



Market and Public Organisation

Is the Private Finance Initiative a Good Deal?

Paul Grout argues that the rules that dictate when to opt for PFI/PPP instead of public provision are flawed. Despite views to the contrary he contends that it is the benefits of private sector provision not public sector that are most likely to be underestimated by the current process.

Issue 6
December
2001

In the last twenty years there has been a two step change in the idea of what the public sector should be expected to provide. The first step, the privatisation programme, although dramatic at the time, now seems relatively tame. Gone are the days when the state was expected to own and operate

networks, such as telecoms and electricity, run airlines or operate automobile businesses. Ownership was passed to the private sector and the role of the state rolled back to its natural function as regulator of residual dominant positions.

The second stage, the Private Public Partnership / Private Finance Initiative approach, is proving just as controversial. The state is now stepping back from buying and owning more traditional assets such as roads and hospitals. Here the state remains the core purchaser but is choosing to buy a flow of services from the private sector and leave the building and ownership of the assets to the private sector. This article is concerned with the PPP / PFI choice.

The traditional, and thorough, way to assess whether private provision is better than public is to look at the costs and benefits of each of the provision alternatives and undertake an appraisal. These days, however, few cost benefit studies of this type are undertaken. They are thought to be too complex and imprecise. Instead a value for money test is undertaken. It is this simplified approach that is the source of the problem. This article argues that the rules that dictate when to opt for private instead of public provision are flawed and that it is the benefits of private sector provision not public sector that are most likely to be underestimated by the current process.

Contents

Page 1
Is the Private Finance Initiative a Good Deal?
Paul Grout

Page 4
Raising Standards in the NHS:
The Use of Team-based Rewards
Marisa Ratto

Page 7
The Lure of False Profits:
The Economics of Optimism
David de Meza and John Maloney

Page 10
European Competition Law:
A New Enforcement System
Thibaud Vergé

Background

A large part of the controversy surrounding PFI projects stems from distrust of the government's case that private provision of infrastructure and public services is really better. The rules dictating when to opt for private instead of public provision have aroused suspicion and there is a view amongst the sceptics that the government is trying to 'fiddle' the books to make private provision appear beneficial.

Some of this is based on old chestnuts such as the notion that, since the government can borrow more cheaply than the private sector, it must be better to have government provision. In reality, regardless of the project, the government can borrow more cheaply than the private sector simply because they are more likely to repay than anyone else is. In the last resort, no matter how unsuccessful a project, the government is able to increase taxation and hence ensure that it can repay debt. This is why the public sector can borrow more cheaply, not because public ownership makes a project less risky and cheaper than it would be in the private sector.

Indeed, the difference between the required return on private sector debt and public sector debt is greatest where risk is greatest. Hence the false view that the government can provide cheap money would predict that the public sector should really be setting its sights on dot.com and swaptions markets rather than investing in the safer options of hospitals and roads. A rather implausible conclusion.

However, we cannot be dismissive of concerns about the rules that determine the public/private choice. People are right to think that something is 'funny' about the way that the public/private decision is judged. In fact, the current rules are flawed and distort the decision. However, it is almost certainly the case that the failure of the current rules works in exactly the opposite direction than is commonly believed. It is the private not the public sector that is harmed by the current process.

The value for money approach

A value for money test asks the following question. Does the Treasury pay less for private provision than public provision? If the answer is yes then the private provision route is taken. In the case of a road, for example, the test compares two cash flows. On the one hand, it calculates the cost if the public sector

procures and owns the road. This is the 'public sector comparator'. The alternative assesses how much the Treasury would have to pay to a private consortium for every vehicle using the consortium's road for the next 40 years. This is the PFI alternative. In the public provision case the government is buying a road whilst in the PFI alternative the government is buying services. In the latter case the private consortium owns the road.

The fact that one is comparing a public sector comparator that measures the cost of buying assets, on the one hand, with a PFI contract that measures the cost of buying services, on the other, is like comparing apples with oranges. Of course, this is not an insurmountable problem if one recognises that one is comparing apples with oranges. The problem with the Government's existing value for money tests is that they fail to distinguish between the apples and oranges, instead they treat everything as a fruit cocktail.

It is helpful to think of the problem in terms of a private sector analogy, say the building and letting of an office block by a private company. If the project just breaks even, i.e., rents just cover costs, then the company could measure the scale of the project either by referring to the cost of buying the block or to the value of all the expected rents. That is, one can calculate the cost of hiring a contractor to build the block or alternatively one can calculate the present value of all the future rents that may be received over the life of the building.

