FISCAL POLICY UNDER UNCERTAINTY

Jamie Murray

Department of Economics

University of Sheffield

A THESIS SUBMITTED FOR THE DEGREE

Doctor of Philosophy

OCTOBER 2018

Abstract

Fiscal policy is interesting again.

As interest rates hit rock bottom during the global financial crisis, the burden of lifting demand fell upon governments. This thesis explores three related issues that have surfaced in the aftermath. Once the storm passed, decisions had to be made about how fast and when to withdraw stimulus — chapter two shows that implementing austerity may not always be plain sailing. How much belt tightening is needed after a crisis? That depends crucially on the amount of slack in the economy – revenue will grow and spending will fall naturally as the economy recovers. The third chapter quantifies the challenge of measuring the output gap in real time and shows how this carries over to estimates of structural borrowing. Finally, the next crisis cannot be too far away. But when should a government junk its budget plans and loosen the purse strings? The trigger for that is the topic of chapter four.

In chapter two, I present a simple estimated model of the New Zealand economy which is used to assess the sensitivity of the impact multiplier and output losses associated with fiscal consolidations to uncertainty over model parameters. I find that, in normal times, the fiscal multiplier can be expected to lie between 0.1 and 0.5, with a central estimate of 0.3. Uncertainty over the output effects of fiscal tightening can be attributed to several model parameters and it is found that a bad outcome is likely to be worse than a good outcome is to be better — output risks are skewed to the downside. Sensitivity analysis reveals that if monetary policy in New Zealand were to be constrained by the zero-lower bound, the fiscal impact multiplier would rise substantially, consistent with the empirical evidence for other OECD countries in that position.

In chapter three I present a range of measures of slack in the UK economy. It is shown that output gap uncertainty is substantial. Revisions owing to the arrival of new data are on average of the same magnitude as the output gap itself, while uncertainty arising from data revisions is found to make a smaller contribution. Model uncertainty is pervasive and all types of output gap uncertainty carry over to measures of structural borrowing. Assuming the Brainard Principle holds, the enormous margin of uncertainty over the cyclical position of the economy calls for caution in fiscal policymaking.

Chapter four assesses a novel rule that was introduced in the UK in 2015. It gave the British government fiscal flexibility whenever GDP growth warranted it. This rule lasted just a year, but it had features worth exploring. I apply solution methods for models with occasionally-binding constraints to assess the demand stabilisation properties of state-contingent fiscal rules.

First it is shown that fiscal flexibility can make recessions shallower. Second, it is suggested that GDP growth, rather than measures of the output gap, is probably the best indicator for triggering fiscal flexibility.

JEL classification: E62, E43, F33, F41, C50, E27, E32

Keywords: linear perturbation, fiscal impact multiplier, DSGE, monetary policy, potential output, output gap, uncertainty, real time, fiscal policy, productivity, Ricardian equivalence, SVAR, consolidation, uncertainty, zero lower bound

For Helen and in memory of Wendy Perren.

Acknowledgements

Thanks to Christoph Thoenissen, my supervisor, for putting me on the right path at various times and, of course, for all the flat whites. I'm indebted to Tugrul Vehbi, Girol Karacaoglu and anonymous referees for support and guidance during my time in New Zealand working on chapter two. To Andy King, Stephen Nickell, Robert Chote and the Office for Budget Responsibility advisory panel for helpful suggestions on chapter three. And to Dan Hanson for inspecting endless simulations with me and bouncing ideas back and forth on chapter four.

Finally, thanks to the New Zealand Treasury, the Office for Budget Responsibility and Bloomberg for allowing time for me to work on this.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Motivation

Each of the chapters of this thesis was intended to answer a specific question relevant to policy at the time of writing.

Chapter two concerned the possible impact of fiscal tightening in New Zealand when the government was embarking on a substantial program of consolidation. Austerity carries big risks, but those are lessened when monetary policy has room to provide a cushion. As it happens, the Reserve Bank of New Zealand did have room to provide a cushion and, in retrospect, the fiscal consolidation looks to have been a success. Chapter two was peer-reviewed by anonymous academic referees from around the globe and published in 2013 as a New Zealand Treasury Working Paper - Parameter Uncertainty and the Fiscal Multiplier.

Chapter three addressed a specific problem, too. As slack in the economy is used up, there is a boost to revenue and spending falls, relative to GDP. So how much fiscal effort is required to put the public finances on a stable footing depends crucially on estimates of the output gap. The trouble was that the British government depended on just two measures of slack and neither appeared credible at the time - both pointed to overheating when there was strong evidence of spare capacity in the labor market.

The third chapter developed a robust multivariate Kalman filter model of potential output to supplement other indicators of slack, quantified the margin of uncertainty in both real time and in retrospect and showed how output gap uncertainty carries over to estimates of the structural budget deficit.

Chapter three was peer-reviewed by the Office for Budget Responsibility's panel of external advisors and published in 2014 as an OBR Working Paper - Output Gap Measurement: Judgement & Uncertainty. The model developed in chapter three is still used today as an important input to the OBR's forecast process.

Chapter four, again, responded to an important and relevant policy question. When should governments junk their fiscal rules and support the economy instead? In 2015 the U.K. government offered a testable benchmark policy prescription - that fiscal policy can be relaxed when growth dips below 1%. Still, if GDP growth slows because potential output slows, this rule could let governments spend when they shouldn't. If however growth decelerates for cyclical reasons, maybe looser policy is the right course of action?

The fourth chapter tested whether GDP growth or the output gap works best as a trigger for fiscal relaxation - the former seems to be just as good if not better.

1.2 Thesis Overview

Chapter Two

The global financial crisis and fiscal responses to it created large and persistent structural deficits - those which persist once output has returned to potential - in many advanced economies. As economic recoveries began, governments had to decide how fast to withdraw stimulus and shrink underlying deficits. Too fast and the recovery would be choked off in its infancy. Too slow and there was seemingly a risk of punishment from creditors.

In normal times fiscal multipliers appear to be relatively small. And if tightening is signalled in advance, monetary policy can work to offset the drag on demand. But there are occasions on which that will not be possible. If, for example, monetary policy is less effective than usual. The fact is that policymakers do not know what the impact of fiscal measures will be at any given point in time. The models used to judge those impacts generally assume stable parameters. After all they are supposed to be 'deep' - determined by structural features of the economy and preferences that do not change over time.

To illustrate the risks, chapter two presents a small, estimated model of the New Zealand economy. To estimate the model I use Bayesian techniques. This allows for the inclusion of prior beliefs, drawn from the literature, about the model parameters but also brings to bear what the data have to say. I use a Kalman filter to estimate the likelihood function and the Metropolis-Hastings algorithm to generate draws from the posterior distribution.

I find that the mean estimates of the model parameters do not always lie in the centre of the distribution, which means the data are consistent with a skewed distribution. That's useful information because it gives an indication of where the risks to that estimate lie. To illustrate those risks I run fiscal policy simulations and vary the model parameters using the confidence intervals from the Bayesian estimation process.

What I find is that monetary policy is crucial, which chimes with the other literature on the topic. The parameters governing both the degree of monetary activism and the sensitivity of output to movements in interest rates have skewed distributions. In other words, consolidate the public finances at a time when monetary policy is unable to provide a cushion or at a time when households and businesses are less incentivised to spend by lower interest rates, and the impact could be very substantial.

The other key parameter is the degree of Ricardian equivalence - how much households and businesses tighten their belts when fiscal policy tightens. If the economy is depressed and households are unable to dip into savings, for example, fiscal tightening might have a bigger impact on growth.

Chapter Three

Governments need economic forecasts if they are to plan how much they are going to spend in years to come. To understand how big the economy will be and, therefore how much revenue will flow into the Treasury, three things are needed. An estimate of trend growth, an estimate of the output gap, and a forecast for how close demand will be to the economy's supply potential. In truth, all forecasts embody these features, either explicitly or implicitly.

Chapter three focuses on the second judgement - the margin of slack in the UK economy. As the unemployed return to work, more income taxes are paid and less is spent on social welfare. It's obvious that this is important to forecasts of the public finances. But there are other forms of slack, too. What if demand is so weak that employers shorten shifts? Or if some factory production lines are closed down until such a time as demand is sufficiently strong to merit switching them back on? All of this is slack, and it all matters to the public finances.

In chapter three I develop a measure of the output gap using a multivariate Kalman filter. It incorporates a wide range of information about slack, including inflation capacity utilisation and joblessness. And it leaves scope for the application of judgement about which of these indicators is most important or reliable. I show, however, that all estimates of the output gap are associated with trade-offs — those that tend to be revised by less may give a misleading signal more often, and those which are more flexible in their interpretation of new data tend to be revised a lot.

The conclusion is that a wide range of evidence should be considered in judging the outlook for the economy. It is also shown that output gap uncertainty carries over substantially to estimates of structural borrowing and that revisions to estimates have more to do with the arrival of new data than the refinement of earlier data.

Chapter Four

History shows that governments are pragmatic in times of crisis. The UK Treasury was very quick to act during the global financial crisis, abandoning its fiscal rules and administering stimulus to the economy. It was obviously the right thing to do faced with such a big shock. What if the next shock were to be smaller? Would the government be criticised for deviating from its fiscal rules when tightening in a downturn would make things worse? Or might a modest slowdown be exploited to help the government pump prime the economy before an election.

In 2015, the UK government cleared up what it defined as significant shock, worthy of delaying fiscal tightening: growth of less than 1%. It no longer uses that benchmark, but the question remains - is it a good one?

Chapter four uses a solution method for models with occasionally-binding constraints and applies that to a model with a state-contingent fiscal rule. It is shown that downturns are shorter when fiscal policy is relaxed – perhaps an obvious conclusion. Still, it is also shown that GDP, which is more quickly available and less prone to revision, is probably a better trigger for fiscal flexibility than the output gap.

Contributions of the thesis

These are threefold:

- Fiscal tightening can be costly, but the risks are vastly reduced if monetary policy can act as a cushion to demand;
- Output gap uncertainty is pervasive in the U.K. and it's best to draw upon a range of information to inform estimates of slack and not to reject any simply due to revision properties; and
- A growth trigger to permit fiscal flexibility can help shorten downturns and looks at least as good if not better for identifying major shocks than are estimates of the output gap.

Structure of the thesis

The material in this thesis is presented in five chapters including the introductory chapter.

Chapter two presents an estimated model of the New Zealand economy. Simulations are run using the confidence intervals for each estimated parameter to illustrate the risks surrounding fiscal consolidations. Chapter three develops a suite of estimates of the output gap, culminating in a multivariate Kalman filter for the U.K. The revisions properties of each are assessed and it is shown how output gap uncertainty carries over to estimates of the structural deficit. Chapter four sets out a regime-switching model of the U.K. economy with a growth knockout to determine when fiscal policy may be relaxed. It is shown that a growth trigger shortens downturns and looks at least as good an indicator of significant shocks as the output gap. Chapter five concludes.

2

Parameter uncertainty and the fiscal multiplier in a small open economy

2.1 Literature Review

The global financial crisis, which began in 2008, led to a significant deterioration in the fiscal positions of many governments, with a number of advanced economies running substantial budget deficits in the years that followed. Because much of the loss of output associated with the financial crisis is judged to be permanent, this has led to governments running persistent structural deficits – those expected to remain once the economic cycle has run its course and output has returned to its steady-state growth path.

Many governments have responded to the deterioration of their fiscal positions by planning large consolidations – usually a mix of spending cuts and tax increases, with most balanced towards the former. A natural question to ask is to what extent might these plans reduce aggregate demand in the economy and, in doing so, slow its cyclical recovery? Besides explaining the origins of the financial crisis and the implications for policy settings, answering this question has become one of the major focuses of macroeconomists in recent years.

Estimates of the size of the fiscal impact multiplier range widely, as do the techniques used to assess them. Estimation methodologies tend to fall into two categories: the structural vector autoregression (SVAR) approach, pioneered by Blanchard & Perotti (2002), and dynamic stochastic general equilibrium (DSGE) modelling, as recently applied by Davig & Leeper (2011). The former approach draws inferences from statistical relationships identified in the data. To reveal the underlying relationships, a number of assumptions about the way the economy functions are applied during the estimation process. The DSGE approach involves the specification of a model, derived from economic theory, and the calibration of that model's parameters either via estimation or through the application of judgement. The size of the impact multiplier is then derived from the simulation properties of the model.

During the financial crisis, the IMF Lall *et al.* (2009) published estimates of the size of fiscal impact multipliers for a number of advanced economies, which averaged around 0.5. Using a DSGE approach, Mountford & Uhlig (2009), also find the multiplier to be around 0.5. While the original SVAR estimate of Blanchard & Perotti (2002) is consistent with an impact multiplier of around unity. Another approach, recently applied by Blanchard & Leigh (2013), has been to decompose forecast errors made during periods of fiscal retrenchment into the part related to exogenous shocks and the part related to the assumed fiscal impact multiplier. The estimate associated with this method is consistent with a fiscal impact multiplier of around unity.

Ilzetzki *et al.* (2013) apply the SVAR methodology using a large data set which includes a number of economies with different characteristics. They find that the multiplier depends critically on the degree of development, the monetary policy framework and the degree of openness. Crucially, their estimate of the multiplier is not significantly different from zero for countries with a flexible exchange rate and they find the multiplier is smaller for more open economies.

Corsetti *et al.* (2012) also find that the monetary policy and exchange rate regime are important in determining the effect of fiscal policy. But the exchange rate is found to appreciate in response to a positive government spending shock. In most models, the exchange rate plays a stabilising role by boosting output at times of fiscal tightening by bringing about a fall in the relative price of domestically-produced goods. The finding calls into question the assumed transmission mechanism and role of the exchange rate. New Zealand is a small open economy with a flexible exchange rate. Taking an SVAR approach, Parkyn & Vehbi (2014) find a statistically significant impact multiplier of 0.3 associated with a change in government spending, rising to 0.6 when debt dynamics are excluded.

A recent study, Ramey (2019), asking what has been learned from the last ten years of fiscal research compares three different and widely applied methods of multiplier estimation. It is found that that the results vary widely. In other words, model uncertainty is pervasive. But correcting for some known shortcomings that have emerged in the literature in recent years suggests a band for expenditure multipliers of between 0.6 and unity.

Still, that finding is by no means universally accepted.

One conclusion that can therefore safely be drawn from the expansive literature on the subject is that the size of fiscal multipliers is extremely uncertain. And as Castelnuovo & Lim (2019) put it, in a recent and exhaustive survey of the literature, the consideration of state-dependent multipliers has become the norm. The bottom line remains that policy makers need to have some view about the likely effects of discretionary fiscal policy and what the risks surrounding it are. With this in mind, I ask ‘Under what conditions might the impact of a fiscal tightening be bigger or smaller?’

To answer this question I estimate a small, reduced-form model of the New Zealand economy, using Bayesian methods. I then conduct fiscal policy simulations by varying a number of the key model parameters and assess the output effects using two metrics. The first is the fiscal impact multiplier, which represents the degree to which a fiscal consolidation might slow GDP growth and widen the output gap.[1] But to explore the broader effect on social welfare, I also consider the cumulative output loss associated with a fiscal tightening. This takes into account both the degree to which a consolidation might reduce output and the time it takes to return to its steady state growth path. The intention is to give a quantified estimate of the risks associated with fiscal consolidations based on the degree of uncertainty about the way the New Zealand economy functions. The remainder of this paper is structured as follows. I discuss the choice of modelling methodology in Section 2.2, before explaining the theoretical underpinnings governing the dynamics of the model. Section 2.8 is concerned with the estimation of the model, including the choice of priors. Section 2.10 sets out the key findings of the fiscal consolidation simulations before Section 2.13 assesses the implications of the results for policy making. Section 2.14 concludes.

2.2 Modelling methodology

There are a number of ways in which to develop a model of the economy, the suitability of which depends upon its intended use. One type of modelling method is the DSGE approach, referred to above. Such models, often used by central banks, are typically quite large and strictly adhere to the prescriptions of their microeconomic foundations. That is to say that, whether it’s a model with three equations or twenty, the laws of motion of the economy are governed by the optimising behaviour of agents operating within it. DSGE models tend to fall into two categories – real business cycle models, which treat all deviations from steady state as optimal responses to shocks, and, so-called, New

[1] In this paper the fiscal impact multiplier is defined as the change in the output gap over a period of one year associated with a 1 per cent of potential GDP fiscal tightening.

Keynesian models, which attribute some of the deviation to nominal rigidities. The model presented here can be considered a reduced-form version of the latter.

The strength of micro-founded models is that the equations are based on optimising behaviour and so should be robust to changes in policy – the dynamics of the model are driven by 'deep' or structural parameters.[1] However, models featuring forward-looking equations tend to underperform simple autoregressive models and, quite often, little empirical support is found for the hypothesised underlying relationships.[2] This is partly because many of the variables move with a degree of inertia that is inconsistent with the adjustment paths implied by forward-looking, rational expectations models. In these models, it is the rational but immediate adjustment of households' expectations to innovations which implies jump responses, which are rarely seen in the data. To overcome this problem, many DSGE models feature adjustment costs and other mechanisms introduced with the intention of replicating the inertial responses of the data.

All economic models are misspecified, since they represent a simplification of reality. And while the model presented in this paper is deliberately small, with the objective of parsimony and tractability in mind, it is also likely to be particularly prone to misspecification. Recognising this, I attach structural interpretations only loosely to the parameters of the reduced-form model, since many of them capture broader influences on the variables to which they pertain.

The four key variables of interest are the output gap, bank rate, the inflation rate and the real exchange rate; given by the investment-saving (IS), Taylor, Phillips and real uncovered interest parity (RUIP) relations respectively. In what follows, I present the baseline functional forms adopted for the core equations and identify where the assumptions are consistent with the microeconomic theory upon which they are founded and how that translates into the functional form of the reduced-form model. For reference, a complete set of the equations that constitute the model can be found at the end of this section and descriptions of the data used to estimate are presented in the annex.

2.3 IS relation

The IS equation relates output in the economy to deviations of the real interest rate from the level consistent with stable output and inflation in the medium term. Equations of this form are a staple of macroeconomic modelling and appear, in some form, in all New-Keynesian models.

[1]As opposed to being based on statistical relationships identified in the data, which may not be stable over time

[2]See Fuhrer & Rudebusch (2004).

The standard forward-looking consumption IS relation is given by,

$$c_t = \alpha_c c_{t+1_{|t}} - \alpha_r + \epsilon_t^c \tag{2.1}$$

Where

$$r_t = i_t - \pi_{t+1_{|t}} \tag{2.2}$$

Equation 2.1 represents the baseline consumption Euler equation that arises from the representative household's optimisation problem. It has been log-linearised around its steady state so c_t represents the deviation of consumption from its steady-state growth path, $c_{t+1_{|t}}$ is the expected deviation of consumption from its steady state, conditioned on information available at time t, r_t is the real interest rate gap and ϵ_t^c is an independent, identically-distributed consumption shock. The nominal interest rate and expected rate of inflation are given by i_t and $\pi_{t+1_{|t}}$, respectively.

The consumption Euler equation simply states that, in equilibrium, the representative household is unable to increase its utility by shifting consumption between periods — that is, the marginal utility of consumption today is balanced with the discounted marginal utility of consumption tomorrow.

Such an equation implies the immediate adjustment of output as households update their expectations. In practice, consumption appears to react quite slowly to changes in interest rates, for example, and a number of studies attempt to explain this behaviour. One such endeavour is the habit formation model of Fuhrer (2000). Fuhrer postulates that the utility derived from consumption depends both on the absolute level of consumption and the level of current consumption relative to past consumption - that households do not like consuming less than they have been and initially resist changes, before eventually adjusting. This modification was shown to substantially improve the fit of the model.[1]

Other work, predominantly concerned with why the behaviour of consumption appears to invalidate the permanent income hypothesis, such as Muellbauer (1988), suggests that households may be myopic in their consumption choices. Campbell & Mankiw (1989) offer the hypothesis that households do not have the resources to engage in producing full forecasts and so it is optimal for them to use a rule of thumb when updating their consumption plans in response to income shocks.

That lagged output improves the fit with the data is important, but whether one accepts the habit formation story, the rule of thumb hypothesis or simply assumes that households are less forward-looking than is often suggested, is less important for the specification of the IS relation. In empirical work, an assumption of habit formation or myopia in

[1] See, for example, Giannoni & Woodford (2004) for a formal derivation of the habit formation-augmented NKIS relation

household consumption choices is not uncommon and both Batini & Haldane (1999) and Smets & Wouters (2003) allow for it in their respective models of the UK and the euro area economies. Indeed, neither Carlin & Soskice (2010) nor Ryan *et al.* (2009) include expected output in their baseline IS relations for the UK and New Zealand respectively.

With this in mind, I introduce persistence to the output gap process by assuming a degree of external habit formation in consumption (given by α_c) — while acknowledging that the true source could be myopia, rule of thumb behaviour or a failure of rational expectations, resulting in an equation of the form,

$$c_t = \alpha_c c_{t-1} + (1-\alpha_c)c_{t+1_{|t}} - \alpha_r r_t + \epsilon_t^c \tag{2.3}$$

To capture the effect of discretionary fiscal policy on the economy I allow for the possibility of non-Ricardian behaviour, in a similar way to Ratto *et al.* (2007). Ricardian behaviour states that, faced with a reduction in taxes, for example, households will tend to save the associated additional income since they know it heralds higher taxation or lower spending in the future. The effect on permanent incomes is zero and, therefore, so is the output response. In this model I allow for non-Ricardian behaviour by specifying the proportion of households who are affected by discretionary fiscal policy.

The consumption of Ricardian households does not respond to changes in public spending and taxation so changes in the fiscal balance do not feature in the consumption equation and this is given by the standard IS relation presented in 2.4, but Ricardian consumption is denoted by a superscript R.

$$c_t^R = \alpha_c c_{t-1}^R + (1-\alpha_c)c_{t+1_{|t}}^R - \alpha_r r_t + \epsilon_t^c \tag{2.4}$$

Non-Ricardian households spend all the additional income/reduce consumption by the full extent of the fiscal tightening and so the change in the fiscal stance is introduced to the consumption equation, where represents the fiscal impulse - a similar measure to the change in the cyclically-adjusted budget balance.[1],

$$c_t^{NR} = \alpha_c c_{t-1}^{NR} + (1-\alpha_c)c_{t+1_{|t}}^{NR} - \alpha_r r_t + f_t + \epsilon_t^c \tag{2.5}$$

The proportion of non-Ricardian households is indexed by the parameter, α_{NR}, giving rise to the aggregate consumption equation

$$\begin{aligned} c_t = {} & (1-\alpha_{NR})[\alpha_c c_{t-1}^R + (1-\alpha_c)c_{t+1_{|t}}^R - \alpha_r r_t + \epsilon_t^{Rc}] \\ & +\alpha_{NR}[\alpha_c c_{t-1}^{NR} - (1-\alpha_c)c_{t+1_{|t}}^{NR} - \alpha_r r_t + f_t + \epsilon_t^{NRc}] \end{aligned} \tag{2.6}$$

[1] This is defined as the change cyclically-adjusted budget balance plus the change in the level of capital expenditure as a share of GDP. A full methodology for the construction of the data set can be found in Philip & Janssen (2002) Broadly, this measure is intended to capture the change in the fiscal position arising from discretionary policy measures.

Separating out the term capturing the change in the fiscal stance,

$$
\begin{aligned}
c_t = (1-\alpha_{NR})[\alpha_c c^R_{t-1} + (1-\alpha_c)c^R_{t+1_{|t}} - \alpha_r r_t + \epsilon^{Rc}_t] \\
+\alpha_{NR}[\alpha_c c^{NR}_{t-1} - (1-\alpha_c)c^{NR}_{t+1_{|t}} - \alpha_r r_t + +\epsilon^{NRc}_t] + \alpha_{NR} f_t
\end{aligned}
\tag{2.7}
$$

and then aggregating the remaining Ricardian and Non-Ricardian terms gives:

$$
c_t = \alpha_c c_{t-1} + (1-\alpha_c)c_{t+1_{|t}} - \alpha_r r_t + \alpha_{NR} f_t + \epsilon^c_t \tag{2.8}
$$

Based on its structural interpretation, the indexing coefficient, α_{NR}, should be bounded by 0 and unity. But there how much output changes in the short run for a given discretionary policy measure may depend on the precise policy package. There are several ways to bring about a structural adjustment in the public finances. These include revenue measures, such as consumption or income tax changes, and spending measures, such as changes in departmental expenditure or welfare policies. In practice, each of these measures is likely to be associated with a different multiplier, since they tend to affect different groups in society, for example. In this sense, households, in aggregate, might be less 'Ricardian' in their response to some measures than to others.

A comprehensive analysis would estimate different α_{NR} parameters for different types of policy measure. This would not be practical here, since I focus only on New Zealand and the time series with which I am working are relatively short. There simply is not enough variation in the series to provide reliable estimates at a granular level. Instead, I focus on the overall (average) fiscal impact multiplier and use changes in the government's cyclically-adjusted budget balance to estimate its size. This is consistent with the estimated parameter relating to the direct effects of a fiscal policy package of average composition.

Furthermore, because the effects of fiscal policy on the net trade position are not explicitly articulated in the model, the effects of such 'leakages' are reflected in the estimate of the parameter α_{NR}. In what follows, I describe this coefficient as the degree of non-Ricardian behaviour but readers should be aware that, due to the reduced-form nature of the model, this parameter captures more than this structural parameter alone — another way to think of it may be as the direct fiscal impact multiplier, before any offset from monetary policy, for example.

In this model, the fiscal policy stance is determined exogenously and follows an autoregressive process,

$$
f_t = \zeta_f f_{t-1} + \epsilon^f_t. \tag{2.9}
$$

This specification is intended to account for the autocorrelation introduced by interpolating annual fiscal data. To get from the consumption Euler equation to the IS equation I assume that the behaviour of the consumer can explain whole-economy behaviour. This is a common assumption in small models of the economy but is not completely satisfactory given, in particular, the contribution of business and inventory investment to the cyclical volatility of output.

Without deriving the behaviour of firms explicitly from microeconomic foundations here, it suffices to say that the change in output associated with firms' responses to changes in real interest rates is in the same direction as that implied by the response of households. Intuitively, if the real rate of interest falls, this lowers the cost of borrowing and increases the overall rate of return of an investment project. Therefore, any profit-maximising firm has a greater incentive to invest.[1]

There are a number of extensions to these simple theories, which highlight the role of uncertainty and irreversible costs in the investment decision — see Leahy & Whited (1995) and Pindyck & Solimano (1993), for example. Like habit formation in consumption, these extensions serve to increase the persistence of the model. While these theories are not articulated within the modelling framework here, the cyclicality of business investment and its contribution to output volatility should already be captured by the reduced-form parameters of the IS relation.

Aggregating the consumption Euler equation to the whole economy level gives equation 2.10. In the spirit of Gali & Monacelli (2005), I also include a term for changes in the trade-weighted real effective exchange rate, which is intended to capture the effect on output of changes in relative prices which serve to shift the allocation of resources to and from the export-facing sector,

$$y_t = \alpha_y y_{t-1} + (1-\alpha_y) y_{t+1_{|t}} - \alpha_r r_{t-1} - \alpha_q \Delta q_{t-1/t-4} + \alpha_{NR} f_t + \epsilon_t^y.\text{[2]} \tag{2.10}$$

y_t is the output gap, $y_{t+1_{|t}}$ is the expected output gap at time t and r_{t-1} is the real interest rate gap, Δq_{t-1} is the change in the lagged real expected exchange rate gap and ϵ_t^y is an independent and identically-distributed aggregate demand shock. I include four lags of the change in the real exchange rate to allow output to respond slowly to changes in relative prices.

[1]Tobin's q theory of the investment decision, Tobin (1969), operates in a similar way. Lower expected interest rates decrease the rate at which income streams are discounted, increasing the valuation of companies' net assets. When the market value of assets exceeds the book value, there is a profit opportunity and companies expand their investment until such a time that book prices are equal to market prices.

2.4 Phillips curve

The New Keynesian Phillips Curve (NKPC) relates current inflation to expectations of future inflation and marginal cost pressures. That the inflation process is forward-looking follows from the price-setting behaviour of firms, which is assumed to follow Calvo (1983). The basic premise is that in each period a firm has a fixed probability that it will keep its prices unchanged, so firms set prices now with a view to the future because they know that they may not be able to change their prices in the subsequent period.[1] The probability of changing/not changing price each period is independent of the time elapsed since the firm last changed its price, and this attribute simplifies the aggregation of individual firm behaviour to the whole-economy level. This gives an equation of the form,

$$\pi_t = \beta_\pi \pi_{t+1_{|t}} + \beta_y y_t + \epsilon_t^\pi \tag{2.11}$$

where π_t is the rate of inflation and $\pi_{t+1_{|t}}$is the expectation of inflation conditioned on information available at the current time.

I assume that real marginal cost pressures drive the inflation process, consistent with Clarida *et al.* (1999) and that these cost pressures are well-represented by the output gap, y_t. There are other measures which could be used - Batini *et al.* (2005) use the labour share of income in their estimate of the Phillips curve, which has the advantage of being directly observable.[2] But using the labour share for forecasting with this model would not be possible because it does not capture the evolution of the labour market, so the output gap is preferred. The error term, ϵ_t^π, is an independent, identically-distributed inflation shock.

As with the IS relation, the purely forward-looking version of this equation fits the data poorly — failing to capture the observed inertia of inflation. The equation specification implies that persistence in either movements in the output gap or changes in inflation expectations could produce an inertial path for inflation, but leaves open the possibility of large jumps. It also implies that inflation should lead the output gap, which is the opposite of what we observe in the data; both empirical evidence and conventional wisdom suggests that monetary policy affects inflation only with a lag, rather than instantaneously.[3]

A model that does not adequately capture the persistence of inflation would not fit the data and have misleading simulation properties. Therefore, in what follows, I relax

[1] Note that this probability is independent of the general level of inflation. This seems unlikely, and has implications for the model, such as the potential non-neutrality of money.

[2] It is also the case that, under certain assumptions, the labour share (the average product of labour) is proportional to real marginal cost in an economy characterised by a Cobb-Douglas production function.

[3] See Rudd & Whelan (2007) for a detailed discussion of this point.

the restrictive assumption that households and firms are completely forward-looking and anchor expectations of inflation to the middle of the Reserve Bank's target range.

The hybrid version of the New Keynesian Phillips Curve, used in a number of empirical estimates of the equation Galı & Gertler (1999) modifies the standard NKPC formulation by allowing a proportion of firms to use a rule of thumb when setting prices, consistent with a degree of indexation in price setting. This modification provides a theoretical justification for the presence of an inflation lag in the first order condition of the NKPC. Intuitively, the inclusion of lags of inflation serves to act as a proxy for the rational expectation of future values of the driving variable. The resulting equation therefore includes a backward-looking term and a coefficient, β_π, that determines the weight placed on past inflation relative to inflation expectations in the inflation process,

$$\pi_t = \beta_\pi \pi_{t-j} + (1 - \beta_\pi)\pi_{t+j_{|t}} + \beta_y y_{t-j} + \epsilon_t^\pi \tag{2.12}$$

The restriction placed on the inflation coefficients summing to unity (effectively imposing a discount factor of one) means that money is super-neutral in this model. It also implies that the coefficient, β_π, can be interpreted directly as the proportion of firms in the economy that set prices in a backward/forward looking manner.

In this paper I take a slightly different approach to the Gali & Gertler set-up and adopt the prior expectation that agents in the economy expect that monetary policy is able to return inflation to target at some time horizon (typically assumed to be around two years). Because the target is a constant, expectations drop from the equation.

$$\pi_t = \beta_\pi \pi_{t-j} + \beta_y y_{t-j} - \beta_q \Delta q_{t-j} + \epsilon_t^\pi \tag{2.13}$$

To allow for the effect of exchange rate pass-through to prices, I include lags of the change in the real trade-weighted exchange rate.[1] That the Phillips curve can be augmented in this way is demonstrated formally in Batini *et al.* (2005) and Gali & Monacelli (2005).

Finally, linearising around the target rate of inflation gives the Phillips curve in 'inflation gap' terms.

$$\pi_t = \beta_\pi \pi_{t-1} + \beta_y y_t - \beta_q \Delta q_{t-1/t-4} + \epsilon_t^\pi \tag{2.14}$$

2.5 The real exchange rate

In specifying the dynamics of the real exchange rate, I begin by setting its medium-term anchor. In the long run, the nominal exchange rate is thought to move in such a way that

[1] It is assumed that the real exchange rate is proportionate to the terms of trade — i.e. that the elasticity of substitution between domestically-produced and foreign goods is equal to unity.

prices between two countries are equalised. That is to say that capital will flow between countries such that relative unit labour costs are equalised and that the relative demand for currency acts to push the nominal exchange rate so that price differentials gradually erode, albeit with a wedge arising from transport costs. Such a medium-term relationship is known as relative purchasing power parity (PPP) where, the nominal exchange rate is given by the ratio of domestic to foreign prices and a permanent wedge. Linearising around this steady state and taking logs gives the long-run PPP steady state condition,

$$e_t = p_t^* - p_t \tag{2.15}$$

The short-run dynamics of the nominal exchange rate, e_t, are given by the uncovered interest rate parity condition (UIP),

$$e_t = e_{t+1_{|t}} + i_t - i_t^f \tag{2.16}$$

where the nominal exchange rate gap is given by the expected nominal exchange rate one period ahead and the relative interest rate between New Zealand and a foreign country.

Substituting in the real exchange rate identity, $q_t = e_t + p_t - p_t^*$, gives the real-UIP (RUIP) condition,

$$q_t - (p_t^f - p_t) = q_{t+1_{|t}} - (p_{t+1_{|t}} - p_{t+1_{|t}}^f) + i_t - i_t^f. \tag{2.17}$$

And solving for the real exchange rate gives

$$q_t = q_{t+1_{|t}} + (i_t - \pi_{t+1_{|t}}) - (i_t^f - \pi_{t+1_{|t}}^f) + \epsilon_t^{qs} \tag{2.18}$$

Gali & Monacelli (2005) include a similar equation in their open-economy model and convergence with a steady state is achieved by iterating forward, such that the expectations term drops out of the equation. Because they linearise around long-run PPP this is consistent with agents in the economy expecting long-run PPP to hold in subsequent periods. In practise, deviations of the observed real exchange rate from that consistent with long-run PPP can persist for long periods of time and it is thought that convergence with PPP is slow. To allow for this empirical observation I include a convergence parameter, $kapp_q$, which weakens the pull from the steady state in the short term. Compared with the G&M model, this specification increases the responsiveness of the exchange rate to interest and inflation shocks, consistent with the high degree of volatility associated with the New Zealand Dollar.

Taking the RUIP equation, linearising around long-run PPP and introducing a convergence term gives the equation,

$$q_t = q_{t+1_{|t}} + (i_t - \pi_{t+1_{|t}}) - (i_t^f - \pi_{t+1_{|t}}^f) - \psi_q q_{t-1_{|t}} + \epsilon_t^q \tag{2.19}$$

The equation is based on the notion that, if a real interest rate differential exists, the real rate of return on domestic and foreign assets is equalised by movements in the exchange rate. The assumption of convergence with long-run PPP is consistent with more sophisticated models. For example, in macro-balance models of the exchange rate, short-run dynamics are typically governed by some version of real or nominal uncovered interest parity. But, in the medium term, the real exchange rate moves to stabilise a country's net international investment position.[1]

In this model, foreign interest rates follow an autoregressive process and, in steady-state, are equal to the steady-state domestic nominal interest rate,

$$i_t^f = \zeta_{if} i_{t-1}^f + \epsilon_t^{if} \tag{2.20}$$

For the purposes of including changes in the real exchange rate in the IS relation it is necessary to have a forecast of foreign inflation. This too follows an autoregressive process and is assumed to have a steady-state rate consistent with the domestic inflation target, a similar assumption to that made by Carlin & Soskice (2010),

$$\pi_t^f = \zeta_\pi \pi_t^f + \epsilon_t^{\pi f} \tag{2.21}$$

Given the share of primary goods in New Zealand's exports, it is perhaps unsurprising that commodity price changes influence the exchange rate. To allow for the effect of persistent commodity price changes on the exchange rate, I introduce persistence to the shock term,

$$\epsilon_t^q = \psi_\epsilon \epsilon_{t-1}^q + \epsilon_t^{q2}. \tag{2.22}$$

2.6 Central bank reaction function

Taylor (1993) observed that the conduct of monetary policy can be well-captured by a simple rule relating interest rates to inflation and the output gap. Following Taylor's paper there began a concerted academic effort to assess this class of policy rules and their implications for optimal monetary policy. However, some form of Taylor's original rule, which is entirely backward-looking, remains the default specification for the behaviour of the central bank in many economic models.

The IS and Phillips relations described above operate with a lag. That is to say, it takes time for interest rates to affect the output gap and, in turn, for inflation to respond

[1]Larger-scale DSGE models, such as that of Harrison & Oomen (2010) ensure that the real exchange rate converges with its steady-state value by applying a risk premium to net foreign asset holdings in the UIP equation. Such an assumption is also consistent with relative PPP holding in the medium-term.

to the output gap. The lag structure embodied in these equations means that monetary policy should be conducted with a view to the future. Therefore, given the involvement of the central bank in forecasting the economy and the lags associated with the conduct of policy, I specify a forward-looking form of the Taylor rule which is consistent with the other equations in the model — the Bank's expectations are assumed to be model-consistent.

As well as being a reasonable empirical description of the conduct of monetary policy, Svensson *et al.* (1997) and others have shown that the Taylor class of rules can also be derived from the inflation targeting central bank's optimisation problem. Simply allowing for the lag structure associated with the monetary transmission mechanism gives a forward-looking Taylor rule of the form,

$$i_t = \bar{i}_{t-j} + \delta_y y_{t+j_{|t}} + \delta_\pi(\pi_{t+k_{|t}} - \pi^*), \tag{2.23}$$

where i_t is Bank Rate, $\bar{i}_{t-j}$ is the equilibrium nominal rate of interest,$y_{t+j_{|t}}$ is the output gap forecast at the relevant time horizon and $\pi_{t+k_{|t}}$ is the forecast deviation of inflation from target.[1,2]

Unlike the IS and Phillips relations, I do not include an exchange rate term in the specification of the Taylor rule. In this model, the central bank responds to movements in the exchange rate only indirectly, via its effect on output and domestically-generated inflation. This is consistent with both New Zealand's Policy Targets Agreement and the Taylor (2001) finding that the inclusion of exchange rates does little to improve the stabilisation of output and inflation and is possibly detrimental.

A substantial literature also exists on the observed inertia of interest rate setting by central banks around the world, see for example Goodfriend (1991). In what follows, I adopt the same approach as Galı & Gertler (1999)), which is to assume the presence of a policy rate smoothing parameter in the central bank's reaction function. They suggest this smoothing arises from a desire to avoid the credibility costs associated with large policy reversals, a desire to minimise disruption to capital markets and the time it takes build a consensus to support a policy change.[3]

Later discussions have identified ways in which interest rate smoothing might be optimal for a central bank in the presence of parameter uncertainty. Svensson (1999), for

[1]The specification is slightly different from the original Taylor rule but consistent with Nelson & Nikolov (2004).

[2]I use effective bank rate in place of actual bank rate to account for the effects of credit spreads and unconventional monetary policy on lending rates to the wider economy.

[3]In a rational expectations context, Woodford (2000) also shows that it can be optimal for a central bank to move the current policy rate less in response to demand and inflation shocks if, at the same time, the changes are characterised by a high level of persistence. This way, agents in the economy expect interest rates to be lower for longer once they have been cut, in turn lowering longer-term interest rates as well as short rates.

example, shows that parameter uncertainty for an inflation-targeting central bank dampens the policy response, confirming what Brainard (1967) first described. Söderström (2002) extends this analysis to a dual-mandate central bank with output in its loss function. He finds that uncertainty over inflation dynamics tends to heighten the response to inflation deviations (in case expectations become unanchored) but uncertainty over output dynamics encourages caution.

Regardless of the precise motive, the inclusion of central banks' smoothing of policy rates in their reaction functions significantly improves the fit with the data. Equation 2.24 captures interest rate inertia as in Clarida *et al.* (1999),

$$i_t = (1-\delta_i)i_t^* + \delta_i i_{t-1} \tag{2.24}$$

where i_t is the interest rate set, ψ is the smoothing parameter, i_t^* is the interest rate implied by the reaction function (absent smoothing) and i_{t-1} is the interest rate set in the preceding period.

