Models of Public Sector Information Provision via Trading Funds

Professor David Newbery
Faculty of Economics
Cambridge University

Professor Lionel Bently
CIPIL, Faculty of Law
Cambridge University

Rufus Pollock
Mead Fellow in Economics
Emmanuel College, Cambridge University

February 26, 2008

Commissioners

This study was commissioned jointly by the Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury in July 2007.

Acknowledgments

The authors would like to thank Christopher Whitcombe for excellent research assistance which made a major contribution both to the quality, and timely completion, of the report.

Copying and Redistribution

Contents

Chapter 1

Executive Summary

This study has analyzed the impact of adopting different models for the provision of public sector information by trading funds. Its basic task has been to examine the cost and benefits for society, and the effects on government revenue, of four different charging policies: profit-maximization, average cost (cost-recovery), marginal cost and zero cost; both on their own and when interacted with various data distinctions such as raw versus value-added, and unrefined versus refined.

The study focused on the six largest trading funds by data provision: the Met Office, Ordnance Survey, the UK Hydrographic Office, the Land Registry, Companies House and the Driver Vehicle Licensing Agency. Starting from the general theoretical framework set out in Chapter 3, Chapter 4 reviewed the general empirics, followed by a trading fund by trading fund analysis in Chapter 5. The main results were as follows:

- For many products, limitations of data and/or ambiguities about the nature of the good itself (particularly the data/service divide) would have made the analysis so speculative as to be of little value. In these cases a 'conservative' approach has been adopted of simply assuming such products would be left 'as is', with their pricing and access policies unchanged.

- Nevertheless, for each trading fund, there was a set of products with sufficient information available to allow an adequate analysis. In general, these were digital products of a 'bulk' or 'wholesale' kind, corresponding fairly closely to the 'unrefined' category as defined by the OFT.

- The digital nature of these products has implied marginal costs of approximately zero and hence the marginal and zero costs regimes were essentially identical. Furthermore, limitations of data mean that it has been infeasible to reliably estimate the effects of moving to profit maximization. This is not however necessarily a great deficiency for reasons discussed in greater detail in Section 5.1.4. Specifically (a) these are unrefined products for which it would therefore be difficult to apply a profit-maximization regime without raising a host of serious competition issues (b) if moving to marginal cost from average cost increased social welfare then this implies than moving from profit-maximization to marginal cost would raise social welfare.

- Thus the analysis has generally been confined to comparing the existing average cost (cost-recovery) regime with marginal cost.

- **Performing this comparison on the subset of products suitable for analysis, it was found that, in most cases, a marginal cost regime would be welfare improving** – that is, the benefits to society of moving to a marginal cost regime outweighed the costs. Further details of which products would be affected, and what the benefits and costs would be, are provided by the trading fund summaries in the Conclusion (chapter 7).

- For registration based trading funds (DVLA, Companies House and the Land Registry) it likely that this change in charging policy could be made without the need for government to provide additional funds as any shortfall could be made up from the registration side of their activities. For the other trading

funds some direct assistance, beyond that already provided, would be required. In the case of the UKHO and the Met Office the sums involved would be limited (around £1m) but in the case of Ordnance Survey would be substantially larger (though the benefits in this case would be commensurably bigger).

- A change in charging regime **should not** have a detrimental impact on the performance of trading funds in terms of efficiency or data quality, providing a suitable governance and regulatory regime is put in place (and this is desirable in any case). Trading fund performance is primarily determined by the governance/regulatory structure under which it operates and this structure should (and can) be chosen independently of charging policy. Chapter 6 considers these questions in detail and includes discussion of ways to strengthen the existing regulatory structure so as to minimize risk of adverse consequences from any changes.

- It is important to emphasize that having an adequate governance/regulatory regime in place is absolutely central to realizing the potential benefits from change (and also for delivering value for money even under the present charging arrangements). Thus, getting this right should be one of the first items for consideration whether or not any restructuring does take place (and will be essential if additional subsidies are required under a move to marginal cost pricing). The substantial experience with regulation both in the UK and abroad in recent years should assist greatly in performing this task.

- To sum up, socially optimal policy would involve (a) leaving the charging regime for many (probably most) products unchanged (b) moving to marginal (zero[1]) cost charging for a subset of products, roughly approximating to the bulk 'unrefined' digital category. Note that the items in the first category

[1]For these digital products marginal cost is approximately zero.

would primarily be refined products built on unrefined data. These would now be in commercial competition with other suppliers – as such suppliers would now have access to unrefined data at marginal cost. Thus one added benefit of adopting the marginal cost pricing scheme suggested by the analysis is that it would immediately address the competition concerns raised by the OFT as, a fortiori, outside organizations would now have access to 'unrefined' ('upstream') data on the same terms as the trading fund itself.

Chapter 2

Introduction and Background

2.1 Introduction

This study was commissioned by HM Treasury in July 2007. It aims to inform Government decisions on the charging policies of Trading Funds that obtain some portion of their revenues through the sale of Public Sector Information (PSI). The study focuses on an economic analysis of existing and alternative models for public sector information provision by Trading Funds.

A study of this kind has been recommended in several recent Government and independent publications on PSI charging policies, including HM Treasury's Cross Cutting Review of the Knowledge Economy (2000), the Government Response (2007) to the OFT's study on the Commercial Use of Public Information (CUPI), and the Power of Information review (2007).[1]

Currently, in accordance with recommendations[2] in the Cross Cutting Review, raw[3] information from across government is priced at marginal cost by default.

[1] See, Treasury (2000), paragraph 5.11; Office of Fair Trading (2006), The Government Response to the Office of Fair Trading Study (2007), paragraph 11; and The Power of Information (2007), Recommendation 9.

[2] Cross Cutting Review of the Knowledge Economy (2000), paragraph 5.20.

[3] See Section 2.3 for more on the distinctions between unrefined/raw data and refined/value added data.

However, Trading Funds have greater flexibility in how they charge for PSI and are not required to adhere to marginal cost pricing. Both the OFT's CUPI study and the Power of Information review recommended that Trading Funds' charging policies should be reformed. The government responded that any such reforms would be made on the basis of careful analysis of the costs and benefits of various charging models to producers, consumers and the wider information market. It is to these costs and benefits that the present study attends.

2.2 Trading funds

A Trading Fund is an operation of a government department that has been established by a Trading Fund Order in accordance with the Government Trading Funds Act 1973 (as amended by the Government Trading Act 1990). A Trading Fund may be established where a Minister of the Crown judges that the revenue of an operation could "consist principally of receipts in respect of goods or services provided in the course of the operations in question", and that setting one up would lead to "improved efficiency and effectiveness of the management of those operations". Trading Funds are required by statute to recover principally their costs (i.e. to recover a majority of their costs) through income derived from operations within the trading fund.

As mentioned above, the Cross Cutting Review excluded Trading Funds from its recommendation that raw PSI is priced at marginal cost.[4] The prices that Trading Funds charge for PSI reflect the cost recovery requirements set by statute. Table 2.1 lists those operations which currently have Trading Fund status according to the official Trading Fund Orders.[5]

[4]Though given the definition of 'raw' information this exclusion might have had little effect since most trading fund data would not have been considered raw.

[5]This list is derived from Statutory Instruments made available by the Office of Public Sector Information, and differs from the list provided in Appendix D of the Cross Cutting Review of the Knowledge Economy (2000).

Trading Fund
ABRO
Central Office of Information
Companies House
Defence Aviation Repair Agency
Defence Science and Technology Laboratory
Driver and Vehicle Testing Agency
Driver Vehicle and Licensing Agency
Driving Standards Agency
Fire Service College
HM Land Registry
Medicines and Healthcare Products Regulatory Agency
Meteorological Office
OGCbuying.solutions
Ordnance Survey
Patent Office
Queen Elizabeth II Conference Centre
Registers of Scotland
Royal Mint
UK Hydrographic Office
Vehicle and Operator Services Agency

Table 2.1: List of Trading Funds

This study restricts itself to the six largest Trading Funds by revenue from information provision. These are, in order of sales of information: Ordnance Survey, the Met Office, the UK Hydrographic Office, HM Land Registry, the DVLA and Companies House. One reason for this restriction in scope is that many of the trading funds listed above do not derive income from the sale of PSI, or PSI based goods or services. For example, the Defence Science and Technology Laboratory does not derive any its income through the sale of information. The Trading Funds chosen account for over 70% of the estimated total annual income from UK Public Sector Information Holders (PSIHs).

Ordnance Survey is Great Britain's national mapping agency. In 2006/2007 it had an income of around £110m from the supply of information, which accounts for over a quarter of the estimated total income of £390m from the supply of information of UK PSIHs.[6] It produces and sells large quantities of Geographical Information (GI) and a variety of derivative value-added products and services in addition to high resolution paper and digital maps.

The Met Office provides information and services related to the weather and the environment. It had a total income last year of around £170m of which £90m comes directly from the sale of information and related services.

The UK Hydrographic Office is part of the Ministry of Defence and a major provider of navigational products and services. It had a total income of around £80m in 2006/2007.

HM Land Registry is responsible for recording and maintaining the register of property dealings in England and Wales. While its total income in 2006/2007 was one of the largest of any trading fund only a small part of that came from the provision of information.

The DVLA is responsible for vehicle registration and the collection and enforce-

[6]The Commercial Use of Public Information, Annexe A: Survey of Public Information Holders, paragraph 3.6, p.12.

ment of Vehicle Excise Duty. Like HM Land Registry their revenues in 2006/2007 were among the largest of any trading fund but again, only very small proportion of these revenues derived from the sale of data. Finally, Companies House is the official government register for UK companies. In 2006/2007 their total income was around £72m, the majority of which came from registration and other activities rather than the sale of data.

2.3 Data Distinctions

In discussion of PSI charging policy various different data distinctions have been proposed. In this section these different distinctions, and their origins, are briefly outlined. The two main sets are:

1. Raw versus Value-Added.

2. Unrefined versus Refined.

2.3.1 Raw versus Value-Added

These distinctions were first proposed in the Treasury (2000) and are formally set out on p. 14 of *Charges for Information: When and How* (Treasury, 2001) as follows:

> Raw data (or Crown copyright "Material", with a capital M, as the HMSO Class License puts it) was defined in the Review of Government Information as "information collected, created, or commissioned within Government which is central to Government's core responsibilities. The supply of selected components of a raw data package, exactly as in the package is raw data supply, but the supply with further analysis, summarisation etc, or of data at a different level of aggregation to that used by Government, is not raw data for the purposes of this report but is

> value-added information." See also the supplementary note in Annex 5 of this guidance note. Raw data is not synonymous with raw material, or with unchecked data. For example, the raw material for value-added services may, or may not, be raw data.
>
> ...
>
> Value-added information was defined in the Review as "information where value is added to raw data enhancing and facilitating its use and effectiveness for the user, for example through further manipulation, compilation and summarisation into a more convenient form for the end-user, editing and/or further analysis and interpretation, or commentary beyond that required for policy formulation by the relevant government department with policy responsibility. It also includes supplying retrieval software, or where work on material is included as part of the compilation of related data, and where there is not necessarily a statutory or operational requirement for Government to produce the material.

To summarize: *raw* data or material comprises information which is central to Government's core responsibilities and which is in its most basic state (that is without any additional manipulation, analysis etc), while *value-added* information is "information where value is added to raw data enhancing and facilitating its use and effectiveness for the user, for example through further manipulation, compilation and summarisation into a more convenient form for the end-user, editing and/or further analysis and interpretation."

2.3.2 Unrefined versus Refined

The definitions of 'unrefined' and 'refined' come from the Office of Fair Trading (2006, p.5, para 1.5):

> PSIHs are usually the only source of the basic information they hold. There are good reasons why this is the case, such as: high fixed collection costs, government funding for collection and privileged access, perhaps through statutory collection powers. We refer to this basic information, which cannot be substituted directly from other sources, as unrefined information. Once a PSIH does something with the unrefined information which could also be performed by another organisation, such as a private business, if it were given access to that unrefined information, it becomes refined information.

That is, *unrefined* information is information which cannot be substituted directly from other sources while *refined* information is information which could be provided by another organisation should it have access to the underlying unrefined information. Thus, refined information supplied by a trading fund can be seen as being, at least potentially, in competition with information from other suppliers. By contrast, for unrefined information the trading fund is the sole source and faces no competition, actual or potential, in its supply.

2.4 Charging Policies

The four basic data charging policies that will be considered by this review are:

1. Profit-maximizing: setting a price to maximize profit given the demand faced by the trading fund.[7] Where the product being supplied does not face competition then this will naturally result in monopoly pricing.

[7] One also occasionally hears reference to 'market-based pricing'. It is not entirely clear what this means since several of these pricing strategies involve attention to the structure of the demand curve (that is the price/demand trade-off displayed by the market). However our interpretation is that is intended to indicate that the trading-fund behaves as any other 'normal' market participant would and sets a price to maximize profits given the underlying demand curve.

2. Cost-recovery pricing: setting a price equal to average long-run costs (including, for example, all fixed costs related to data production).[8]

3. Marginal-cost: setting a price equal to the marginal cost of supplying data (that is, simply the cost of actually transmitting the data to someone).

4. Zero-cost: setting a price equal to zero.

The question of reuse and redistribution are also of great importance in evaluating data supply policy but are not explicitly covered by the charging regime alone. We therefore should state clearly that in the first two cases, those of profit-maximizing and cost-recovery pricing, it is our assumption that, for any given product, a trading fund would be at liberty to impose any conditions on reuse and redistribution of its data permitted by the underlying intellectual property rights existing in that material.[9] While in the second two cases, that or marginal-cost and zero-cost pricing, it is our assumption that a trading fund would be making the data 'openly' available so that anyone who acquired data would be free to reuse or redistribute in any way they saw fit.[10] In particular, it should be made clear that the impact of using share-alike licenses[11] will not be considered formally, nor will the possibility of discriminating by fields of endeavour, for example by permitting free (zero-cost or marginal-cost) use of material for non-commercial purposes but

[8]There are a various subtleties as to what exactly cost-recovery entails which are discussed further below.

[9]In the EU these IPRs would primarily sui generis rights in databases provided for under Council Directive 96/9/EC, though there might also be some additional form of copyright protection available both for information that constituted an original database and information that constituted original literary or artistic works under section one of CDPA (1988).

[10]Note that this would not exclude the imposition of conditions entirely. For example, licensors might wish to impose 'integrity' conditions to ensure that the data was clearly marked as only coming indirectly from the original source and therefore potentially no longer having the same authority (such a provision already exists with the PSI 'click-use' license. Licensors might also wish make certain 'public-interest' restrictions. For example, the Land Registry already prohibits usage of its data for unsolicited mail-shots.

[11]Such licenses are popular in open-source and open-knowledge community where material is made freely available for use, reuse and redistribution but with the 'share-alike' proviso that any derivative work is distributed under the same 'open' terms as the original material.

following a cost-recovery or profit-maximizing pricing policy for commercial users.[12]

Finally, while most of these pricing strategies are self-explanatory there are a various subtleties in relation to cost-recovery which merit elucidation. For example, consider the items listed below, any one of which can have a substantial impact on the form of the average-cost curve, and hence on the evaluation of the performance of cost-recovery regime.