But the latter is a very different animal from the former. The 'build cost' is a front-end cash flow; the majority spread over two years or so and may be a relatively fixed figure, depending on the procurement contract and exposure to cost overruns. In contrast, the present value of rents is spread over forty years or so and will be massively more risky. The demand for space will depend on the state of the economy and so rents will be lower when the economy faces hard times. As common sense suggests, and any finance textbook shows, future rents should be discounted at a higher rate because of this uncertainty. However, there are even bigger differences to worry about. For example, if in a particular year the offices have something wrong with them then there will be no rents arriving at all in that year. This again suggests that the future rents should attract a higher

comparing a public sector comparator with a PFI contract is like comparing apples with oranges

we cannot be dismissive of concerns about the rules that determine the public/private choice

discount rate. The salient point is that if the company does not take account of these factors in the discount rate then the project will look as if it is very profitable when it is not. That is, the present value of expected rents will be grossly over estimated.

The analogy with the value for money test is clear. The public sector comparator is analogous to the cost of hiring the contractor to build the office. The PFI alternative is equivalent to calculating the cost to the Treasury of paying all the future rents but failing to take account of the risk. As we have seen, the present value of future rent is overestimated and suggests to the Treasury that building the office itself is a falsely cheaper alternative. In the PFI context the public sector comparator appears better value than it really is.

Four critical concerns

A simple application of economic theory identifies four areas where the apples need sorting from the oranges – the failure risk, the inherent risk of services, the treatment of quality and the role of profitability.

First, the current rules take no account of the fact that the payment to the contractor may not be made, i.e., failure risk is not recognised. The inability to deal correctly with failure risk is a major problem for PFI projects. The scale of the problem stems from the ferocious nature of the PFI incentive schemes. Most payment on delivery schemes look a bit like the ‘movie’ problem. If the film takes months longer to shoot than intended then the future returns from the film get pushed further away and have lower value as a result. This creates a strong incentive to finish on time. But most PFI projects have far more powerful incentive schemes built into them. If a project is supposed to deliver services for 25 years and comes on stream five years late then the contract does not push the whole process back five years in time but only pays out for 20 years. Indeed, the consortium loses more than 20% of the value since the years that get no payment are the initial years, those with the highest present value. This is a powerful incentive scheme, good in one respect since it imposes powerful incentives to deliver the services on time, but under current rules creates a massive underestimate of the value of private delivery. In short the existing methodology does not take proper account that the costs are incurred by the public sector under conventional

procurement whether or not the desired benefits are received, whereas under the PFI the public sector only pays to the extent that it receives these outputs. That is, the expected cost of most PFI projects is much lower than appears in the value for money test.

Second, the inability to recognise the inherent risk of services leads to overestimates of the cost of private provision. For example, the volume of vehicles on a road will be prey to the state of the economy in the same way as the volume of office take up. As with the office block, this risk needs to be recognised. Of course, the risk will be project specific and what is appropriate in the road context may not be appropriate for MOD contracts. What is certain is that the omission of failure risk and the inherent service risk in the discount rate overestimate the cost of private sector delivery in comparison to the public sector comparator. Because the PFI projects are so long lived small differences in discount rates may have a huge impact on the attractiveness of PFI projects and so we may be far away from recognising the true benefits of private provision.

Third, the value for money tests fail to recognise the quality of benefits delivered. A common criticism of the PPP/PFI programme is that they have failed to come up with really innovative approaches. This is not altogether surprising if these do not receive their true value in the assessment. It is far easier to replicate the public sector project at a lower cost.

Finally, there is a real issue of how to treat profitability. With the current system the ability of the private sector to deliver more cheaply than the public sector is only seen to have value when it reduces the Treasury cost. Whether this makes any sense or not goes to the heart of the reasons for adopting a mixed private and public delivery system in the first place. If the sole driver is cost reduction then the current approach to profitability is correct. But if the PPP/PFI model should be thought of as part of the overall productivity drive, and to my mind it makes sense to see it this way, then better productivity is the true test. How the spoils get split is important but not the sole issue. If productivity is an objective we ought to recognise this more formally in the assessment process.

the inability to recognise the inherent risk of services overestimates the cost of private provision

because PFI projects are long lived the consequences of small differences may be huge

Where do we go from here?

The empirical significance of the four points raised in the previous section is currently unknown but two points appear to be fairly clear.

- One, the net impact of the failure of the current system to recognise these effects seems to work almost universally against the private sector, although the true answer can only be known when the correct assessment is undertaken. I agree with the view that something is wrong with the current approach but the idea that there is a fiddle to help the private sector is wrong. The bias looks to be strongly in the opposite direction.
- Second, the failure to recognise these effects is unlikely to be small beer. I have argued that the Treasury should use different discount rates for PFI projects and the public sector comparator. Because PFI projects are long lived the

consequences of small differences may be huge. For example, if the PFI discount rate is wrong by one per cent, then the costs of private provision may be overestimated by around 14% on a forty-year project.

