The Economics of Measurement [1]

Report for NMS Review

By

G M Peter Swann

Professor of Economics and Management of Innovation
Manchester Business School
University of Manchester

Final Draft, 9th June 1999

[1] This is a slightly extended (and slightly revised) version of the original draft circulated on 17th May 1999. It responds to helpful comments from some members of the NMS review committee, and from some members of the PA Consulting Group team. I would particularly like to thank: Mark Lanyon, Ray Lambert, Roy Crouch, John Hobday, Geraldine Alliston, Paul Temple, Ian Miles, Ian Bradley – but none of these are responsible for remaining errors. This report does not claim comprehensive coverage of the literature, nor does it capture all the insights from the earlier studies commissioned by DTI and NMS. It is also expected that some of the observations made here may be refined during the course of the NMS Review.

TABLE OF CONTENTS

1. Scope of the Report

I have been asked by DTI to prepare a report on the fundamental economics of measurement to help in the NMS review. The scope of the report is defined as follows (annex to contract, page 8):

> "the aim of the paper is to improve understanding of the economic role of the NMS, the fundamental rationale for public funding, and contribute to the methodologies for identifying and valuing the economic and social benefits that it generates."

Accordingly, this report is split into three main parts: a review of the economics of measurement (Part 2), a review of the fundamental rationale for public funding (Part 3), and a contribution to methodology (Part 4). Part 5 addresses a number of points and questions raised in response to the first draft of 17th May 1999.

Part 2 examines the economics of measurement from three perspectives. These are:

(1) a functional and sectoral approach which describes where measurement happens;
(2) an economic map which describes the economic role of measurement activities;
(3) a "common pool" or "virtuous circle" perspective.

We argue that it is important to distinguish between research in measurement, the development of tools and infrastructure, and the operational use of measurement, since the economic issues are different. We also note that when one speaks of economic benefits from measurement, it is important to distinguish between the benefits that accrue directly to the user paying for a measurement service for his own use, and the wider social benefits accruing from public investment in measurement programmes.

We note that economists have given relatively little attention to the economic benefits of scientific measurement, though there are some related areas of the economics literature that are helpful. The economic analysis of standards has some important lessons for the economics of measurement. In addition, economic analysis of R&D, "common pools", asymmetric information and diffusion is also useful.

Part 3 sets out the basic economic rationale for public funding of some measurement activities. We argue that the most important rationale stems from the likely market failure in private provision of measurement infrastructure. However, we also argue that there may be a "strategic" rationale for public funding in this area. All three of the generic causes of market failure (externalities, scale economies and asymmetric information) may be relevant here, but we argue that externalities are the most important issue requiring attention in the NMS review.

Part 4, drawing on a description of the problem and the techniques available for empirical analysis, moves towards an agenda for empirical research in the NMS review. We explain why the measurement of externalities is so difficult, and then propose a methodological approach which combines direct measurement, case studies, engineering economics and (if possible) some econometric research.

Part 5 comments on seven broad points raised at the NMS review meeting on 24th May 1999, and in subsequent correspondence. These relate to: the international dimension of measurement standards; whether it is better to be a leader or follower in measurement; the link from measurement (and standards) to innovation; NMS as a technology transfer institution; metrics for centres of excellence; whether the overall scope of the NMS should be broad (do a bit of everything) or narrow (do a few things very well); and what is the meaning of the term "customer" in this study.

Part 6 concludes.

2. The Economic Role of Measurement

Some discussion about the economics of measurement is confusing because the term is used to cover a variety of rather different activities. The first aim in this part of the report is to develop a simple taxonomy of measurement activities, but one that captures the main differences. Then we shall explore each of these in greater depth, focussing in particular on its economic role.

2.1 What is Measurement?

As an economist, I would not presume to tell the NMS what is involved in measurement. What I try to do here, however, is to describe measurement activities in a way that brings out their economic significance. In what follows we map measurement activities in three ways. One could be called a functional and sectoral map: what sorts of measurement activity are there, and where do they take place? The second maps the economic role of measurement activities by locating them within a simple model of a market economy. The third views measurement activities as part of a "virtuous circle" where producers draw on several "common pools" of measurement (and other) resources in order to add value.

A Functional and Sectoral Map of Measurement Activity

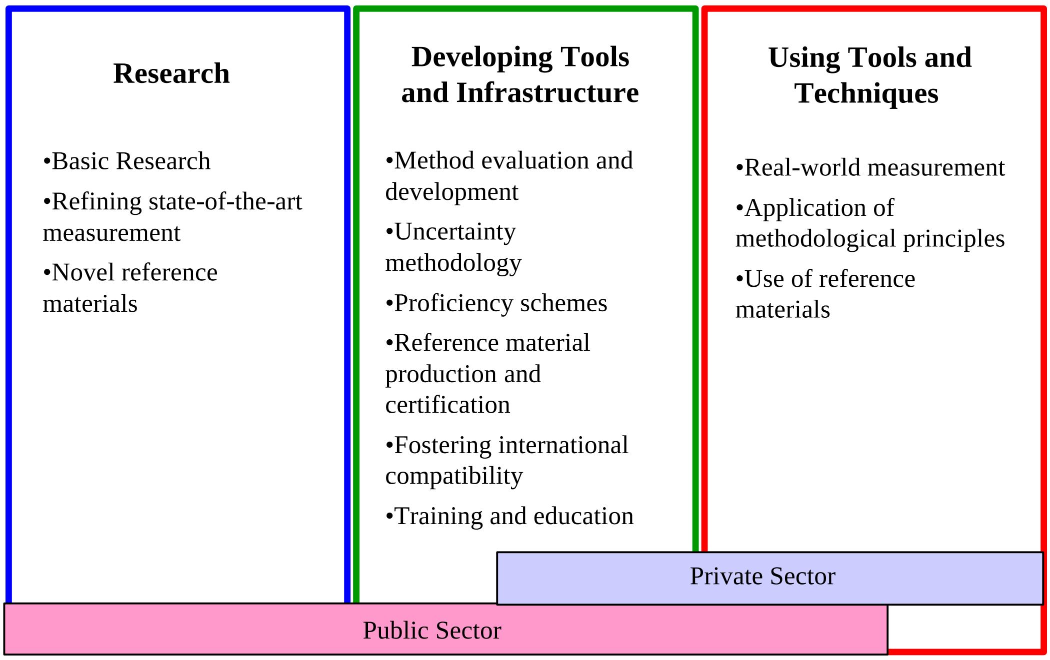
The functional and sectoral map is shown in Figure 1. It identifies three different types of measurement activity, and notes that some of these take place in the private sector and some in the public sector.

The three columns represent rather different activities. Figure 2 gives a more detailed indication of what each column comprises. There is research in measurement: basic research, refining state of the art measurement and discovery of novel reference materials. There is the development of measurement tools and infrastructure to carry out measurement, including: method evaluation and development, proficiency schemes and the production and certification of reference materials. And third, there is the day-to-day use of tools and techniques for real-world measurement.

Figure 1

Sector		Type of Measurement Activity		
		Research	**Developing Tools and Infrastructure**	**Operational Use of Tools and Techniques**
Sector	**Private (1)** Suppliers of Measurement Tools and Services			
	Private (2) Commercial Users of Measurement			
	Public (1) NMS			
	Public (2) Other			

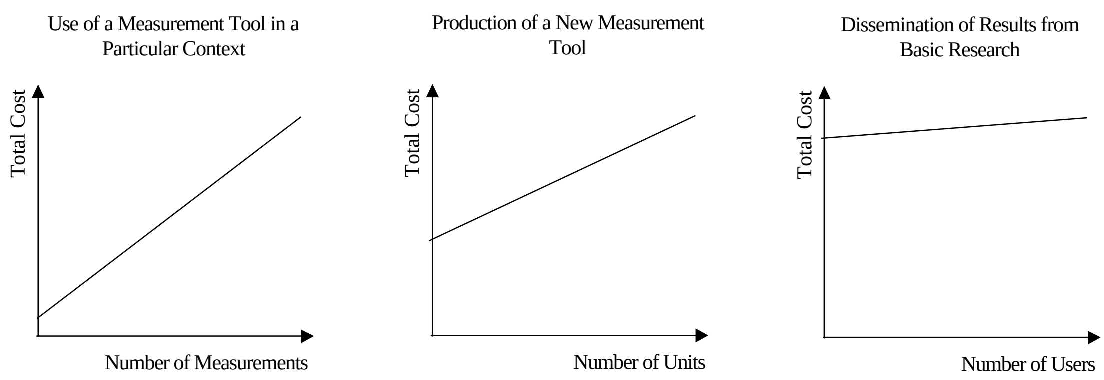
Figure 2



Research is pursued mainly in the public sector - in Universities and Government funded research labs. By contrast, the development of measurement tools and infrastructure may be located in both public and private sectors. And the day-to-day use of tools and techniques is again located in both public and private sectors - in large public labs, but also in labs within large companies, small specialist analytical labs, and even by the individual at home (home self-test medical kits, for example).

From an economic perspective the three columns represent rather different sorts of activity. They are differentiated in terms of their cost functions - see Figure 3. Research had large fixed costs, but its results can be disseminated (at least in some form) quite widely at relatively low marginal cost.[2] Developing new tools also has significant fixed costs, and economies of scale, but the marginal production cost per tool is not trivial. The use of measurement for day to day purposes has probably the lowest degree of economies of scale. These activities are also differentiated in terms of their generality or specificity, basic research having the greatest generality, and measurement activities relevant to a particular operation having the greatest specificity.

Figure 3



Moreover the rationale for investing in each of the three stages is very different. In making a business case to company management as to why they should invest more of their own funds in *using* measurement within their company, the case would ultimately have to address the *bottom line*. Does this investment deliver sufficient benefits in terms of increased sales, market share, or increased profit to justify the cost? Some companies would have a fairly good ‘gut feel’ for the answer to this, but the basic economics of private R&D could be used to analyse the benefits of such an investment. If there is a case to be made, it will rest on the argument that private investment will yield benefits to the firm in excess of the costs they incur.

Of course, some investment in measurement is undertaken not so much with an eye on profitability, but as a necessary pre-requisite to being in business at all. If a chemical

[2] The existence of these economies of scale raises the question of whether measurement activities might be funded on a European rather than a national basis (Evans et al, 1991). However it is not clear that economies of scale will be realised by such policies, as different institutions play "musical chairs" in the quest to avoid being cut.

company's analysis does not serve to convince the regulatory authorities world-wide that its products are safe, efficacious and quality-controlled, the company will not be able to trade.

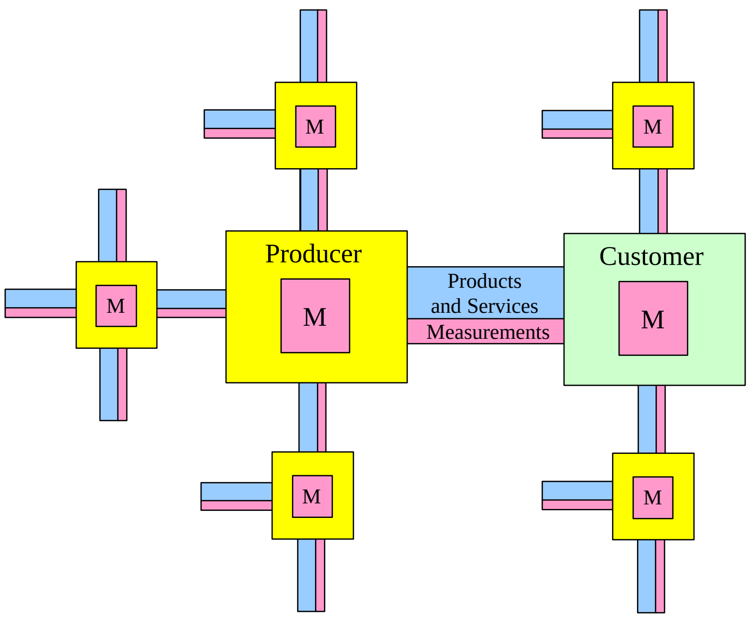
This first case, however, would be different from the case to be made to government for public investment in measurement activity. The latter case should not simply rest on demonstrating that this investment will yield a benefit to business or customers in excess of the cost to the public purse. For then the obvious retort is: "Why does this activity need to receive *public* funding? Why doesn't the private sector fund this?" Instead, the typical economic case for industrial policy rests on the market failure argument. We turn to that in Part 3 of the report.

The case for public investment in measurement will also vary according to which category of measurement activity is being discussed. The case for the support of research in measurement is likely to be similar to the traditional argument for supporting basic research in general. This will be discussed in Part 3 of the report. When we move nearer to the market, these arguments often lose force.

An Economic Map of Measurement Activity

The first map describes measurement activities and where they are located, but it does not give a clear picture of where measurement fits into the broader framework of a market economy. Figure 4 describes a simple economic map of measurement activity.

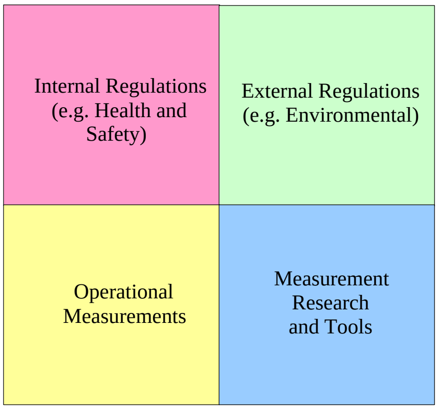
Figure 4



Measurement is located here in two places. First, it is located inside each organisation, where it is represented as a smaller box inside the producer or customer boxes. And second, it is part of the exchange between organisations, alongside traded goods and services. We shall argue that the economic conditions governing these two differently located measurement activities are rather different.

We can open up the measurement box within each organisation, and this is what we find.

Figure 5

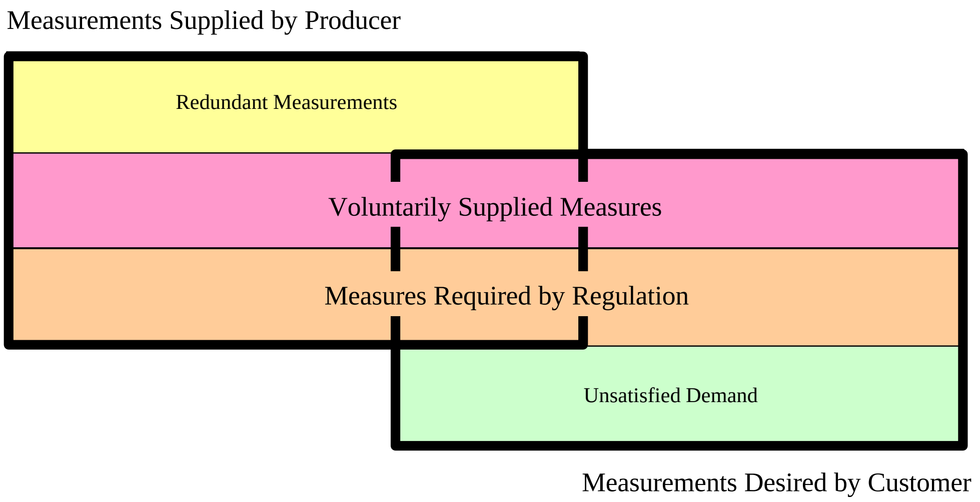


Here there are mandatory measurement activities, to conform with those regulations that govern operations within the company (e.g. health and safety regulations) and also with those regulations that govern the effects of operations on the outside world (e.g. environmental regulations). There are operational measurements as part of the production process, and there may be activities directed towards improving measurement tools, and perhaps even some research.

Clearly, all of these are important to the company's success. But rather like any other investment, we could argue that it is the company's own business to spend the appropriate amounts on the relevant measurement activities. If it does not, it may not meet health and safety regulations, environmental regulations, or it may simply put itself at a competitive disadvantage.

Now when we open up the measurement channel between organisations, Figure 6 shows what we find. It shows a Venn diagram, with a set of measurements supplied by the producer and a set of measurements desired by the customer. For reasons that will become clearer in Part 3, these may not overlap perfectly. There may be some measurements quoted by the supplier that are - from the customer's point of view - unimportant or even redundant. But more important, there may be some measurements that the customer would like but which the producer either cannot or will not supply. Why should than be? We explore this in more detail in Part 3, but the basic idea is that producers do not generally have much incentive to publish measurements that show their product or service in a poor light.

