Gauging capacity pressures within businesses

By Colin Ellis and Kenny Turnbull of the Bank's Monetary Analysis Division.

This article discusses the measurement of capacity pressures within businesses — a key influence on the outlook for inflation. The degree of pressure on capacity relative to normal is likely to affect businesses' costs and prices. A variety of different methods are presented, each with their own advantages and drawbacks. Ultimately, no measure of capacity pressure is perfect, and the policymaker's judgement is crucial.

When setting interest rates, the Monetary Policy Committee (MPC) focuses on the outlook for CPI inflation in the medium term. One influence on CPI inflation in the short to medium term is the degree of capacity pressure in the economy. This article focuses on how this capacity pressure can be estimated, using several different techniques. Of course the degree of capacity pressure is not the only factor that will influence inflation in the short to medium term. While capacity pressure captures the balance of demand and supply in the product market, the degree of slack in the labour market will also influence inflation, as will movements in the terms of trade.

What are capacity pressures?

Put simply, the degree of capacity pressure faced by businesses — often referred to as capacity utilisation — reflects the intensity of production, or how 'hard' businesses are working their factors of production. Capacity pressure measures the balance between actual output and how much output companies would produce if their equipment (capital) and workers (labour) were employed at their normal intensities.(1) Capacity utilisation therefore depends not just on how much demand there is for market sector output, but also on companies' production techniques — how they use capital and labour to make that output.

Rises in capacity pressure are likely to be fairly short-lived: if a company faces persistently strong demand, it is likely to respond by hiring more workers or investing in new capital equipment, thereby increasing its 'normal' capacity. Similarly, if businesses experience persistently weak demand they may cut back on labour or investment. But in the short term, it may be difficult or costly for companies to change their capital stock or the number of employees. As a result, changes in demand are likely to be reflected by movements in capacity utilisation, at least in the short term. In addition, if businesses have a degree of pricing power they may take advantage of stronger demand to raise prices, thereby limiting the responsiveness of output to the stronger demand. But this may not be sustainable in the longer term, due to price competition from other businesses.

How do capacity pressures affect inflation? When utilisation is above normal workers are likely to be working harder, by putting in longer hours or more effort. Those workers will expect to be paid accordingly, for example through overtime pay, which is often higher than normal pay rates. At the same time, companies may respond to strong demand by using equipment more intensively — with the result that the costs of maintaining that equipment are likely to rise and/or the equipment will depreciate more quickly.

As a result of these factors, higher capacity utilisation is likely to go hand in hand with diminishing marginal productivity — that is, as utilisation rises it is likely to result in smaller and smaller increases in production.

So when capacity utilisation rises, companies tend to face higher costs. This implies that the marginal cost of production — that is, the cost of producing an extra unit of output — rises with the degree of capacity pressure. This rise in marginal cost will tend to put upward pressure on businesses' prices. As such, if businesses are working harder than usual — capacity pressure is above normal — there is likely to be upward pressure on prices and hence inflation.

How can capacity pressures be measured?

There are several different ways of gauging the degree of capacity pressure in the economy. This article examines three

(1) Note that capacity pressure does not directly incorporate slack in the labour market, or the difference between actual and potential employment. Barwell *et al* (2007) describe the determinants of potential employment.

broad types of measurement: models of production; statistical filtering; and business surveys.

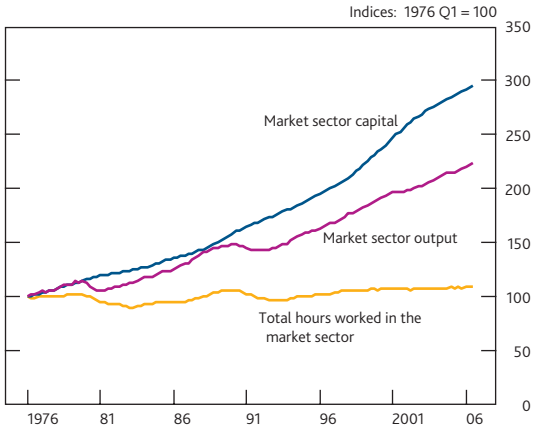
Using models of production to gauge capacity

Capacity utilisation reflects the balance between demand and supply in the economy — what companies are actually producing, and what they would 'normally' produce given the people and equipment that they employ. So one way to measure capacity pressures is to look at output data, and try to separate sustainable rises in output — often called 'trend' increases in output — from rises that reflect utilisation.