So where should we go from here? A good first step would be to conduct a study of the sensitivity of existing projects to differences in discount rates between the public sector comparator and the PFI model. As far as I understand this has never been done. I anticipate that the effect is likely to be considerable. Of course, looking at existing projects somewhat misses a lot of the point. It is the very projects that failed to get going, which may well include many high tech projects, where the effects may be greatest. We need a reassessment of the assessment procedures. We cannot engage in a sensible PPP/PFI debate until we have a clear idea of the financial benefits.

Raising Standards in the NHS The Use of Team-Based Rewards

*As part of the drive to raise standards in the NHS the government is experimenting with team-based rewards. Here **Marisa Ratto** summarises the lessons from the economics literature and argues that the pilot schemes should differ according to team type and offer highly varying levels of rewards.*

The NHS Plan identifies teamwork as important in improving healthcare and suggests that it should be promoted by the use of extra team-based rewards. However, the Plan is somewhat vague about the form of the teams and the extra rewards. Only a few examples of teams are provided, some of them based on patient experience. It is also implicitly assumed that financial rewards are likely to promote better outcomes in the NHS. The definition of rewards is, however, deferred to the results of some team bonus schemes currently being piloted in a number of NHS trusts. One example is the Norfolk and Norwich Acute Hospital Trust, where all staff will qualify for team-performance bonuses. Another example is the Northumbria Healthcare Trust, where the idea of using team-based rewards to raise standards will be

tested for one particular disease, and stroke and heart disease teams will be offered extra bonuses. Here we look at the economics literature on teams for guidance as to whether financial team-based rewards will necessarily improve healthcare.¹ But first it is important to note a couple of things about teams and team pay.

First, teamwork is usually defined by the technology of production, in that the only way of producing the output is pulling together the (unobservable) contributions of different individuals. This is complementarity in production. A simple example is a sports

¹ A longer discussion of these issues is available in 'Team-Based Incentives in the NHS: An Economic Analysis' and is available at <http://www.bris.ac.uk/Depts/CMPO>.

team. But teamwork may also be promoted when individual tasks are independent, to exploit some positive features of joint production that are not available if individuals produce by themselves (an example being grouping sales reps who work separately into a team for the purposes of pay). Second, team compensation is based on group output: each team member will receive some share of this and the share may not necessarily be related to his/her own output. This feature distinguishes group rewards from individual performance related pay.

Some insights from the economics literature on team-based rewards

According to the economics literature, the issues of how to reward and how to define a team are related, in that the optimal incentive scheme depends on the type of team. In a seminal paper, Hölmstrom shows that when a team is defined by complementarities between the agents, if total output is fully shared among the team, members will tend to free-ride. This is because a lower effort by a team member decreases joint output but the private marginal cost of lower effort (the decrease in the agent's share in total output) is less than the social marginal cost (the decrease in total output borne by the whole team). Hölmstrom shows that the mechanism for inducing appropriate levels of effort is the compensation system. Team members should not receive the entire value of output if it falls below some target level. The threat of not distributing all the output serves as an incentive mechanism. This threat is credible if a third party, not involved in the production process, claims the residual.

Hölmstrom goes on to suggest that teams may be usefully defined and rewarded even in the absence of complementarities in production. This is in those cases where agents are subject to common uncertainty (for example, fluctuations in demand) that influences the measurement of their individual effort. In this case, it is optimal to link an individual's reward to the contributions of other agents exposed to the same uncertainty. This is done by comparing the outputs of agents, in a manner similar to yardstick competition. Rewards are then based on relative performance in the team.

Teams could also be *created* by designing a system of rewards that fosters co-operation. Itoh analyses cases when it may be optimal to design a system of rewards that links each

individual's reward to those of other workers. Hence group rewards may be optimal even when teams are not defined by the production function. Common uncertainties or the need to promote some particular features of teamwork, like co-operation, may require the rewarding of individuals not only for their own contribution but also on the basis of group performance.

The definition of teams in the NHS

The literature also shows that the compensation scheme varies according to the nature of a team and the definition of a team has to come prior to the specification of the optimal system of incentives. The definition of teams in the NHS is not straightforward. The approach suggested in the NHS Plan, to define a team around a patient, has significant limitations. Only in a few cases are there precise protocols and guidelines describing the procedures to be taken, so that a patient's care pathway is well defined and teams can be identified around the patient. In most cases patients with the same condition can follow different care pathways, due to the uncertainty of the diagnosis and treatment. In different hospitals patients with the same condition may follow different care paths. Therefore, defining a team around a patient would be rather ad-hoc.