Figure 6



Some of the overlap will be measurements that the producer is obliged to publish in order to satisfy trading regulations. Some by contrast will be measures that the producer is not obliged to publish, but is happy to do so. But this model also illustrates how the degree of standardisation of these measurements may be very important. If the customer is processing measurement data from a variety of sources, or if he wishes to preserve an option to switch suppliers frequently, then standardisation of measurement is highly desirable.

This observation has led some (including the present author) to argue that the economics of measurement has much in common with the economics of standards. Perhaps it would be more precise to say that *some* aspects of the economics of measurement are *very similar* to the economics of standards - while others are *not*. Those measurement activities that are done to ensure that production processes or products and services meet regulatory requirements, and those that are done to enable trade between suppliers and customers certainly have a great deal in common with the economics of standards.

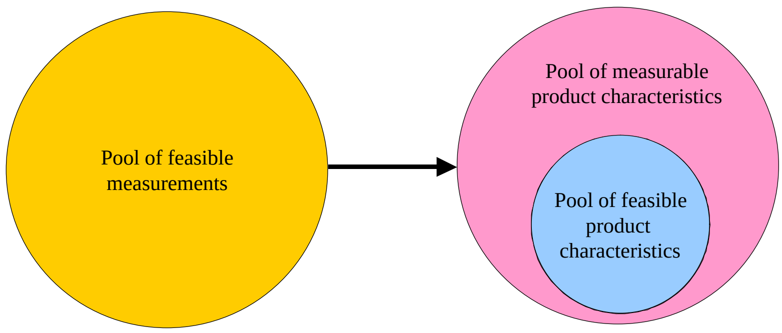
On the other hand, those measurement activities that are done for internal operational purposes have a rather different economic character. When we focus on the business case for operational measurement, this may be directed at increasing the efficiency of production, for example. Here, a useful analogy is with the *economics of investment* in cost reducing process innovations.

And finally, as noted already, those measurement activities directed at developing new tools and techniques and at measurement research have a quite different character again. Some of these are be akin to R&D activities (reverse engineering and synthesising new materials, for example) - and hence some aspects of the *economics of R&D* are relevant here.

A Common Pool or Virtuous Circle

A third perspective on the economic role of measurement will be useful in what follows. At its simplest, it is a "common pool" model. This is summarised in Figure 7. Cumulative experience in measurement and metrology and the work of those who develop measurement tools defines a "common pool of feasible measurements". The more that can be measured, the more characteristics that can be measured, and hence the size of this pool determines the size of another "pool of measurable product characteristics". And while not all measurable characteristics can necessarily be embodied in products (or services), it is reasonable to expect that as the pool of measurable characteristics increases, so also does the (smaller) pool of feasible characteristics. In short, anything that expands the pool of feasible measurements can be expected to expand the pool of feasible product characteristics.

Figure 7



In fact Figure 7 is one small part of a more complex chain of interactions. One chain is shown in Figure 8.

As drawn, each interaction in this diagram is bi-directional. The strength of these interactions may not be symmetric, but we want to emphasise that the chain model described here is far from the much-criticised linear model. So there is both demand pull and measurement push in this model. On the other hand, we have suggested above (Figure 6) that demand pull may work imperfectly here, because producers may not be willing to supply all the measurements required by customers. In that case, the feedback back along the chain from the customer may be dissipated quite quickly. For that reason, the interactions in this chain may work better when there is a virtuous circle - as for example in Figure 9.

Figure 8

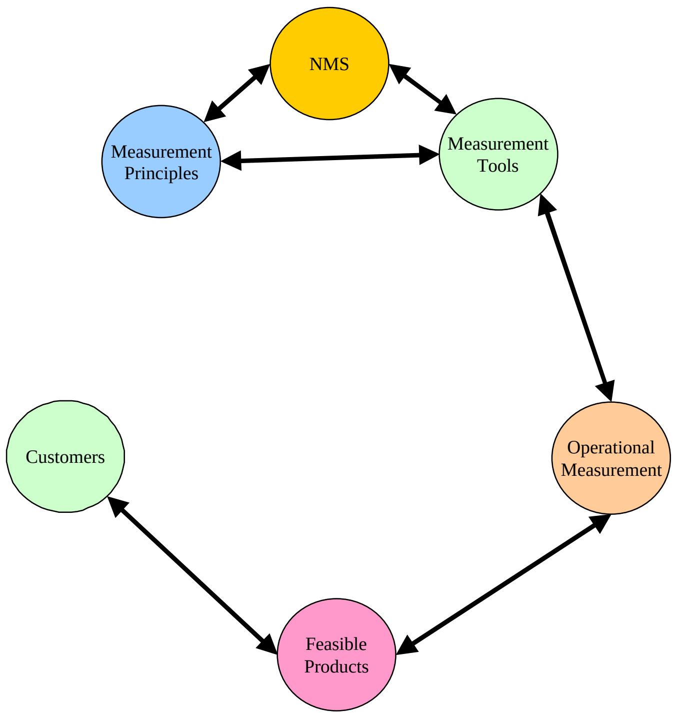
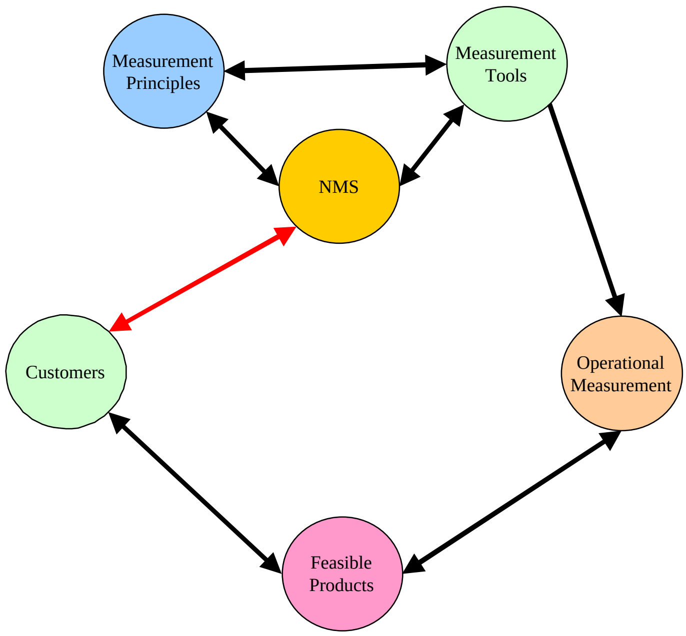
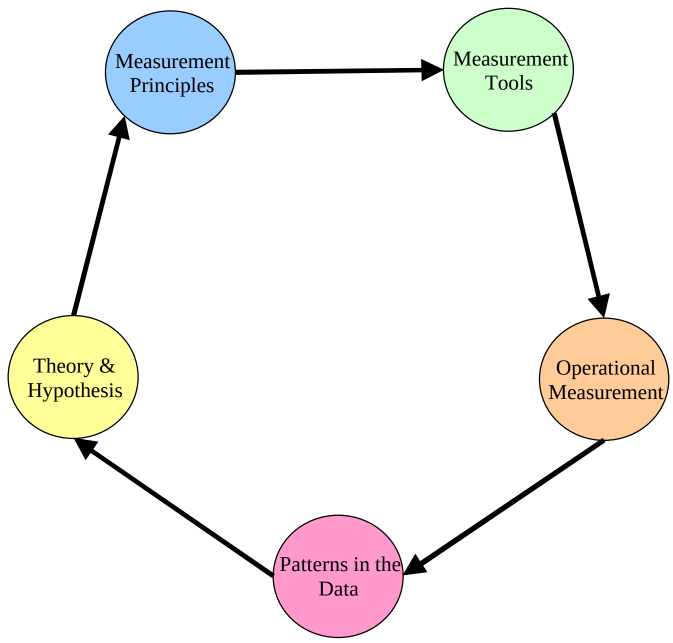


Figure 9



This may not be the obvious way to complete the circle - I'd be interested in DTI and NMS views on that. But the point is important because, as we shall see below, the effectiveness of basic research in generating market returns depends on the number and magnitude of the feedback loops in these virtuous circles. So long as the feedback is positive, then within reason, the more loops the better. This result is demonstrated more formally in Section 2.4.

Figure 10



Indeed, this is only one of the chains or virtuous circles in which measurement activities are located. Figure 10 shows a virtuous circle between measurement and research. A striking example of this is in one of the case studies in Swann et al (1997). It describes how a training and measurement tool was created to assess the performance of trainee surgeons at laparoscopic surgery.[3] But in addition to providing data on individual performance, one user (with responsibility for training) also found some interesting correlations in the data that could suggest a linkage between the personal and physical characteristics of the surgeon, and his or her aptitude for laparoscopic surgery. This observation completed the virtuous circle and launched another orbit, seeking new measurement principles and so on.

* * * *

So while there is very little economic literature on the economics of measurement *per se*, it will be useful to review a number of concepts from the economic literature on standards, R&D and investment, and virtuous circles.

[3] That is, "keyhole" or "minimally-invasive" surgery.

In this part of report we have limited our discussion to standards, common pools/virtuous circles and R&D. In Parts 3 and 4, we shall also focus on issues of externalities, asymmetric information and risk, economies of scale, and on diffusion theory. But it may be valuable over the remaining stages of the NMS review to add more detail to the picture of measurement activity presented here - what activities are included in measurement, who does them, and for what? This may indicate that other areas of the industrial economic literature may be relevant. If so, it would be appropriate to extend the models in Figures 8-10 above which link production activities and the market. For example, we can ask of each measurement function whether it: increases the speed of a particular activity; reduces the cost of a particular activity; improves the quality of a particular activity; reduces the risk from a particular activity; or if it facilitates communications or transactions between different links in the chain.

2.2 The Economics of Standards [4]

Let us start with a brief summary of the economics of standards. Then we can assess which parts of this are most relevant to the economics of measurement.

One important distinction to be made at the start is between formal public standards – such as those defined by ISO or BSI, and informal de-facto standards, which are proprietary dominant designs that emerge as de facto standards after a process of market competition. Which is better? There is no one answer. Some argue that the public standards-setting process takes too long, and that as a result the public standard – when it arrives – is too late to be useful. Accordingly, they argue, market-defined standards are becoming more important. Others argue that the market defined "standard" is not a standard, because it is not open and is not in the public domain. Rather, these are proprietary designs that dominate – or often monopolise – a market. And this monopolisation can clearly be a matter for concern.

In the economic analysis of standards, the standard is classified in terms of what it does. The literature recognises four broad types of effect that a standard can have.

- to define interfaces and compatibility
- to define minimum quality (e.g. safety standards)
- to achieve a reduction of variety
- to define standards of information and product description

Any one standard may achieve more than one effect, but this classification is important because these four different mechanisms have different implications for the economy. Table 1 summarises this.

Compatibility or interface standards help to expand market opportunities because they generate what economists call *network externalities*. These are, loosely speaking,

[4] This section draws on my earlier report for LGC (Swann, 1997), but also uses some additional material.

benefits that a consumer or user enjoys from being part of a large network of users. There are two broad categories of network externality: direct and indirect.

The telephone network is a good example of a network that generates *direct* externalities. The value of being a subscriber to a telephone network depends in an obvious and direct way on the number of other subscribers. If there are few other users, then the utility of the network is limited.[5]

By contrast a good example of *indirect* network externalities are those that arise for the owner of a common model of car. The owner may regret his lack of distinction in owning a car that is commonplace, but it is very convenient (and relatively inexpensive) when he comes to have his car serviced. The user doesn't care about the large size of the network per se (he may even wish it were smaller), but he is pleased to enjoy the good service network and competitive supply of spare parts.

Table 1
Effects of Standards

	Positive Effects	**Negative Effects**	**Necessity for Standard to be "Public" ?**
Compatibility / Interface	Network Externalities	Monopoly/Security	Monopolisation if not public
Minimum Quality/ Quality Discrimination	Correction for Gresham's Law; Reduced Transaction Costs	Regulatory Capture; "Raising Rival's Costs"	Essentially a Public / Co-operative Activity
Variety Reduction	Economies of Scale	Reduced Choice	Not Necessary
Information Standards	Facilitates Trade; Reduced Transaction Costs	Regulatory Capture	Essentially a Public / Co-operative Activity

[5] It is unlikely that utility is a linear function of network size. Indeed, a linear relationship only emerges under quite strict assumptions. But utility would normally be a monotonic increasing function of network size - unless consumers desire exclusivity in the networks to which they belong.

When network externalities are important to the buyer, the wise supplier will generally produce a product that conforms to the "industry standard". If the product is idiosyncratic and does not conform to the industry standard, the buyer will not generally find it attractive. The supplier may face more competition in conforming to the standard, but it is better to have a share of a large market than a monopoly of a tiny one. In some of the standards races that have taken place in network technologies (such as personal computers, audio recording media, video cassette recorder formats, and so on) the winner is not necessarily the "best" technology from the perspective of technological performance. The winner is the one that has been most effective and building a wide network of followers, and of support products from third party producers (e.g. software) that conforms to his standard.

Such "industry standards" may not be standards in the sense understood by standards professionals. They are not defined by committee, but rather are proprietary designs that win a position of market dominance - and hence earn the title of *de facto* (or informal) standards. But if a compatibility standard is a proprietary design rather than a public document, then the owner of the proprietary design can develop undue and undesirable monopoly power. For that reason, it can be preferable if standards are public rather than proprietary.

Minimum quality or quality discrimination standards can help to overcome Gresham's Law - the proposition that *bad drives out good*. If buyers cannot distinguish high quality from low quality before purchase, then it is hard for the high quality seller to sustain a price premium. In the absence of this premium, and if the high quality seller's costs exceed those of the low quality seller, then the former may not be able to survive. Bad sellers (who sell low quality produce) drive out good quality sellers by undercutting them. A system of public (or at least co-operative) standard setting and certification can help to overcome this. If this exists and is well understood then the buyer can confidently distinguish high quality from low quality before purchase, and then the high quality seller can sustain a price for his superior product. Standards are not the only way of overcoming Gresham's law - but they can be one of the most effective.[6]

Minimum quality or quality discrimination standards can - more generally - reduce what economists call transaction costs. If the standard defines the product in a way that reduces buyer uncertainty, then first the risk to the buyer is reduced, and second there is less need for the buyer to spend time and money evaluating the product before purchase. Consider a commodity market, for example: how could it exist in the absence of standards? Traders must be able to buy and sell large volumes without even viewing their trades. This is only possible if there is complete confidence about what it is that is being traded. That presumes a clearly defined standard grade, and certification that all produce traded meets that grade.

[6] It is recognised that some companies trade on their reputation, and can sustain a price premium for their products that are of quality well above this minimum threshold. It is also recognised that *ex post* restitution (a guarantee, for example) may be a satisfactory substitute for a certified minimum quality standard.

Do these minimum quality standards need to be "public"? They may not necessarily need to be defined by public agencies, but they must be co-operatively defined and certified to ensure that all remnants of Gresham's Law are overcome.

Variety reduction standards perform a different function. They seek to exploit economies of scale by minimising the wasteful proliferation of minimally differentiated models. So for example, high street stores stock suits in a limited range of standard sizes to exploit economies of scale. This may involve a certain compromise for some "non-standard" customers, and it is always possible to by a "perfect fit" in Saville Row, but at a price. The trade-off operating here is between choice and price. Variety reduction standards do not *need* to be defined publicly: the economies of scale can in principle be obtained with an idiosyncratic model range.[7]

Standards of information and product description are usually treated as a distinct category from the above, but for many purposes it is sufficient to treat these as a hybrid of the above three categories. Take the example of different grades of petrol: four-star, unleaded, and super-unleaded. These are standards of product description that also offer the other three features. Most motorists are confident that one type of four-star is *compatible* with another, and so can fill up at a Shell garage one week and BP the next.[8] Equally, these grades satisfy certain quality standards. And of course there are major economies of scale in distribution from the limited range of petrol grades.

Those near-market measurements that are carried out to confirm that a product is what it is supposed to be would appear to have much in common with this type of *product description* standard. The producer can confirm that the product to be sold is indeed what he expects it to be, and that reduces the risks (of compensation or litigation) to him, and also the risks to the buyer. In principle, the buyer can buy with confidence and without the need to carry out his own independent test that the product is what it is supposed to be. As such, this sort of certified measurement can help to reduce transaction costs, and hence make markets work better.

