RECOMMENDED READING

How long does it take for monetary policy to have its peak impact?

MASD

Main messages

- Evidence from the empirical macro literature across a range of countries finds that the peak effect on GDP and inflation from a change in monetary policy tends to come in the second year.
- But, on balance, there is some evidence that the peak effects could be felt earlier in the UK.
- There is also evidence of time variation in the timing of the peak impact, although this is mainly associated with the move to inflation targeting around the mid-1990s.
- Using an estimated time-varying structural model, we find some evidence of time variation in the exchange rate channel, although this does not affect the *timing* of the peak impact.

1. Introduction

Conventional wisdom is that the effect of a change in Bank Rate takes time to propagate. For example, recent Open Letters have stated: "Monetary policy takes time to affect the economy: its peak effect on inflation is generally estimated to occur with a lag of somewhere between 18 and 24 months" and Federal Reserve Chair Yellen has recently made a similar remark saying: "We need to keep in mind the well-established fact that the full effects of monetary policy are felt only after long lags." Despite these rules-of-thumb there remains considerable uncertainty over the precise time-lag and it could be affected by a range of factors and vary over time.

This note summarises the findings from our stock-take of the lags in the monetary transmission mechanism.[1] Specifically this work tackles four key questions:

i. How much consensus is there about the timing of the peak impact following a change in monetary policy?
ii. Is there important variation across different countries?
iii. Is there any evidence that the timing of the peak effect may have varied over time?
iv. Given the prominence of the exchange rate in Committee discussions, what is the interaction between the exchange rate channel and the timing of the peak effects of monetary policy?

Background materials on each of these issues can be found in the 3 notes and 2 slide-packs (which provide all the underlying evidence) referenced in the Appendix.

2. Our results come from a range of methods, drawn from a range of countries.

Our focus is on the dynamics of the economy following a change in Bank Rate. We therefore draw on evidence based on macro studies and methods. In doing so, two issues need to be addressed. First, isolating the causal effect of a change in monetary policy is tough. Second, structural features of the economy that affect the speed of policy

[1] Last year a Special Project Team examined the risks around the first Bank Rate increase. Many of the channels discussed in this work are relevant here, although the SPT did not specifically focus on the dynamics in the transmission mechanism. Some of the new econometric work in this note, however, builds on models and results used in the SPT's work.

transmission may vary over time and across countries. Understanding the degree of time variation in the past may therefore be particularly useful when assessing the risks around the impact of changing policy going-forward.

To be as robust as possible against the first issue we review studies that have employed a range of approaches. Specifically we consider studies that have employed traditional Vector Autoregression (VAR) methods (based on an identifying restriction that monetary policy does not affect the economy in the period in which it is altered); the Romer and Romer approach based on readings of the FOMC minutes and Greenbook forecasts; high frequency identification using financial markets data around policy events[2]; and, VAR evidence using sign restrictions (which restrict the initial effects on key variables to be in a particular direction, motivated by theory).

The second issue regarding time and cross-country variation is tackled in a number of ways. First, evidence across a range of countries is considered. Our literature review does not, therefore, simply rely on seminal studies that are typically for the United States. In addition, where possible, we draw on studies which have explicitly considered time variation in the transmission mechanism. We then provide new time-varying estimates for the U.K. by extending earlier work conducted by last year's Special Project Team (see We also make use of time-varying estimates from a fully structural model developed at the Bank by (2015).[3]

Interpreting macro estimates of the effects of monetary policy

Taking a macro approach is natural given the focus on the effects of policy over time. But, in establishing causality, macro methods require some source of exogenous variation in interest rates. The literature has, necessarily, focused on the response to unexpected changes in interest rates – so called policy 'shocks'.