Team production clearly exists in the NHS and the existence of several types of (and possibly overlapping) teams suggests that there are strong benefits to be had from team production. Different types of teams include teams across disciplines, teams across organisations, hierarchical teams within an organisation (where senior and junior members of staff work together). Some of these may be small, others large². Among these categories, our analysis suggests that small teams, hierarchical teams and teams across disciplines seem to be more suitable for incentives, in that the positive aspects of teamwork (such as co-operation and mutual monitoring) outweigh the negative aspects (such as free-riding and professional jealousies). Many of the teams that currently

² Following Hölmstrom, the distinction between small and large teams depends on the degree of observability of each team member's contribution. In small teams final output is significantly affected by each agent's contribution, so that a decrease in effort by one member will significantly decrease final output and will be easily detected.

the definition of a team has to come prior to the specification of the optimal system of incentives

operate in the NHS, such as wards and firms, have these characteristics.

Is a case for teamwork necessarily a case for financial rewards?

However, the fact that team production operates well in certain teams does not necessarily mean they are the most suitable to reward financially. First, there is the issue of the definition of output. Hölmstrom emphasises that the efficacy of financial rewards is conditional on the identification of a good measure of performance. This measure needs to separate out that which is due to the effort of the agent and that which is due to uncertainty in the production process.

Performance indicators have been set by the Department of Health to measure performance against targets. These are set on the basis of the levels of service already being achieved by the best organisations. It is proposed that these be linked to financial rewards for teams (in this case the whole organisation). However, there is a danger that if the rewards are too strongly linked this might lead to collusion between NHS organisations to lower the targets. These collusive activities may be difficult to detect because the monitoring and measuring of performance is delegated to Regional Offices and the Commission for Health Improvement (CHI), who do not have any direct control of the funds to be used for promoting better performance. This practice is in strong contrast with the classic economics literature on teams, which suggests that the body which gives out rewards should have no part in the production process, and hence should have no role in measuring or defining the output, but should be in control of the funds.

Second, teams may be motivated by non-financial factors. Members of a team may have their own internal incentive arrangements, which entirely depend on the internal structure of the team and cannot be reinforced simply by altering agents' compensation. In this case, implicit incentives (such as organising the production process in such a way that it is possible for team members to monitor each other, or motivating workers by giving them more job independence) may be more effective than explicit financial rewards. Further to this, explicit rewards may conflict with implicit ones.

Third, the response of individuals to being rewarded financially for performance needs to be considered. Financial rewards may reduce intrinsic motivation. Itoh has shown that increasing rewards for group performance where previously workers were only rewarded according to individual performance does not necessarily lead to greater co-operation and may reduce all levels of effort. In other words, introducing small financial benefits for team production where previously there were none does not necessarily lead to co-operation and the financial cost of inducing co-operation may be quite substantial. Financial rewards may also attract workers with less intrinsic motivation.

Fourth, teams overlap, in that the same individual can be a member of different teams. Moreover, people can change teams over time. Financial rewards may be less desirable in such instances because tasks performed in one team may be more easily measured than others, so that individuals may devote more effort to the tasks that are better rewarded, or the measures of performance in different teams may be inversely correlated. This issue is similar to the multi-tasking problem that arises in the context of individual financial rewards in a public sector setting (Dixit, see footnote 1). Individuals may also seek to change to teams that are better performers, so disrupting existing patterns of production.

In conclusion, the design of the optimal incentive scheme to induce teamwork in the NHS requires, as a first task, the clear identification of teams and of their points of strength. The current pilot study ought to be doing this. Secondly it is necessary to identify in what circumstances financial rewards will have a beneficial effect. This requires tying team rewards to particular types of teams and piloting very different levels of team rewards. The lesson from the literature is that team production may operate better in certain types of teams, but this does not imply that they are the best to reward financially. Financial rewards do not automatically improve performance in a team. Task assignment, the physical organisation of the production process and the creation of a good working environment (e.g. recreational facilities for staff) may better motivate team members and increase performance.

the efficacy of financial rewards is conditional on identifying a good measure of performance

implicit incentives may be more effective than explicit financial rewards

The Lure of False Profits

The Economics of Optimism

David de Meza and John Maloney discuss their recent research on bias in personal financial forecasting. Over-optimism is endemic amongst the self-employed and they argue that the government should adjust its policies to account for the unrealistic expectations of success among certain sectors of the population.

Why do fruit machines have nudge and hold buttons? Because the tendency of players to believe they can beat the odds shifts into overdrive when they perceive themselves to be in control. Our contention is that such unrealistic optimism is pervasive in economic life. It shapes the way economic institutions evolve and has profound implications for economic policy. When economists give such advice the starting point is invariably that individuals are the best judges of their own welfare. Break this link and all recommendations must be reconsidered.