Imagine that a new technique is devised which makes it possible to measure a product attribute that has not previously be measurable. And take the case where this attribute is desirable - so that some at least will be willing to pay a higher price for products with a higher score in terms of this attribute. The arrival of this measurement technique benefits manufacturers who are especially good at producing a product that is rich in the attribute. It allows them to demonstrate the superior quality of their product, and charge a higher price for it.[9]

---

[7] But it is probably true to say that a high street store selling a range of suits in idiosyncratic sizes will not perform well, because its standard does not achieve *compatibility*.

[8] Sometimes manufacturers state that their product can use standard replacement parts but also imply that it will work best with "genuine" parts. This marketing strategy has been effective in the past, but perhaps less so now.

[9] This does not imply that all product characteristics relevant to consumers are precisely measurable. Some purchases are made with reference to rather subjective qualities that are not precisely quantifiable. Consequently, to sustain a price premium may not depend on precise measurement. We could say therefore that the arrival of this new measurement technique may be a sufficient condition to sustain a price premium but need not be a necessary condition to sustain all price premia.

In this case, there is a benefit to the immediate user of this measurement technique, and hence it might appear that discussion of externalities and spillovers is a diversion. In practice, however, we cannot be sure that a small group of immediate users will capture all the benefits of the measurement technique. Product discrimination standards can help to establish that a product is distinctly better, but like patents, they also convey information to the potential competitor. In a competitive market, new entrants may also benefit from this new measurement technique - and indeed the fact of their entry to the market and their competition reduces the benefit enjoyed by the original proximate users of the new measurement technique. Equally, some of the benefit may then be passed on to consumers - or if this product is an input to the production of many other products and services, the benefits are duly passed on to a much wider constituency.

There is an extensive literature on the economics of standards, to which the present author has made several contributions. Because the focus of the present report is on measurement – and not on standards – I shall attempt to survey it any further. But if members of the NMS review team would like to explore this further during the course of the review, I can point them to the relevant parts of the literature.

Apart from case studies of standards races, there is relatively little systematic empirical evidence on the macro-economic effects of standards. One recent econometric study (Swann et al, 1996) examined the effects of BSI standards-setting activity on British trade performance. This was an econometric study. If we try to control for some of the other influences on British exports and imports, what is the effect of standards-setting activity - if any? The study measured BSI standards setting activity in each industrial sector by comparing relevant BSI standards counts with relevant DIN (German) standards counts.

Table 2
Hypotheses about Macroeconomic Effects of Standards

Hypothesis	Effect on Exports	Effect on Imports	Empirical Support - if any - in Swann *et al* (1996)
Trade Creation	+	+	Receives the most support
Competitive Advantage	+	–	Some support
Trade Reduction	–	–	No support
"Competitive Disadvantage"	–	+	No support

The study considered four alternative perspectives on the effect of standards on trade. The most prominent in the economics literature is the so-called intra-industry trade (or trade creation) hypothesis, and the competitive advantage hypothesis. As Table 2 shows, the first expects a positive effect of standards activity on imports and exports,[10] while the second considers that standards activity enhances exports but reduces imports. These are the two hypotheses receiving the most support in the study.

The critics of standards-setting institutions, who consider them to be bureaucratic and unhelpful to trade, have advanced two other hypotheses. The first of these is that standards reduce trade because they increase barriers to imports and make the products of one country less marketable in export markets. The second ("competitive disadvantage") hypothesis asserts that standards impose costs on those that have to observe them, but do them no good. Accordingly, so this argument goes, standards lead to larger imports from countries unencumbered by standards and get in the way of exports. These hypotheses receive no support from the empirical results summarised in Swann et al (1996).

[10] Cunningham et al (1991) also argue that such an outcome can be expected for public investment in measurement.

This empirical evidence is in line with the general presumption in the economic theory literature that the emergence of standards - whether compatibility, minimum quality or product description standards - is helpful to the operation of markets as it reduces transaction costs.

The Relevance of the Standards Metaphor to the Economics of Measurement

Just how relevant is this discussion of standards to the economics of measurement? As we noted before, those measurement activities that aim to ensure that production processes or products and services meet regulatory requirements have a great deal in common with standards. Here the most relevant standards concept is probably the minimum quality standard, though the compatibility standard concept may also be relevant. Those measurement activities that are done to enable trade between suppliers and customers again have a great deal in common with the standard. Here the most relevant standards concept is probably the interface or compatibility standards, though again the minimum quality standard may be relevant.

Moreover when we turn to discuss the rationale for public investment in measurement in Part 3 of the report, it will have quite a lot in common with the rationale for public investment in standards, though there are some differences.

In short, we can conclude - as John Barber did some years ago (Barber, 1987) - that the economics of measurement and the economics in standards have much in common. Barber spoke of *measurement standards* to refer to those parts of measurement activity that fulfils a similar function to standards, and that remains useful now.

2.3 Common Pools and Virtuous Circles

We have grouped these two together because, while clearly distinct, they do have a number of features in common.

Common Pools

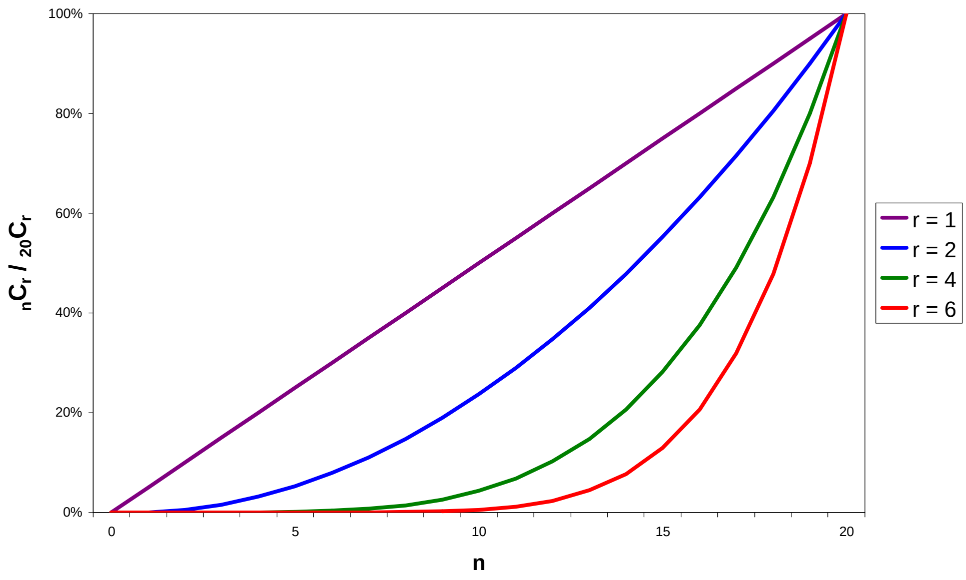
The basic idea of a common pool is that the producer adds value by combining together resources from a common pool. And the number of combinations that can be made accelerates rapidly as the size of the pool increases. Why is this? Because from elementary probability theory we know that the number of ways of drawing r items from a population of n - when the order of sampling does not matter - is given by:

$$nCr = n! \ / \ \{ r! \ (n-r)! \}$$

As Figure 11 shows, when r = 1, this is a linear function of n. But for larger r, this function accelerates sharply as n increases.

As I have argued elsewhere (Swann, 1990), we should not see product innovation as simply the improvement of particular product characteristics. As important, and often much more so, is the incorporation of new characteristics into products. For that reason there is a combinatorial character to product innovation. How many variants can be produced depends (very roughly speaking) on the number of characteristics that can be combined.

Figure 11



In my report for LGC, I presented a simple model using this approach to plot the economic benefits of investment in measurement. The model makes no pretension to generality. Moreover, it rests on strong assumptions that will not apply in the majority of measurement cases. But it is useful as an indication of the common pool approach at work. For full details, reference is made to Swann (1997), but the basic structure is as follows.

The model is based on an economic technique called *characteristics* analysis. This starts from the unremarkable observation that products are differentiated by the features (or characteristics) they embody. If each of these characteristics is treated as an axis of a multidimensional space, then it is possible to summarise the spectrum of competing products in a product *space*.[11] The use of a space analogy turns out to be a very powerful one, because the behaviour of firms and customers in product space shows some striking similarities to their behaviour in geographical space.[12] It is found

---

[11] This is an unusual space, since the units of each dimension may be quite different. Euclidean measures of distance are rarely of any relevance. Moreover, while some characteristics may be continuous variables, others are binary.

[12] For example, innovative firms often try to avoid highly congested product spaces by locating in a less congested part of the space, or more interestingly by creating a new area of the product space.

that product spaces are usually of quite a high dimension, and moreover the dimensionality sometimes increases over time.[13]

The extent to which firms and customers can envisage and understand the space within which products are located depends on measurement. It is hard (or even impossible) to establish that one product is superior to another on some characteristic, unless that product characteristic can be measured. And as we shall see in Part 3 of the report, Gresham's Law asserts that if the customer cannot distinguish between products that offer high quality on some characteristic and products that offer low quality, then the producer of high quality will not be able to sustain the price premium he needs to survive.

Accordingly, the advent of a new measurement method to measure a particular characteristic will serve to open up a new dimension of product space. When producers can demonstrate reliably that their products are indeed superior to the competition in a particular dimension, then if demand for quality is strong enough, they will be able to sustain the necessary price premium. This means that it is feasible to trade in this new part of product space, and that will be beneficial to producers and to customers who seek that sort of quality.[14]

Hence in this model, the steady growth of new measurement activities brings value because they open up product space, to the benefit of producers and customers.

The proportion of possible product locations that are supported by measurement is a function of the cumulative investment in measurement, and the relationship is similar to that presented in Figure 11. When $r = 1$, the relationship is linear - as indicated above. For larger r, the function is very skewed. With high values of r, indeed, it is only when the vast majority of characteristics can be measured that the real benefits of the measurement system are felt. Or to put it another way: if measurements are used to support complex new products with many characteristics, the real benefits from the measurement programme don't accrue until near the end.

However the full picture is more complex. It doesn't necessarily follow that the economic benefits are proportional to the share of product locations that are supported. The reason is simple: it is unlikely that each measurement step is of equal significance. As the programme proceeds, it opens up new areas of product space, but it is not clear that these are all of equal value.

Suppose instead that the measurement programme is organised so that the most important measurement activities are tackled first, and the less important ones are scheduled later. For simplicity, indeed, we assume that the measurement activities are ordered neatly in declining order of importance. In this case, Swann (1997) shows that

---

There are exceptions, of course, when firms seek to cluster together - as, for example, when the network externalities from so doing are strong.

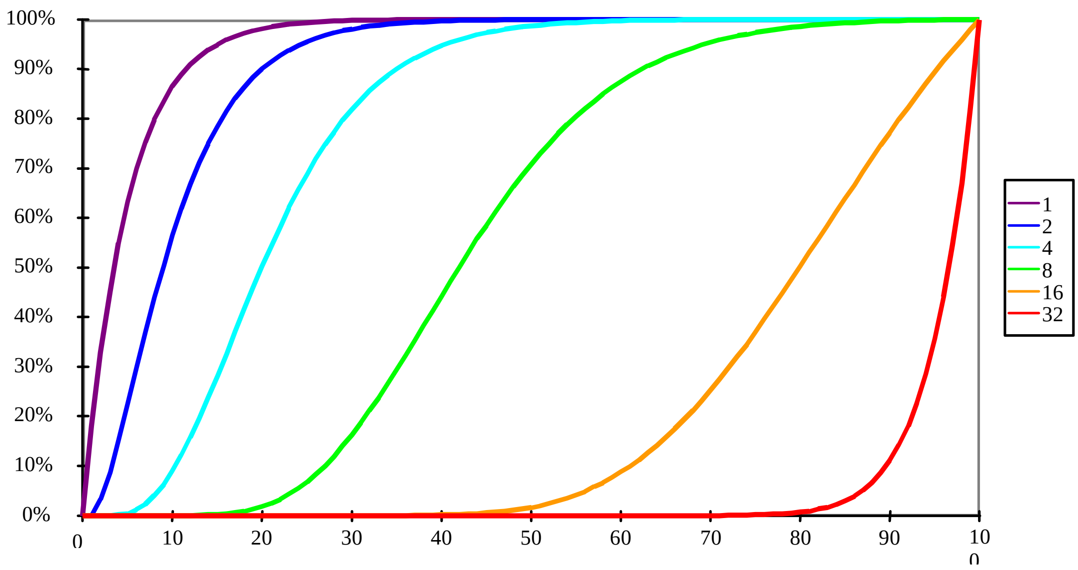
[13] The studies that have looked at this question find that 5-10 dimensions are not uncommon, and in some markets the dimensions of competition can be much higher. One study of the Japanese camera market found that an average of 3-4 new features per annum had appeared over a 20-year period - though these would not all be of equal importance.

[14] The distribution of benefit between producers and customers depends on the extent of competition.

the cumulative (weighted) economic benefits at any time (as a percentage of the total by the end of the programme) follows the relationship illustrated in Figure 12.

Figure 12

Economic Benefits of Investment in Measurement



The incidence of benefits from a measurement programme over the lifetime of the programme depends critically on the value of r. For single-dimensional products, most of the benefits from the measurement programme occur at the start. As the programme matures, it may make further characteristics measurable, but this is of little commercial significance. By contrast, when products are very complex (large r) then the majority of benefits accrue at the end of the programme. This is not to deny that some of the product dimensions that become available towards the end of a programme are in themselves of lesser value. But when r is large, the number of product locations becoming available grows at a prodigious rate towards the end of the programme. For intermediate values of r (r = 4 or 8 for example) we find a s-shaped curve relating economic benefits to investment.

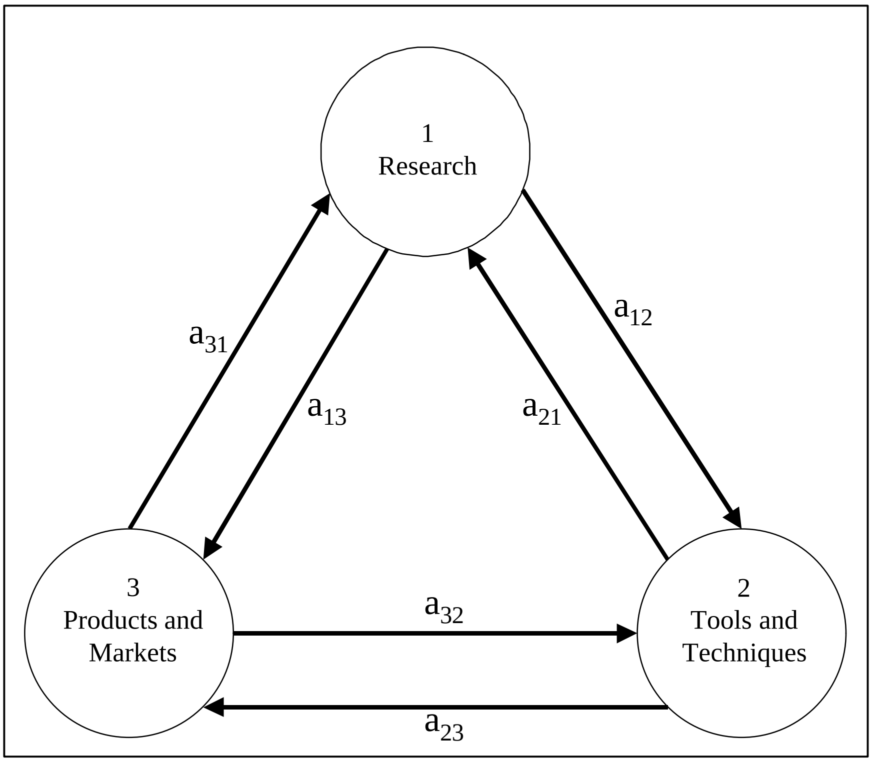
As stressed at the start, this is not a general-purpose model. It is perhaps most relevant to the development of a measuring device (e.g. a sensor, an instrument or a test kit). It is probably less relevant to cases where new market opportunities arise from the discovery of a new effect of a material, rather than from an enhanced ability to measure. To handle these other cases, a somewhat different model would be required.

One apparent implication of this model is that while benefits may accrue fastest from those measurements that support the least complex products (low r), the ultimate benefits may be greatest from measurements that support more complex products occupying a higher-dimensional characteristics space. Swann (1997) explores this argument, and examines what it may imply about the comparative returns from investment in physical measurement and in chemical or biological measurement.

