stationarity tests that allowed for structural breaks in the variance, and unit root tests that did not allow for structural breaks in the long-term drift. The results for the different specifications that allow for breaks in the variance and specifications that do not allow for structural breaks in means are available upon request. Allowing for structural breaks in the variance did not alter the p-values of any of the tests substantially.

Table 6
Correlation with the Unemployment Rate

	Sample Period	AR(1) Model Output Gap	UC-HP Model Output Gap	Model-Avg. Output Gap
United States	1948Q1-2016Q1	0.05	-0.14	**-0.64**
Australia	1978Q1-2015Q4	0.06	-0.01	**-0.43**
Canada	1960Q1-2015Q4	-0.01	-0.02	**-0.34**
Germany	1991Q1-2016Q1	-0.03	-0.11	**-0.24**
France	1978Q1-2016Q1	-0.01	0.05	**-0.25**
Italy	1983Q1-2016Q1	-0.07	0.27	**-0.20**
Japan	1955Q3-2016Q1	0.02	-0.05	**-0.06**
Korea	1990Q1-2016Q1	-0.21	0.08	**-0.69**
New Zealand	1977Q3-2016Q1	0.00	0.19	**-0.47**
United Kingdom	1983Q1-2016Q1	-0.16	0.20	**-0.44**

Note: Bold denotes the most negative correlation for each economy.

Table 7
Correlation with Capacity Utilization

	Sample Period	AR(1) Model Output Gap	UC-HP Model Output Gap	Model-Avg. Output Gap
United States	1967Q1-2016Q1	-0.08	0.27	**0.49**
Australia	1989Q3-2016Q4	0.14	0.39	**0.65**
Canada	1987Q1-2015Q4	-0.47	0.54	**0.76**
Germany	1960Q1-2016Q1	-0.19	**0.64**	0.37
France	1976Q1-2016Q1	-0.20	0.33	**0.71**
Italy	1968Q4-2016Q1	-0.21	**0.47**	0.03
Japan	1978Q1-2016Q1	0.17	0.39	**0.53**
Korea	1980Q1-2016Q1	-0.26	0.37	**0.74**
New Zealand	1977Q3-2016Q1	-0.25	0.28	**0.57**
United Kingdom	1985Q1-2015Q1	-0.26	**0.56**	0.45

Note: Bold denotes the most positive correlation for each economy.

Table 8a
Correlation Between Model Averaged Output Gaps
Longest Available Sample (Full Sample)

	US	AU	CA	DEU	FRA	IT	JP	KR	NZ	UK
US	•									
AU	0.451	•								
CA	0.660	0.628	•							
DEU	0.297	0.040	0.033	•						
FRA	0.074	0.033	0.232	0.115	•					
IT	0.193	0.066	0.117	0.229	0.374	•				
JP	0.468	0.128	0.314	0.314	0.384	0.137	•			
KR	0.317	-0.082	0.006	0.313	-0.010	-0.017	0.203	•		
NZ	0.254	0.530	0.282	0.003	-0.174	0.201	0.098	0.034	•	
UK	0.283	0.144	0.400	0.219	0.451	0.101	0.493	0.043	0.033	•

Table 8b
Correlation Between Model Averaged Output Gaps
Before 1984

	US	AU	CA	DEU	FRA	IT	JP	KR	NZ	UK
US	•									
AU	0.354	•								
CA	0.591	0.605	•							
DEU	0.269	0.070	-0.056	•						
FRA	-0.201	0.001	-0.111	-0.188	•					
IT	-0.010	-0.112	-0.131	0.047	0.209	•				
JP	0.357	0.141	0.093	0.153	-0.239	-0.136	•			
KR	0.218	-0.196	-0.197	0.349	-0.416	-0.352	-0.123	•		
NZ	-0.123	0.499	0.191	0.294	0.151	0.293	0.335	-0.524	•	
UK	0.232	0.195	0.070	0.356	0.152	-0.045	0.331	-0.024	-0.029	•

Table 8c
Correlation Between Model Averaged Output Gaps
1984Q1-2016Q1

	US	AU	CA	DEU	FRA	IT	JP	KR	NZ	UK
US	•									
AU	0.532	•								
CA	0.796	0.686	•							
DEU	0.167	-0.210	0.018	•						
FRA	0.514	0.048	0.431	0.331	•					
IT	0.292	0.056	0.194	0.239	0.560	•				
JP	0.731	0.117	0.446	0.441	0.657	0.316	•			
KR	0.175	-0.135	0.082	0.106	0.133	-0.042	0.328	•		
NZ	0.288	0.485	0.291	0.049	-0.253	0.016	0.060	0.188	•	
UK	0.657	0.239	0.680	0.293	0.595	0.378	0.614	0.090	0.046	•

The pairwise correlations in each case are based on the shorter available sample period in Table 1.