This underlying view of human nature is not original. Back in 1776 Adam Smith noted how

‘The over-weening conceit which the greater part of men have of their own abilities ... Their absurd presumption in their good fortune ... [means] ... The chance of gain is by every man more or less overvalued and the chance of loss by most men undervalued and by scarce any man valued more than it is worth’

Then as now, actors and barristers exemplified these traits, perceiving a low probability of very high earnings, but not imagining it to be as low as it really was.

A substantial body of modern psychological research validates Smith’s view. In particular, it distinguishes between excess optimism about events you cannot affect and excess optimism about those you can. Empirical work finds the second kind involves even less realistic expectations than the first. Which suggests that such optimistic illusions might find a special home among those starting up businesses. First, almost by definition, entrepreneurs see themselves as in charge. Second, starting a new venture involves evaluating many novel eventualities, so there is every opportunity

to downplay the less favourable. Unrealistic optimism flourishes, intensifying the tendency for those starting businesses to be self-selected as having excessively positive expectations.

The unrealistic optimism of the sector is augmented as things start to go wrong (more than a third of new businesses fail within three years). People wonder if they ought to exit and get a job working for someone else. Some do just that. The ones left behind in self-employment will be the super optimists. They see that returns aren’t great at the moment but they hang in, Micawberish, buoyed by their belief that things will soon improve.¹

Of course there are some spectacular successes. They are dangerous role models. The mundane failures attract much less attention but are far more numerous. No doubt there are those for whom it makes sense *ex-ante* to start-up a business, but far fewer than the number of those actually doing so.

Some recent studies provide indirect support for this view. There is evidence that, on average, starting a small business lowers income relative to remaining in paid employment. The difference is so great as to be difficult to reconcile with risk loving, the presence of undeclared income, or the non-pecuniary benefit of being ones own boss. Other research finds that many independent inventors who receive external assessments that say that their discoveries are not commercially viable, nevertheless proceed to develop them. The average return is massively negative. Finally, an ingenious

¹ Further implications are developed in de Meza, David, and Clive Southey (1996). ‘The Borrower’s Curse: Optimism, Finance and Entrepreneurship’ *The Economic Journal*, 106, 375-86.

*unrealistic
optimism is
pervasive in
economic life*

experiment has found that people are more likely to enter a competition when the payoff depends on their skill relative to others than if it depends on chance. In fact, the skill-based format results in so many entrants that on average their returns become negative.

To investigate further we turned to the British Household Panel Survey. Every Christmas it asks a sample of several thousand people whether they expect to be financially better or worse off, or stay the same over the following year. It also asks them how they actually fared in the year just gone. Because the same people are questioned year after year (over an interval covering more or less a full business cycle) it is possible to compare their predictions with what actually happened. The over-optimists will repeatedly expect to be better off but find themselves worse off or, at best, have stayed put.

The overall picture is one of unrealistic optimism and, sure enough, the self-employed are even more over-optimistic than those in paid employment. Amongst the self-employed 4.6 times as many expected to be better off but ended up worse off than were pleasantly surprised to find themselves better off after forecasting a decline. For employees the ratio is 2.9. Were expectations rational, both ratios would be unity. You might think that risks faced by employees and the self-employed are rather different. As long they keep their job, employees might expect greater stability, making accurate prediction easier. This was not an issue in practice as the proportion of incorrect forecasts differed little between the two groups. Moreover, the issue is the bias in forecasting not its precision.

Lots of other interesting facts concerning economic optimism emerge. Smokers are more likely than non-smokers to overestimate their income, but the biggest single factor explaining optimism is gender. On average women expect better than they get, but compared with men, women are realistic. Evolutionary psychologists could have a field day with this finding.

University graduates are less prone to wishful thinking, but is this cause or effect? We cannot answer this at present, but

speculate that realists are more likely to see the need for education. Alternatively, it could be that graduates have learnt to process evidence better. A test would be to see whether the divergence between the optimism of first year students and their non-student contemporaries is yet greater for third years. Similar issues of interpretation arise from the finding that readers of the Financial Times are more realistic than those of any other newspaper.

Economists faced with evidence of systematic divergences from rational expectations will naturally wonder whether people learn from their mistakes. One test is whether unrealistic optimism declines with age. The old might conceivably have acquired wisdom over the years and anyway have few surprises left in store. We do find that the old are indeed less prone to optimistic error but there is also some evidence that the very young also seem to have relatively realistic prospects. The clear sight of youth deteriorates until the mid-thirties, the age at which people are most prone to overestimate their prospects. It's significant that studies of *happiness* - as opposed to optimism - identify the mid-thirties as a uniquely unhappy bunch. People spend their twenties getting sadder and sadder and their forties pulling themselves together. In between come the tragic thirties.