Virtuous Circles

To see the virtuous circle at work, we shall look at a simple structure in which the phenomenon can be illustrated without undue complication. We start by adapting the simple structure used by Callon et al (1992). Suppose that we identify three broad sectors as follows: (1) Research, (2) Tools and Techniques, and (3) Products and Markets. In Figure 13 below, each box represents one of these three broad sectors, and the arrows between these boxes represent the flow of valuable knowledge or goods and services.

Figure 13



The magnitude of these flows is summarised by the parameters a_{ij}, which are assumed to be non-negative. These represents the long run effect of sector i on sector j. Loosely speaking these represent:

a_{12} Research discoveries valuable to development of tools and techniques
a_{13} Research discoveries directly valuable to markets
a_{21} Tools and Techniques aiding research - e.g. shaping of research agenda
a_{23} Goods and measurement services sold in final markets
a_{31} Customer input to defining the research agenda

a_{32} Customer input to design of tools and techniques

As it stands this is a fully simultaneous model. For this fully interactive model, we can show (see Appendix to Swann, 1996) that the long run effect of research on markets is:

$$\frac{\text{[Direct effect of 1 on 3]} + \text{[Indirect Effect of 1 on 3 via 2]}}{1 - \text{Feedback}_{1\text{-}2} - \text{Feedback}_{2\text{-}3} - \text{Feedback}_{1\text{-}3} - \text{Feedback}_{1\text{-}2\text{-}3} - \text{Feedback}_{1\text{-}3\text{-}2}}$$

The direct effect of 1 on 3 is simply a_{13}. The Feedback terms represent the strength of the feedback loop between two (or three) sectors. For the loop from sector 1 to sector 2 and back, $\text{Feedback}_{1\text{-}2} = a_{12}*a_{21}$. And for the three stage loop from 1 to 2 to 3 and back again, $\text{Feedback}_{1\text{-}2\text{-}3} = a_{12}*a_{23}*a_{31}$.

The conclusion is very clear. We should not judge the effect of research on the market just by the direct effect. We should also add on the indirect effect via tools and techniques. But even more important, if these feedback loops are substantial, then the denominator of the equation may be quite small. In that case the long run effect, taking account of all feedback loops is substantially in excess of the numerator alone.

While the results are messier for more complex chains and loops, the basic character of the result is the same. This is why we stressed the importance of feedback loops while discussing virtuous circles in Section 2.1.

2.4 The Economics of Research

There is a huge literature on the Economics of basic research and of R&D. We cannot do justice to it here. [15] However, it is useful briefly to pick out a few themes in that literature, which will surface again later in the report.

Research is an investment, and an investment that can enjoy very high returns. But the returns are also highly variable. These rates of return are, moreover, very context specific. To quote an average is not very helpful when the variance is so very wide. For example, Griliches (1992) survey found rates of return on publicly funded agricultural research between 11 and 83 per cent.

Measurement can be seen as a form of investment, too. The firm invests in measurement to achieve a particular result, to learn some information, to establish that a product is what it purports to be, or to reduce some elements of risk. However, I am not aware of any studies that have looked at the economic returns to investments in measurement per se.

[15] Two recent, and comprehensive surveys are by Martin et al (1996) and Swann (1996).

Since the returns from R&D are so unpredictable, it is seen as a relatively risky form of investment. Risk averse investors may be concerned at such risky investment, and this can lead to what is seen by some as under-investment in R&D.[16] Equally, one could expect the returns to investments in measurement to be risky. However, measurements are often made to *avoid* risk. For that reason we need to compare the total risk that obtains when there is *no* investment in measurement with the total risk when there *is* investment in measurement.

One of the most important features of research as an investment is that it is rich in spillovers. Griliches (1992) assessed the importance of spillovers by comparing private rates of return on R&D (to the company funding the research) to social rates of return. But it is fair to say that measuring spillovers is a difficult matter - an issue to which we return in Part 4. Jaffe (1989) and Jaffe et al (1993) found that the degree of spillovers is spatially determined. In the US context, most spillovers took place within a SMSA (standard metropolitan small area), some within a state, and few spread beyond that. Some research using bibliometrics has found similar geographical patterns.

Another strand of the literature argues that if R&D has to be done rapidly - in response to a major competitive threat, for example - the cost of achieving the desired result will be higher than it would in calmer times.[17] Finally, some theorists have taken an *information theoretic* approach to investment in analytical measurement (e.g. Eckschlager and Stepánek, 1985). Interestingly, while information theory enjoyed some prominence in various areas of economics in the nineteen-sixties and nineteen-seventies, these have not been of lasting significance and the perspective has not found widespread use within economics.

[16] The "short-termism" debate stems from this observation.

[17] The model underlying this assertion is quite a sophisticated one, but the basic idea is this. The R&D manager in a hurry must pursue several research lines in parallel, including many that turn out to be unprofitable, whereas in calmer times he would pursue different research lines sequentially, leading to less wasted effort.

3. The Rationale for Public Funding

This part of the report looks at the rationale for public funding of (some) measurement activities. The analysis that follows examines *whether* there is a case for public funding, and if so *what programmes should be prioritised* ? From what I learnt in an early discussion with the NMS Review team,[18] the short-term issue for discussion and analysis is not so much "whether?" but "what?" But while the discussion that follows may seem more relevant to "whether" it is, in fact, equally applicable to "what?"

3.1 Generic Rationale

The economic justification for industry policy tends to be one of three sorts. First, that policy is required because there is *market failure* requiring some sort of correction, or at least compensating activity. Second, policy is needed to *regulate* private monopolies (though this could be seen as a special case of the first). Third, a “strategic” rationale - for example, those programmes designed to give a new industry a boost (or “kick-start”) so that it moves onto a faster growth curve.

My assumption from the start has been that the justification in the case of NMS is the first sort. But while completing Part 2 of this report, it strikes me that the “strategic” rationale may actually be quite relevant for certain NMS activities. So for most of this part of the report we shall focus on market failure, but we shall return at the end to look at the "strategic" rationale.

Economists tend to identify three generic causes for market failure. The first is that *externalities* (whether positive or negative) drive a wedge between private and social returns from a particular private investment. If externalities are positive, some socially desirable investments will not appear privately profitable, so the market does not support enough activity. If externalities are negative, some socially undesirable investments nevertheless appear privately profitable, so the market supports too much activity.

The second is where economic activities are subject to *increasing returns*. In that case there is no unregulated market outcome that is also economically efficient. If perfect competition is sustained, then production does not exploit the increasing returns, so costs are not minimised. If monopoly is allowed to emerge, the monopolist may be able to exploit the increasing returns, but is liable to restrict output to keep up prices.

The third is that *asymmetric information* between buyers and sellers can make it impossible to find a price at which to trade that is acceptable both to buyers and to sellers. One example of this is Gresham’s law which asserts that “bad drives out good”. The presence of "bad" products in a market, and the inability of the buyer to distinguish bad from good *ex ante*, means that the supplier of good withdraws his produce from the market as he cannot raise a satisfactory price (Akerlof, 1970).

---

[18] Meeting at DTI, 151 BPR, 25 March 1999

One important economic function of the NMS is of course to correct the information asymmetries described above. We shall devote a section to this below. Moreover there are clearly aspects of the NMS where increasing returns apply, so that is also relevant. In particular, we saw in Part 2 that measurement research had large fixed costs.

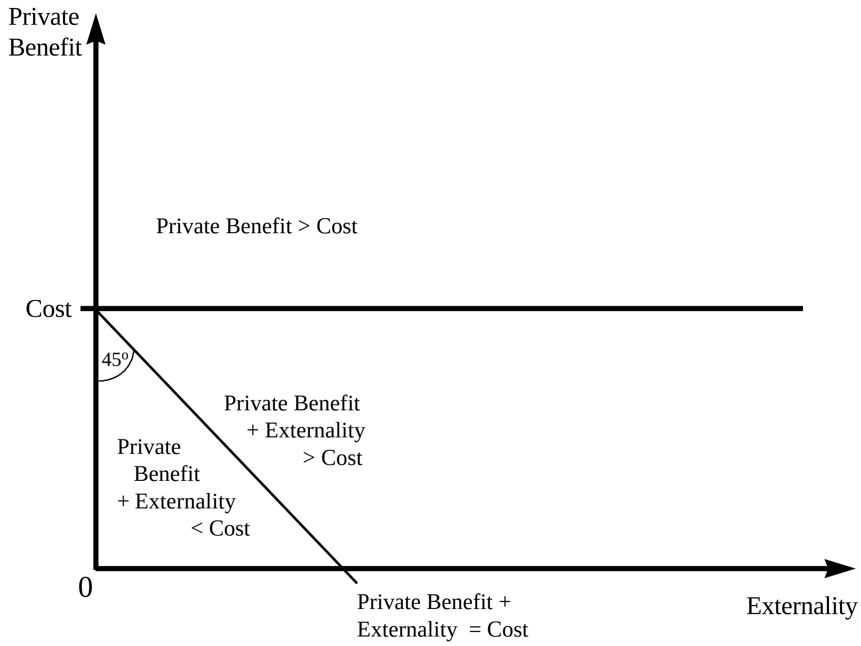
However, our working hypothesis is that the most important source of market failure that needs to be properly understood in this context, is the existence of externalities (mostly positive) from some privately financed measurement activities. John Barber's report of 1987, and my earlier report for LGC put much emphasis on the relevance of externalities in the measurement context. However, it would be timely to give further attention to the form of, and reasons for externalities.

The structure of this part of the report is as follows. Section 3.2 explains why externalities are important in deciding NMS priorities. Section 3.3 explains why externalities arise and why the rate of *internalisation* (as we shall call it) varies from context to context. Section 3.4 goes back to the asymmetric information rationale for NMS activity, and explains why customer-driven requirements for measurement may be important. Section 3.5 examines the "strategic" rationale. In conclusion, Section 3.6 assesses the relevance of each rationale from the analytical perspectives of Part 2.

3.2 Why do externalities matter in assessing NMS priorities?

A series of diagrams will be helpful in developing this argument. Figure 14 can be used as a map of the private benefits, costs and externalities from particular NMS projects.

Figure 14



Any particular project can be located as a point in this map. Projects located above the horizontal line labelled “cost” are privately profitable, while those below the line are not privately profitable. But that is not the end of the matter. The diagonal line, at 45° to the vertical axis, shows all the projects for which the total social benefit - i.e. private benefit plus externality - is equal to cost. Projects above and to the right of this line are socially worthwhile, while those below and to the left are not.

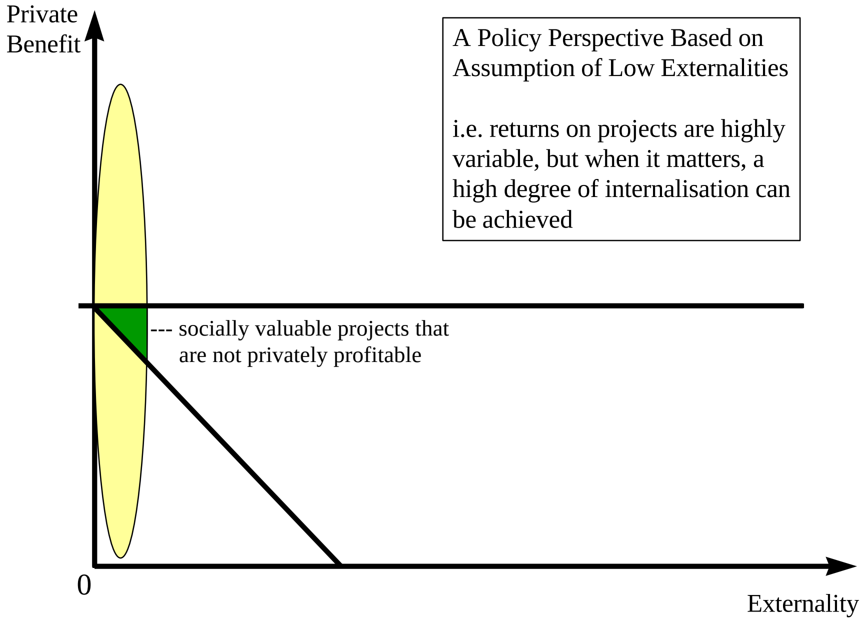
Accordingly, Figure 14 is divided into three regions. Above the horizontal line are projects that are both privately and socially profitable. These should happen anyway, without any additional activity on the part of NMS. Below and to the left of the diagonal line are projects that are neither privately nor socially profitable. These should not take place. But in the third region - above and to the right of the diagonal line but below the horizontal - are those projects that are socially worthwhile but not privately profitable. These are the projects on which NMS activity should, arguably, be focused.

Assumptions about the joint distributions of private benefits and externalities

One may have little idea about the actual empirical magnitude of externalities. Nevertheless, to make any assertions about policy priorities, one must make some assumptions about the form of externalities - even if these assumptions are implicit rather than explicit.

Figures 15 to 18 summarise four different sorts of assumptions that might be made - and I have encountered different people who hold each of these opinions. In each case, Figure 14 is overlaid by an ellipse showing the assumed joint distribution of private benefits and externalities. These ellipses show the locus of (say) 99% of all projects.

Figure 15

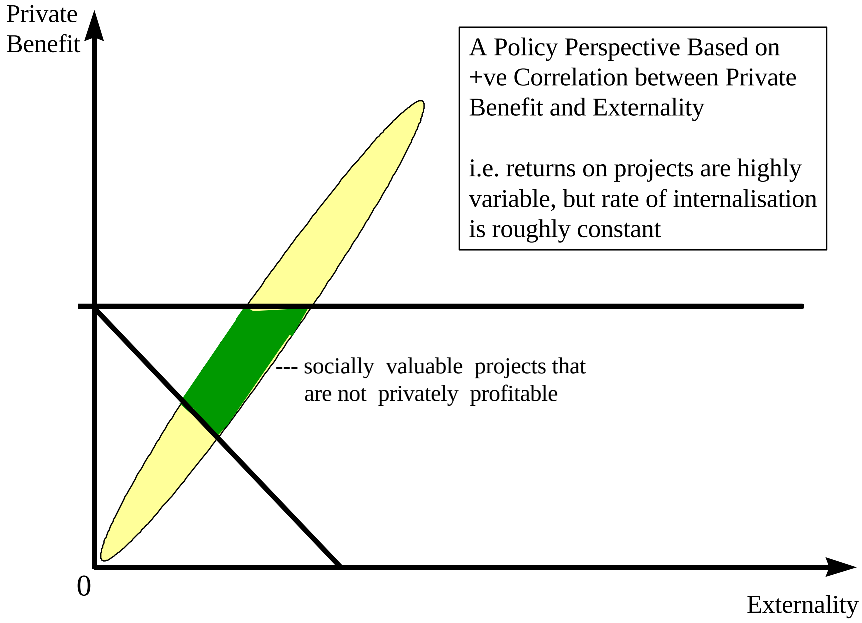


In Figure 15, the assumption is that even if the returns on different projects are highly variable, there are few externalities. In other words, those investing in the project are very successful at internalising these returns - at least, when it matters to achieve this. From such a perspective, of course, externalities are really not an issue. Figure 15 shows a small dark triangle representing socially valuable projects that are not privately profitable. These represent such a small share of all projects, that it will not be a matter for huge concern, for most people, if these are overlooked.

The logical extension of this perspective would appear to be that there is no role for public intervention - at least not to correct market failures deriving from externalities.

In Figure 16, by contrast, it is recognised that externalities exist, but the assumption is that the ratio of externalities to private benefits is roughly constant. The distribution of projects is tightly clustered around a line from the origin, and the rate of internalisation (i.e. the slope of this line) is constant. From this perspective, externalities are somewhat more important. A dark trapezium represents the projects that are socially worthwhile but not privately profitable. And compared to Figure 15, there are rather more of these.