Table 9a
Granger Causality Tests for Model-Averaged Output Gaps
Longest Available Sample (Full Sample)

	US	AU	CA	DEU	FRA	IT	JP	KR	NZ	UK
US	•	✓✓✓	✓✓✓						✓	✓✓✓
AU		•							✓✓✓	
CA	✓✓✓	✓✓✓	•		✓✓✓				✓✓✓	✓✓
DEU				•	✓✓	✓✓		✓✓✓		
FRA	✓				•		✓✓			
IT					✓✓	•				
JP	✓✓			✓	✓✓✓	✓✓	•			✓✓✓
KR	✓							•		
NZ		✓✓							•	
UK					✓✓✓			✓✓		•

Table 9b
Granger Causality Tests for Model-Averaged Output Gaps
Before 1984

	US	AU	CA	DEU	FRA	IT	JP	KR	NZ	UK
US	•	✓✓✓	✓✓✓						✓	✓
AU		•	✓						✓✓✓	
CA		✓✓✓	•						✓	
DEU				•				✓✓✓		
FRA					•		✓✓			
IT						•				
JP		✓					•			✓✓
KR								•		
NZ		✓✓							•	
UK										•

Table 9c
Granger Causality Tests for Model-Averaged Output Gaps
1984Q1-2016Q1

	US	AU	CA	DEU	FRA	IT	JP	KR	NZ	UK
US	•	✓✓✓	✓✓✓		✓✓✓	✓✓				✓✓✓
AU		•	✓		✓✓				✓✓✓	✓
CA	✓✓✓	✓✓✓	•		✓✓✓	✓✓✓				✓✓✓
DEU	✓			•	✓✓	✓✓		✓✓		
FRA	✓				•	✓✓✓				
IT			✓✓✓		✓✓✓	•				
JP	✓✓✓		✓✓✓		✓✓✓	✓✓✓	•			✓✓✓
KR					✓			•		
NZ		✓✓							•	
UK	✓✓✓	✓	✓✓✓		✓✓✓	✓✓✓				•

Notes: Results are based on pairwise Granger Causality tests with 2 lags of quarterly data. A checkmark denotes that the output gap in the row economy "causes" the output gap in the column economy. One checkmark denotes significance at the 10% level, two checkmarks denote significance at the 5% level, and three checkmarks denote significance at the 1% level. See the data description in the text for details on economy abbreviations. The pairwise regressions in each case are based on the shorter available sample period in Table 1.

Table A.1

Summary of Data and Data Sources

Economy	Gross Domestic Product	Inflation	Unemployment	Capacity
United States	Quarterly, real, SAGDPC1 FRED	Quarterly, SA JCXFE FRED	Monthly, SA UNRATE FRED	Monthly, SA TCU FRED
AU	Quarterly, real, SA OECD LNBRQSA (ABS)	Quarterly, SA OECD	Monthly, SA GLFSURSA, ABS	Monthly, SA National bank survey NAB Data
CA	Quarterly, real SA OECD VOBARSA	CPI, CPI Core Monthly (SA, NSA) StatCan	Monthly, SA OECD MEI	Quarterly, SA StatCan (NAICS) Series Code 029-002
DEU	Quarterly, Real, SA OECD LBRQRSA	CPI, CPI Core Monthly (SA, NSA) OECD MEI	Monthly, SA OECD MEI	Quarterly, SA OECD MEI
FRA	Quarterly, real, SA OECD LBRQRSA	CPI, CPI Core Monthly (SA, NSA) OECD MEI	Monthly, SA OECD MEI	Quarterly, SA OECD MEI
IT	Quarterly, real, SA OECD VOBARSA	CPI, CPI Core Monthly (SA, NSA) OECD MEI	Monthly, SA OECD MEI	Quarterly, SA OECD MEI
JP	Quarterly, Real, NSA Cabinet Office	CPI, CPI Core Monthly, NSA OECD MEI	Monthly, SA Cabinet Office	Quarterly, SA Japan Ministry of Economy
KR	Quarterly, Real, SA OECD VOBARSA	CPI, CPI Core (NSA, SA)	Monthly, SA OECD MEI	Quarterly, SA KOSTAT
UK	Quarterly, Real, SA OECD VOBARSA	CPI, CPI Core, SA OECD MEI	Monthly, SA OECD MEI	Quarterly, SA Office of National Statistics (Business Tendency Survey)

All monthly series were converted to quarterly frequency using arithmetic averages. The series that were not seasonally adjusted by the source were seasonally adjusted using the X12 filter. To facilitate comparison with previous studies, we had a preference for OECD VOBARSA GDP series, except when an alternative measure was available for a much longer sample. In all cases when we used a series other than the VOBARSA measure, the correlation with the VOBARSA measure for the overlapping sample periods was above 0.97. Similarly, we had a preference for the OECD Main Economic Indicator (MEI) harmonized unemployment rate, except when an alternative measure was available for a much longer sample. In the case of the US, the FRED series match the preferred OECD measures.

Table A.2
Structural Break Tests

Economy	Number of breaks	Break Dates (Test Statistics and significance level)	
US	1 (2)	2000Q3 (19.10***)	1973Q1 (6.88) p-value =0.13
Australia	0	-	
Canada	1	1974Q2 (20.278***)	
France	1	1974Q2 (65.82***)	
Germany	1 (2)	1973Q2 (15.871***)	1991Q2 (4.95) p-value=0.11
Italy	1	1974Q1 (48.127***)	
Japan	2	1973Q2 (131.695***)	1991Q3 (19.87***)
Korea	1	1997Q3 (26.07***)	
New Zealand	0	-	
UK	0 (1)	1973Q1 (6.07) p-value=0.15	

Note: The table reports the results of the Bai-Perron (1998,2003) sequential test. We consider trimming of 15% of the sample from its end points and between breaks for admissible break dates. The table reports the number of breaks, the test statistics and the significance level (with three stars corresponding to significance at the 1% level, two starts corresponding to significance at the 5% level, and one star corresponding to significance at the 10% level). In the cases when the Bai-Perron test selected a smaller number of breaks than the number of breaks commonly imposed in the literature, we list the maximum number of breaks we considered in parenthesis, and the p-value for the additional break dates below the test statistics.