Substituting the generalised Taylor rule in to equation 2.21 as the i_t^* term gives the central bank reaction function with policy rate smoothing equation 2.25,

$$i_t = (1-\delta_i)\bar{i}_t + (1-\delta_i)\delta_y y_{t+j_{|t}} + (1-\delta_i)\delta_\pi(\pi_{t+j_{|t}} - \pi^*) + \delta_i i_{t-1} + \epsilon_t^i \tag{2.25}$$

I choose a forecast horizon of 6 months for the output gap and a year and a half for inflation, consistent with conventional wisdom over the transmission mechanism of monetary policy. Linearising equation 2.25 around the steady-state interest rate and inflation target gives the 'nominal interest rate gap' equation,

$$i_t = \delta_i i_{t-1} + (1-\delta_i)\delta_y y_{t+2_{|t}} + (1-\delta_i)\delta_\pi(\pi_{t+6_{|t}}) + \epsilon_t^i \tag{2.26}$$

2.7 Credit spreads

Since the onset of the recent financial crisis, a rise in the perceived degree of risk associated with lending and borrowing has significantly widened the gap between the interest rate set by the Reserve Bank of New Zealand and the price of credit available to the wider economy. As a result, monetary policy has subsequently taken serious account of the effect of credit spreads on the behaviour of agents in the economy.

While the existence of a credit spread is not important to the running of fiscal policy simulations (since there is no hypothesised relationship between credit spreads experienced by the wider economy and discretionary fiscal policy) it is important to the estimation of the model. To exclude the effect of higher credit spreads would miss important information relevant to the monetary policy decision and the real interest rate faced by households and businesses.

The inclusion of credit spreads in the model presented here is based on a simple principle: the Reserve Bank is ultimately concerned with the interest rates paid by household and firms in the economy. So if the spread of interest rates experienced by agents in the wider economy over policy rates is higher than usual, this implies that the Bank would set policy rates lower than usual. Therefore, rather than targeting policy rates, in this model, the Bank takes credit spreads into account directly and targets an adjusted policy rate, described here as Effective Bank Rate, i_t^e.

By extending the baseline New-Keynesian model of the economy to include a measure of credit spreads, Cúrdia & Woodford (2016) show that agents in the economy respond in a similar fashion to increases in borrowing rates arising from changes in the default risk premium as they would to an increase in Bank Rate.[1] Importantly, the C&W model shows that, so long as central bankers take credit spreads into account, the Taylor class of policy rules remains optimal in choosing the stance of monetary policy.

To construct a measure of the credit spread, I use a selection of quoted household borrowing and deposit rates and subtract from those the relevant reference rate of interest.[2] For example, I take the average interest rate quoted for a 2-year fixed-rate mortgage and subtract from this the two-year government bond rate. This gives the spread over expected policy rates at the relevant time horizon.[3]

In this paper, the model is presented in terms of deviations around a steady-state. Therefore, the credit spread series should also be expressed in terms of deviations around a steady-state. For simplicity, I assume that the steady-state credit spread is stationary around its long-run average value.

The evolution of the credit spread is given by

$$cs_t = \theta_{cs}(cs_{t-1}) + \epsilon_t^{cs} \tag{2.27}$$

where is assumed to follow an autoregressive process that reverts to an equilibrium mean value of zero.[4]

The effective interest rate is defined as,

$$i_t^e = i_t + \tau_{cs} cs_t \tag{2.28}$$

[1]The authors create a model which assumes that banks are able to finance themselves by issuing deposits which must attract the same rate of interest as government bonds of the same maturity to avoid arbitrage opportunities. In this paper I assume that the relevant spread is over the cost of borrowing, as set by the central bank. This approach is motivated by the observation that the Bank targets a policy rate defined in terms of very short-term government borrowing rates.

[2]Ideally, a measure of credit spreads would also include corporate sector borrowing and deposit rates, but there is little data available with which to construct such a measure.

[3]Insofar as the two-year government bond rate is a good proxy for expectations of policy rates.

[4]i.e. it is exogenous, as in the Curdia-Woodford model.

where the coefficient, τ_{cs}, allows for the possibility of a rise in credit spreads affecting consumer behaviour more or less than a corresponding move in bank rate.

Credit spreads enter the model in two places: the IS relation and the Taylor rule. The real interest rate gap, which features in the IS relation, becomes,

$$r_t = i_t^e - \pi_{t+1_{|t}} \tag{2.29}$$

and Taylor rule becomes,

$$i_t = \delta_i i_{t-1} + (1-\delta_i)\delta_y y_{t+2_{|t}}(1-\delta_i)\delta_\pi(\pi_{t+6_{|t}}) + (1-\delta_i)\delta_{cs} cs_t + \epsilon_t^i \tag{2.30}$$

The Taylor rule is augmented to allow the central bank to respond to deviations of the credit spread from its steady state. When the coefficient, δ_{cs}, is equal to unity, the Bank treats the increase in the credit spread as equivalent to an increase in bank rate. Values not equal to unity allow for a partial or excess response of policy rates to credit spreads. Since the bank's expectations are model consistent, the effect of the credits spread on the economy and the bank's policy response are constrained to be consistent with one another,

$$\delta_{cs} = \tau_{cs}. \tag{2.31}$$

2.7.1 Model equations

All equations are log-linearised around a steady state — i.e. output gap, inflation deviation from target...

$$y_t = \alpha_y y_{t-1} + (1-\alpha_y) y_{t+1_{|t}} - \alpha_r r_{t-1} - \alpha_q \Delta q_{t-1/t-4} + \alpha_{NR} f_t + \epsilon_t^y \tag{2.32}$$

$$\pi_t = \beta_\pi \pi_{t-1} + \beta_y y_t - \beta_q \Delta q_{t-1/t-4} + \epsilon_t^\pi \tag{2.33}$$

$$i_t = \delta_i i_{t-1} + (1-\delta_i)\delta_y y_{t+2_{|t}} + (1-\delta_i)\delta_\pi(\pi_{t+6_{|t}}) + \epsilon_t^i \tag{2.34}$$

$$q_t = q_{t+1_{|t}} + (i_t - \pi_{t+1_{|t}}) - (i_t^f - \pi_{t+1_{|t}}^f) - \psi_q + \epsilon_t^q \tag{2.35}$$

$$cs_t = \theta_{cs}(cs_{t-1}) + \epsilon_t^{cs} \tag{2.36}$$

$$i_t^f = \zeta_{if} i_{t-1}^f + \epsilon_t^{if} \tag{2.37}$$

$$\pi_t^f = \zeta_\pi \pi_t^f + \epsilon_t^{\pi f} \tag{2.38}$$

$$f_t = \zeta_f f_{t-1} + \epsilon_t^f \tag{2.39}$$

$$\delta_{cs} = \tau_{cs}. \tag{2.40}$$

$$i_t^e = i_t + \tau_{cs} cs_t \tag{2.41}$$

$$i_t = \delta_i i_{t-1} + (1-\delta_i)\delta_y y_{t+2_{|t}}(1-\delta_i)\delta_\pi(\pi_{t+6_{|t}}) + (1-\delta_i)\delta_{cs} cs_t + \epsilon_t^i \tag{2.42}$$

$$\epsilon_t^q = \psi_\epsilon \epsilon_{t-1}^q + \epsilon_t^{q2}. \tag{2.43}$$

2.8 Estimation

While some models are parameterised using estimated coefficients, others are calibrated to fit certain aspects of the data. With a model this small, incomplete specification is unavoidable — there are features of recent economic history that cannot be explained within the very narrow modelling framework considered here. But this does not mean it cannot be used for quantitative assessment. It simply implies that accepting the estimation results without some sensitivity to information that is available outside the small model would likely lead to bias.

Likewise, simply choosing the model parameters by applying judgement or with reference to theory would ignore the useful information contained within the data. Bayesian estimation serves as a bridge between calibration and estimation – the selection of priors allows for the incorporation of additional information available to the modeller, while the process of maximum likelihood estimation extracts some value from the data. In practice, the priors serve to guide the maximum likelihood estimate by placing more weight on certain areas of the parameter space. And the chosen prior variance acts to determine the weighting between the prior and the unconstrained maximum likelihood estimate contained within the posterior estimate.

My choice of prior distributions for the model parameters is informed by other studies relating to the New Zealand economy including, Lubik & Schorfheide (2007), Ryan *et al.* (2009), Parkyn & Vehbi (2014) and Karagedikli *et al.* (2013). I also refer to the simulation properties of the New Zealand Treasury Model and Reserve Bank of New Zealand forecasting models – see Szeto *et al.* (2003) and Beneš *et al.* (2009).

There is good reason to suspect that the dynamics of the model associated with the exchange rate are muddied by commodity prices. International evidence such as that presented in IMF (2012b) suggests that domestic output is positively correlated with commodity prices for exporters of primary goods, such as New Zealand. Because the exchange rate is also correlated with commodity prices the exchange rate is positively correlated with output — which is not consistent with theory.

The introduction of commodity prices - which are so persistent as to be indistinguishable, in practice, from a random walk - would affect the stability of the reduced-form modelling and VAR results. Instead, my priors over the exchange rate — drawn from evidence from other, larger models of the New Zealand economy, that include the effect of commodity prices — are imposed more strictly than priors relating to other parameters of the reduced-form model.

With regard to the Bayesian estimation process, I use data from the final quarter of 1993 to the third quarter of 2012 – avoiding New Zealand's disinflationary period but

Table 2.1: Estimated Model Parameters and Priors

Param.	Equation	Description	Prior value	Prior dist.	Prior Deviation	Post value	Post lower conf int.	Post upper conf int.
	IS	Degree of habit formation	0.80	Norm	0.03	0.87	0.83	0.91
α_y								
α_r	IS	Interest rate elasticity of demand	0.25	Norm	0.04	0.09	0.06	0.13
α_q	IS	Exchange rate elasticity of demand	0.025	Norm	0.0025	0.022	0.018	0.026
α_f	IS	Proportion of Ricardian Households	0.65	Norm	0.08	0.44	0.30	0.57
ϵ_t^y	IS	Demand shock	0.50	Inv Gamma	0.10	0.67	0.58	0.75
β_π	PC	Calvo coefficient	0.20	Norm	0.05	0.22	0.14	0.29
β_y	PC	Output gap sensitivity of inflation	0.25	Norm	0.10	0.26	0.15	0.29
β_q	PC	Exchange rate sensitivity of inflation	0.025	Norm	0.0025	0.025	0.021	0.029
ϵ_t^π	PC	Inflation shock	1.40	Inv Gamma	0.1	1.4	1.23	1.55
δ_i	TR	Interest rate smoothing parameter	0.75	Norm	0.02	0.79	0.76	0.82
δ_y	TR	Output gap sensitivity of interest rates	0.50	Norm	0.20	0.57	0.28	0.85
δ_π	TR	Inflation sensitivity of interest rates	1.50	Norm	0.40	1.48	0.84	2.09

Table 2.2: Estimated Model Parameters and Priors continued.

Param.	Equation	Description	Prior value	Prior dist.	Prior Deviation	Post value	Post lower conf int.	Post upper conf int.
ϵ_t^i	TR	Interest rate shock	0.60	Inv Gamma	0.05	0.70	0.63	0.77
$\delta_{cs} = \tau cs$	TR/IE	Spread equality with base rate	1.00	Norm	0.10	0.98	0.83	1.14
ψ_q	RUIP	Error correction coefficient	0.02	Norm	0.005	0.022	0.014	0.030
ψ_ϵ	RUIP	Shock persistence	0.85	Beta	0.1	0.86	0.81	0.91
ϵ_t^{q2}	RUIP	Exchange rate shock	0.70	Inv Gamma	0.20	0.61	0.43	0.78
θ_{cs}	CS	Spread persistence	0.80	Norm	0.10	0.79	0.71	0.87
ϵ_t^{cs}	CS	Credit spread shock	0.45	Inv Gamma	0.05	0.43	0.38	0.48
ζ_{if}	IF	Foreign interest rate persistence	0.85	Beta	0.10	0.86	0.81	0.91
ϵ_t^{if}	IF	Foreign interest rate shock	0.40	Inv Gamma	0.05	0.36	0.32	0.40
ζ_π	PF	Foreign inflation persistence	0.15	Norm	0.05	0.12	0.04	0.19
$\epsilon_t^{\pi f}$	PF	Foreign price shock	2	Inv Gamma	0.50	2.29	1.99	2.58
f_t	F	Fiscal policy shock	0.14	Inv Gamma	0.05	0.13	0.11	0.15

making use of data over the recent recession. The likelihood function is estimated using the Kalman filter and the Metropolis-Hastings algorithm is used to generate draws from the posterior distribution. 100,000 draws are run with the first 25,000 discarded as burn in. Tables 2.1 and 2.2 presents both the choice of priors and the posterior estimates of the model parameters.[1]

The data used in the estimation process were sourced and transformed as follows. The output gap is presented as a percentage of real GDP and is sourced from Szeto *et al.* (2013). The interest rate is the deviation from the sample average and sourced from the Reserve Bank of New Zealand. The real exchange rate is a composite of weighted bilateral rates, sourced from the Reserve Bank of New Zealand and Statistics New Zealand. Inflation is quarterly annualised CPI inflation sourced from statistics New Zealand and seasonally adjusted. The credit spread is the deviation of sample average of composite lending rates weighted by shares in lending - each sourced from the Reserve Bank of New Zealand. Foreign policy interest rates are sourced from Reuters and trade weighted and expressed as a deviation from the sample mean. The same is true of foreign inflation, which is also seasonally adjusted. The fiscal impulse is the change in cyclically-adjusted budget balance. This is sourced from the New Zealand Treasury and interpolated to a quarterly frequency.

2.9 Model Evaluation

The estimation results do not throw up a huge number of surprises. The persistence of the output gap is a little higher than the prior but still some distance from being a random walk. And, of course, monetary policy acts as an additional force to bring demand in line with the economy's supply potential. The posterior estimate of the interest rate elasticity of demand is rather lower than the prior, suggesting monetary policy has a smaller impact on aggregate spending than expected. Still, it has an important stabilising role to play in the model.

The impact of the exchange rate on demand is broadly in line with the prior. The degree of non-ricardian behaviour is a little lower than the prior. That's an interesting finding. The effect of tax and spending changes on output varies depending on the type of measure being implemented and this parameter can be thought of as the fiscal impact multiplier. It's worth remembering, then, that the estimated parameter will reflect the impact of the mix of fiscal policies implemented on average in the past. The width of the confidence intervals will also reflect the possibility that the impact has varied over the past, depending on the composition of measures.

[1] Estimation is conducted using the DYNARE software package.

To evaluate the fit of the model I first estimate a two-lag SVAR model, which allows the data to speak with the minimum number of identifying restrictions applied. The data series included are the inflation deviation from target, the output gap, the real interest rate gap and the exchange rate gap, which is also the Cholesky ordering of the variables. Fiscal policy enters the SVAR exogenously since it does not theoretically depend on any other variable in the model.

I then compare the impulse responses from the SVAR with those of the estimated reduced-form model. I find that, broadly speaking, the impulse responses are consistent with one another. In particular, the humped responses of inflation to the output gap and the output gap to interest rates receive good empirical support. The dynamics of the exchange rate and the associated influences on output are less well-supported. Overall, the SVAR impulse responses generally support the dynamics of the reduced-form model – that's illustrated in Figures 2.1,2.2,2.3 and 2.4.

To provide an alternative model against which to compare the impulse responses, I estimate a sign-restricted VAR (SRVAR) in a similar way to Jääskelä & Jennings (2011). The associated impulse responses are presented alongside the responses of the other models below). The restrictions are presented in Table 2.3.

The interest rate response to itself is assumed positive given the persistence of interest rates - that's also reinforced by the smoothing parameter that features in most estimated reaction functions. Inflation is assumed to respond negatively to rates, and the exchange rate jumps as predicted by UIP. The output gap is also persistent, justifying the positive restriction on itself and higher output causes the labor market to tighten, lifting wages and costs. Inflation and exchange rate shocks prompt replies in line with the general theoretical predictions of a New Keynesian model.

The restrictions serve to eliminate both the exchange rate and price puzzles associated with the VAR estimates and leaves the magnitude of many of the impulse responses broadly consistent with those from the reduced-form model[1]. However, the approach also introduces an excess sensitivity of the exchange rate to endogenous factors, probably reflecting misspecification and the omission of relevant variables in the estimation process.

Finally, I am also interested to see whether the cross-equation restrictions of the reduced-form model have a significant bearing on its dynamics. To investigate whether this is the case I run a stochastic simulation of the model and record the data that is generated. I then estimate another SVAR using that simulation data - the identification is achieved using the same ordering as in the SVAR described above - and compare the

[1]The drawback of this approach is that the sign restrictions appear to make the responses more immediate, since they apply to the first lag of the VAR — this is also a feature of the sign-restricted estimates presented in Jaaskela and Jennings)

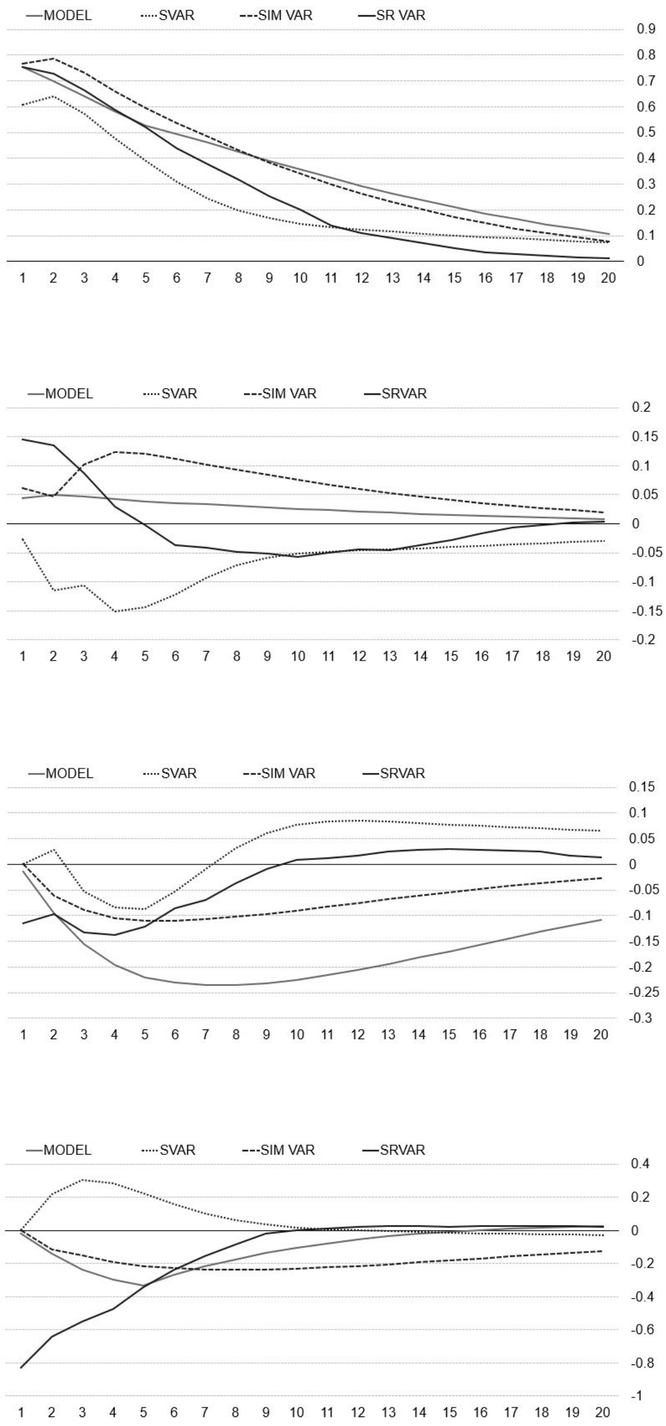


Figure 2.1: First figure - response of output to an output shock. Second figure - response of output to an inflation shock (1 standard deviation). Third figure - response of output to an interest rate shock. Fourth figure - response of output to an exchange rate shock – (1 standard deviation).

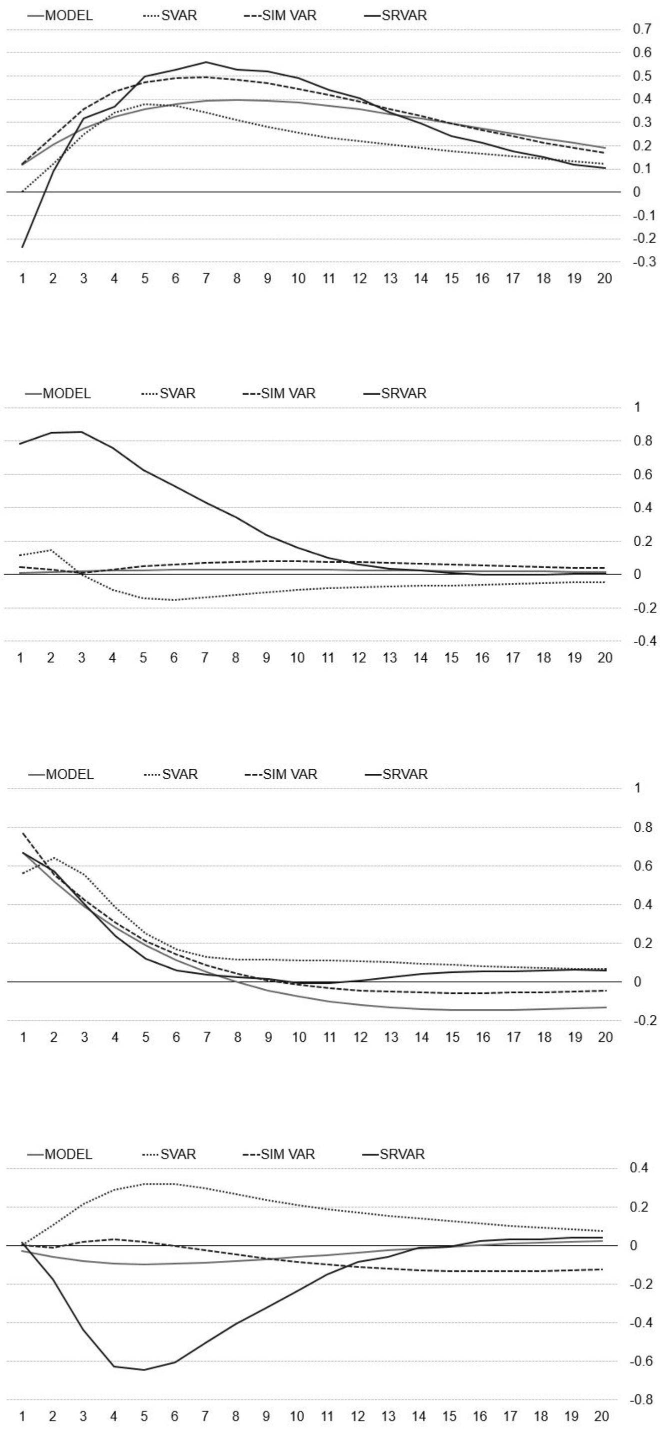


Figure 2.2: First figure - response of inflation to output shock. Second figure - response of inflation to an inflation shock (1 standard deviation). Third figure - response of inflation to interest rate shock. Fourth figure - response of inflation to an exchange rate shock – (1 standard deviation).

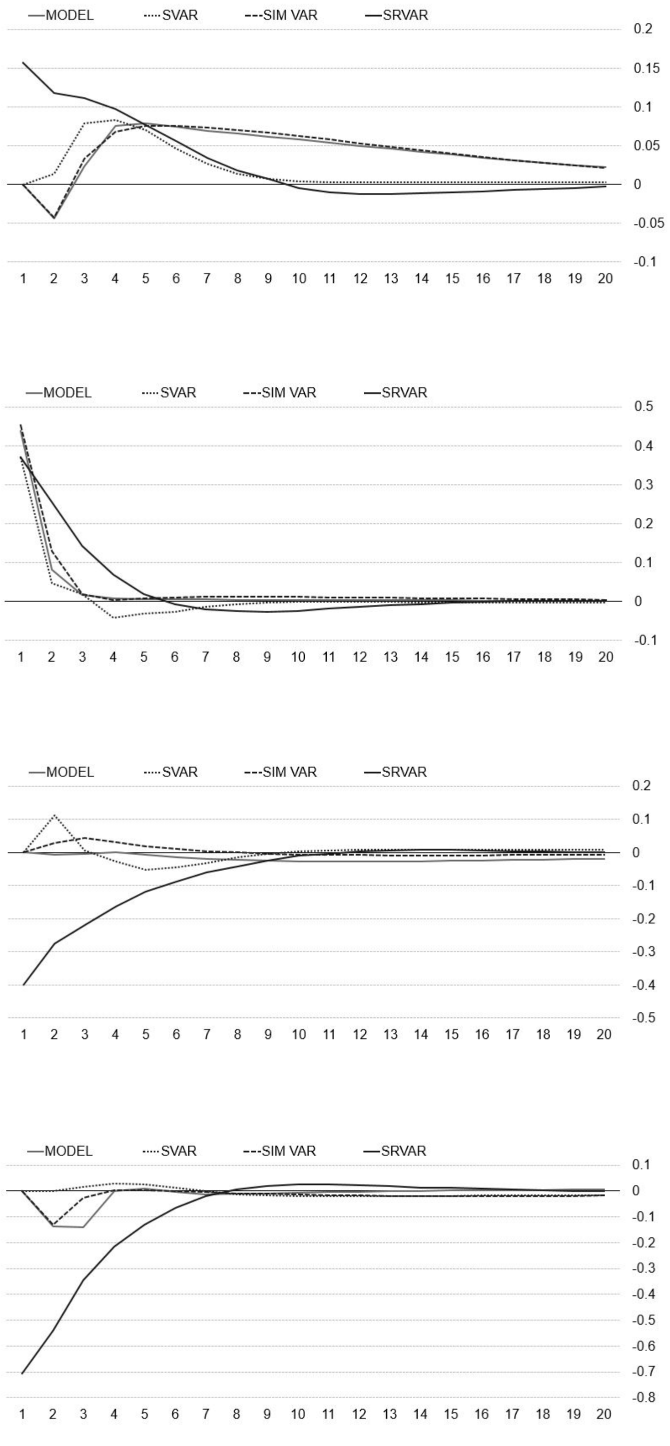


Figure 2.3: First figure - response of interest rate to output shock. Second figure - response of interest rate to an inflation shock (1 standard deviation). Third figure - response of interest rate to an interest rate shock. Fourth figure: response of interest rate to an exchange rate shock (1 standard deviation).

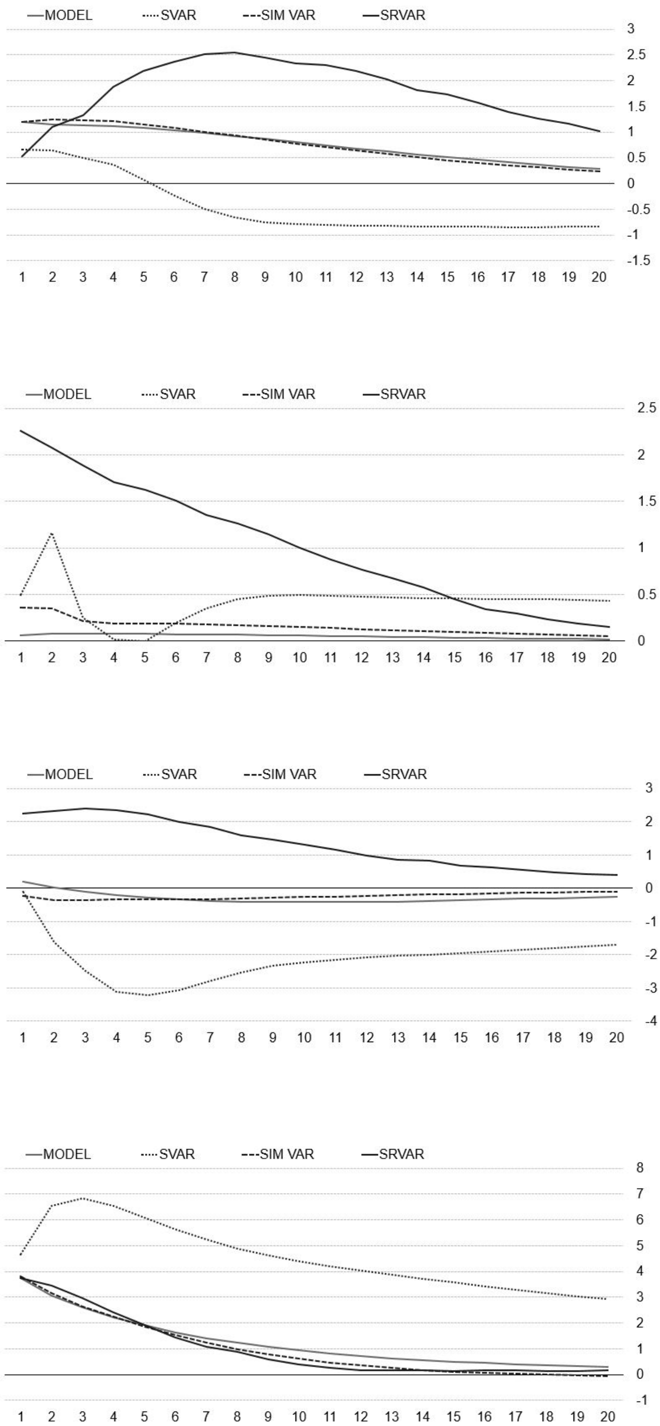


Figure 2.4: First figure: response of exchange rate to an output shock. Second figure - response of exchange rate to and inflation shock (1 standard deviation. Third figure - response of exchange rate to an interest rate shock. Fourth figure response of exchange rate to an exchange rate shock (1 standard deviation).

Table 2.3: The table shows the identifying restrictions for the sign-restricted VAR. Column to the left shows the variable that is being shocked. The row to the top shows the type of shock. The cells show the sign restriction pertaining to that shock.

Shock to:	Interest rate	Output	Inflation	Exchange rate
Interest rate	Positive	-	Negative	Positive
Output	-	Positive	Positive	-
Inflation	Positive	-	Positive	Positive
Exchange rate	-	Negative	Negative	Positive

impulse responses with those of the estimated model. I find that the impulse responses are broadly consistent with one another, suggesting the findings presented later are not overly dependent on the cross-equation restrictions of the reduced-form model.

2.10 A baseline fiscal consolidation

In this section, I first set out a fiscal policy simulation using the estimated model parameters, which serves as the baseline case against which other simulations are compared. I then vary the model parameters one at a time, to show the sensitivity of the fiscal impact multiplier and associated cumulative output losses to those parameters. The variations are proportionate to the confidence intervals obtained during the Bayesian estimation procedure, which serve as proxies for parameter uncertainty.

The baseline fiscal consolidation scenario presented here is based on a four-year consolidation program equal to a one per cent of GDP tightening in each year, starting one quarter after it is announced. The full consolidation is, therefore, expected by the central bank, foreign exchange market participants and households. Figures 2.5(a) and 2.5(b) illustrate the dynamics of the four key model variables in response to a scheduled fiscal consolidation.

The figures show that the exchange rate depreciates at the point of announcement, while interest rates fall with a degree of inertia. The level of prices rises initially, reflecting the increase in import costs associated with the currency depreciation. The exchange rate depreciation provides support to output over the first year; thereafter the output gap opens up and exerts downward pressure on the annual rate of inflation. The output gap continues to widen over the period of fiscal consolidation, implying a negative GDP growth effect and a positive fiscal impact multiplier.

Figure 2.5 shows that the fiscal impact multiplier is small in the first year as much of the effect of the consolidation is offset by the output effects of currency depreciation. Thereafter, the multiplier rises to around 0.4 before shrinking to around 0.3 toward the end of the consolidation period. The average fiscal impact multiplier over the consolida-

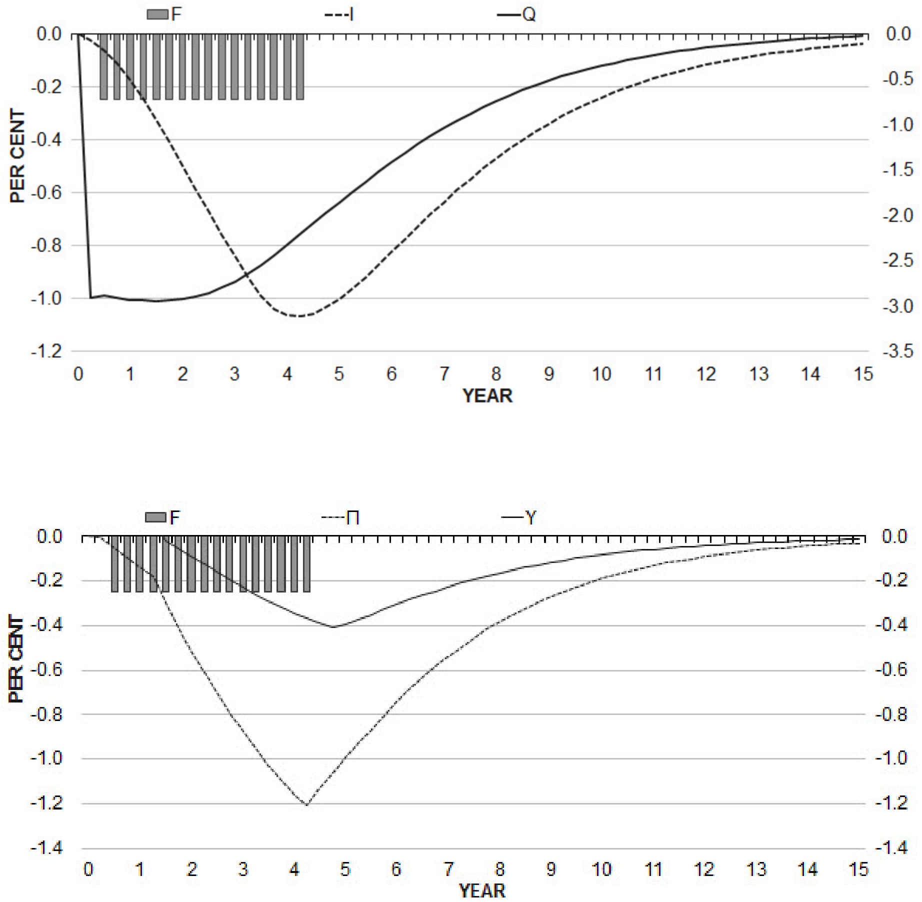


Figure 2.5: Top: Interest and exchange rates In both figures the bars show the size of the fiscal tightening being administered as a percentage of GDP. In the top figure the dark line shows the movement of the exchange rate in reply and the dashed line shows the response of interest rates. In the bottom figure the solid line is inflation and the dashed line is the output gap. **Bottom: Output and inflation**

tion period is also 0.3.

The opening of the output gap is associated with a cumulative loss of income of around 6 per cent of annual GDP — illustrated in Figure 2.6. This output will never be recovered since there is no offsetting positive output gap following the consolidation — i.e. the level of output returns to its steady-state growth path and remains there.

The risks associated with speed of consolidation are not quantified here. Varying the pace of the fiscal consolidation by, for example, compressing the duration over which it takes place to two years does not yield a larger estimated cumulative output loss or a larger fiscal multiplier. The implication is that fiscal policy can be set without any regard to speed of consolidation and, in effect, makes achieving fiscal balance today just as costly as achieving it over four years — this is a limitation that arises from the linearity

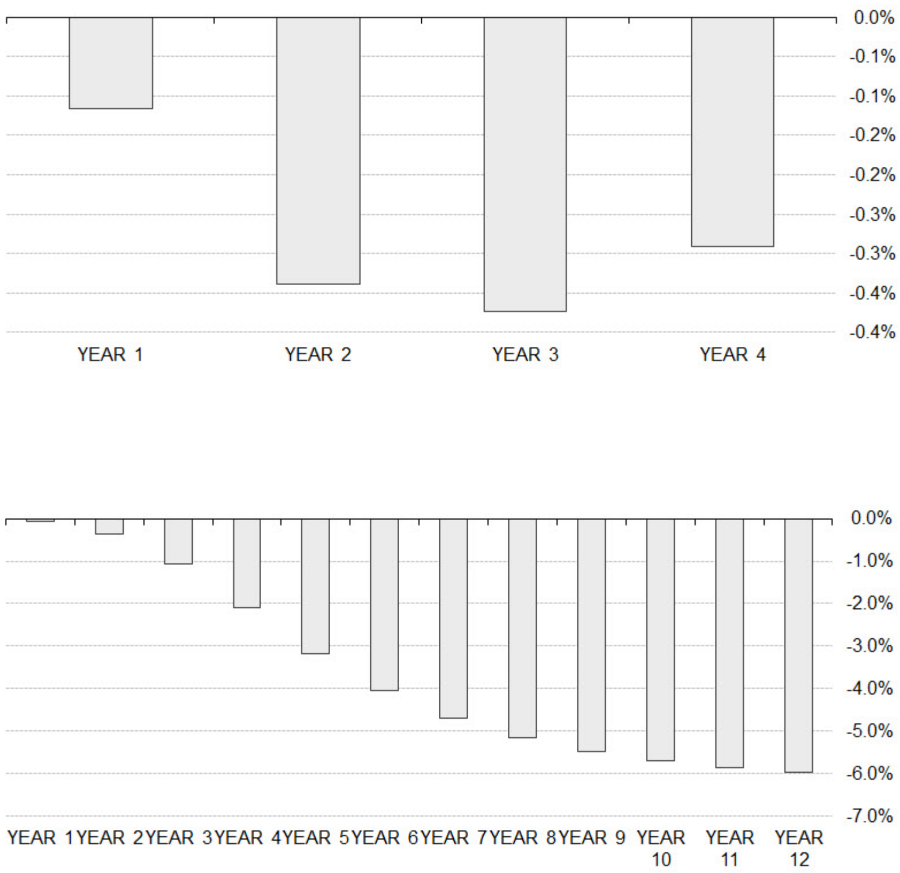


Figure 2.6: **Top: Fiscal impact multiplier** The top figure shows the size of the fiscal impact multiplier in each of the years when a four-year tightening program is administered. The bars show the impact as a percentage of GDP. The bottom figure shows the sum of output losses made in each year as a percent of annual GDP. **Bottom: Cumulative output loss**

of the model.

Adjusting the baseline scenario so that the consolidation starts in a year's time (rather than the next quarter) affects the profile of GDP growth over time. This is because the exchange rate adjusts at the point of announcement, not when the consolidation begins, which means the associated benefits to growth are felt before consolidation begins. But this simply brings demand forward from later years and, overall, the cumulative gains associated with preannouncement are small — equal to around 0.1 per cent of annual GDP over the period in which the consolidation takes place.

2.11 Measuring parameter uncertainty

I have used the confidence intervals from the Bayesian estimation for the relevant parameters as proxies for parameter uncertainty. Table 2.4 ranks the model parameters by the cumulative output losses (relative to baseline) associated with a one standard deviation variation from their respective estimated values[1]. The mean estimates for each of the model parameters do not always lie precisely at the centre of their estimated confidence intervals.

2.11.1 Some parameters that matter

The **degree of monetary activism** is captured by the responsiveness of interest rates to both the output gap and deviations of inflation from target. The persistence of the output gap is affected by variations in both of these parameters and the risks posed by both the fiscal impact multiplier and cumulative losses are skewed to the downside.

The **interest rate elasticity of demand** represents the willingness of households to swap consumption today for consumption tomorrow. It is important because it determines how effective a given interest rate change will be in stimulating aggregate demand and variations in this parameters alter the persistence of the output gap. The results show that the risks posed to both the cumulative output losses and associated fiscal impact multipliers are skewed to the downside when this parameter is varied.

The **degree of non-Ricardian behaviour**[2] introduces considerable uncertainty over the likely effect of fiscal consolidation on the economy. This is not surprising since it is the scalar for the size of the initial shock. The results are consistent with a broadly symmetric loss/gain in cumulative output and varying this parameter has a roughly equal effect on the impact multiplier when varied in both directions.

The output gap paths associated with independently varying each of the four parameters above by one standard deviation of the parameter estimate are presented in Figure 2.7(a) while Figure 2.7(b) shows these in terms of deviations from the baseline scenario. The simulation shows that a higher degree of non-Ricardian behaviour would reduce output sooner and more sharply than would a lower interest rate elasticity of demand or interest rate sensitivity to the output gap/inflation. This is reflected in estimates of the fiscal impact multiplier illustrated in Figure 2.8.

[1] To some extent, the confidence intervals reflect the priors for the standard error of the parameters but the data also influences these ranges.

[2] Which includes the effect of other leakages

Equation	Description	Cumulative output loss -1 s.d.	Cumulative output loss +1sd	Cumulative output loss -1sd vs baseline	Cumulative output loss +1sd vs baseline
TR	Interest rate sensitivity to output gap	9.2	4.4	3.0	-1.7
IS	Interest rate elasticity of demand	8.7	4.6	2.5	-1.6
IS	Degree of non-Ricardian behaviour	4.3	8.0	-1.9	1.8
TR	Interest rate sensitivity to inflation	8.2	4.8	2.1	-1.4
PC	Inflation sensitivity to output gap	6.7	5.6	0.6	-0.6
PC	Inflation persistence	6.2	5.9	0.0	-0.2
RUIP	Error correction coefficient	6.0	6.1	-0.2	-0.1
IS	Exchange rate elasticity of demand	6.1	6.0	-0.1	-0.1
PC	Exchange rate sensitivity of inflation	6.1	6.0	-0.1	-0.1
TR	Interest rate smoothing parameter	6.0	6.2	-0.2	0.0
IS	Degree of habit formation	6.1	6.0	0.0	-0.2

Table 2.4: Cumulative output losses.