Figure 16

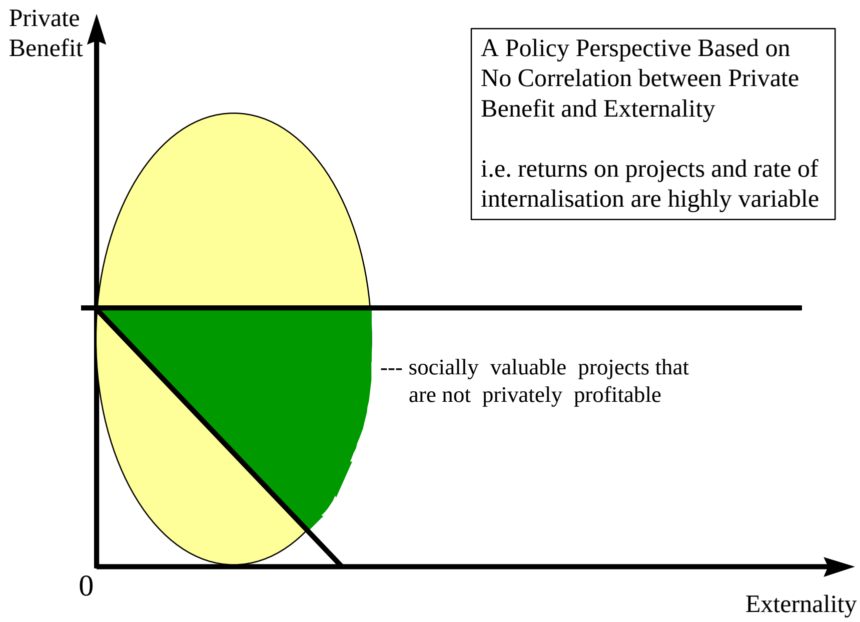


Nevertheless, given this strong correlation between private benefits and externalities, the ranking of projects by private value will produce a similar ordering to the ranking of projects by social value. As a result, if NMS uses private value as a measure of priority, that would be reasonable enough.[19] But once again, this perspective poses the question of whether projects above the horizontal line actually need public support.

In Figure 17, the assumption is that returns on projects are highly variable, and moreover that the rate of internalisation is also very variable. In that case, the dark triangle showing projects that are socially profitable but not privately profitable is large - bearing in mind of course that the area of the ellipse is also larger. Moreover the majority of projects below the horizontal line are "dark". This is in contrast to Figure 16, where about half the projects below the line were not socially worthwhile.

[19] It seems that some of the policy perspectives that have been expressed in relation to earlier NMS work embody the assumptions behind Figure 16.

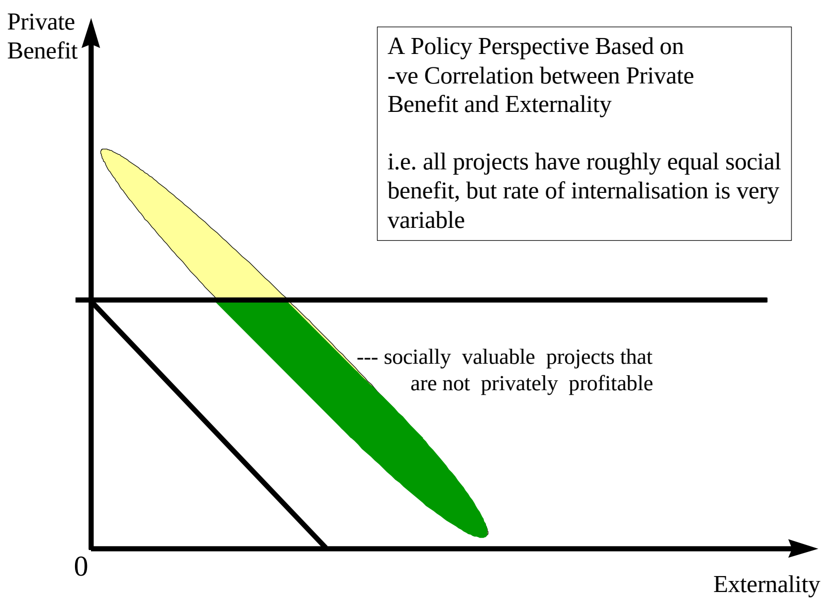
Figure 17



Here, the NMS has a lot of work to do, and because there is no correlation between private value and externalities, the private benefit from a project is not a good indicator of its social value. If this perspective seems plausible, then it is essential to find out some more about externalities.

Finally, Figure 18 illustrates a perspective that might seem to be pathologically anti-free-market, but I have certainly encountered people of less extreme views who take such a perspective - whether explicitly or implicitly.

Figure 18



Here, the assumption is that we have a portfolio of projects all of roughly equal social value - hence the projects are all scattered around a line at 45° to the axes. However, the rate of internalisation is extremely variable, so that instead of a positive correlation between private benefits and externalities, we actually get a negative correlation. This happens because in those cases with high internalisation, we have high private benefits and low externalities, while in other cases of low internalisation, we have low private benefits and high externalities.

In this case, all the projects below the horizontal are "dark", so the NMS has the greatest amount of work to do. Moreover, it makes no sense to use private benefit as an indicator of social value in setting priorities.

Which of these figures corresponds to the real world? Or can we say which applies in which circumstances? In short, the answer is that we just don't know. It will depend on the selection of projects, to be sure, but an important objective of the consultants' empirical work should be to establish which of these pictures is most common.

* * * *

Having seen how sensitive policy prescriptions can be to the assumed character of externalities, the reader will understand why I consider it important that the present NMS review tries to make some progress in identifying and assessing externalities.

At the risk of labouring the point, it is worth stressing that the above observations have an important implication for the choice of an empirical methodology to value measurement activities. Any technique that focuses on the "bottom line" value of a project to the proximate users will find it harder to assess any spillover benefits. As such it can only generate an *incomplete* measure of the economic benefits from measurement.

Moreover, it is worth drawing out one further implication of the above. Suppose that measurement activities are allocated between private and public sector as follows. Those where the private return is in excess of the cost are purely privately financed, while those where the full social return exceeds the cost but the return to proximate users does not will be funded by the public sector. If an *incomplete* methodology is used, then two results will follow:

- private sector projects will be measured to have a higher benefit to cost ratio than public sector projects

- public sector projects will not appear to be worthwhile

These "findings" will resonate with some political perspectives, perhaps. But the "findings" would have nothing to do with the true social merit of the different projects; they would simply follow from the unavoidable measurement error inherent in using an incomplete empirical methodology.

In conclusion, however, we should add that these "findings" stem from particular assumptions about the allocation of projects between private and public sectors. Another possibility that needs further thought here is the possibility that public sector provision may crowd out private provision. To limit this costly and undesirable outcome, it is important to ensure that public sector priorities focus on the activities that are rich in spillovers.

3.3 Internal or External?

Section 3.2 shows why externalities are important in assessing priorities within NMS. But why do they arise at all, and why should the rate of internalisation vary from context to context?

John Barber's early paper (1987) gives a good account of why externalities are important for some sorts of measurement activity. My paper to LGC also comments on this. The basic point is if a single company or group of companies invest to create a new or better method of measuring a product characteristic, that will yield financial benefits to then, but will also benefit some other companies who are not party to the project. To the extent that these spillover benefits accrue to companies who are external to the project, and are not charged for these benefits, then the spillovers can be called externalities.

Barber argues that some aspects of measurement have the character of "pure public goods". There exist some projects generating such widespread spillovers, and it is so hard to exclude beneficiaries unwilling to pay, that the project is best handled as a "pure public good". This is financed in the same way as company accountants treat a generally beneficial overhead. That is, that they do not attempt to apportion costs to users in proportion to some index of use, but instead finance the project from a "top-slice" on general revenues. In the public context, that means financing the project from general taxation.

A typically "Coaseian" [20] response to this observation would be: why do the project managers not require those third parties that enjoy the spillovers to join the group paying for the project, and hence *internalise* these spillovers? Indeed, economists in this tradition argue that there are fewer pure public goods than is commonly supposed. Even the lighthouse - often cited as the purest form of a public good - is no such thing, they assert. They point out that when lighthouses were first built in Britain, they were financed by tolls on ships when they arrived in port, not out of general taxation. This mechanism works because ship-owners unwilling to pay for the benefits yielded by the lighthouse are excluded from using a complementary service (the port) even if they cannot be excluded from using the services of the lighthouse. A modern equivalent to this might be to finance a sector-specific measurement activity from a general levy on members of a specific trade association. While those in the industry unwilling to contribute in this way could not perhaps be excluded from use of the new measurement approach, nevertheless they could be excluded from the trade association - and it is this fear which persuades them to pay up.

[20] So called after followers of the Nobel Prize-winning economist, R. Coase (1960, 1974).

Indubitably however, it can be a costly business to track down all beneficiaries from spillovers and charge them for the benefits they receive. Sometimes the benefits of doing this justify the costs; often they do not. My view on the Coaseian argument is this. The argument is correct in a world of perfect information and zero transaction costs. While there are indubitably some schemes for internalising some externalities in real worlds, there are in practice still - and always will be - a very large number of externalities that evade any levy of this sort.

For NMS, however, it is important to understand what factors influence the rate of internalisation that can be achieved. Figure 19 is an attempt to illustrate why it may sometimes be easy to track down beneficiaries from spillovers, but sometimes may not. Suppose that a project to create a new measurement method will benefit those who seek to introduce a particular characteristic in their product or service. And suppose that amongst a population of companies, about a third would find it helpful to use this characteristic. Suppose also that these companies can be represented in a two dimensional map. The axes of this map could represent geographical space, or they could represent a more subtle competitive space. But assume at any rate that proximity in this map implies corporate proximity. So for example, all the members of a particular trade association would be clustered together in a particular part of the map.

Figure 19 represents the results of two simulations from a model of diffusion across this population of companies. In this figure, companies that use the characteristic (and hence the measurement activity) are shown in yellow (pale), while those who do not are shown in blue (dark). In the top part of Figure 19, diffusion essentially follows an *epidemic*[21] process, where companies are very likely to adopt characteristics and processes that are used by their neighbours. As a result, use is clustered into contiguous or coherent blocks. In the lower part of Figure 19 by contrast, diffusion is essentially determined by a "*probit*" (or firm-specific) diffusion process. Here firms are less concerned by the behaviour of their neighbours, but more by the benefits they themselves would enjoy by adopting a new characteristic. As a result, use is not clustered, but instead is spread in a patchwork fashion across the whole population.

Consider the task faced by those who seek to levy charges on those beneficiaries to a spillover, and hence internalise the spillover. If he faces a map as in the *top half* of the diagram, he has a relatively easy task to locate beneficiaries. They are almost all in contiguous block, and some at least may belong to the same trade association. So it looks as though it may be relatively easy to finance this activity by *club* subscriptions.[22] There is still the task of designing a mechanism that excludes these beneficiaries from some related benefit if they do not pay for that. But at least the task of identifying and negotiation with the beneficiaries is fairly straightforward.

Consider, by contrast, the task facing him if he has a map as shown in the lower half of Figure 19. Here it is a very hard task just to locate the beneficiaries, let alone to

[21] The meaning of the terms "epidemic" and "probit" is explained in Section 4.4.

[22] Where we use the term "club" in the local public-good sense employed by Tiebout (1956) and Buchanan (1965).

survey them. In such a dispersed collection of companies, it is unlikely that many belong to the same trade association. In such a setting, the inclination would be to give up and accept that this is close to being a pure public good, and that the Coaseian solution is denied by sheer transaction costs.

This example is simplistic but it demonstrates two important points:

- First, that the extent to which spillovers can be internalised rather than leak out as externalities depends on the spatial distribution of the beneficiaries - where we use the term "spatial" in the broadest socio-economic sense, and not just in a geographical sense.

- Second, that just as the spatial pattern of diffusion of use of a new measurement technology depends on the precise character of the diffusion process, so also does the achievable rate of internalisation of spillovers. In this simple example, it appears that *epidemic* diffusion (where there is much higher probability of contagion from neighbouring companies) leads to a much higher achievable rate of internalisation than *probit* diffusion. However, it would be unsafe to generalise from this simple example, and these issues need further careful attention.

At any rate, these observations offer an interesting angle on how we can assess the importance of externalities for NMS priority-setting. Externalities may be of least concern in stable environments, where the most likely group to use the new measurement approach are spatially contiguous, and indeed where there are very few beneficiaries from other sectors or regions. In short, externalities may be of least significance where beneficiaries are distributed as in the top half of Figure 19 - because there it is possible to achieve very high rates of internalisation.

By contrast, externalities are clearly important in contexts where the beneficiaries are not a tight-knit socio-economic group, where *de novo* entry and cross-entry are important, and where indeed the benefits from a particular measurement project spread across a number of sectors. In this case, we have a pattern of externalities where beneficiaries are widely dispersed, as in the lower half of Figure 19. Indeed, the prevalence of these forms of entry may be a good indirect indicator of a measurement system generating important social benefits - see Section 4.6.

Where would this second pattern be most common? Certainly, where (in contrast to Jaffe et al's research summarise in Part 2) diffusion is not mediated by socio-economic proximity, and where many firms are innovators in the sense used by Koestler (1969) and Simon (1985) - that is, they bring together insights from diverse knowledge bases. This would be relevant where the mass *broadcasting* media play an important role in diffusion - so that diffusion can spread across the competitive space with ease.[23] To the extent that those who make use of new characteristics, and the measurement methodologies that support these, are Koestler/Simon innovators, then they will be very dispersed in spatial terms - and hence we can expect a very low rate of internalisation. In short, the world facing NMS is becoming more like the lower half of

[23] An *important* distinction arises between diffusion driven by broadcasting (i.e. one-way) and diffusion driven by *communication* (two way).

Figure 19, and less like the top half. So we can expect that NMS will have more and more work to do.

We conclude this section by noting an implication of the above that may seem surprising to some. "Club" solutions presume stability and familiarity. "Club" solutions, moreover, are only sustainable if they do not promote a pattern of measurement usage that disrupts this stability and familiarity. In contrast, publicly funded measurement systems encourage the sorts of innovation by outsiders that disrupts this stability and familiarity. So while measurement competes for part of the innovation budget, it must be seen as a complementary activity to innovation. Without necessary measurement techniques, there can be no success in innovation. Those who do not do well on new characteristics will not invest to create relevant measurement methodologies. Measurement "clubs" may be captured by incumbents, who could resist new product dimensions that threaten their competitive position. So a public measurement infrastructure is necessary for the most radical innovative advance.

3.4 Some Thoughts on Asymmetric Information

As noted at the start of this paper, asymmetric information (between buyers and sellers) is one of three generic causes of market failure. It is conceptually distinct from the externality as a source of market failure, although attempts to correct asymmetric information can sometimes lead to an associated market failure resulting from externalities. This happens when the better-informed vendor incurs investment costs himself to improve the information available to the buyer, but where there are externalities from this investment.

It is unlikely that one could ever measure directly the importance of asymmetric information - for indeed, the act of measuring would to some degree correct the asymmetry. Nevertheless, it is important to understand the circumstances in which asymmetric information could arise, because these could priority areas for NMS activity.

For simplicity, consider a single-characteristic product. Suppose that the measured value of the characteristic is z, where the measurement follows a normal distribution with (true) mean μ and variance σ^2. Moreover, while the producer knows the actual accuracy σ^2 of this measurement technique, the customer only has an estimate of this accuracy, s^2.

So the actual distribution of measurements is as follows:

$$z \wedge N\ (m, s^{\ 2})$$

but the customer thinks it is:

$$z \wedge N\ (m, s^{\ 2})$$

Suppose that the customer's utility is given by:

$$U = z\left(1 - f\, s^{\,2}\right)$$

But because the customer is ill-informed, he thinks the product is worth:

$$U = z\left(1 - f\, s^{2}\right)$$

There are three possible cases:

1. If $\sigma^2 > s^2$, then the customer is bearing undue risk, without knowing it. This could mean he ends up buying a product that he would not buy if he knew the true risk.

2. If $\sigma^2 < s^2$, then there is an Akerlof-type market failure.[24] The buyer only has a relatively inaccurate measurement apparatus at his disposal, and cannot identify good from bad. As a result, fearing that the seller will abuse his superior information, and sell him a "lemon",[25] he doesn't buy.

3. If $\sigma^2 = s^2$, then there is no asymmetric information. The seller is not in a position to exploit the customer's ignorance. If however, σ^2 is higher than it needs to be - because the latest (more accurate) measurement technologies are not being used - then the market will not be as big as it could be.

A seller of a good quality product (A) facing case 2 would have an incentive to ensure that the customer is better informed about the true quality of product A - even if the seller of a lower quality product would not. The seller of A might achieve this in a variety of ways (e.g. by obtaining independent certification that product A meets a given standard). Moreover, the seller of A would (in case 2 or 3) have a particular incentive to invest in improved measurement methods that reduce σ^2 and s^2.

On the other hand, a seller facing case 1 might not feel a strong incentive to improve the customer's information. So while it would be good for the customer to be better informed,[26] the seller would not necessarily gain from making an investment in this direction. While this argument is incomplete, and would need some further development before we can draw any firm conclusions, the example is suggestive of some circumstances in which the seller does not gain from investing to better inform the customer.