Table A.3
Prior Distributions for Model Parameters

	Parameter Description	Model(s)	Prior
μ	Unconditional mean growth	All except UC-HP and BB	$N(1,3^2)$
γ_0	Growth in expansion regime	BB	$N(2.5,3^2)$
$-\gamma_1$	Impact of other regime	BB	$Gamma(\frac{15}{2},\frac{5}{2})$
λ	Bounceback coefficient	BB	$N(0,0.25^2)$
$-\tau$	Mean of transitory shocks in other regime	UC-FP	$Gamma(\frac{15}{2},\frac{5}{2})$
ϕ_j	AR parameter at lag j	All except UC-HP	$TN(0,(0.25/j)^2)_{[\lvert z\rvert>1,\phi(z)=0]}$
p_{00}	Expansion regime continuation probability	BB, UC-FP	$Beta(1,20)$
p_{11}	Other regime continuation probability	BB, UC-FP	$Beta(5,15)$
ν	Degree of freedom for Student t errors	All except UC	$Gamma(\frac{1}{2},\frac{0.1}{2})$
$\frac{1}{\sigma_e}, \frac{1}{\sigma_\eta}, \frac{1}{\sigma_\omega}$	Precision for independent shocks	All except UCUR and UCUR-FP	$Gamma(\frac{5}{2},\frac{2}{2})$
$\Sigma_{\eta\omega}^{-1}$	Precision for correlated shocks	UCUR and UCUR-FP	$Wishart(5,2\times I_2)$

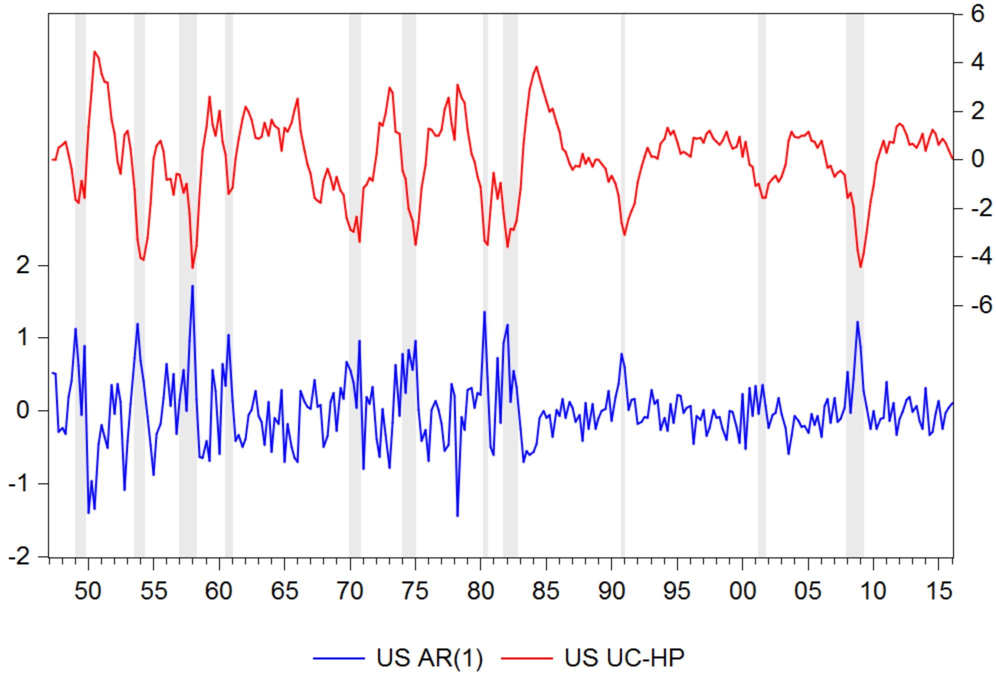


Fig. 1 – Output gaps based on competing models of US real GDP (NBER recessions shaded)

Note: The output gap for an AR(1) model for 1947Q2-2016Q1 is in blue and the output gap for a UC-HP model for the corresponding sample period is in red.