Monetary policy is, of course, set in response to economic conditions to meet the inflation target. Changes in policy that are systematic and anticipated – by following a Taylor rule, for example - may have different observed effects than a surprise change. But one reason for still considering "shocks" is that, under some assumptions, the *marginal* effect of a systematic change is equivalent to the effects of a shock. If we are willing to assume that the effects of a systematic change can be measured relative to a no-change scenario (a deviation from what was already expected), then the marginal impact of following the rule is captured by the effect of a shock. This is the approach taken in Bernanke, Gertler and Watson (1997) where they decompose the systematic component of monetary policy in the face of oil price shocks.

Furthermore, while the literature focuses on "shocks", it is not necessarily the case that we should think of these as purely random disturbances in policy. Sims (2011) has a useful interpretation of shocks from a Bayesian perspective arguing "There is therefore no paradox in supposing that econometricians and the public may have probability distributions over policy maker behaviour, while policy makers themselves do not see their choices as random".

In summary, there are still good reasons – over and above econometric necessity - to focus on the response of the economy to surprise changes in monetary policy. But it is worth keeping in mind that the estimates covered below are still conditional on the assumed monetary policy rule. The overall effects of a policy change are determined by both the overall monetary framework and other structural features of the economy.[4]

[2] There is also a larger literature on event studies around policy changes but these look at the effect on financial markets around the policy decision rather than tracing out the dynamic effects on the real economy over time. While an important literature, we do not consider these papers in this work.

[4] One crucial ingredient is the response of the policy rate after the initial disturbance: the "policy rule" is likely to affect all the other responses. In reviewing the evidence we therefore always consider the dynamics of the policy rate after the initial change. Reassuringly there is some

3. The literature tends to find a peak effect on GDP and inflation after 4-8 quarters with prices falling (relative to the counterfactual) by the end of the first year. But there is variation across countries.

3.1 Highly-cited, seminal, papers from the US tend to find lags of around 6-7 quarters for GDP and inflation

The most highly cited and seminal research has tended to be for the United States. We reviewed 19 papers for the US.[5] Many of these are considered seminal papers in the literature. In all studies considered - both for the US and for other countries below - the estimates are based on the response of the level of GDP or other real variables (in per cent relative to trend) and the response of the price level (relative to trend) or quarterly inflation. Where annual inflation is considered, this is specifically mentioned.

Averaging across studies, **the peak effect on output, as measured by industrial production or GDP, tends to be in the second year** (Table 1). That said there is evidence of quicker effects; some studies find peaks effects after as little as 2 quarters. Furthermore, focusing on the precise peak may be misleading. Even though the peak effect comes later, output tends to start responding relatively quickly and, in some cases, a sizable proportion of the effect at the peak has come through earlier.

For papers that explicitly consider inflation, **the peak effect on quarterly inflation is 7 quarters**. More commonly papers report the response of the price level relative to trend and it is a common feature of US results that, unlike GDP, prices do not start falling (relative to trend) significantly until the second year. There is, however, important variation across methodologies. For example, the high frequency approach (using financial markets surprises) in Gerlter and Karadi (2015) produces a significant effect on prices immediately.

The relative dynamics of different variables tell an intuitive story about the stages of the transmission mechanism. Gertler and Karadi (2015) find that the **peak pass-through to mortgage spreads, the excess bond premium and commercial paper spreads is either on impact or after 1 quarter but the GDP peak still comes after 6 quarters**, suggesting that financial market movements come early in the transmission mechanism. Bernanke, Boivin and Elliaz (2005) find that the effect on housing starts and new orders comes through faster than on GDP or prices, with the peak effect at 2 quarters. Finally, **the effect on unemployment, employment, consumption and investment tends to reflect the estimates for output.** In addition, Boivin, Kiley and Miskin (2010) find that the peak impact on consumer credit comes after 13 quarters and capital utilisation after 8 quarters.