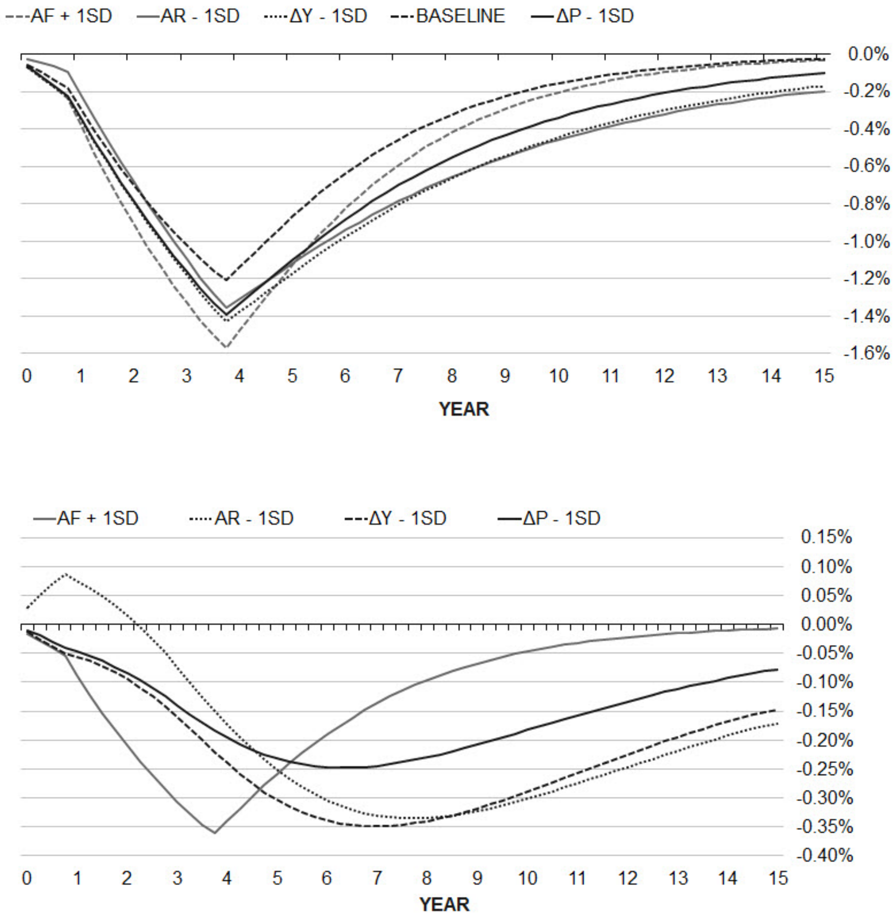
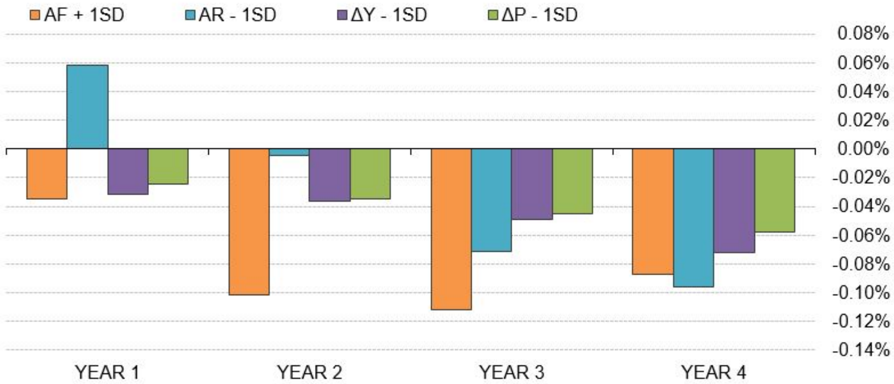


Figure 2.7: Top: Output gap simulations Both figures show the different responses of output to the same fiscal shock under different model parameters. In separate simulations one standard deviation is added or subtracted to the estimated model coefficients – the degree of Ricardian equivalence, the interest rate elasticity of demand, the persistence of the output gap and the persistence of inflation. The top figure shows the estimated responses. The bottom shows the difference between those responses and the baseline simulation. **Bottom: Output gaps relative to baseline**

2.11.2 Some parameters that matter less

Several parameters affect the distribution of GDP growth over the consolidation period but not the cumulative output losses associated with it. Exchange rate variables, for example, affect the impact multiplier in the first year but work in the opposite direction in later years as the initial depreciation is offset by appreciation to achieve parity with the rest of the world.

The parameters that have this sort of influence include the degree of interest rate smoothing, the degree of habit formation, the pass-through from the exchange rate to inflation and the elasticity of output with respect to the exchange rate. None of the stan-



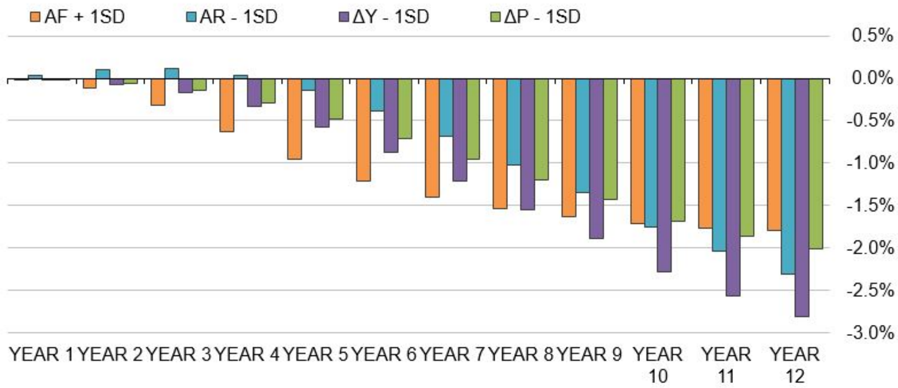


Figure 2.8: Top: Fiscal impact multiplier vs baseline The top figure shows how the fiscal impact as a percent of GDP varies under different assumptions for the model parameters. The variations in those parameters are the same as in the previous figure. The bottom figure shows the sum of output losses in each year as a percent of GDP. **Bottom: Cumulative output loss vs baseline**

dard deviation variations of these parameters affects the fiscal impact multiplier by more than 0.1 ppts in any single year. The sensitivity of inflation to the output gap also has a relatively small influence on cumulative output losses and the fiscal impact multiplier, although it does not fall into the same category as the other variables described in this section.

2.11.3 How much better or worse might it be?

The four parameters which are most important in determining the overall output losses are the sensitivity of output to interest rates, the sensitivities of interest rates to both output and inflation and the degree of non-Ricardian behaviour. If these parameters are independent, unbiased and normally distributed, the likelihood of all four of these pa-

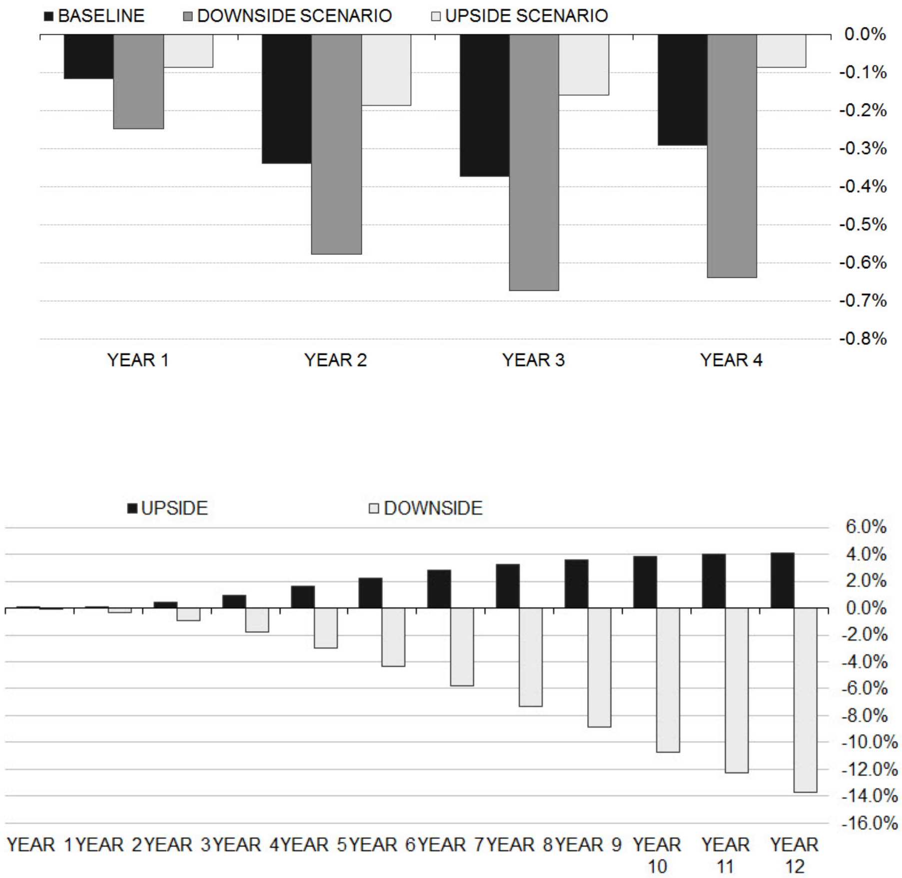
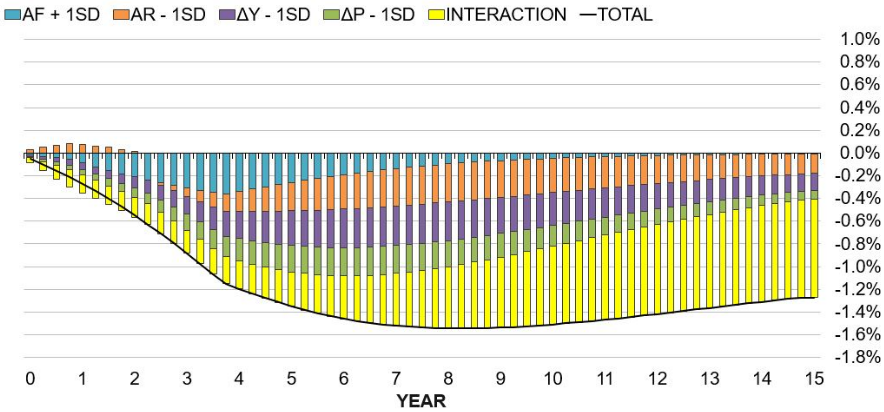


Figure 2.9: **Top: Fiscal impact multiplier vs baseline** The top figure shows how the impact of fiscal policy on the level of GDP - in percent - varies depending on how model parameters are varied. This figure shows the output of three simulations. A baseline, one in which parameters are shifted in a way that softens the impact and another in which the shift deepens the impact. The bottom figure shows the sum of annual losses as a percentage of GDP. **Bottom: Cumulative output loss/gain vs baseline**

rameters lying at one end of their respective confidence intervals simultaneously is rather small. Nonetheless, it is an interesting thought experiment to ask what the output path might be if they did. Likewise, it is informative to explore the implications for output if these parameters were to lie at the favourable end of their respective confidence intervals.

The results of this exercise suggest that the cumulative output loss associated with a one standard deviation shock to each parameter (to the side consistent with output losses) would be substantial - more than tripling the total cumulative output loss associated with the consolidation. And the fiscal impact multiplier is also larger, peaking at 0.7 in Year 3, compared with 0.4 in the baseline scenario. The average multiplier is 0.5 in this scenario compared with 0.3 in the baseline case. These results are illustrated in Figure 2.9.

A key finding of this study is that a bad outcome is likely to be worse than a good out-



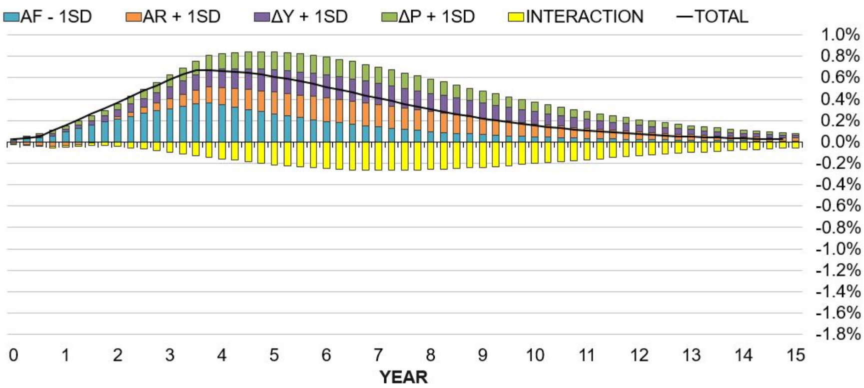


Figure 2.10: **Top: Downside scenario** Both figures show the impact of varying model parameters for the same fiscal tightening as a percentage of GDP **Bottom: Upside scenario**

come is to be better, suggesting that risks associated with the consolidation are skewed to the downside. This finding reflects the underlying distributions of the parameter estimates which are found to be skewed themselves[1].

However, the interaction between the increased degree of non-Ricardian behaviour (which widens the output gap) and the increased persistence of the output gap arising from variations in the other model parameters means that the cumulative loss of output is greater than the sum of losses associated with varying each of the model parameters independently. This is illustrated in Figure 2.10 which shows the contributions to the wider output gap from independent variations in model parameters and the contribution of the interaction between them.

[1]The model itself is linear, but the degree by which the parameters are varied depends on the confidence intervals of the estimated parameters, which reflect their distributions.

2.12 Fiscal policy at the zero-lower bound

The central estimate of the average fiscal multiplier associated with a four-year fiscal consolidation is 0.3, consistent with the findings of Parkyn & Vehbi (2014). The results of the sensitivity analysis conducted here suggest that the fiscal impact multiplier is likely to be larger and the cumulative output losses substantially greater if certain model parameters differ from the central estimates. But these variations do not cause the estimated fiscal impact multiplier to reach unity – the peak impact multiplier over the consolidation period is 0.7 in the downside scenario while the average impact multiplier is 0.5.

Partly, these findings may reflect the choice of priors and so the parameters may simply not be varied by enough[1]. However, the results suggest that the functioning of the economy would have to be greatly different to the model specified here for the fiscal impact multiplier to be as large as is found by Blanchard & Leigh (2013). Given the importance of monetary policy in determining the size of the fiscal multiplier, it is useful to test whether the zero-lower bound of nominal interest rates is able to reconcile the differences. To that end, I run a simulation in which the nominal interest rate is held fixed over the period of consolidation and monetary policy is unable to stimulate aggregate demand.

In this scenario the fiscal impact multiplier is substantially higher, peaking at 0.9, and averaging 0.7 over the consolidation period. Assuming that the parameter capturing the degree of non-Ricardian behaviour lies at the unfavourable end of its distribution increases the peak multiplier to 1.2 and the average multiplier to unity — Figure 2.11. These estimates are fairly consistent those of Blanchard & Leigh (2013), suggesting that much of the difference between estimates of the fiscal multiplier in New Zealand relative to other OECD countries is due to monetary policy constraints.[2]

2.13 Implications for policy research

The key findings of the research presented here are that uncertainty surrounding the effects of fiscal consolidations on output can be attributed to several model parameters and that a bad outcome is likely to be worse than a good outcome is to be better. Overall, the evidence suggests that policy makers should be sensitive to the prevailing economic environment when determining the fiscal stance because cumulative output losses can vary substantially in some situations — particularly when monetary policy is constrained by the lower bound of nominal interest rates.

[1] See Leeper *et al.* (2017) for a discussion of this possibility.

[2] Note that the assumption that the lower bound binds is consistent with the belief that other measures taken by central banks to stimulate aggregate demand have had limited positive effects.

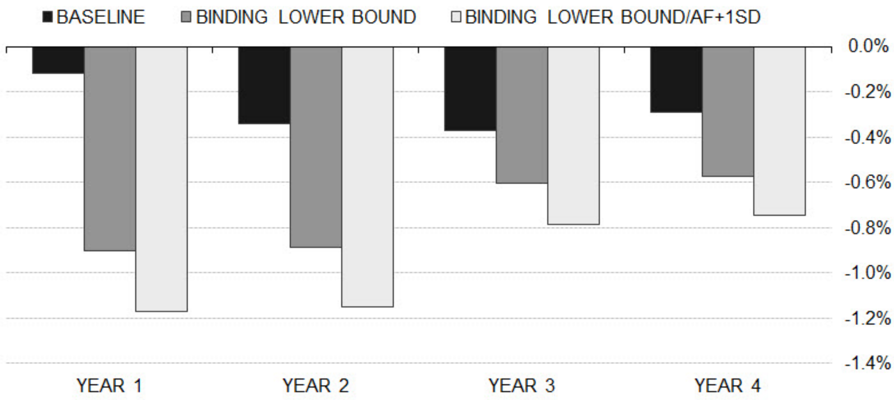


Figure 2.11: Figure shows the impact of fiscal policy tightening in each year of the simulation as a percentage of GDP under different assumption about whether the lower bound of interest rates binds.

The responsiveness of aggregate demand to changes in interest rates is a key determinant of the output losses associated with any fiscal consolidation. This is thought to be related to a structural parameter — the elasticity of intertemporal substitution — which, in turn, is often considered to be stable and not to fluctuate over the cycle. This is of particular interest at the current juncture since little is known about how the structural position of household and corporate balance sheets might affect those agents' willingness to bring consumption forward in response to lower interest rates. Further work into the validity of this assumption would improve our understanding of the effects of fiscal consolidations on the economy.

The degree of monetary activism — how much a central bank might be expected to move interest rates in response to an announced fiscal consolidation — is also an important determinant of the effect of consolidations on output. And, in extremis, when monetary policy is constrained by the zero lower bound of interest rates, the effects of fiscal tightening are likely to be much larger. The implication is that central banks and fiscal authorities should coordinate their activities closely if the worst outcomes are to be avoided.

The degree of non-Ricardian behaviour, which (due to the reduced-form nature of the model) also includes the effect of trade leakages, is of particular importance. When setting policy governments should consider whether the particular mix of measures is likely to affect households likely to exhibit more or less Ricardian behaviour. A government should also consider whether the package it designs is likely to be more or less prone to leakage, reflecting the import intensity of certain areas of expenditure, for example. This reduces the uncertainty over this parameter to the extent that the central estimate can be thought of as being based on the average effect of a number of packages which differ in their precise make-up.

A broader question is how this information should shape a government's policy choices. In the parallel literature on optimal monetary policy, uncertainty over model parameters implies an inertial response of interest rates to shocks. The issue for fiscal authorities is rather more complicated since the benefits of reducing structural budget deficits need to be balanced against the output losses associated with fiscal consolidations — I leave this area unexplored but it stands to reason that the asymmetric output losses arising from parameter uncertainty might incentivise a degree of gradualism in policy setting.

2.14 Conclusion

In this paper, I have presented and estimated a small model of the New Zealand economy. I then ran a number of fiscal consolidation scenarios and used the results to show the sensitivity of the fiscal impact multiplier and the associated cumulative output losses to uncertainty over the model parameters.

The key findings are that uncertainty surrounding the effects of fiscal consolidations on output can be attributed to several model parameters and that a bad outcome is likely to be worse than a good outcome is to be better. I find that, if monetary policy were to be constrained by the zero-lower bound, the estimated fiscal impact multiplier for New Zealand would be broadly consistent with estimates of the fiscal multiplier in a number of other OECD countries in that position.

Overall, the evidence suggests that fiscal policy makers should be sensitive to the prevailing economic environment when determining the fiscal stance and work closely with central banks if the worst outcomes are to be avoided.

3

Output Gap Uncertainty in the United Kingdom: A Fiscal Perspective

3.1 Introduction

In thinking about the sustainability of the public finances, it is important to consider how the cyclical position of the economy might be affecting revenues and spending. When the economy is operating below its full capacity, elevated unemployment suppresses income tax revenues and boosts spending on out-of-work benefits, for example. Likewise, an overheating economy inflates revenues, since higher wages are needed to tempt more people into the workforce or encourage them to work more hours, and lowers spending on some benefits. The most commonly described measure of spare capacity or overheating is the output gap — the difference between actual output and an estimate of underlying potential output.

Recognising the role played by these cyclical factors, some governments aim to achieve balance of a cyclically-adjusted measure of the public finances over a chosen time horizon. Or, instead, such a gauge is used to decide how much tightening or loosening is needed to deliver a specific fiscal target some years ahead.

In practice, cyclically adjusting the public finances is not a simple task: first, the output gap is not directly observable, is inherently uncertain and is prone to substantial revision; second, even if the cyclical position of the economy could be known with certainty, the sensitivity of revenues and spending to it would still have to be assessed. This paper is concerned with the issue of output gap measurement and uncertainty[1].

The paper begins with an analysis of issues surrounding the appropriate definition of the output gap for the purpose of assessing fiscal sustainability. It goes on to quantify three sources of output gap uncertainty – data revisions, the arrival of new data and the use of different models. The paper ends with analysis of the big problems policymakers face in the use of output gap estimates.

One key conclusion is that governments should choose the method of output gap estimation that best fits their priors over the processes driving trend growth. Blindly applying statistical filters, in whatever form, invites repeated and avoidable errors. Even so, it is also shown that output gap uncertainty is huge and the arrival of new data will generally prompt big revisions to estimates of the structural deficit.

3.1.1 Conceptual issues

The output gap is the difference between actual output and potential output — the maximum level of output that could be achieved while maintaining stable inflation over a given time horizon. It depends on how many people are available to work and how many hours they are willing to put in (labour); the number of buildings, machines and computers that are available to work with (capital); and the efficiency with which they can be combined (productivity).

3.1.2 Time horizon

In both the academic literature and in public discourse, spare capacity is often viewed from the perspective of a central bank, rather than that of a fiscal authority. Normally this is not particularly significant, but there is a distinction between the two that may be more important when the output gap is large.

In setting borrowing plans over the coming years (five in the UK), fiscal authorities are interested in how long it might take for spare capacity to be soaked up over a relatively long time horizon. Central banks, on the other hand are concerned with what resources can be put to use without stoking inflation over a much shorter period.

In the short-term, for example, lifting output when the long-term unemployment rate is very high might prompt inflationary pressure if those workers are ill-equipped for the roles available. The demand for labour might instead be satisfied by encouraging existing workers to put in more hours by paying them more, creating inflationary pressure.

In time, however, the long-term unemployed might reasonably be expected to re-train and output could be lifted in a way that does not trigger excess inflation but does lift tax

[1]The sensitivity of the public finances to the cycle is the subject of Helgadottir *et al.* (2012)

receipts and lower benefit spending, improving the fiscal position.

It might be worth considering, then, that measures of spare capacity used by fiscal policymakers should differ from those at central banks. A failure to understand that it might take longer for the public finances to recover than for inflationary pressure to build could prompt policymakers to overestimate the size of the structural deficit. That could lead to excessive fiscal tightening and slower economic recoveries, particularly when monetary policy is constrained by the zero lower bound.

Many measures of the output gap with a monetary policy emphasis, applied in the academic literature, can be adjusted to take account of this difference in perspective. For example, bottom-up estimates of potential output, such as those derived using a production function, require explicit judgements over the equilibrium rate of unemployment, trend hours worked and activity. And judgements surrounding smoothing parameters, for example, can be adjusted when using statistical filters to produce top-down estimates.

In what follows, the methods of output gap estimation are consistent with being from the perspective of the fiscal authority rather than the central bank.

3.1.3 Uncertainty

The output gap cannot be measured directly or known with certainty, even with the benefit of hindsight. Revisions to estimates of the output gap can, therefore, be significant. They come from three sources:

- end-point uncertainty arises because the future path of output is unknown and it may contain information about the cyclical position of the economy now. This matters more for some estimation methods than others, largely reflecting the assumptions that underpin them and the extent to which information from the future is used to inform current estimates of the output gap;

- data uncertainty arises because the information available at the time is not the final vintage of that data. It may become more accurate with time as more information from that time period becomes available and measurement methods improve. Some methods are more sensitive to this than others, dependent on the degree to which revisions are attributed to the level of potential output or the output gap; and

- model uncertainty reflects mainly changes to our understanding of how the economy functions. Generally, methods with a richer economic structure would be more susceptible to this source of uncertainty but, as our understanding of the process governing the growth of productivity evolves, for example, so might our view on the volatility of potential output — which would affect the estimates from all the methods presented in what follows.

Uncertainty matters because it could be relevant to policy. It has long been considered, Brainard (1967) that uncertainty over the way the functioning of the economy functions could affect the way in which monetary policy is set. That same principle must surely be relevant to fiscal policy as well.

The susceptibility of output gap estimates to different sources of revision is examined throughout this paper. The effect of new data is assessed by comparing the real-time and ex-post estimates for a variety of the methods presented; the influence of revisions to data is assessed by comparing output gap estimates using current and real-time vintages of the data, while model uncertainty is reflected in the wide range of output gap estimates produced using different methods.

3.1.4 Assessing performance

Since the potential output of the economy cannot be measured directly, we will never know whether the estimates we construct are accurate. This makes it difficult to assess the benefits of one method of estimation over another. Ideally, we would like a measure that can be calculated accurately in real time (i.e. is not revised much) and appears plausible with hindsight. It is tempting to rank output gap measures only on their tendency to be revised as the benefit of hindsight informs past estimates. However, it would be easy to win this competition using a method that sets actual output equal to potential output at all times and is never revised. Unfortunately, such a method would be useless for forecasting, since, among other things, it would implicitly assume that the unemployed never return to work. So we need to look beyond just tendency for revision when assessing the performance of measures.

For central banks, estimates of the output gap can be assessed based on their performance in explaining that element of inflation thought to depend on demand pressures. For fiscal authorities (and fiscal watchdogs), a sensible metric might be the extent to which the output gap explains cyclical variations in the public finances. But there is an obvious circularity here — how do we know the cyclical parts of inflation and the fiscal balance without already having an estimate of the cycle? Insofar as the estimated output gaps that best explain inflation are likely to have been estimated using the Phillips curve, for example, then they will be ranked (possibly undeservedly) top.

Policymakers typically consider a wide range of evidence when forming a view on the margin of spare capacity in the economy, supported by a narrative that reflects their subjective interpretation of economic developments. An alternative approach to quantitatively evaluating various output gap measures (and having to place a weight on tendency to revision relative to theoretical coherence) is to ask whether they are consistent with our broader understanding of economic history. Clearly, a method indicating significant overheating during the Great Depression might have been a rather unhelpful guide to subsequent economic prospects.

3.2 Univariate Methods

Univariate methods are those that utilise just the output series itself. In what follows, real, non-oil, gross value added per capita is chosen as the measure of actual output.[1] The motivation for this is that the size of the working-age population is unlikely to be closely related to the cyclical position of the economy. Scaling by the population prevents demographic trends from introducing noise to the estimated output gaps. The oil sector is also excluded, since it accounts for a tiny percentage of employment but a more significant share of output — oil output is very volatile because it is affected by maintenance procedures, for example. In what follows, the logs of actual output, potential output and the output gap are given by y_t, y_t^{*t} and c_t respectively and are related by the identity presented in Equation 1.

$$y_t = y_t^{*t} + c_t \tag{3.1}$$

3.2.1 Linear de-trending

The simplest method for estimating the path of potential output is to assume it is a straight line. More complicated is deciding the dates between which it should be drawn. Figure 3.1 show the estimates of the output gap associated with drawing a straight line from the first quarter of 1965 and measuring deviations of actual output from it.

The first figure plots estimates of the output gap in real time and ex-post — where real time means using only data up to the quarter in question and ex-post estimates make use of the whole sample. The revisions are shown to be very large — this is because, as time passes, more is known about the historical rate of growth and the slope of the line is updated accordingly Figure 3.2. Towards the end of the sample, the ex-post estimates (made using the full sample) converge with the real-time estimates of the output gap. This method is relatively sensitive to the chosen starting point — shifting it 5 years forward, for example, reduces the size of the output gap in the final quarter of 2013 from -6.4 per cent to -6.0 per cent.

[1] Where the population is defined as those aged 16 years or over, as recorded by the Labour Force Survey (LFS). Prior to the LFS (which began in 1971), annual total population figures are used, available from the ONS, interpolated to a quarterly frequency and spliced to the later data.

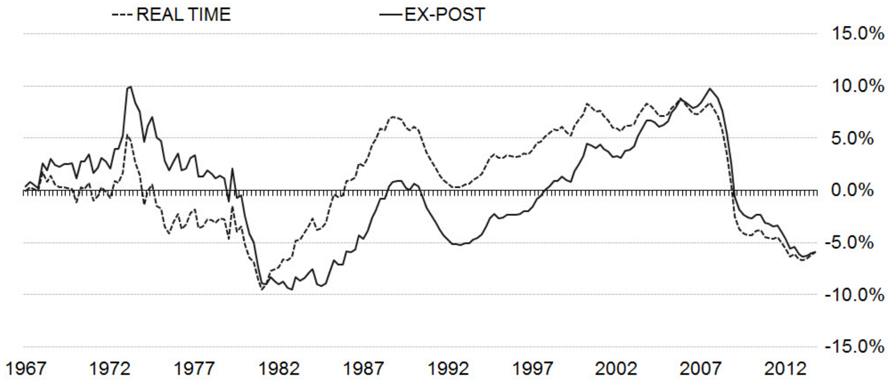


Figure 3.1: Figure shows real-time and ex-post estimates of the output gap as a percent of GDP using the linear de-trending method.

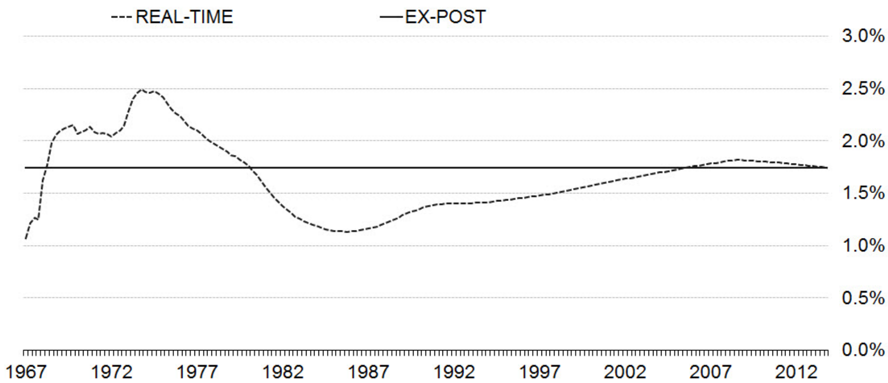


Figure 3.2: Figure shows real-time and ex-post estimates of year-over-year potential GDP growth using the linear de-trending method.

3.2.2 Hodrick-Prescott (HP) filter

The HP filter — Hodrick & Prescott (1997) — is based on two beliefs:

- output does not deviate too far from its trend level (cycles are not too big); and
- the growth rate of potential output is relatively smooth (potential output is not too volatile).

$$\sum_{t=1}^{T}(\frac{1}{\sigma_1^2}(c_t)^2 + \frac{1}{\sigma_2^2}(\Delta y_{t+1}^* - y_t^*)^2) \tag{3.2}$$

The filter chooses y_t^{*t} such that loss function (2) is minimised, where σ_1^2 is the variance of the output gap and σ_2^2 is the variance of trend growth. The user of the HP filter can specify the relative weight placed on the two beliefs by constraining the ratio of the two variance terms to be equal to a specific value, given by λ — as shown in Equation 3.

$$\lambda = \frac{\sigma_1^2}{\sigma_2^2} \tag{3.3}$$

It is clear from Equation 2 that penalising the smoothness of potential output is the same as minimising the sum of squared residuals from the equation:

$$\Delta y_{t+1}^* = \Delta y_t^* + \epsilon_{2,t+1} \tag{3.4}$$

And, in minimising the sum of squared deviations of actual output from trend output, we are minimising the sum of squared residuals from the equation:

$$c_t = \epsilon 1, t. \tag{3.5}$$

The HP filter can also be represented in state-space - the signal equation is given by the identity presented in Equation 1, the state equation for potential output is given by Equation 7 (a manipulation of Equation 4) and the state equation for the output gap is given by Equation 8:

$$\text{Signal :} y_t = y_t^* + c_t \tag{3.6}$$

$$\text{State :} y_{t+1}^* = 2y_t^* - y_{t-1}^* + (\frac{\epsilon_{1,t}}{\lambda^{1/2}}) \tag{3.7}$$

$$\text{State :} c_t = \epsilon 1, t \tag{3.8}$$

The lambda term has been applied such that the error terms are constrained by the relative weight placed on output gap minimisation and potential output growth smoothness. The state-space representation of the HP filter can be solved using the Kalman algorithm, which provides one-sided and two-sided estimates of the output gap (i.e. real-time and ex-post estimates) — Kalman (1960).

The value represented by the choice of lambda is the only explicit judgement associated with the HP filter and is one subject to significant debate in the literature. The authors, Hodrick and Prescott, posit that a value of 1600 is appropriate for quarterly data reflecting the view that this tallies with their subjective assessment of the US business cycle. Pedersen (2001), for example, argues that a value of 1000 is better. Of course, there is no reason why the UK business cycle must be characterised by the same degree of persistence as in the US or that the business cycles today should typically be of the same length as in the past. And the appropriate value of lambda (which determines the stiffness of the filter) will vary with the time-horizon of the output gap being measured.

While it is the choice of lambda that often gets the most attention, the HP filter is also consistent with a number of other implicit assumptions and judgements:

- the data-generating process governing the evolution of potential output is assumed to be a random walk with stochastic drift;
- from the above, the best indicator of tomorrow's potential output growth is today's potential output growth;
- the output gap is an independent and identically distributed random variable, so the best guess of tomorrow's output gap, given today's output gap, is zero (no expected persistence);[1] and
- shocks to demand are not correlated with shocks to supply.

A number of these implicit judgements are important. Of particular significance, and not too well understood by most practitioners is that trend growth is assumed to vary at random rather than tend towards a specific value. This distinction sounds arcane but it extremely important - it materially affects the end-point properties of the filter. Hamilton (2017) goes as far as to suggest the HP filter should never be used.

The assumption that demand shocks are uncorrelated with supply is also contentious, since it rules out the possibility of periods of protracted cyclical weakness permanently lowering the level of output. DeLong *et al.* (2012), for example, take a different view — they posit that a large negative output gap can have very persistent effects on the level of potential output, via hysteresis effects in the labour market or reduced investment lowering the capital stock, for example. The IMF IMF (2012a), in its advice to the UK government, used an estimate that a negative output gap of 1 per cent might lower the level of potential output by around 0.1 per cent a year and found that much of the effect in the UK is accounted for by labour market hysteresis.

The way in which hysteresis effects influence estimates of the output gap depends on the time horizon of the output gap measure in question. For example, hysteresis effects in the labour market will tend to push up the rate of unemployment consistent with stable inflation in the medium term, but may affect the long-term structural rate by less. So estimates of the output gap on a medium-term basis would be likely to report less slack in the economy than those aiming to capture a longer-term measure of spare capacity. Hysteresis effects are also important to forecasts of potential output, although that is not the subject of this paper.

At the end of the sample, there are no future data available to assist with estimating the current level of potential output — the so-called end-point problem. All two-sided filters suffer from this problem and are revised once future data become available. Filters

[1] In practice, persistence of the output gap is introduced via the smoothing parameter - similar results could be obtained by instead assuming the output gap is an autoregressive process and calibrating the parameters of that equation.

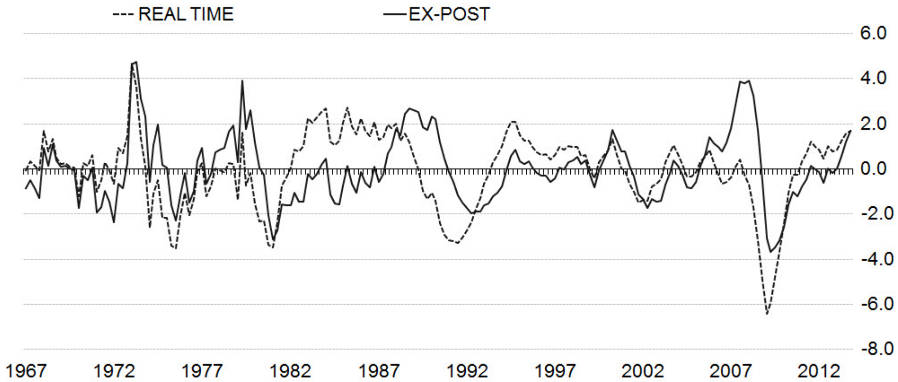


Figure 3.3: Figure shows real-time and ex-post estimates of the output gap as a percent of GDP using the HP filter.

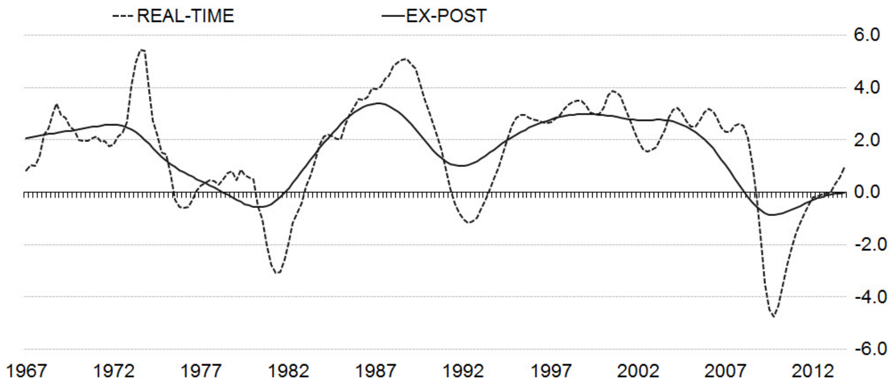


Figure 3.4: Figure shows real-time and ex-post estimates of year-over-year potential GDP growth using the HP filter.

respond differently to the end-point problem depending on the assumptions that underpin them. The specific assumptions underpinning the HP filter mean that, at the end of the sample, potential output growth tends to be biased down when the output gap is negative and biased up when it is positive. It is shown later that other filters may respond differently, but do not necessarily offer a more favourable balance of characteristics.

Figure 3.4 illustrates the HP filter estimates of the output gap (with lambda set to 1600) in both real-time and ex-post. Perhaps the most striking thing about this figure is that revisions tend to be largest around recessions and particularly over the most recent recession. In 2007, this measure would have said the economy was operating at around its trend level of output. It now says the economy was overheating by around 4 per cent of potential GDP — and that potential output growth began to slow in the early 2000s.

This specific reinterpretation reflects the loss function of the filter — because sharp movements in potential output growth are heavily penalised, the two-sided filter begins the slowdown ahead of the recession, to avoid making a very large adjustment when the crisis hits. The one-sided (real-time) version of the filter cannot see this coming so is forced to allocate the shocks to potential growth and the output gap as they become apparent. It is worth considering that the assumptions underpinning the HP filter are consistent with a specific view of the process governing potential output.

The average revision associated with the HP filter over the time period selected is around 1.4 percentage points (a little larger than its average absolute size of 1.1 per cent). The latest estimate of the output gap provided by the HP filter (i.e. those for the final quarter of 2013) do not change materially when the sample start point is shifted forward five years.

The distribution of actual output is consistent with booms gathering pace gradually but recessions being more abrupt. This real-time asymmetry can partly be ameliorated by placing more weight on estimates of past potential growth than the HP filter does, which can be achieved by altering the objective function of the filter.

3.2.3 Prior-constrained (PC) filter

Like the HP filter, the prior-constrained (PC) filter[1] is based on two beliefs — the first is the same but the second differs:

- output does not deviate too far from its trend level (cycles are not too big); and
- the growth rate of potential output does not differ too much from its historical average rate of drift.

$$\sum_{t=1}^{T}\left(\frac{1}{\sigma_1^2}(c_t)^2 + \frac{1}{\sigma_2^2}(\Delta y^*_{t+1} - \text{drift}_t)^2\right) \tag{3.9}$$

The PC filter chooses y_t^{*t} such that the loss function given by Equation is minimised. σ_1^2 is the variance of the output gap and σ_2^2 is the variance of trend growth deviations from its historical rate of drift. Like the HP filter, the user of the PC filter can specify the relative weight placed on the two beliefs by constraining the ratio of the two variance terms to be equal to a specific value, given by k — as shown in Equation 10.

$$k = \frac{\sigma_1^2}{\sigma_2^2} \tag{3.10}$$

The parameter k is set to 625 in what follows — as a rough guide, this setting implies that shocks to the output gap are around five times as large as those to the level of potential

[1]The PC filter is applied by Benes & N'Diaye (2004) and Laxton & Tetlow (1992), for example.

output. The PC filter is estimated in state-space using the Kalman algorithm. To assist this algorithm in its search, it is helpful to set starting values for the unobserved states — these are set to be consistent with the output gap being closed at the start of the sample. In practice, this only affects estimates of the output gap very early in the real-time sample — both the real-time estimates at the end of the sample and the ex-post estimates across the whole sample are not sensitive to the choice of initial conditions.