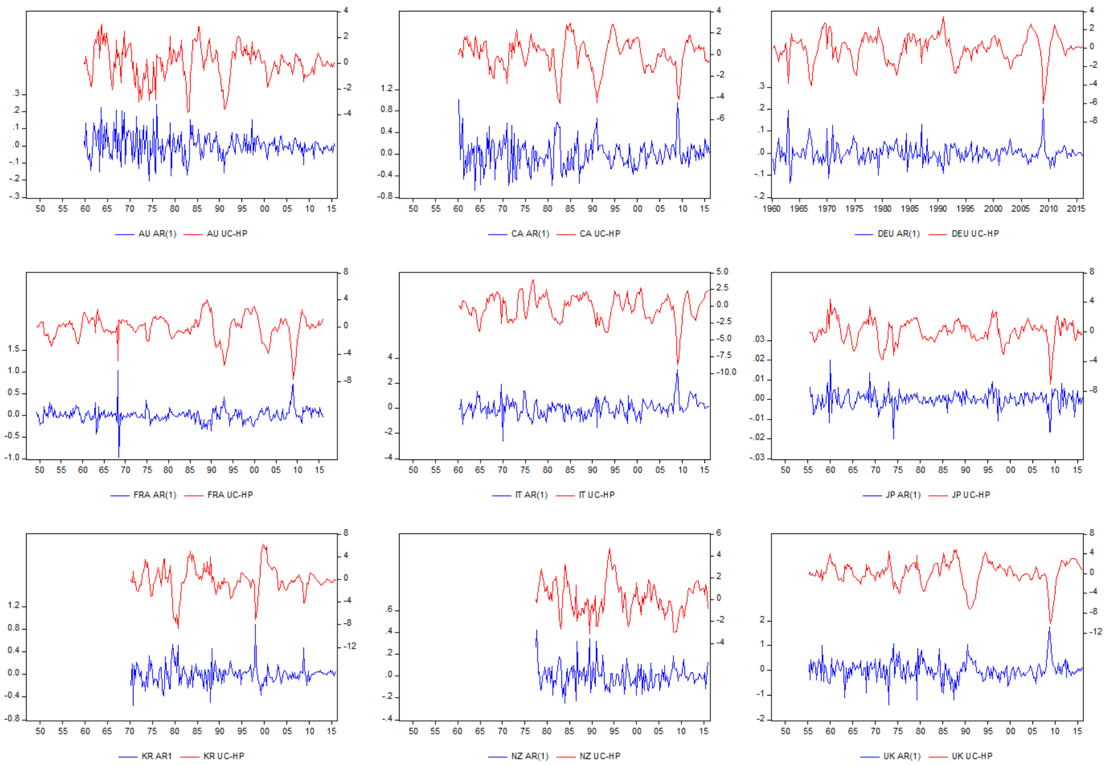


Fig. 2 – Output gaps based on competing models of real GDP for selected industrialized economies

Notes: From the top left and by row, the economies are Australia, Canada, Germany, France, Italy, Japan, Korea, New Zealand, and the UK. The output gap for an AR(1) model is in blue (left axes) and the output gap for a UC-HP model is in red (right axes). The horizontal axis runs from 1947Q2-2016Q1. See Table 1 for details of the available sample period for each economy.

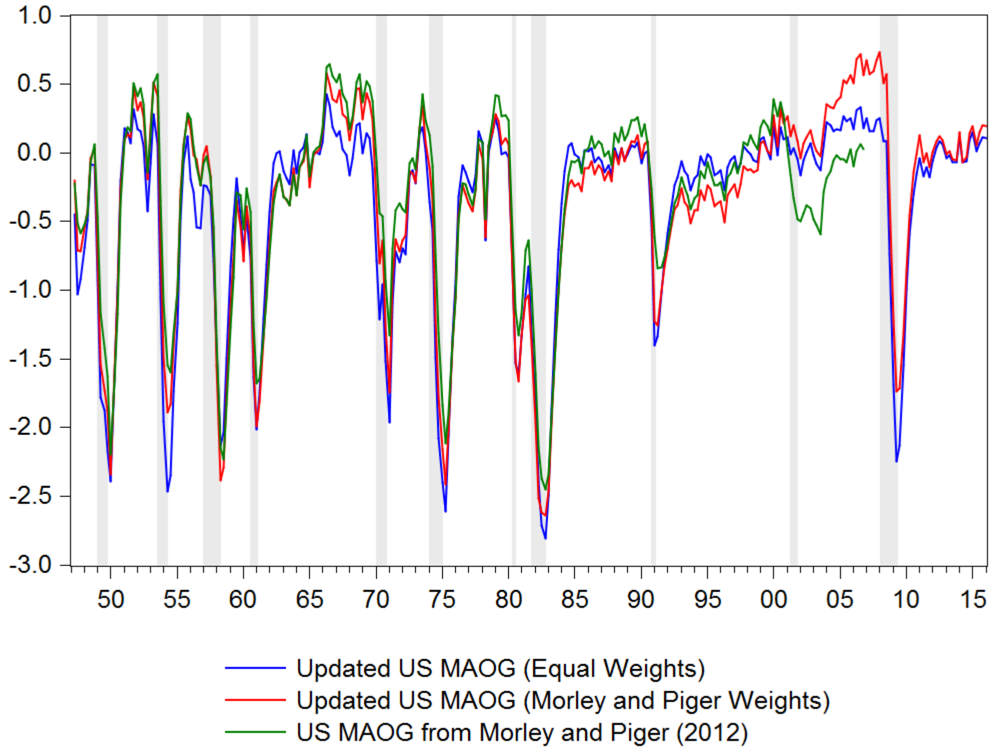


Fig. 3 – Model-averaged output gap for US real GDP for different weighting schemes, estimation methods, and sample periods (NBER recessions shaded)

Note: The model-averaged output gap for the 1947Q2-2016Q1 sample based on equal weights and Bayesian estimation is in blue, the model-averaged output gap for the vintage 1947Q2-2006Q4 sample from Morley and Piger (2012) based on BMA weights and MLE is in red, and the model-averaged output gap for the 1947-2016Q1 sample based on BMA weights and MLE is in green.