One way to do this is to use measures of companies' capital and labour — their inputs to production — to estimate how much output those factors would normally produce.

Over the past 30 years, capital has tended to rise at a faster pace than market sector output, while total hours worked in the market sector have risen at a slower rate, as increases in employment have been offset by falls in average hours worked (**Chart 1**). The capital series in **Chart 1** measures the productive flow of services from capital, rather than the asset value of the capital stock: the former is more relevant for calculating production capacity (see Oulton (2001)). Importantly, to gauge how much output these resources could normally produce, some assumptions need to be made about production. The basic tool economists use is the production function, which describes how capital and labour are combined to produce output.

Chart 1 Output, capital, and hours worked in the market sector



One key feature of production is companies' ability to swap between capital and labour in production — the elasticity of substitution in production. A common assumption is that this elasticity of substitution is equal to one (Cobb-Douglas).(1) However, previous work suggests that the elasticity of substitution is probably lower than unity in the United Kingdom: most estimates range from around 0.3 (see Harrison *et al* (2005)) to around 0.5 (see Barrell and Pain (1997)).(2)

Apart from the elasticity of substitution, it is also important to consider technological progress. Typically, when a production function is estimated on actual output data, it suggests that output has risen by more than can be accounted for by measured changes in capital and labour. Part of this could reflect measurement issues.(3) But, even when these are considered, some unexplained rise in output normally remains. This is called 'total factor productivity' or 'multi-factor productivity'.

These productivity measures tend to rise over time. So in order to gauge capacity pressures from them, the measures are separated into temporary and permanent components. The permanent component is often ascribed to technical progress — that is, increases in the efficiency with which companies combine capital and labour to make output. Technical progress is frequently assumed to be constant — and so efficiency is assumed to rise in a straight line.

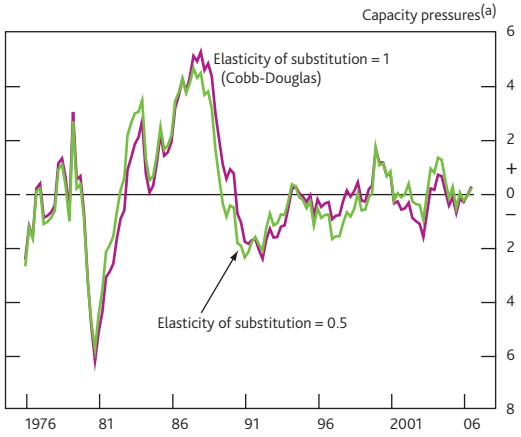
Equipped with all these factors — labour, capital, technical progress and the elasticity of substitution — a production function can be used to gauge the evolution of 'normal' supply capacity. The difference between actual output and this measure of normal capacity is then a measure of capacity utilisation.

Before doing this it is important to consider which concept of output to use. Often, measures of capacity are calculated using data on GDP. But GDP includes public sector output. And, quite apart from the measurement difficulties that arise when measuring public sector output, capacity pressures in the public sector are unlikely to have direct implications for CPI inflation, which applies mainly to the prices of market sector goods and services. So the relevant measure of capacity pressure within businesses should relate to the market sector, rather than the economy as a whole.(4) Of course, there may still be implications for inflation from public sector activity, which arise via the labour market (see Hills *et al* (2005)).

Chart 2 shows two measures of capacity utilisation, based on two different elasticities of substitution. The measures broadly track each other over time. This is because the implied time trends (proxying technical progress) will differ depending on the elasticity of substitution. Over a long period of time, this difference in technical progress broadly accounts for the difference in how capital and labour affect capacity.(5)

(1) This implies that a 1% fall in the price of capital relative to labour is matched by a 1% increase in the amount of capital used in production relative to labour (and *vice versa*). For more information on the elasticity of substitution in the United Kingdom see Ellis and Groth (2003).
(2) Chirinko *et al* (2004) estimate an elasticity of 0.4 for the United States.
(3) For example the measurement of labour skills: see Groth *et al* (2004).
(4) For more details on this measure see Churm *et al* (2006).
(5) Larsen *et al* (2002) describe a more complex model-based approach to gauging technical progress and utilisation rates.

Chart 2 Measures of utilisation based on production functions



(a) Percentage deviation of output from estimated 'normal' capacity.