In summary, those who control particular measurement technologies will not have an incentive to create measurement technologies that demonstrate dimensions in which they are weak. This is especially relevant when the rate of addition of new characteristics is high, and hence there is a strong need for new measurement techniques to handle hitherto unmeasured characteristics.
Moreover, when we take a slightly more subtle view of what characteristics actually are, this argument can become even more important. In most of the literature on

---

[24] Following the pioneering paper by Akerlof (1970). Actually, in Akerlof's analysis, $\sigma^2 = 0$ and $s^2 > 0$, so outcome 2 above is the inevitable one in his paper.
[25] American slang for a car that looks OK at auction, but breaks down as soon as you get it home.
[26] We are assuming here that the popular saying, "what you don't know can't hurt you" is false.

product characteristics, following Lancaster (1971), it is assumed that the characteristics themselves are defined by the producer of the product. Bacharach (1991) takes a rather different approach. In his analysis, characteristics are - in effect - answers to questions about the product. These questions may indeed be highly predictable ones, posed by (and answered by) the producer. But they may also be questions posed by the consumer.

In this latter case, there may well be circumstances in which it is not in the producer's interest to invest in measurement methods that improve the reliability of answers. For if the question from the consumer is one which makes the producer uncomfortable, it is probably easier for the producer if available measurement methods are none too accurate. As Rabinovich (1993) argues, it is generally important that those who measure do not overstate the accuracy of their measurements - that seems obvious enough. But he argues that it is also important that they do not *understate* the accuracy of their measurements. Here is a case where the consumer needs to hear the full story, to avoid the risk of purchasing the wrong product, but it is better for the seller if that bad news cannot be articulated.[27]

In short it may well be that no producer-oriented club will have an incentive to invest in measurement methods to answer awkward questions from assertive customers. This is a priority, *a fortiori*, for NMS activity, because with a few notable exceptions, it has generally proven very hard for consumers to achieve a "club solution" in such cases.

This is why the "return route" from customers to producers in Figures 8 and 9 (in Section 2.1 above) may be rather weak. That is why Figure 9 made much of the potentially important role for NMS to complete the virtuous circle by offering a "return route" from the customer to the research agenda.

It would clearly be hard to create a macro-level index of the degree to which product characteristics in an industry are consumer-defined rather than producer-defined. Nevertheless, this last argument is suggestive of particular product markets in which consumers are voicing concern over characteristics that they wish to have measured even if producers do not measure these. It could well be important for NMS to prioritise the advance of measurement techniques relevant to these cases.

3.5 A Strategic Perspective on NMS Activity

As noted at the beginning of Part 3, my initial expectation was that the "strategic" rationale for public funding was probably not relevant here. We think of the "kick start" as something done with young industries or new technologies, where ultimately we see a self-supporting growth in the industry, but where the industry needs an initial "kick start" along its growth path. So for example, DTI programmes such as MAP

[27] A related argument is the following. It would be a good thing for the general public if a new technique for toxicity measurement were made public. However, if this new technique is created by a company that obtains strong competitive advantage from exclusive use of the technique, that company will not make it public. This competitive advantage might arise, for example, when use of the technique is necessary for regulatory approval. I am grateful to Geraldine Alliston for this observation.

(for microprocessors) and the VR Awareness Programme (part of ISI) can be seen in this way. That seems a rather different sort of programme from continuing involvement with NMS.

I feel this basic intuition is right. But for completeness it is worth a moment's reflection on the “kick start” metaphor. For any economic system that can be represented as a "virtuous circle" or as a "common pool", it is natural to ask whether a “kick start” may be relevant. Can it help to accelerate the virtuous circle? Or could early contributions to the pool help to accelerate growth of innovations drawing on that pool?

If DTI and NMS consider that Figures 8-10 are a useful representation of the role of measurement, then it could still be helpful to see publicly funded measurement activities as having a “strategic” role. One aspect of this is it is helpful to strengthen the feedback loops in these diagrams. The second is that it is helpful to support some pioneering activities that generate an additional and ongoing increase in traffic within the virtuous circle.

I offer this as an issue deserving a little further thought and discussion, but accept that DTI and NMS will probably prefer instead to focus their attention on the dominant "market failure" rationale.

3.6 Conclusion: How relevant are these rationales to Measurement?

Part 2 of this report set out three approaches to describing the role of measurement. The first was functional and sectoral, and while it is useful for categorising those measurement activities that take place, it does little to extend our economic understanding. The other two approaches categorised measurement activities in terms of their economic effects:

- The Economic Map
- The Common Pool / Virtuous Circle

In the economic map, some aspects of measurement have a similar economic effect to standards. As we discussed in Part 2, the argument for public funding of standards is that this promotes trade, increases network externalities and corrects market failure. If standards setting is left to the market, we either get monopolisation or we don't get a standard. The "economic map" perspective suggests a similar case for public funding of measurement. Yes, such public institutions may be susceptible to “capture” by large organisations that can afford to send staff to participate in such committees. (The present author has examined this question in earlier work on standards committees.) However, if standards setting (and, we suggest, measurement) the degree of monopolisation is likely to be an order of magnitude greater.

From the perspective of the common pool and the virtuous circle, measurement standards have the following effects: (1) they make it easier for innovators to exploit the common pool; and (2) they strengthen the feedback loops in the virtuous circle.

The first consequence is similar to the role of common standards in extending network externalities - and that is an important part of the rationale for publicly funded standards setting. The second consequence is important in the light of observations in Section 3.4. I suspect there is good reason to doubt that radical measurement ideas proposed by customers will find an easy "return route" through markets and producers. In the light of that, I suspect NMS has an important role to provide the "return route" illustrated in Figure 9. This may not be how NMS perceives its role at present - but I think this observation is worthy of further discussion.

4. Measurement Methodology

Amongst economists, applied econometrics is perhaps the most widely used empirical technique. It is a key part of the economist's education, and some economists rarely use any other empirical techniques. The economist working in a business school, however, is reminded that amongst the full range of social sciences and business studies, a much wider range of empirical techniques are available.

The following list is not complete but gives an idea of the range available:

- Econometrics
- Simulation and Visualisation
- Engineering Economics
- Experimental Economics
- Direct Measurement: Tools and Techniques
- Paper Trails (e.g. patent citations, bibliometrics)
- Surveys and Questionnaires
- Economic and Business History
- Case Studies
- Interviews and Ethnography
- Delphi: The “Oracle” in his/her own words

In my report to LGC, I argued that four of these (econometrics, engineering economics, direct measurement, and case studies) would probably be most useful for better understanding the economic effects of publicly funded measurement activities. I think that perspective remains broadly correct, though in the light of some discussions with DTI and NMS it is possible that simulation and "paper trails" could also be useful.

The aim of this Part of the report is three-fold:

a) to give the reader a rough understanding of the empirical techniques that may be needed;
b) to sketch out the empirical research problem, and explain why it is so difficult;
c) to move towards a suggested strategy for empirical work.

Section 4.1 briefly describes the empirical techniques that are advocated here, and Section 4.2 summarises their relevance in this context. Section 4.3 describes why it is so difficult to trace and measure externalities. Section 4.4 explains why an understanding of the social and economic processes of diffusion is needed to trace externalities. Section 4.5 suggests some categories into which to place externalities. Section 4.6 then sets out a strategy for empirical work.

The strategy described in 4.6 is not meant to be exhaustive, and will certainly require further discussion with DTI and NMS, but I feel that it is the right way to match available techniques to the research problem. To some extent it is an ideal, and may well involve more time and resources than available. The rationale for pursuing two streams is this. Many social scientists have argued that researchers should pursue two different approaches to measuring a phenomenon, where the strengths and weaknesses of different techniques are mutually offsetting. In particular, there is a strong case that applied economics research should seek, wherever possible, to combine methods of indirect inference with direct measurement. Direct measurement is obviously attractive because it aims to get directly at X, the phenomenon of interest. But sometimes direct measurement of X is difficult (as here with externalities), and hence it is also useful to make indirect inferences about the likely size of X from measurements of related phenomena that are easier to identify. Or as Polonius has it: "by indirections find directions out".[28]

4.1 Empirical Approaches

For those readers who are unfamiliar with the research methods of applied economics, it may be useful if I summarise the four main techniques that I envisage being of use to DTI and NMS. This section reproduces (with a few editorial changes) Section 5 of my report to LGC (Swann, 1997). Those familiar with these techniques, or who have already read that earlier report can turn now to the next section (4.2).

Direct Measurement

The empirical approach used by Scientific Generics could be described as direct measurement. They have attempted to measure benefits directly by polling some of those who have benefited. Admittedly, this direct measurement is subjective assessment by the user rather than a hard objective measure, but it is probably as good a sort of direct measurement as can be done in the circumstances.

The advantage of direct measurement is that it approaches the magnitude of interest directly, and without relying on inference from correlations between different economic data, or on deduction from some theoretical model. The disadvantages are:

(1) the mechanisms identified may not be exhaustive;
(2) that it is hard to achieve a comprehensive coverage of all beneficiaries, and as a result the technique has an inbuilt bias
(3) respondents may not understand how a mechanism yields benefits for them, and may not realise that some benefits they enjoy are attributable to this mechanism.
(4) users may focus on "bottom line" effects for their companies and miss the wider economic benefits of these mechanisms.

If we accept the arguments presented in Parts 2 and 3 above, it seems probable that the benefits of measurement are not simply captured by a few principal users. The beneficial "market-creation" effects extend more widely, to a wide range of firms and customers. Indeed if we concentrate on how measurement affect the profitability of proximate users is partially to miss the point. If measurement has a role in market creation, and in opening up markets, that does not necessarily translate into an increased profitability for proximate users. Indeed, the growth of standards, when it facilitates the growth and opening-up of markets, may in fact reduce the profitability of

[28] Hamlet, act II, scene I.

some users by increasing competition. But despite that, the benefits of open markets to end-users should exceed any losses of monopoly rent to some suppliers.

Engineering Economics

"Engineering economics" is a loose term used to define an approach to economic analysis where the form of the economic model of a process - and indeed some of the parameters of that model - are derived from the engineering fundamentals of the process. It is easiest to appreciate this by means of an example. It is possible to write down a reasonably accurate model of the cost per bit of memory in a semiconductor chip, as a function of the total number of bits on the chip. This is derived from the basic physics of lithography and the engineering details of the lithography process. A few cost parameters need to be estimated from accounting data, perhaps, but the rest of the parameters follow from engineering and physical measurements summarising the "state of the art" in lithography. To identify the form of such a function, and the relevant parameters, by direct cost measurement or econometric analysis would be very difficult.

In the present context, the "engineering economics" approach would try to describe some of the cost-benefit relationships from a detailed "engineering" model of the uses to which measurement is put. While the Scientific Generics approach is *direct measurement*, this is *indirect deduction* from a theoretical model.

Econometric Studies

In contrast to the above techniques that are, respectively, direct measurement and indirect deduction, the econometric technique is one of *indirect inference*. In this context, the technique attempts to make inferences about the effects of measurement from correlations between data on measurement and other macro-economic data.

The econometric technique is relevant to the case where a set of variables X are thought to influence Y, but we wish to isolate the effect of one variable in particular - say X_1. The econometric approach seeks a set of data in which there is substantial variation in Y and X (over time, perhaps, or across firms or countries) and then tries to disentangle the effects of different parts of X on Y using multiple regression - or related techniques.

There are a number of obvious difficulties with such an approach. First, the technique can only work if there is some *independent* variation in X_1 - that is, variation in X_1 which is not correlated with other elements in X. In practice, however, it is common to find in much economic data that different elements of X will be inter-correlated - which makes it hard to disentangle the different effects. Second, in many economic contexts, Y may influence X just as much as X affects Y, though recent advances in the subject have helped to unpick the direction of causation. And of course, correlation need not imply causation.

Despite these problems, econometric analysis is still useful because it does not suffer from some of the problems inherent in direct measurement. First, with the appropriate technique, it should be possible to establish the full economic effect of X_1 on Y. The incomplete nature of estimates based on direct measurement need not be a problem here. Moreover, there is no requirement that survey respondents understand subtle questions or subtle effects: the technique can identify effects that may be hidden to the eye. Finally, moreover, if the appropriate data are to hand, the technique offers a relatively inexpensive empirical method.

Case Studies

While the preceding techniques use direct measurement, indirect deduction, or indirect inference, the case study is really an art rather than a well-codified scientific method. The most advanced case study methodology is the *ethnographic* technique - developed in anthropology. This involves the researcher in participant observation over a long period (often years). The researcher tries to build up a detailed picture of how different social and economic phenomena interact. The technique has been applied by sociologists studying the social and economic effects (and the "social construction") of new technologies. It has also been used in the study of "laboratory life". I do not know of any studies that have applied this technique to measurement per se. The technique is very costly, and is sometimes criticised for telling us only about one case, which may not generalise to other cases, and also for being highly subjective.

However, those immersed in this case study tradition usually take a sceptical view of the extent to which the techniques summarised above are really "objective". Moreover, they would argue that since social and economic behaviour is several orders of magnitude more complex and subtle than most physical phenomena, then techniques borrowed from physical and natural sciences cannot give useful measures of social and economic values. Whatever one thinks of this position, it seems indisputable that a few good case studies are always a vital complement to quantitative studies, because the former open up subtleties and complexities that are lost to the latter.

* * * *

Table 3 summarises the advantages and disadvantages of these four generic approaches. Each has is strengths, but each also has some weaknesses. There is no one *best technique*: the appropriate tool for a piece on applied work depends on the setting, and often these different techniques complement each other in a valuable and constructive way.

Table 3
Methods of Applied Economics

	Advantages	Disadvantages
Case Studies (unstructured interviews)	Good at unravelling complex mechanisms, and gaining understanding of subtle effects Easily understood by the user	Hard to generalise from a small sample of cases cost per case is high
Direct Measurement (structured questionnaire)	good at quantifying specific effects in a large sample direct measurement avoids some of the problems of indirect inference	respondents may not understand questions respondents may not recognise that "X has an important effect on Y"
"Engineering Economics " Approaches	Avoids problems of respondent misunderstandings "an objective model of what effect X has on Y"	Models may be (usually are) too mechanistic models may not tally with user perceptions of the underlying process
Econometric Analysis (Indirect Inference)	Avoids problems of user misunderstanding and assumed theoretical models "let the data talk" relatively inexpensive to obtain a macro view	Many problems with indirect inference e.g. : correlation does not imply causation hard to unravel different effects of many variable factors

4.2 Applicability of these Techniques

Which of these techniques are applicable here? The short answer is that all four may be relevant. But a fuller answer is that none is complete on its own, but each has something distinctive to offer.

Direct Measurement: The work by Scientific Generics has shown what can be achieved by direct measurement. They have been able to make an estimate of the value of different measurement activities for a variety of proximate users. But this approach is less good at identifying externalities. Indeed, it is arguable that direct measurement on its own will always have severe difficulties in locating externalities, still less quantifying them. One needs to know who the beneficiaries are before one can question them, and indeed some beneficiaries may not actually realise the extent to which they benefit from externalities. Some of those who do, moreover, may not necessarily be willing to discuss the fact!

***Econometrics*:** The merit of econometric work (as described in my report to LGC) is that it offer ready quantification, and also that in principle it can capture externalities from investment as well as the direct internal benefits to those who pay for it. This approach has generated some useful results in the context of standards (e.g. Swann et al, 1996). We have discussed whether a similar methodology might be applied in the context of the NMS. It is unclear whether it would be easy to obtain such good data on measurement activities. But in any case, the assumptions that would have to be made to complete such an analysis would probably render it less interesting for the present purpose. In particular, the assumption would have to be that a 'typical' measurement activity has a given 'typical' effect. However, for the purpose of reviewing priorities in the NMS, this is not very helpful. After all, the goal is to distinguish between the *different effects* of different measurement activities.

***Case Studies*:** The case study methodology has the merit of being very good at identifying, describing and perhaps even quantifying the mechanisms by which measurement investments may generate externalities, and the identity of the beneficiaries. The shortcoming is that cases are costly, per case, and so one is usually left with the problem of making inferences from a very small sample. However, it seems to me that a case study element to the NMS review would be valuable. It would seem to be a good idea to identify a small number of measurement projects, and for each of these to explore at length how they work, by what mechanisms they generate externalities, and who benefits. The merit of even a small number of these would be to make sure that the team has a good understanding of the economics of at least a few specific projects. Perhaps, indeed, a few preliminary cases exist already that could be extended?