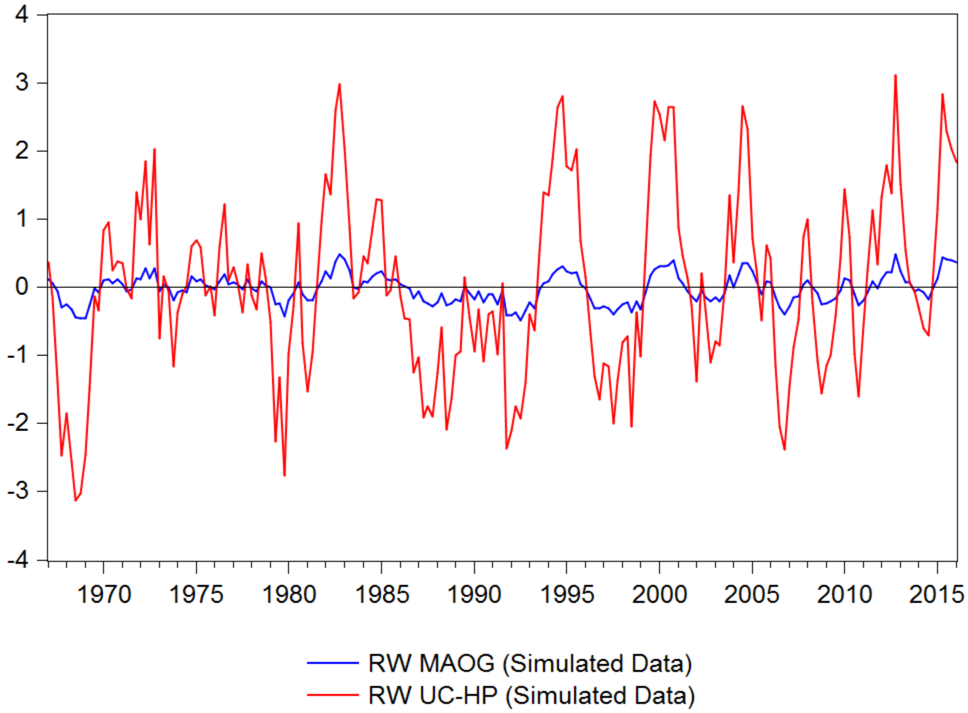


Fig. 4 – Model-averaged output gaps for a simulated random walk

Note: The model-averaged output gap for a simulated random walk of a sample length corresponding to the length of the observed sample for U.S. GDP is in blue and the output gap for a UC-HP model for the same simulated random walk is in red.

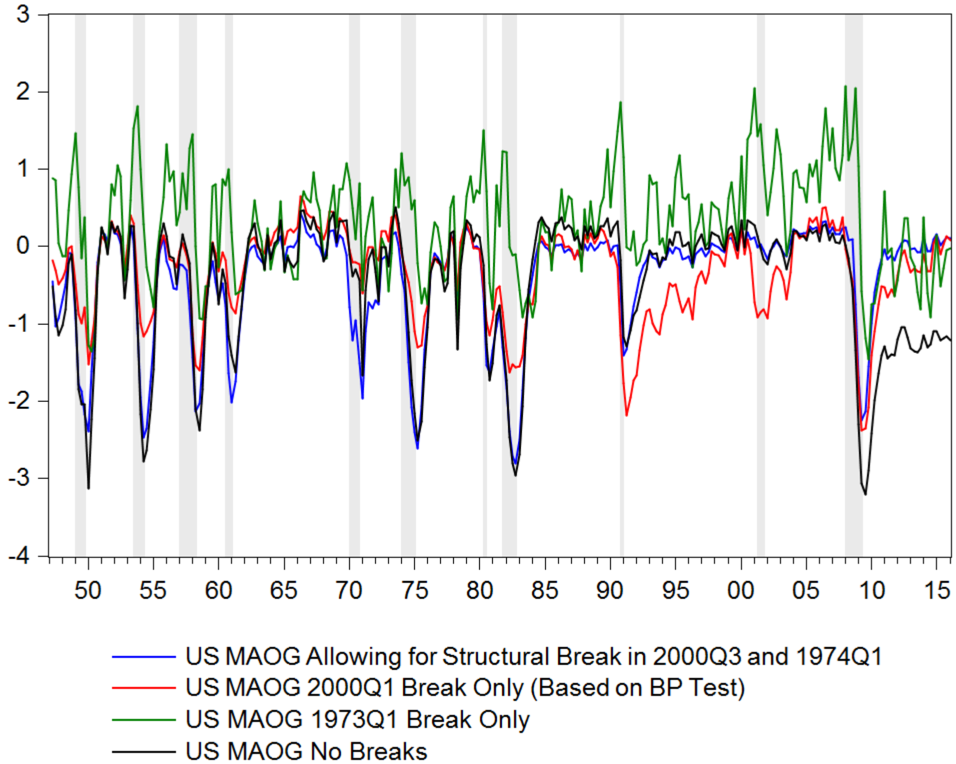


Fig. 5 – Model-averaged output gap for US real GDP with and without structural break in long-run growth (NBER recessions shaded)

Note: The model-averaged output gap for US real GDP for 1947Q2-2016Q1 allowing for structural breaks in long-run growth in 1973Q1 and 2000Q3 is in blue, the model-averaged output gap for US real GDP for the corresponding period assuming a break in 2000Q3 only is in red, the model-averaged output gap for US real GDP for the corresponding period assuming a structural break in 1973Q1 only is in green, and the model-averaged output gap for US real GDP for the corresponding sample period, but assuming no structural breaks is in black.