- Over what period are investments amortized?
- Should all types of expenditures be included as costs – for example, should the marketing budget be included or money spent by a trading fund on advocacy?
- Does cost-recovery include making a specified return on investment?[13] This leads to further complications as the exact size of the underlying asset base of a trading fund is often unclear – valuing information assets is notoriously problematic due to their heterogeneity and the exact status of a given piece of information from an accounting point of view may be uncertain.[14] This ambiguity obviously gives the trading funds significant leeway in their pricing policy should they wish to use it. For, by changing their assessment of their capital base, they can increase or decrease their target rate of return and thereby subsume any increase or decrease in income under the heading of cost-recovery. A similar effect can also be achieved by a other means such as simple changing internal variable or capital expenditures – well-known issues with cost/capital based regulation. Together these may render the differences

[12]Neither the data, nor the time available would permit us to address these questions in a formal empirical framework. Nevertheless some of these items may be addressed informally in the discussion of regulation in Chapter 6 below.

[13]At present the trading funds are generally expected to make a return on capital employed at or above a minimum threshold that varies between approximately 3.5% (e.g. Companies House and DVLA) and 5.5% (e.g. Ordnance Survey).

[14]Since 1999/2000, the accounts of the Ordnance Survey have been qualified by Comptroller and Auditor General of the Houses of Parliament precisely because they do not share the OS's opinion that the costs of setting up and maintaining the data in the National Geographic Database need not be capitalized.

between cost-recovery and profit-maximizing pricing rather less in practice than they are in theory.

Finally it is important to discuss the impact of these different charging policies on overall trading fund income and government revenues. At present, trading funds are not like ordinary companies in having shares which may be owned by anyone. Instead the UK Government acts as a 100% shareholder and may receive annual 'dividends' from a trading fund depending upon its performance.[15] Hence, any increase or decrease in the revenues of trading funds may have an impact on revenues received by the UK government. Furthermore, both the marginal and zero cost charging policies involve a trading fund in explicitly setting prices which will *not* generate revenues sufficient to cover total costs. Thus, if it is expected that a trading fund will maintain, at least approximately, both the quality and extent of its data production it will be necessary for this shortfall between revenues and costs to be made good from some other source – most likely the Exchequer.[16] In this report this likelihood will be taken as an assumption, that is, where a charging policy results in income being insufficient to cover *production* costs the shortfall will be made up by a subsidy from the UK central government. We shall return to a discussion of the issues raised by this assumption in Chapter 6 below.

[15]To be completely accurate OS, UKHO and Met Office are Executive Agencies, financed through a Trading Fund. They are not companies in any sense, but remain Government bodies and there are no shares. Nevertheless since surpluses are returned to (and shortfalls made up by) Government the basic point remains the same.

[16]The exact extent of this shortfall may be uncertain. For example, a move to marginal cost pricing might make it unnecessary to incur expenditures on marketing or legal expertise that are necessary for a profit-maximizing or even cost-recovery oriented trading fund. It is also possible that a trading fund, particularly one following a average-cost based charging policy, may *over*-provide quality due to their orientation towards revenues rather than profits.

Chapter 3

Theory

3.1 Introduction

The main task of this report is to consider how different charging policies would affect the following 'outcome' variables:

- Consumer surplus
- Producer surplus
- Government Revenue/Expenditure
- Total welfare (taking account of the cost of government funds)

From a policy-makers point of view the last of these, total social welfare, would usually be the most significant since it is an overall measure which incorporates all of the other changes into a single value, usually presented in monetary terms for convenience of comprehension.

3.2 Theoretical Underpinnings

The theoretical underpinnings of the calculations conducted in this report can best be understood by the digram presented in Figure 3.1 which resembles those presented in Appendix C of Treasury (2000). This shows a (linear) demand function along with the marginal and average cost curves for a single good. As illustrated the cost curves correspond to a good displaying constant marginal costs and a non-zero fixed cost of production. In addition to the demand and cost curves also shown is the price (along with the associated output) set by a profit-maximizing firm facing this demand curve. Note that, the particular functional forms and parameters have been chosen simply for illustrative purposes and do not necessarily indicate those that will be used in doing calculations – though, of course, the natural division of costs into fixed and marginal will be retained.

Using this diagram one can easily see the value of each outcome variable under the different regimes. Before examining the impact of each of these regimes let us first formally state the definitions of each of these variables.

First, note that producer surplus equals profits that is revenue minus costs (fixed as well as variable). Turning to consumer surplus, using the partial equilibrium approach adopted here, this will equal the area under the demand curve which is above the price being set. Meanwhile government revenue/expenditure meanwhile is equal to producer surplus (if producer surplus is positive then this will be government revenue while if negative it will a subsidy). Finally total social welfare is the sum of consumer and producer surplus though perhaps with some adjustments to take account of the distributional effects of particular policies (equivalently: the cost of raising tax revenues – this is discussed in greater detail below). In addition to these outcome variables one should also mention deadweight loss which, for a given regime with prices above marginal cost, is the loss in surplus due to setting that higher price.

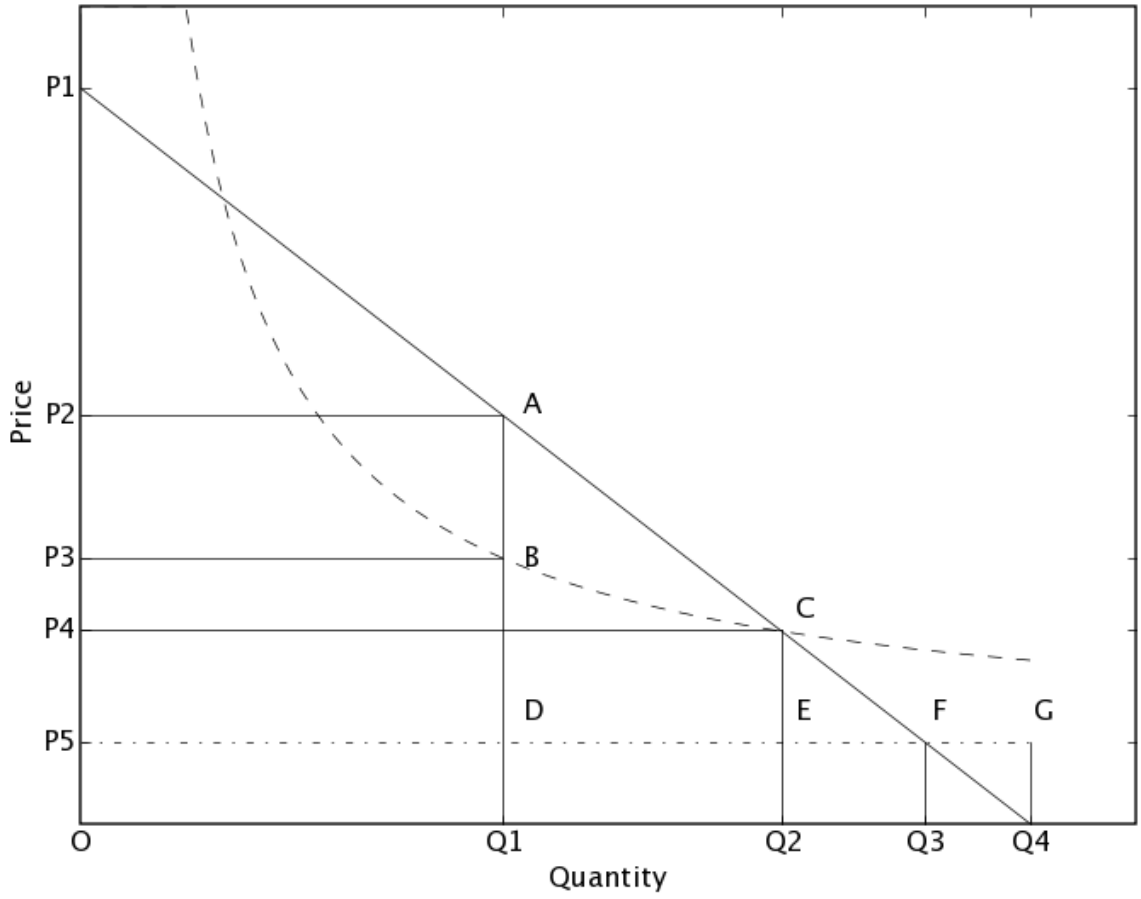

Figure 3.1: Illustrative supply and demand functions using linear demand, for a product with fixed costs and constant marginal costs. Marginal cost (dot-dashed) and average cost curves (dashed) are shown. In addition the prices and output under the 4 regimes of (a) a profit-maximizing: P_2, Q_1 (b) cost-recovery (average-cost): P_4, Q_2 (c) marginal-cost: P_5, Q_3 (d) zero-cost: $0, Q_4$.

	Profit Maximizing	**Average Cost**	**Marginal Cost**	**Zero Cost**
Consumer Surplus	P_1AP_2	P_1P_4C	P_1FP_5	P_1Q_4O
Producer Surplus	P_2ABP_3	0	$-P_4CEP_5$	$-P_4CEP_5$ $-P_5GQ_4O$
Deadweight Loss	ADF	CEF	0	$-FGQ_4$

Table 3.1: Outcomes Under Different Charging Regimes With Reference to Figure 3.1. Government Revenue has been omitted as it is equal to producer surplus. Note that the rectangle P_4CEP_5 exactly equals

To illustrate, Table 3.1 below explicitly relates each outcome variable to a particular area under the demand curve in Figure 3.1 for each of the four possible pricing regimes discussed above.

As already mentioned, the most important value is total social welfare which aggregates all of the others together. Normally, this would simply be the sum of consumer and producer surplus but here one needs to be a little careful because producer surplus can be negative. In that case it is being assumed that any deficit is to be made up by a government subsidy.[1] In this case, it is necessary to take account of the benefits those government funds would otherwise have generated (if they were not being used for the subsidy). The basic approach for performing this kind of cost/benefit analysis is well known. It involves taking uncommitted government funds as the numeraire and then adjusting the surplus from the project under consideration using the appropriate social weights to reflect the different values of public and private costs and benefits. Further discussion of this question, including a formal incorporation into the analysis and an estimate of the appropriate weights to use, can be found in Section 3.4 below.[2]

[1]The only other alternative would be to assume that the good simply is not produced because the Trading Fund expects to make a net loss.

[2]Alternatively one can take money in consumer's pockets as the numeraire. In that case one needs to use the marginal cost of public funds was one (that is, how much does raising one pound of government funds cost general society at the margin).

When considering the theoretical approach to take, the decision must necessarily be informed by the type (and amount) of data that is going to be available. In this case information on prices, sales and costs of a product (or group of products) for a period of a few years will be provided. What variables does such data allow us to identify? Firstly, note that the cost information will be disaggregated into fixed and variable costs. Thus for the given level of sales (output) reported one knows both total fixed costs and marginal costs. Next, note that there will be at least one reported price/output pair.[3] Finally, to locate these points with respect to the diagram presented in Figure 3.1 requires an explicit assumption as to the charging policy currently being followed. Specifically, if a trading fund is pursuing a strategy of profit maximization then the marginal cost value will correspond to point D, the price/output values to A (P_2, Q_1) and fixed costs to P_3BDP_5 (note that as fixed costs are constant this must equal the area of the rectangle P_4CEP_5); while if a trading fund is following an average-cost (cost-recovery) strategy then the marginal cost value will correspond to point E, the price/output values to C (P_4, Q_2) and fixed costs to P_4CEP_5. There is some uncertainty as to exactly which of these two policies is being pursued. Treasury (2000) implicitly assumes that average-cost (cost-recovery) pricing is being used. However conversations with HM Treasury suggested that there is no obligation on trading funds to restrict themselves solely to cost-recovery. Meanwhile discussions with trading funds indicated that they generally were pursuing a strategy of cost-recovery. Nevertheless, on balance, it seems the assumption of average-cost pricing seems the more reasonable. Thus, in what follows, it will be assumed that trading funds follow a cost-recovery strategy with the implications for identification just discussed.

Having identified the 'positive' aspects of what can be learned from the data,

[3]There may be more than one such pair, corresponding to different years, but one would need to be cautious in interpreting any temporal variation since with such large changes in the external technological environment it would be hard to identify whether the source of the variation lay in the demand or supply curve (or both).

	Profit Maximizing versus Average Cost	**Average Cost versus Marginal Cost**	**Marginal Cost versus Zero Cost**
Δ Consumer Surplus	P_2ACP_4	$P_4CEP_5 + CFE$	P_5FQ_4O
Δ Producer Surplus	$-P_2ABP_3$	$-f$	$-P_5GQ_4O$
Δ Total Welfare	$ACB' - \alpha P_2ABP_3$	$CFE - \alpha f$	$-\alpha P_5FQ_4O - FGQ_4$

Table 3.2: Outcome Differentials With Reference to Figure 3.1. α denotes the marginal excess burden (i.e. the marginal cost of public funds minus one). f is the fixed cost of production (equal to the rectangle P_4CEP_5).

what of the 'negative' side, that is the restrictions imposed? The main point of note is that very little information about *demand* will be available – as noted above one may only be able to identify a single point (C) on the demand curve. As a result little or nothing can be said about the overall shape of the demand curve and it is even unlikely that there will be sufficient data to estimate accurately elasticities. This has two consequences. First it means that, where absent, plausible values of such variables will have to be inferred from the existing literature. Second, it will be better to focus on *changes* in outcomes between regimes rather than on the *absolute value* of outcomes. This is because estimating the result of changes between regimes requires less global knowledge of the demand curve than obtaining absolute values.[4]

Specifically, charging regimes will be compared pairwise and in the order given (which is a natural one corresponding to decreasing price and increasing output). To be explicit, a profit maximizing regime will be compared to an average cost regime, an average cost regime to a marginal cost one and finally a marginal cost regime to a zero cost one. Using this approach Table 3.1 becomes Table 3.2. This breaks the differences between regimes with reference once again to the regions shown in

[4]For example, the area under the demand curve between O and Q_1 is common to all regimes. Thus it will be irrelevant to any comparative analysis but will be essential to any estimate of absolute values. This is an area that is 'off-equilibrium' and hence one for which very little information will be available and so the shape of the demand curve in this region will be particularly uncertain.

Figure 3.1. The task then is to use the data available to estimate the size of each of the regions which appear in Table 3.2, focusing particularly on those in total welfare row.[5]

3.3 Additional Considerations

3.3.1 The Demand Curve and Social Welfare

The approach laid out uses the standard partial equilibrium approach of equating areas under the demand curve with social surplus. But one needs to ask here whether, in this case, demand accurately reflects surplus. Note that this is *not* about the standard question as to whether using the uncompensated (Marshallian) demand curve is a good approximation to the compensated demand curve (see Willig (1976); Hausman (1981)) but whether the demand curve systematically misrepresents willingness-to-pay and hence welfare for the reason that trading funds are often *not selling information direct to consumers but to other firms who in turn provide products to consumers.*

Consider, for example, the case where a trading fund is selling to a downstream firm who is in turn a monopolist in its own market. In that case, with royalty-based pricing, one would have a classic case of 'Cournot complements' and attendant double marginalization. In that case the demand curve seen by the trading fund would under-represent underlying demand and welfare changes.[6]

A similar, but additional, effect will also arise if downstream firms have fixed costs.[7] In its essence the effect arises from the existence of the Dupuit triangle.