3.2 Evidence from other countries is similar, although the real effects may be felt earlier

There is a sizable literature for other countries, although typically these employ conventional VAR methods. We considered 25 papers covering Germany, France, Italy, Spain, Ireland, Canada, Australia, Japan and the Euro Area as a whole.[6] More detail can be found in the background summary note and the reference slide-pack

The peak effect on output tends to be early in the second year, slightly earlier than in the US. Virtually no study has the peak effect on output at more than two years. There also does not appear to be systematic differences in the

---

general consistency across the studies considered. The rule itself, of course, plays a further role in stabilising the economy and anchoring expectations.

[5] More detail can be found in the background note and all the underlying empirical estimates can be found in the slide pack

[6] The criteria for determining this particular list of studies is discussed in more detail in the background note.

response of output across countries. For the Euro Area as a whole the average peak effect comes after 4 quarters, with relative homogeneity among the major EA countries. The peak effect for Germany and Italy is, on average, after 5 quarters; for France and Ireland it is 4 quarters. The average is also 4 in Canada and 5 quarters in Japan (Table 1). As in the US, it's worth noting that while the peak tends to be later, output still starts responding relatively quickly.

In terms of prices and inflation, **prices start falling significantly, relative to trend, by the end of the first year.** Relative to the US, this tends to be slightly faster. For example, the EA average is 2 quarters for prices to start falling (although results for individual countries can be longer). Only a few studies report inflation and it is hard to get a clear steer from the literature. The inflation numbers in Table 1 therefore need to be interpreted with care, especially where there is only one study per country. Boivin et al (2008) find the peak of **the annual CPI inflation response can be as long as 15 quarters for a range of European economies** although the peak effect on PPI inflation is faster at around 5 quarters.

Prices appear to start falling, relative to the counterfactual, marginally earlier than in the US studies. Consequently, even if the peak effect is further out, **monetary policy can still have a material effect even in the shorter term.** On the other hand, the limited results available for inflation suggest a more sluggish response than in the US (or the UK, as discussed below).

The response of other real variables such as consumption and investment reflect the lags in GDP or industrial production. Boivin et al (2008), again the most detailed study, finds similar time lags for consumption, investment, exports and employment across a range of European economies. Generally, **unemployment and employment indicators also take between 4 and 8 quarters to peak** (although 2 quarters in Canada).

3.3 There are fewer UK studies but the lags are comparable to the wider literature, although slightly shorter than in the US

Results for the UK are sparse compared to other countries and there are 4 main studies we consider. The results for output suggest the peak is felt after 5 quarters, on average. This is comparable to results for other European countries but slightly faster than results for the US. But there is still variation across studies. New evidence we present in the next section from a factor VAR finds a peak effect on GDP after around 6-8 quarters. In the effect on industrial production comes through after around 3 quarters (this persists for a further year).

Estimates for the effect on prices and inflation are harder to come by and are less consistent for the UK. Two of the papers considered find price puzzles (where prices increase strongly after a monetary contraction). In two studies the inflation impact seems to come through relatively quickly. The updated factor VAR (discussed below) suggests a peak effect on quarterly CPI inflation after 2-3 quarters, with the effect on annual inflation coming after 4-6 quarters. In their full sample Cloyne and Hürtgen (2014) find that prices typically do not respond for 6 quarters. But the effect comes through faster in the post-1992 period with prices falling (relative to trend) by the end of the first year and the peak effect on annual inflation coming after 5-6 quarters (which then persists).[7]

In summary, there appears to be broad similarities in the timing of the peak impact of monetary policy across countries, especially for output. But there is evidence that the timing of the peak effect in the UK (and some other

[7] Recent ID work has considered the effect of foreign monetary policy shocks on the UK. Interestingly, the timing of the peak effects sit well with the results presented in this note. For more detail see .

European countries) is slightly faster than in the US.[8] That said, care must be taken in drawing strong conclusions for the UK given the few studies available.