The state-space representation of the PC filter is given by Equations 11 to 14.

$$\text{Signal} : y_t = y_t^* + c_t \tag{3.11}$$

$$\text{State} : y_{t+1}^* = y_t^* + \text{drift}_t + (\frac{\epsilon_{1,t}}{k^{1/2}}) \tag{3.12}$$

$$\text{State} : \text{drift}_t = \text{drift}_{t-1} \tag{3.13}$$

$$\text{State} : c_t = \epsilon 1, t \tag{3.14}$$

As with the HP filter, the smoothing parameter,k, is important, but the PC filter is also consistent with a number of other implicit assumptions and judgements:

- the data-generating process governing the evolution of potential output is assumed to be a random walk with constant drift;
- from the above, the best guess of tomorrow's potential output growth is the historical rate of drift;[1]
- the output gap is an independent and identically distributed random variable, so the best guess of tomorrow's output gap, given today's output gap, is zero (no expected persistence); and
- shocks to demand are not correlated with shocks to supply.

The assumptions outlined above are similar in nature to those of the HP filter, with the main difference being that the HP filter is consistent with stochastic drift (shocks occur both to the level of potential output and its underlying growth rate) while the PC filter is consistent with constant drift (shocks occur only to the level of potential output, not its growth rate). In principle, any prior over the drift term can be incorporated, including structural breaks to the growth of potential GDP.[2] The remaining assumptions are subject to the same criticisms as described in the HP section.

[1]The one-sided filter calculates the historical average rate of drift up to the current time period and updates this as it moves forward. The two-sided filter makes an estimate of the drift term over the whole sample.

[2]Incorporating a structural break during the recent financial crisis does not affect the output gap estimate, rather it changes the path of supply shocks. So this judgement is more important for forecasting potential GDP than it is for estimating the output gap.

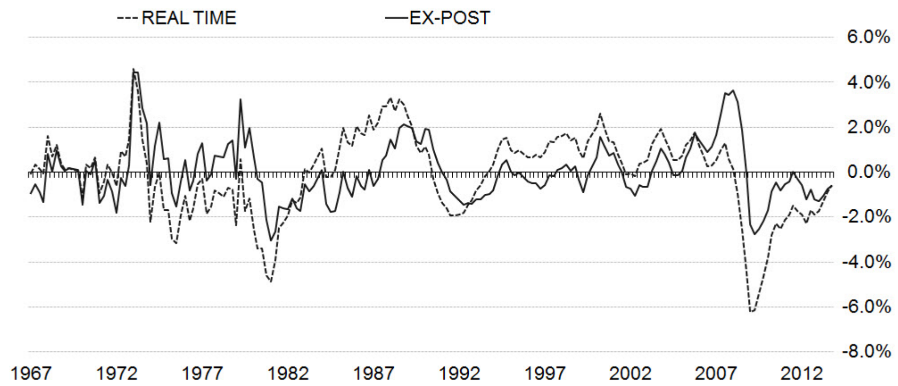


Figure 3.5: Figure shows real-time and ex-post estimates of the output gap as a percent of GDP using the prior-constrained filter.

With an appropriate choice of scaling parameter, the PC filter gives identical two-sided (ex-post) estimates of the output gap to those provided by the HP filter but it provides different real-time estimates (at the end of the sample). Both filters choose potential output such that their loss functions are minimised. The HP filter is based on a more flexible assumption for the dynamics of potential output than is the PC filter. So minimising the first part of its objective function (closing the output gap) is easier, because it is less costly to do so by adjusting its estimates of potential output growth.

It is important to recognise that the assumptions underpinning different filters significantly affect their real-time properties. The HP filter is far more likely to signal a closed output gap at the end of the sample and be revised subsequently than is the PC filter — a feature thought by many to be undesirable. But, by placing more weight on past growth as a guide to future growth, the PC filter is slower to respond should there be structural breaks in the growth rate of potential output, and subsequently be revised for this reason — again, an undesirable property. Because of this, the real-time estimates of different filters might be more reliable at different times and care should be taken to consider other evidence.

Like the HP filter, the PC filter penalises volatility of trend growth so revisions to its output gap estimates, with the benefit of hindsight, imply more overheating in the economy before the recent financial crisis and slower trend growth in the years that preceded it. The key distinction between the real-time HP and PC filters is that the latter is more likely to interpret growth rates above the historical average as being unsustainable in real time than would the former. With the benefit of hindsight, the PC filter estimates are very close to those of the HP filter.

Starting the sample later has a small effect on an estimate of the output gap now -

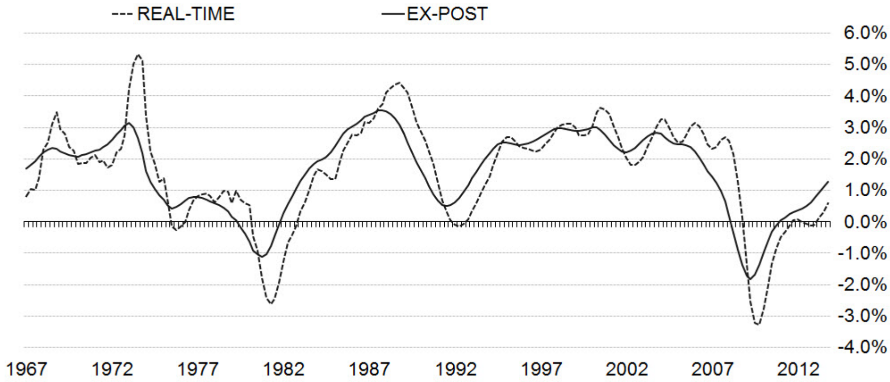


Figure 3.6: Figure shows real-time and ex-post estimates of year-over-year potential GDP growth using the prior-constrained filter.

shifting the sample on five years alters the current output gap estimate by 0.1 percentage points, for example. Largely this is because the filter uses estimates of historical trend growth to inform its estimates, which is affected by the choice of sample period.

3.2.4 Beveridge-Nelson (BN) decomposition

The trend-cycle decomposition of Beveridge & Nelson (1981) presents output as an autoregressive integrated moving average (ARIMA) process. They postulate that the permanent component of the series is equal to the long-run forecast of output, taking into account its mean rate of change — which identifies the trend and cycle components of output.

The BN decomposition depends upon a number of assumptions:

- output growth is stationary;
- the trend is equal to the long-run forecast of the series;
- both trend and cycle are affected only by a common shock; and
- the ARIMA specification is correct.

The most important assumption is that all movements in the trend and cycle components of output are driven by a common (unidentified) shock. The shock to potential output is assumed to be negatively correlated with the cyclical shock — when the shock pushes potential output up, it pushes aggregate demand down. This is a fairly restrictive assumption but one possibility is that the shock could be accounted for by movements in productivity — an interpretation which could be consistent with a 'real business cycle' view of the world.[1] The results are also sensitive to the specification of the ARIMA model - Canova (1998) shows that the inclusion of more or fewer lags can greatly influence the resultant output gap estimates.

I estimate the output gap using an ARIMA (2,1,0) specification and the estimates are presented in Figure 3.11. The combination of assumptions described above gives an output gap that is generally of smaller amplitude than its comparators, while the Beveridge-Nelson estimates of potential output are more volatile than actual output.

3.2.5 Christiano-Fitzgerald (CF) filter

The Christiano-Fitzgerald (CF) filter is a band-pass filter, formulated in the frequency domain. It works by filtering out data according to its frequency, decomposing a time series into trend, cycle and noise. In what follows, anything with a frequency below two years is considered noise, between 2 and 8 years is cycle and over 8 years is trend. This is a typical convention, but, like the HP filter, there is no strong evidence to bring to bear on the choice of cut-offs, so this choice is a judgement.

The CF filter makes use of the entire sample to estimate the cycle and is subject to the same end-point problem as the other filters — the absence of future data makes it prone to revision. However, Nilsson & Gyomai (2011) find that the CF filter revisions tend to be a little smaller than those for the HP filter. The cost, though, is that it is less likely than the HP filter to pick up signals of turning points. They judge that the HP filter is more appropriate to the OECD's short-term forecasting needs than is the CF filter.

Furthermore, Estrella (2007) compares the performance of a range of univariate filters and finds that the HP filter performs best but, crucially, only in cases when its assumptions are consistent with the true process being examined. As with the BN decomposition, because the CF filter has no state-space representation, real-time gap estimates are not presented, but the ex-post estimates are included in comparison Figure 3.11.

3.2.6 Hamilton filter

The HP filter depends on very restrictive assumptions about the underlying data generating process for trend output and the cycle. Yet the Hamilton filter also embodies assumption which may not match the process driving GDP either. Hamilton (2017) illustrates this fact and shows that the HP filter is ill-suited to the wide range of problems to which it is often applied by practitioners of econometrics. A better solution, Hamilton says, is to use a simple regression model to separate trend from cycle. The logic being

[1] See Kydland & Prescott (1982) for a description of real business cycle theory.

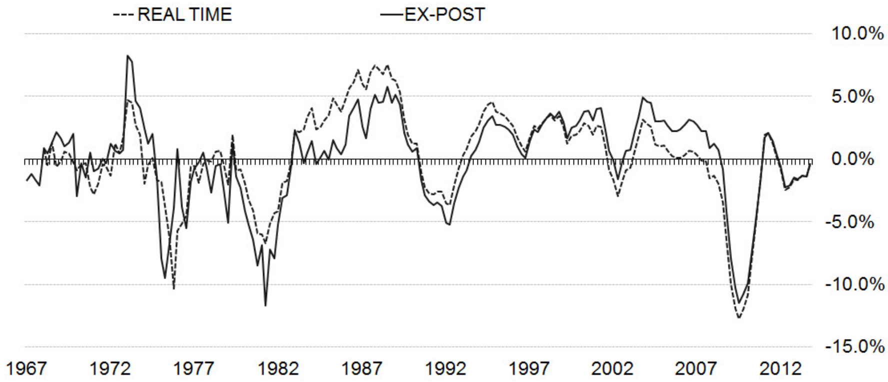


Figure 3.7: Figure shows real-time and ex-post estimates of the output gap as a percent of GDP using the Hamilton filter. Real time estimates are produced with a rolling regression, ex-post estimates make use of the full sample

that the difference between the realised values and forecasts for the variable in question should be a good proxy for cyclical deviations.

Hamilton proposes estimating a regression model of the following specification for quarterly data:

$$y_t = c + \beta_1 y_{t-8} + \beta_2 y_{t-9} + \beta_3 y_{t-10} + \beta_4 y_{t-11} + \epsilon_t \tag{3.15}$$

Where the difference between the predicted value of output less the actual value is the cyclical component of the series being detrended. This specification may address some of the problems of using the HP filter and perhaps it is the best method for application to a wide range of problems - I cannot say. But when it comes to estimating the output gap in the United Kingdom, it doesn't seem to offer huge advantages over other methods.

Implicitly, the trend rate of growth is determined by the estimates of the coefficients on lagged output and the constant. Implicitly, then, it is assumed when applying this filter that trend growth in the future will be the same as that estimated on average over the past. That differs to the HP filter but it is fairly similar to the assumptions embodied in the specification of the PC filter.

Because the estimated parameters will vary as new data arrive, this means the Hamilton filter estimates are also subject to revision. As Figure 3.7 shows, the output gap revisions are material when applying the Hamilton filter. Crucially, on the eve of the financial crisis, the Hamilton filter would have said that the economy was running with a margin of slack, while later vintages would instead indicate some overheating - the same as would have been recorded by the HP and PC filters.

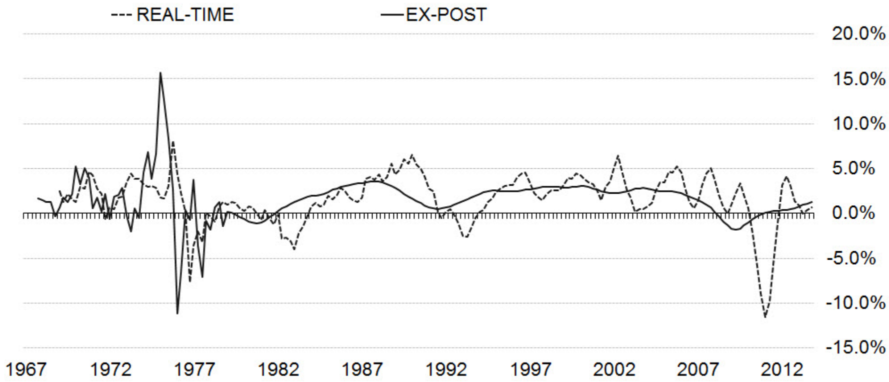


Figure 3.8: Figure shows real-time and ex-post estimates of year-over-year potential GDP growth using the Hamilton filter.

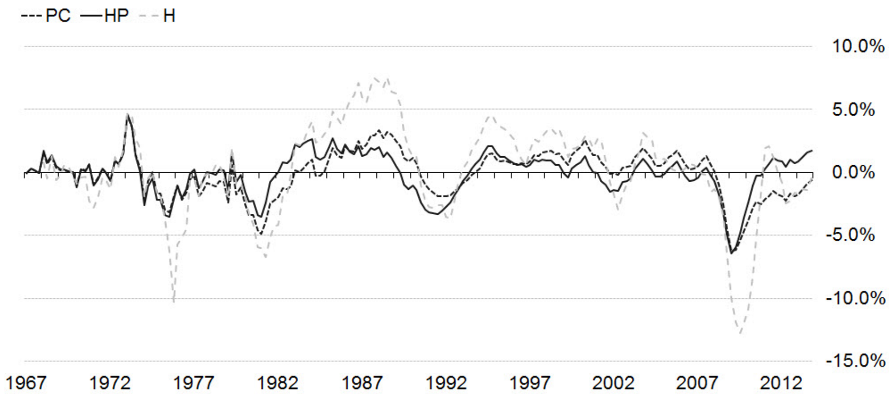


Figure 3.9: Figure compares real-time output gap estimates made by the HP, PC and H filters - percent of GDP

3.2.7 Comparison of univariate methods

The three univariate filters estimated in real time are the HP, PC and Hamilton methods, illustrated in Figure 3.9 and Figure 3.10. It is clear to see that potential growth is anchored more closely to the historical average using the PC filter than with the HP filter, though the broad pictures painted by both are similar. Potential growth is substantially more volatile when estimated using the Hamilton filter.

The resulting output gap series illustrate the specific point that the HP filter assumptions lead to bias at the very end of the sample. The very large negative output gap serves to bias down potential growth and so the HP filter persistently estimates a rapidly closing or even positive output gap at the end of the sample — most obvious during the present recovery period. Again the Hamilton filter produces an estimated output gap of bigger

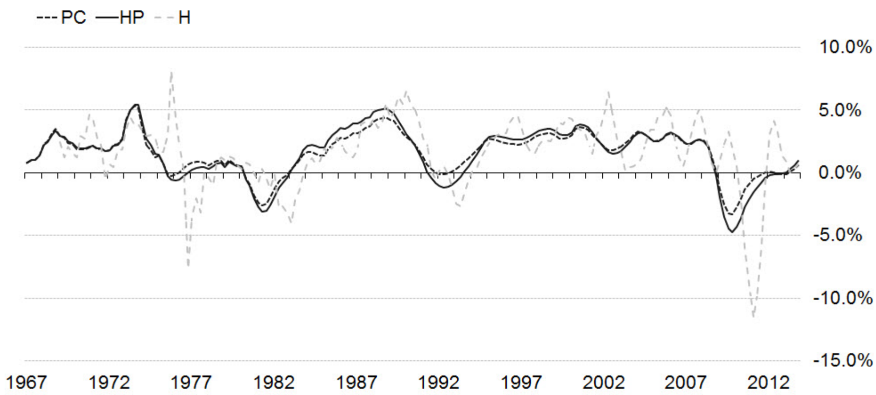


Figure 3.10: Figure compares real-time potential growth estimates made by the HP, PC and H filters

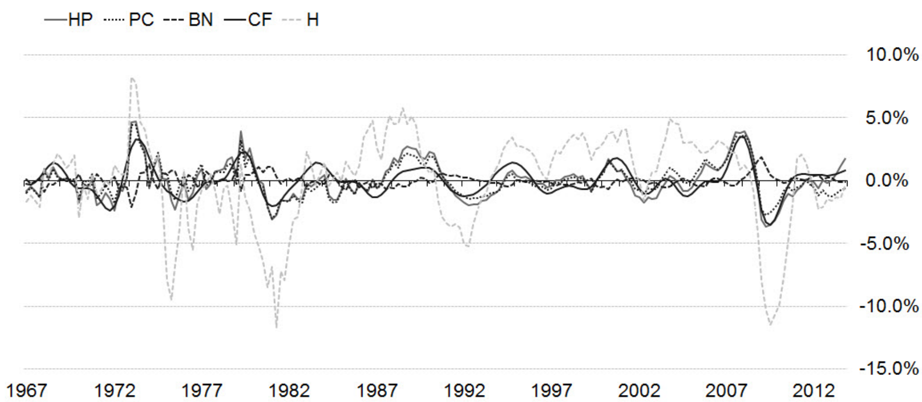


Figure 3.11: Figure compares ex-post output gap estimates made by the HP, PC, H, BN, and CF filters - percent of GDP

amplitude during the last recession in the sample.

Crucially, on the eve of the financial crisis the Hamilton filter would have said that the economy was running with a margin of slack, while later vintages would instead indicate some overheating - the same as would have been recorded by the HP and PC filters. In other words, whether the Hamilton, HP or PC filter had been deployed in mid-2008, policymakers would not have drawn vastly different conclusions.

The ex-post estimates of the output gap and potential growth include the BN decomposition and CF filter — presented in Figure 3.11 and Figure Figure 3.12. Most striking is how the different assumptions lead to such a wide range of estimates. Clearly, the BN decomposition is a different class of model altogether, consistent with significant potential output volatility (more so even than actual output). The Hamilton filter also produces

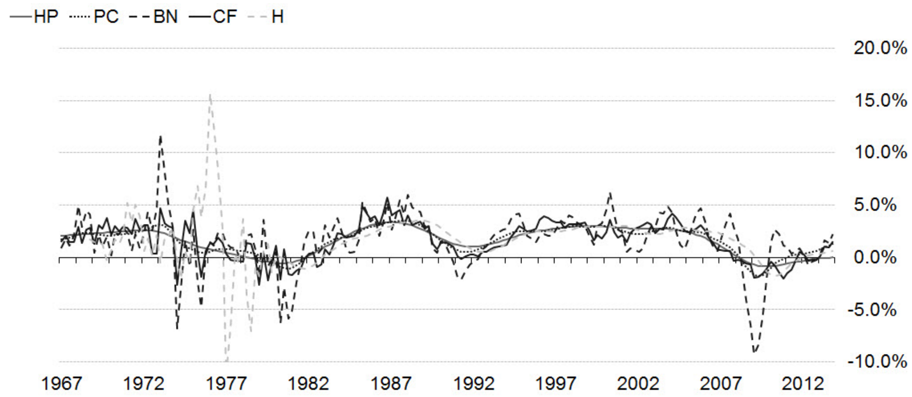


Figure 3.12: Figure compares ex-post potential growth rate estimates made by the HP, PC, H BN and CF filters.

very volatile estimates.

The PC and HP filters present similar estimates (except at the end of the sample, when the ex-post estimates converge with the real-time estimates). The CF filter is consistent with a smoother output gap series but more volatile potential output (since the volatility of actual output — the noise identified by the filter — must be assigned somewhere), although a different set of cycle-length assumption would give a different set of estimates.

In summary, the results show that estimates of the output gap obtained using univariate filters are very sensitive to the assumptions underpinning them. As Estrella (2007) notes, filters must be carefully selected for any particular application and it seems likely that no single method would accommodate all circumstances well.

3.3 Multivariate methods

This section covers a number of methods which make use of more than one variable. It starts by setting out a suite of models based on the multivariate PC filter, which is adapted from the multivariate HP filter presented in Laxton & Tetlow (1992). It then presents alternative methods including principal components and an aggregate composite of survey indicators.

Ultimately, the univariate filters set out in the previous section rely on judgement over the amplitude of the cycle and strong priors over the dynamics of potential output. To refine those judgements, we need more information. Often, though, taking on more information requires explicit assumptions over how it should influence estimates of the output gap, as is shown below.

3.3.1 The multivariate PC filter

The standard multivariate filter, presented by Laxton & Tetlow, augments the objective function of the HP filter with the sum of squared residuals from another signal relationship. In what follows, the same approach is taken but the PC filter is used instead, to avoid the specific type of end-point bias associated with the HP filter. The objective function of the filter (15), therefore, looks much like (9), but also includes the sum of squared residuals, , from a relationship that includes the output gap - that is, the filter also chooses the path of the output gap that most improves the fit of the hypothesised relationship.

$$\sum_{t=1}^{T}(\frac{1}{\sigma_1^2}(c_t)^2 + \frac{1}{\sigma_2^2}(\Delta y^*_{t+1} - \mathrm{drift}_t)^2 + \frac{1}{\sigma_3^2}(\epsilon_{3,t})) \tag{3.16}$$

And so, the multivariate PC filter is based on three beliefs, the first two of which are common to the standard PC filter:

- output does not deviate too far from its trend level (cycles are not too big);
- the growth rate of potential output does not differ too much from its historical average rate of drift; and
- other indicators, for example unemployment, tell us something about the cyclical position of the economy.

While taking on information from other sources is likely to ameliorate, to some extent, the end-point problem, judgements are still required. Instead of one smoothing parameter to choose, there are now two. These are used to decide, first, the variability of potential output and, second, how much weight to place on the structural relationship of choice. In practice, this allows us to select a smoothness of potential output that is consistent with the time horizon appropriate for use by a fiscal authority. But it also allows us to make use of other information, which might be more relevant to a medium-term concept of potential output (such as inflation). The two parameters are given by:

$$k = \frac{\sigma_1^2}{\sigma_2^2} \tag{3.17}$$

and

$$\psi = \frac{\sigma_1^2}{\sigma_3^2} \tag{3.18}$$

While the scaling parameter in Equation 16 sets the variability of potential output relative to the sum of squared residuals of the output gap equation, the weight placed on other information, the specification and the fit of the structural equation of choice will also influence the overall cyclicality of the estimated output gap series. So, in what follows, the cyclicality of the resultant series is inspected relative to the baseline case and any major differences reported. The variability of potential output is tied down judgementally by setting k.

The weight placed on information from the structural relationship is determined by the size of ψ relative to k. As a baseline case, ψ and k are set equal to one another. Alternatively, judgement about how much weight to place on other information can be applied by adjusting the second parameter, given by Equation 17.

In what follows, the PC is augmented filter by one of three possible relationships,[1]:

- a Phillips curve;
- a version of Okun's law; and
- a capacity utilisation equation.

3.3.2 Philips curve-augmented PC filter

Inflation could contain useful information about slack in both the labour and product markets. First, tightness in the labour market could see the pass-through of higher wage demands to prices. Second, excess demand in the product market could affect firms' mark-up decisions. To make use of this information, the model presented below includes a reduced-form version of a structural relationship between inflation and spare capacity in the economy, known as the New Keynesian Phillips curve — similar to that presented in Gali & Monacelli (2005) and presented in Equation 18:

$$\pi_t = \beta_1 \pi^e_{t+1} + (1 - \beta_1)\pi_{t-1} + \beta_2 c_{t-1} + \epsilon_{3,t}. \tag{3.19}$$

π_t is the deviation of inflation from its steady state rate, π^e_{t+1} is the expected deviation one period ahead and π_{t-1} is lagged inflation. This formulation introduces persistence to the inflation process, such that it better matches the observed data. The neutrality of money is preserved by constraining the sum of the coefficients on future and lagged inflation to unity. The influence of the cycle, c_{t-1}, is captured by the coefficient β_2. In this model, inflation expectations are anchored to the steady state inflation rate, so the expected deviation of inflation from steady-state is zero (i.e. $\pi^e_{t+1} = 0$ in equation 3.19).

The chosen measure of inflation is CPI inflation, adjusted for the estimated effects of VAT measures (which are not directly related to the cycle) and the influence of food and

[1] In some ways, this method is similar to the MV method presented in Beneš *et al.* (2010) although here the model parameters are calibrated and weight is placed on the information, rather than using Bayesian methods to restrict the parameter space and specify priors over shocks to the measurement equations.

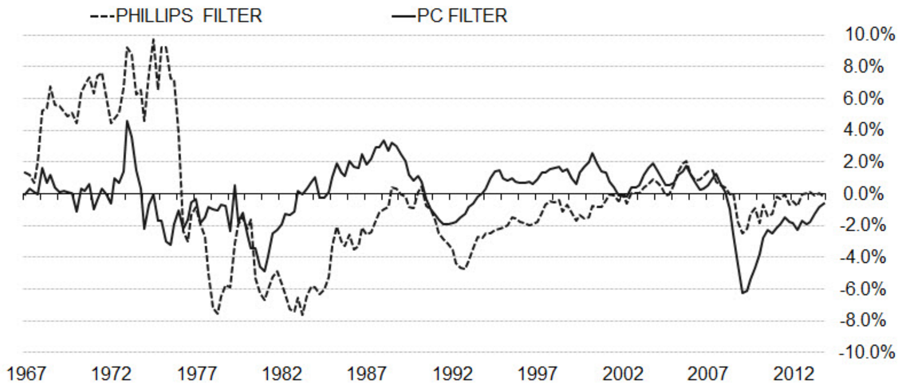


Figure 3.13: Figure shows real-time estimates of the output gap as a percent of GDP using the prior-constrained and Phillips-augmented filter.

oil costs, which are volatile and exogenous. The model parameters are calibrated in line with Murray (2012).

The steady state inflation rate is derived by applying the PC filter to the inflation series and the data-generating process is assumed to be a random walk without drift (the drift term is assumed to be zero in the objective function of the filter). Real-time estimates of the output gap use real-time estimates of steady-state inflation and likewise ex-post estimates are based on the ex-post steady-state estimate. The state-space model is given by Equations 19 to 23.

$$\text{Signal} : y_t = y_t^* + c_t \tag{3.20}$$

$$\text{Signal} : \pi_t = (1 - \beta_1)\pi_{t-1} + \beta_2 c_{t-1} + (\frac{\epsilon_{1,t}}{\psi_\pi^{1/2}}) \tag{3.21}$$

$$\text{State} : y_{t+1}^* = y_t^* + \text{drift}_t + (\frac{\epsilon_{1,t}}{k^{1/2}}) \tag{3.22}$$

$$\text{State} : \text{drift}_t = \text{drift}_{t-1} \tag{3.23}$$

$$\text{State} : c_t = \epsilon 1, t \tag{3.24}$$

The real-time and ex-post output gap series estimated using the Phillips curve to inform the judgement are compared with those of the standard PC filter in Figure 3.13 and Figure 3.14. The real-time estimates illustrate the uncertainty associated with real-time estimates of the trend inflation rate and, to some extent, the challenge of removing exogenous shocks from the series — particularly oil price shocks in the 1970s. The ex-post estimates look more sensible, since the trend inflation rate is known with more certainty over the difficult periods.

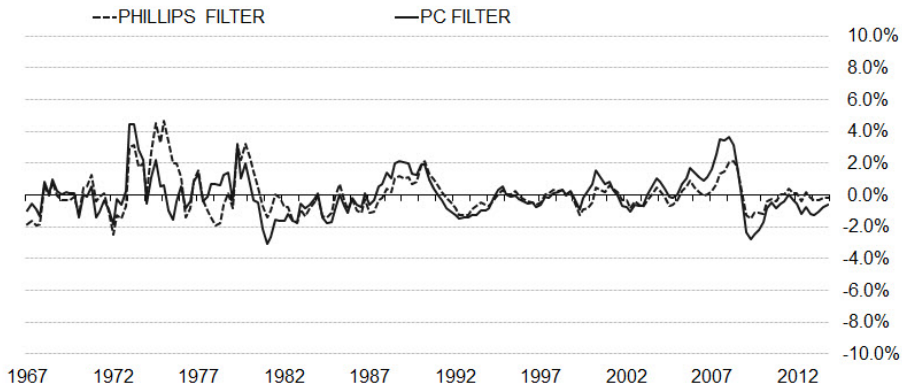


Figure 3.14: Figure shows ex-post estimates of the output gap as a percent of GDP using the prior-constrained and Phillips-augmented filter.

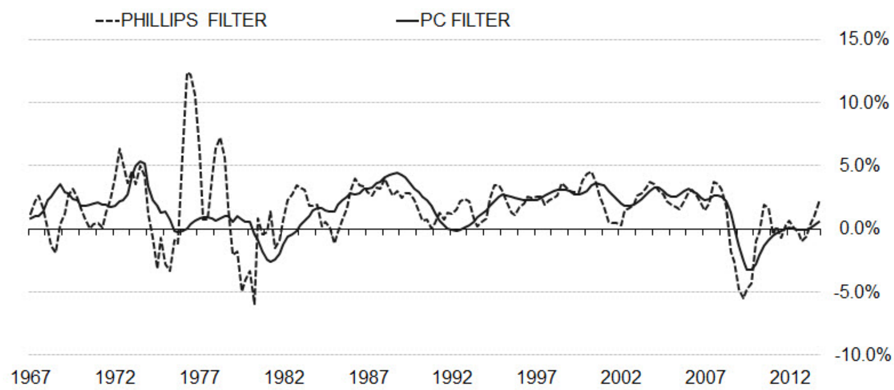


Figure 3.15: Figure shows real-time estimates of the year-over-year potential growth rate using the prior-constrained and Phillips-augmented filter.

When inflation is taken into account, the late-1980s boom looks a bit smaller than the PC filter implies ex-post. The low and stable inflation environment from the late 1990s serves to anchor the Phillips estimates of the output gap close to zero, but the inclusion of inflation tells us little about the output gap over that period on an ex-post basis. Overall, the ex-post estimates from the Phillips curve are relatively close to those of the basic PC filter, but the weakness of inflation before the last crisis and absence of strong disinflationary pressure afterwards tends to reduce both the estimated size of the boom and bust.

The additional information from inflation tends to increase the volatility of trend growth, relative to the estimates provided by the basic filter. This is because less weight is placed on minimising deviations from the historical rate of growth, in order to take on

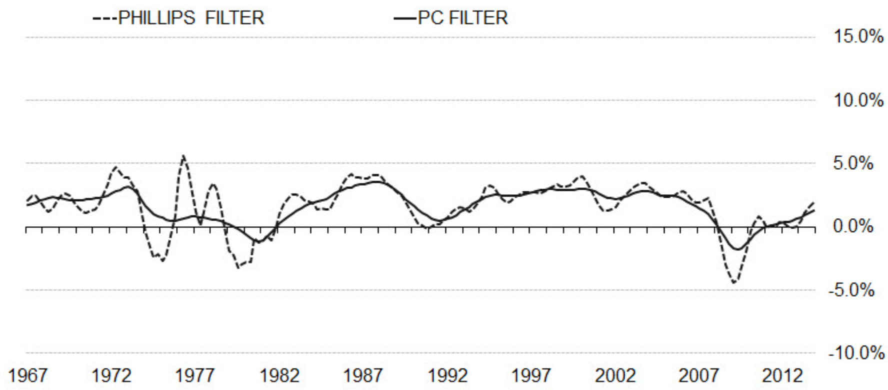


Figure 3.16: Figure shows ex-post estimates of the year-over-year potential growth rate using the prior-constrained and Phillips-augmented filter.

information from inflation. In extremis, a weight could be chosen that would make the output gap look exactly like the inflation deviation from steady-state and the potential growth rates would be much more volatile to reflect this.

The real-time estimates provided by this method are probably only reliable towards the end of the sample, when the monetary policy regime became more settled. The ex-post estimates are useful insofar as they help to incorporate information from inflation into estimates of the output gap while broadly preserving the features of the cycle. But incorporating such information requires a number of assumptions and implicit judgements, including but not limited to:

- the hypothesised Phillips relation being correctly specified;
- the calibrated coefficients of the Phillips curve being correct;
- stability of the relationship over time;
- the filter-based estimate of steady-state inflation being accurate; and
- the prior weight placed on its information content being appropriate.

These assumptions are highly uncertain - there is ongoing contention over how the Phillips curve should be specified - see Rudd & Whelan (2007) for example. The coefficients are poorly identified and circularity is introduced because they are typically estimated using output gap estimates. The relationship may not be stable over time - see Iakova (2007) for evidence on the flattening of the Phillips curve. The steady-state is estimated using a relatively stiff filter but there have been significant changes to the monetary policy framework over the sample period. And finally, given the uncertainties, it is unclear how much weight should be placed on the Phillips curve.

Neither changing the starting values nor shifting the sample on five years make a material difference to the estimates of the gap at the end of the sample period. Partly, the small difference reflects the fact that the key parameters are calibrated rather than estimated so model uncertainty does not have an influence when the sample is changed. The average output gap revision is around 2.6 percentage points, which is very large, but this largely reflects uncertainty over the trend rate of inflation. By selecting a sample over which the monetary policy framework was stable (1995 onwards, for example), this falls to 0.8 percentage points.

3.3.3 Okun's law augmented PC filter

To take account of slack in the labour market, we can make use of another relationship — that between cyclical unemployment and the output gap. This simple relationship is set out in Okun (1963) and a version of it is presented in Equation (24), where is the cyclical deviation of unemployment from its natural rate and $\varepsilon_{4,t}$ is an i.i.d shock:

$$uc_t = \beta_3 c_t + \epsilon_{4,t}. \tag{3.25}$$

To ensure consistency across the various measures presented that use the unemployment gap as an indicator, the real-time and ex-post estimates of the structural rate (the NAWRU) are produced using a common methodology, described in the Annex. These measures of the NAWRU are used to inform the estimate of the output gap. Like the Phillips model, the Okun relationship is added to the state-space model of the PC filter by introducing the following signal equation:

$$\text{Signal}: uc_t = \beta_3 c_t + \left(\frac{\epsilon_{1,t}}{\psi_u^{1/2}}\right) \tag{3.26}$$

Figures 14 and 15 compare output gap estimates from the Okun-augmented PC filter with those from the basic PC filter. The Okun filter sees more slack in the late 1970s and early 1980s than does the PC filter, which accords with elevated unemployment over that period and ultimately, the ex-post estimates from the PC filter are revised closer to the Okun estimates. The late 1980s boom appears to be more pronounced in real-time, using the Okun filter and, again, the PC filter shows a sharper boom at the end of the 1980s ex-post.

Making use of the Okun relationship, the elevated rate of unemployment widens the estimated output gap in the period since the recent recession and makes the boom that preceded it look small. Overall, the results suggest that it is important to recognise such

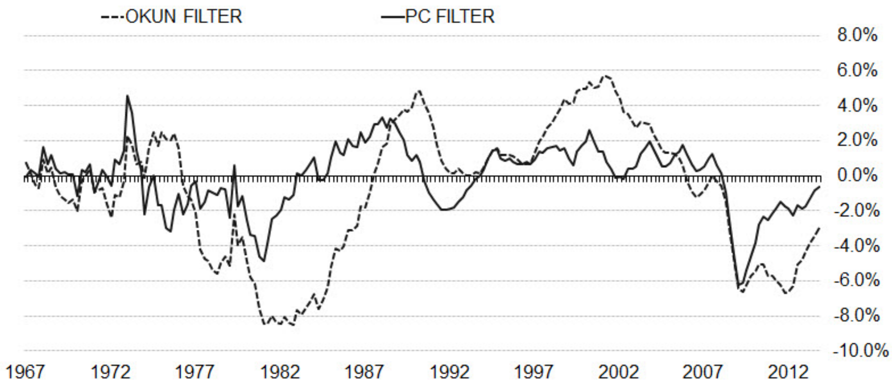


Figure 3.17: Figure shows real-time estimates of the output gap as a percent of GDP using the prior-constrained and Okun-augmented filter.

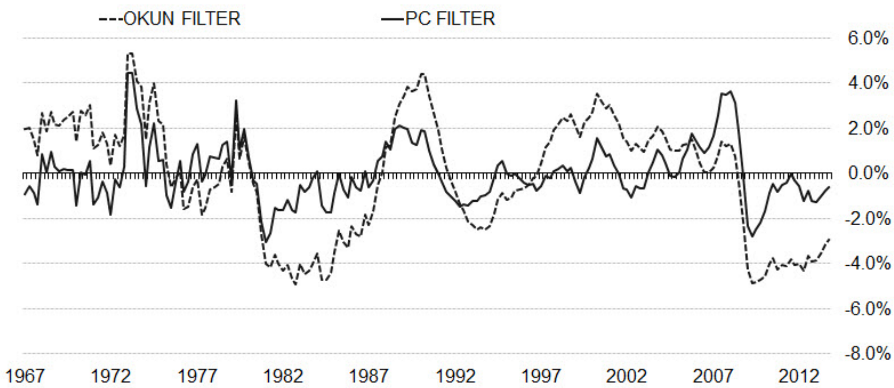


Figure 3.18: Figure shows ex-post estimates of the output gap as a percent of GDP using the prior-constrained and Okun-augmented filter.

an obvious source of spare capacity in the economy when forming a view of the output gap.

But, as with the Phillips relation, it is important to remember that the estimates depend heavily on a number of assumptions:

- the hypothesised Okun relation being correctly specified;
- the estimated coefficients of the Okun relation being correct;
- stability of the relationship over time; and
- the filter-based estimate of the NAWRU being accurate.

Again, these assumptions are highly uncertain. The Okun relation has various specifications and, while the coefficient of -0.5 is estimated over the whole sample period,

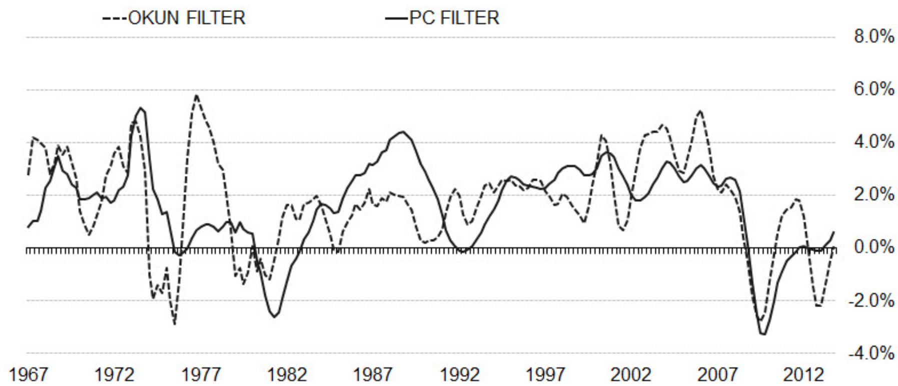


Figure 3.19: Figure shows real-time estimates of the year-over-year potential growth rate using the prior-constrained and Okun-augmented filter.

there is evidence to suggest that it may vary — the recessions of the 1980s and 1990s were associated with far larger increases in unemployment than was the latest recession. If the Okun coefficient were now lower, then the estimated output gap would be biased in a way that made it appear larger.

Moving the sample on five years has a small effect on the estimates of the current output gap (a bit less than 0.1 percentage points), while the estimates of the output gap are revised, on average, by around 1.3 percentage points — a bit less than the average magnitude of the output gap (around 1.7 per cent).

3.3.4 Capacity utilisation-augmented PC filter

The third augmentation relies on a posited, non-structural relationship between capacity utilisation indicators and the output gap. This is intended to capture slack within firms.

$$capc_t = \beta_4 c_t + \epsilon_{5,t} \tag{3.27}$$

This time, the signal equation is given by,

$$\text{Signal :} capc_t = \beta_4 c_t + (\frac{\epsilon_{1,t}}{\psi_{\text{cap}}^{1/2}}) \tag{3.28}$$

The capacity utilisation data are sourced from the Confederation of British Industry (CBI) and pertain to manufacturers only — services firms, which account for the bulk of activity, have not been surveyed over the long time series available for the manufacturing sector.[1] For simplicity, the weight placed on the equation in explaining the output gap is set equal to the smoothing parameter, as with the Phillips and Okun filters.

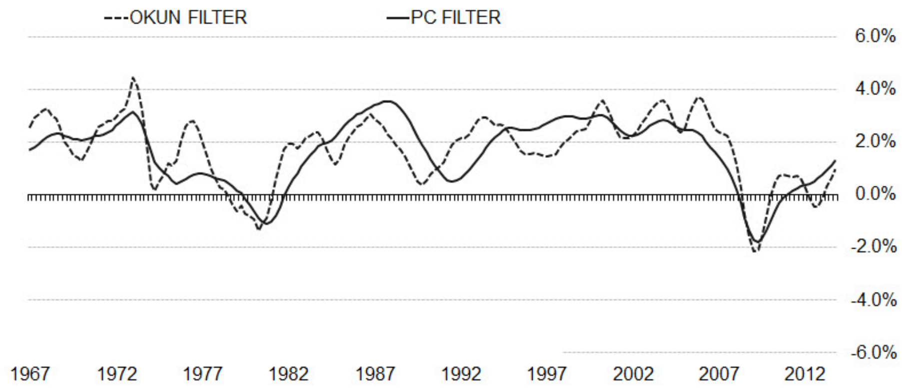


Figure 3.20: Figure shows rex-post estimates of the year-over-year potential growth rate using the prior-constrained and Okun-augmented filter.