[5]In fact knowing the values for the regions in the total welfare row in this case allow one to infer all the other values.

[6]Note that this effect still occurs if the downstream market is an oligopoly rather than a monopoly though the degree of double-marginalization will decrease as the level of competition increases.

[7]This effect occurs whether the tariff used by a trading fund is a royalty or a fixed fee – unlike the case of 'Cournot complements'.

To understand how, imagine there are a large number of downstream firms each demanding one unit of the product but each with different fixed costs. The trading fund's demand curve then arises from aggregating across all these downstream firms. Pick a point on the trading funds demand curve, p, q say, and consider an increase of δp in the price charged[8] resulting in some reduction δq in purchases. Now this reduction in demand corresponds to some number of downstream firms who cease to purchase (and hence cease production). Consider one of these firms and let initial revenue be R and C their total costs (excluding the payment for data). Then one must have $R - C \approx p$ (since $R - C < p + \delta p$ and $R - C \geq p$). What about the surplus generated by this firm? Its producer surplus is zero ($R - C - p = 0$) but consumer surplus, denoted CS, is almost certainly not zero. Thus, from the point of view of society current total surplus produced by this firm is $p + CS$. However using the demand curve of the trading fund all that would be recorded is the p coming from the payment for data.[9]

To illustrate, consider the following simple explicit example. Suppose each downstream producer is a monopolist[10] facing the linear demand curve: $P = k_i(1 - Q)$ where k_i is an indicator of product quality (so maximum demand is always 1 but willingness to pay increases in quality). Assume that there is a uniform distribution of product quality in some interval $[0, k]$, and that, for simplicity, the only costs for these downstream producers are those arising from the purchase of data (which takes the form of a one-off payment). Then a downstream producer with quality k is willing to pay up to $k/4$ (their maximum potential gross income) for the trading fund data. Thus, aggregating, the trading fund will face a linear demand curve

[8]Since only one unit of the product is demanded here this is necessarily a fixed fee. However this argument can easily be extended to the more general case where demand is variable and the trading fund sets a nonlinear tariff.

[9]This, of course is in the extreme case where the firms who no longer purchase simply cease operation. However the basic point still holds in the more realistic case where a rise in the price of the trading fund's product causes them to substitute it with another (necessarily inferior) input.

[10]Imagine, for example, the case of monopolistic competition in downstream markets.

with unit negative slope and intercept at $k/4$ which understates the actual value of any change in price by a third – that is the while willingness-to-pay at a point p, q on the trading fund demand curve is obviously p (equal to the monopoly profit of the marginal downstream firm) the actual surplus associated with this point is $3p/2$ because of the additional Dupuit (consumer surplus) triangle on the downstream firm's demand curve.

To sum up, the basic point to take from both this examples is that, if users of a trading fund's information products are not end consumers but other firms, then there is good reason to think that the demand curve seen by the trading fund will significantly *underestimate* the welfare benefits (costs) of lower (higher) prices.

3.3.2 Dynamics and Innovation

Another limitation of the basic approach laid out above is that it is entirely static in nature. The demand curve is shown only at a particular point in time with no allowance for how it might change, and, in particular, how changes in *present* prices might affect both *future* demand and related markets. Specifically, lower prices for data today, by increasing access and usage, might stimulate the rate of innovation by the producers of complementary goods – for example, cheaper geodata may lead to more rapid improvement in the quality of the software and hardware components of Geographical Information Systems (GIS). This in turn, not only increases the surplus generated in those related industries but stimulates demand for data in the future. Alternatively, increased access to data due to lower prices may lead to or assist the development of entirely new products and services – for example, Weiss (2004) argues that marginal cost access to weather data in the US was a large factor in the development of the multi-billion dollar weather derivatives industry. Similarly trading fund data may be reworked or incorporated with other information sources

by other firms to produce new products.[11]

It is quite possible for such effects on welfare to be large, much larger in fact than those arising from purely static considerations related to the underlying product's demand curve.[12] Furthermore this may be true even if the current costs of access are relatively low – at least relative to the potential benefits. This is somewhat surprising since normally one would imagine that the cost of a particular piece of information should place a (rough) upper limit on the value of the innovations which it enables.[13] However there are a variety reasons why this basic logic fails. The simplest example is to consider a chain of cumulative innovations in which an innovator at each stage can only extract some fraction r of the total surplus generated by the subsequent innovator. In this case for a chain of length N the initial innovator only receives r^N of the actual surplus generated (and so conversely an innovator with a willingness to pay of only X for a piece of data may be generating a surplus of X/r^N). Another possibility is that the innovation effort is distributed across many different firms or individuals ('componentized' innovation – as an explicit example one could think of an open-source project working to produce GIS software). In this case if each agent needs access to the underlying data supplied by the trading fund in order to contribute to the project the total costs may become so high as to be prohibitive.[14] Of course, this problem could potentially be addressed by having a special license of some kind. However this would entail a significant administrative burden (and

[11] One might also think of trading fund data as having some of the aspects of an *experience* good.

[12] Of course, in doing such a calculation, one would need to be cautious about how one allocated these 'spillover' benefits. Just because the data provided by a trading fund is used in (or is even central to) the activities of a particular firm does not mean one can allocate all the surplus generated to the availability of that data.

[13] To give a concrete illustration, suppose a particular set of geodata costs £1000 and there is a potential innovator who has an idea for a new product based on that geodata worth £X. If $X > 1000$ then the innovator should be willing to pay for access to the geodata. This suggests that only innovations worth less than £1000 are lost when the price is at this level.

[14] Suppose the innovation is worth V and the cost of data is X and their are N participants. If these N participants were all in a single firm which could obtain a single development license the cost is X, willingness to pay is V and the project is undertaken if $V \geq X$. However if the participants are distributed and must all buy their own license then the willingness to pay of any individual would is only V/N and the project is only undertaken if $NX \leq V$.

the associated transaction costs), for example, in ensuring that the licensed data did not 'leak' out of the authorised group.

Other more complex models found in the cumulative innovation literature also generate analogous results (see e.g. Bessen and Maskin (2006); Pollock (2006)). Their common theme is the presence of 'multiplier' (or 'spillover') effects whose relevance to this case is that the willingness-to-pay as encapsulated in the demand curve of the initial firm (in this case the trading fund) may significantly understate the true long-term benefits arising from access to that information.

3.3.3 The Multiplier

Both of the two previous sections provide reasons to think that using the basic demand curve may lead to underestimates of the gains from lower prices – equivalently, underestimates of the deadweight losses of higher prices. This would imply that, when doing cost/benefit style calculations of social welfare, one would need to scale up the welfare related to demand increases by some form of 'multiplier'. Therefore a parameter representing the 'multiplier' will be included in the calculations below. Empirical estimates for the 'multiplier' is an issue dealt with in Section 4.4 below.

3.4 Expressions for the Outcomes of Interest

This section converts the basic analysis above into explicit equations characterising the difference between regimes in terms of the key underlying variables (listed in Table 3.3). The numeraire for all of these calculations will be government funds (and not funds in the hands of consumers). That is, all values calculated will be relative to a pound in the hands of the government rather than a pound in the hands of a consumer. To convert from government funds back to consumer funds one need only multiply by a given constant ($1/\theta$ specifically – see table 3.3). Thus, the choice

of numeraire has no effect on the signs of any particularly value and therefore on choice of policy, but simply acts to scale a particular value by a constant. Taking government funds as the numeraire seems the natural approach here given their centrality in the calculations – it is government funds that will be used in paying any subsidy. Furthermore, this is consistent with the Government's own cost benefit guidelines as laid out in the 'Green Book' (HM Treasury, 2003).[15]

Additionally all calculations will be done to give average costs/benefits *per year*. If a reader requires total net present value of costs/benefits all that is required is to multiply by $1/(1-\delta)$ where δ is the relevant discount factor.

Additionally, a few general assumptions will be made and these are laid out here.

First, on the producer side, that (a) marginal costs for a given product are constant and (b) total costs equal the sum of a single fixed cost plus marginal costs. Both of these assumptions are minor. In the first case, given that one is dealing with information products (frequently in digital form) it seems reasonable to take marginal costs as constant rather than increasing – and at best it is likely that only a single marginal cost figure will be available. In the second case similarly, given the nature of information products, a basic division into fixed and marginal costs should be sufficient – and again a simple division into fixed and marginal costs is all that is likely to be possible with the data provided.

Next turning to the demand side, the very limited availability of data make necessitate some assumption about the shape of the demand curve. The approach adopted here will be to assume that in the region of interest the demand curve may be approximated by a linear function and thus that the elasticity of demand captures sufficient information for us to calculate changes in consumer and producer surplus. For small changes in prices such an approximation is quite reasonable. Of course here the price changes under consideration are likely to be quite substantial. In this

[15]See below in Chapter 4 and Appendix A.2.

case using a more convex inverse demand function (e.g. $p = 1/q$) or a more a concave one (e.g. $p = k - q^2$) might lead to changes in the surplus estimates. Nevertheless, given the data constraints an assumption of linearity seems a reasonable first-order approximation.

Next, turning to the question of a change in overall costs as the pricing regime changes it will be assumed that there is no change. This is important because it allows us to base the calculations simply on changes in revenue and output between regimes. For example, with this assumption when moving from average costs to marginal costs the shortfall that the Government would need to 'make up' to the trading fund equals the loss of revenue. In reality it is likely that such a change in pricing regime would also be accompanied by a change, likely a reduction, in costs – for example, if marginal costs are zero (so the information is provided free) there may be significantly less administrative overhead in relation to billing, contract monitoring, enforcement etc.[16] However, while such cost changes might not be non-negligible, the easier approach is to 'assume them away' – at least for the time being – for three reasons. First, such cost changes would be hard to calculate given the data available. Second, such cost changes are likely to be 'second-order', that is small relative to the main effects. Third, and perhaps most decisively, such an omission is 'conservative', in the sense that it biases the results towards the existing (average-cost) regime. While inserting 'bias' is never first-best, inserting a 'conservative' one could be seen here as a reasonable 'second-best' – and where a marginal cost (or zero) price cost regime is found to be preferable, this 'bias' would be irrelevant in the sense that it would make the preferability 'stronger'.

Finally, transaction costs, whether in modifying charging policy or those incurred in running a particular regime, have been ignored. The primary reason for this is that there is absolutely no evidence on which to base estimates of their magnitude.

[16]There is some more discussion on this specific point in Section 3.4.4 below.

However if these were a priori thought to be large, this might not seem a strong reason to set them at zero. Thus it is worth making some further remarks. First, in the case of the costs of transition one would imagine these to be reasonably small, at least compared to the magnitude of the other sums involved. Second, for general transaction costs it would seem reasonable to assume that starting from an average cost regime (the current state) these would either go down due to less need for monitoring and enforcement – when moving to marginal or zero prices, or would be relatively unchanged – when moving to profit maximization. Thus, ignoring them can either be seen as having little effect (move to profit-maximization) or as instilling a 'conservative' bias in favour of the existing regime. Again, inserting such 'bias' is not first-best, but given its 'conservative' nature it could be seen as a reasonable 'second-best', particularly as it compensates for ignoring any transition costs.

3.4.1 Summary of Key Variables

The key variables used in the analysis below are set out in Table 3.3.

3.4.2 Average Cost (Cost Recovery) Versus Marginal Cost

The comparison of average cost versus marginal cost is made first for several reasons. First, and most importantly because it is the estimate least subject to error. This is because, as discussed above, the data points available are located closest to the regions relevant to this comparison. As one moves away from these values uncertainty increases and the estimates become more unreliable. Second, it is this comparison that appears to be of most interest to policy-makers – for example, Treasury (2000) only considers the choice between these two options.

The costs of a change from average cost to marginal cost equals the subsidy now required for the trading fund. This, in turn, equals the drop in revenue resulting

Name	Variable
θ	Distributional weight for the project under consideration.
$1+\alpha$	The marginal cost of public funds. Note that θ and α are linked via $1+\alpha = 1/\theta$.
π	Proportion of 'consumer surplus' that is actually producer surplus. As already discussed, much of the demand faced by a trading fund comes not from consumers themselves but from other producers.
τ	The corporate tax rate, and hence the proportion of producer surplus that is returned to the government as tax.
δ	The discount factor.
T	Time delay in realizing increased usage. This takes account of the fact that reductions in prices may take some time to have an effect. Note that costs, whether to government or others, have an immediate impact.
F	Revenue net of marginal costs (equal to fixed cost under cost recovery).
g	The proportion of revenue derived from government sources.
c	Marginal cost.
λ	Demand curve 'multiplier'. Note that $\lambda \geq 1$.
$\epsilon = \epsilon(p)$	Elasticity of demand at price p. When no price is specified this is the elasticity with the price and output set under a cost-recovery regime.
p, q	Price and output under average cost pricing (point C in figure 3.1).
p^c, q^c	Price and output under a marginal cost regime (point F in the figure). Note that $p^c = c$
p^m, q^m, c^m	Price, output and average costs at the profit-maximizing level (point in A in the figure). Note that $c^m = c + f/q^m$.
q^0	Output under a zero cost regime (the price in this case, obviously being zero).
$\Delta p = p - p^c$ $\Delta q = q^c - q$	The (absolute) change in price and quantity as a result of moving from average cost to marginal cost pricing.
$\Delta p^m = p^m - p$ $\Delta q^m = q - q^m$	The (absolute) change in price and quantity as a result of moving from an average cost to profit-maximizing regime.
$\Delta p' = p^m - c^m$	The difference between the profit-maximizing price and average cost.
$\Delta q^0 = q^0 - q^c$	The change in output when moving from marginal cost to zero-cost pricing.
$\beta = g + \tau\pi(1-g)$	Defined for convenience.
$\gamma = \lambda\epsilon\Delta p/2p$	Defined for convenience. (Roughly, this is the additional surplus from a reduction in price of Δp).

Table 3.3: Key Variables.

from this change. Using the variables defined above:

$$C = q\Delta p = F$$

Note that the second part of this equation follows from the fact that, under average cost pricing per product, the change in trading fund surplus ($q\Delta p$) should exactly equal total fixed costs.

On the other hand, the benefits are:

$$B = F + S = F + \delta^T DWL$$

The F figure arises because the reduction in trading fund income is transferred directly across to outside producers and users. In addition there is a gain to society due to increased usage which is captured by S (new surplus) term. This in turn is equal to the DWL (Deadweight Loss Loss)[17] multiplied by δ^T to take into account the fact that this benefit may not be realized for T years.

The next step is to ensure that both of these terms are represented in terms of the numeraire which is government funds. Now obviously government expenditure or income need not be modified but any gains to those outside government need to be scaled by the distributional weight θ. Thus the first step is to breakdown the costs and benefits into those accruing to the government, and those accruing outside government, whether to consumers (consumer surplus) or to producers (producer surplus). The breakdown is show in Table 3.4.2. The important point is that the values in the unweighted section are not necessarily commensurable since they are not expressed with respect to the same numeraire. Those in the second 'Weighted' section have been corrected with the necessary distributional weights to ensure they are all expressed in terms of the numeraire used (government funds).