Table 1: Average time to peak effect across studies in the literature by country (quarters)

Country	Output peak	Inflation peak	OR price level starts falling	Exchange rate peak	Wage inflation peak	Consumption peak	Investment peak	Employment or unemployment peak	Exports peak	Bank loans peak	Policy rate changes sign
UK	5	4[9]									9
US	6	7	4			5	5	8		13	6
EA	4	12	2	4	9	7	6	8	5		5
France	4	15[10]	4	7	10	8	6	6	4	8	5
Germany	5	15	2	4	15	9	6	6	6	6	8
Italy	5	15	4	9	11	7	7		3		6
Spain	6	12	6	3	4	7	6	6	3		6
Ireland	4	18	1	1							2
Australia	14	11		1							
Canada	4	1	5	1				2			5
Brazil	1		2								3
Japan	5		6								3
Average	5	11	4	4	10	7	6	6	4	9	5

4. **There is some evidence of time variation in the lags in the UK transmission mechanism, although the major change is around 1993.**

Estimates using a time-varying parameter factor augmented VAR (TVP FAVAR)

The main message from the previous section is that the peak effects of monetary policy on output and quarterly inflation tend to be in the second year, although policy still has important effects at shorter horizons. But it is easy to imagine that these effects might vary across time as the structure of the economy changes, or if policy is conducted differently over time. Consequently, looking at the average effects in the literature as a guide to the impact of future policy decisions could be misleading.

Whether the timing of the peak effect varies over time is explored using a factor augmented structural VAR with time varying coefficients and stochastic volatility. As mentioned in Section 3, the model is an updated version of the time-varying coefficient factor augmented VAR with stochastic volatility developed for the UK by and (2014) and published in the *Economic Journal*. This was originally re-estimated and updated for the Special Project Team in 2014 and covered in . For brevity, this section focuses on the

[8] The broad similarities between the US and the UK might appear surprising given the prevalence of fixed rate mortgages in the US. This issue is the subject of recent research by (2015) who find that the response of mortgage repayments in the US is more sluggish than the UK but the overall magnitudes and dynamics of the effects (on the aggregate and on household consumption and income) are very similar. The paper finds the fixed versus variable distinction seems to play a smaller role in the aggregate than one might originally have thought. Other general equilibrium effects therefore seem at least as important as the cash flow consequences of different mortgage contracts.

[9] Results for UK annual inflation are between 4 and 6 quarters. As noted in the text, the average in this table is only based on two studies for the UK.

[10] Only one study reports inflation for the individual EA countries, which potentially explains why the lags are so long for France, Germany, Italy, Spain and Ireland.