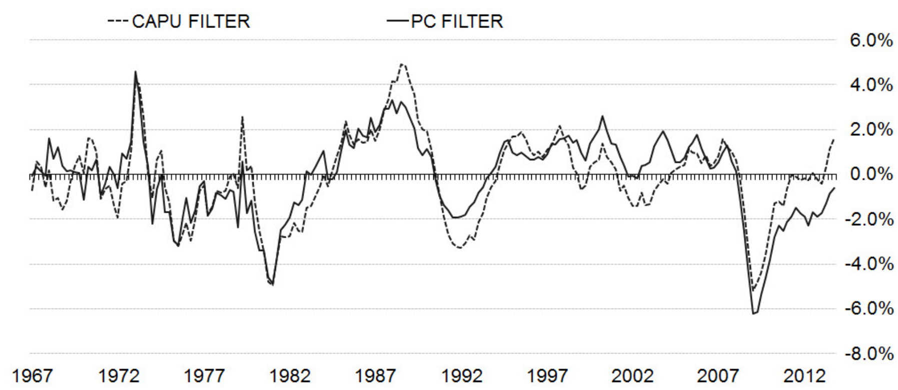


Figure 3.21: Figure shows real-time estimates of the output gap as a percent of GDP using the prior-constrained and capacity utilisation-augmented filter.

The inclusion of capacity utilisation serves to make the 1980s boom appear larger than it does in the basic PC filter. And, because capacity utilisation was close to usual levels before the latest recession, it makes the positive output gap smaller. It also points to the economy operating significantly above trend from 2010 onwards. Again, this model depends on a number of non-trivial judgements and assumptions:

- the hypothesised capacity utilisation relation being correctly specified;
- the freely-estimated coefficient being correct;
- stability of the relationship over time; and

[1]As with the other methods, the capacity utilisation data have been pre-filtered to extract trend and cycle - both in real time and ex post.

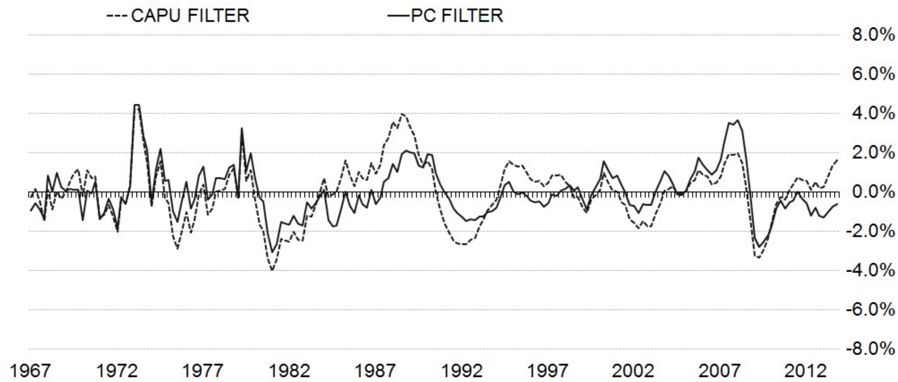


Figure 3.22: Figure shows ex-post estimates of the output gap as a percent of GDP using the prior-constrained and capacity utilisation-augmented filter.

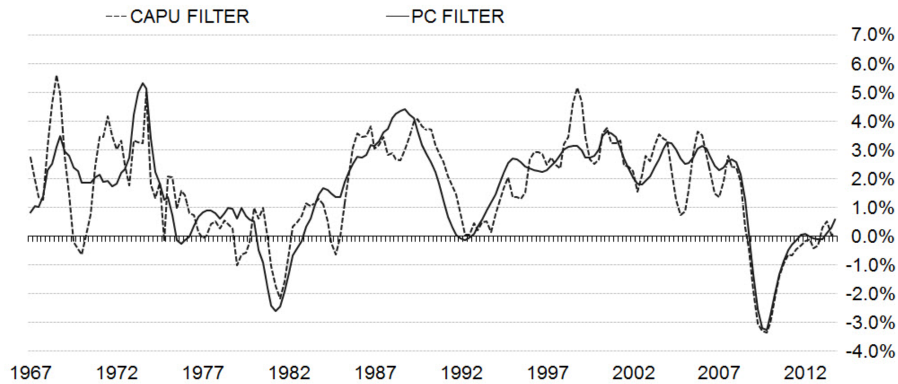


Figure 3.23: Figure shows real-time estimates of the year-over-year potential growth rate using the prior-constrained and capacity utilisation-augmented filter.

- the steady-state survey balance being correctly estimated.

Again, these assumptions are highly uncertain — there is no structural link between survey balances and the output gap, rather it should be treated as an indicator. It is also difficult to know exactly how respondents interpret the survey question. When asked, they may be thinking of a very short-term notion of spare capacity and ignore mothballed capacity that could be brought back online in the medium term. This may lead to biases in the short term that later unwind, but these would not be captured in the methodology described above. Furthermore, the survey captures the number of firms operating below/above capacity, rather than the extent to which firms are operating above/below capacity — the latter is what is relevant to output gap estimation.

Shifting the sample start date forwards by five years has a small effect on the size of

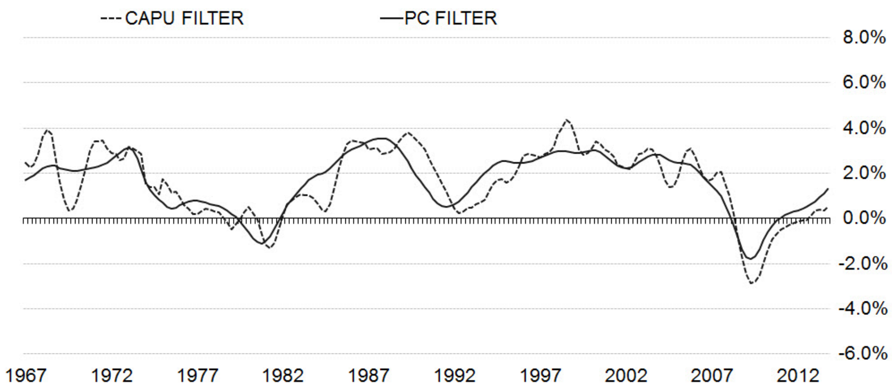


Figure 3.24: Figure shows ex-post estimates of the year-over-year potential growth rate using the prior-constrained and capacity utilisation-augmented filter.

the output gap estimate for the final quarter of 2013 — around 0.1 percentage points. The revisions are small because the capacity utilisation data are not revised and abstract from uncertainty over the steady state values.

3.3.5 Multivariate filter model

The final incarnation of the PC filter in this paper takes information from all three of the output gap relationships described above and forms a multivariate filter that is includes the following signal equations:

$$\text{Signal}:\pi_t = (1-\beta_1)\pi_{t-1} + \beta_2 c_{t-1} + (\frac{\epsilon_{1,t}}{\psi_\pi^{1/2}}) \tag{3.29}$$

$$\text{Signal}:uc_t = \beta_3 c_t + (\frac{\epsilon_{1,t}}{\psi_{\mathrm{u}}^{1/2}}) \tag{3.30}$$

$$\text{Signal}:capc_t = \beta_4 c_t + (\frac{\epsilon_{1,t}}{\psi_{\mathrm{cap}}^{1/2}}). \tag{3.31}$$

The parameter governing the volatility of potential output, k, is set to 625 but there are now three other parameters to select based on the weight placed on them in explaining the output gap. The weights chosen here are all equal — setting the parameters ψ_π, ψ_{cap} and ψ_{u} to 625.

Figures Figure 3.25 and Figure 3.26 illustrate the effect of including additional information from hypothesised relationships between the output gap and inflation, unemployment and capacity utilisation. The weight placed on additional information in forming a view on the output gap serves to reduce the end-point problem of the filter, insofar as the

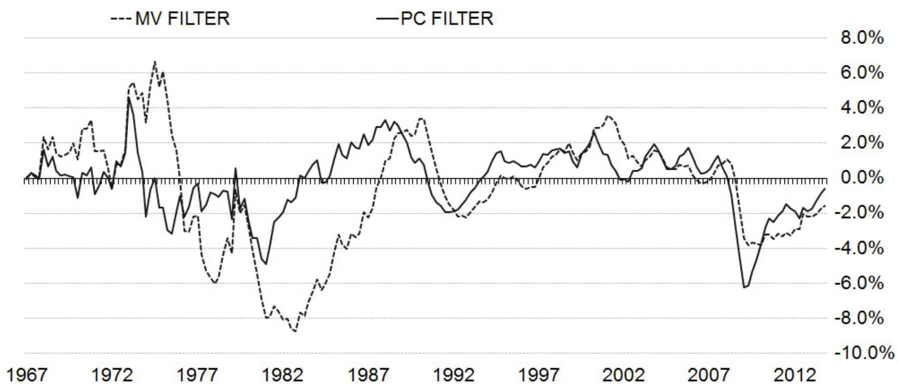


Figure 3.25: Figure shows real-time estimates of the output gap as a percent of GDP using the prior-constrained and multivariate filter.

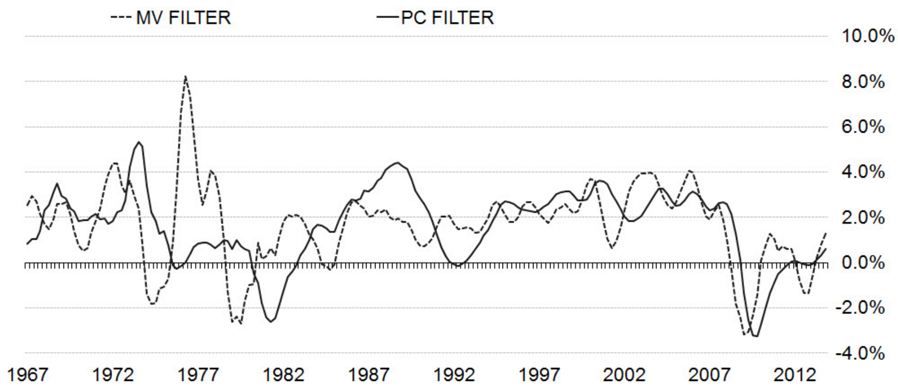


Figure 3.26: Figure shows real-time estimates of the year-over-year potential growth rate using the prior-constrained and multivariate filter.

MV model places less weight on GDP data from the future and so is less susceptible to revision when it comes to bear.

The most notable difference between the MV and PC estimates is that the size of the pre-crisis boom looks substantially smaller, because much of the data used to augment the filter were consistent with output being close to trend. The information does suggest, however, that output was further above trend in the late 1980s and the late 1990s than the basic filter would suggest.

Potential growth is somewhat more volatile using the MV filter, because less weight is placed on preserving its smoothness once other data are taken into account. In principle, a higher weight could be placed on minimising the deviation of potential growth from trend and this would serve to make the MV filter estimate closer to the PC estimate. Ultimately it is a matter of judgement as to how volatile potential output should be, since

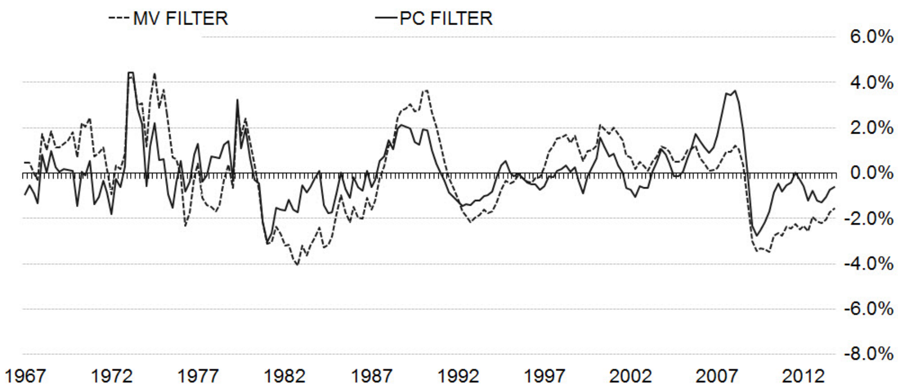


Figure 3.27: Figure shows ex-post estimates of the output gap as a percent of GDP using the prior-constrained and multivariate filter.

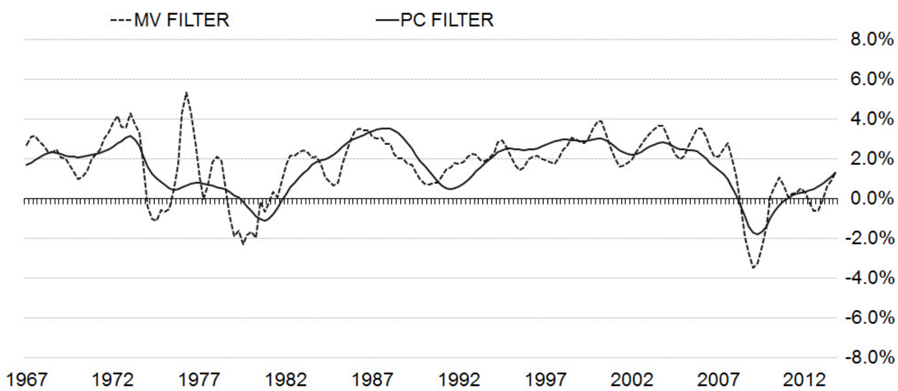


Figure 3.28: Figure shows ex-post estimates of the year-over-year potential growth rate using the prior-constrained and multivariate filter.

it cannot be observed.

While the MV filter makes use of more information, the judgements and assumptions required to make the most of it begin to add up, at the cost of reduced transparency. As well as the weights placed on the value of each relationship, their specification and underpinning assumptions all affect the resulting output gap series. So, in some sense, the weight placed on them is arbitrary. The myriad uncertainties and assumptions suggest that a method such as this, and variants of it, should be considered as part of a suite of indicators rather than providing a single point estimate.

3.3.6 Principal components analysis

Principal components analysis (PCA) is a statistical technique that attempts to draw a common signal from a range of de-meaned, standardised cyclical indicators - which can consist of survey measures of capacity utilisation, recruitment difficulties or inflation, for example. The estimation process involves assigning weights to each of the data series such that the resultant output gap series explains as much of the variability of the dataset as possible. A detailed treatment of the PCA can be found in Jolliffe (2002).

Once a principal component has been estimated, its mean and standard deviation (a scaling parameter) must be chosen. These choices are the main judgements upon which PCA estimates of the output gap depend. One method would be to set the standard deviation such that the amplitude of the business cycle accords with some prior judgement. Another would be to try to replicate the standard deviation of another output gap series.

Figure 27 illustrates the path of the output gap implied by the PCA, where the scaling parameter and mean have been set in line with the standard deviation and mean of the PC-filtered series (to maintain comparability). The PCA method estimates the output gap directly from the cyclical indicators, so it makes no effort to smooth trend growth - the additional volatility in trend growth relative to the PC filter is presented in Figure 3.30.

The PCA measures the output gap directly using data that are typically not revised, which means that the principal component it produces in real time is unlikely to differ very much from that which it estimates on an ex-post basis. However, the scaling parameter and mean might well change with the benefit of hindsight as these judgements are updated. It is difficult to generate a real-time series of estimates using the PCA method, but to get an indication of the scope for the PCA to be revised two vintages of its estimates are considered.

The figure shows the PCA estimate that would be generated in real time using data available up to and including the final quarter of 2007, with the standard deviation and mean set to those from the PC-filter estimate of the output gap up to that point. It also shows the PCA estimate using data up to the final quarter of 2013, with the standard deviation and mean set to those of the PC filter up to the same point. It is clear from the figure that the size of the pre-crisis boom now appears significantly larger than it did at the time, for example. So the PCA is not very sensitive to data revisions but is sensitive to changes in judgement over the scaling parameter and mean, with an average revision of around 1.2 percentage points between the two vintages shown here.

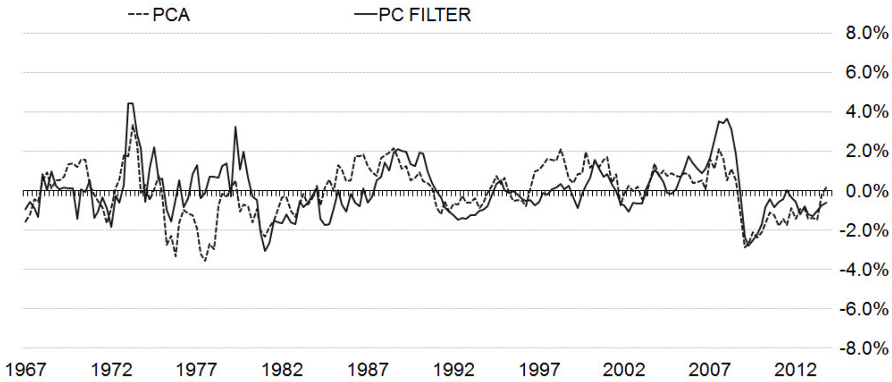


Figure 3.29: Figure shows ex-post estimates of the output gap as a percent GDP using the PC filter and PCA method.

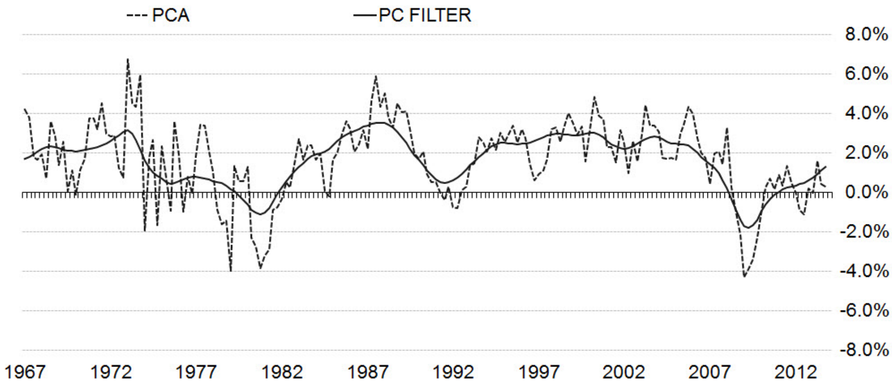


Figure 3.30: Figure shows ex-post estimates of year-over-year potential growth using the PC filter and PCA method.

3.3.7 Aggregate composite

The aggregate composite (AC) measure of the output gap uses similar data to that of the PCA, described above. But, rather than estimating the weight to be placed on each series based on the correlation of the dataset, these are set explicitly by the user. The aggregated composite is a weighted average of survey indicators of capacity utilisation and recruitment difficulties — where the weights are based on factor income and sector shares.

Like the PCA, the AC must be adjusted and scaled either subjectively or to match the mean and standard deviation properties of another output gap series. To construct the estimates presented in Figure 3.31 and Figure 3.32, the mean and standard deviation is set so that it is consistent with the PC filter estimates. The series is also spliced together

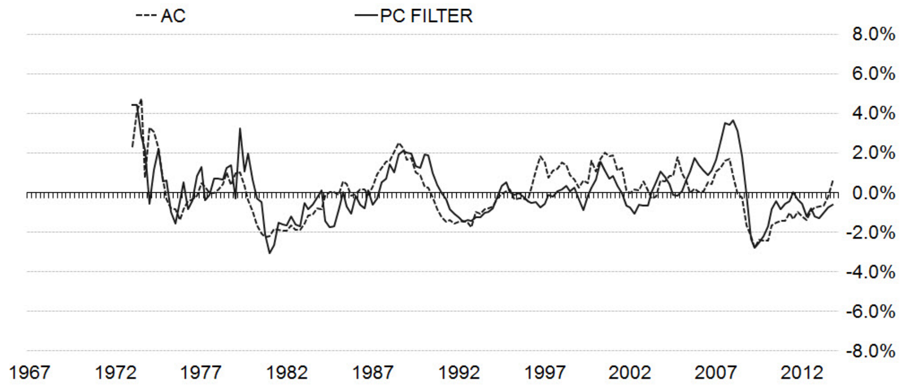


Figure 3.31: Figure shows ex-post estimates of the output gap as a percent GDP using the PC filter and AC method.

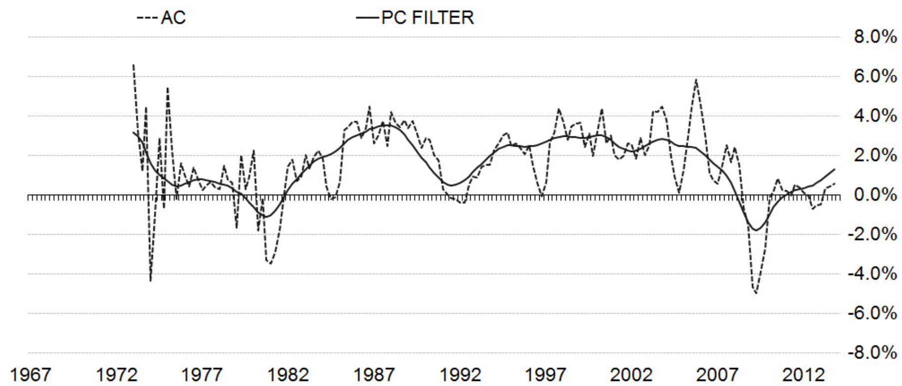


Figure 3.32: Figure shows ex-post estimates of year-over-year potential growth using the PC filter and AC method.

in the middle, since survey data for the services sector is unavailable prior to 1995 and the estimates are calculated using only manufacturing data prior to that point.

Overall, the results are fairly similar to the PCA measure of the output gap and potential output growth. Like the PCA, the AC does not place any explicit weight on smoothing potential output, but the amplitude of the cycle is constrained to be the same as the PC filter estimates. The revisions properties are likely to be very similar to the PCA.

3.3.8 Credit-Augmented Principal Components Analysis

Before the 2008 financial crisis, there was little evidence of overheating in the UK economy — wage growth, inflation and unemployment were all at rates consistent with historical averages. Yet, most analysts and commentators now agree that there was something unsustainable about the strong growth in the years before the crisis hit — most estimates of the output gap imply a significant and permanent hit to the level of potential productivity.

The period before the crisis saw significant growth of credit. To the extent that the financial cycle and the business cycle are correlated, financial variables could be used to inform estimates of the output gap. Indeed, output gaps augmented in this way may better-explain the cyclical behaviour of tax receipts, since some are related to asset prices and financial transactions. However, there are a number of practical problems in capturing the influence of credit.

First, credit can rise as a share of GDP for structural reasons — deregulation in the 1980s increased mortgage availability while falling real interest rates in the 1990s made mortgage payments more affordable, each boosting owner-occupation rates and so net mortgage lending. For this reason, the Basel Committee on Banking Supervision (BCBS) excludes some forms of lending in its preferred measure. However, even excluding loans secured on dwellings, loans for direct investment and derivatives, the liabilities of households, corporations and not for profit entities rose from rose from about 130 per cent of GDP in 2000 to around 170 per cent of GDP in 2007, before falling back during the recession.

Second, there is no direct relationship between credit and GDP (it is not a factor of production), and there is no structural relationship between the credit gap and the output gap. Instead it is treated as an indicator and how much emphasis to place on it requires judgement.

The BCBS-recommended credit gap measure, which the Bank of England publishes and the FPC uses to inform its policy decisions, is replicated in Figure 3.33. It is a real-time series produced using the HP filter with the smoothing parameter set to 400,000 — for further details see BOE (2014) and BaselCommittee (2010). The series indicates a significant financial boom in the late 1980s and a smaller one before the recent financial crisis. The ex-post series is also presented, which shares the same broad features as the real-time series but indicates that the 1980s credit boom starting somewhat later. It is also consistent with a bigger credit cycle in 2007/08 than it would have suggested using data available to policy makers at the time.

More broadly, this version of the credit gap does not tally especially well with conventional wisdom over the cyclical position of the economy. Because balance sheet movements are a relatively slow process, the credit gap is very persistent and recoveries can start long before leverage begins to rise. So a credit gap may help to inform judgements about the sustainability of growth and risk in the financial system but might be less useful as a real-time indicator of slack in the economy and, therefore, the scope for growth as

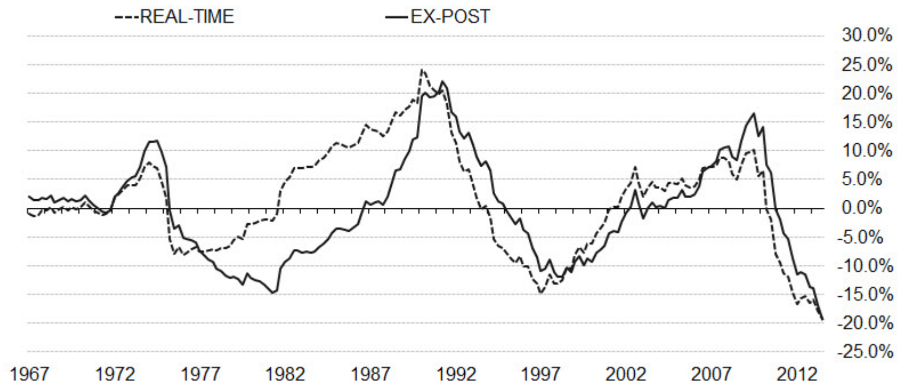


Figure 3.33: Figure shows Credit Gap as a percent of GDP

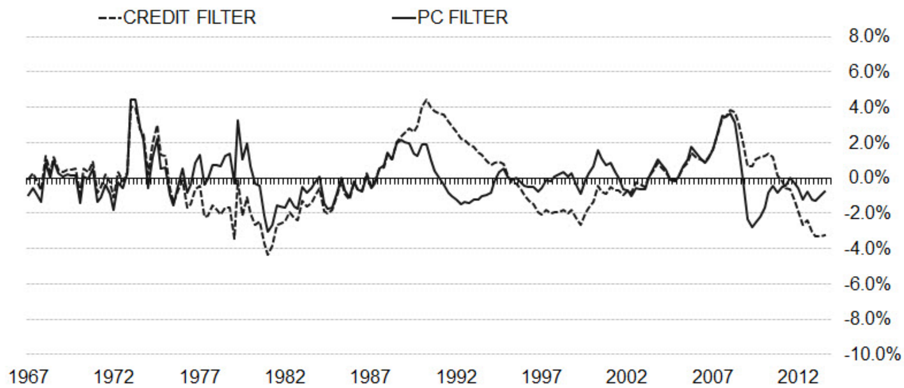


Figure 3.34: Figure shows ex-post estimates of the output gap as a percent GDP using the PC filter and credit filter.

it is taken up. This persistence is evident in the output gap estimates produced using an augmented PC filter, in real time and the revised, ex-post estimates.

An alternative method of taking on board information from the credit gap, when estimating the output gap, is to add it as a relevant indicator to a principal components analysis. However, PCA estimates are not drastically affected by the inclusion of the credit gap - largely this is because its low correlation with other variables means it is assigned a low weight when a common signal is extracted. There are still further ways in which to incorporate financial information into estimates of the output gap — see Borio et al (2013), for example — but each must grapple with the differing frequencies of the financial and output cycles.

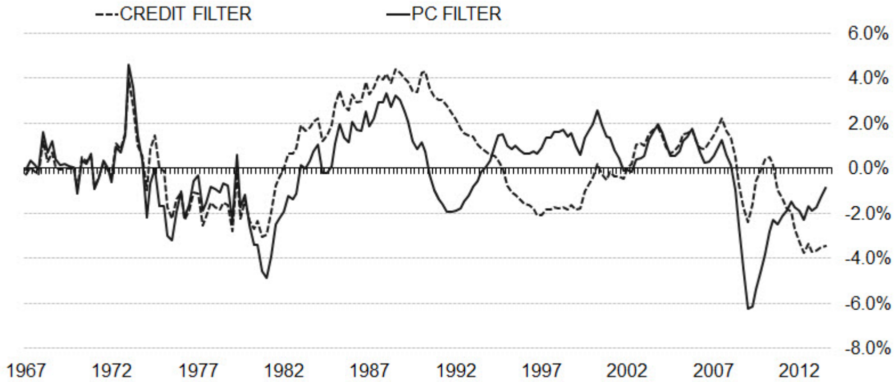


Figure 3.35: Figure shows real-time estimates of the output gap as a percent GDP using the PC filter and credit filter.

3.3.9 Production function approach

A production function is simply an equation that relates inputs to the production process to outputs. That is to say, the level of potential output is a function of labour supply, the capital stock and the maximum efficiency with which they can be combined (total factor productivity (TFP)). To estimate the output gap, the actual level of output is compared with this potential level.

There are two key judgements associated with production function estimates of the output gap. The first is the choice of production technology — an economic theory that relates the inputs to the outputs. How might output be expected to change with the addition of another unit of capital or labour, for example? The second is the choice of method for estimating the potential levels of the factor inputs and total factor productivity, which are to be aggregated using the specified production technology.

The production technology can take many forms, but, in what follows, the baseline Cobb-Douglas approach is taken, Douglas (1948), and the standard assumptions are used. It is assumed that the production process is characterised by a function of the form:

$$y = A(L^{\alpha}K^{(1-\alpha)}) \tag{3.32}$$

where A is TFP, L is labour input, K is the capital stock, α and $(1-\alpha)$ are the output elasticities of labour and capital, respectively. This particular function is consistent with a number of assumptions:

- constant returns to scale;
- the marginal productivity of each factor being proportionate to its average productivity;

- technology being Hicks neutral - technological improvements increase the returns to labour and capital in equal measure; and

- under the additional assumption of perfect competition in the product market, factors are paid their marginal products, the steady-state labour share of income is, therefore, stable and can be used to calibrate the elasticity of factor inputs with respect to output.

The functional form of the Cobb-Douglas production function is largely one of convenience, since it is easy to work with. An alternative would be the, more general, Constant Elasticity of Substitution (CES) production function, which could be consistent with a trending labour share of income, for example.[1] However, as Miller (2008) shows, there is little evidence that supports the use of one over the other when the objective is to forecast GDP or factor shares, provided the labour share of income is stable. Absent cyclical fluctuations, the UK labour share of income appears broadly stable over the past few decades, so the more tractable Cobb-Douglas specification is used.

Now that a production technology has been selected, it is necessary to estimate the trend series for labour, capital and TFP. Almost all practitioners of this approach do this by using filters of some sort, so it stands to reason that the estimates it produces share many similarities with those obtained via the methods described earlier in this paper (rather than constituting an estimation technique in its own right).[2] It does, however, enable the user to interrogate the estimates with economic theory, by decomposing the output gap into contributions from TFP and the factors of production.

In what follows, labour supply is defined as potential hours worked and comprises potential average hours worked and potential employment. Potential employment, itself, consists of potential participation in the labour market as a share of potential population less the structural rate of unemployment - the NAWRU.

To estimate the trend series used to construct the production function, I:

- assume the population gap is zero at all times by setting potential equal to actual;

- apply a stiff PC filter to the activity rate to identify potential participation;

- take the NAWRU estimates used by the MV and Okun-augmented PC filters — estimated using the PC filter and a wage equation;

- apply a stiff PC filter with drift to actual average hours to obtain potential average hours;

[1] Barnes *et al.* (2008) finds evidence using firm-level data that would support the use of a CES production function.

[2] The European Commission's approach is set out in D'Auria *et al.* (2010), for example.

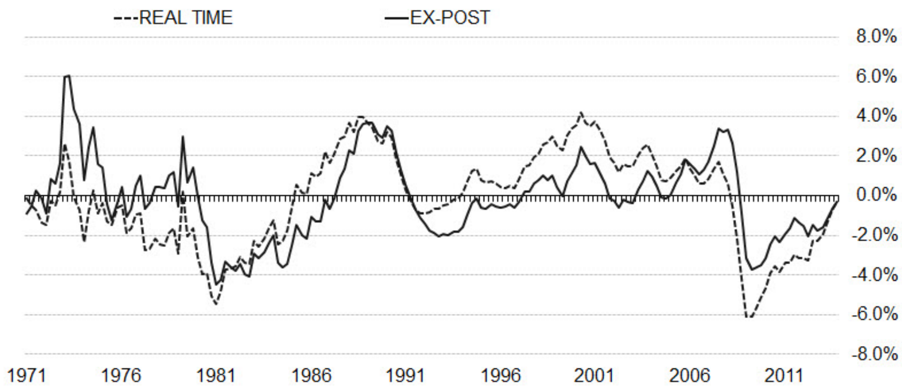


Figure 3.36: Figure shows real-time and ex-post output gap estimates from a production function as a percent of GDP

- assume the potential capital input is equivalent to utilising the existing capital stock; and
- apply a PC filter with drift to the level of actual TFP to obtain potential TFP.

Clearly, the assumptions underpinning the production function estimates of potential output are numerous and susceptible to exactly the same sorts of issues identified in earlier filter applications. And any number of filter or specification choices could be made that could be expected to yield different estimates. The output gap is estimated by comparing the level of potential output with the actual value. Furthermore, the actual values of the factor inputs and TFP can be compared with their potential values, taking the production technology into account, to give contributions to the output gap from each — the principal purpose of the exercise.

Figure 3.36 and Figure 3.37 illustrate the output gap and potential output growth paths produced by the production function approach. In many ways, they are unremarkable since they look similar to estimates produced by some of the methods described above. But the advantage of using a production function estimate is that we can decompose the output gap into contributions from the factor inputs and TFP.

Figure 3.38 and Figure 3.39 illustrate two approaches to decomposing the output gap into contributions. The first is perhaps the most intuitive, since it breaks the gap down into straight-forward contributions from employment, average hours and labour productivity deviations from trend.

The figures shows that the estimate of the output gap is made up of a negative contribution from employment and a smaller contribution from labour productivity lying below their potential levels, partly offset by a positive contribution from average hours exceed-

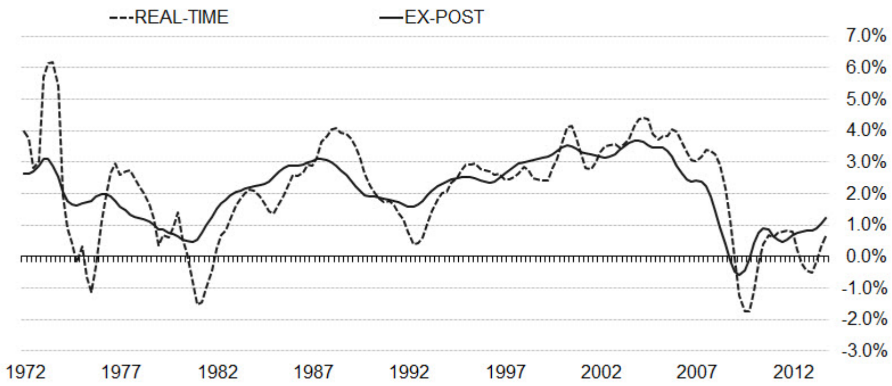


Figure 3.37: Figure shows real-time and ex-post estimates of potential growth from a production function

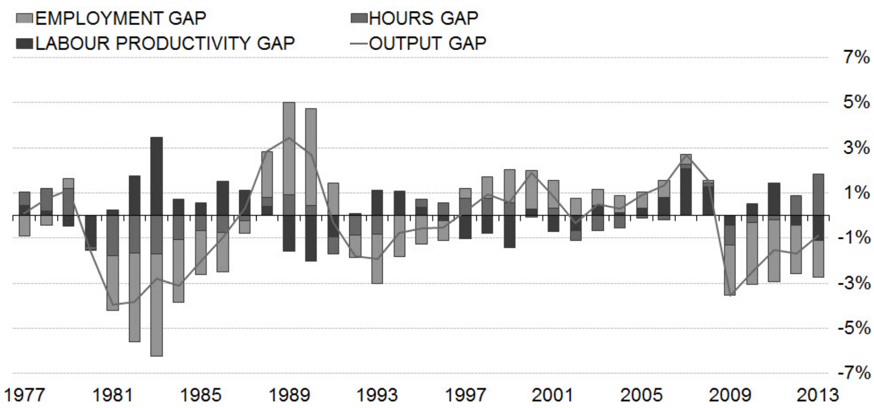


Figure 3.38: Figure shows the output gap decomposed into contributions from average hours, employment and labour productivity in percentage points

ing its long-run downward trend. More generally, this approach is consistent with the employment gap driving much of the cyclical variation of output with smaller contributions from other sources.

Another, more detailed decomposition gives us an indication of the TFP and capital intensity contributions to the output gap. This method shows that much of the current labour productivity gap is accounted for by TFP falling short of its potential level. Capital per worker (K/L) appears now to be making a neutral contribution to the output gap, having contributed positively since the recession started.

The production function can also be used to infer what other methods of output gap estimation say about the level of potential TFP. In what follows the contributions of potential labour and capital input used to construct the production function are subtracted

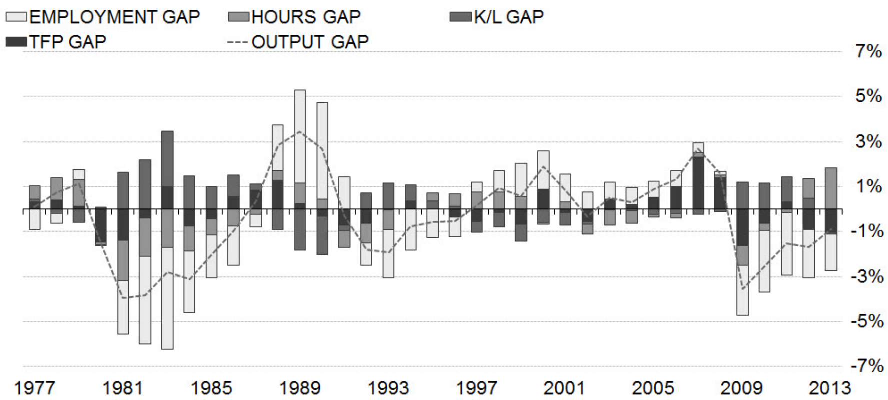


Figure 3.39: Figure shows the output gap decomposed into contributions from average hours, employment and total factor productivity and the capital to labour ratio in percentage points

from the level of potential output associated with each method, leaving estimates of potential TFP.[1]

Figure 3.40 shows that each of the HP, MV, PCA, PC and production function (PF) methods is consistent with a significant and permanent fall in the level of potential TFP over the crisis period (although the HP filter dates the start of the slowdown as pre-crisis). Interesting too, is that most of the methods are unable to explain the weakness of productivity over 2012, instead, attributing it to weaker potential TFP growth. While breaking production function estimates of the output gap into factor contributions, and considering the TFP implications of other estimation methods, may give some insight into how the output gap might have evolved, the results should be considered in the context of the wider uncertainties surrounding output gap estimation.

3.3.10 Summary Statistics for the Estimation Methods

The table below describes a few key statistics, where available, for the methods presented above, as well as some metrics for a production function estimated using market sector data (described in the annex). Some methods are excluded from the summary statistics, including the AC model, the BN decomposition, the CF filter, the H filter the market sector production function model and the PCA model.

[1] To facilitate this comparison, the production function and labour-market augmented filters use the same estimate of the NAWRU.

Table 3.1: Table showing statistical properies of a range of output gap estimation methods

Estimation Method	Average Absolute Output Gap Revision	Average Absolute Potential Growth Rate Revision	Standard Deviation of Output Gap	Standard Deviation of Potential Growth Rate
Capacity Utilisation Model	0.6	0.4	1.6	1.5
Credit-Augmented Filter	1.2	0.7	1.9	1.7
HP Filter	1.4	1.0	1.5	1.3
Linear Trend Model	3.8	0.3	5.1	0.0
Multivariate Filter Model	1.3	0.6	2.0	1.6
Okun's Law Model	2.0	0.9	2.6	1.3
PC Filter	1.3	0.7	1.3	1.3
Phillips Curve Model	2.6	1.0	1.2	1.9
Production Function Model	1.5	0.8	2.1	1.0
AC Model	NA	NA	1.3	2.0
BN Decomposition	NA	NA	0.5	2.7
CF Filter	NA	NA	1.2	1.7
Market Sector Production Function Model	NA	NA	2.4	1.2
PCA Model	NA	NA	1.3	2.0
Hamilton Filter	1.6	2.3	3.8	2.5
Average	1.7	0.7	2.2	1.3
Average Magnitude of Output Gap		1.0		
End Point Uncertainty (Average Revision Due to Benefit of Hindsight)		1.7		
Model Uncertainty (Average Range of Estimates)		3.6		
Data Uncertainty (Arising from Revised Statistics)		0.7		

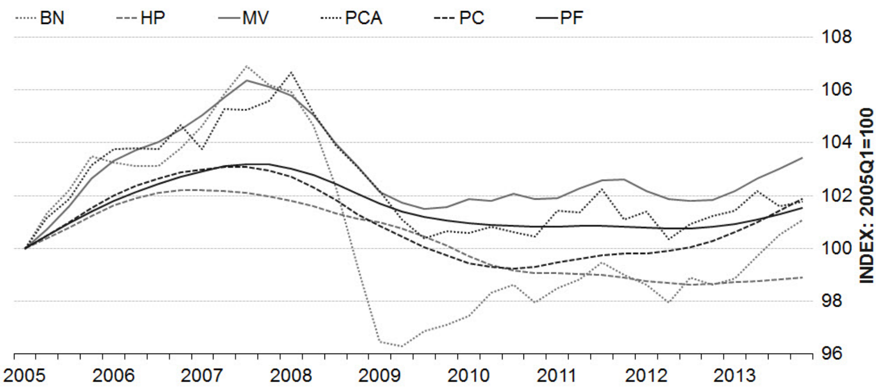


Figure 3.40: Figure shows level of potential TFP implied by the output gap estimates from a range of different methods

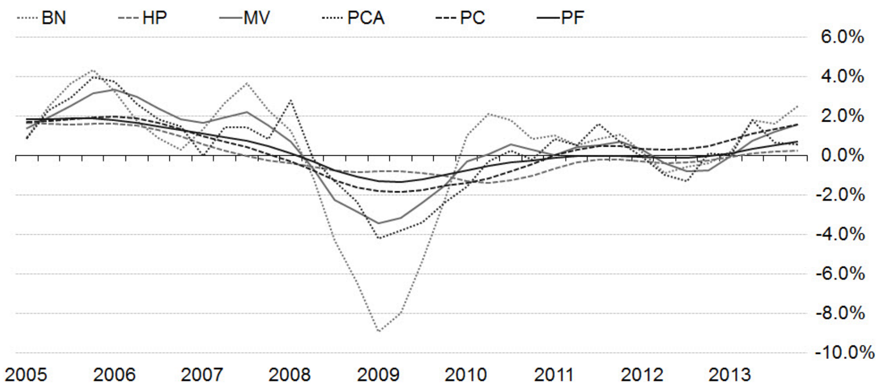


Figure 3.41: Figure shows growth of potential TFP implied by the output gap estimates from a range of different methods

3.4 Quantifying Uncertainty

3.4.1 Model Uncertainty

This section summarises the estimates presented in the preceding sections. The many model assumptions and possible judgements required to estimate the output gap are borne out in the wide range of estimates produced — illustrated in Figure 3.42.[1] And, of course, this range could expand with the inclusion of other methods.