If people at this age are the most over-optimistic *and* the least happy, this singles out thirty-odd as the age where life's disappointments crowd in fastest. Very often it is the age when you realise you're not going to do everything with your life that you dreamed of; as people adjust to this reality and scale down their goals, they lose their over-optimism and return to relative happiness.

If unrealistic optimism is indeed the norm it may explain many puzzling economic phenomena, including the urge to merge even though subsequent results are typically disappointing, and the overreaction of stock markets to the appointment of new CEOs and to other news. The common use of stock options in executive remuneration packages is inconsistent with standard incentive theory, for with rational expectations it is normally optimal that pay rises with performance throughout the range. If

the self-employed are even more over-optimistic than those in paid employment

unrealistic optimism may explain the urge to merge, the common use of stock options and financial 'pecking orders'

managers overestimate the chance of success, stock options, which are valuable in precisely these states, become a cheap way to make a job attractive.

These ideas also have implications for corporate finance. Optimistic managers believe that failure is a lot less likely than participants in an efficient capital market do. As a result, managers view the risk premium attached to debt as excessive. This induces a preference to finance new investment by means of internal cash flow rather than by issuing risky debt. From the manager's point of view, selling equity is even worse than financing through risky debt. Debt provides the outside financiers with the maximum possible protection should the firm perform badly whilst maximising the payoff to the owners of the firm when it succeeds. Shares do indeed share returns in all states. If the financiers purchase equity, their returns are more sensitive to firm performance than were they to hold debt. So when there is a divergence of opinion as to the likelihood of success, managers will view their company's shares as even more underpriced than its debt. Thus a 'pecking order' emerges whereby internal finance is considered best, then debt, then equity. There is considerable evidence that this is how firms behave in practice, though to do so is at odds with the classic Modigliani and Miller theorem. The usual explanation is the presence of asymmetric information. Biased information seems at least as likely. Unlike the asymmetric information interpretation, the consequences of an optimism-induced pecking order may not be all bad. Optimistic managers sometimes want to undertake projects (including mergers) that they falsely believe will be successful. These undertakings may nevertheless be abandoned due to the mistaken belief that the required external finance is overpriced, to the benefit of all.

What does unrealistic optimism imply for policy? There are many areas where overestimating the chance of preferred outcomes may lead to regrettable outcomes. For example, people may under insure, take excessive stakes in the business for which they work, or fail to make contingency plans for downturns. Government should consider

policies that, whilst not too intrusive, counteract such tendencies.

Let's think of the great issues. Keynes was unequivocal,

'If animal spirits are dimmed and spontaneous optimism falters, leaving us to depend on nothing but a mathematical expectation, enterprise will fade and die.... It is our innate urge to activity which makes the wheels go round'

Peter Mandelson, when in charge of the DTI, shared this view believing

'Britain needs to foster a pro-entrepreneurial, risk-taking culture like that in the US'

No hypocrite he.

In contrast, Adam Smith felt measures should be taken to curb those afflicted by irrational exuberance. For example, interest rates should be capped because at high rates money

'...would be lent to prodigals and projectors, who alone would be willing to give this high interest. Sober people ... would not venture into the competition'.

In fact it is not necessarily true that under *laissez-faire* optimism leads to more businesses being formed than if there was a greater degree of realism. The potential number of over-optimists may lead to such a deterioration of quality that lenders may find lending to some groups unprofitable at any interest rate. However this does not mean it is appropriate to subsidise such 'redlined' groups for they will, indeed, go on to perform poorly on average.

Despite the dot.com debacle the Government is bullish about new businesses. It notes that more new jobs are created in the sector than elsewhere in the economy. It is less receptive to the fact that more jobs are destroyed there than anywhere else. Myriad measures have been introduced to promote start-ups. What animal spirits are being promoted? The perspective we take here is that the government is in effect clearing the path by which lemmings access the sea.

over-optimism may so lower quality that it is unprofitable to lend to some groups at any interest rate

European Competition Law

A New Enforcement System

*In 1996 the European Union began a wide-ranging review of competition law. The proposed changes will devolve much of the decision-making from the Commission to national authorities. In this article **Thibaud Vergé** assesses the changes and argues that the current debate is giving insufficient weight to the function of the competition authorities.*

The last five years have seen major changes in European competition policy towards agreements between firms (e.g. franchising, exclusive territories or dealing). In 1996, the EC initiated a vast reform of the policy towards vertical restraints¹. The main effect has been the adoption of the New Block Exemption Regulation in December 1999. This regulation defines the types of agreements exempted from analysis by a competition authority and those that are illegal per se (e.g. price restraints, price-fixing or market-sharing agreements) and provides detailed guidelines for case by case analysis with greater emphasis on economic analysis, when the market share of the firms involved exceeds 30%. Following adoption of this regulation, the EC initiated an important reform of the enforcement system by issuing the *White Paper on Modernisation of the Rules Implementing Articles 85 and 86 of the EC Treaty*.²

In this article, I review the current regulation and the proposed change and highlight the results of recent theoretical analysis. The economic theory provides some justification for this change, but also shows that the reform may not be radical enough.