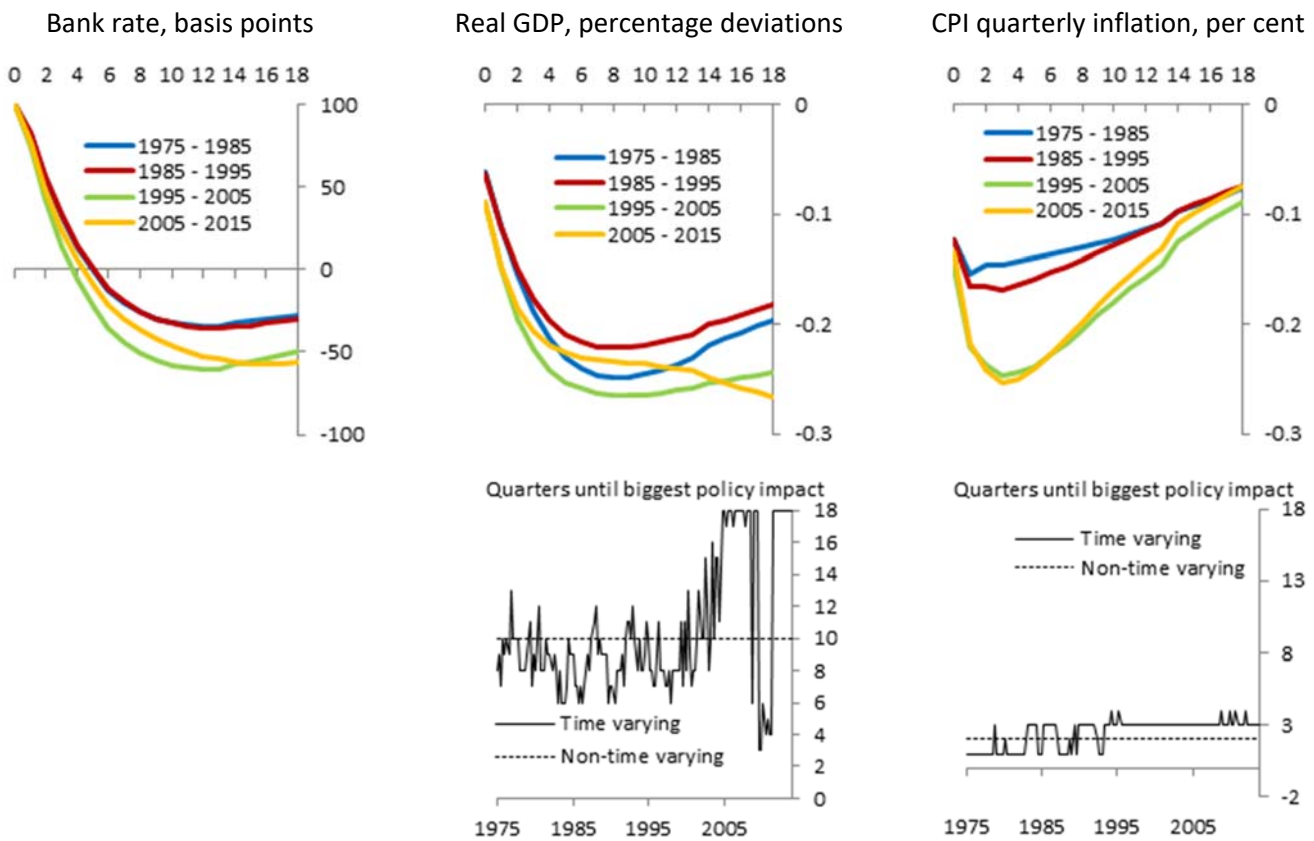
response of GDP and inflation. More details on the methodology can be found in the background note . The background note also shows the responses for a wider range of variables such as the effect on other GDP components, house prices, the exchange rate, import prices, lending and the FTSE All Share Index.

As in the original et al (2014) paper, the clearest sign of time variation is the difference in the responses before and after the mid-1990s for CPI inflation. In terms of the magnitudes for the inflation targeting period, a 1 percentage point in the interest rate, that is unwound over time, produces a 0.25pp fall in quarterly CPI inflation (at peak). The GDP response does vary over time but is more stable than inflation, with the peak percentage fall being around -0.2.[11]

Variation in the timing of the peak impact can be seen in the second row of Figure 1. This shows the peak quarter for all the impulse responses over time. For GDP, this fluctuates between 6 and 12 quarters for much of the sample, and in recent years the persistence has increased, with the response still not peaking before the end of simulation. But the GDP profile after 2005 is relatively flat after around 6 quarters and the precise peak is sensitive to small changes in the profile. The majority of the effect on GDP has therefore come through by quarter 6, suggesting that the results post-2005 are not too different to those pre-2005.

The timing of the peak impact on quarterly inflation is faster than on the level of GDP, occurring between 1 and 3 quarters. That said there is a protracted effect on inflation throughout the period. The timing of the peak effect on inflation also does not seem to have dramatically changed over the sample. Converting the responses into the implied effect on annual inflation suggests that peak is slightly later at 4-6 quarters.