The range of output gap estimates was around 4.7 percentage points over 2013 as a whole, which compared with a range of estimates made by external analysts at the end

[1]For the purposes of constructing the figure, estimates associated with the linear trend, Hamilton and credit gap methodologies are excluded.

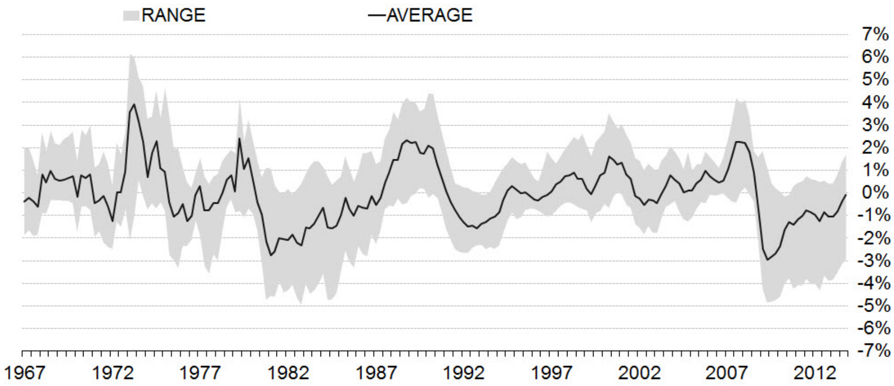


Figure 3.42: Figure shows the range of output gap estimates from a variety of selected estimation methods. The black line is the mean estimate and all are expressed as a percent of GDP

of 2013 of 5.7 percentage points.[1] There must, therefore, be a more diverse population of definitions, methods and judgements applied by other analysts than is presented here, indicating that it is worth considering a broader range of evidence when reaching a judgement on spare capacity.

It has been suggested that aggregating estimates from a range of models may reduce sensitivity of the measures to model-specific bias - Armstrong (2001). The mean of model estimates is presented in Figure 3.42, but it is worth considering that this, in some sense, is arbitrary because it places an equal weight on each measure and there is no reason to suppose each is equally plausible.

Nonetheless, the figure shows that the range of uncertainty varies substantially over the sample period and that the average of estimates rarely lies in the centre of that distribution. Overall, the range of estimates provided by these methods alone is about 3.6 percentage points on average between 1972 and 2013 and that's with the benefit of hindsight — they use the full run of data available now.

3.4.2 End-point Uncertainty

Figure 3.43 illustrates the tendency of each measure to be revised, though this should not be the only criteria one considers when assessing the performance of output gap measures. The average (excluding the linear methodology) is some 1.7 percentage points. While the overall magnitude of revisions to output gap estimates is very large, it is also variable. The figure shows that there are substantial differences in the revisions properties of the models presented, with estimates generated by linear de-trending much more likely

[1] See HM-Treasury (2013).

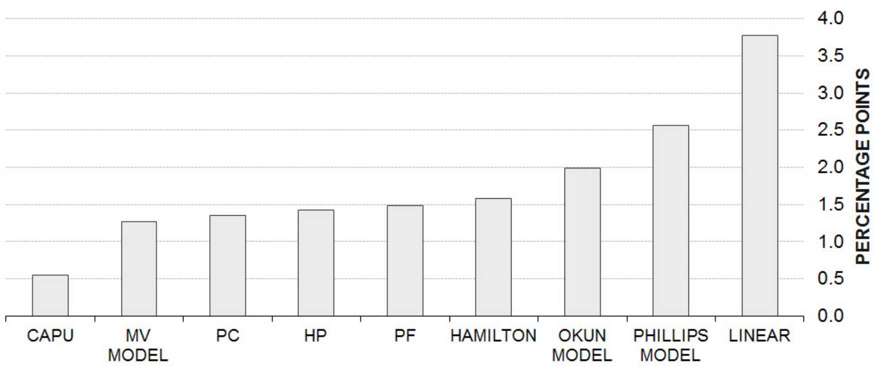


Figure 3.43: Figure shows the mean revision between real-time and ex-post output gap estimates for a range of methods

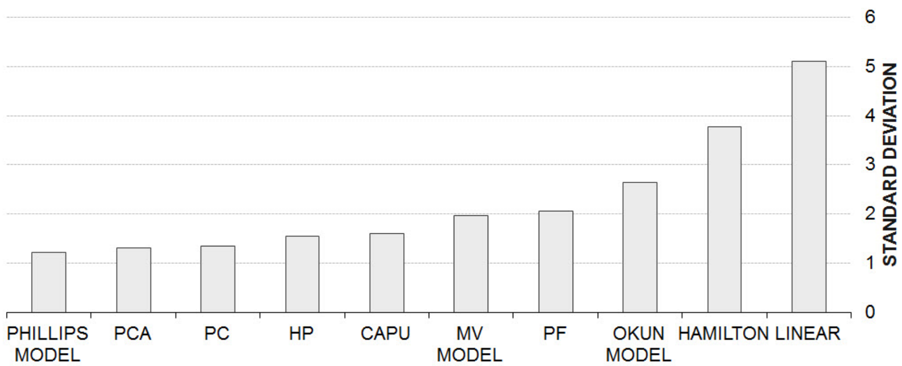


Figure 3.44: Figure shows the standard deviation of output gap estimates for a range of methods

to be revised than the multivariate filter, for example. Likewise, the volatility of the output gap and potential growth vary significantly between methods.

Figure 3.46 shows that, generally speaking, the lower the standard deviation of potential output growth, the more prone is an output gap measure to revision. Linear detrending, for example, lies in the south-east corner - it assumes a constant potential output growth rate, but its output gap estimates are revised the most. Another way of thinking about it is that a linear growth model places the most weight on information from the future (the whole sample determines the slope of the line) in determining its estimates of the output gap. So, naturally, as more data become available it makes big revisions to its earlier estimates. Other methods place less weight on future information and, therefore,

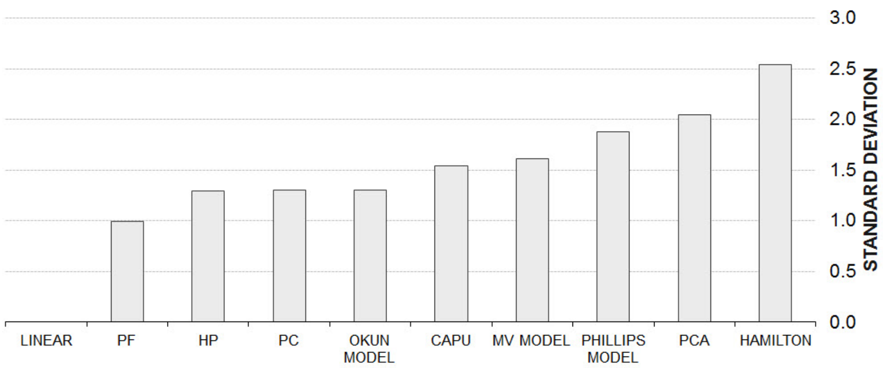


Figure 3.45: Figure shows the standard deviation of potential growth estimates for a range of methods

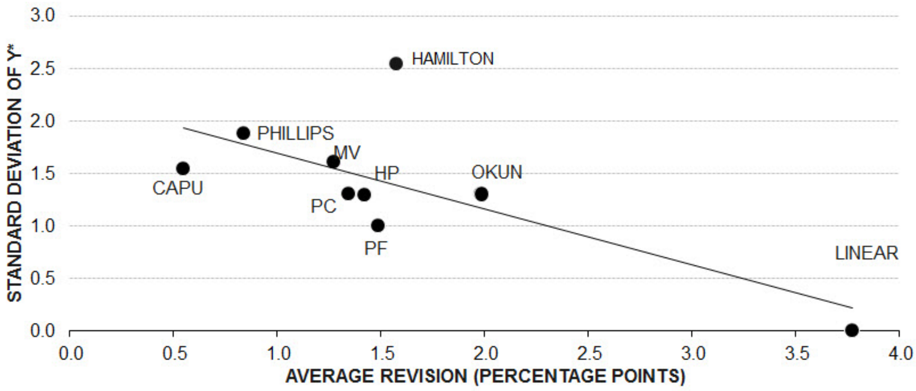


Figure 3.46: Figure plots the mean revision against the standard deviation of potential growth for a range of output gap estimation methods

they tend to be revised less[1].

The methods presented above are by no means intended to capture all the methods of output gap estimation available to forecasters, but it does capture a wide range, which is reflected in the variability of estimates over time. There are other methods not considered here, such as the multivariate Kalman filter, as applied by Konuki (2008). In practice, this is very similar to the MV (PC) filter, presented above. The parameters of a multivariate Kalman model must be calibrated, often under the assumption that they are deep parameters - i.e. stable and based on microeconomic evidence. The PC filter also depends upon calibration and some of the parameters could be considered reduced-form

[1]The statistics for the Phillips Curve method are over a sample from 1995-2013, as uncertainty over the steady state of inflation in the earlier years of the sample causes significant revisions that are not considered relevant to this comparison.

representations of structural relationships.

The MV (PC) filter requires explicit judgements over smoothing parameters and the weights placed on explaining the path of potential output, while a Kalman model requires explicit priors over the size of shocks to its measurement equations. Likewise, estimates produced using models that utilise Bayesian methods — such as Beneš *et al.* (2010) — require priors over the size of model parameters and shocks as well as their underlying distributions. Fundamentally, since potential output is a notion constructed by economists, all methods used to estimate it need to be told how volatile it is - there is no escaping this key judgement.

The key point to take away from this analysis is that the choice of model really matters to the estimates made. The reason they differ so much is because they embody different assumptions about the process that drives trend output and the how the indicators used reflect deviations of the economy from it. Some models do not fit with commonly held views about these processes – the HP filter is one. Yet the Hamilton filter also embodies assumption which may not match the process driving GDP either.

There is one question all researchers should ask before applying any method to separate trend from cycle. Are the assumptions embodied in the chosen method appropriate for the process being modelled?

3.4.3 Data Uncertainty

It is a well-documented finding that optimal ex-post monetary policy prescriptions appear different from those made in real time and that this mostly reflects the unreliability of output gap estimates in real time — see Orphanides & Norden (2002). In these papers, the authors assess the revisions properties of a range of output gap estimates for the United States, including linear de-trending, the HP filter, an Okun relation and an unobserved components model augmented to include the Phillips curve. Across all measures, the role of data revisions is found to be relatively small in explaining policy errors, when compared with those arising from the end-point problem.

Marcellino & Musso (2011) find that output gap revisions for euro area economies are substantial and of equal magnitude to the output gap estimates themselves - they find that data revisions account only for a small part of this. While Cayen & Van Norden (2005) find that revisions to output gap estimates for Canada are significant but can mainly be attributed to data revisions.

To see which of these findings hold for the UK, revisions to HP filter estimates of the output gap from two sources are assessed — from new data and revised data. The sensitivity of output gap estimates to data revisions is assessed first. To this end, an HP filter is applied to both the real-time GDP data and the current vintage and the results

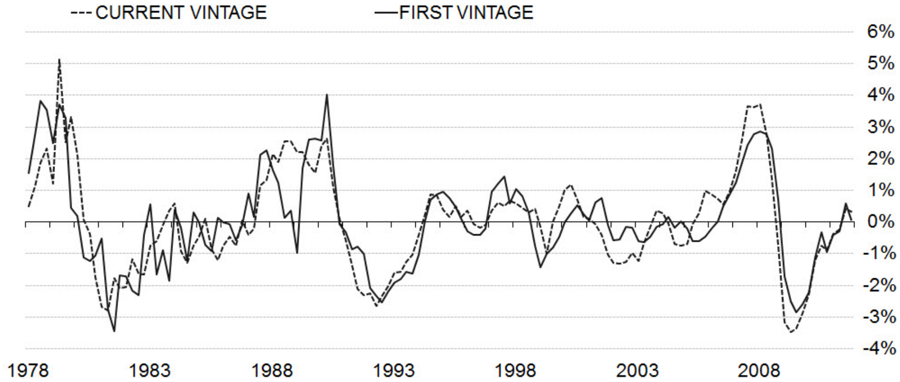


Figure 3.47: Figure shows HP-filtered estimates of the output gap using the first available and current vintage of GDP data as a percent of GDP

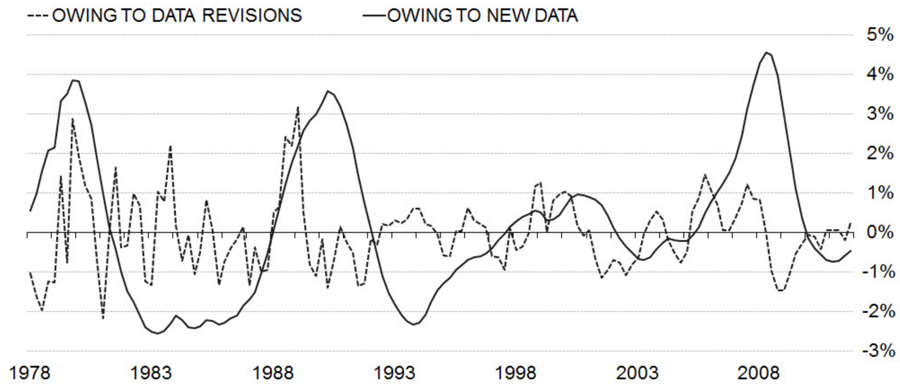


Figure 3.48: Dashed line shows output gap revision owing to revisions to the first estimate of GDP, while solid line shows the revision to the output gap reflecting the arrival of new data points in the series.

are compared. The average absolute output gap revision arising from GDP data revisions from the first quarter of 1978 to the final quarter of 2011 is around 0.7 percentage points.

I then compare the one- and two-sided HP filter estimates applied to the current vintage of GDP data — the average revision owing to the arrival of new data (which is called the benefit of hindsight here, although that would also include updates to model parameters, for example) is found to be around 1.4 percentage points. The influence of data revisions on output gap estimates is presented in Figure 3.46, while the sources of output gap revision are illustrated in Figure 3.47.

The results suggest that, on average, the benefit of hindsight is about twice as useful as better data when it comes to improving HP-filtered estimates of the output gap — although this conclusion would probably be affected by the choice of estimation method.

The results obtained here lie between those for Canada, where output gap revisions are mostly attributable to data revisions, and the US and euro area, where data revisions play a much smaller role. It is unclear why data revisions should play a different role in UK output gap revisions than in other countries. UK GDP prior to the recent financial crisis now look a bit stronger than early vintages implied, and the recession a bit deeper. To the extent that revisions like this are not features of the data in other countries, the proportion of output gap revisions owing to data revisions will differ from the UK. Of course, history will continue to be rewritten as methodological improvements come to change our understanding of the past.

3.4.4 Summarising Uncertainty

Overall, the results show just how slippery a concept the output gap is. Uncertainty comes from three sources:

- a failure to pin down the process that underlies potential growth — Overall, different judgements over how the economy functions drives a gap between selected measures of the output gap that has averaged 3.6 percentage points since 1970;
- the benefit of hindsight — knowing how the economy evolved after a shock, for example, tells something about the sustainability of the economy before it. On average, revisions from this source were about 1.7 percentage points since 1970; and
- revisions to past data are a smaller source of uncertainty than the benefit of hindsight or model uncertainty, associated with revisions to output gap estimates of about 0.7 percentage points on average since 1978.

3.5 How Smooth is Potential Output?

A common feature of all the output gap estimation methods summarised above is that they are underpinned by assumptions, explicitly or implicitly, about the smoothness of potential output. The Hodrick-Prescott (HP) and prior-constrained (PC) filters, for example, require explicit judgements over the choice of smoothing parameter.

The amplitude of the cycle obtained using principal components depends on the choice of scaling parameter, which determines the medium-term variation in potential output. Directly estimating the output gap with principal components also allows very high frequency movements in the survey data and measurement error to feed through to the level of potential.

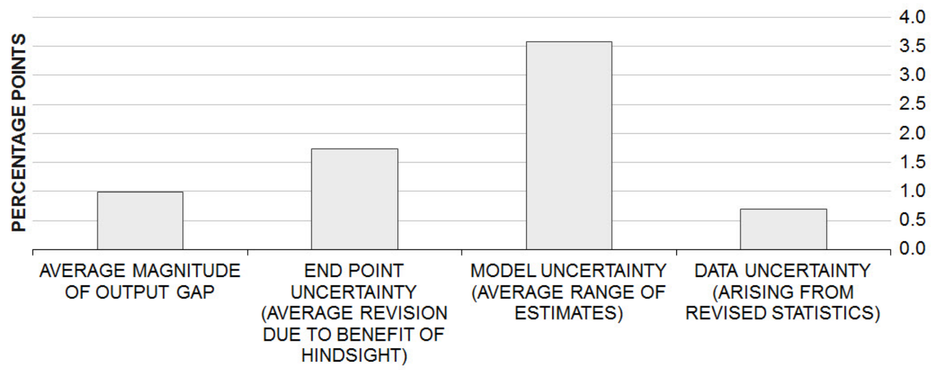


Figure 3.49: Figure shows different sources of output gap uncertainty. Bars represent mean revision owing to that source.

The production function approach allows the user to apply more disaggregated judgements to arrive at an overall smoothness. For example, the structural rate of unemployment, which depends on things like the degree of wage indexation to inflation and union intensity, might be less variable than potential total factor productivity (TFP) over time.

Yet, once decisions have been made about the smoothness of the other components of potential output, one is left with the crucial judgement over the smoothness of potential TFP. Unfortunately, the economics profession has little to say about the processes that determine it, which makes judgements about its potential level, volatility and likely growth rate dependent on historical averages — the key judgement being how much weight to place on the recent past relative to longer time periods.

For example, most economic commentators viewed the sustained period of strong productivity growth in the UK over the 2000s as being structural. And, with no theory to suggest otherwise, it was hard to make the opposite argument. During the 2008 financial crisis, TFP fell sharply, and has barely grown in the subsequent six years. This performance has prompted some to revisit their assumptions about the sustainability of growth in the preceding years. This is illustrated in Figure 3.50, which shows the output gap path the OECD published before the crisis, in 2007, and the one it published in 2014. And to a lesser extent in Figure 3.51, which shows the Treasury and, latterly, OBR estimates.

The OECD, which uses a production function method, estimated with a two-sided filter, now judges that there was a significant positive output gap in 2007 of around 5 per cent. This interpretation is consistent with potential TFP growth having slowed substantially in the years before the crisis started. Partly, this reflects the properties of the HP filter, which penalises sharp movements in potential output. An alternative interpretation, from the multivariate (MV) filter or principal components analysis (PCA), for example, is that potential TFP continued to grow at roughly its usual rate up until the crisis began,

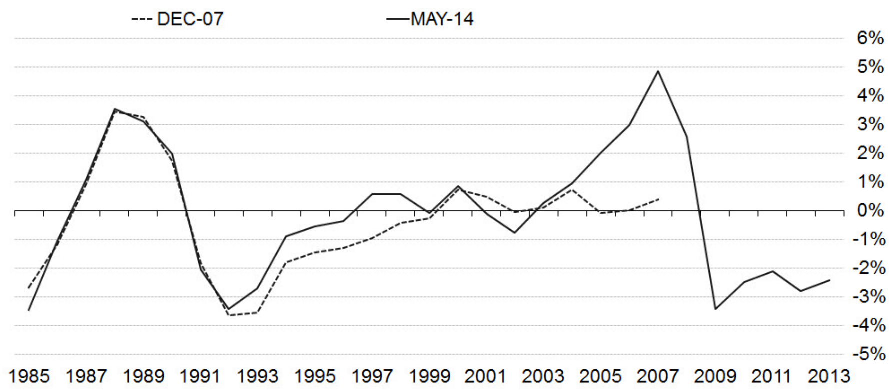


Figure 3.50: Figure shows two different vintages of the OECD's output gap estimate for the UK.

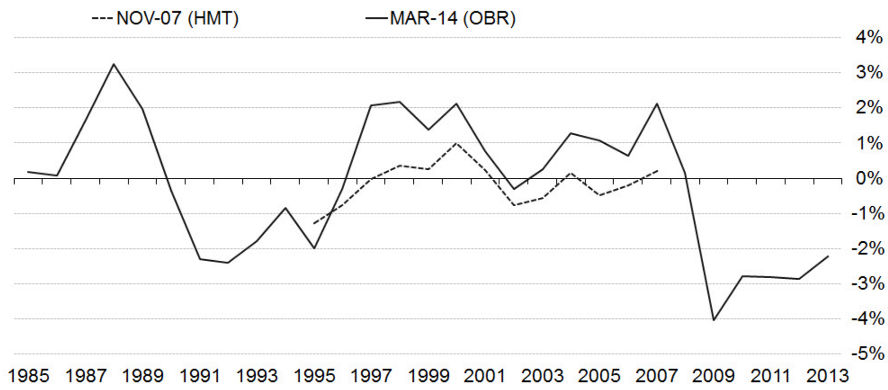


Figure 3.51: Figure shows two different vintages of the UK government's output gap estimates

but fell sharply during it.

So which interpretation seems more plausible? Potential TFP growth is considered to reflect advances in technology and improvements in efficiency (such as better process management, for example). It is difficult to understand how a substantial TFP gap could open up, since it implies excessive utilisation of technology available at the time — working the capital stock at unsustainable rates, for example. Yet there was no evidence of this in the period before the crisis.

But it is also hard to see how the state of technology and efficiency can suddenly regress so far — have we forgotten how to do things we could do previously? Or is it that some of the things we were doing, which seemed to make things more efficient at the time, turned out to be less useful than we thought (some activities of the shadow banking sector, for example)? All we really have to go in is that actual TFP fell and has barely grown for six years and either the gap or level hypothesis could be correct.

So, it is important to recognise that different methods for estimating the output gap are consistent with different interpretations of the process driving TFP. The implication is that output gaps should not just be ranked on their tendency for revision: a method that assumes potential TFP can fall sharply will tend to be revised by less than a method that assumes it cannot. Given the scant theory on the subject, there is no strong case for favouring one view of the world over the other, so both should be considered when making a judgement and neither discarded on the basis of tendency for revision.

3.6 Mapping to the Public Finances

The UK Government sets fiscal policy such that a measure of balance is restored to the public finances over a rolling five-year horizon. In doing so it takes into account the effect of the cycle on revenues and spending, which requires an estimate of the output gap. In practice, every forecast of GDP is underpinned by an assessment (implicit or explicit) of spare capacity and, therefore, how much scope there is for above-trend growth as it is taken up. This section considers the sensitivity of cyclically-adjusted measures of the fiscal aggregates to output gap mismeasurement.

Recent years have seen the emergence of new literature concerned with the reliability of cyclically-adjusted measures of fiscal aggregates. As reported in Tereanu *et al.* (2014) a number of papers have assessed output gap uncertainty and some explore the implications for measures of the public finances:

- Koske & Pain (2008) find that revised output gap estimates account for cyclically-adjusted public borrowing revisions of around 0.4 percentage points, on average, across a range of OECD countries;
- Bouis *et al.* (2013) find that output gap revisions average 1 to 1.5 percentage points in OECD countries, but that underlying fiscal balances are not very sensitive to this;
- Hughes Hallett *et al.* (2012) find that revisions to cyclically adjusted public borrowing estimates owing to output gap revisions average around 1 percentage point in most euro area economies; and
- Ley & Misch (2013) find that revisions have substantial effects on measures of the structural balance.

In what follows, the measurement of cyclically adjusted public sector net borrowing (CAPSNB) is considered. It is first perhaps useful to set out the wide band of uncer-

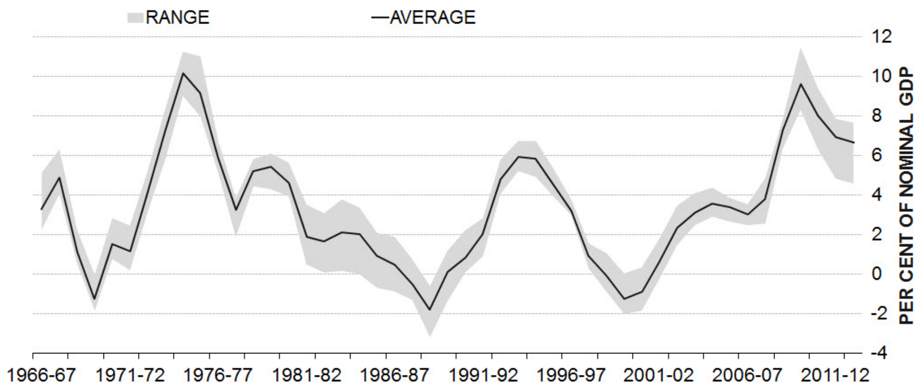


Figure 3.52: Solid line is the cyclically-adjusted public sector net borrowing balance as a share of GDP. The shaded area represents the range of estimates using different methods of output gap estimation. Note:Data are cyclically-adjusted PSNB from 1994-95, and the cyclically adjusted public sector borrowing requirement before that.

tainty surrounding CAPSNB — this is illustrated in Figure 3.52, which presents estimates based on the suite of output gap measures presented in this paper and ex-post data. Uncertainty over point estimates of the output gap carries over to CAPSNB via the cyclical-adjustment coefficients, roughly 0.7 for 1.[1]

Uncertainty over point estimates of the CAPSNB is compounded by uncertainty over how estimates might change in future. CAPSNB can be revised for three reasons, of which this section considers the first two:

- the public sector finances and output data can be revised;
- estimates of the output gap can change; and
- estimates of the sensitivity of the public finances to the cycle can be updated.

So what might be the greater source of uncertainty? To assess this, the CAPSNB is first estimated using initial estimates of PSNB as a share of GDP and a real-time measure of the output gap — in this case given by the one-sided PC filter. This serves as a baseline case.

To assess sensitivity to output gap revisions, it is shown how the estimates change when the CAPSNB is calculated using an ex-post output gap (a two-sided PC filter). Finally, the CAPSNB is estimated using both the latest fiscal data and an ex-post output gap, to capture the effect of data revisions. The results are presented in Figure 3.53.

The dark bars in Figure 3.53 can be interpreted as total revision owing to revised fiscal data (and in-year forecast errors), while the lighter bars are revisions because the

[1] See Helgadottir *et al.* (2012)

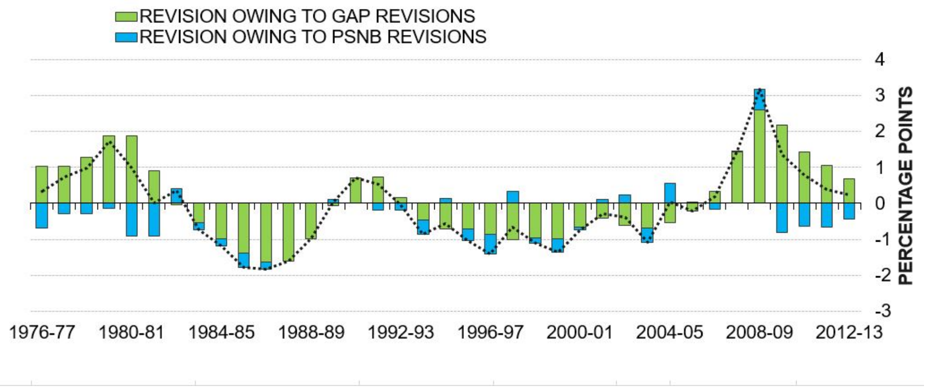


Figure 3.53: Figure decomposes the overall revision to estimates of structural borrowing into that caused by output gap revisions and that reflecting actual borrowing data revisions.

output gap path is now thought to be different — the purple and blue bars sum to the total CAPSNB revision. The figure shows that the estimates of structural borrowing through the financial crisis presented here are revised much higher once data revisions and revised output gap paths are taken into account — with the latter playing the dominant role. This is consistent with the findings of Tereanu et al, who state that ”during the crisis years, the estimates of CAPB were considerably worse for most countries after they were re-estimated following budget execution”.

On average, the magnitude of revisions to CAPSNB owing to output gap revisions is around 1 percentage point, far larger than the average revision arising from revised data of 0.3 percentage points.[1] So output gap uncertainty is likely to be the larger source of uncertainty over the structural fiscal position in real time, although selecting other measures of the output gap with different revisions properties could affect this conclusion. This result is consistent with the findings of Hallett et al (2009), who find that CAPB revisions owing to output gap revisions are around 1 percentage point on average.

The average output gap revision across the range of methods applied to UK data earlier is around 1.3 percentage points. This is roughly consistent with the findings of Tereanu et al, who find that the average revision for EU countries is around 1.5 percentage points. Using cyclical adjustment coefficients summing to 0.7, it would be reasonable to assume that revisions to CAPSNB would be close to 1 percentage point using that range of methods.

It is worth noting that the analysis above suggests that revisions to the CACB tend to be largest around turning points, also consistent with the conclusion reached by Tereanu et al. This is problematic, since these are precisely the moments when policymakers most

[1]For simplicity, the revision to the output gap does not take into account revisions to the path of GDP so output gap uncertainty may be an even larger source of uncertainty.

need them to be reliable.

This section has shown that output gap mismeasurement is a significant source of uncertainty over the real-time fiscal position and probably a larger one than revisions to the public finances data. But it is worth remembering that, when fiscal policy objectives are set in the future, the output gap is only one part of what is needed to forecast the cyclically-adjusted fiscal position. One also needs to make a judgement about how fast the economy might be able to grow on a sustainable basis. This paper is not concerned with that issue, although it is likely to be as important.

3.7 Conclusion

This paper began by emphasising that the policy horizon should be taken into account when formulating the definition of the output gap. This reflects the possibility that fiscal authorities and central banks may have different perspectives, particularly when output gaps are large. Output gap estimates are prone to revision, not least because the output gap is a concept invented by economists that cannot be observed, only estimated. Three sources of output gap revision were then identified — those arising from the arrival of new data, revisions to past data and changes to model specification.

It has been shown that revisions owing to the arrival of new data are large and, on average, tend to be of the same magnitude as the output gap estimates themselves. Output gap revisions owing to revised data are also significant, while the range of output gap estimates produced by the handful of methods presented here is substantial. Overall, the level of uncertainty about the size of the output gap is high and it is shown that this carries over to estimates of the cyclically adjusted fiscal position, while revisions to public sector finances data also contribute to a smaller degree.

Along the way, it has been illustrated that the assumptions underpinning various methods for estimating the output gap are consistent with different views of how the economy functions — in particular, about the time series properties of potential productivity, about which little is known. Whether the views underpinning a methodology are explicit or implicit, they represent the application of judgement to the estimation of spare capacity. No methodology can be made totally free from judgement.

The important conclusion from this analysis is that no statistical filter should be applied blindly. It is all too easy to apply an HP filter or a form of the Kalman filter to data without asking whether the underlying process is consistent with the, often very restrictive, assumptions embodied in the filter. The HP filter may have a use but, as Hamilton suggests, detrending GDP is probably not one of them. Indeed, detrending GDP may not even be the best application of the Hamilton filter.

All medium-term forecasts of the public finances embody an assumption, implicit or explicit, about spare capacity. Yet, given the substantial degree of output gap uncertainty, there may good reason to tread cautiously when setting fiscal policy. Such an approach would be consistent with the Brainard Principle of policy conservatism. Further work in this area might assess how the optimality of fiscal rules is affected by output gap uncertainty.

3.8 Annex

3.8.1 Estimating the structural rate of unemployment

Some of the methods presented in this paper make use of the structural rate of unemployment to estimate the output gap. To ensure consistency between those methods, the structural rate of unemployment is estimated in advance of output gap estimation. To do this, it is assumed that the structural rate is a random walk without drift and the PC filter is applied to it. This filter is augmented to include information from real wages, adjusted for underlying movements in productivity growth — the estimated structural rate can, therefore, be thought of as a long-term non-accelerating wages rate of unemployment (NAWRU).

The long-run NAWRU is estimated in state space using the Kalman filter and is given by Equations 32 to 38:

$$\text{Signal}:u_t = u_t^* + uc_t \tag{3.33}$$

$$\text{State}:u_t^* = u_{t-1}^* + (\frac{\epsilon_{1,t}}{k^{1/2}}) \tag{3.34}$$

$$\text{State}:uc_t = \epsilon 1, t \tag{3.35}$$

$$\text{State}:cw_t = \text{prod}_t^* + \beta_5 uc_t + \beta_6 uc_{t-1} + \epsilon 1, t \tag{3.36}$$

$$\text{State}:\text{prod}_t = \text{prod}_t^* + \text{prod}_t^* \tag{3.37}$$

$$\text{State}:\text{prod}_t^* = \text{prod}_{t-1}^* + (\frac{\epsilon_{1,t}}{\psi_{\text{prod}}^{1/2}}) \tag{3.38}$$

$$\text{State}:\text{prod}c_t = \epsilon 1, t \tag{3.39}$$

I posit a simple relationship between the growth of real product wages, as measured by average earnings deflated by the gross value added deflator at factor cost, the unemployment rate gap and a measure of trend productivity growth – to capture underlying movements in labour productivity (unrelated to the cycle) that could influence wages.

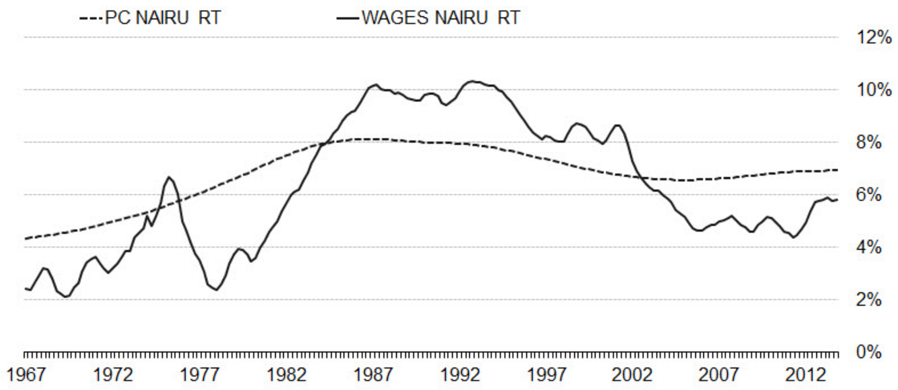


Figure 3.54: Figure shows a simple PC filter estimate of the NAWRU and another in which that filter is augmented with wages data – both are real-time estimates.

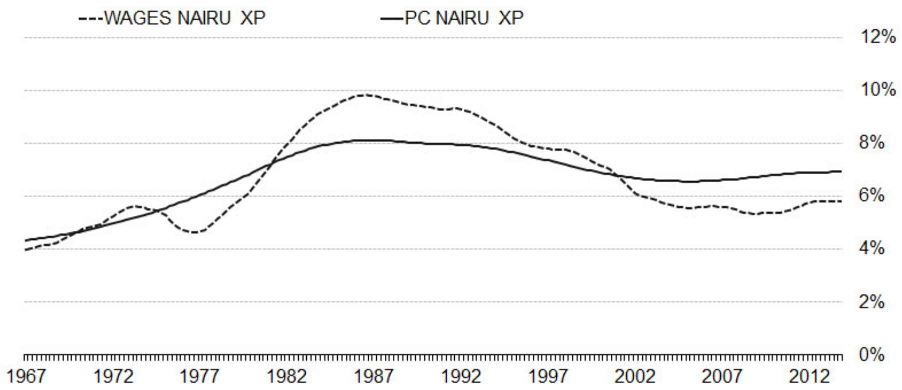


Figure 3.55: Figure shows a simple PC filter estimate of the NAWRU and another in which that filter is augmented with wages data – both are ex-post estimates.

The labour productivity trend is jointly estimated with the NAWRU. The coefficients are calibrated drawing from evidence, including Greenslade *et al.* (2003).

Figure 3.54 and Figure 3.55 show the effect of including wages on the NAWRU estimates, relative to using the naeve univariate PC filter. Because real wages growth was low and stable in the 2000s, the PC filter is able to identify that the NAWRU must have fallen in the preceding years. Likewise, the weakness of real product wages growth during and following the recession helps the filter to identify more cyclical weakness in unemployment, preventing a significant increase in the long-run NAWRU estimate.

This, combined with the stiffness of the filter, is consistent with the view that little has changed in the structure of the labour market over that period, and that the long-term unemployed will eventually find their way back into employment, for example.

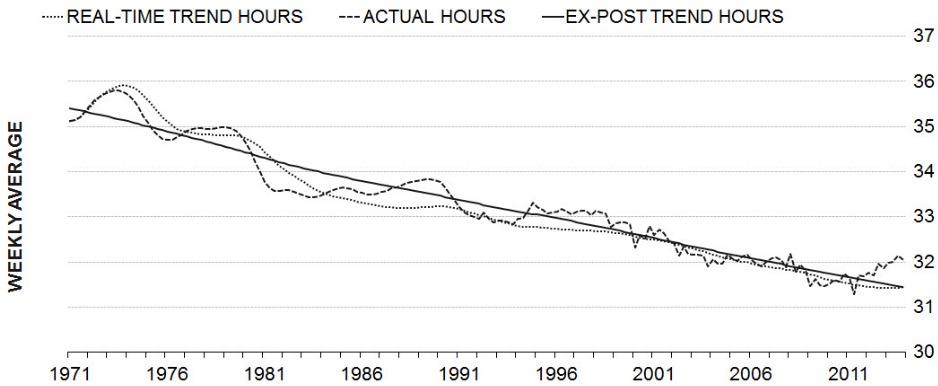


Figure 3.56: Figure shows the filtered real-time and ex-post estimates of trend average hours, together with the actual data.

3.8.2 Estimating potential average hours

Average hours have drifted downwards for centuries, as productivity has risen and additional income has been substituted for more leisure time. To model trend average hours, the PC filter is applied with drift to the average hours series using a stiff smoothing parameter. The real-time and ex-post results are illustrated in Figure 3.56. A key uncertainty is the extent to which the latest increase in average hours reflects a permanent response to a loss of permanent income relative to prior expectations. The estimate presented here is consistent with average hours eventually falling back to the long-run trend.

3.8.3 Estimating potential activity

Potential activity has also been estimated using the PC filter but, unlike average hours, without a drift term — the estimates are presented in Figure 55. Further information relating to demographic influences on participation rates could be introduced via a cohort model, for example, but that is not explored here.

3.8.4 Taking account of market sector information

Many of the traditional methods used to estimate the output gap are based on inferring supply from the balance of demand. But demand is really a market sector concept, so it would make sense to estimate the output gap on this basis. In what follows, it is shown that variations in the output of the non-marketed sector of the economy contribute to output gap volatility when measured on a whole-economy basis. This is best investigated using the production function approach.