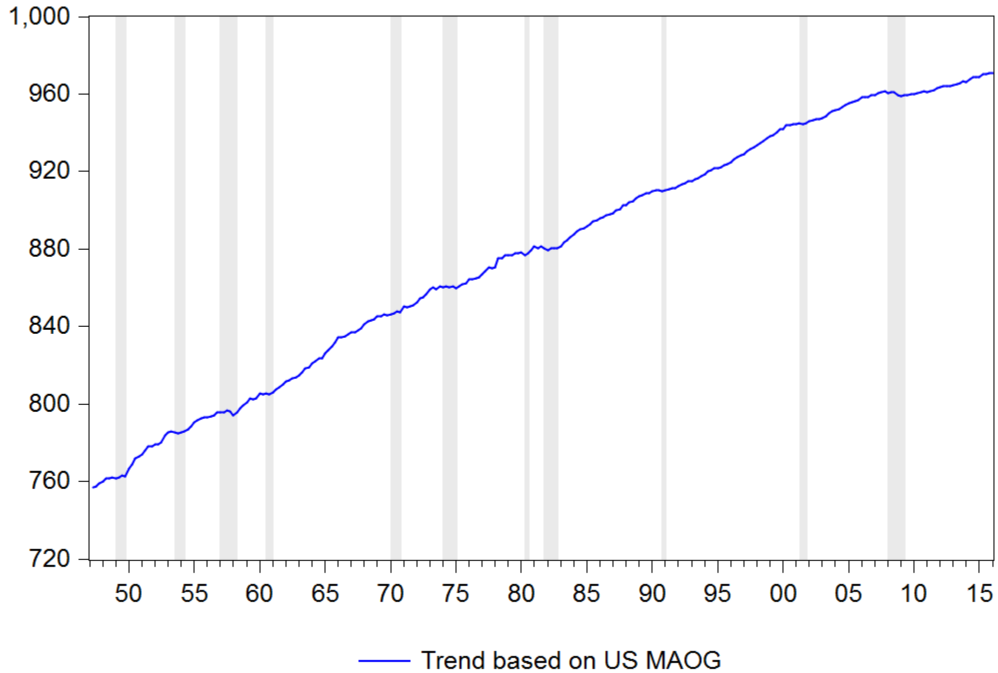
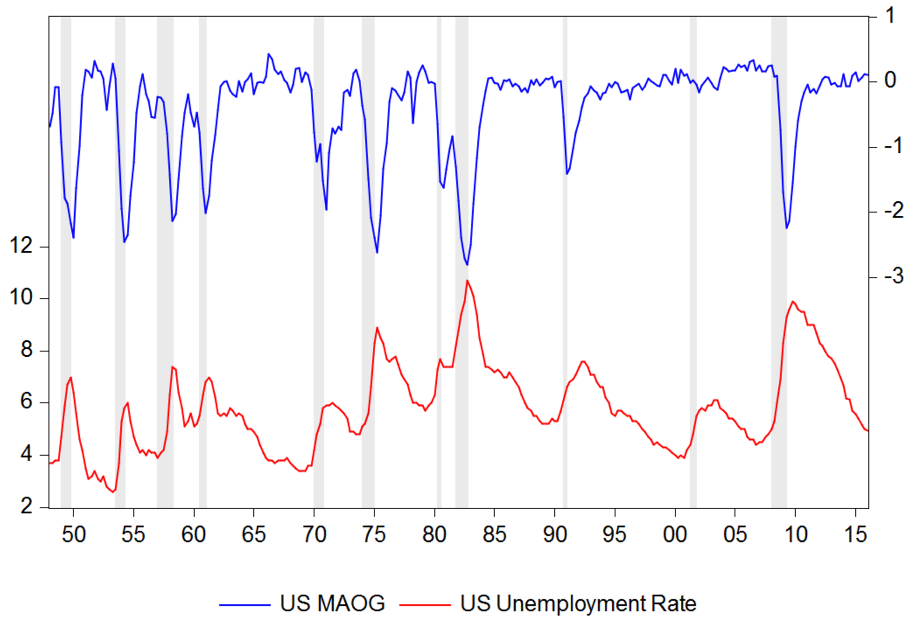


Fig. 6 – Estimated trend in US real GDP based on model-averaged output gap (NBER recessions shaded)

Note: The trend estimate is calculated as the difference between 100 times log US real GDP and the US model-averaged output gap for 1947Q2-2016Q1.