Figure 1: Response of GDP, quarterly CPI and Bank Rate to a change in monetary policy, and the peak impact date



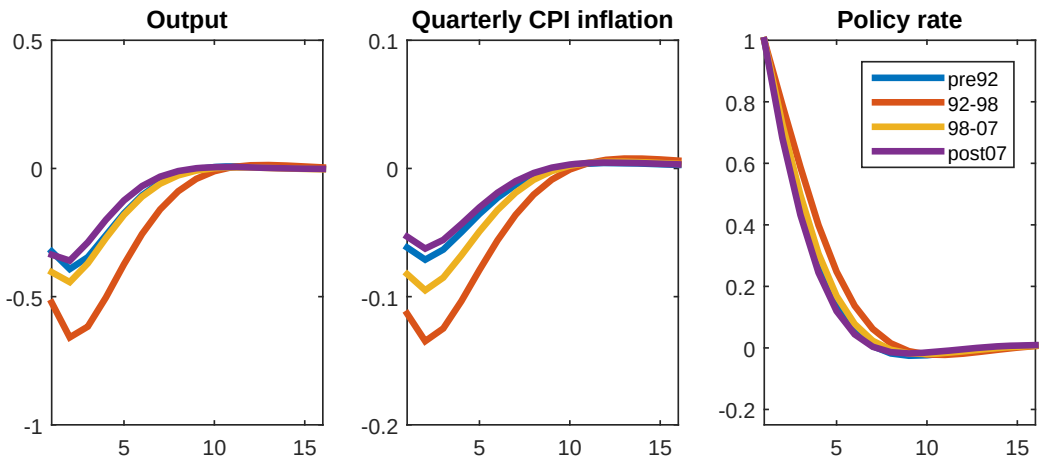
[11] It is also worth noting that these overall *magnitudes* are smaller than in, for example, (2014) who find a peak effect on GDP of around -0.5 and annual inflation of nearly 1 percent.

Results from an estimated structural DSGE model

Additional evidence comes from recent Bank work by (2015) which estimates a time varying version of COMPASS, the Bank's estimated structural DSGE model. This allows us to explore time variation in the effects of monetary policy – as in the empirical FAVAR above – but also provides estimates of how the structural characteristics of the economy may have changed over time.

Figure 2 shows the time-varying response of GDP and quarterly CPI inflation to a 1 percentage point increase in Bank Rate where the policy innovation unwinds according to the Taylor Rule in the model. There is some time variation in the response of GDP and inflation over time, but not in the timing of the peak effects. Monetary policy has its peak effect on both inflation and GDP after around 2 quarters. The effect on annual inflation is slightly later at around 4 quarters. This is clearly towards the lower end of the estimates in the empirical literature discussed in Section 3[12] although there are a few studies that find similar lags. These peak timing estimates are also similar to those from the FAVAR above.

Figure 2: Time varying effects of a monetary policy shock in COMPASS



5. The role of exchange rate in the monetary transmission mechanism has varied over time, although the effect on GDP and inflation seems relatively modest.

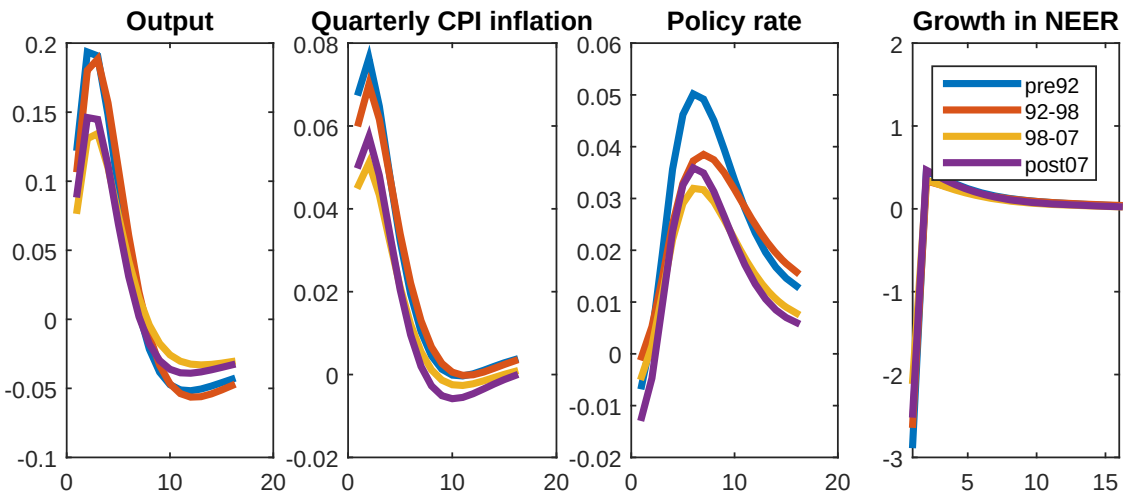
The role of the exchange rate in the monetary transmission mechanism has featured in Committee discussions at various points in recent years. We can also use the time-varying version of COMPASS to explore whether the role of the exchange rate in the monetary transmission mechanism has changed over time. For consistency, the exercises in this section are all conducted using the time varying version of COMPASS.