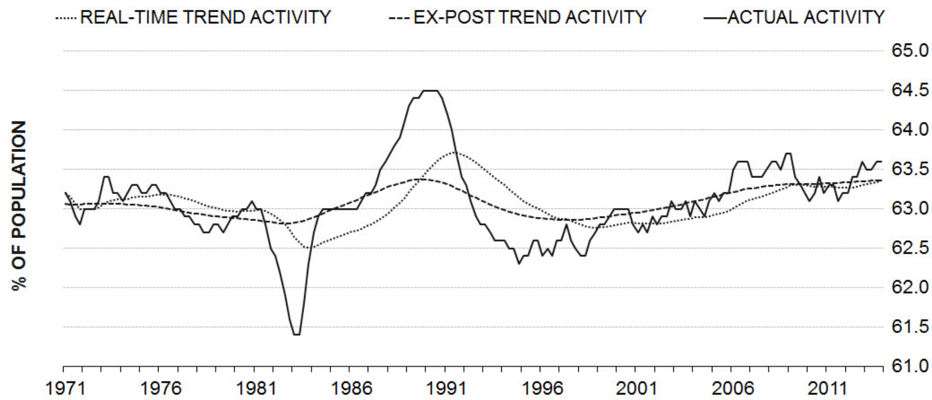


Figure 3.57: Figure shows the filtered real-time and ex-post estimates of trend activity, together with the actual data.

To arrive at a whole-economy equivalent measure for the output gap, non-marketed TFP, labour input and capital are assumed to be at their trend values at all times. Actual values for market sector unemployment, hours and activity (labour input) are combined using the production function methodology with estimates of the capital stock and market-sector, non-oil GVA to arrive at series for market sector TFP. These series are then de-trended to arrive at an estimate of the market sector output gap.

Figure 3.58 and Figure 3.59 show that using market sector data does affect estimates of the whole-economy output gap. This is largely because it excludes any information from the government sector, which, theoretically, should not be cyclical, and avoids the issues associated with the measurement of its output.[1] The differences are, though, typically small, particularly over the past 25 years.

Figure 3.60 shows that the level of actual TFP in the market sector fell by substantially more in the most recent recession than in the economy as a whole, which is to be expected because of the way in which public sector productivity is measured - market sector TFP fell nearly $8\frac{1}{2}$ per cent from peak to trough, compared with around 6 per cent in the economy as a whole. It also fell more sharply in 2012 than did whole economy productivity. For this reason, trend (filtered) estimates of market sector TFP appear to continue falling, long after trend whole economy TFP stabilises.

The conclusion to draw from this is that the way in which public sector productivity is measured could muddy the picture of the process we really want to examine — it is the growth of market sector TFP which matters most for projections of receipts, for example. But the price paid for this more disaggregated information is greater complexity and

[1] See Atkinson (2006) for a full discussion of the challenges associated with measuring government output in the UK.

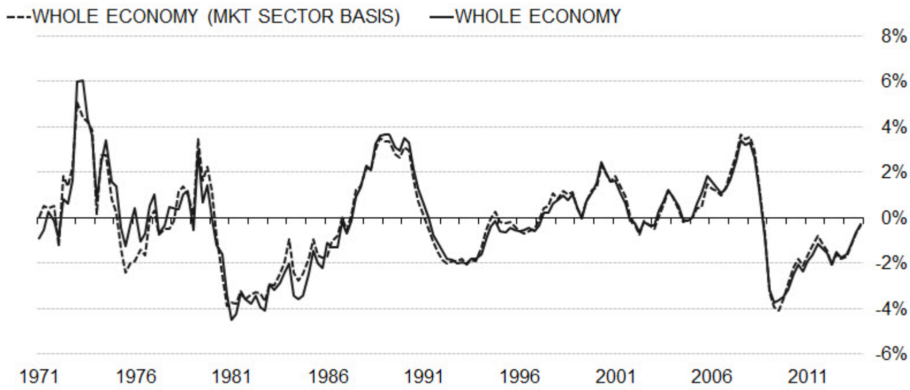


Figure 3.58: Figure shows two estimates of the output gap using different measures of economic activity.

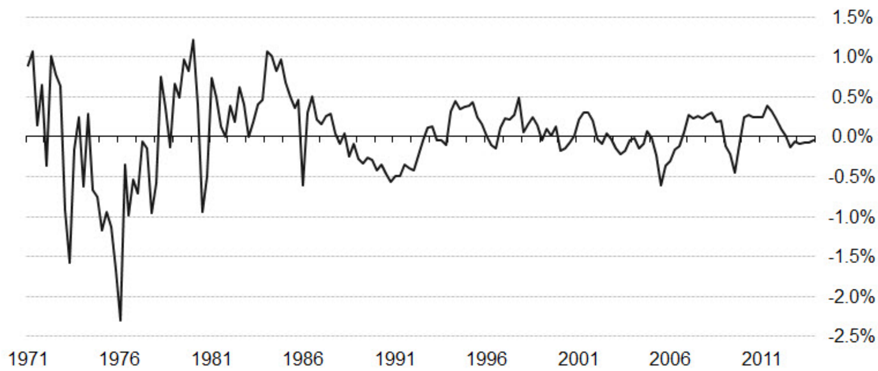


Figure 3.59: Figure shows the differences in output gap estimates between whole economy and market sector only data.

dependency on the assumptions required to construct a market sector data set.

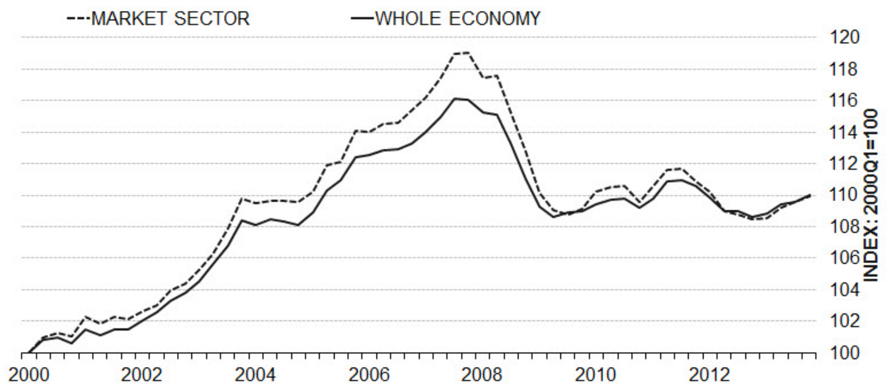


Figure 3.60: Figure shows two estimates of total factor productivity using different measures of economic activity.

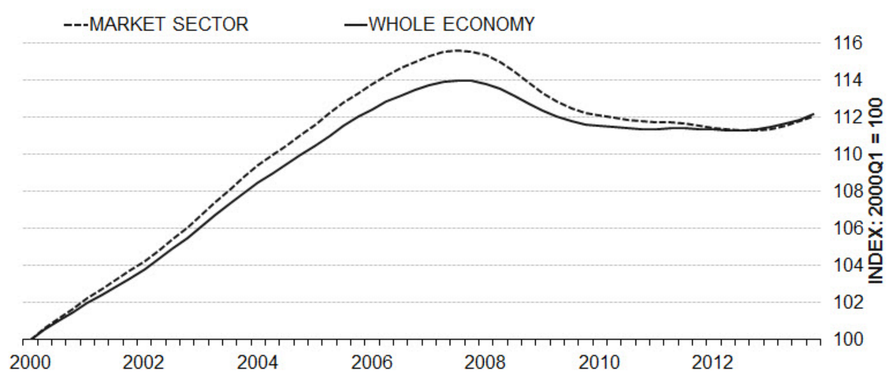


Figure 3.61: Figure shows two estimates of potential total factor productivity using different measures of economic activity.

4

Modelling State-Contingent Fiscal Rules in the United Kingdom

4.1 Introduction

Fiscal rules come and go. That's particularly true in the United Kingdom where many incarnations have made their way through parliament in recent years. Finding one that lasts would be a leap forward, especially since there is only very weak global evidence that they are binding - Heinemann *et al.* (2018).

In what follows, I focus on a rule that was extremely short-lived - it was proposed by the government in 2015, HM-Treasury (2015) and abandoned a year later, HM-Treasury (2016). It had two interesting features which are worthy of consideration, even now the rule has been discarded:

- it explicitly gave the government flexibility to delay fiscal consolidation when the economy is hit by a significant negative shock; and
- it defined a significant negative shock as being when GDP growth slows or is forecast to slow to 1% on a rolling 4 quarter-on-4 quarter basis.

I explore the merits of these features in turn. First, by asking whether this type of flexibility helps dampen the impact of shocks — intuition suggests that it should. Second, by assessing how using a GDP growth threshold might differ from using estimates of the output gap for the identification of significant shocks.

4.1.1 The Golden Rule — 1998 to 2008

Before 1998 Britain was a lawless place, fiscally speaking. It wasn't until Gordon Brown arrived at the Treasury that the government bound itself by rules of its own choosing. The experience ever since then has been that when the circumstances change, the goalposts move.

The original rules were twofold.

- The golden rule: over the economic cycle, the government will borrow only to invest and not to fund current spending. In other words it will run a surplus on the current budget, on average.

- The sustainable investment rule: over the economic cycle, the ratio of net public sector debt to GDP will be set at a 'stable and prudent' level, defined by the Chancellor as no more than 40% of GDP.

The logic of the rules is reasonable. Both are intended to ensure that public spending is spread fairly across generations - none should have to deal with the consequences of another's profligacy. That borrowing only had to be balanced over the cycle meant there was also scope to support the economy at times of weakness. The rules also permit spending on investment, which tends to be squeezed out in favour of current spending if it is not protected.

A couple of issues, noted at the time Emmerson *et al.* (2001) are that the debt level chosen is arbitrary — there's no particular reason to favour 40% over other figures. The definition of investment is also narrow — why should it be limited to investment in capital when spending on human capital such as health and education also offer returns on investment?

Another is that the backward-looking nature of the rule meant that if surpluses were run in early years of an economic cycle, deficits could be run in later years. That may be sustainable over the economic cycle but it means the public finances end the cycle in a poor position — as was the case on the eve of the financial crisis.

So how did the rules fair? Not so well.

The loophole was the economic cycle, which is a subjective concept open to exploitation. The forecasts for revenue growth also proved too optimistic, creating deficit bias. Both were easy to manipulate because the prisoners were also the guards – the Treasury held all the keys.

Budget surpluses run in the Labour government's first term meant a deficit could be run in later years and the overall current budget would average zero. Even then, in 2005, the deficit was big enough to put that in jeopardy. The solution was to shift the timing of the cycle earlier. By saying the economic cycle began in 1997 rather than 1999, more years of surplus were captured within it, making the target easier to hit.

With regard to the next cycle, the Treasury was serially optimistic on the scope for a lift to revenue as spare capacity was taken up. Again, part of the trouble was that the economy was judged to be operating below trend when, even at the time, most economists thought it was running at or close to full capacity.

The result was that when the global financial crisis struck, Britain was running a substantial structural budget deficit that was larger than its peers. The general conclusion, Riley & Chote (2014) is that fiscal policy ahead of the global financial crisis was not the reason the public finances deteriorated when it struck. But weaknesses in fiscal management left the government with much less room to manoeuvre than it perhaps should have had.

4.1.2 Temporary Operating Rule — 2008 to 2010

As the U.K. economy plunged into crisis, an operating rule was put in place. The November 2008 objective was to "set policies to improve the cyclically-adjusted current budget each year, once the economy emerges from the downturn, so it reaches balance and debt is falling as a proportion of GDP once the global shocks have worked their way through the economy in full" HM-Treasury (2008).

The report contained a number of stimulus measures to support the economy during the downturn and it was imagined that the current budget would not be balanced until the 2015/16 financial year. In other words, pragmatism saw to it that the right fiscal policy was pursued during the crisis, but that came at the expense of the fiscal rules that went before it.

A natural question to ask is whether there is a specification that could explicitly permit fiscal flexibility during a downturn, such that the rules do not have to be junked? If binding at other times, such a rule would be more credible. The 2015 rule comes close, but there were plenty of others in between.

4.1.3 Fiscal Responsibility Act — Feb 2010 to June 2010

This one didn't last long. The Labour government's last Budget, HM-Treasury (2010), refined plans for the path of public spending and taxation in subsequent years and coded them into law. The fiscal objective was for borrowing to fall as a share of GDP in each of the subsequent years — but even on those plans the government would still be running a substantial deficit even five years hence.

4.1.4 Fiscal Mandate and Supplementary Target - 2010 to 2015

The arrival of a new Coalition government marked a major change to fiscal management. The new fiscal rule was to achieve cyclically-adjusted current balance, by the end of the rolling, five-year forecast period, and for public sector net debt to be falling at a fixed date of 2015-16.

In principle, the rule was similar to the golden rule — it offered limited flexibility during fiscal shocks and protected investment spending. But the main difference was that it was forward rather than backward looking. To police these targets an independent fiscal institution was also created, the Office for Budget Responsibility. The fudging of the economic cycle and revenue forecasts would be a thing of the past.

So how did this rule fair? Again, not so well.

The trouble was that deficit reduction proved far harder than anticipated with very weak potential growth following the financial crisis creating a substantial fiscal drag. Because the deficit target was always five-years ahead, the point at which structural balance was achieved kept slipping backward, and the debt target was missed.

The other problem was that big consolidation can always be pencilled in for a couple of years in the future. The target looks like it will be met, but only if implausible paths for public spending were realised.

Finally, the government was unable to fudge the forecast for revenue or the economy, but it was able to make big spending cuts look like a normal baseline - this greatly reduced the transparency of policymaking. As the OBR noted, the baseline assumption for spending in the medium term — years in which there were no detailed fiscal plans — changed often and became much more complicated.

4.1.5 A Conditional Fiscal Rule - 2015 to 2016

The first Budget of the new Conservative government saw a the introduction of a conditional fiscal rule. It was completely unique, compared with what went before it. For the first time the rules specified not just the fiscal objective - which was for outright surpluses - but when it would be allowed to deviate from it. As a recap the rules specified:

- a target for a surplus on public sector net borrowing in 2019-20, and a supplementary target for public sector net debt to fall as a share of GDP in each year from 2015-16 to 2019-20

- a target, once a surplus is achieved in 2019-20, to run a surplus each subsequent year as long as the economy remains in normal times To gauge when the economy is outside of normal times, a significant negative shock was defined as reducing real GDP growth to less than 1% (on a rolling 4 quarter-on-4 quarter basis).

It's worth remembering that this rule was announced during a period of significant consolidation and intense discussion on the appropriate stance of fiscal policy. A big contribution to this debate is Portes & Wren-Lewis (2015), who draw various conclusions from the literature on fiscal policy to devise specific policy recommendations.

They advocate for fiscal rules as a solution to deficit bias and suggest that deficit targeting is superior to debt targeting – that, the reasoning goes, is because deficit targeting should allow fiscal adjustments to be less sharp and debt can be brought under control in the medium-term by altering the deficit targets. The authors also conclude that the rule should give fiscal policy space to support the economy when negative shocks occur — particularly when interest rates are as low as they can go.

The key features of the government's fiscal rule, implemented shortly after the paper was published, look broadly aligned with the recommendations made.

The Institute for Fiscal Studies, Emmerson *et al.* (2016) made an assessment of the fiscal rule in the year after it was announced. The analysis used forecast and outturn data to identify periods of significant shock, finding that fiscal flexibility would have been granted during each of the past four decades. This assessment said when deficit reduction could be abandoned but made no formal attempt to measure the benefit to demand stabilisation of doing so.

Overall, then, no detailed quantitative assessment has been made of the government's fiscal rule. But related research suggests it would be a good rule to investigate and that methods have been developed making it possible to test whether it might dampen demand shocks.

A second question worth asking is whether GDP growth is the best indicator to trigger fiscal flexibility. The output gap is already used to adjust the public finances for the effects of the economic cycle — helping to determine how much fiscal effort is needed to reach a particular target for the headline deficit in years to come.

The output gap is also used in the formulation of monetary policy and the literature records a debate between whether the output gap or growth might be the more appropriate indicator for setting interest rates Walsh (2003). The drawback with the output gap is that real-time estimates can be extremely misleading — that's been shown empirically by Orphanides & Norden (2002) for the U.S., Rünstler (2002) for the euro area and for the U.K. in Murray (2014).

As the IFS notes, the output gap is also a difficult concept to explain to the public and it is also a challenge to estimate objectively. So why should it be considered as a trigger variable: because growth indicators have drawbacks too.

First, the cause of the growth shock could matter. If it is because aggregate demand in the economy has weakened, this calls for fiscal flexibility, to get output back to potential. If instead, growth has stuttered because the economy's supply potential has temporarily weakened, the growth rule would allow a necessary fiscal response to be delayed for no obvious reason.

Second, should the economy's potential growth rate permanently slow to below 1%, the fiscal rule would permit the government to abandon deficit reduction indefinitely. That's particularly important now that trend productivity growth appears to have slowed.

Third, it's possible that growth could fall below trend for cyclical reasons, prompting a protracted spell of higher unemployment, but not trigger the growth knockout. In that case fiscal support might be warranted but not granted. The IFS observes that actual data suggest this isn't a particularly common occurrence.

4.1.6 A Return to Deficit Targets - 2016 to...

The U.K.'s decision to leave the EU and a new Chancellor at the Treasury prompted yet another set of fiscal rules. The overarching objective of which is to "return the public finances to balance at the earliest possible date in the next Parliament". HM-Treasury (2016). The formal definition of a significant shock went, the target slipped backward and the cyclically adjusted measure of borrowing returned.

Overall, the new rule has delivered substantially more flexibility but is considerably less onerous than those that went before it. It is also less interesting from a research standpoint. In what follows I therefore focus on the conditional fiscal rule introduced in 2015.

The questions I attempt to answer are:

- How useful is fiscal flexibility during a downturn?
- Which is a better trigger for fiscal flexibility, growth or slack?

4.2 Methodology

The simulations are conducted in three stages. The first stage involves estimating demand and supply shocks to the U.K. economy using a Kalman filter, Kalman (1960). The second step is to use a model to compare the performance of two fiscal rules when faced with these shocks — one that provides the government with fiscal flexibility when growth slows, and another which does not. The third stage is to assess the trigger for fiscal flexibility. I compare the performance of one rule which gives flexibility when growth slows and another which provides flexibility when the output gap turns negative.

4.2.1 Part 1: Modelling Approach

Demand and supply shocks to the U.K. economy

Demand and supply shocks to the economy affect the public finances in different ways. A negative demand shock prompts a cyclical weakness that might be expected to reverse as tax revenues recover and spending returns to a more normal share of GDP. A negative supply shock leaves the economy permanently smaller than it otherwise would have been, creating a structural deterioration in the public finances that won't improve without remedial fiscal effort.

So to understand how different fiscal rules might have performed in the past, shocks to the economy must, as a minimum, be decomposed into either demand or supply shocks. There are a number of possible paths forward at this juncture. A straightforward approach would be to use the HP Filter, Hodrick & Prescott (1997), specify a law of motion for aggregate supply and demand and recover the disturbances from those equations.

Instead, I have used a multivariate Kalman filter. Any filter can be expressed in state space, including the HP filter, and solved using the Kalman algorithm. It is therefore not the Kalman algorithm that is desirable, in and of itself. Rather it is the flexibility of state space which is useful. It allows for much more sophisticated filters to be created and solved easily. In this case I have sought to include information beyond the path of actual GDP that is relevant to slack and should have some bearing on its estimation. Specifically, I have included indicators such as unemployment, inflation and survey measures of capacity utilisation.

This filter strikes a balance between drawing from a wider range of relevant data and using so much that availability of that data become a problem. The specification I have used means shocks can be recovered all the way back to 1965.

This filter is also one of the key methods used by the British government to inform the official estimates of the output gap that are used to produce the public finances forecast and estimates of structural borrowing. So, compared with the HP filter, the method I choose is a step closer to the real world.

A further consideration is that the model selected has different properties regarding the end-point problem. It is well known, Orphanides (2002), that estimates from the HP filter tend to be revised - and substantially - when new data arrive. This is true of all filters, but the HP filter is more likely to allow the level of trend to move to meet the level of actual GDP at the end of the sample than is the filter specified here. At least that's what Murray (2014) shows.

The reason is that there is no prior assumption made by the HP filter that the level of trend GDP tends to rise over time. The only assumptions made for the purpose of estimation by the HP filter is that cycles aren't too big and that trend growth isn't too volatile. Introduce the assumption that trend output is likely to keep rising over time and the revisions properties change substantially - open output gaps are more likely at the end of the sample period.

Filters featuring these sorts of assumption were used by Beneš *et al.* (2010) and Laxton & Tetlow (1992), for example. A multivariate version has since been used by Blagrave *et al.* (2015) for a range of countries and Alichi *et al.* (2017) for the euro area and U.S. In both cases, the multivariate filter was found to be more reliable in real-time than univariate statistical methods such as the HP filter.

The model used is illustrated in the following state-space representation.

$$\text{Signal}: gdp_t = gdp_t^* + y_t \tag{4.1}$$

$$\text{State}: gdp_{t+1}^* = gdp_t^* + \text{drift}_t + \left(\frac{\epsilon_{1,t}}{k^{1/2}}\right) \tag{4.2}$$

$$\text{State}: \text{drift}_t = \text{drift}_{t-1} \tag{4.3}$$

$$\text{Signal}: y_t = \epsilon_{1,t} \tag{4.4}$$

Actual output is constrained to equal the sum of potential output, gdp^*, and the output gap Kappa is the ratio of the variance of shocks to demand and supply and therefore determines how smooth potential GDP is. This specification of the Kalman filter is known as the prior-constrained filter as it contains a prior belief about the process driving potential GDP: it's assumed to be a random walk with drift. Without the inclusion of this assumption, this specification of the Kalman filter is virtually identical to the HP filter - and most differences appear towards the end of the estimation period in any case.

From here, additional signal equations are added to the Kalman filter to incorporate more relevant information about the economic cycle. The equations chosen are a Phillips relation, an Okun relation and an equation linking capacity utilisation to the output gap.

The filter is estimated from 1965 up to the third quarter of 2017, with the first and last couple of years of data discarded. Further detail on the multivariate Kalman filter used to estimate demand shocks to the economy is presented in Murray (2014).

The Kalman filter estimates of the output gap can be used to decompose the actual GDP growth rate of the economy into cyclical and structural components - those are illustrated in Figure 1.

Figure 2 goes further, plotting the shocks to supply and demand recovered from the Kalman filter estimates. As is plain to see, the 1960s and 1970s were a more volatile period in the U.K. than during recent times - with the notable exception of the global financial crisis in 2009.

Since we are trying to understand how different fiscal rules might have performed faced by the same sequence of shocks as occurred in the past, it's important to account

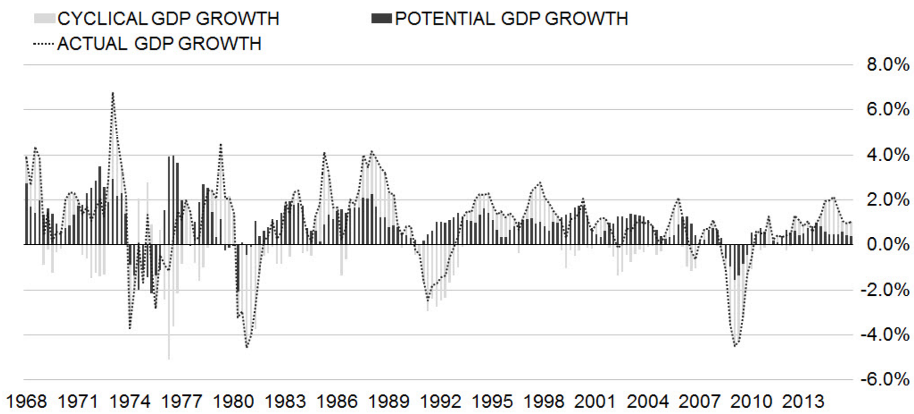


Figure 4.1: Figure shows the contributions to actual GDP growth from growth in demand, depicted by dark bars, and growth of supply, light bars. The estimates are derived from a Kalman filter model.

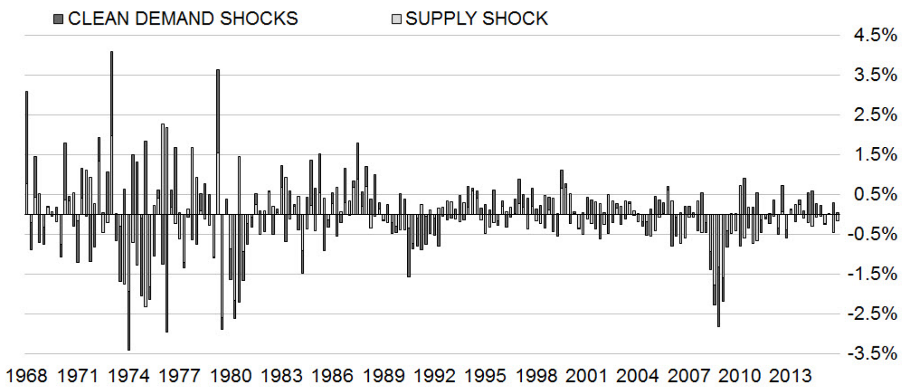


Figure 4.2: The figure plots non-fiscal demand and supply shocks derived from a Kalman filter.

for the influence of fiscal shocks in the past. In other words, it's important to clean the data set to isolate just the shocks that fiscal authorities might have to deal with again.

To remove the influence of fiscal policy, the change in the structural deficit is regressed on the demand shocks recovered from the Kalman filter. What's left, the residuals from that equation, are assumed to be clean. The key assumption here is that changes to the structural deficit are a suitable proxy for discretionary fiscal policy.

4.2.2 Solution Method

To assess how the fiscal rule might have performed in the past, an economic model is needed — and its solution method must satisfy two requirements.

- the shocks to which it is subjected must be unanticipated, which calls for a stochastic rather than deterministic model. The government did not see the global financial crisis coming, so it mustn't be able to in the model.

- the model must permit asymmetric responses. Under the fiscal rule, the government responds differently to positive shocks from negative ones.

It doesn't matter at this stage quite which model is used, but the state-contingency of the policy rule creates a challenge. The fiscal rule is asymmetric - deficit reduction happens when growth is strong, but not when growth is slow. Any single model with state-contingent policy rules that cannot be differentiated has no log-linear approximation, meaning standard solution methods cannot be used.

To address this I specify two separate models and switch between them. In one, the fiscal policy rule is for deficit reduction, in the other no effort is made to shrink the deficit. Both policy rules are differentiable and so simulations can proceed, one quarter at a time and switching between models as and when growth drops below the knockout threshold.

To run simulations I have used tools more usually applied in the literature on monetary policy - in particular in the analysis of the zero-lower bound of interest rates. Out of necessity, new methods have been developed that allow for the simple solution of models with occasionally-binding constraints - Guerrieri & Iacoviello (2015) have been instrumental in driving this field forward and opening it up to a broader set of researchers.

The authors show that a model with occasionally-binding constraints can be solved using a piecewise linear perturbation method. The difference, when compared with the naive solution, is that agents are able to form expectations over when the regime will switch back. This is much more realistic than an assumption that fiscal laxity will continue indefinitely.

The Guerrieri and Iacoviello algorithm depends on two assumptions. That no shocks are expected in future - a standard assumption - and that if a jump is made from the reference to alternative regime, that a jump back will occur at some point in the future.

The algorithm computes estimates of how long it will take to return to the reference regime using a guess and verify approach depending on the starting conditions in any given quarter and the shock experienced in that quarter. When applied, this solution method is shown to match more analytically sophisticated but computationally demanding methods such as dynamic programming.

4.2.3 Other Applications of the Solution Method

The OccBin toolkit has been used to answer a range of policy questions. The creators of the toolkit themselves have published research applying this solution method, including

Guerrieri & Iacoviello (2017). In that analysis it is shown how the impact of loosening of collateral constraints had a smaller effect on household consumption in the run up to the global financial crisis, than the sudden binding of those constraints had during the downturn. That asymmetry helps explain the dynamics of the financial crisis. Perhaps the most obvious application is to the zero-lower bound of monetary policy. In one regime monetary policy is effective, but when rates approach zero, the central bank is no longer able to provide additional support to aggregate demand.

An early research effort on this topic is Fernández-Villaverde *et al.* (2015), who use a simple model together with the OccBin toolkit to understand how large shocks to the U.S. economy need to be to arrive at the zero-lower bound and how long the constraint might persist. They also uncover early results showing that the fiscal impact multiplier is around 1.7 when rates are at the ZLB, compared with 0.5 in normal times.

Other investigations into fiscal policy at the lower bound using the OccBin toolkit have been published, too. Flotho (2018) find that fiscal policy plays an important role in stabilising the economy when monetary policy is constrained and that a properly designed policy can shorten the spell spent subject to the constraint. Bhattarai & Egorov (2016) explore the implications of trade elasticities for optimal fiscal policy when rates are at the ZLB.

Finally, an OECD analysis, Claveres & Stráskỳ (2018), explores the benefits to demand stabilisation of a pooled unemployment insurance mechanism for the euro area when monetary policy is constrained by the lower bound of interest rates. Again, the OccBin toolkit is used to solve this non-linear model. It is shown that such a scheme would materially reduce output losses, especially in the periphery, when the economy experiences a negative shock.

In short, there is plenty of research using the OccBin toolkit to assess constraints on the operation of monetary policy and the role of fiscal policy when rates are at the ZLB. That's not the case for analysis of constraints on fiscal policy, such as the state-contingent rules explored here.

4.3 The Model

For both regimes I use a simple reduced-form model of the U.K. economy. The ideal model of the economy would accurately capture the key mechanisms that determine the impact of fiscal policy on the economy. For a start, that means it must have a representation of the fiscal impact multiplier - the force with which changes to tax and spending depress or boost output.

It must then capture the dynamics of how the economy adjusts to those shocks. Monetary policy is one of the key adjustment mechanisms for the economy - the ideal model would replicate the Bank of England's broad understanding of how interest rate changes affect growth and inflation. Monetary policy also affects the economy via shifts in the exchange rate, so it is important to capture that channel of transmission also.

I have chosen a small, tractable model that broadly matches the simulation properties of monetary policy shocks in COMPASS. In other words, the model captures the cushioning impact of monetary policy on an economy suffering fiscal shocks well. But it does not attempt to model all shocks that might affect the U.K. economy. The general approach to parameterisation is to draw reasonable estimates from the literature for the U.K. economy, where possible, or elsewhere if estimates are unavailable.

Aggregate demand in the economy is given by an IS equation, relating the output gap, y, to lags and expectations of itself, real interest rates, r, changes in the exchange rate, q, discretionary fiscal policy measures, f, and a stochastic shock, ϵ^y. The parameters α_y, α_r, α_q, and α_{NR} determine the degree of forward-looking behaviour, the interest rate elasticity of demand, the responsiveness of output to the exchange rate and the fiscal impact multiplier respectively.

$$y_t = \alpha_y y_{t-1} + (1 - \alpha_y) y_{t+1_{|t}} - \alpha_r r_t - \alpha_q \Delta q_{t-1} + alpha_{NR} f_t + \epsilon_t^y \tag{4.5}$$

The degree of forward-looking behaviour plays an important role in determining the persistence of shocks to demand and the adjustment of the economy to them. Empirical evidence shows that adjustment occurs more slowly than purely forward-looking models would suggest. For example, the habit formation model of household consumption, which implies slow adjustment to income shocks, is a better fit with the data than a forward-looking model, Fuhrer (2000).

The chosen parameter for the degree to which agents in the economy are forward looking is 0.85. That, in combination with other parameter choices drawn from the literature is consistent with the general response of the economy to interest rate shocks contained in COMPASS. It is also close to the parameter chosen by Batini & Haldane (1999), which they consider to be empirically plausible.

The COMPASS model contains a habit formation parameter, which is 0.7. But, in that modeI, additional persistence appears to be achieved by using forcing process equations in the specification of shocks. This means overall persistence is likely to be a little higher and that is the reason for the bigger parameter selected. In other words, persistence is a feature of both models, but it is introduced in different places in COMPASS compared to the model used here.

A crucial assumption that governs the response of the economy to fiscal policy is the interest rate elasticity of demand. That is a function of structural parameters in COMPASS, but in this reduced form model it is captured by a single parameter.

The evidence on the effectiveness of monetary policy in the U.K. is a little thin, but a recent contribution to the literature, Cesa-Bianchi *et al.* (2016) suggests the impact of monetary policy shocks on demand and inflation is substantial. It also contains a summary of related research for the U.K. and U.S. — the estimates vary widely.

The parameter related to the interest rate elasticity of demand in this model is set at 0.1. This is consistent with Cesa-Bianchi *et al.* (2016), who find that a one-percentage point jump in the cost of borrowing, maintained for a year, depresses output by 0.1% in the following quarter, reaching a peak impact of 0.4% on the level of output. Again, together with other parameters of the IS-relation, the impact of a monetary policy shock on the economy is qualitatively similar to that produced by the Bank of England's forecasting model, COMPASS - see Burgess *et al.* (2013).

How currency movements affect the economy is captured in the IS relation, but the right coefficient is harder to judge. Britain's decision to withdraw from the EU is a useful illustrative example of this problem. The pound depreciated by between 10% and 20%, depending on the starting point, in the run up to and following the Brexit vote. But this has not translated into a surge in exports and net trade.

Ordinarily, a depreciation can lift the prices of exported goods and services in local currency terms — maintaining prices in foreign currency terms — and this thickens profit margins. The bigger profits make it desirable to lift production, in order to make more. Yet, businesses do not yet seem to have invested in additional capacity.

The reason may be that the depreciation may reflect a long-term hit to the economy - trade will be harder after Britain leaves the EU and so the currency has adjusted to reflect that. In years to come, exporters will find it harder to sell and so this limits the incentive to raise production.

Meanwhile, the weaker currency has raised import costs, dragging on household consumption growth. In this case it is perfectly possible that the depreciation has depressed output rather than lifting it. The source of the shock, may therefore matter, meaning the choice of parameter ought to reflect the shocks assessed using the model.

In the simulations, I'm interested in how monetary policy and the exchange rate cushion the impact of fiscal policy, not how risk premium shocks can affect demand. So this argues for a negative coefficient - a lower exchange rate should lift demand. The magnitude of the coefficient is broadly consistent with Murray (2012), which reviews evidence from single equation estimates and a structural VAR.

The sections above have concerned how demand might adjust to a general shock. But what about the immediate impact of a discretionary fiscal tightening or loosening? That is captured by the fiscal impact multiplier. If a government cuts spending, the mechanical impact on GDP may not be one-for-one because fewer imports will be consumed, for example. But the withdrawal of demand could also prompt belt tightening by other agents in the economy, lifting the impact.

The choice of parameter is an important one and it depends on a wide number of influences. For example, the type of spending measure is likely to be signifiocant — OBR (2010) put the impact of investment spending at unity, with welfare and departmental spending at 0.6; but closer to 0.3 for income tax changes. More recent estimates OBR (2015) put the impact of departmental spending cuts a little lower at 0.45, reflecting the impact of prices adjustment.

The timing may also matter. Auerbach & Gorodnichenko (2011) find that the fiscal multiplier can be small during expansions but significantly bigger than unity during recessions. Blanchard & Leigh (2013) also find that the multiplier can be large. And Murray (2012) shows that the distribution of estimates of the impact multiplier is skewed toward being larger rather than smaller in New Zealand, a small open economy.

Another way to think about the fiscal impact multiplier is in terms of Ricardian equivalence. If every agent in the economy is rational and optimising, the impact of a fiscal shock should be small. After all, a tax cut now means a tax hike later on — households should save the windfall in readiness for the burden that will be borne later on. But if some proportion of households does not optimise, as seems to be the case, fiscal policy does have an impact.

The evidence shows that Ricardian Equivalence does not hold. Haug (2016) tests equivalence using a measure of tax shocks developed by Romer & Romer (2010). The results show that Ricardian Equivalence is rejected and that the associated fiscal multiplier varies over time. Hayo & Neumeier (2017) also reject full Ricardian Equivalence on the basis that many people surveyed in Germany did not adjust their behaviour to account for the run up in debt during the global financial crisis.

Fiscal policy is subject to strict assumptions in COMPASS. Full Ricardian Equivalence is assumed, meaning the impact of discretionary fiscal policy is not captured directly. Instead, the Bank of England depends on a suite of other models to inform its judgement on the impact of spending and tax shocks. This is one other reason why COMPASS is not suitable for the simulation exercise undertaken here.

Since much of the recent and substantial fiscal consolidation has been focused on changes to departmental spending, which has a larger multiplier than tax measures, and it seems likely that the impact of fiscal policy may have been bigger than the OBR expects it to be in the future, at certain times in the past, I have chosen an impact multiplier of 0.6. A spending cut of 1% of GDP therefore has an immediate impact of 0.6% on the level of GDP. In time, that impact shrinks as monetary policy reacts to the shock and the economy adjusts.

Inflation is divided into two types — domestically-generated and that originating from beyond the country's borders. A Phillips curve relates domestically-generated inflation, π^D, to a lag of itself, the output gap and a stochastic shock, $\epsilon^{\pi D}$. The parameters β_π and β_y determine the degree of forward-looking behaviour in price setting and the sensitivity of inflation to slack respectively. The equation is presented as deviations from the inflation target.

While the Phillips curve contains a substantial forward looking element, it is assumed that inflation expectations are anchored on the target. Since this is assumed to be a constant, it means expectations drop out when the equation is log-linearised around its steady state. In other words, the monetary policy framework in the U.K. is assumed to be credible and stable in the model.

$$\pi_t^D = \beta_\pi \pi_{t-1}^D + \beta_y y_{t-2} + \epsilon_t^{\pi D} \tag{4.6}$$

The persistence of inflation — seasonally-adjusted and on a quarterly basis — is lower than that of the output gap and the parameter chosen for lagged inflation, 0.6, reflects this. The persistence of headline inflation is lower still, thanks to the temporary influence of shocks from import prices and energy costs on quarterly price changes.

The sensitivity of domestically generated inflation to the output gap is the parameter determining the slope of the Phillips curve. The parameter chosen is close to that used by the OBR in its small model of the economy, Murray (2012).

Headline inflation, π_t, includes external influences on prices owing to movements in the exchange rate over the past year and additional shocks, denoted by ϵ^π. The sensitivity of inflation to the exchange rate is determined by , which partly reflects the import intensity of consumer goods and services. Other external shocks, such as those to the price of energy are captured by the error term.

$$\pi_t = \pi_t^D \beta_q \Delta q_{t-1/t-4} + \epsilon_t^\pi \tag{4.7}$$

The Bank Rate, i, is assumed to follow a forward-looking Taylor rule. It is a function of lagged interest rates — to capture inertia related to policy uncertainty — the expected output gap, the expected deviation of domestically-generated inflation from target and can be subject to a stochastic shock, ϵ^i.

In the model, the persistence parameter is set at 0.9, reflecting the substantial degree of inertia in policymaking. That's very close to the parameter estimated in COMPASS. The forecast horizon is selected based on the transmission mechanism of monetary policy and the fact that the Bank of England targets forecasts of inflation, not actual inflation.

The inclusion of the output gap and inflation in the central bank's reaction function are consistent with one another - a closed output gap will lead to inflation settling on the

2% target. It's worth knowing then that when the economy is experiencing a demand shock, it is the sum of the coefficients on output and inflation that determine the amount of monetary activism.

When the economy experiences other shocks as well, there is the possibility that a persistent period of above target inflation will coincide with weakness in demand. In this case the central bank must balance its desire to return inflation to target with a need to keep output stable. Such a trade-off has been assessed by Carney (2017), with reference to the sterling-induced period of above-target inflation and weak demand.

Because I am primarily interested in the impact of shocks that do not create this trade-off, the precise weighting is not too important to the results. The parameter chosen is a half.

$$i_t = \delta_i I_{t-1} + (1-\delta_i)\delta_y y_{t+2_{|t}} + (1-\delta_i)\delta_\pi(\pi_{t+4_{|t}}) + \epsilon_t^i \tag{4.8}$$

The real interest rate is the difference between nominal rates and expected inflation.

$$r_t = i_t - \pi_{t+1_{|t}} \tag{4.9}$$

Foreign interest rates are exogenous to the model and follow an autoregressive process:

$$i_t^f = \zeta_{if} i_{t-1}^f + \epsilon_t^{if} \tag{4.10}$$

As does the deviation of inflation from target in other countries:

$$\pi_t^f = \zeta_\pi \pi_t^f + \epsilon_t^{\pi f} \tag{4.11}$$

The real exchange rate is determined by relative interest rates and prices. The short-term dynamics are driven by a real uncovered interest rate parity condition and the medium-term anchor is purchasing power parity, toward which the exchange rate gradually converges. The equation is linearised around that steady state and the speed of convergence is determined by the parameter, ψ_q. The exchange rate can be subject to shocks, which are denoted by ϵ^q.

The choice of convergence speed is based on the general view in the literature - see Rogoff (1996) - that the half-life of shocks to PPP is somewhere between three and five years. The parameter chosen here is consistent with a half-life of four years.

$$q_t = q_{t+1_{|t}} + (i_t - \pi_{t+1_{|t}}) - (i_t^f - \pi_{t+1_{|t}}^f) - \psi_q q_{t+1} + \epsilon_t^q \tag{4.12}$$

Public sector net borrowing as a share of GDP, pb_t, is simply the sum of the cyclical component of public borrowing, cb_t, and the structural part, sb_t.

$$pb_t = cb_t + sb_t \tag{4.13}$$

The cyclical budget deficit is a function of the average output gap over the past four quarters multiplied by the cyclical adjustment coefficient, θ_{CA1}, and the average output gap in the four quarters before that, multiplied by the cyclical adjustment coefficient, θ_{CA2}. The cyclical adjustment parameters are drawn from Helgadottir et al (2012).

$$cb_t = \theta_{CA1}(y_{t/t-3t}/4)\theta_{CA2}(y_{t-4/t-7t}/4) \tag{4.14}$$

The structural budget deficit is a function of itself and discretionary fiscal measures. The structural balance is subject to two shocks The first is fiscal and denoted by ϵ_t^{sb}. The other reflects shifts in the denominator - potential GDP and is denoted by ϵ^{y^*}.