These production-based measures of utilisation are subject to a great deal of uncertainty, not least about the measurement of capital, labour and technical progress. In particular, the assumption that technical knowledge advances at a constant rate is likely to be unrealistic. Indeed, research suggests that the rate of technical progress does indeed vary over time.(1) So considering alternative ways of measuring capacity pressures where technical progress can (potentially) vary is also important. The production function approach could be adapted to incorporate time-varying technical progress, although that time-varying behaviour would have to be determined somehow.

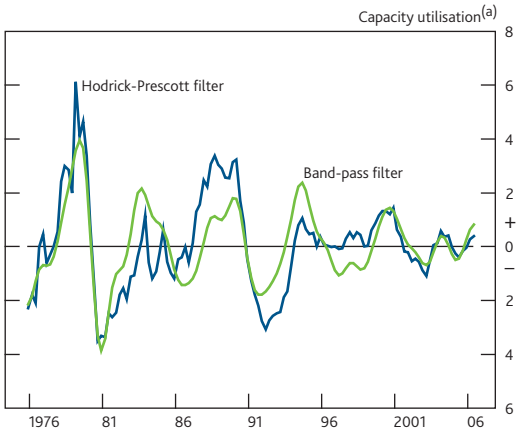
A statistical approach: filtering output data

The previous section described how capacity utilisation can be measured by focusing on the inputs companies use in production. But these inputs may be mismeasured: for example, capital services data depend on a variety of assumptions about depreciation. So examining alternative approaches, such as using statistical filters, may be useful.

Essentially, a statistical approach takes the output data described earlier and plugs it into a statistical filter — typically a smoothing mechanism. This 'smoothed' measure of output is then a proxy for normal capacity — once again, the difference between actual output and this smoothed measure is the measure of capacity utilisation. The logic for this approach is based on the earlier observation that short-term changes in demand are likely to be reflected in capacity utilisation and also that businesses typically find it difficult to adjust their factors of production in the short run. So 'normal' supply capacity should be less volatile than output, or alternatively sustainable (or 'trend') increases in output should be smoother than actual changes. However, those sustainable increases in output could vary over time, unlike the 'constant rate' technical progress measures described earlier.

There are many types of statistical filters, but two of the most commonly used are the Hodrick-Prescott (HP) and band-pass (BP) filters. The HP filter basically 'smoothes' data by calculating a moving average of the time series. The BP filter is a little more subtle, and is based on the notion that time series can be divided into components of different frequency: very slow-moving components, intuitively associated with the notion of a trend; fast-moving ones, associated with 'noise', or seasonal factors; and components that are between these two, which are often associated with the business-cycle fluctuations.(2) **Chart 3** shows measures of capacity utilisation based on these two filters.

Chart 3 Measures of utilisation based on statistical filters



(a) Percentage deviation of market sector output from filtered trend.

Of course, these measures are very uncertain, just like those based on production functions. These types of filters tend to produce a 'smoothed' series that is very close to the original series at the end of the time series, often known as the 'end-point' problem. For monetary policy makers, where the latest data are particularly important, that is a key concern. The implicit assumption that 'normal' supply is less volatile than demand will attribute any sudden change in output to utilisation, even if the change in output actually corresponds to a change in supply capacity.

Indeed, one strand of the literature has suggested that fluctuations in output growth may not reflect deviations in output from its potential — or capacity utilisation from normal.(3) Instead, changes in output growth could reflect an equilibrium response of the economy to real shocks, such as changes in potential supply, rather than companies working harder than usual. The production function and filtering approaches to measuring capacity pressure would then be misleading, as they would not allow for sudden changes in supply.

(1) See for example Kydland and Prescott (1982) and Ellis (2006).
(2) For more information on this filter, see Christiano and Fitzgerald (2003).
(3) See Kydland and Prescott (1982) and Hall (2005).

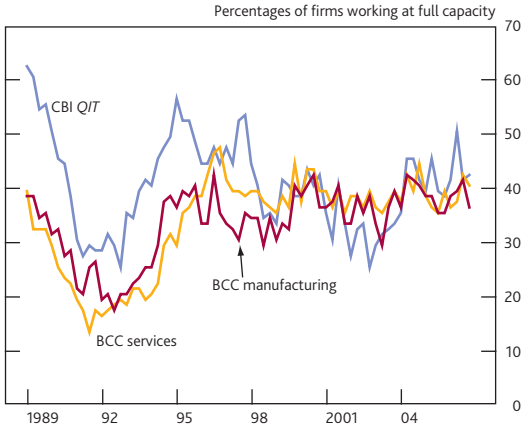
However, other work has shown that changes in output growth, at least in the short term, are more likely to be driven by demand shocks than supply shocks: see Blanchard and Quah (1989). This suggests that, over time, potential supply is less volatile — smoother — than demand, as embedded in both the production function and statistical filter measurement approaches. But it is still possible that supply capacity could change more suddenly than these measures assume, which could cause them to be misleading.