[17]The term 'loss' is a little confusing here since it is actually what is gained when moving to a marginal cost regime.

Item	**Expression**
Unweighted (no common numeraire)	
Cost:	$-F$
Benefit: B	$F+S$
o/w Govt	gF
o/w Non-Govt Surplus: H	$(1-g)F+S$
– o/w Consumers	$(1-\pi)H$
– o/w Producers	$\pi(1-g)H$
— o/w Tax to Govt	$\tau\pi H$
— o/w Producers	$(1-\tau)\pi H$
Weighted (numeraire = govt funds)	
Δ Government: ΔG	$-F+gF+\tau\pi H$
Δ Consumer Surplus: ΔCS	$\theta(1-\pi)H$
Δ Producer Surplus: ΔPS	$\theta(1-\tau)\pi H$
Δ Total Welfare: ΔW	$\Delta PS+\Delta CS+\Delta G$

Table 3.4: Theoretical Breakdown of Surplus

Substituting for Producer Surplus, Consumer Surplus and Government income/expenditure, and rearranging one has the following expression for the change in Total Welfare:

$$\Delta W = \theta F\left(\frac{H}{F}(1+\alpha\beta)-(1+\alpha)(1-g)\right)$$

Where $1+\alpha = 1/\theta$ and $\beta = \tau\pi$. The final step is to determine the size of the gain from increased usage: DWL. The exact size of the change will depend on the shape of the demand curve. Using the linear form for the demand curve the expression for the deadweight loss takes a particularly simple form as follows (where ϵ is the elasticity of demand at point C):

$$DWL = \lambda\text{Triangle CFE} = \lambda\frac{1}{2}\Delta p\Delta q = \lambda q\Delta p\frac{\Delta q}{2q} = \lambda F\frac{\Delta q}{2q}$$

As Δp and p are known and estimates for ϵ may be available but those for Δq may not it will be useful to rewrite this using the fact that $\Delta q/q = \epsilon(p)\Delta p/p$.[18]

[18]Note that this would normally be only an approximate equality but for the case of linear demand it is exact.

ΔG	$-F((1-g)(1-\beta)-\beta\delta^T\gamma)$
ΔCS	$\theta F(1-\pi)(1-g+\delta^T\gamma)$
ΔPS	$\theta F(1-\tau)\pi(1-g+\delta^T\gamma)$
ΔW	$\theta F\left((1+\alpha\beta)\delta^T\gamma-\alpha(1-g)(1-\beta)\right)$

Table 3.5: Average Cost vs. Marginal Cost: Expressions for Outcomes of Interest

Defining $\gamma \equiv \lambda\epsilon\frac{\Delta p}{2p}$ one has:

$$DWL = F\lambda\epsilon\frac{\Delta p}{2p} = F\gamma$$

Thus, $H = (1-g)F + \delta^T F\gamma = F((1-g)+\delta^T\gamma)$. Substituting one has:

$$\begin{aligned}\Delta W &= \theta F\left((1+\alpha\beta)(1-g+\delta^T\gamma)-(1+\alpha)(1-g)\right)\\ &= \theta F\left((1+\alpha\beta)\delta^T\gamma-\alpha(1-g)(1-\beta)\right)\end{aligned}$$

In terms of decision-making all that matters is whether the change in social welfare is positive or negative ($\Delta W <> 0$). Since the term outside of the brackets is always positive it follows that ΔW is greater than zero, and hence that marginal cost pricing delivers higher social welfare than average cost pricing, if and only if:

$$\gamma \equiv \lambda\epsilon\frac{\Delta p}{2p} \geq \frac{\alpha(1-g)(1-\beta)}{\delta^T(1+\alpha\beta)} \tag{3.4.1}$$

Finally, the general expression for the outcomes of interest are summarized together in Table 3.4.2.

3.4.3 Profit Maximizing versus Average Cost

For the case of moving from a profit-maximizing to an average cost regime the analysis is exactly as above except that F should be taken to be the change in trading fund income (this is because the government's role as 100% shareholder and provider of a subsidy mean that all changes in trading fund income translate through directly to government income). Specifically one should set $F = \Delta p^m q^m$.

Note also, that the same approach can also be used to compare profit-maximizing and marginal cost regimes. In this case one should set $F = (p^m - p^c)q^m$. This immediately implies that if marginal cost pricing is better for social welfare than average cost pricing then marginal cost pricing is also better than profit-maximization pricing. To see this algebraically, note that marginal cost is better than average cost if:

$$\lambda\epsilon(p)\frac{\Delta p}{2p} \geq \frac{\alpha(1-g)(1-\beta)}{\delta^T(1+\alpha\beta)}$$

Now the elasticity at the profit maximizing price is greater than at average cost price: $\epsilon(p^m) \geq \epsilon(p)$ and since $p^m > p$ it also the case that: $\frac{p^m - p^c}{2p^m} > \frac{\Delta p}{2p}$. Thus one has that:

$$\lambda\epsilon(p^m)\frac{p^m - p^c}{2p^m} \geq \lambda\epsilon(p)\frac{\Delta p}{2p} \geq \frac{\alpha(1-g)(1-\beta)}{\delta^T(1+\alpha\beta)}$$

But this is precisely the conditions under which marginal cost is better than profit maximization.

3.4.4 Marginal Cost versus Zero Cost

Here the good is going to be supplied for less than its marginal cost. Normally this should immediately imply that the welfare impact of such a change were negative – providing a good for less than its opportunity costs must necessarily be inefficient (a fact made worse here by the fact that such provision is funded out of government tax revenue). However the existence of the multiplier means this is not quite as

ΔG	$-F((1-g)(1-\beta)-\beta(1+\delta^T\gamma)+2\gamma/\lambda)$
ΔCS	$\theta F(1-\pi)(1-g+\delta^T\gamma$
ΔPS	$\theta(1-\tau)\pi(1-g+\delta^T\gamma$
ΔW	$\theta F\left(\gamma((1+\alpha\beta)\delta^T-2(1+\alpha)/\lambda)-\alpha(1-g)(1-\beta)\right)$

straightforward as intuition might suggest.

Formally, setting F to equal initial revenue: $F = p^c qc$, noting that the subsidy actually required is not just F but $F + 2p^c\Delta q^0 = F + 2F\epsilon\Delta p/2p = F + 2F\gamma/\lambda$, and substituting produces the values for the outcomes of interest listed in Table 3.4.4.

Note that all terms will be negative, and hence the change in welfare will be negative, if $(1+\alpha\beta)\delta^T - 2(1+\alpha)/\lambda)$. Observing that $(1+\alpha\beta) \leq 1+\alpha$ it would be sufficient if $2 \geq \lambda\delta^T$. Of course ignored in this is the possibility that there may be transaction costs involved in charging (e.g. related to access control, billing etc). These costs will be present with marginal cost pricing but absent under zero cost. Should the difference between marginal cost pricing and zero cost pricing be small then this issue could well be significant. Given that the good under consideration is digital data and that internet bandwidth is already cheap and getting cheaper the possibility that marginal cost prices could be close to zero is not an idle one. Furthermore, recall that it has been assumed that both marginal cost and zero cost pricing allow users to redistribute and reuse the data freely. As a result, a trading fund might only need to supply a small fraction of total demand and this would greatly reduce the cost burden.[19] That said, when marginal costs are close to zero this whole issue becomes moot for then the marginal cost and zero cost regimes are essentially identical.

[19]It might also permit more efficient distribution mechanisms, for example using peer-to-peer type technologies such as bittorrent. These sorts of mechanisms are already being experimented with by content producers such as the BBC as a way to distribute high volumes of information more cheaply.

Chapter 4

Empirics I: General Variables

4.1 Introduction

The aim of this chapter and the following is to apply data to the models derived in the previous one. Specifically, data from trading funds together with the existing literature will be used to estimate the key parameters of those models and thereby to obtain values for how different charging policies affect outcomes. This chapter focuses on providing general estimates or ranges for parameters that are not specific to a particular trading fund while the following one uses these together with specific data from each trading fund to derive estimates of charging policies on the key variables of interest (social welfare, government subsidy etc).

4.2 Distributional Weights and the Social Value of Public Funds

As previously discussed the numeraire for surplus calculations will be uncommitted government funds. It is then necessary to compute θ, the distributional weight for the specific project under consideration. Roughly the logic here is that uncommitted

public funds could either be used for lowering the price of PSI or for some other government purposes. These uncommitted funds, by definition, have a weight of one.[1] This is almost certainly not true for the (consumer and producer) surplus generated by the project under consideration, and the appropriate distributional weight will depend on how the project's benefits are realized across the population which in turn depends on the existing distribution of income, the degree of inequality aversion, the marginal utility of consumption and the income elasticity of demand for PSI data.

Here, it will be assumed that the benefits from lowering the price of PSI are received in proportion to income. Specifically, the income elasticity of consumption of trading fund PSI is assumed to be one. This is a fair assumption given that general consumption is (approximately) proportional to income. Using data from the UK 2003-04 Expenditure and Food Survey (for National Statistics, 2005) gives a log income variance of 0.47. The Green Book (HM Treasury (2003)) quotes Cowell and Gardiner (1999) conclusion that "most [studies] imply values of the elasticity of marginal utility of just below or just above one". It also states that "Pearce and Ulph (1995), in their survey of the evidence, estimate a range from 0.7 to 1.5, with a value of 1 being defensible." This study therefore uses a range of 0.7 to 1.5 with a point estimate of 1. Consequently θ ranges from 0.718 to 0.857, with a point estimate of 0.802.[2]

Note that the inverse of θ will equal the marginal cost of public funds (MCPF).[3] The MCPF approach is less well-grounded in welfare economics but, as just noted, should give similar results. It is based on taking general consumption as the numeraire and it was the approach adopted by HM Treasury in its original analysis of

[1] That is, there is at least one project in the government portfolio where £1 of expenditure generates benefits equivalent to £1 equally distributed across the population.

[2] This estimate applies for a utilitarian social welfare function. For further details of the exact calculations see Appendix A.2.

[3] Assuming the tax regime is optimal.

these questions (Treasury, 2000). In particular, the point-estimate of the inverse of θ is 1.25, which is very similar to the estimate used by HM Treasury in Appendix C of Treasury (2000).[4] Appendix A.3 gives a general review of the MCPF literature and estimates.

4.3 Elasticity of Demand

4.3.1 Introduction

The formal definition of price elasticity of demand is given in the expression below where p is the price of data, q is output and where δ represents an infinitesimal increase in the variables:[5]

$$\epsilon = -\frac{(\delta q/q)}{(\delta p/p)}$$

Intuitively one can think of the price elasticity of demand as the percentage increase in demand for trading fund data for a one percentage point decrease in price. Similarly, a change from average cost to marginal cost pricing (or vice-versa) allows one to elicit the elasticity of demand. However as the price changes can be quite substantial, the elasticity will depend upon whether one uses the old price and output pair, p_0, q_0, or the new price and output pair, p_1, q_1. This is best illustrated with an example, say where a price rise results in revenues increasing by 40 percent and output decreasing by 40 percent. Let R_0 and R_1 be old and new revenues respectively. The relation between old and new revenues and old and new prices can be expressed as:

$$R_1 = \frac{7}{5}R_0 \tag{4.3.1}$$

$$q_1 = \frac{3}{5}q_0 \tag{4.3.2}$$

[4]Based on Ruggieri (1999), they use a range of 1.2 to 1.3.

[5]Note that usually there would be no negative sign at the front here. However, for convenience, and to match with the definition used in the theory section above, the elasticity of demand has been defined so as to (normally) be positive rather than negative.

Now assume that the demand curve is linear. If one defines the elasticity using the old price output pair then:

$$\epsilon_0 = -\frac{(q_1 - q_0/q_0)}{(p_1 - p_0/p_0)} \tag{4.3.3}$$

$$R_1 = p_1 q_1 = \frac{7}{5} p_1 q_1 = R_0 \tag{4.3.4}$$

Substituting (4.3.2) into the (4.3.4) gives:

$$p_1 = \frac{7}{3} p_0 \tag{4.3.5}$$

Similarly substituting (4.3.5) into (4.3.3) one finds that $\epsilon_0 = --3/10 = 0.3$. However if one defines the elasticity using the new price output pair then:

$$\epsilon_1 = -\frac{(q_1 - q_0/q_1)}{(p_1 - p_0/p_1)} \tag{4.3.6}$$

Substituting (4.3.5) into (4.3.6) one finds that $\epsilon_1 = --7/6 = 1.17$.

For a linear demand curve the high price, low output pair generates a significantly higher elasticity. The demand curve may not be linear. Demand may be more inelastic at higher prices, where there are a few large businesses who simply have to have the data and so are willing to pay a very high price. However the demand may also be particularly elastic lower down the demand curve, where a substantial amount of experimentation with the data may take place. Both these effects would reduce the effect on elasticity estimates with different price, output pairs. Nonetheless it is important to note that where there are substantial price changes, there will also be a significant range in elasticity estimates using the same underlying data. Where necessary it would seem more appropriate to use the mid-point as opposed to either the upper or lower bounds. With this theory in mind the literature on the

price elasticity of the demand for similar products to Trading Fund data is reviewed. Remember when reading the sections below that the elasticity has been defined so as to normally be positive rather than negative.

4.3.2 Direct estimates of the elasticity of PSI

Treasury (2000) argue that the price elasticity of demand is likely to be inelastic: 'The demand for information should be price inelastic because it is an input into a production process and a small proportion of production costs. Suppose the price elasticity of demand for a product was 2 and information costs were as high as 10 per cent of total costs. Then a good approximation to the derived demand elasticity for the information input into this product is (10 per cent times 2=) 0.2.' However this assumes complete cost pass through. That is if information costs go up by 20 percent then the final product price will go up by 2 percent (10 percent of 20 percent). However this may not necessarily be the case. For example increasing costs of the information input may no longer make it profitable to use it as part of the final product. This could lead to a substantial increase in the demand elasticity of the information. Alternatively the increased cost of the information may have no effect on the final product price; the input mix could be adjusted to sustain the same price. In this case the demand elasticity for information could be even less elastic.

Weiss (2004) argues that the price elasticity for information is likely to be high in most cases: 'Only when use of the information is mandatory or somehow indispensable might the demand be less elastic.' One such argument in favour of high elasticity is that businesses will make do with lower quality substitutes, if prices are set too high. In order to deduce a likely range of elasticities, direct evidence on the demand for similar information products is first reviewed.

The OFT report (OFT, 2006) provides a good starting point in their survey of the

elasticities of demand for information in other countries. It suggests an elasticity of 0.3 for New Zealand national mapping data based on evidence quoted by Longhorn and Blakemore (2004): 'Rhind reviewed data charging outcomes after New Zealand had imposed a rigorous cost recovery programme on national mapping, noting a reduction in sales between 1989 and 1994 of 60%, although income was 25% greater in real terms.' However this calculations appears to be using the lower bound – using the same calculation as in section 4.3.1 one finds that 0.3 is the lower bound and that the upper bound is around 2.2.