The current enforcement system

Undertakings and agreements between firms are considered by European Authorities under articles 81 and 82 of the EC Treaty. These articles set out general rules applicable to restrictive practices (art.81) and abuses of

dominant position (art.82). Article 81(1) prohibits all agreements which 'may affect trade between Members and which have as their object or effect the prevention, restriction or distortion of competition.'. However, an agreement can be granted exemption under article 81(3) as long as it 'contributes to improve the production or distribution of goods or to promote technical or economic progress, while allowing consumers a fair share of the resulting benefit.'

In 1962, the Council adopted Regulation 17 defining the system of enforcement of articles 81 and 82. Regulation 17 created a centralised authorisation regime. Under this regulation, the EC had exclusive power to grant exemptions under article 81(3), while national authorities and courts were only allowed to apply article 81(1). To be granted exemption firms were required to notify their agreements to the EC. The agreements were null and void unless they received negative clearance (formally or informally) from the EC.

The choice of a centralised authorisation system was justified by the need to establish a common culture of competition in Europe at a time when competition policy was not widely developed within the EU. These procedures have been applied by the EC for almost 40 years without any major change. This article asks whether this system is still relevant today and how it could be improved.

Weaknesses

Although this system has proved useful to develop coherent law enforcement (as decisions were centralised at EC level) and to diffuse a competition culture within each of the Member States, it has been heavily criticised.

¹ Any contract between firms operating at different levels (e.g. a manufacturer and a retailer) which does not consist of a linear tariff (a constant per unit price).

²Articles 85 and 86 of the Treaty of Rome were renumbered when the Treaty of Amsterdam came into force and are now articles 81 and 82 respectively.

The main criticism is that the authorisation system is highly bureaucratic and creates long delays. Regulation 17 requires a response from the EC for each notified case and the EC has quickly been overwhelmed with pending files: the EC has only made around 20 formal decisions a year and 4 years after the regulation came into force about 37000 cases were pending. The EC has tried to reduce the backlog by implementing changes to limit the number of new notifications and speed up the analysis. It therefore defined more precisely its judging criteria (introducing the concept of 'appreciable effect' on competition, allowing minor cases to be removed; issuing general notices to clarify exemption conditions and instituting block exemption regulations) and developed the 'comfort letter' system (an informal decision which allows delays due to required publication in the Official Journal, and mandatory translation into the 11 official EU languages, to be avoided). Even though notifications are now stable (at around 200 new cases per year) and the number of decisions is increasing (50% more in 2000 than in 1996), delays are still very long (about 4 years for a formal decision, 18 months for a comfort letter). Moreover, these figures mask the quality of the decisions taken by the EC, as more than 90% of the cases are closed informally (through settlement or a comfort letter), thus not providing the firms with the expected legal security.

Other criticisms have focussed on the potential inefficiencies of the notification regime. Competition policy experts, lawyers and economists do not always believe that consumers' protection and the promotion of competition, two of the initial goals, have been achieved. Indeed less than 1% of notifications have led to an infringement decision and this could suggest that the current system is not sufficiently deterrent. An alternative interpretation could be that the agreements harmful to competition are not notified and slip through the net, since the EC has to use its resources on notifications and cannot investigate other cases.

The proposed reform

These concerns, combined with rapid economic changes and the expected enlargement of the European Union, pushed the EC to examine reforms to the system. This was initiated by the publication of the White Paper and eventually led to a proposal for a Council Regulation submitted to the Council

of Ministers in September 2000. The proposal intends to replace the current centralised administrative authorisation system with one in which not only the EC but also the national courts and competition authorities will be able to apply Article 81 in full. However, the fact that the national authorities will be able to apply it as a whole does not mean that they will be allowed to grant exemptions. The current notification system will be removed and replaced by a directly applicable exception rule. Under this new framework, agreements would no longer have to be notified to be validated and would be audited by the national competition authorities after implementation by firms.