Measuring capacity pressures using business surveys

Capacity utilisation is clearly hard to measure using official data. One alternative is to look at other sources of information, in particular evidence from surveys.

Some surveys ask businesses directly whether they are operating at full capacity. Two of the best-known surveys that ask about capacity are the British Chamber of Commerce's (BCC's) *Quarterly Economic Survey*, and the CBI's *Quarterly Industrial Trends* (*QIT*) survey. **Chart 4** shows the capacity balances from these surveys.

Chart 4 Survey measures of capacity utilisation



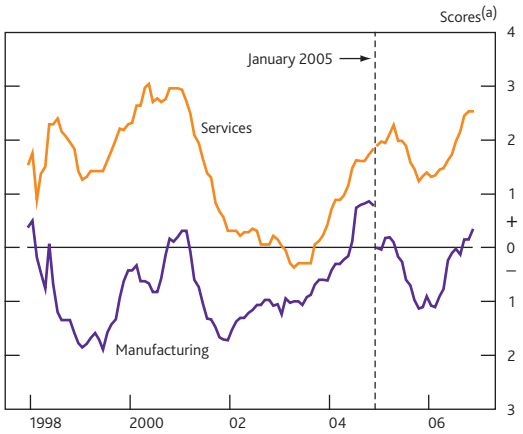
Sources: BCC and CBI.

There are also other data available on capacity pressure. As part of their regular reporting on economic conditions, the Bank's regional Agents produce a set of quantitative 'scores', which give a guide to conditions in the economy.(1) These include scores on capacity pressures in the manufacturing and service sectors (**Chart 5**).

Other surveys can also be informative. The CBI/Grant Thornton *Service Sector Survey*, and the CBI/PWC *Financial Services Survey*, do not include questions that directly ask about capacity pressure. But they do ask about how current demand compares to normal demand (**Chart 6**). If we assume that businesses set 'normal' capacity to meet 'normal' demand, then these questions can also provide information about capacity pressures in parts of the economy.

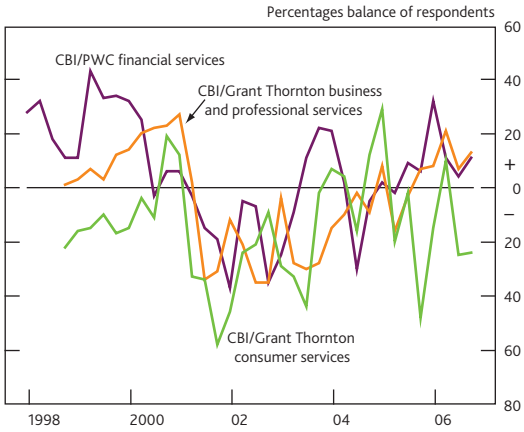
In addition, we can glean something from the CIPS/RBS *Report on Services*. This survey does not include a capacity utilisation question, or one on demand relative to normal. But it does include a question on changes in outstanding business, defined as work placed but not yet completed (**Chart 7**).

Chart 5 Agents' scores for capacity pressures(a)



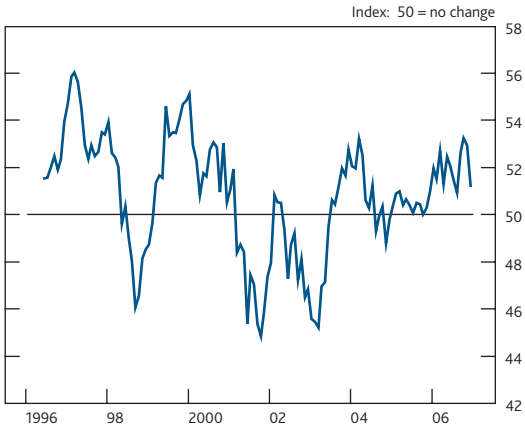
(a) A score of +5 indicates severe capacity pressure; a score of -1 indicates pressures are a little lower than normal. Prior to January 2005, the scores were based on current capacity utilisation relative to normal. Since January 2005, they relate to capacity constraints over the next six months.