How much the structural balance deteriorates when the potential size of the economy experiences a negative shock is given by τ. The logic is the same as that pertaining to the cyclical adjustment coefficients. Spending rises as a share of GDP and the tax take falls when growth slows. The difference is that it doesn't automatically get better. The coefficient chosen is 0.7.

$$sb_t = sb_{t-1} - f_t + \epsilon_t^{sb} + \tau\epsilon^{y^*} \tag{4.15}$$

Except the fiscal policy rules, the remaining equations are identities included only to allow the GDP growth condition to operate. The change in the log of potential GDP, gdp^*, between the current time period and four periods ago is given by g, the long-term trend GDP growth rate analogous to the drift parameter of the multivariate Kalman filter presented in Murray (2014), and the sequence of shocks to the level of potential GDP over the preceding 4 quarters.

$$\Delta gdp_t^* = g + \epsilon_t^{y^*} + \epsilon_{t-1}^{y^*} + \epsilon_{t-2}^{y^*} + \epsilon_{t-3}^{y^*} \tag{4.16}$$

The growth of actual GDP is simply the above combined with the change in the output gap over the same period. This is also illustrated back in Figure 1.

$$\Delta gdp_t = \Delta gdp_t^* + (y_t - y_{t-4}) \tag{4.17}$$

The indicator for the fiscal rule, $\text{rule}gdp_t$, is the 4-quarter rolling average of the annual GDP growth rate, given by:

$$\text{rule}gdp_t = \Delta gdp_{t/t-3t}/4 \tag{4.18}$$

The two models that capture the reference and alternative regimes differ only in the specification of the fiscal rule. The fiscal rule in question instructed the government to balance the books over five years when growth is above 1% and provides fiscal flexibility when the pace of expansion drops below that. The two equations which govern fiscal effort attempt to replicate this rule.

4.3.1 Fiscal Regime-Switching Model with a Growth Knock-Out

In the reference regime, when the economy is growing normally, fiscal policy eliminates a proportion of the structural deficit or surplus in every quarter. Though the rule refers to the headline deficit, the cycle has usually run its course within five years, in which case targeting the structural deficit is consistent with targeting the headline deficit.

The speed of deficit reduction is determined by and the parameter is chosen reflecting the revealed preference of the Coalition government — to eliminate the structural deficit in a five-year period.

$$f_t = \zeta_{sb} sb_t + \epsilon_t^f \tag{4.19}$$

In the alternative regime, when growth is weak, the automatic stabilisers are permitted to function and no effort is made to reduce the structural deficit.

$$f_t = \epsilon_t^f \tag{4.20}$$

A constraint is specified to determine when the alternative regime is in place.

$$\text{rule}gdp_t < 1\% = \text{TRUE} \tag{4.21}$$

The relaxation condition determines when the reference regime returns.

$$\text{rule}gdp_t > 2\% = \text{TRUE} \tag{4.22}$$

This latter constraint is not actually described in the fiscal rule itself, but some assumption over when fiscal tightening can resume is needed. The choice of 2% reflects the notion that the economic recovery should be underway before tightening begins — growth of that speed is roughly consistent with the trend rate in the U.K. A choice of, say 2.5%, would delay further the implementation of consolidation.

4.3.2 A Fiscal Regime-Switching Model with an Output Gap Knock-Out

The model is identical to the model described above, but the equations needed to calculate rule GDP are dropped. The 4 quarter rolling average of the output gap is used to switch between the reference and alternative regimes. Given the U.K. economy's trend rate of growth, an output gap of about -1% is comparable to the growth knock out. It is also assumed that the recovery is almost complete before returning to the reference regime.

$$\text{Constraint}: (y_{t/t-3})/4 < -1\% = \text{TRUE} \tag{4.23}$$

$$\text{ConstraintRelax}: (y_{t/t-3})/4 > -0.1\% = \text{TRUE} \tag{4.24}$$

In both models, no effort is made to assess the implications for debt or the feedback of that to the economy. Experience suggests that higher but sustainable debt has not had a particularly damaging influence on the U.K. economy. That doesn't mean it would always be costless, as it is possible that debt servicing costs rise sharply once a threshold is met. Reinhart & Rogoff (2010) found evidence that countries with debt in excess of 90% of GDP tend to grow more slowly than those that do not. The U.K has skirted that level but not breached it since the global financial crisis. Subsequent studies have also cast doubt on the existence of that threshold. Égert (2015) finds that the nonlinear relationship between debt and growth is "quasi inexistent". Still, this feature could be incorporated as an occasionally binding constrain. I leave it as a possible avenue for further research.

Why not use a bigger model, like COMPASS? One reason that is not possible is that the Bank's main model does not capture the impact of fiscal policy on the economy very well.

But there are other reasons to use something else - the purpose of the model is very different. COMPASS is intended to capture an enormous range of shocks to the U.K. economy, which are relevant to the conduct of monetary policy. I am interested in just a few.

For example, how wage growth responds to cost mark-up shocks among importing firms may be interesting to policymakers, but it's of no relevance to gauging the impact of fiscal policy on the economy. The additional complexity of COMPASS comes therefore comes at the significant expense of tractability. Complexity makes it harder to interrogate the simulation results.

Tractability is a good thing, but not if it compromises the ability of the model to approximate the real world. In other words, is a small model too small to be useful? The answer is no. Del Negro & Schorfheide (2012) show that the forecast performance of a medium-sized model like COMPASS, but in this case the Smets & Wouters (2003) model used for a time by the European Central Bank, isn't vastly different to a small-scale model. The latter performed slightly better at longer forecast horizons, while the former did better at shorter horizons.

Separately, it's worth considering evidence on model selection from another strand of literature - Structural Vector Auto-Regression Modelling. SVARs are typically used for assessing the sources of shocks to the economy, rather than for policy analysis. , there is a link between the two — many DSGE models can be represented as SVARS, albeit with fairly onerous restrictions Giacomini (2013).

SVAR models are often small, featuring just three or four variables, but they can often explain a big proportion of the variability of the data. Even early examples in the literature, such as Shapiro & Watson (1988) show that the bulk of business cycle fluctuations can be explained by just a handful of shocks. This suggests that a parsimonious approach to modelling has empirical support.

As a very rough gauge, it is possible to assess the fit of the model to the data by comparing the volatility of simulated variables with those observed in reality. There are some caveats, of course. First, neither fiscal policy rule implemented in the simulations is comparable to how policy was conducted in reality over the past half century - the rule changed over time, for example. And I have already stripped out the influence of discretionary fiscal policy from the shock sequence applied to the model. This renders meaningless a direct comparison output gap volatilities – and indeed they differ materially.

Yet in other respects, this diagnostic is helpful. The volatility of inflation in the simulations should be well below that observed in reality because only economic demand and supply shocks, not import cost or oil price shocks are implemented in the simulations run below. And, reassuringly, that is what we see. The standard deviation of the inflation gap is 0.71 in the growth knockout simulations run below, considerably smaller than the 2.32 standard deviation observed in reality.

A more meaningful comparison perhaps is of the volatility of the interest rate gap. A big difference here would imply that monetary policy is conducted very differently in the model to how it is in real life. The standard deviation of the interest rate gap in the model, when a fiscal growth knockout is in place is 1.82, while it is 1.87 in real life – the two are therefore pretty close. In the simulations that ask more of monetary policy because fiscal policy tries to eliminate the deficit even during a downturn, the standard deviation of the interest rate gap rises above that observed in real life, to 2.26. This makes some logical sense, too.

4.4 Simulations

4.4.1 Part 1: The benefit of fiscal flexibility with a growth knock out

To assess the potential benefits of fiscal policy flexibility I assess how the growth knock out rule performs in the face of unexpected supply and demand shocks that actually

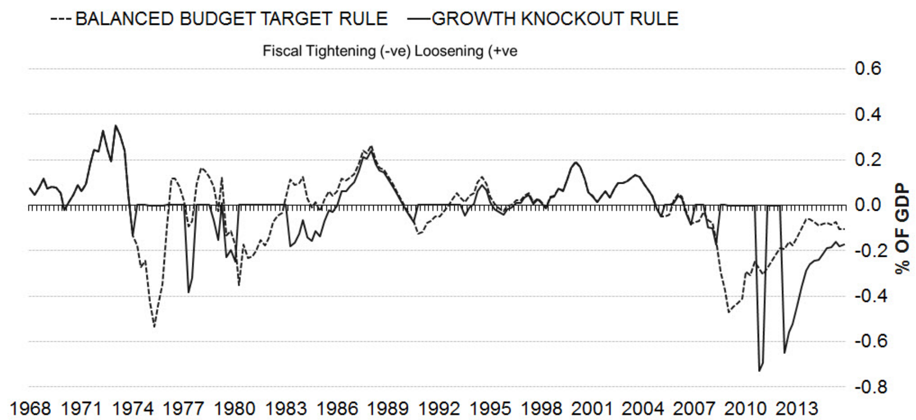


Figure 4.3: The figure shows how different policy rules affect the stance of fiscal policy. The solid line is net fiscal loosening or tightening as a percent of GDP under a balanced budget rule. The dashed line is under a regime in which policy tightening can be abandoned when growth slows.

occurred in history (from 1968 to 2015). The results are compared to a simulation in which the model is the same but no knockout exists and the government targets a balanced structural deficit at the five-year horizon. The results are plotted in the figures below and there are several noteworthy findings.

First, compared with targeting a balanced budget, the growth knockout rule affords significant fiscal flexibility. I find that the growth knockout might have been in operation around a quarter of the time between 1968 and 2016. That flexibility would have been granted during economic downswings in every decade.

A few of specific periods stand out. First, the growth knockout would have applied during the early 1980s, which is an interesting episode in Britain's fiscal history. It is now thought that austerity, especially, in 1981 was an error Nickell (2006) and that the 364 economists who signed an open letter of complaint, were right to do so. Had the growth-knockout rule been in place then, the letter might never have been written because austerity would have been delayed.

Significant flexibility would also have been granted during the global financial crisis, but also during the euro crisis of 2011. Again, these are interesting periods of Britain's past. The actual fiscal response to the crisis was to ease policy, just as the growth knockout would have allowed. Consolidation was then initiated in 2010, again, as the growth knockout would indicate. But, had the growth knockout been in place, the tightening could have been abandoned in 2011 as the euro crisis was raging and would not have been resumed until the worst of it was over. Instead, the government pressed ahead with discretionary fiscal tightening, mostly via spending restraint.

So a growth knock-out rule would likely have afforded fiscal flexibility at times that

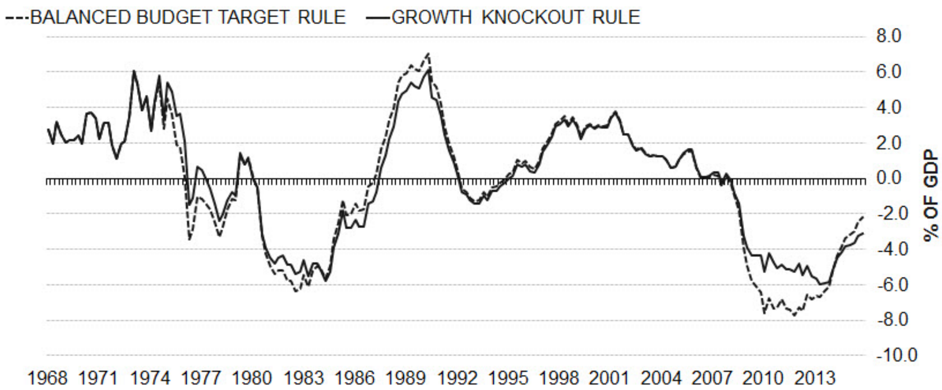


Figure 4.4: The figure shows two paths of the output gap under different fiscal regimes

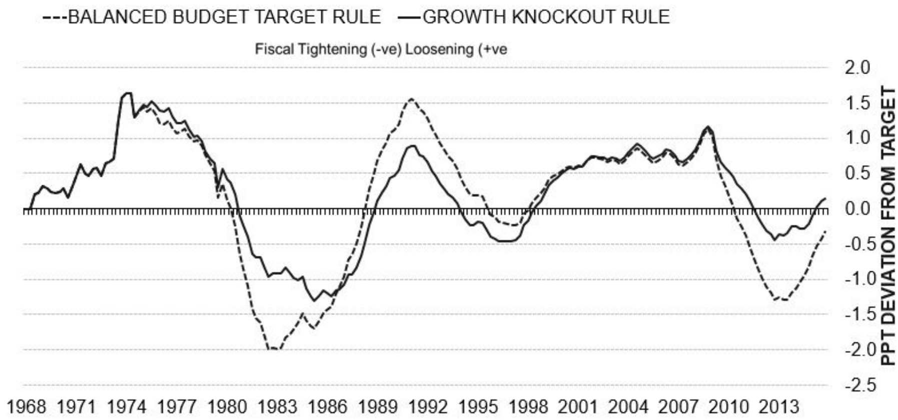


Figure 4.5: The figure shows two paths of the inflation deviation from target under different fiscal regimes

it may have been needed. But what difference would it have made? The simulations show that the margin of spare capacity is smaller during downturns when fiscal policy was more flexible. That is consistent with what economic theory would suggest.

A separate finding, which follows from the narrower margin of economic slack during downturns afforded by fiscal flexibility, is that inflation deviation from the target would be smaller as well. That is illustrated in 4.5.

4.4.2 Part 2: The benefit of fiscal flexibility with an output gap knock out

What the simulations shows is that, by and large, the knockout would be triggered by the output gap at approximately the same times as it would have been by GDP growth. That suggests that the distinction between demand and supply shocks to growth isn't too

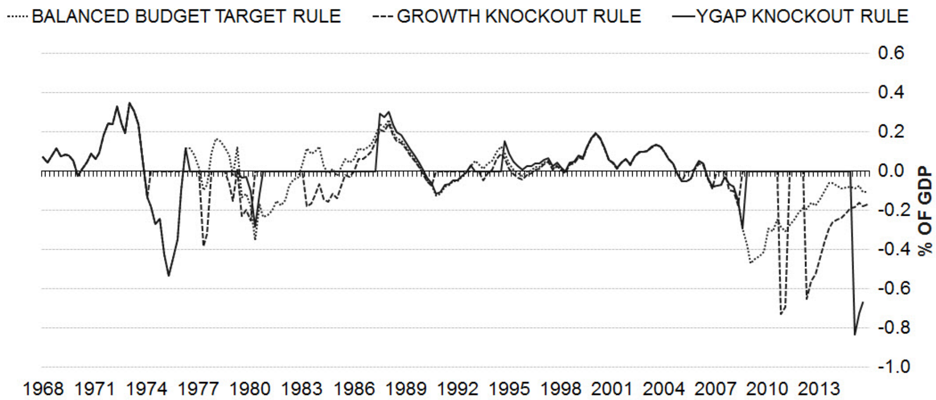


Figure 4.6: The figure plots discretionary fiscal policy tightening or loosening as a share of GDP under three regimes – a balanced budget rule, a growth knowckout rule and an output gap knockout rule.

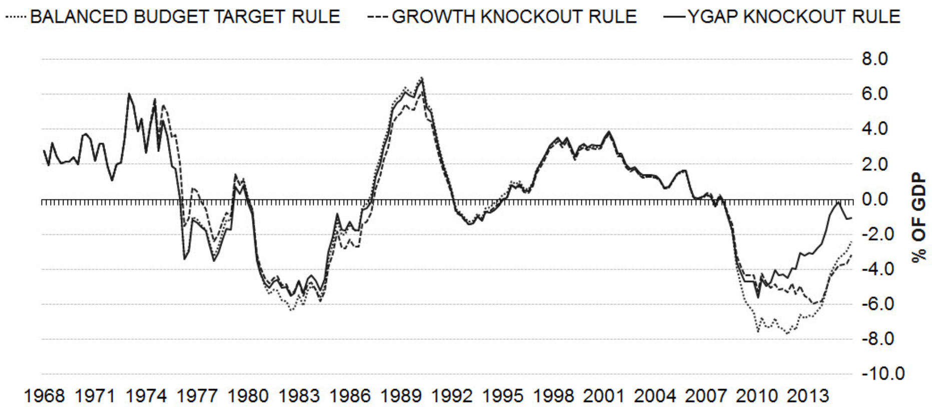


Figure 4.7: The figure shows that the output gap would have closed sooner if the margin of spare capacity is used to determine when fiscal tightening begins after a negative shock. The output gap under all three policy regimes is plotted as a percent of GDP.

important — that may be because demand and supply shocks tend to go in the same direction during deep recessions. The path of fiscal tightening under all three regimes — structural balance targeting, growth knockout and output gap knockout are illustrated in Figure 4.6.

A second observation is that the output gap rule would have delayed fiscal consolidation right the way up to 2015, following the global financial crisis. That's a reflection of the assumption that the output gap is closed before tightening occurs, which is arbitrary. As the chart below illustrates, the output gap would have closed far sooner following the financial crisis if fiscal consolidation had been delayed.

Of the two findings, it is the former that is more important. That the output gap doesn't offer any obvious benefits compared with GDP growth as a trigger for fiscal flexibility is useful, since GDP is far easier to measure and communicate than the output gap. It is less useful to know that the recovery would have been swifter if the output gap trigger has been used because it depends crucially on the assumption about how long the government waits before initiating fiscal tightening.

Neither finding means the output gap should be disregarded as an indicator altogether. Some estimate of spare capacity, implicit or explicit, is always needed to judge how much fiscal effort will have to be made to achieve headline budget deficit targets — that's because revenue and spending are affected by the economic cycle.

4.5 Conclusion

It has been shown that a state-contingent fiscal policy rule can be embedded in a small model of the economy and solved using linear perturbation methods. The resulting simulations show that a growth knockout affording fiscal flexibility in the face of significant shocks to the economy can make recessions shallower.

Though there are theoretical benefits to using the output gap as a growth trigger, rather than GDP growth, in practice the differences appear to be quite small. It may therefore be simpler and more transparent just to let GDP growth determine when fiscal flexibility is warranted.

Given the literature on the merits of deficit rather than debt targeting, the presence of parameters in the rule that could be flexed to manage debt levels in the long-term and the finding that the knockout helps stabilise demand, the government's rule doesn't look too bad.

One problem throughout the modelling exercise is that the government has stated when the deficit targeting can be relaxed, but not when it should bind again. Reasonable assumptions have been made for the purposes of running the simulations - that the government doesn't tighten into an ongoing downturn - but the rule would probably benefit from stating that clearly.

The flexibility of the linear perturbation solution algorithm means there are plenty of avenues for further research. The models used here assume that the fiscal multipliers are the same both during recessions and in normal times - further research could flex this assumption.

Another path forward would be to assess the interaction between fiscal and monetary policy at the zero lower bound. In the analysis presented here it is assumed the central bank finds other effective ways to ease policy when rates are at rock bottom - that may not be a reasonable assumption.

5

Conclusions and future work

5.1 Summary of findings

In chapter two I presented and estimated a small model of the New Zealand economy. I then ran a number of fiscal consolidation scenarios and used the results to show the sensitivity of the fiscal impact multiplier and the associated cumulative output losses to uncertainty over the model parameters.

The key findings are that uncertainty surrounding the effects of fiscal consolidations on output can be attributed to several model parameters and that a bad outcome is likely to be worse than a good outcome is to be better. I find that, if monetary policy were to be constrained by the zero-lower bound, the estimated fiscal impact multiplier for New Zealand would be broadly consistent with estimates of the fiscal multiplier in a number of other OECD countries in that position.

Overall, the evidence suggests that fiscal policy makers should be sensitive to the prevailing economic environment when determining the fiscal stance and work closely with central banks if the worst outcomes are to be avoided.

In chapter three I identified three sources of output gap revision - the arrival of new data, revisions to past data and shifts in model specification. I showed that it is the arrival of new data that matters the most - the benefit of hindsight allows us to say with much greater accuracy whether growth was sustainable or not. This uncertainty carries over to estimates of structural borrowing, complicating fiscal policymaking.

I also demonstrated that the assumptions underpinning some methods mean they are more likely to be revised and, while other methods are less prone to changes, they may not be any more useful in forecasting the path of actual GDP - so no method should be discarded based only on its tendency to be revised.

A wide range of information should be brought to bear on the assessment of slack in the economy. In chapter three I develop a multivariate Kalman filter model, which incorporates data on GDP, inflation, capacity utilisation and unemployment, while allowing for the application of judgement on the underlying process driving productivity growth and the relative importance of each signal from the data — it is still in use by the U.K. government some four years later.

In chapter four I showed that a state-contingent fiscal policy rule can be embedded in a small model of the economy and solved using linear perturbation methods. The resulting simulations show that a growth knockout affording fiscal flexibility in the face of significant shocks to the economy can make recessions shallower.

Though there are theoretical benefits to using the output gap as a growth trigger, rather than GDP growth, in practice the differences appear to be quite small. It may therefore be simpler and more transparent just to let GDP growth determine when fiscal flexibility is warranted. On balance, the knock-out mechanism introduced in the U.K. in 2015 looks like a worthwhile innovation and it is a pity it was abandoned a year later.

5.2 Final remarks

Finance ministries must not operate in a vacuum. There are very clear benefits to macroeconomic policy coordination between fiscal and monetary authorities, as chapter two illustrates. One possible logical step forward would be for a process to be put in place by which the central bank assesses its capacity to offset fiscal tightening before that adjustment takes place.

Chapter three demonstrated that output gap uncertainty is a major headache for policymakers. It also showed that all forecasts for the economy embody a judgement about slack, whether explicit or implicit. Rather than dismiss output gaps as too unreliable to be of use, it would more constructive to keep on refining the methods used to estimate it.

Fiscal rules are intended to bolster credibility, yet they tend to be broken. Specifying the circumstances in which that is ok might mean they endure a little longer than they have in the past. The findings of chapter four suggest that simple is probably best when it comes to defining a trigger for fiscal flexibility. Central banks spend a lot of time explaining how they will respond in different situations. Why shouldn't finance ministries define reaction functions too?

References

ALICHI, A., BIZIMANA, O., LAXTON, M.D., TANYERI, K., WANG, H., YAO, J. & ZHANG, F. (2017). Multivariate filter estimation of potential output for the united states. *International Monetary Fund*. 114

ARMSTRONG, J.S. (2001). Principles of forecasting: a handbook for researchers and practitioners. *Springer Science & Business Media*, **30**. 88

ATKINSON, A.B. (2006). Measurement of government output and productivity. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **169**, 659–662. 103

AUERBACH, A. & GORODNICHENKO, Y. (2011). Fiscal multipliers in recession and expansion, nber working paper 17447. *NBER Working Papers*. 120

BARNES, S.N., PRICE, S. & BARRIEL, M. (2008). The elasticity of substitution: evidence from a uk firm-level data set. *Bank of England Working Paper 348*. 82

BASELCOMMITTEE (2010). An assessment of the long-term economic impact of stronger capital and liquidity requirements `https://www.bis.org/publ/bcbs187.pdf`. 79

BATINI, N. & HALDANE, A. (1999). Forward-looking rules for monetary policy. *Monetary policy rules - Chicago Press*, 157–202. 12, 118

BATINI, N., JACKSON, B. & NICKELL, S. (2005). An open-economy new keynesian phillips curve for the uk. *Journal of Monetary Economics*, **52**, 1061–1071. 15, 16

BENES, J. & N'DIAYE, P.M. (2004). A multivariate filter for measuring potential output and the nairu application to the czech republic. *International Monetary Fund Working Paper 04/45*. 54

BENEŠ, J., BINNING, A., FUKAC, M., LEES, K. & MATHESON, T. (2009). Kitt: Kiwi inflation targeting technology. *Reserve Bank of New Zealand, Model Documentation*. 23

BENEŠ, J., CLINTON, K., GARCÍA-SALTOS, R., JOHNSON, M., LAXTON, D., MANCHEV, P. & MATHESON, T. (2010). Estimating potential output with a multivariate filter. *International Monetary Fund, Working Paper 10/285*. 64, 91, 114

BEVERIDGE, S. & NELSON, C.R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary economics*, **7**, 151–174. 57

BHATTARAI, S. & EGOROV, K. (2016). Optimal monetary and fiscal policy at the zero lower bound in a small open economy. *Dallas Fed Working Paper, no. 260*. 117

BLAGRAVE, P., GARCIA-SALTOS, R., LAXTON, D. & ZHANG, F. (2015). A simple multivariate filter for estimating potential output. *International Monetary Fund, Working Paper 15/79*. 114

BLANCHARD, O. & PEROTTI, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *the Quarterly Journal of Economics*, **117**, 1329–1368. 7, 8

BLANCHARD, O.J. & LEIGH, D. (2013). Growth forecast errors and fiscal multipliers. *American Economic Review*, **103**, 117–20. 8, 41, 120

BOE (2014). Inflation report February 2014. 79

BOUIS, R., COURNÈDE, B. & CHRISTENSEN, A.K. (2013). Implications of output gap uncertainty in times of crisis. *OECD Working Paper 1081*. 96

BRAINARD, W.C. (1967). Uncertainty and the effectiveness of policy. *The American Economic Review*, **57**, 411–425. 20, 48

BURGESS, S., FERNANDEZ-CORUGEDO, E., GROTH, C., HARRISON, R., MONTI, F., THEODORIDIS, K., WALDRON, M. *et al.* (2013). The bank of england's forecasting platform: Compass, maps, ease and the suite of models. *Bank of England Working Paper, 471*. 119

CALVO, G.A. (1983). Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, **12**, 383–398. 15

CAMPBELL, J.Y. & MANKIW, N.G. (1989). Consumption, income, and interest rates: Reinterpreting the time series evidence. *NBER macroeconomics annual*, **4**, 185–216. 11

CANOVA, F. (1998). Detrending and business cycle facts. *Journal of monetary economics*, **41**, 475–512. 58

CARLIN, W. & SOSKICE, D. (2010). A new keynesian open economy model for policy analysis. *CEPR Discussion Paper 7979*. 12, 18

CARNEY, M. (2017). Lambda. *Speech at the London School of Economics*. 122

CASTELNUOVO, E. & LIM, G. (2019). What do we know about the macroeconomic effects of fiscal policy? a brief survey of the literature on fiscal multipliers. *Australian Economic Review*, **52**, 78–93. 9

CAYEN, J.P. & VAN NORDEN, S. (2005). The reliability of canadian output-gap estimates. *The North American Journal of Economics and Finance*, **16**, 373–393. 91

CESA-BIANCHI, A., THWAITES, G. & VICONDOA, A. (2016). Monetary policy transmission in an open economy: new data and evidence from the united kingdom. *Bank of England Working Paper 615*. 119

CLARIDA, R., GALI, J. & GERTLER, M. (1999). The science of monetary policy: a new keynesian perspective. *Journal of economic literature*, **37**, 1661–1707. 15, 20

CLAVERES, G. & STRÁSKỲ, J. (2018). Euro area unemployment insurance at the time of zero nominal interest rates. *OECD Working paper 2018/46*. 117

CORSETTI, G., MEIER, A. & MÜLLER, G.J. (2012). What determines government spending multipliers? *Economic Policy*, **27**, 521–565. 8

CÚRDIA, V. & WOODFORD, M. (2016). Credit frictions and optimal monetary policy. *Journal of Monetary Economics*, **84**, 30–65. 21

D'AURIA, F., HAVIK, K., MC MORROW, K., PLANAS, C., RACIBORSKI, R., ROGER, W., ROSSI, A. *et al.* (2010). The production function methodology for calculating potential growth rates and output gaps. *European Economy, Ecomomic Papers 420*. 82

DAVIG, T. & LEEPER, E.M. (2011). Monetary–fiscal policy interactions and fiscal stimulus. *European Economic Review*, **55**, 211–227. 8

DEL NEGRO, M. & SCHORFHEIDE, F. (2012). Forecasting with dsge models: theory and practice. *Handbook of forecasting*, **2**. 125

DELONG, J.B., SUMMERS, L.H., FELDSTEIN, M. & RAMEY, V.A. (2012). Fiscal policy in a depressed economy [with comments and discussion]. *Brookings Papers on Economic Activity*, 233–297. 52

ÉGERT, B. (2015). Public debt, economic growth and nonlinear effects: Myth or reality? *Journal of Macroeconomics*, **43**, 226–238. 125

EMMERSON, C., FRAYNE, C. & LOVE, S. (2001). The government's fiscal rules. *IFS briefing note 16*. 108

EMMERSON, C., JOHNSON, P. & JOYCE, R. (2016). Ifs green budget 2016. *Institute for Fiscal Studies. https://www.ifs.org.uk/publications/8154*. 111

ESTRELLA, A. (2007). Extracting business cycle fluctuations: What do time series filters really do? *New York Fed staff report 289*. 58, 62

FERNÁNDEZ-VILLAVERDE, J., GORDON, G., GUERRÓN-QUINTANA, P. & RUBIO-RAMIREZ, J.F. (2015). Nonlinear adventures at the zero lower bound. *Journal of Economic Dynamics and Control*, **57**, 182–204. 117

FLOTHO, S. (2018). Interaction of fiscal and monetary policy in a monetary union under the zero lower bound constraint. *Annals of Operations Research*, **260**, 159–196. 117

FUHRER, J.C. (2000). Habit formation in consumption and its implications for monetary-policy models. *American Economic Review*, **90**, 367–390. 11, 118

FUHRER, J.C. & RUDEBUSCH, G.D. (2004). Estimating the euler equation for output. *Journal of Monetary Economics*, **51**, 1133–1153. 10

GALI, J. & GERTLER, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, **44**, 195–222. 16, 19

GALI, J. & MONACELLI, T. (2005). Monetary policy and exchange rate volatility in a small open economy. *The Review of Economic Studies*, **72**, 707–734. 14, 16, 17, 64

GIACOMINI, R. (2013). The relationship between DSGE and VAR models. *CeMMAP working papers, CWP21/13*. 126

GIANNONI, M. & WOODFORD, M. (2004). Optimal inflation-targeting rules. *NBER chapter 9557, University of Chicago Press*, 93–172. 11

GOODFRIEND, M. (1991). Interest rates and the conduct of monetary policy. In *Carnegie-Rochester conference series on public policy*, vol. 34, 7–30, Elsevier. 19

GREENSLADE, J.V., PIERSE, R.G. & SALEHEEN, J. (2003). A kalman filter approach to estimating the UK NAIRU. *Bank Of England Discussion Paper no. 12*. 101

GUERRIERI, L. & IACOVIELLO, M. (2015). OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics*, **70**, 22–38. 116

GUERRIERI, L. & IACOVIELLO, M. (2017). Collateral constraints and macroeconomic asymmetries. *Journal of Monetary Economics*, **90**, 28–49. 117

HAMILTON, J.D. (2017). Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics*. 52, 58

HARRISON, R. & OOMEN, O. (2010). Evaluating and estimating a dsge model for the united kingdom. *Bank of England Working Paper 380*. 18

HAUG, A. (2016). A new test of ricardian equivalence using the narrative record on tax changes. *MPRA Paper no. 75452*. 120

HAYO, B. & NEUMEIER, F. (2017). The (In)validity of the Ricardian equivalence theorem - findings from a representative German population survey. *Journal of Macroeconomics*, **51**, 162–174. 120

HEINEMANN, F., MOESSINGER, M.D. & YETER, M. (2018). Do fiscal rules constrain fiscal policy? a meta-regression-analysis. *European Journal of Political Economy*, **51**, 69–92. 107

HELGADOTTIR, T., CHAMBERLIN, G., DHAMI, P., FARRINGTON, S. & ROBINS, J. (2012). Cyclically adjusting the public finances. *Office for Budget Responsibility (OBR) Working Paper No*, **3**. 46, 97

HM-TREASURY (2008). Pre-Budget Report 2008. *The Stationery Office*. 109

HM-TREASURY (2010). Budget 2010: Securing the recovery, economic and fiscal strategy report and financial statement and budget report. 109

HM-TREASURY (2013). Forecasts for the UK economy: a comparison of independent forecasts December. *The Stationery Office*. 88

HM-TREASURY (2015). Charter for budget responsibility: Autumn 2015 update. *The Stationery Office*. 107

HM-TREASURY (2016). Charter for budget responsibility: Autumn 2016 update. *The Stationery Office*. 107, 112

HODRICK, R.J. & PRESCOTT, E.C. (1997). Postwar us business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 1–16. 50, 113

HUGHES HALLETT, A., KATTAI, R. & LEWIS, J. (2012). How reliable are cyclically adjusted budget balances in real time? *Contemporary Economic Policy*, **30**, 75–92. 96

IAKOVA, M.D.M. (2007). Flattening of the Phillips curve: Implications for monetary policy. *International Monetary Fund, paper number 07/76*. 67

ILZETZKI, E., MENDOZA, E.G. & VÉGH, C.A. (2013). How big (small?) are fiscal multipliers? *Journal of monetary economics*, **60**, 239–254. 8

IMF (2012a). United Kingdom 2012 Article IV Consultation. *IMF*. 52

IMF (2012b). World economic outlook: growth resuming, dangers remain. *IMF Washington, DC*. 23

JÄÄSKELÄ, J.P. & JENNINGS, D. (2011). Monetary policy and the exchange rate: Evaluation of var models. *Journal of International Money and Finance*, **30**, 1358–1374. 27

KALMAN, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, **82**, 35–45. 51, 112

KARAGEDIKLI, Ö., RYAN, M., STEENKAMP, D. & VEHBI, T. (2013). What happens when the kiwi flies? the sectoral effects of the exchange rate shocks. *CAMA Working Paper 73/2013*. 23

KONUKI, T. (2008). Estimating potential output and the output gap in slovakia. *International Monetary Fund*. 90

KOSKE, I. & PAIN, N. (2008). The usefulness of output gaps for policy analysis. *OECD Working Paper 621*. 96

KYDLAND, F.E. & PRESCOTT, E.C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, 1345–1370. 58

LALL, M.S., CARDARELLI, M.R. & ELEKDAG, S. (2009). Financial stress, downturns, and recoveries. *International Monetary Fund*. 8

LAXTON, D. & TETLOW, R. (1992). A simple multivariate filter for the measurement of potential output. *Bank of Canada Ottawa*. 54, 62, 114

LEAHY, J.V. & WHITED, T.M. (1995). The effect of uncertainty on investment: Some stylized facts. *NBER Working Paper 4986*. 14

LEEPER, E.M., TRAUM, N. & WALKER, T.B. (2017). Clearing up the fiscal multiplier morass. *American Economic Review*, **107**, 2409–54. 41

LEY, E. & MISCH, F. (2013). Real-time macro monitoring and fiscal policy. *The World Bank*. 96

LUBIK, T.A. & SCHORFHEIDE, F. (2007). Do central banks respond to exchange rate movements? a structural investigation. *Journal of Monetary Economics*, **54**, 1069–

1087. 23

MARCELLINO, M. & MUSSO, A. (2011). The reliability of real-time estimates of the euro area output gap. *Economic Modelling*, **28**, 1842–1856. 91

MILLER, E. (2008). An assessment of CES and cobb-douglas production functions. *Congressional Budget Office*. 82

MOUNTFORD, A. & UHLIG, H. (2009). What are the effects of fiscal policy shocks? *Journal of applied econometrics*, **24**, 960–992. 8

MUELLBAUER, J. (1988). Habits, rationality and myopia in the life cycle consumption function. *Annales d'Economie et de Statistique*, 47–70. 11

MURRAY, J. (2012). A small model of the UK economy. *Office for Budget Responsibility, Working Paper No 4*. 65, 119, 120, 121

MURRAY, J.P. (2014). Output gap measurement: judgement and uncertainty. *Office for Budget Responsibility Working Paper No. 5*. 111, 113, 114, 123

NELSON, E. & NIKOLOV, K. (2004). Monetary policy and stagflation in the uk. *Journal of Money, Credit and Banking*, 293–318. 19

NICKELL, S. (2006). The budget of 1981 was over the top, Bank of England speech. *Were*, **364**, 54–61. 127

NILSSON, R. & GYOMAI, G. (2011). Cycle extraction: A comparison of the phase-average trend method, the hodrick-prescott and christiano-fitzgerald filters. *OECD Statistics Working Paper 2011/04*. 58

OBR (2010). June Budget 2010. *The Stationery Office*. 120

OBR (2015). Office for Budget Responsibility: Economic and fiscal outlook July. *The Stationery Office*. 120

OKUN, A.M. (1963). Potential GNP: its measurement and significance. *Yale University, Cowles Foundation for Research in Economics New Haven*. 68

ORPHANIDES, A. & NORDEN, S.V. (2002). The unreliability of output-gap estimates in real time. *Review of economics and statistics*, **84**, 569–583. 91, 111

PARKYN, O. & VEHBI, T. (2014). The effects of fiscal policy in new zealand: Evidence from a var model with debt constraints. *Economic Record*, **90**, 345–364. 8, 23, 41

PEDERSEN, T.M. (2001). The Hodrick–Prescott filter, the Slutzky effect, and the distortionary effect of filters. *Journal of economic dynamics and control*, **25**, 1081–1101. 51

PHILIP, R. & JANSSEN, J. (2002). Indicators of fiscal impulse for New Zealand. *New Zealand Treasury. Working Paper*, **2**, 30. 12

PINDYCK, R.S. & SOLIMANO, A. (1993). Economic instability and aggregate investment. *NBER macroeconomics annual*, **8**, 259–303. 14

PORTES, J. & WREN-LEWIS, S. (2015). Issues in the design of fiscal policy rules. *The*

Manchester School, **83**, 56–86. 111

RAMEY, V.A. (2019). Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? 8

RATTO, M., ROEGER, W. *et al.* (2007). Fiscal policy in an estimated open-economy model for the euro area. *European Economy Economic Papers, no. 266*. 12

REINHART, C.M. & ROGOFF, K.S. (2010). Growth in a time of debt. *American Economic Review*, **100**, 573–78. 125

RILEY, J. & CHOTE, R. (2014). Crisis and consolidation in the public finances. *Office for Budget Responsibility Working paper, 7*. 109

ROGOFF, K. (1996). The purchasing power parity puzzle. *Journal of Economic Literature*, **34**, 647–668. 122

ROMER, C.D. & ROMER, D.H. (2010). The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. *American Economic Review*, **100**, 763–801. 120

RUDD, J. & WHELAN, K. (2007). Modeling inflation dynamics: A critical review of recent research. *Journal of Money, Credit and Banking*, **39**, 155–170. 15, 67

RÜNSTLER, G. (2002). The information content of real-time output gap estimates: an application to the euro area. *ECB Working Paper 182*. 111

RYAN, M., SZETO, K.L. *et al.* (2009). An introduction to the New Zealand Treasury model. *New Zealand Treasury Working Paper 0902*. 12, 23

SHAPIRO, M. & WATSON, M. (1988). Sources of business cycle fluctuations. *NBER Working Papers, 2589*. 126

SMETS, F. & WOUTERS, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, **1**, 1123–1175. 12, 125

SÖDERSTRÖM, U. (2002). Monetary policy with uncertain parameters. *Scandinavian Journal of Economics*, **104**, 125–145. 20

SVENSSON, L.E. (1999). Inflation targeting: some extensions. *Scandinavian journal of Economics*, **101**, 337–361. 19

SVENSSON, L.E. *et al.* (1997). Inflation targeting in an open economy: strict or flexible inflation targeting? *Reserve Bank of New Zealand Working Paper G97/8*. 19

SZETO, K.L., GARDINER, P., GRAY, R., HARGREAVES, D. *et al.* (2003). A comparison of the NZTM and FPS Models of the New Zealand economy. *New Zealand Treasury Working Paper no. 03/25*. 23

SZETO, K.L. *et al.* (2013). Estimating new zealands output gap using a small macro model. 26

TAYLOR, J.B. (1993). Discretion versus policy rules in practice. In *Carnegie-Rochester* *conference series on public policy*, vol. 39, 195–214, Elsevier. 18

TAYLOR, J.B. (2001). The role of the exchange rate in monetary-policy rules. *American Economic Review*, **91**, 263–267. 19

TEREANU, E., TULADHAR, M.A. & SIMONE, M.A. (2014). Structural balance targeting and output gap uncertainty. *International Monetary Fund*. 96

TOBIN, J. (1969). A general equilibrium approach to monetary theory. *Journal of money, credit and banking*, **1**, 15–29. 14

WALSH, C. (2003). Speed limit policies: The output gap and optimal monetary policy. *American Economic Review*, **93**, 265–278. 111

WOODFORD, M. (2000). Interest and prices. *Princeton University*. 19