Davies and Slivinski (2005) suggest that the elasticity for demand of weather forecasts is 0.3 based on evidence by Lazo and Chestnut (2002). However this paper only measures direct household demand for improving day-to-day weather forecasts through stated preference surveys. This should therefore be treated as a lower bound since it excludes demand for weather data coming from intermediaries and the private sector.

Using a study for Bedrijvenplatform (2000) that claims 'lowering the price of public sector geographic data by 60 per cent would lead to a 40 percent annual turnover growth', Office of Fair Trading (2006) deduce an elasticity of demand of 1.7. Interpreting turnover as revenue one finds an upper bound elasticity of 4.17 and a lower bound elasticity of 0.48 using the same calculations as in section 4.3.1.

Under the *Making Information Freely Available* initiative, Statistics New Zealand is in the process of making a wide range of products and data available for free.[6] For example, Digital Boundaries Files on CD and StreetLink files were distributed for free from July 6th 2007.[7] Digital Boundaries Files previously cost around NZ$3,300 for the standard five-yearly census pattern, or NZ$25,212 for the annual detailed file. StreetLink Files previously cost NZ$6000 for first supply and then NZ$2000 for

[6]The policy press release is available at `http://www.beehive.govt.nz/Documents/Files/Statistical\%20Info\%20FAQ.pdf` for details.

[7]See press statement available at `http://www.beehive.govt.nz/ViewDocument.aspx?DocumentID=2998`

annual updates. As of August 28th 2007 around 250 copies of Digital Boundaries CDs and 75 StreetLink files have been provided.[8] This is a two-fold and ten-fold increase in Digital Boundaries Files and Street Link Files respectively in the 6 weeks after charges were withdrawn compared with what Statistics New Zealand sold in the past three and a half years.

Using these immediate changes in demand would imply very high elasticities. However the initial surge of requests could be a consequence of a backlog of demand for the data at zero-cost and so the annual uptake is likely to stabilise at a much lower level. Bearing this in mind it seems appropriate to use this recent demand to approximate the average annual uptake. Doing so and using equation (4.3.3) one finds an upper bound elasticity of around 6 and 34 for Digital Boundaries and StreetLink Files respectively. Small Area Population Estimates which previously cost around NZ$250 were made free to download on August 28th 2007. By September 14th 2007 there had been 184 accesses by unique visitors compared to around 75 customised jobs per year previously. Again using this recent uptake to approximate the new annual output and using equation (4.3.3) one finds an upper bound elasticity of around 1.5. These estimates are still likely to be too high since the high surge in demand may include a large number of users who are unlikely to find the data of use, but request it at no cost to see if it may be suitable.

The Australian Bureau of Statistics made information free on their website towards the end of 2005. Table 4.3.2 shows the total products download statistics from 2003-2007.[9] Figure 4.1 graphs usage of ABS statistics over this time period.[10] It is clear that there is a significant increase in the usage of data once it was made

[8] See statement available at http://www.beehive.govt.nz/ViewDocument.aspx?DocumentID=30426

[9] This data is available in Table 13.3 in the ABS Annual Report at http://www.abs.gov.au/AUSSTATS/abs@.nsf/39433889d406eeb9ca2570610019e9a5/FBF88ADA798ABCA1CA257371001411C3?opendocument

[10] Available at http://www.epsiplus.net/content/download/7380/88070/file/3_3_ePSIplus_TM2_Pricing2_QUT_11107.pdf

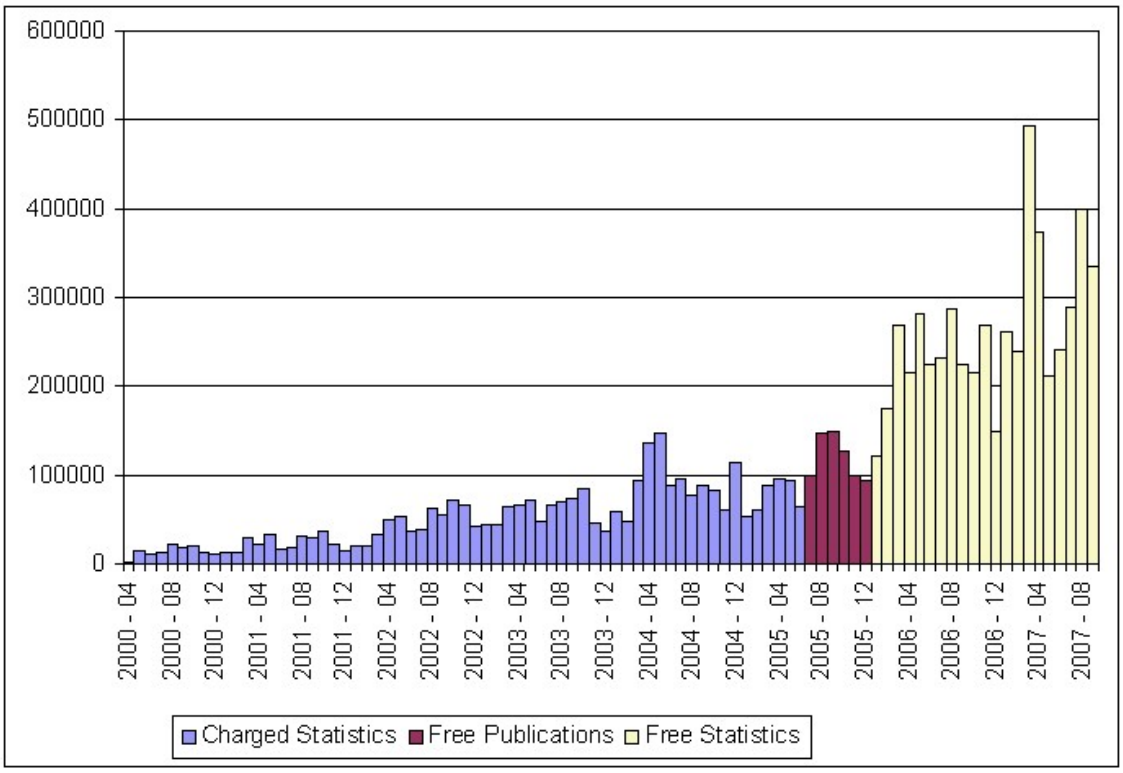

Figure 4.1: Australian Bureau of Statistics 'Dissemination of Statistics'.

	2003-04	**2004-05**	**2005-06**	**2006-07**
Reported	948,956	962,872	1,868,280	4,501,530

Table 4.1: Product Downloads from ABS website

freely available. Comparing the average dissemination of 2003-2005 with 2005-07 estimates (crudely) gives an elasticity of 2.33.[11]

The Office of Spatial Data Management in Australia conducted a wider programme to make available fundamental spatial data across a range of agencies for free or at marginal cost. [12] The policy was announced in September 2001 and implemented over a 6 month period so that by February 2002 agencies were provid-

[11]Using the 2007 values rather than an average 2006-2007 would give an even higher elasticity of around 3.5. Thus the long-run elasticity might well be even higher – though of course one would need to then make efforts to detrend for the effect of technical advance and general growth in demand.

[12]This list of fundamental spatial datasets is listed on the Data Schedule available at `http://www.osdm.gov.au/schedule/schedule\_search.jsp`

Year	Scheduled Dataset Units Delivered	Trend 1
2000-01	(75,310)	-
2001-02	75,310	75,310
2002-03	83,049	108,597
2003-04	52,565	156,597
2004-05	219,821	225,813
2005-06	862,530	325,622

Table 4.2: Office of Spatial Data Management Scheduled (free) Datasets Delivered. The figures in brackets are estimates. Trend 1 uses a growth rate of 44.2%.

ing data for free online, or at marginal cost in CD format. Table 4.3.2 details the delivery figures for scheduled datasets (i.e. those that fell under the new policy).[13] Unfortunately no data was available for the period before the pricing policy was announced. However, if one makes the conservative estimate that data delivered in 2000-01 was no more than in 2001-02 and compares this to uptake in 2005-06 gives an elasticity of upper bound elasticity of 10.45. Of course this does not take into account any general increase in demand due to other factors. One approach would be to detrend using the ABS figures since their data did not become freely available until 2005. Using the ABS data a reasonably generous estimate for the growth rate 2001-2005 (for non-free data) would be around 44.2%. The effects of applying this growth rate is shown in Table 4.3.2 as Trend 1. Comparing the 2005-06 value in Trend 1 with the reported value suggests an elasticity of 1.65.

[13]Figures from 2001-02 are quoted by OSDM as from the fundamental dataset. This is understood to be all data listed on the Data Schedule as 'This Policy is premised on the view that all fundamental spatial data should be freely available at no more than marginal cost of transfer in order to maximise the net economic and social benefits arising from its use' (`http://www.osdm.gov.au/fund\_pricing.html`). OSDM also state that the 'Australian Government spatial datasets that are available under the terms of the Policy on Spatial Data Access and Pricing ('the Policy') are listed on the Schedule.' (`http://www.osdm.gov.au/schedule/schedule\_search.jsp`). This policy states that ' Fundamental spatial data will be provided ... at no more than the marginal cost of transfer...' (`http://www.osdm.gov.au/policy/accessPricing.html`)

4.3.3 Evidence from similar sectors

Direct evidence on the price elasticity of information is limited. Estimates are derived from large changes in prices which result in a large range of elasticities. Furthermore the sources of demand data are mostly from the supplier. Therefore, in order to supplement the direct evidence on the price elasticity for information the related body of evidence on telecommunications is also examined. Many analogies can be drawn between the two sectors making them suitable for comparison. Both are related to innovation and new technology. Both serve as inputs into other activities. Both sectors display spillover (multiplier) effects – a business does not only directly benefit from using email, but also benefits others in the ease with which they can communicate. Telecommunications is also a route through which information can be distributed and hence they are intrinsically related. The internet for example offers access to a wide range of information. Part of the demand for access to the internet will therefore reflect the demand for this information, and so the elasticities in each sector can be compared.

Hausman, Pakes, and Rosston (1997) finds a price elasticity of 1.61 and 0.51 for the introduction voice messaging and mobile phones respectively in the united states. Goolsbee (2006) finds an average price elasticity of demand for broadband of 2.75 at an average price of $40 per month for a range of metropolitan areas in the US. Goolsbee and Klenow (2006) takes into account the opportunity cost of ones time to deduce the value of using the internet and so estimates a price elasticity of 1.6. Kridel, Rappoport, and Taylor (2002) find a price elasticity of broadband of about 1.8 at $49.95 a month. Hackl and Westlund (1996) finds a range of price elasticities of demand for international telecommunications in Sweden from 0.09 to 1.25.

4.3.4 Conclusion

Clearly, the elasticity will vary depending upon the product under consideration. To allow for this three ranges for the elasticity of demand for PSI will be used (these are absolute values – the elasticity of demand is of course negative):[14]

- Low: 0-0.5 (midpoint: 0.25)
- Medium: 0.5-1.5 (midpoint 1.0)
- High: 1.5-2.5 (midpoint 2.0)

The evidence presented above suggests that, in general, the medium or high range would be the most appropriate for PSI products of the kind that will be examined in this report. For example, the experience of the ABS and OSDM in Australia suggest elasticities in the high range (or even above) as does the evidence from New Zealand.[15] Of course, as already stated, products vary substantially and assignment of an elasticity category will always be taken on a case by case basis when performing the trading funds analyses below.

4.4 The Multiplier

The theoretical analysis in Section 3 provided some reasonable a priori grounds for believing that the 'multiplier' could be significant. However, it would obviously be important to have empirical evidence for the significance and magnitude of the 'multiplier'. Unfortunately, there is, at present, very little empirical evidence available. This is perhaps not surprising given the empirical difficulties to be faced and

[14]These ranges should be interpreted as reasonably short-run elasticities. Over the long term elasticities are likely to be higher as new uses and applications for data are found.

[15]HM Treasury's analysis in Appendix C of Treasury (2000) appeared to think that a value of around 0.2 would be reasonable. However that study had no direct evidence available to it and the material which has accumulated since (presented above) would suggest that this is likely to be a significant underestimate.

the general lack of the detailed time-series firm-level data which would be required. However there are some suggestive individual items as well as a body or more 'anecdotal' evidence that can be drawn upon.[16]

Weiss (2004) argues that marginal cost access to weather data in the US was a large factor in the development of the multi-billion dollar weather derivatives industry (and that its limited availability has retarded developments in the EU). An analogous argument for general weather services is made in a recent paper by Richard Pettifer, general secretary of PRIMET.[17] It argues, that particularly by comparison with the US, the EU weather marketplace is seriously underdeveloped. It goes on to argue that much of the potential, but unrealized value, lies in the 'small unit value sector of the market place which is extremely price sensitive.' Furthermore, and of more relevance to this section, realizing the potential value of those markets would involve the development of new products and services based on cheaper access to the data collected by national meteorological services.[18]

Turning to geographic data, again hard data is sparse. Returning to Australia, (Spatial Information Industry Action Agenda, 2001) presents evidence that reducing the price of access to geographic information had a significant impact on use and, more importantly, reuse: 'The most important impact has been the dramatic increase in the volume of data sold. In Victoria, the number of licences or "seats" has

[16]There is, of course a significant literature on spillovers in R&D, particularly from public to private R&D. For example Jaffe (1989) and Mansfield (1995) both provide evidence of large spillover effects in this area.

[17]*Towards a Stronger European Market in Applied Meteorology.* PRIMET is the association of Private Meteorological Services. Obviously, their particular interest in this area should be taken into account when considering the arguments made in the document.