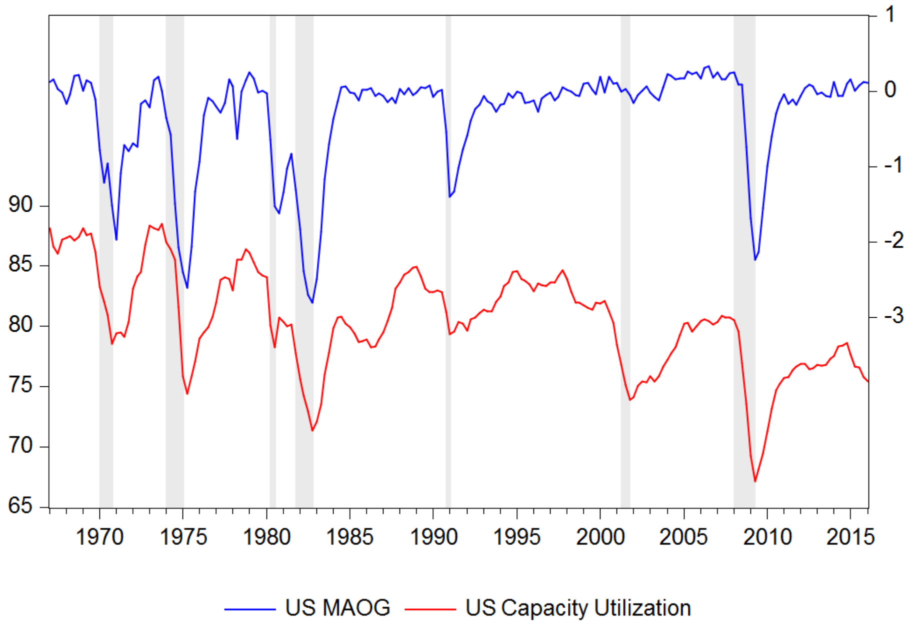


Fig. 7 – Model-averaged output gap for US real GDP and other measures of economic slack (NBER recessions shaded)

Notes: In the top panel, the model-averaged output gap for US real GDP for 1948Q1-2016Q1 is in blue and the unemployment rate for the corresponding sample period is in red. In the bottom panel, the model-averaged output gap for US real GDP for 1967Q1-2016Q1 is in blue and capacity utilization for the corresponding sample period is in red.

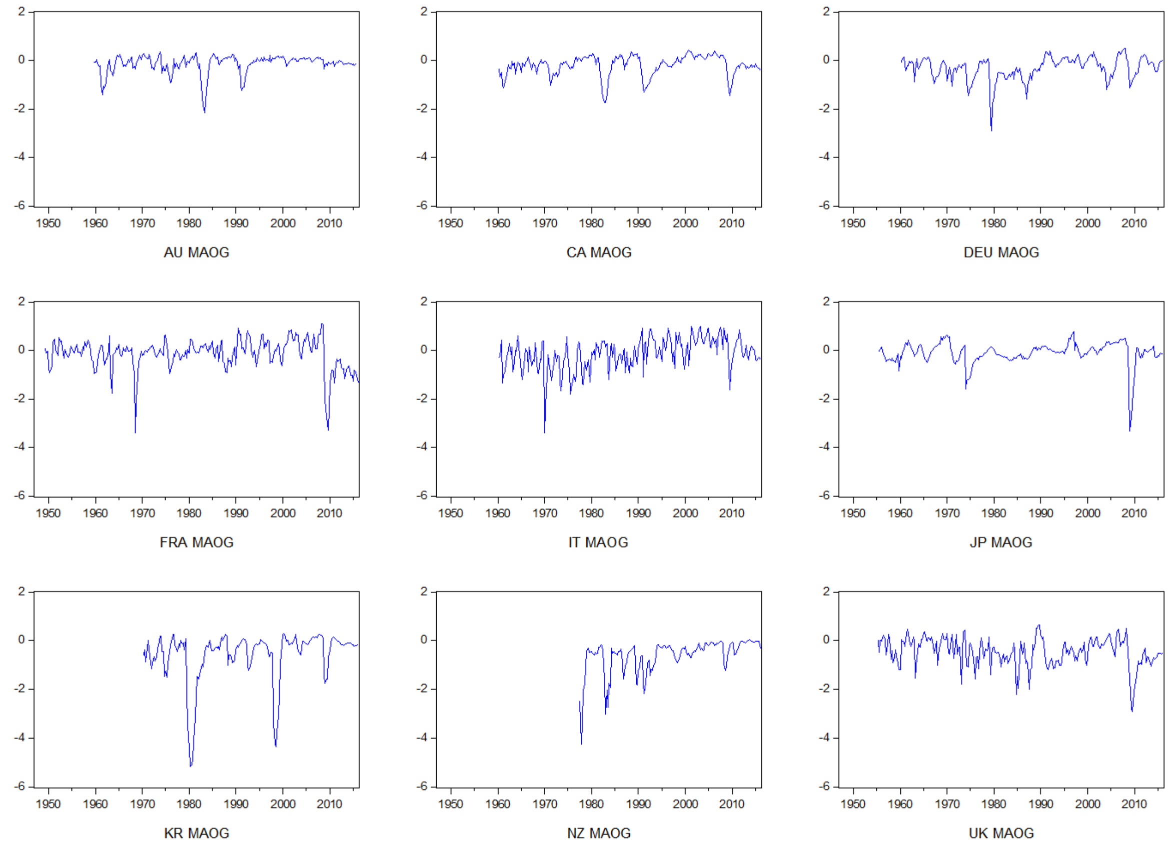


Fig. 8 – Model-averaged output gaps for real GDP from selected industrialized economies

Notes: From the top left and by row, the economies are Australia, Canada, Germany, France, Italy, Japan, Korea, New Zealand, and the United Kingdom. The horizontal axis runs from 1947Q2-2016Q1. See Table 1 for details of the available sample period for each economy.