Chart 6 Survey measures of demand relative to normal



Source: CBI.

Chart 7 CIPS/RBS outstanding business in the service sector



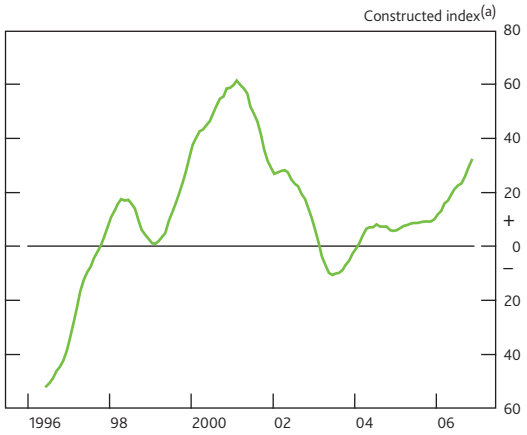
Source: CIPS.

(1) See Ellis and Pike (2005).

This outstanding business index tells us about the balance between demand and supply, and hence capacity utilisation. When outstanding business is rising, demand is running ahead of supply, and so capacity utilisation will be rising. But when outstanding business is falling, demand is running behind supply: so capacity utilisation will be falling. Adding up the consecutive changes in outstanding business(1) yields an indicator of capacity pressure. In order to derive the indicator, it is also necessary to set a reference date corresponding to a period where capacity pressure is close to normal. This can be gauged from other surveys.

The resulting CIPS/RBS-based index is shown in **Chart 8**; it indicates a peak in service sector capacity pressure around 2001, followed by a subsequent fall and more recent pickup. This is broadly consistent with other surveys (eg **Chart 5**). So the CIPS/RBS survey can also provide a gauge of capacity pressures in the service sector.

Chart 8 A capacity measure based on CIPS/RBS outstanding business



Sources: CIPS and Bank calculations.

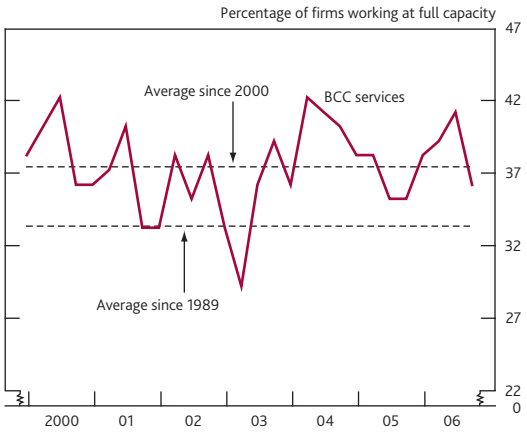
(a) Reference date where index is zero is March 2004, based on other services surveys (BCC, CBI/PWC, CBI/Grant Thornton and Bank's regional Agents' scores) being near their averages since 1999.

More generally, the issue of what level to 'benchmark' survey evidence against is important — it determines whether pressures are judged to be higher or lower than normal. This issue is related to the question of whether production function or statistical models incorrectly attribute changes in output to utilisation, rather than a sustainable change in 'trend' productivity. Crucially, in order to determine whether capacity pressures are higher than normal, we need some gauge of what 'normal' is. The production function approach determines this using the measured factor inputs and the assumption about constant technical progress. The statistical filters essentially impose that 'normal' capacity levels are smoother than actual output. But what benchmark of normal is appropriate for survey evidence?

One gauge of 'normal' would be to take an average of the series. But that average could be sensitive to the sample period. By choosing an average over a long period, its sensitivity to any particular data point or any single economic cycle is reduced. But it may be more appropriate to focus on the average measure over a shorter time period, for example if 'normal' capacity usage has changed. Evidence from contacts of the Bank's regional Agents suggests that companies have become used to working with smaller margins of spare capacity over the past ten years, partly because they are more able to 'flex' capacity. As such, comparing a capacity survey balance today to its average over the past 17 years may give a misleading picture of capacity pressures.

Chart 9 shows two averages for the BCC measure of capacity utilisation: one since the series began (1989), and the other since 2000. The two averages are very different — in particular, comparing the survey balance to the average taken over a longer time period suggests that capacity pressures have been above 'normal' for almost all of the past six years, associated with constant upward pressure on prices in the service sector. That is because the average it uses as a benchmark includes the marked fall in output in the early 1990s.