***Engineering Economics*:** The engineering economics approach has something in common with the case study but instead of expensive data collection it summarises the mechanisms by which measurement activities impact on the economy by means of a series of models, and then deduces the sorts of externalities that arise by reference to these models. The reliability of any quantitative indicators derived from these depends on the accuracy of some key parameters, and that may not be easy to establish here. But as a qualitative indicator this approach is very useful, and allows us to go further than a few case studies - for fixed resources.

We shall see in Section 4.6 how an ideal strategy would combine these for approaches. But first, we need to reflect on the particular character of the empirical problem in hand.

4.3 Measuring Externalities: Why is it difficult?

In an often-cited paper on the search for spillovers, one of the great applied econometricians, Zvi Griliches (1992) of Harvard University, concluded that spillovers were often very important, but were very hard to quantify.

It is worth reflecting on this observation. For simplicity, take the case of positive spillovers - that is, where one person's investment creates benefits for others who are not party to the investment. An economist in the Chicago (i.e. ultra free-market) tradition might argue as follows:

> "There is a strong incentive for business-people to internalise positive spillovers. Either it is easy to do this, or it is difficult. If it is easy, then we can rely on business people to internalise. Then when this has been done, there will be no externalities left. If it is difficult, because business-people can't track down all the externalities, then how can the researcher expect to do so? In short, the conclusion would be that externalities are either unimportant or impossible to identify."

Actually, this answer is incomplete because it fails to make an important distinction between: (a) the ability to track down where the spillovers are accruing; and (b) the ability to contrive a contractual arrangement to internalise these spillovers. If business-people are to internalise spillovers they must have (a) *and* (b). The researcher, by contrast, only needs to have (a).

A simple example will illustrate the point. Suppose that a developer buys a derelict house in an otherwise well-to-do street, and improves that house. It is likely that there will be positive externalities from this to immediate neighbours, and perhaps all the way along the street. The developer knows where the beneficiaries are. He can knock on their doors, indeed! But it is rare to see such developers contract with neighbours to internalise these externalities. The issue is not one of locating the beneficiaries. Nor is it because of the impossibility of measurement. For here, a surveyor could give a good idea of the amenity value of having a house in good condition next door, rather than a derelict house.[29] No, the difficulty arises because it is difficult for the developer to find any way to make the neighbours pay towards this externally, for if they believe the developer will go ahead anyway, they have no incentive to reveal the value of their amenity windfall.

Our perspective, hence, is that the pessimistic conclusion of the (hypothetical) Chicago economist overstates the case. It *is* possible for important externalities to exist *and* for the researcher to track them down. But to do this the researcher needs to know where to look, and what to look for.

---

[29] Or an economist with sufficient data on house prices and characteristics could use "hedonic" regression analysis to estimate a shadow price for this amenity. A lot of applied econometric work in environmental economics has taken such an approach.

4.4 Where do we start looking?

There is something in common between tracking spillovers and the popular representation of detective work on television and in novels. In what follows, this metaphor is rather useful. In tracking the spread of spillovers, there are three essential and preliminary steps. These may seem obvious, but it does no harm to spell them out. First, we need to be clear exactly *what* is "spilling over". Second, we need to be clear about the mechanisms *by which* and channels *through which* it spills over. Third we need to know *where* these channels lead, and so where we should start to look for spillovers.

Unless we have a clear understanding of these three, we are unlikely to be successful in tracking spillovers. To pursue the detective analogy, we need to know what has gone, how it got out, and what route it might have followed. It is impossible to find something if we do not have an accurate description of the missing item. Equally, one cannot search everywhere, so we have to narrow the area for search. A starting point is to look around the paths leading from the point at which the "spillage" started.

To understand what form externalities take, a natural starting point is to describe the character of *internalities*. The work that NMS has already commissioned, from Scientific Generics for example, makes an important contribution here. But we have to remember that the form in which externalities manifest themselves may not be the same as the form of internalities. The spillovers may take a form that makes them very hard to recognise. Indeed, those who benefit from these spillovers may not even realise they are doing so.

In trying to understand the "what" of spillovers, it would be useful to gather as much information from those who are most closely acquainted with the "what". This will not just be users: it will also be producers. To what extent have those involved in the production side of measurement codified what forms they think the benefits may take? Sceptics may say that producers have an incentive to exaggerate spillovers, and that may be right. However, they may be able to give as good a description of the "missing item" as anyone.

To understand the *mechanisms and channels* through which spillovers occur, it is natural to ask whether the spillover emanates from the producer or from the original user. Or indeed, the spillover may emanate from both, or from the very transaction.

The producer may be the direct source of a spillover if the results of an investment can be copied directly by other parties. Alternatively, the producer may be an indirect source of spillovers. This could happen, for example, because as is often remarked, intellectual capital is an asset that appreciates with use.[30] If the producer is commissioned to carry out a particular project, this will both draw on and extend the producer's intellectual capital. The extension may create benefits for subsequent clients.

[30] In contrast to most physical assets, which depreciate with use.

The original user may be the source of the spillover. This is the thinking underlying the epidemic model of diffusion, as we shall see in a moment. The original user is, in a sense, a demonstrator of the new measurement technique. Those who see this user trading profitably on the strength of it will learn from the observation.

Finally, to understand the channels through which a spillover will spread, we need to understand the social and economic processes through which diffusion takes place. Diffusion models can conveniently be grouped into three categories. (Those familiar with the economics of diffusion may wish to jump ahead to the start of Section 4.5).

Epidemic models of Diffusion

Economists use the, “epidemic”, to indicate that there is a certain similarity between the diffusion of a new technology amongst a population of users and the spread of an infectious disease amongst a population of people who do not have resistance. In its simplest possible format, the epidemic model assumes (as with the infectious disease) that the rate of new cases (of adoption) is proportional to the *product* of the number infected (who use the technology) and the number who are not infected, but could be (the potential future adopters). This simple form yields the familiar s-shaped diffusion curve, showing how the number of adopters increases over time. Simple mathematical considerations show that the rate of diffusion is fastest towards the middle of the diffusion process, and in the simplest version of the model the fastest rate of adoption per period is when 50% (exactly) of the target population are 'infected'.

Other more subtle 'epidemic' diffusion models generate slightly different patterns, but the basic story is the same. More complex models recognise that more complex patterns of interaction are required before one user will 'catch' the technology from another. The social process which spreads the use of a technology is taken to be the exchange of information amongst potential users, and the pioneer plays an important *demonstrator* and educational role in this. In more subtle models of this sort, it is not just the *number* of pioneers that count, but also their *quality*. Moreover, the effect that the pioneer has on subsequent adopters is also dependent on his experience with the technology: good experiences promote diffusion, while bad experiences retard diffusion.

In the very long term the level of diffusion approaches (asymptotically) a maximum or *saturation* level of diffusion. This technique is useful for projecting how fast the uptake will be, given a certain saturation level of diffusion, but it does not *in itself* help us to estimate the saturation level. Moreover, speeds of diffusion will vary enormously from one technology to another - compare the diffusion of the car and a particular pop record, for example.

The "Probit" Model of Diffusion (1): Changing Characteristics of the Buyer

Another force leading to diffusion is related to rising income, increasing turnover or other characteristics of the buyer. As income per capita rises gradually over time then consumption of some products will rise. At an early stage, buyers may be limited to those on relatively high incomes, but as the average level of incomes rise, more and more buyers enter the market. The same principles can be applied to the diffusion of new technologies amongst firms. For some expensive new technologies, early buyers may be the larger firms (in a particular industrial sector), but as the smaller firms grow, they too would be potential buyers.

This approach to the analysis of diffusion can also be used with other consumer or firm characteristics. As customers become more educated they are more likely to buy certain products. Likewise, as firms become more concerned to seek new sources of competitive advantage in an ever more competitive market, they may be more likely to buy into some new technologies.

In general, diffusion promoted by growth in incomes, revenues or other buyer characteristics tends to be a slower process than diffusion attributable to other effects. The reason for this is simply that the rate of growth of incomes is typically quite a small percentage rate. GDP per head rises at perhaps 3% per annum on average, and most consumers' incomes rise no faster than that. Large companies do well to grow any faster than that. Some smaller companies of course can grow a good deal faster, but they are outnumbered by companies that don't. And once again, this approach is useful as one of the factors determining the rate at which a technology diffuses, but does not in itself indicate the likely saturation level of diffusion.

The "Probit" Model of Diffusion (2): Changing Characteristics and Price of the Product

In many high technology markets the main driving force behind the diffusion of a new technology is that as it becomes cheaper and/or better, ever more buyers are willing to adopt. Moore's Law is a classic example of this. The law states that the rate of increase in the number of 'gates' per semiconductor chip is doubling every 18 months or so, and also implies an exponential decline in the price per component. As the price per bit falls over time, so ever more users are drawn into the market. The speed of diffusion here depends on the speed of price decline and the variance of the willingness to pay of different buyers. As noted above, these rates of change provide a more rapid impetus to the diffusion of the technology than rising incomes or turnover.

A subtler version of this story gives due account to the role of improvements in performance as well as reductions in price. The usefulness of this more subtle picture is that it can distinguish between the effects of improving quality (at given price) and falling price (at given quality). It can also show how the market is segmented between base quality, superior quality and top quality versions of the technology, and how this segmentation would change over time. The rate of diffusion in this process, again, depends on the speed at which prices fall and quality improves.

Strategic or Game-Theory Approaches to Diffusion

In the approaches described above, a company's or consumer's decision to buy is assumed to be independent of what others decide. However, for both for the individual customer and for the industrial buyer, this need not be so. This observation opens up two important strands of diffusion analysis.

The first arises where the fact that others are already using a technology makes a new buyer more likely to adopt. This could arise where are benefits from adopting a technology that has a large installed base of users. The existence of a large installed base is often a powerful signal to some buyers that this technology is no longer the risky prospect that it appeared at first. This feature makes for a very powerful and rapid diffusion process, where if one technology gets ahead of the field, the market will quite rapidly lock into that technology. This approach is used to model the emergence of *de facto standards* or *dominant designs* in markets.

This feature of rapid take off can also arise where firm B faces competitive pressure to adopt a new technology because their rival firm A is using it with good effect to take market share away from B. While firm B may not be especially interested in adopting the technology on its own merits, they may become interested if firm A exploits the technology for competitive advantage. This can lead to a 'me too' attitude to adoption, and a very rapid rate of diffusion.

This process of diffusion due to competitive pressure, operating in addition to the other processes described above can further accelerate the process of diffusion. One major feature of technological change noted above is that firms tend to find new competitors emerging from new and unexpected directions. That means the technology adoption decisions of an ever-greater number of firms can impinge on the decision faced by one particular firm. This proliferation of competitors can make diffusion due to competitive pressure a very pervasive as well as a very rapid process, and will continue to raise the level at which the market reaches saturation.

So far, we have just described the case where the fact that others already use a technology makes the new buyer *more likely* to adopt. For completeness we should mention that it is well recognised in consumer buying behaviour that consumption may be as much influenced by the behaviour of the *distinction group* from whom the consumer wishes to distinguish himself, as by the behaviour of the *peer group* with which he wishes to associate. When this feature is important, it tends to suggest that there are likely to be clear upper limits to diffusion, covering far less than 100 per cent of the potential population. While these last effects may seem most relevant to the buying behaviour of the final consumer, they do also arise in the context of industrial buying.

* * * *

If we know which of these diffusion processes is most prevalent in a particular case, then we can form a clear idea of where we should start searching for spillovers. In the example of the beneficent developer given at the start, the beneficiaries from spillovers are geographically adjacent. That gives a clear guide to a search strategy. But other diffusion processes may give rather different patterns. We saw in Figure 19 how epidemic diffusion could lead to contiguous groups of adopters, while probit diffusion could result in much more dispersed adoption.[31] But when the character of the diffusion process is well understood, it helps to define a clear search strategy.

All the above suggests that we need to give much attention to the views of those who are closest to the measurement business, if we are to understand the what, how and

[31] Of course, the pattern may appear dispersed or contiguous depending on the definition of the search space.

where of spillovers. Regardless of whether those closest to the process may exaggerate, we need to have their guidance on what to look for, how it got out and where it might have gone. That exercise would then be supplemented by an intensive search along the relevant axes. If nothing is found, then perhaps there is nothing to find. But let us give ourselves the best chance possible of finding something.

4.5 Assessing and Measuring the Externalities

In our search for externalities, it is helpful to index externalities according to three criteria. Then we shall discuss briefly what sort of empirical measurement methodologies might be most relevant.

First, it will clearly be useful to index externalities according to the type of measurement activity from which they emanate. The work by Scientific Generics identifies seven types of NMS activity. In my report to LGC, I used a simpler three-stage classification of measurement activities: **(a)** research; **(b)** developing tools and measurement infrastructure; and **(c)** using tools and techniques. It seems probable that the extent of spillovers will be highest in (a), intermediate in (b) and lowest in (c). In particular, the spillovers from measurement research - and especially basic research - will tend to be greater than the spillovers from the final use of measurement.

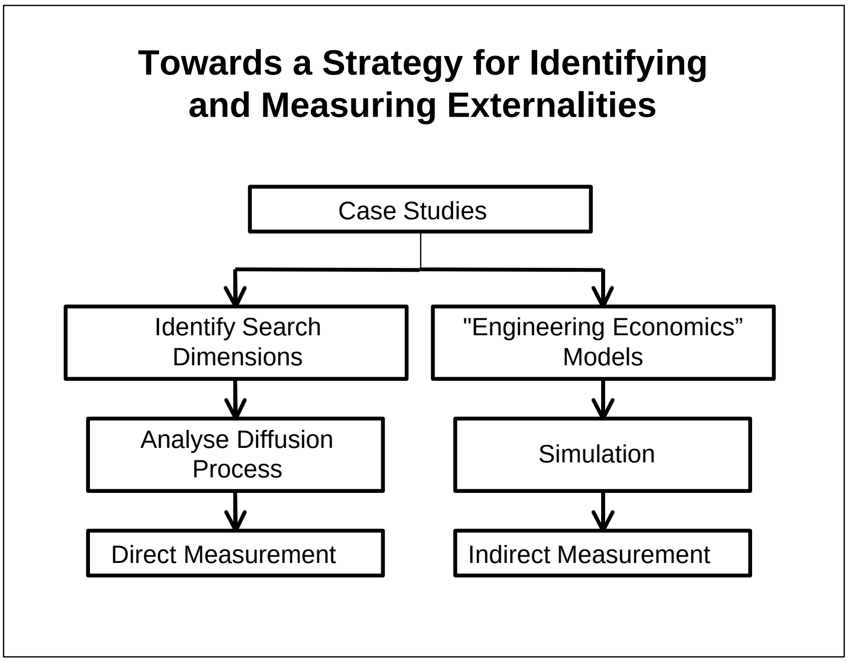
Second, it is useful to index externalities according to the mechanisms by which measurement activities create such externalities. Some of the discussion above explains how new means of measuring characteristics can help to extend markets and even create new markets. The point here is that many who benefit from the creation of a new market may have little idea of their debt to measurement research.

Third, it is useful to index spillovers according to the parties that benefit from them. For example, if these spillovers are mainly of benefit to *incumbents* in the same sector, then it may be possible (as we saw above) to use a “club” solution to achieve high internalisation. On the other hand, if the spillovers accrue to firms in other industries, to *de novo* entrants or cross-entrants, to customers and indeed to final consumers, then the achievable rate of internalisation will be lower.

4.6 Towards a Strategy for Identifying and Measuring Externalities

I suggest that the strategy should consist of two parallel streams of study. These can be described as a route towards direct measurement of externalities and a route towards indirect measurement. The two streams are not, of course, completely independent and things learnt in one will help to guide the other. Moreover, some components in each stream already exist, and relatively little new work on these components may be required.

Figure 20



Detailed case studies of particular measurement projects and their spillovers are a pre-cursor to both streams of work. The aim here is to gain a clear understanding of exactly how measurement activities create economic benefits for a small number of specific projects (perhaps 6). What is the mapping from measurement activities to the characteristics that they help to measure? What are the mechanisms by which measurement is beneficial (e.g. market creation)? These cases would provide the foundations for both streams of subsequent work.

The left-hand stream could be called the “direct measurement” stream, while the right hand stream is “indirect inference”.