[18]Specifically, according to the document: "[T]his potential market [the small unit value, high potential demand] is not reached by the large government owned players because their high fixed costs and politically sponsored operating constraints prevent them from delivering the end user price and flexibility this market demands. It is not fully penetrated by the small, private sector companies largely because the exploitation of the monopoly supply position of the government owned players in respect of the raw material necessary to permit the development of suitable products at appropriate market prices. The data are subject to wholesale pricing that is too high and in some cases there is a failure to supply the data in a timely fashion (or at all), while re-use license terms can render it impossible fully to exploit the non rival nature of the data."

increased from around ten before the price reductions to about 600. In Queensland, over 75 licences to distribute and value-add to the data have been issued, whereas under the previous arrangements no whole-of-state sales were made at the then commercial rate.' Meanwhile, Bedrijvenplatform (2000), looking at the Netherlands suggested that a substantial portion of the benefits from cheaper geodata would arise from the development of new products and services. In the UK, the Ordnance Survey themselves commissioned Oxera in 1999 to estimate the value of the economic infrastructure 'built on' OS data.[19] The resulting report gave an estimate that around £79-136 Billion of Gross Value Added came from activities for which the Ordnance Survey's geographic information was a primary input. Of course this figure does not tell one much directly about the multiplier since the fact that many businesses use (or even depend) on OS data does not indicate how large the spillovers are or how much innovation is occurring. Nevertheless the report is indicative of the fact that geographic information is widely used, particularly as an input into intermediate products and services, which in turn suggests the multiplier could be quite significant.[20]

Finally, analogies can also be drawn with the spill-overs in other sectors. The Power of Information review (Mayo and Steinberg, 2007) itself provided several examples. For example, in medical studies such as Rodgers and Chen (2005) and Ziebland (2004) on breast cancer and Hellinger (2002) on HIV, it was found that access to medical information on the internet allowed users to cope better with a resulting reduction, in some cases, in treatment costs. On a different tack, Hampton (2007) finds that members of 'wired' neighbourhoods are more likely to know each

[19]See `http://www.ordnancesurvey.co.uk/oswebsite/aboutus/reports/oxera/index.html`

[20]The argument that there are large potential gains from increased access to and reuse of PSI can be found in the PIRA report prepare for the European Commission back in 2000 (PIRA, 2000) – with a similar set of points made in OECD, Working Party on the Information Economy (Directorate for Science Technology and Industry) (2006). As with most material the contentions are based more upon analogy with the United States, and a general consideration of the market, than any 'hard' data – not surprising given how difficult 'hard' data would be to obtain.

other and Lomax (2005) finds that providing clear information with medication can improve patient adherence to medical advice by 16-33%. One could argue that similar spill-over would be present for some of the products considered below. For example, easier access to DVLA data could enable more and better HPI checks, leading to a greater return of stolen vehicles and a reduction in theft. Or consider the Land Registry's data on property boundaries where better access could make it easier for planners of construction projects to contact those owning neighbouring land.[21]

4.4.1 Conclusion

Turning this diverse, and predominantly anecdotal evidence into an exact estimate for the 'multiplier' is clearly impossible. Furthermore, the multiplier will vary across products (just as the elasticity will). Thus the simplest way to proceed is to create 3 basic ranges. Recalling that the multiplier has a lower bound of 1 corresponding to no multiplier effect on welfare, a suitable set would be:

- Low: 1-2 (midpoint: 1.5)
- Medium: 2-4 (midpoint: 3.0)
- High: 4-10 (midpoint: 7.0)

Given the great uncertainty about the exact value for the multiplier any assignment for a particular product will necessarily be substantially speculative. Thus it should be emphasized (as discussed further below) that all welfare calculations will be checked for robustness using a multiplier of one (i.e. no effect). This way, while

[21] One could multiply these examples of 'potential' applications almost indefinitely. Easier access to current and historical weather data might help those researching climate change. Better access to geographic information would enable greater citizen understanding and participation in the planning at the local, regional and national level. Increased freedom of reuse would greatly multiply the potential for specific groups, whether those with disabilities such as the blind or with particular interests such as walkers, to add value to basic geographic data whether via annotation or integration with other sources of data.

the multiplier is incorporated into the analysis one can also be sure that it is not 'driving the results'.

4.5 The Corporate Tax Rate

The main rate of corporation tax is 30%. There are however special rates for unit trusts and open-ended investment companies of 20%. In addition the small companies rate is 20%. [22] A reasonable range for τ is therefore 0.2-0.3 and a mid-point of 0.25 will be used where necessary.

4.6 Proportion of gain in Surplus attributed to Producers

There is no data known to the authors which would allow one to calculate π, the proportion of any gain in surplus (due to lower prices) which can be attributed to producers. That said, the existence of competition and the basic shape of most demand curves would suggest that π should be under 50%. Given this, and lacking any further information, of 0.3 seems a not unreasonable value to use. Any reader unhappy with this choice should keep in mind that the impact of π on the results will be very limited. Furthermore, any changes from the status quo will checked for robustness checked using the most disadvantageous value of π for that scenario.

4.7 The Discount Rate

The Green Book (HM Treasury, 2003) finds a short term social time preference rate of 3.5%. A discount rate of 3.5% is therefore applied in the analysis. That is $\delta = 0.035$.

[22] Details are available at `http://www.hmrc.gov.uk/ctsa/ct\_rate\_band.htm`

4.8 The Time Delay

Estimates of the time delay, T, can be based on the experience of implementing marginal cost pricing policies in Australia and New Zealand. In the case of New Zealand, increased usage appears to be almost instantaneous. For the Australian Bureau of Statistics, looking at Figure 4.1 in Section 4.3 that the response occurred fairly rapidly, and a substantial effect was already apparent in 1-2 years.[23] In the case of the OSDM it is evident from Table 4.3.2 in Section 4.3 that the use of the free data grew substantially between its introduction in 2001-02 and 2005-06 (a 4 year period). Taken together these experiences would suggest that a value of T in the range of 0 to 3, with a mid-point value of 1.5, would not be unreasonable.

4.9 Costs

When analysing cost data from trading funds it will be important to make some distinctions. First, between costs arising due to activities related and unrelated to the collection or dissemination of data. Here, the interest is only in the former. Second, when looking at data related costs one must distinguish between:

- Costs related to the production of the data. This includes costs related to the collection of data and its maintenance over time.
- Costs related to the distribution of data.
- Costs related to sales and marketing or to the provision of value-added services.

When considering these basic categories several important questions arise. First, what sorts of activities get classed as value-added services rather than production or distribution? For example, should the provision of an online search facility by a

[23]Since no stabilisation of take up has yet been observed the time delay could be higher over the long-term, but T is measuring the time delay in first increased usage

trading fund be classed as related to the second category (distribution) or the third (sales etc)? This is important because expenditure in this third category is excluded from our analysis.[24]

The approach here, will be to take production and distribution costs as relating solely to the data (the nonrival good). Costs related to providing value-added services such as an online search facility will be excluded. Thus, distribution costs will only include the costs of providing the data in its most basic (and perhaps bulkiest form): crudely the cost of providing data dumps and making them available for download with reasonable ease and speed.[25]

A second important question arises in relation to trading funds who are collecting 'registration' data such as Companies House and the Land Registry. For these trading funds the data they hold comes primarily from legally required registrations – as compared with trading fund such as Ordnance Survey or the Met Office who clearly need to collect data themselves. Fees are charged to those registering and the income from these fees account for the majority of their income. In this case what is the appropriate way to take account of costs (and income) for data whose collection (production) is partially or largely paid for by its original 'producers'. In particular there is some ambiguity about how production costs should be computed. For example, suppose total costs of collecting and storing data were £50 million (including overheads) and registration fee income is £45 million. Should production costs, for the purposes of this analysis, then be computed as £50M or £5M? The first quantity is the true full cost of collecting, processing and maintaining the data, but, the second figure gives the correct net figure once registration fee income is taken account of. For the purposes of this analysis it would seem that it is this second figure which is the more relevant because the concern is with charging policy. Our

[24]Note that of course this does *not* mean that expenditures related to valued-added data or information is outside of the analysis.

[25]These costs would also include ensuring that the data changes were also made available on a prompt and regular basis as well as providing sufficient documentation for the data to be usable.

focus is on the charges (and costs) for providing access to data out, not for collecting it in. In particular, when considering, for example, a change from average cost to marginal cost pricing it is this second figure which gives the amount of government subsidy that would be required not the first one.

The third, and final, question to consider is how these categories relate to the fixed and marginal costs which are the key inputs needed for comparison of charging regimes. Here the situation should be fairly straightforward. The marginal cost is the expenditure incurred providing an extra copy of a piece of data to a user (with fixed costs essentially being everything else). As production costs do not vary with the number of users it is clear all production costs come into the fixed category. It is tempting to then assign all distribution costs to the marginal category but one needs to be careful. Many aspects of distribution involve fixed costs and display economies of scale. For example, buying an additional server to handle downloads, purchase of bandwidth or even bulk purchase of traditional mailing facilities usually involving declining average costs. Thus it may be necessary to break down distribution costs into fixed and marginal components. Regarding the third category, costs incurred in relation to sales and marketing and value-added services: these will be excluded from the cost computations since, as already mentioned above, these expenditures are not related to the basic production or provision of data. Thus, total fixed costs will be made up of production costs plus whatever proportion of distribution costs are determined to be fixed with marginal costs coming from the remaining marginal proportion.

Chapter 5

Empirics II: Trading Fund Summaries

5.1 Introduction

As will become apparent below, even among the five trading funds selected for analysis there is very substantial variation. For example, while some trading funds are primarily registration based, with the sale of information a more or less important side activity, other trading funds derive almost their entire revenue from the provision of information.

In addition, there was little consistency in the amount or format of data provided by trading funds to the study. Hence quite apart from the actual variation in structure and operation of each trading fund, each set of data requires special attention to address its particular strengths and weaknesses. Thus rather than pursue a 'one size fits all' approach each trading fund will be analyzed individually, with the overall results aggregated where possible and appropriate in the conclusion.

5.1.1 Which Products?

That said, before commencing the analysis on a trading fund by trading fund basis, it is worth noting certain general issues which are common to almost all of those studied here.

Most importantly, something should be said regarding the selection of products to be included in the analysis (and hence those to be excluded). For example, as already discussed above in section 4.9, it would seem best to exclude 'service' products from consideration. However the line can quickly blur, for while it is obvious that 'consultancy' would count as a service and not a data product what about a search service built on top of an underlying database of information? The central difficulty if one were to include such cases is that much of the cost providing this 'product' comes from the 'service' side (e.g. building and maintaining the search utility) and not from the data side (collection, storage etc).[1] The inclusion of such items without adequately distinguishing the two aspects would obviously substantially distort the results of any analysis – for example any calculation of subsidies needed under marginal cost pricing would now include the subsidy to the 'service' as well as for the basic data production. In order to avoid this, wherever this problem is encountered and the data is lacking (as is almost always the case), the basic approach is to exclude these products from the analysis.

A second, related, issue arises from the form of cost information available, in particular the general lack of marginal cost information (no trading fund was able to provide marginal cost information directly though in some cases it was possible to obtain information that would allow one to conjecture an approximate value). The solution to this problem adopted here is to focus on those products where a reasonably accurate figure for marginal cost can be obtained a priori. The obvious

[1] One might think that the refined/unrefined distinction could be one way of dealing with this issue but unfortunately this is isn't always the case. For example see the discussion of Companies House below.

case where this is possible are digital products which are supplied in a standardized (and perhaps bulk) form. In these cases it is reasonable to conjecture that the (basic[2]) additional cost of supplying another user is the cost of bandwidth, which in most cases is so small as to be taken as approximately zero.

A final case to consider is that where a trading fund is supplying a product consisting or using data which the trading fund does not 'control'.[3] Here the approach is simple: all such products will be excluded from the analysis. The reason is also simple: without such 'control' it is not clear that a given change in charging policy (whether to reduce or increase prices) could be implemented at all, and, even if it could, would involve detailed negotiations with 3rd parties regarding royalty rates, conditions on provision etc, which in turn would likely have a very substantial (but hard to predict) effect on the costs (and benefits) of the change.

To summarize, the focus will, in general, be on products that conform to the following criteria:

- Data only (no 'service' products or data products with a 'service' component).
- Digital (so no paper products).
- The trading fund has sufficient rights to dispose of the data how it wishes (so for example the product does not contain substantial 3rd party data).
- Products that are standardized or made available in reasonable bulk (hence per item additional overhead beyond basic distribution costs is minimal).
- Reasonably specific revenue information is available for that product (or product category) – otherwise no overall welfare figures cannot be calculated.

[2]This would exclude items such as support, special assistance and advice etc. as these are all 'service' items which one would expect a trading fund to continue to charge for even if the data itself had no (or a very low) charge.

[3]'Control' need not entail having collected the data, or being the IP owner, but means that the trading fund does have sufficient rights do do whatever it wishes with the data without having to seek permission or pay a 3rd party a fee.

Where a product (or set of products) are excluded from analysis by these criteria, a **conservative approach will be adopted in that it will be assumed that such products will remain under the current charging policy** ('conservative' therefore, in the sense of leaving things unchanged).

5.1.2 Costs vs. Revenue

Another general issue arises from the fact that for most of the trading funds studied average cost pricing does not occur at the product but only at the aggregate or division level. That is to say, for a given product, or even for a given product category, there is often quite a substantial difference between revenues and costs. This raises the question as to whether to use revenues or costs when doing calculations. Given the approach taken in section 3.4 when deriving the algebraic expressions for the welfare impact of regime changes in general revenues will be used rather than costs. This also has the advantage that revenues per product are, in general, known while costs frequently are not.

5.1.3 Data Distinctions

For the purposes of analysis each trading fund sought to classify relevant products both in terms of the raw versus value-added distinction and in terms of unrefined versus refined distinction. It should be emphasized that while every trading fund sought to do its best to apply the definitions provided, due to uncertainties and time limitations, any classification should be considered as preliminary and subject to future revision in the light of further discussion and evaluation. For the Ordnance Survey and Companies House in particular, the classifications have arisen out of informal exchanges with the authors and should be considered as the interpretation by the authors of these agencies' views, rather than the views of the agencies themselves.

At this point some general comments on the results of the classification will be provided with further specific remarks in the individual trading fund sections below. First, application of raw versus value-added was almost impossible due to variations and inconsistencies in usage. Key terms such as 'public task' were either ill-defined or even undefined. Even when known, trading funds often did not seem to interact this requirement in a consistent way with the additional requirement for 'rawness' – i.e. that data be in its most 'basic state'.

The unrefined versus refined distinction fared much better from a classificatory point of view though there were still some rather anomalous applications.[4] Rather, the issue here was that there were a large number of refined products for which sufficient data was unavailable to do an analysis. Thus, it was not possible to judge affect of applying a particular charging policy to all refined products. By contrast there was a reasonably good fit between the unrefined distinction and the set of products that can be analyzed (see previous sections). Hence a reasonable estimate of applying a change to unrefined products can be made.

5.1.4 The Form of the Analysis

As the form of the analysis takes a standard form it is worth laying it out once in general here to avoid repetition in each individual trading fund section. The first point to make is that the focus here (and in the sections below) will, in general be on a single comparison, that of average cost to marginal cost pricing. On what basis are the other two possible comparisons excluded?

The case of marginal cost to zero cost is the simpler. As discussed above, the

[4]While the OFT obviously had good reason for choosing the use of the term 'refined' and 'unrefined' it does have some rather unfortunate connotations. Specifically 'refining' has obvious suggestions of 'processing' or 'distilling'. But for trading fund data, while, in general one would expect 'unrefined' data to be fairly 'unprocessed' this need not necessarily be so (after all, almost all data has been processed to some degree to get it into a database). Given the availability of other terms such as 'upstream' and 'downstream' (or even 'contestable' or 'non-contestable') which appear to be as close (if not closer) to the OFT's intended meaning and which have less room for ambiguity it might be worth considering using them in future rather than 'unrefined' and 'refined'.

focus will be on products whose marginal costs can be known a priori, in particular (bulk) digital products whose marginal costs can be taken to be approximately zero. Hence, in this case, the marginal cost and zero cost regimes are in practice identical and there is no need to examine the zero price case.