Chart 9 Capacity pressure in the service sector



Source: BCC.

Collating survey evidence

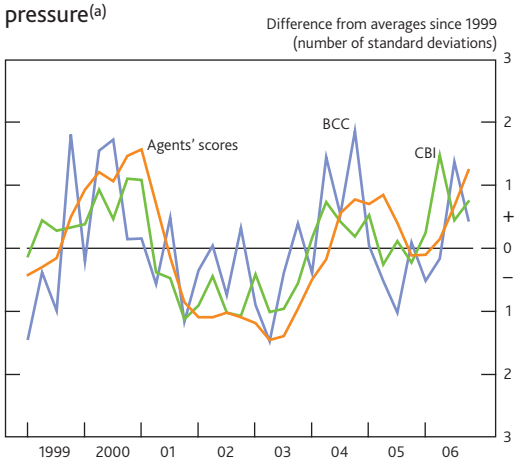
The previous section illustrated a number of capacity measures based on survey information. One option for summarising these data is to construct aggregate balances from each survey by weighting together the sectoral balances using output weights. The resulting balances are shown in **Chart 10**. All three measures suggest that capacity pressure peaked around 2001, and then fell back before rising in recent years.

An alternative approach to deriving economy-wide measures of capacity utilisation is to use statistical techniques to extract a common trend from the various sectoral surveys. Such techniques typically assume that each sectoral measure can be thought of as the sum of a common economy-wide

(1) In the CIPS/RBS survey, 50 is the benchmark 'no change' level. So this was subtracted from the CIPS series (the blue line minus the black line in **Chart 7**) before cumulating.

component, a sector-specific factor, and a measurement error. The assumption of a common trend across measures is given some support by the fact that sectoral surveys appear to move in broadly similar ways over time (see for example **Charts 4** and **5**). But there are different techniques for extracting such a trend, for example taking a linear transformation of data or using an estimated model. So it is important to cross-check these estimates against those produced using other techniques.(1)

Chart 10 Aggregate survey balance of capacity pressure(a)



Sources: Bank of England, BCC and CBI.

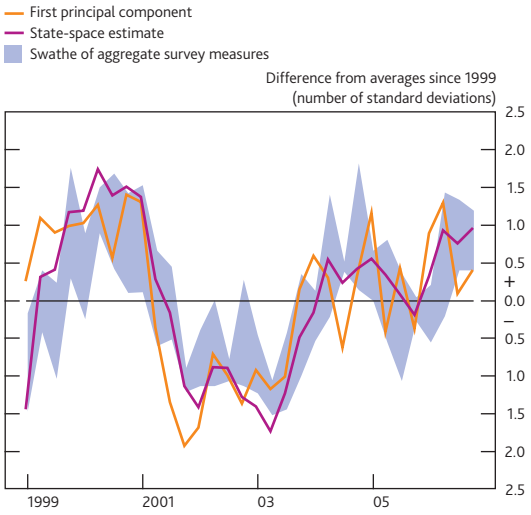
(a) The measures are produced by weighting together different sectoral surveys using nominal shares in output.

One technique for extracting common information from different series is called principal component analysis (PCA). PCA uses a linear transformation of the data series to identify the common component that underlie those different data series. Essentially, the technique 'chooses' a weighted combination of the different survey series so as to extract the most variable components from the data.(2) In this instance, there are nine different survey measures of capacity across all sectors: two from the BCC survey; four from various CBI surveys; one from the CIPS services survey; and two from the Bank's regional Agents. Using these nine measures, it is possible to generate the principal components using PCA.

Another means of gauging 'underlying' capacity pressure from the range of surveys is to use a state-space approach. State-space models allow unobserved variables — in this instance 'underlying capacity' — to be modelled using observed data, such as the survey measures. In practice such a model could take many forms — but for the illustrative purposes of this article a relatively simple model is used: each of the observed surveys is assumed to be the sum of the 'underlying capacity' measure and another unobserved component, which could vary across the different survey measures.(3)

Chart 11 shows the first principal component from the data series, and a state-space estimate of underlying capacity, together with a 'swathe' showing the range (minimum to maximum) of the aggregated whole-economy capacity measures in **Chart 10**. The principal component and state-space estimates match the swathe fairly closely. This suggests that the assumption underpinning the principal component and state-space measures — namely that each survey captures both economy-wide and sector-specific (or survey-specific) pressures — may not be unrealistic. The latest reads from all three measures in **Chart 11** suggest that capacity pressures may be a little higher than their recent average.

