The cyclicality of mark-ups and profit margins for the United Kingdom: some new evidence

Summary of Working Paper no. 351 Clare Macallan, Stephen Millard and Miles Parker

In the United Kingdom, the Bank of England's monetary policy objective is to deliver price stability, which is defined by the Government's 2% inflation target. The Monetary Policy Committee does not seek to keep inflation at precisely 2% every month, which would be unrealistic, but instead aims to set interest rates so that inflation returns to the target within a reasonable time period. That makes it clear that understanding the dynamics of inflation after a disturbance (or 'shock') is absolutely critical. In particular, we need to understand how prices and wages move in response to economic events. The key aspect of this is the way in which firms set prices as a mark-up over costs. The mark-up is of course an important determinant of firms' profit margins, but also determines how movements in costs are translated into the prices, and therefore inflation. So understanding the dynamics of the mark-up, the subject of this paper, is fundamentally important if we want to understand the dynamics of inflation, and therefore maintain price stability.

Consequently, we assess how mark-ups and profit margins move over the business cycle within the United Kingdom, at both the aggregate and industry level. Economic theory does not give a clear answer to this question. In particular, the relationship between mark-ups and output is likely to depend on what factors are causing them to move. And the response of the mark-up to any particular shock will depend on how sticky are wages and prices, as well as on whether or not the 'desired' mark-up responds to the shock. Among our conclusions is that mark-ups and margins appear to move procyclically, increasing in booms. If so, this means that the dynamics of inflation differ between demand and supply-driven causes of inflationary pressure. Supply-side cost pressures may feed into inflation relatively slowly.

Under certain assumptions, the mark-up will be inversely related to the labour share (equivalent to real unit labour costs, the cost of wages per unit of output, adjusted for changes in the price level). We find that the private sector labour share moves countercyclically, ie the labour share is low when output is high. This suggests that the aggregate mark-up moves procyclically, ie is high when output is high. This result is in line with evidence from the United States. But in contrast to studies using US data, we found that more sophisticated measures of the mark-up also move procyclically. Industry-level data support the finding that the mark-up is procyclical. We also find evidence suggesting that the average mark-up has fallen since around 1992. Since the average mark-up set by a particular producer will be inversely related to the degree of competition that he faces (more competition puts downward pressure on prices), this result suggests that competition has increased in most industries since 1992. We consider whether this increase in competition can be attributed to an increase in 'globalisation', or more precisely, the expansion of developing economies (in particular, China and India) and their increasing importance in the world trading system. But we do not find convincing evidence to suggest that globalisation explains the fall in average mark-ups.

We then consider profit margins as measured by the 'profit share': profits divided by the total value of output (turnover). We find that the aggregate market sector profit margin moves procyclically and that the cyclical behaviour of profit margins is largely the same across industries. In particular, we did not find strong evidence that margins are more procyclical in less competitive industries. However, both mark-ups and margins started to fall in 1997, in line with survey measures of capacity utilisation, whereas GDP growth did not peak until 2000. That this fall in margins occurred more in the export-intensive industries suggests that the explanation lies in firms trying to remain competitive in the face of the large appreciation in sterling from 1996 to 1998. Finally, we found that the market sector profit margin has trended upwards since 1970. This is in contrast to the downward movement in mark-ups observed over the same period. The explanation for this divergence seems to be the marked rise in the capital-output ratio since the 1970s: that change would help to increase measured profits, offsetting the decrease caused by the fall in the mark-up.

Understanding the dynamics of mark-ups and margins is crucial to understanding the dynamics of inflation. If, as our results suggest, mark-ups move procyclically then it must be either that wages are stickier than prices or that firms' *desired* mark-ups move procyclically. If desired mark-ups do move procyclically, this will have ramifications for how we might expect the economy to respond to shocks. In particular, an increase in demand will lead to a rise in mark-ups whereas a rise in costs would lead firms to reduce their desired mark-ups. And that reduction in desired mark-ups would lead to a delaying of the pass-through of the rise in costs into an increase in inflation.

An agent-based model of payment systems

Summary of Working Paper no. 352 Marco Galbiati and Kimmo Soramäki

A large share of all economic transactions is ultimately settled via money transfers between banks, taking place on 'large-value payment systems' (LVPSs). In 2006, the annual value of interbank payments made in the European system TARGET totalled €533 trillion (about \$670 trillion), amounting to more than 50 times the value of the corresponding countries' gross domestic products. The sheer size of these transactions, and their importance for the functioning of the economy, explains why policymakers are interested in LVPSs, and in the behaviour of their participants.

In the past, most payment systems worked on a deferred, net settlement basis. During a business day the banks would exchange promises of payments, deferring the actual transfer of funds to the end of the day, when only net positions were settled. The advantage of this arrangement was that only net debtors had to actually provide funds, and only in a quantity sufficient to cover their *net* position. Because net positions are typically small (compared to gross payments), the system as a whole would require little liquidity to function. Today instead, most LVPSs work on a gross settlement basis: there is no netting, and a payment obligation is legally discharged only when the corresponding full amount is transferred across accounts held at a central bank. This apparent backward step, strongly encouraged by monetary authorities worldwide, was motivated by credit risk concerns. Suppose indeed that, in a net system, at the end of the day a bank is unable to make good its final position. Its creditors may face losses too large to be sustained, so their payments too might have to be cancelled, creating a domino effect with significant consequences for financial stability. Gross settlement eliminates this risk but requires more liquidity, as the benefits (not only the risks) of netting are foregone. These arguments suggest that the provision of liquidity is an essential issue to modern payment systems.