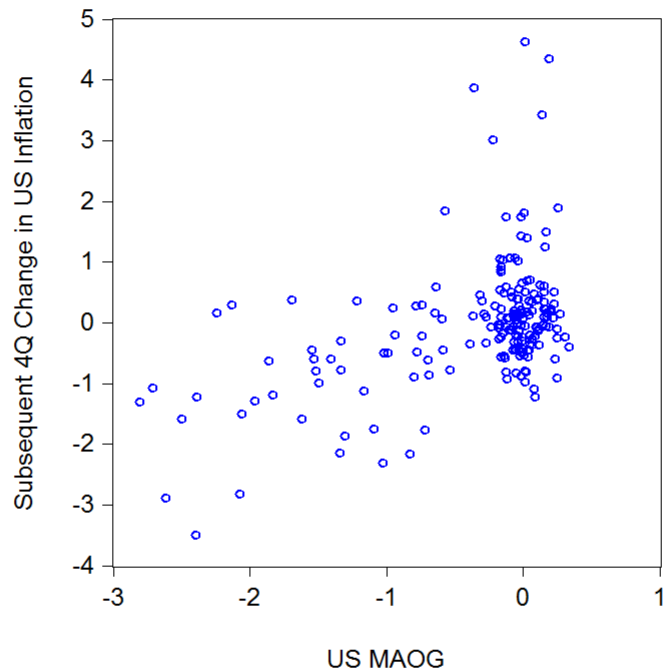


Fig. 9 – US Phillips Curve based on model-averaged output gap

Note: The scatterplot is for the sample period of 1960Q1-2015Q1 based on availability of the core PCE deflator measure of US inflation.

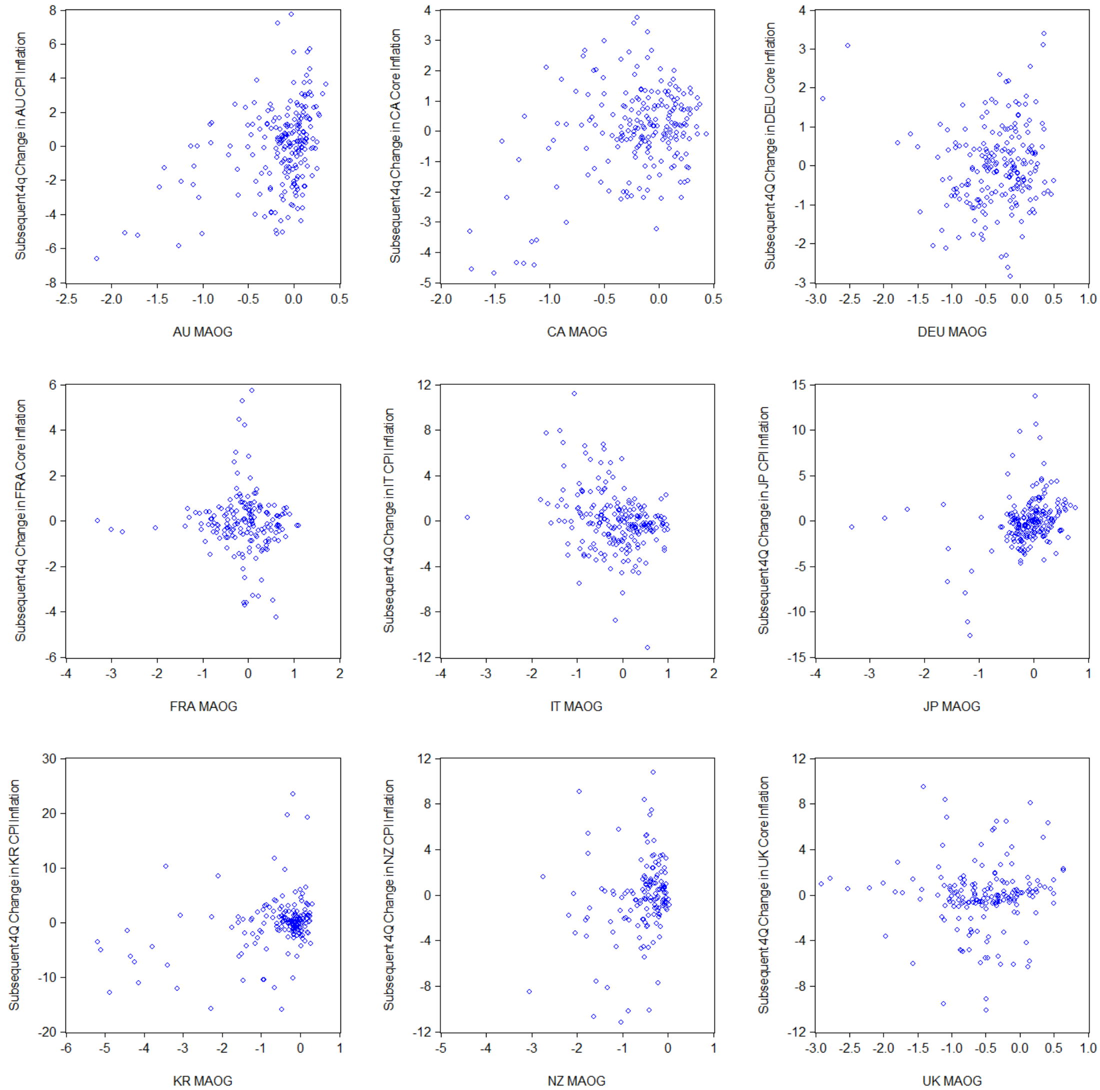


Fig. 10 –Phillips Curves based on model-averaged output gaps for selected industrialized economies

Notes: From the top left and by row, the economies are Australia, Canada, Germany, France, Italy, Japan, Korea, New Zealand, and the United Kingdom. See Table 3 for details of the sample period for each economy and the data description in the text for the corresponding inflation measure.