Chart 11 Different capacity measures based on survey data



Sources: Bank of England, BCC and CBI.

Conclusions

The degree of capacity pressure — essentially, how 'hard' companies are working — is a key influence on the outlook for inflation. Capacity pressure can be estimated using various approaches. This article has examined three different ways of measuring this capacity pressure, using: models of production; statistical filters; and survey evidence. Each approach has advantages and drawbacks, so it is important to look at a range of estimates. The estimates using these three different approaches have moved in a broadly similar way over time.

A common problem with this analysis is the risk of assigning a change in output to a change in utilisation, rather than a permanent change in supply capacity. But these methods can shed light on an unobservable — but key — element in the monetary transmission mechanism. As such, although no measure of capacity pressure is perfect, they are very important in helping policymakers form their own judgements.

(1) Astley and Yates (1999) describe a model-based approach to generating a capacity utilisation measure that combines information from output and survey data.
(2) For more information on PCA, see Jolliffe (1986).
(3) For more information about state-space models, see Greenslade *et al* (2003).

References

Astley, M and Yates, T (1999), 'Inflation and real disequilibria', *Bank of England Working Paper no. 103*.

Barrell, R and Pain, N (1997), 'Foreign direct investment, technological change, and economic growth within Europe', *Economic Journal*, Vol. 107, pages 1,770–86.

Barwell, R, Bell, V, Bunn, P and Gutiérrez-Domènech, M (2007), 'Potential employment', *Bank of England Quarterly Bulletin*, Vol. 47, No. 1, pages 60–69.

Blanchard, O and Quah, D (1989), 'The dynamic effects of aggregate demand and supply disturbances', *The American Economic Review*, Vol. 79, No. 4, pages 655–73.

Chirinko, R, Fazzari, S and Meyer, A (2004), 'That elusive elasticity: a long-panel approach to estimating the capital-labor substitution elasticity', *CESIFO Working Paper no. 1240*.

Christiano, L and Fitzgerald, T (2003), 'The band-pass filter', *International Economic Review*, Vol. 44, Issue 2, pages 435–65.

Churm, R, Srinivasan, S, Thomas, R, Mahajan, S, Maitland-Smith, F and Tily, G (2006), 'Measuring market sector output in the United Kingdom', *Bank of England Quarterly Bulletin*, Vol. 46, No. 4, pages 404–14.

Ellis, C (2006), 'Elasticities, markups and technical progress: evidence from a state-space approach', *Bank of England Working Paper no. 300*.

Ellis, C and Groth, C (2003), 'Long-run equilibrium ratios of business investment to output in the United Kingdom', *Bank of England Quarterly Bulletin*, Summer, pages 177–87.

Ellis, C and Pike, T (2005), 'Introducing the Agents' scores', *Bank of England Quarterly Bulletin*, Winter, pages 424–30.

Greenslade, J, Pierse, G and Saleheen, J (2003), 'A Kalman filter approach to estimating the UK NAIRU', *Bank of England Working Paper no. 179*.

Groth, C, Gutiérrez-Domènech, M and Srinivasan, S (2004), 'Measuring total factor productivity for the United Kingdom', *Bank of England Quarterly Bulletin*, Spring, pages 63–73.

Hall, R (2005), 'Separating the business cycle from other economic fluctuations', paper presented at the Jackson Hole symposium on August 25–27.

Harrison, R, Nikolov, K, Quinn, M, Ramsay, G, Scott, A and Thomas, R (2005), *The Bank of England Quarterly Model*, Bank of England.

Hills, B, Thomas, R and Yates, T (2005), 'The impact of government spending on demand pressure', *Bank of England Quarterly Bulletin*, Summer, pages 140–52.

Jolliffe, I (1986), *Principal component analysis*, Springer-Verlag, New York.

Kydland, F and Prescott, E (1982), 'Time to build and aggregate fluctuations', *Econometrica*, Vol. 50, No. 6, pages 1,345–70.

Larsen, J, Neiss, K and Shortall, F (2002), 'Factor utilisation and productivity estimates for the United Kingdom', *Bank of England Working Paper no. 162*.

Oulton, N (2001), 'Measuring capital services in the United Kingdom', *Bank of England Quarterly Bulletin*, Autumn, pages 295–309.