There is some evidence of time variation in the effect of exchange rates on inflation and GDP….

As a first step in understanding the role of the exchange rate, we look for evidence of time variation in the effects of exchange rate shocks over time. Figure 3 simulates the effect of an exchange rate risk premium shock in the estimated model. This disturbance moves the nominal effective exchange rate on impact by 2.5 per cent. Figure 3 shows that there is some time variation in the response of GDP and inflation to movements in the exchange rate and the largest time variation is pre and post 1998. There is also time variation in the policy response which is, of course, important in determining the overall outcomes. These results also reflect the fact that, as noted by , (2015), the standard deviation of the exchange rate shock has changed over time.

[12] This is a common finding in estimated structural models. For example, Christiano, Eichenbaum and Evans (2005) require a range of mechanisms to induce enough inertia for their model to match their empirical impulse response functions.

Figure 3: Time varying effects of a beneficial foreign risk premium shock in COMPASS



…But changes in the exchange rate channel do not appear to have a material impact on the timing of the peak effect, although there are differences in the overall magnitudes.

To explore the role of the exchange rate in the monetary transmission mechanism the (possibly time varying) effects of monetary policy are simulated assuming no change in the exchange rate. This experiment holds the exchange rate fixed for the whole period. The contractionary effects of an increase in Bank Rate are attenuated when the exchange rate is not allowed to appreciate, as would be expected. This can be seen by comparing the first row of Figure 4, where the exchange rate is not allowed to respond, with the second row which repeats the results from Figure 2. While there is a difference in magnitudes, time variation in the role of the exchange rate does not appear to have significantly affected the *timing* of the peak response of GDP or inflation.

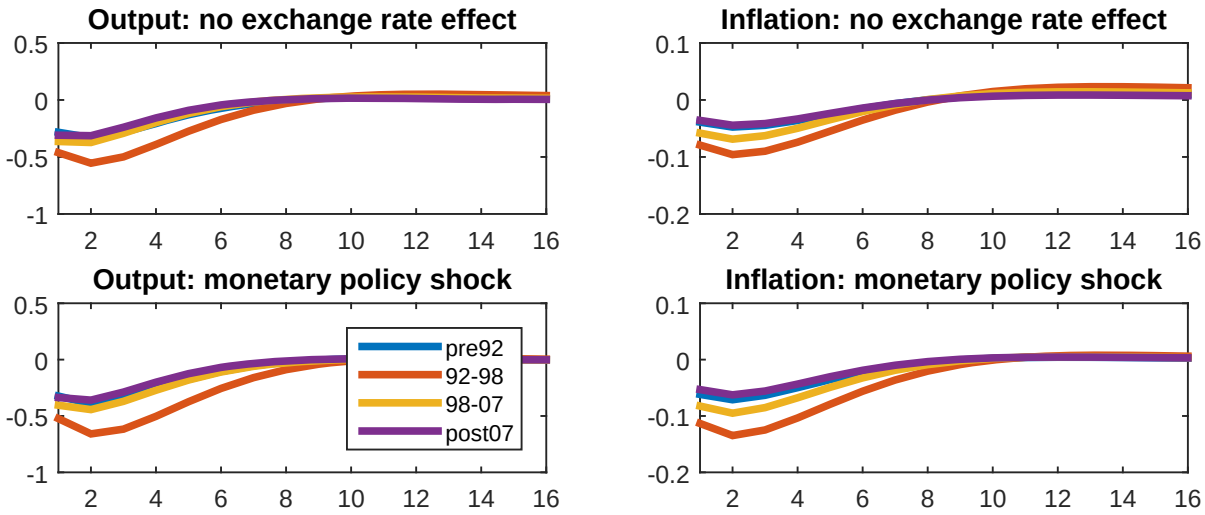
To quantify the role of the exchange rate for the *magnitude* of the responses over time, the average response of output and inflation over the period in the two scenarios is compared.[13] Allowing the exchange rate to respond produces a 22%-32% larger effect on GDP, on average, over the response period. The role of the exchange rate is smallest after 2007 (22%). For inflation, the effects would be 40%-50% larger and these numbers decline over time (from 50% pre 1992 to 40% since 2007). The findings are also broadly similar when considering the peak effect at 2 quarters. The role of the exchange rate therefore seems to be important but there is evidence that the role of the channel has diminished in recent years.