Direct Measurement

This stream has three stages:

1. ***Identify the dimensions of search***. As described above this will involve clarifying the “what”, “how” and “where” of spillovers. That will involve in depth interviews with those most closely involved in the measurement business - producers and

users. What sorts of benefits do they think they create? The aim is to form a clear idea of how the spillovers get out, and through what channels they diffuse.

2. ***Clarification of diffusion processes***. As described above, tracking down spillovers requires a good understanding of the details of the diffusion process. We need to know which of the commonly used models of the diffusion process are relevant in each case. When we know what sort(s) of process is (are) at work here, we can project where the spillovers are most likely to be found, how widely dispersed they are, and what are the implications of that for the internalisation rate. That gives us a "road map" for stage 3 of this stream.

3. ***Direct measurement of spillovers***. The earlier work of Scientific Generics has made considerable progress in measuring the internalities. Broadly speaking, the method for measuring externalities is similar: it involves questionnaires/interviews of third party beneficiaries. As stressed above, however, the questioning here may need to be rather subtle, because the form in which externalities manifest themselves may not be the same as the form of internalities. The spillovers may take a form that makes them very hard to recognise. Indeed, those who benefit from these spillovers may not even realise they are doing so. Methods of direct measurement could include: counting characteristics, citations to measurement standards, visits to web sites, and so on.

Indirect Inference

This stream also has three stages:

1. ***Build "engineering economics" models of each of these case study projects***. My earlier report for LGC (Swann, 1997) showed how such models can be constructed. The aim is to draw out some of the broader implications of a particular measurement activity, including some of the indirect effects of spillovers from measurement. When we have an indication of what form these can take, then we can draw up a list of related economic magnitudes that we should seek to measure.

2. ***Carry out simulation with these models.***

3. ***Indirect measurement***. This involves measuring some of the related phenomena which may not immediately indicate the magnitude and location of spillovers, but when interpreted from the perspective of "engineering economics" models can give an indirect indication. For example, these could include a market structure approach to measuring externalities. As indicated above, the rate of internalisation will probably be lower when there is strong *de novo* entry, or cross-entry into a market.

5. Other Issues

Participants at the meeting of the NMS review team on 24th May 1999 suggested some additional points that would be worth discussing, briefly, in the final version of this report. This section collects these comments.

5.1 International Dimension

The discussion above refers to a measurement system for a single economy. It says very little about how the measurement system(s) for several economies should be arranged. To what extent can economies of scale be realised by running an integrated measurement system for a collection of countries (e.g. for the EU)? To what extent does trade globalisation mean that different national systems must be co-ordinated?

Economies of Scale

We touched on the first question in section 2.1 and footnote 2. We argued that in principle there probably are strong economies of scale in basic research on measurement, but economies of scale are less important when we consider the development and especially the use of new measurement tools. However, just because there is a possible economy of scale in basic research programmes does not mean that this scale economy can be realised when different national research programmes are merged. First, there is the issue of whether two merging institutions with different organisational cultures and traditions will be able to exploit scale economies. Much of the economic evidence about the experience of mergers finds that while such mergers may be in the interest of some stakeholders, crude performance calculations suggest that economy-of-scale effects are limited. In addition there is the problem of "musical chairs": when rationalisation is mooted, each party will understandably work hard to ensure that (s)he is not one of those to be "rationalised". This can make it very hard to realise any potential scale economies.

Second, there is the issue of how significant any scale economies will actually be in proportion to costs as a whole. Consider two polar scenarios. In the *first* scenario, the full output of basic research on measurement can all be codified (written down) and disseminated costlessly via a web page, so that any user could download the results and derive value from them. In the *second* scenario, by contrast, the output of basic research is highly tacit, cannot all be codified, and is often embodied in researchers and their tools. For that reason, technology transfer and dissemination require frequent interaction with intermediaries who can translate the results of research to make them relevant in a particular local or national context.

In the first scenario, the cost of doing the research is the dominant cost in this scenario, and the cost of dissemination and technology transfer is minimal. In such a scenario, any strategy that cuts the total costs of basic research will be very important. In the second scenario, however, the cost of doing the research itself is only a part of the total cost, and the cost of dissemination may be as important or more so. In this second case, a strategy that cuts the total costs of basic research will be less important, and if that strategy also increases the costs of dissemination, then it would be counter-productive.

Globalisation and National Systems?

The second issue is whether increasing globalisation means that national systems become less relevant unless they are fully compatible with international systems? And if national systems are fully compatible, what distinctive role do they perform? This question is perhaps more easily understood in the context of national and international standards institutions, so let us start with that.

If globalisation means that standards must be international, and if international (or European) standards bodies can supply an increasing number of international standards, then does that reduce the role of the national standards body? Some may think the answer is obviously, "yes". Actually, this is plainly not the case for *leading* national standards bodies. The reason is that, as many in the standards community have put it, international standards often offer only a lowest common denominator standard. The best national standards bodies will offer standards that exceed this lowest common denominator. The results in Swann et al (1996), for example, found that *distinctive* national standards offered by BSI (that is, standards which were not just copies of the corresponding international or European standard) were more influential in enhancing export performance than *compatible* national standards.

It seems reasonable to apply this same logic to national measurement systems. Certainly, a national measurement system that does not match the lowest common denominator of international measurement is an anachronism. But, especially in the second scenario sketched out in earlier paragraphs (where measurement research and technology is largely tacit), a national measurement system that exceeds this lowest common denominator can offer a distinctive advantage to business in that country.

5.2 Leader or Follower?

The last observation naturally leads into a second question. Should a national measurement system seek to be a leader or a follower? Or, if (as in scenario one above) a "follower" national measurement system performs no significant economic function, we could rephrase the question to ask, do we need a leading national measurement system?

This question is really just a variant on the very general question about whether it is preferable to be a leader or a "fast second". Leaders face higher costs and often face high risks. However, they do have an advantage of lead-time. Clearly, some companies relish their position as leaders, and have no wish to be seconds – not even "fast seconds". Equally, a country's national system of innovation is predicated on a particular place in the competitive cycle; some simply *have* to be leaders while others should not aspire to be leaders.

Should the NMS aim to be a leading national measurement system? If the UK's national system of innovation is such that our competitive position depends on having a distinctive and leading national measurement system, then the answer must be "yes". If, by contrast, our national system of innovation does not require this pre-eminence in measurement, then "no". I realise that this argument replaces one question with another equally difficult one, but I think it makes it plain how questions about the National Measurement System require answers about the economy as a whole.

5.3 Measurement, Standards and Innovation

Participants in the meeting of 24th May asked me to restate and highlight my views on the role of measurement in promoting innovation. My views are these.

The discussion of common pools in section 2.3 emphasised that measurement capability can be important to enable innovative firms to demonstrate that their products and services are indeed superior to the competition. In this sense, measurement capability can help to make innovation worthwhile: you can demonstrate superiority. Without that capacity to demonstrate superiority the incentive for innovation may fade.

However, the value of measurement in promoting innovation is not limited to that. I argued in section 2.1 that measurement plays a role in promoting exchange (as above) but also a role within the company. In discussing the concept of a virtuous circle in section 2.1, I gave an example of where the results of measurement could suggest an interesting hypothesis, and hence lead to further innovation. More generally, measurement plays a central role in a company's innovation process.

I would make a distinction between the role of measurement for innovation, and the role of standards more broadly. The effects of standards on innovation are more complex. As argued in the review of literature on standards (section 2.2), compatibility standards can enhance innovation in network technologies. By this, we mean that compatibility standards by providing opportunities for "one component" companies to supply components of little value in themselves but which are valuable when they can be used in conjunction with a network of existing products and services. On the other hand, markets (and standards institutions) can get locked into inefficient standards that can in some circumstances retard innovation. Equally, minimum quality standards can support innovation if they help to overcome Gresham's Law – that is, they stop bad driving out good. But minimum quality standards set at an unrealistically high level act as a barrier to entry, and that can deter innovative entrants.

5.4 NMS and Technology Transfer

It was argued at the meeting on 24th May that my report understates the role of the NMS as a technology transfer agency. This is a fair comment. To some extent section 5.1 has addressed this point. Measurement Institutions don't just do research; they help to embody the results of that research into people and tools in a way that the results can be used by those who do not fully understand the basic research.

I have said little about technology transfer in my report because I believe that the economics of measurement and standards, in their purest form, is different from the economics of technology transfer. But I fully accept that the NMS is involved in all of these – from measurement research to technology transfer.

5.5 Metrics for Centres of Excellence

At the meeting on 24th May, there was some discussion of appropriate metrics for assessing the performance of technology transfer. I mentioned that in seeking such metrics, the NMS Review needs to keep in mind an observation made by Paul Willman.[32] There are two stages to technology transfer: (1) between the science base and the "science friendly" part of the company (or other using organisation); and (2) between the science-friendly part and the core of the company. When technology transfer falls short of what it might be, critics tend to assume that (1) is the weak link, and focus attention on improved dissemination from the science base. In fact, Willman argues, stage (1) is often done well because both parties speak the same language. Stage (2) is usually the more difficult and the less successful. Moreover, there isn't so much that the science base can do to promote stage 2, because it is internal to the company.

In the present context, this message is important. Any metric based on success in transferring technology from the science base to the core may do injustice to those who have been very successful at stage (1), but have had ill fortune in talking to companies that are bad at stage (2). Equally, some who are only mediocre at stage (1) will be flattered because they have the good luck to pick companies that are good at stage (2).

5.6 Scope of the NMS: Less or More?

Should the NMS seek to do a smaller number of things really well, or should it seek to cover all the bases – accepting that this may compromise quality? There is no simple answer to this. What is essential here is to understand some of the different arguments that point in different directions. We shall consider three perspectives.

Economists have been giving much more attention in recent years to the role of increasing returns, and this is particularly conspicuous in the discussion of clusters. From this first perspective, the policy maker should not spread his budget too thinly. It is better to concentrate on a few programmes and do them well. If the policy maker tries to achieve too many "centres of excellence" the risk is that (s)he ends up with none – because all of them fall below critical mass.

---

[32] This argument is summarised in a Memorandum by London Business School in *Report of the House of Commons Science and Technology Committee*, Volume III, pp. 1-10, London: HMSO (1994).

By contrast, two other perspectives stress the importance of having all the elements in place before a policy measure will work. Perspective two is the argument in section 2.3 (around Figure 12) which suggests that the economic benefits of measurement only take-off when most characteristics can be measured. From this perspective, it is not enough to measure a few things very well: it is essential to measure most things. This perspective would tend to imply that the priority for NMS is to offer a basic standard coverage of all dimensions of measurement, and perhaps leave it to the private sector to offer premium services.

A third perspective suggests a similar conclusion. Suppose that the policy maker has a number of levers that (s)he can pull. He knows very little about the effective of each lever except that:

a) each lever has a non-negative effect
b) at least one lever has a positive effect

It can be shown that the maxi-min strategy for a policy-maker faced with this uncertainty is to spend equally on pulling each lever.[33] In short, the policy-maker should do a bit of everything.

The reader may feel perplexed. Understandably so, because the preceding paragraphs have such very different implications for policy. But the main message has to be this: beware of taking one of these perspectives uncritically, without considering the others.

5.7 "The Customer"

A final request was that I should be a little more specific about who and what I mean by "the customer" in this report. But probably the main message from this report is that it is very hard to say exactly who is *the* customer for measurement, and who is *not*. The whole essence of the discussion about externalities is that many may benefit from measurement who don't pay for it, and many do not even realise that they are beneficiaries. And it is precisely because there is no narrowly defined customer that we have market failure, and hence there is a case for the NMS – as summarised in Part 3.

[33] The proof of this is set out in an (as yet) unpublished note by this author. The basic argument is this: if the policy maker wants to be sure of maximising the return to policy intervention in the most pessimistic of circumstances, then (s)he should spend a little on everything. Otherwise, there is a risk that (s)he spends nothing on the one lever that actually works.

6. Conclusions

This paper has examined the basic economics of measurement and the basic rational for public funding of some measurement activities.

Part 2 examined the economics of measurement from three perspectives: a functional and sectoral approach which describes where measurement happens; an economic map which describes the economic role of measurement activities; and a "common pool" or "virtuous circle" perspective. The last two provided a rationale for public funding of measurement activity.

Part 3 set out the basic economic rationale for public funding of some measurement activities. We argued that the most important rationale stems from the likely market failure in private provision of measurement infrastructure. However, we also argued that there may be a "strategic" rationale for public funding in this area. All three of the generic causes of market failure (externalities, scale economies and asymmetric information) may be relevant here, but we argued that externalities are the most important issue requiring attention in the NMS review.

Part 4, drawing on a description of the problem and the techniques available for empirical analysis, suggested an agenda for empirical research in the NMS review. We saw why the measurement of externalities is so difficult, and then proposed a methodological approach which combines direct measurement, case studies, engineering economics and (if possible) some econometric research.

Part 5, an addition to the first full draft filed on 17th May, answered some specific points and questions raised at the NMS Review meeting on 24th May 1999, and in subsequent correspondence.

The perspectives, rationale and methodology set out here require continuing discussion with the NMS Review team. But we believe that this report can provide a useful foundation for the empirical work required by the NMS Review team.

References

Akerlof, G. (1970) "The Market for *Lemons*", *Quarterly Journal of Economics*, 84, 488-500.

Bacharach, M. (1991) "Commodities, Language and Desire", *Journal of Philosophy*, 87(7), 346-368.

Barber, J. M. (1987) *The Role of Government in the Provision of Measurement Standards*, Department of Trade and Industry, March

Buchanan, J. A. (1965) "An economic theory of clubs", *Economica* 32, 1-14.

Callon M., P. Laredo, V. Rabeharisoa, T. Gonard and T. Leray (1992) "The Management and Evaluation of Technological Programmes and the Dynamics of Techno-Economic Networks: The Case of AFME", *Research Policy*, 21, 215-36.

Coase, R. (1960) "The problem of social cost", *Journal of Law and Economics* 3, 1-44.

Coase, R. (1974) "The lighthouse in economics", *Journal of Law and Economics* 17(2), 357-376.

Cunningham M., J. Evans and L. Georghiou (1991) *Benefits to UK Industry of the National Physical Laboratory*, PREST, University of Manchester, May

Eckschlager K. and V. Stepánek (1985) *Analytical Measurement and Information: Advances in the Information Theoretic approach to Chemical Analysis*, Letchworth, Herts: Research Studies Press

Evans, J., M. Cunningham, J.S. Metcalfe and L. Georghiou (1991) *European Aspects of the National Measurement System*, PREST, University of Manchester, July

Griliches Z. (1992) "The Search for R&D Spillovers", *Scandinavian Journal of Economics*, 94(S), 29-47

House of Commons Science and Technology Committee (1994), Report, Volume III, pp. 1-10, London: HMSO

Jaffe A.B. (1989) "Real Effects of Academic Research" *American Economic Review*, 79, 957-970.

Jaffe A.B., M. Trajtenberg and R. Henderson (1989) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations" *Quarterly Journal of Economics* 108, 577-598.

Koestler, A. (1969) *The Act of Creation* London : Hutchinson

Lancaster, K.J. (1971) *Consumer Demand: A New Approach*, New York: Columbia University Press

Martin B. et al (1996) *The Relationship Between Publicly Funded Basic Research and Economic Performance*, SPRU, University of Sussex: Report for HM Treasury, May

Rabinovich, S. (1993), Measurement Errors: Theory and Practice, New York: American Institute of Physics

Simon, H. A. (1985) "What do we Know About the Creative Process?", in R.L. Kuhn (Ed.) *Frontiers in Creative and Innovative Management*, 3-20, Cambridge, MA: Ballinger

Swann, G.M.P. (1990) "Product Competition and the Dimensions of Product Space" *International Journal of Industrial Organisation*, 8(2), 281-295.

Swann, G.M.P. (1996) *The Economic Value of Publicly Funded Basic Research*, Report for Technology and Innovation Division, DTI, March

Swann, G.M.P. (1997) *The Economic Benefits of Chemical Measurement*, Report for Laboratory of the Government Chemist, December

Swann, G.M.P., P. Temple and M. Shurmer (1996) "Standards and Trade Performance: The British Experience', *Economic Journal*, 106, 1297-1313.

Swann, G.M.P., N. Pandit and T. Watts (1997) *Case Study of VR Application: MIST$_{VR}$*, http://www.ukvrforum.org.uk/CaseStudies/mist.htm

Tiebout, C. 1956. A pure theory of local expenditures. *Journal of Political Economy* 64, 416-424.