Experience has shown that the centralised notification system does not serve to safeguard competition. With the enlargement of the EU, it is also expected that a single institution will be unable to ensure that Community rules are complied with. Now that national laws are aligned on Community laws and there is a widespread culture of competition within the EU, decentralisation would not be detrimental for consumers. The objective is not a re-nationalisation of competition law but rather the dissemination of the Community rules. To ensure that all economic agents are treated on an equal basis, the EC will retain the power to examine any case and to overrule national authorities.

Leaving cases to national authorities and moving from a notification regime to an exception rule will also free up resources. According to Jenny, one aim of the reform proposal is to 'redirect the investigatory means of the Commission towards a fight against trans-national horizontal cartels or abuses of dominant position.'

The economic analysis

Now that it is widely agreed that the change is needed, the main question left is the economic impact of such a reform.

Barros has analysed the expected effect of the shift from the notification system to an ex-post control regime on the type of agreements firms will implement. He shows that the proposed reform will result, in general, in firms implementing less restrictive agreements. Two competing forces exist. One the one hand, the shift from ex ante (the current authorisation system) to ex post control (the exception regime) reduces the probability of the agreement being checked

the authorisation system is highly bureaucratic and creates long delays

the proposed reform will result in firms implementing less restrictive agreements

(and possibly prohibited) by the EC. Firms may therefore be tempted to try more restrictive agreements. On the other hand, the proposed reform will lead to a decrease in legal certainty for the firms. The EU competition legislation may be interpreted in different ways by national competition authorities and courts. The increase in the risk of fines may lead to firms choosing a more conservative stance and implementing less restrictive agreements. Barros shows that the latter effect dominates, thereby highlighting a favourable effect of the proposed reform.

A second argument to explain this change relies on knowledge about market organisation and competition and on the control and penalty tools used in the different regimes. Even though regulation 17 does not state it explicitly, there has been a 'gentleman's agreement' between firms and EC that notification granted the firm immunity from fines. The only risk incurred by the firm was to see its agreement declared void. On the other hand, the proposed legal exception regime is mainly based on heavily fining firms if they are caught implementing restrictive practices. This fine can therefore be set to deter harmful agreements (price-fixing) without deterring beneficial undertakings such as research and development agreements. However, this assumes that the EC has sufficient knowledge and is able to distinguish between beneficial and harmful agreements.

If knowledge is poor or if the data collected during an investigation is not informative, it will be impossible to deter just the restrictive undertakings. If the policy is applied too harshly, it deters even beneficial agreements, whereas if it is too soft it has no impact. In those circumstances it is better to choose an ex ante control regime (the mandatory notification system) to get more information.

If knowledge is now good or if the investigation leads to very informative data, the new regime can deter harmful agreements, as the competition authority is better able to detect serious malpractice. In that case, it is better to move to an ex post control regime, as

the absence of fines in the notification system will never deter bad agreements.

The move from the current authorisation regime (regulation 17) to an exception regime (ex post control) can therefore be justified by the improvement in the EC's rulings. In 1962, the EC had poor understanding of market conditions and a notification system was thus optimal. Moreover, it was an easy way to learn, as the firms were required to provide information on market structure. As the number of cases increased, the EC acquired better information and the quality of its judgements improved. Combined with the improvement of the economic analysis this now justifies the proposed move to an ex post control system.

Conclusions

It is commonly agreed that the current authorisation regime is too bureaucratic and that a change is needed to reduce the burden of notification faced by the EC. The proposed move to an ex post control regime moves in that direction, aiming to save resources and focus only on the agreements with a real likelihood of damaging competition.

The economic literature offers some theoretical justification for this change. Barros shows that the change will lead to less restrictive agreements due to the increase in legal uncertainty, whereas Bergès-Sennou et al. justify the move through a better quality of ruling and the deterrent effect of fines.

However, the economic analysis only concentrates on a comparison of the two regimes. Rey has suggested that greater emphasis should be placed on determining the optimal policy and addressing implementation problems. Bergès-Sennou et al. have shown that increasing the scope of the competition authority (e.g. by allowing it to compensate firms when the result of the investigation is positive) would increase the efficiency of the ex post regime. However, as this would give new powers to the competition authorities, I believe that we should open the debate on their role and attributions.

The Leverhulme Centre for Market and Public Organisation was founded in 1998 with the objective of furthering our understanding of an appropriate design for the reform of a wide range of activities on the 'boundaries of the state'. The Centre aims to develop research, further debate and inform policy relating to the public sector and the recently privatised entities. Further information can be obtained from:

*Leverhulme Centre for Market and Public Organisation, Department of Economics, University of Bristol,
Mary Paley Building, 12 Priory Road, Bristol BS8 1TN
Tel: 0117 954 6943 Fax: 0117 954 6997 Email: cmpo-office@bristol.ac.uk Website: www.bris.ac.uk/cmpo/*