The lack of attention to the comparison of profit maximization for a different set of reasons. First, for several of the trading funds it is highly unlikely that such a policy could be implemented due to existing legal and competition restrictions.[5] In particular, it would be hard to imagine that the trading funds, as public sector bodies, could pursue pricing strategies that would attract regulatory attention (and likely censure) in the private sector. Obviously this point only applies to those products where the trading fund had a monopoly (or near monopoly) (crudely, 'unrefined' products in the OFT distinction). However, given the selection criteria almost all products being analysed will be of this 'unrefined' type.[6] Second, and perhaps more importantly, the empirical difficulties involved in comparing profit-maximization with average cost are much more substantial than in comparing average cost and marginal cost. This is because when moving to marginal cost from average cost the new price is known (as the marginal cost is known) while when moving from average cost to profit maximization the new price is not known.[7] For this reason, given the data available, any estimation is likely to be so speculative as to have little policy-making value. Third, and finally, recall that where marginal cost dominates average cost (i.e. delivers higher welfare) then necessarily marginal cost also dominates profit maximization.[8] Thus if it transpires that marginal cost does dominate average cost the case of profit-maximization becomes redundant.

[5]For an example of an explicit legal restriction, as well as further discussion of this issue, see Chapter 6, especially the footnote on page 110.

[6]The service (and retail) type products, for example, which are more likely to be refined are being excluded for the reasons set out at length above.

[7]Furthermore, not only is it unknown but it arises from a complex optimization by the trading fund in response to both current (and possibly future demand if, say, demand expands in response to usage).

[8]For details of the proof see the subsection on profit-maximization in Section 3.4.

5.1.5 Average Cost to Marginal Cost

Moving on now to the actual analysis, for a given product, using the formula derived in Section 3.4, the net change in social value from switching from average cost to marginal cost is:

$$\Delta W = \theta F \left((1 + \alpha\beta)\delta^T \gamma - \alpha(1 - g)(1 - \beta) \right) \tag{5.1.1}$$

where, F is the absolute change in revenue, λ is the multiplier, p, q are original price and output, Δ indicate (absolute) change in variable moving from average cost to marginal cost, $1+\alpha$ is the marginal cost of public funds (which equals the reciprocal of the distributional weight $1/\theta$), and ϵ is the (absolute) value for the elasticity of demand.

Using the results of Chapter 4 one can substitute in for: θ (0.8) and $\alpha = 1/\theta - 1 = 0.25$, δ (1/1.035), T (1.5), τ (0.25) and π (0.3). Also as discussed in Chapter 4, α is the most precisely known of all the parameters with a range of 0.15 - 0.35 and a point estimate of 0.25. Furthermore, for products where the marginal cost is approximately zero, $\Delta p/p = 1$ and so $\gamma = \lambda\epsilon\Delta p/2p = \lambda\epsilon/2$. Substituting these values gives ΔW as:

$$\Delta W = 0.8F(0.97\frac{\lambda\epsilon}{2} - 0.23(1 - g))$$

And a marginal cost regime is superior if:

$$\lambda\epsilon \geq 0.48(1 - g)$$

Frequently, the proportion of revenue from government is 0 ($g = 0$) and then one has:

		λ			
		1	1-2	2-4	4-10
ϵ	0.0-0.5	AC	AC/MC (AC)	AC/MC (MC)	AC/MC (MC)
	0.5-1.5	MC	MC	MC	MC
	1.5-2.5	MC	MC	MC	MC

Table 5.1: Preferred charging regimes in different parameter ranges. AC indicates that average cost pricing is defensible across both ranges. MC indicates that marginal cost pricing is preferable across both ranges. AC/MC indicates that average cost and marginal cost pricing is preferable, but in different parts of the range (the figure in brackets indicates what is preferable at the mid-points of both ranges

$$\lambda\epsilon \geq 0.48 \approx 0.5$$

Thus deciding on the preferable charging regime for a given product reduces to deciding on the value of λ (multiplier) and ϵ (elasticity of demand) appropriate to that product. Here there are 3 broad categories for each of λ and ϵ: low, medium and high and each product is assigned one of these ranges (independently – that is the high range for ϵ can be chosen while the low range for λ can be chosen). Table 5.1 shows how the preferable charging regime will vary according to the categories assigned to a product with marginal costs approximately equal to 0 (and $g \approx 0$). An extra column has been added to show the preferable regime using the minimum value of $\lambda = 1$.

It should be noted that if one uses the upper bound of $\alpha = 0.39$, then marginal cost pricing is only preferable when $\lambda\epsilon > 0.78$. That is average cost pricing may now be preferable within the ranges of λ from 1 to 2 and ϵ from 0.5 to 1.5, but only at the lower ends. On using the mid-point values of these ranges marginal cost pricing is still preferable. Thus it is likely that the results will continue to hold for most of the range.[9]

[9]The only change in preferable regime, at the mid-point estimates, when using $\alpha = 0.39$ is when λ varies from 2 to 4 and ϵ varies from 0 to 0.5. That is, at the mid-points of these ranges, average cost pricing is preferable to marginal cost pricing.

Robustness Checking

As discussed above in the sections on the multiplier (λ) (and to a lesser extent the proportion of surplus going to producers: π), due to the lack of data, there is necessarily very significant uncertainty over the appropriate ranges to use for any given product. Thus, one would be wary of any results indicating a change of regime which depended purely on the value of λ chosen. To guard against this, wherever the welfare calculation indicates a change in regime away from the status quo (specifically from average cost to marginal cost) a robustness check will be performed whereby the same calculations will be done but using the minimum value of $\lambda = 1$ (and the minimum value for $\pi = 0$). As can be checked by running through the algebra with $\lambda = 1, \pi = 0$ (leaving other variables unchanged) the change in welfare moving from average cost to marginal cost is:

$$\Delta W = 0.8F(\frac{\epsilon}{2} - \alpha(1 - g))$$

A marginal cost regime will thus be superior if:

$$\epsilon \geq 2(1 - g)$$

In the welfare tables below this check will be referred to as the 'Robustness Check' (RC).

5.1.6 Registration-Based Trading Funds

At this point it would be appropriate to note that trading funds with registration fees call into question the basic approach set out above for comparison of different charging regimes. Specifically in the preceding analysis it has been assumed that revenue shortfalls arising from a change to a marginal cost (or zero cost) pricing regime would made up by central government (via a subsidy). However for these

sorts of trading funds one could instead assume that any deficit was made up by an increase in their registration fees. Demand for registrations is likely to be very inelastic (at least relative to the demand for data) – especially if registration is a statutory requirement. Hence, raising fees may well be a much more efficient way of covering any shortfall than using general government tax revenue.

This argument has implications about existing charging policies at such trading funds: from a basic Ramsey pricing perspective markups over marginal cost should be proportional to the inverse of the elasticity. If, as just suggested, registrations are very inelastic (and therefore likely to much more inelastic than the demand for data) then costs should be recovered primarily by marking up registration fees (over marginal costs) with the markups on data sales kept very low (so prices are close to marginal cost). Thus, for registration-based trading funds, marginal cost pricing for data is likely to be an preferable to average-cost for simple Ramsey pricing reasons without any need to assume (and calculate the cost and benefits) of subsidies from central government.[10]

5.1.7 Outcomes of Interest

The three main outcome variables that will be reported in the analysis of each trading fund are:

- ΔB: change in producer plus consumer surplus. As these will always be positive when moving from average cost to marginal cost these are also referred to as the 'gross benefits'.

- ΔG: change in net government income. This will be equal to minus the total subsidy provided plus the gain in tax revenue levied on the increase in producer surplus. Normally one will expect ΔG to be negative.

[10]It is noteworthy that this type of pricing rule does *not* appear to have been adopted at the registration-based trading funds examined in this report – in fact, if anything the markups over marginal for data products are well above those for registrations.

- ΔW: the net change in social welfare. As all values are expressed in terms of the same numeraire (uncommitted public funds) this will simply equal $\Delta B + \Delta G$.

Algebraic expressions for each of these variables can be found in Section 3.4. It should be emphasized again that all values are expressed in terms of same numeraire, namely uncommitted public funds.

5.2 Companies House

Companies House is the official government register for UK companies. As such it falls squarely within the registration category of trading funds with the bulk of its income (and costs) arising from its registration activities as can be seen in Table 5.2). The data provided by Companies House to the study was the best of any trading fund with full price/quantity information on all products and a detailed breakdown of costs both by activity and by product area (there was however no marginal cost information).

Activity	**Revenue**	**Cost**	**Surplus**
Registrations	57829	57689	140
Search	12698	11802	896
Others	1694	1405	289
Total	72220	70895	1325
– Of Which Government (search products only)	302	-	-

Table 5.2: Companies House Summary for Year 2006/2007 (in £000s).

5.2.1 Data

The main challenge then with this dataset was to decide which products to include in the analysis (see discussion above). The bulk of Companies House's income (and costs) are incurred in relation to registration activities. These would clearly be

excluded from the analysis. Focusing on the 'search' area alone the breakdown by category is given in Table 5.3.

Subcategory	Type	**Revenue**	**Cost**	**Surplus**
CONFIDENTIAL				

Table 5.3: Break-down of search category into subcategories (in £000s). Type: S=Service, DD=Digital Data, ND=Non-Digital.

Clearly some of these categories have both data and service (search/physical facilities/etc) components: Companies House Direct, WebCheck, Companies House Information Centres (CHIC), Companies House Contact Centre (CEU), Certificates and Copies (physical), CH Monitor. For the reasons already discussed above, most importantly uncertainty as to the cost division between data and service and the associated variation (and uncertainty) in marginal costs and elasticities across products, these categories were excluded from the analysis in order to focus solely on the 'CD Rom' and 'Bulk Data and Image' products.[11] Here the marginal costs of supply can be presumed to be close to zero and there is much greater homogeneity within product categories.

Table 5.4 provides breakdowns of prices, sales and revenues for the products selected for analysis. These products selected for analysis accounted for £926,000 in revenue and £653,000 in costs (out of a total on 'Search' of 12.7m in revenue and 11.8m in costs). Note that several 'bulk' products are no longer listed in the catalogue due to the lack of demand (they are now 'bespoke' products). **... CONFIDENTIAL ...** .

As the table shows, all of the products selected were classified as 'Raw' and 'Unrefined' by Companies House. One might have thought that the natural re-

[11]Recall that cost information is only provided at the product category level so with product variation it is harder to assign costs accurately. Further issues include the fact that CHD combines has a nonlinear pricing structure (fixed registration fee plus per item charges). It is also noteworthy that the CHD category contains around 39 different 'products' while Bulk Data (Mag Tape) and Bulk Image only contain 10 each.

tail/wholesale division between Companies House Direct and 'Bulk Data and Image' would translate neatly to the OFT's refined/unrefined distinction, especially given that that all information available in the CHD products could be produced via extraction from their 'Bulk Data and Image' material (Companies House also stated that they allow free reuse and redistribution of such information). Nevertheless, all of their 'search' products were classified as unrefined because of Companies House's status as a sole source for this information. Given the wording, and clear intent, of the OFT distinction this is perhaps surprising. However, as 'retail' products have already been excluded from direct consideration this is not an issue that merits further debate here, though it would perhaps be something that the government would like to discuss further were such a distinction to be adopted.

5.2.2 Analysis

The basic form of the analysis has already been laid out in detail in Section 5.1.4. What remains is to set values for g, the proportion of government revenue, and for the elasticity of demand: ϵ, and the multiplier λ. For g, the 'conservative' assumption has been made that all of government expenditure on Companies House search products was on those with a 'service' component, that is the expenditure on the 'pure' data products was zero (so $g = 0$).[12]

Turning to the elasticity and the multiplier, for all products other than CDROMs, the 'high' range for elasticity of demand (ϵ) and the 'medium' range for the multiplier have been used. The reasoning behind this is that the usage of all of these products is currently very low – and the price high.[13] Furthermore, it is clear that not only is there significant interest in accessing this data, but its nature clearly offers significant scope for reuse, whether recombined with other datasets, incorporated in

[12]In fact Companies House does make some 'bulk' sales both to the HMRC and to ONS but no precise figures were available so the 'conservative' approach has been taken.

[13]For example, almost all the Bulk products have sales in single figures, and for several items only one or two copies are sold a year while prices are in the thousands or tens of thousands.

Product	RW/VA	URF/RF	Revenue	Sales	List Price	PPM	SPY
CONFIDENTIAL							

Table 5.4: Companies House prices, sales and revenues for 2006/2007. List price: from Companies House website (generally per unit/per annum). PPM: price per month. Sales: total sales per month over the year (the data available provides sales on a monthly basis with the associated per month prices). Revenue: total revenue for the year. SPY: sales per year (a rough estimate obtained by dividing revenue by per unit annual price). 'Mag Tape' items are now, in fact, provided by ftp download.

other material (reports, analyses etc) or simply re-presented in a more convenient form. Thus, it seems plausible that a move to marginal costs (here approximately equal to zero) would result in a substantial increase in demand and thus the high range for the elasticity has been assigned. Similarly on the multiplier side, because of the scope for reuse and recombination the medium range for the multiplier has been assigned. Finally for CDROMs, because both of the products listed are similar (often containing the same data but with some restrictions) to other Bulk products it seems reasonable to think their elasticities would be lower. Furthermore, one would imagine most reusers would prefer ftp access to a CDROM and hence the scope for reuse of the CDROM products is limited. Thus, the low category for the multiplier has been assigned.

The assignments for ϵ and λ, along with the resulting values for the outcome variables, are shown in Table 5.5. With the values used, for all of the products under analysis, a marginal cost regime would be preferable, and this result is robust to the usual checking (see Section 5.1.4 for more on robustness).

5.2.3 Summary

The analysis focused on Companies House's 'bulk' ('wholesale') data, specifically their 'CD ROM' and 'Bulk Data and Image' product categories, and for all other products it was assumed that the pricing regime would remain unchanged.