Real-time gross systems are more 'liquidity hungry' than deferred net systems. However, they allow for liquidity 'recycling': when a bank receives a payment, it can use the received funds to make other payments of its own. To make an analogy, in a football game the ball can be passed between the players many times; similarly, a same unit of liquidity can be used to settle many payments. Consider however what happens if the ball is expensive to buy — maybe no one would like to pay for it in the first place. Unfortunately the analogy carries on to payment systems, where liquidity (the ball) bears a cost for commercial banks. This is an interest cost (typically charged by the central bank) or an opportunity cost (when liquidity is obtained against a pledge of collateral). So, even though just a little liquidity could generate a large volume of exchanges, it is unclear who should provide it. Banks are thus faced with a dilemma: to act as liquidity providers by acquiring costly funds, or to wait for liquidity to arrive from other banks. In the first case a bank does not depend on its partners, and it can promptly execute payments. In the second case, a bank benefits from a free source of liquidity, but is exposed to the risk of delaying payments while waiting for funds to arrive.

This paper develops a dynamic model of liquidity provision in a payment system, where banks face a choice between: a) the costs of borrowing liquidity, and b) the cost of delaying payments. In more detail, the model is a sequence of days. At the beginning of each day, every bank chooses how much liquidity to borrow from external sources. This liquidity is then used to execute payment orders which arrive throughout the day in a random, exogenous fashion (these orders can be interpreted as being commissioned by a bank's external clients, or by some area of the bank, different from the treasury). As long as the bank has sufficient funds, payments are executed as soon as they are received; when instead a bank's liquidity balance reaches zero, payments are queued until incoming payments provide the bank with new funds. Finally, at the end of the day banks receive profits, which depend on the liquidity borrowed, and on the delays suffered in executing payment orders. Day after day, banks adapt their liquidity choices following a particular learning process. As a consequence, the banks' behaviour eventually stabilises, and the banks end up providing an equilibrium amount of liquidity.

The system's equilibrium level depends on the model's parameters. By changing these, we look at the amount of liquidity absorbed by the system in a variety of scenarios, drawing conclusions on the efficiency of the system. We find that, for a wide range of costs, efficiency could be enhanced if banks were to commit more liquidity than they do in equilibrium. This might constitute a rationale for imposing measures that encourage liquidity provision (for example, throughput guidelines). From a different perspective, systems with fewer participants are found to be more liquidity-efficient than larger ones, due to the emergence of 'liquidity pooling' effects, as described by previous studies. These results are found by varying the size of the system but not its structure: it is outside the scope of this work to look at how liquidity choices are affected by changes in the extent of 'tiering' of a payment system (that is, we do not fully investigate the case of banks 'moving out' of the system, and making their payments through other system participants).

The conduct of global monetary policy and domestic stability

Summary of Working Paper no. 353 Andrew P Blake and Bojan Markovic

Inflation and output in the United Kingdom became much less volatile during the 1990s. There is an ongoing debate trying to establish whether this was caused by the absence of major disturbances in this period or by improved UK monetary policy following the introduction of the inflation-targeting framework. Answering this question is important not just for economic historians, but also for a wide range of practitioners and policymakers, who need to assess the prospects for the future volatility of UK inflation and output. Should the future world be affected by larger and less favourable disturbances, will the volatility of UK inflation and output revert to that seen before the 1990s? Or has more skilled policymaking, both home and globally, permanently reduced the capacity of shocks to increase the volatility of UK inflation and output?

When analysing the contribution of improved policymaking, most of the previous work on this question focused on an improvement in domestic monetary policy. This paper, by contrast, additionally examines the importance of an improvement in global monetary policy for the macroeconomic stability of a small open country like the United Kingdom. Specifically, an improvement in monetary policy means a shift from a policy where real interest rates do not increase in response to a rise in inflation to one where they do. To conduct our analysis we use a three-country model calibrated to mimic the United Kingdom, the United States and the 'Rest of the World'. The model allows for the relative price of exports to vary between countries, due to a 'home bias' for domestically produced goods. To examine the consequences of monetary policy that does not increase the real interest rate when inflation rises, we propose a simple new method for analysing 'indeterminate' solutions, a problem encountered when monetary policy does not produce one unique macroeconomic outcome.

Our key finding is that an improvement in global policy cannot sufficiently explain the extent of the fall in the volatility of output and inflation in the domestic economy. In particular, we find that a well-conducted monetary policy abroad cannot substitute for a well-conducted domestic monetary policy. If domestic monetary policy is well conducted (by which we mean that it increases the real interest rate when inflation rises), then an improvement in global monetary policy is unambiguously good for the domestic economy, because the shocks coming from the world economy are reduced.

By contrast, if domestic monetary policy is badly conducted then an improvement in global policy may not necessarily benefit the domestic economy. It can actually increase the volatility of inflation and output. There are two effects at work. A well-conducted monetary policy abroad stabilises output abroad, which stabilises domestic exports and thus output. But because any policy change, whether well conducted or bad, will affect the relative price of exports, there is a second, offsetting effect. Good global policy naturally takes account of and exploits changes in that price, which lead to more volatility in the domestic economy's exports. Good domestic policy would then offset this external source of volatility, but poor policy will not do so sufficiently. The precise effect depends on the specific parameterisation of the model, but in the hypothetical case that we examine had only domestic monetary policy been badly conducted, the second (destabilising) effect would have dominated, and the global policy improvement would have harmed the domestic economy.

We also find that an improvement in global monetary policy reduces the exposure of the domestic economy to disturbances originating abroad, indicating that an inadequate monetary policy abroad amplifies the impact of disturbances originating abroad. Furthermore, an improvement in global monetary policy increases the exposure of the domestic economy to foreign supply relative to foreign demand disturbances. This is because the second (destabilising) relative price effect is particularly important for supply disturbances, while demand disturbances mainly spill over across countries through the first (stabilising) global demand effect.