The previous experiment held the exchange rate fixed using unanticipated exchange rate risk premium shocks. But a limitation of this approach is that agents still expect movements in the exchange rate each period.[14] Another way of exploring the role of the exchange rate channel over time is to compare how the effect of a monetary policy shock varies when we only allow the parameters governing the pass-through of the exchange rate to import prices and CPI to vary over time. This exercise suggests that the impact on GDP varies by at most 4% and inflation up to 9% across the time periods considered in Figure 4. These are of the same order of magnitude as the findings from the previous exercise albeit a bit smaller.

[13] This is done over the period where the responses of inflation and GDP are negative (8 quarters). The computation of ratios is complicated as the responses get close to zero.

[14] This is, however, a problem that affects all conditional projections and forecasts, including from empirical VAR models. This is a further motivation for the second scenario we consider.

Figure 4: Effects of a change in monetary policy when the exchange rate is not allowed to respond (row 1) compared to the unrestricted response (row 2).



6. Concluding remarks and policy implications

When changes in monetary policy take time to have their desired effect this may provide a rationale for changing policy sooner to hit the inflation objectives by a given date. Understanding (i) how long it takes for the effects of policy to propagate and (ii) whether these lags vary over time are key questions for the conduct of monetary policy.

The evidence above suggests that, on average, monetary policy can take 1-2 years to have its peak effects. But non-US evidence suggests these lags could be slightly shorter for the United Kingdom. Two key results are:

1. The 18-24 months rule-of-thumb is reasonable on the basis of US evidence, but could be too long for the UK.
2. There is some time variation in the timing of the peak effect, although this does not seem to have been significant for the UK over the last two decades.

The timing of the peak impact is not, however, the only consideration when assessing the impact of policy. Both GDP and inflation tend to fall in the first year for the UK and, in many cases, these movements account for a sizable proportion of the overall impact at the peak. The profile of the responses, especially for real variables, also means that the precise timing of the peak is not very precisely pinned down. Consequently, an important conclusion is that even if the peak impact is felt further out, monetary policy can still have sizable effects even in the first year.

Appendix A: Full list of background notes

Note	Auhors	Description	Filesite link
When does monetary policy have its peak impact: is there evidence of time variation?		Presents new evidence on the transmission mechanism from a time varying parameter FAVAR for the UK.	
When does monetary policy have its peak impact? Evidence from other countries.		Summary of the international literature review. Accompanied by a slide-pack including all the impulse response functions.	**Note:** **Full evidence slide-pack:**
When does monetary policy have its peak impact? A review of key papers in the literature.		Summary of the seminal and high profile papers in the literature (largely US). Accompanied by a slide-pack including all the impulse response functions.	**Note:** **Full evidence slide-pack:**