Focusing on these two categories, the results suggested that under the likely range of parameter values a change from an average cost to a marginal cost regime would be welfare improving. Specifically, the calculations showed that adopting a marginal cost pricing policy for these particular products would result in gross benefits of around £2.6m with government incurring net costs of around £681k. This figure is calculated on the basis that Government would make up the loss in revenue of 946k

Product	Rev	g	ϵ	λ	ΔB	ΔG	ΔW	AC/MC	RC
Dir - Full snapshot Mag Tape	XXX	0.0	2.0	3.0	64942	-16218	48724	MC	True
Dir Daily updates Mag Tape	XXX	0.0	2.0	3.0	535489	-133727	401762	MC	True
Dissolved Mag Tape	XXX	0.0	2.0	3.0	> 0	0	> 0	MC	True
Mortgage Register Mag Tape	XXX	0.0	2.0	3.0	> 0	0	> 0	MC	True
Directors Reg Dly update Mag tape	XXX	0.0	2.0	3.0	535489	-133727	401762	MC	True
Disqual Orders Mag tape	XXX	0.0	2.0	3.0	3418	-853	2564	MC	True
Dir Full Snapshot Mag Tape	XXX	0.0	2.0	3.0	79753	-19916	59836	MC	True
Co Supp To Lon Gaz Mag Tape	XXX	0.0	2.0	3.0	8545	-2133	6411	MC	True
Bulk Image Annual Subscription	XXX	0.0	2.0	3.0	36792	-9188	27604	MC	True
Bulk Image Annual Accounts	XXX	0.0	2.0	3.0	482510	-120496	362013	MC	True
Bulk Image Annual returns	XXX	0.0	2.0	3.0	510993	-127610	383383	MC	True
Bulk Image Form 288	XXX	0.0	2.0	3.0	136720	-34143	102577	MC	True
Bulk Image Form 287	XXX	0.0	2.0	3.0	17090	-4267	12822	MC	True
Bulk Image New Incorps	XXX	0.0	2.0	3.0	22786	-5690	17096	MC	True
Bulk Image Mortgage	XXX	0.0	2.0	3.0	14241	-3556	10685	MC	True
Bulk Image capital	XXX	0.0	2.0	3.0	19226	-4801	14424	MC	True
Bulk Image Liquidation	XXX	0.0	2.0	3.0	22786	-5690	17096	MC	True
Bulk Image Miscellaneous	XXX	0.0	2.0	3.0	45573	-11381	34192	MC	True
Bulk Image Change of Name	XXX	0.0	2.0	3.0	11393	-2845	8548	MC	True
Cd ROM Directory Subs	XXX	0.0	1.0	1.5	62948	-43300	19648	MC	True
CD Rom Change of Name/Dissolved	XXX	0.0	1.0	1.5	2090	-1438	652	MC	True
Total	945797				2612793	-680986	1931806		

Table 5.5: Companies House Analysis. Rev: revenue, g: proportion of revenue from sales to government, ϵ: elasticity of demand, λ: multiplier, ΔB: change in consumer and producer surplus (excluding government), ΔG: change in government income, ΔW: change in overall social welfare, AC/MC: which regime (average cost or marginal cost) is preferable, RC: robustness check (True=Robust, False=Not Robust). XXX indicates that the contents of a cell have been removed for confidentiality reasons.

suffered by Companies House itself.[14] However given the registration-based nature of Companies House it is feasible that this shortfall could be covered by the registration side of its operations.[15] Being conservative and leaving aside this possibility for the time being, and using the figures for benefits and costs just enumerated, the overall net welfare gain to society from the change envisaged would be approximately £1.9m.

5.3 The Met Office

5.3.1 Introduction

The Met Office provides weather and climate change forecasts for the UK and worldwide. In contrast Companies House, which is a registry-type trading fund, the Met Office is firmly within the data collection and provision category. Table 5.6 below summarizes the sources of Met Office revenues (and some costs). As that shows, one striking features of the Met Office compared to the other trading funds studied here, is that it receives a substantial direct subsidy[16] to support the PWS (Public Weather Service).[17] This subsidy is jointly provided by the Government and the Civil Avia-

[14]Note that for these product categories Companies House's costs here were actually substantially lower than revenues (653k versus 946k). Thus if Government were simply to make up costs rather than revenues the figure would be lower.

[15]Since these would be costs incurred in maintaining and created the registration database it is likely that there would not be any conflict with the existing legal restrictions on Companies House charging policy (see footnote on page 110) – though this is something that might need to be investigated further.

[16]The Met Office have indicated that it is their belief that, since becoming a Trading Fund, they have not been subsidised to undertake any activity but are simply paid by our customers to provide products and services. They therefore argue that they are 'paid', not 'subsidised', by the Government, (which is represented by the PWS Customer Group) to provide the PWS. This seems a rather semantic distinction – after all almost all 'subsidies' from Government are used either to pay for, or reduce the price of, some good or other. Thus, the term subsidy has been retained both in the discussion here and below.

[17]According to the Met Office the PWS provides "a set of services for the benefit of the public, including basic weather forecasts, reports about significant weather events, national severe weather warnings, and information for non-commercial re-use through a customer centre and the national met' library and archive." As such it obviously includes substantial data gathering and processing activity.

tion Authority (CAA) with the Government providing the lion's share. Combining this subsidy with other sources of revenue from government one finds that of total revenue in 2006/2007 of £168m, £125m (75%) came from government.[18]

Revenue Area	Revenue	Percentage of Total	Costs
CONFIDENTIAL			

Table 5.6: Met Office Financial Summary for Year 2006/2007 (in £000s).

The Met Office is unusual, and noteworthy among trading funds in operating a form of internal pricing via a clear division of their wholesale and retail arms. In particular the Met Office's 'retail' (or 'value-added') arm, which sells to both Government and business, explicitly buys data from its wholesale division on the same terms as any other market participant. Thus in the figures presented in Table 5.6 'Wholesale' revenues are broken down into 'external' and 'internal' with 'internal' sales being to the Met Office 'retail' arm and 'external' sales being to other businesses.

5.3.2 Data

For the purposes of analysis the two areas of interest would be 'Wholesale external' and 'Commercial Retail'. Unfortunately, the only information available regarding 'Commercial Retail' beyond the totals shown in the table was a breakdown of revenues by customer type (e.g. 'transport', 'construction' etc). Moreover, as discussed extensively above, there are good reasons to exclude this kind of area from analysis and focus purely on the more 'bulk', pure-data, products.[19]

[18]Other figures supplied put total revenues at the slightly higher figure of 171m. This was close, but not equal to the figure obtained by totalling up the figures from the full breakdown. Since it would be desirable for the figures to reconcile the second, derived, total of 168M has been preferred.

[19]In particular here there is the non-standard/bespoke nature of the products supplied, the substantial 'service' component, the complete lack of cost data, no product disaggregation etc. Note also that, in this case, the division fairly neatly follows the 'raw/value-added' and 'unrefined/refined' divide for the Met Office has classified all of its 'Commercial Retail' activity as 'Value-added/Refined' while all 'Wholesale' products are 'raw' and most are 'unrefined'.

Product Category	Type	RW/VA	RF/URF	Rev	Cust	Description
CONFIDENTIAL						

Table 5.7: Met Office Wholesale Product Category Information for Year 2006/2007 (in £000s). Type = Data Type (DD = Digital Data, S = Service, 3P = 3rd Party). Rev = Revenue (in £000s). Cust = Number of Customers. Customer numbers are actually from 2007/2008 as no detailed breakdown was provided for 2006/2007. However total number of customers (excluding 3rd party data) were 40 in 2006/2007 compared to 43 in 2007/2008 so the difference is likely to be small. More detail on the products in each category can be found in the online price list at `http://www.ecomet.eu/members/UK.htm`.

Focusing on wholesale area alone, Table 5.7 below provides details of wholesale product categories available. As the reader will note there are no prices and no quantities (just customer numbers – with some customers purchasing in multiple categories). The reasons for this are that a) in some categories there are several products b) sales are only recorded by the Met Office as per customer totals and not per-product.[20] Since customers may buy several products and there are bulk discounts the relation of revenues to prices and quantities is not a direct one. Given this the best way to proceed is simply to present the basic data available – that is revenue and customer numbers. For those interested list prices for products in each of the data product categories listed can be found online at `http://www.ecomet.eu/members/UK.htm`.

Looking within the set of wholesale product categories there are three distinct types: Digital Data, Services, and 3rd Party Data. The first two categories are self-explanatory. The third, as indicated by the descriptions in the table, consists of revenue from the sales of data performed by the Met Office but which is sourced from other European weather offices or organizations. For the reasons discussed above both the 'service' type and the 3rd party type products will be excluded from the analysis leaving the pure 'digital-data' bulk products only.

This exclusion of some product categories presents some difficulties because cost information was only available at the level of the entire wholesale operation (including internal data). The basic cost breakdown is presented in Table 5.8. In terms of a split between fixed and marginal costs both the information in this table and discussion with the Met Office indicate that practically all of these costs could be regarded as fixed and that marginal costs were so low as to be approximately zero.[21]

[20]The product category breakdowns available are only possible thanks to the efforts of the Met Office in hand analyzing recent invoices.

[21]Specifically, it was indicated that, in terms of both staff and infrastructure, usage could at least double before any changes would be required. The also stated that regarding bandwidth total usage was around 300GB per day of which around 11GB (2.7%) was due to wholesale operations. At this level of usage bandwidth costs are likely to be under 15p/GB and storage costs would be around

Costs Area	2005	2006	Comments
CONFIDENTIAL			

Table 5.8: Met Office Cost Breakdown for 2005/2006 and 2006/2007 (in £000s).

Given the interest in only one of the three product category types it is useful to also apportion costs in the same way. To do this a simple approach has been taken whereby: a) all 3rd party costs have been allocated to 3rd party data b) remaining costs have been allocated in proportion to the number of contracts in that area (excluding 3rd party contracts). The exact breakdown is shown in Table 5.9 below.

	2006	Comments
CONFIDENTIAL		

Table 5.9: Allocation of Met Office Costs for 2006/2007 (in £000s). Number of contracts in each area extrapolated from 2007/2008 data but difference is unlikely to be large as total external contracts (excluding 3rd party) changed little between the two years (40 vs. 43).

5.3.3 Analysis

The basic form of the analysis has already been laid out in detail in Section 5.1.4. What remains is to set values for g, the proportion of revenue from government, and for the elasticity of demand: ϵ, and the multiplier λ.

For the data under consideration it is clear that g is zero. Turning to the elasticity, comparison with the kinds of PSI goods consider in Section 4.3 suggests that the medium or 'high' range for elasticity of demand would be appropriate. Further taking into account the low levels of current demand and the relative complexity of obtaining data, it seems reasonable to choose the 'high' alternative in most cases (for NWP, its refined status, and the higher number of existing customers has militated

the same level (e.g. Amazon's 3S storage solution offers rates of $0.18/GB or less, with storage costing $0.15/GB). This implies costs from bandwidth usage of a few pounds a day (of course this may exclude setup costs though these are likely to be fixed than marginal).

for selecting the medium range only).

For the multiplier, the greater uncertainty makes a decision more difficult. Weather and climate related data clearly have very substantial scope for reuse (and redistribution – given the size of the datasets). Furthermore, some of activities based on access to this sort of data, such as climate modelling, storm prediction and notification, clearly have associated positive externalities which are large.[22] Thus it seems appropriate here to take the medium, or even the high, range for the multiplier. Without hard data, it is hard to say which is the more appropriate but given the wide range of usages and the possible externalities it seems not unreasonable to apply the high range for the majority of the products under consideration (again for the NWP a lower range is assigned because of its refined nature).

These assignments for ϵ and λ, along with the resulting values for the outcome variables, are shown in Table 5.10. With the values used, for all of the products under analysis, a marginal cost regime would be preferable, and this result is robust to the usual checking (see Section 5.1.4 for more on robustness).

It should be noted that these results ignore any impact of a change in pricing regime on costs. Increased demand might necessitate some increase in capacity though, as discussed, marginal costs are very near zero. On the other side of the equation the move to marginal cost (or zero cost) pricing might reduce some overhead costs both due to general efficiency improvements and, for example, less need for management and monitoring of sales activities (the cost data shown above suggest that a non-trivial proportion of costs arose directly from the 'transaction' costs of monitoring and managing customer relationships).

Also, left to one side, is the possibility that cheaper access to wholesale data might result in greater competition for the Met Office's retail arm resulting in lost profits. When asked, the Met Office put the likely loss at 'somewhere between

[22]One reason, perhaps, why much weather-related information is already made freely *to consumers* – and why government already sees fit to provide substantial subsidies to the Met Office.

Product	**Rev**	g	ϵ	λ	ΔB	ΔG	ΔW	**AC/MC**	**RC**
Radar data	109	0.0	2.0	7.0	616	-46	570	MC	True
Surface data	10	0.0	2.0	7.0	56	-4	52	MC	True
NWP data	237	0.0	1.0	3.0	425	-193	231	MC	True
Climat data	23	0.0	2.0	7.0	130	-9	120	MC	True
Lightning data	11	0.0	2.0	7.0	62	-4	57	MC	True
Total	390				1291	-259	1031		

Table 5.10: Met Office Analysis (in £000s). For meaning of headings see caption to analogous Companies House Table in section 5.5. Note that all products are raw and unrefined except for NWP data which is value-added and refined.

200,000 and 3 million pounds'. This of course would not necessarily affect the net social welfare position since this loss to the Met Office would be likely be more than offset by gains to other producers and the general consumer (in fact the larger the losses to the Met Office the greater the steps being taken by private business to use and reuse the wholesale data implying a higher value for both ϵ and λ). Thus, given the uncertainties in the likely impact on 'Retail' and the indirect consequences for external users and producers, it seemed reasonable to simply leave this issue to one side.

5.3.4 Summary

Given the data available the analysis has focused on the Met Office's 'wholesale' products.[23] The focus was further narrowed down by the exclusion of service products and those which consist primarily (or entirely) of third party data (see Section 5.1.1 for the reasoning behind this). While only accounting for a very small proportion of overall costs and revenue (1% or below) these wholesale data products make available almost all of the basic data the Met Office collects itself.

For this set of products, the results of the analysis suggested that under the likely

[23]The Met Office is unusual among trading funds in explicitly dividing its upstream (wholesale) arm from its downstream (retail) arm with retail purchasing from wholesale on the same terms as any other user.

range of parameter values a change from average cost to marginal cost regime would be welfare improving. In particular, gross benefits would be around £1.2m with costs to government of only around £260k. The actual additional payment required to the Met Office from government would be equal to the loss in revenue which in this case is £390k. This would be on top of the existing £68m in subsidy (payment for the PWS) and £57m in sales coming from government. Thus the change in policy would require a 0.3% increase in current expenditure by government in this area.

Finally, putting together the benefits and the costs implies an overall net benefit to society of £1.03m and a return on investment of approximately 500%.[24]

5.4 Ordnance Survey

5.4.1 Introduction

Ordnance Survey (OS) is Great Britain's national mapping agency and it provides a wide array of geographic information services and products. Like the Met Office the Ordnance Survey is firmly within the data collection and provision category in contrast to other trading funds studied, such as Companies House or the Land Registry. As Table 5.11 shows, Ordnance Survey has total revenues of around £114m in 2006/2007 which makes it the second largest trading in terms of data provision after the Met Office. Unlike the Met Office the Ordnance Survey does not receive any direct subsidy from Government though its indirect revenue from Government via sales is quite substantial amounting roughly to around half of its total income. For the Met Office the figure is around 74% and so overall the Ordnance Survey had higher income from outside of government than the Met Office (£58m vs. £46m)

[24] Adding in the loss of ECOMET royalties – and assuming no corresponding benefit to society since these gains accrue to non-UK citizens elsewhere in the EU – would increase government costs by 56k. Subtracting out the NWP product to leave only unrefined items would reduce government costs by 193k to only 66k but would also reduce gross benefits to 866k and net gains to 800k.

	Revenue	% Public Sector
Large Scale Topo	66097	58
Transport Network Products	3263	53
Mid & Small Scale Digital	**CONFIDENTIAL**	19
1:10k Products	**CONFIDENTIAL**	54
Mid & Small Scale Digital	**CONFIDENTIAL**	19
Mid & Small Scale Consumer Mapping	**CONFIDENTIAL**	8
Address	**CONFIDENTIAL**	42
Imagery & Height	**CONFIDENTIAL**	72
Other Products & Services	**CONFIDENTIAL**	43
Total	113639	49

Table 5.11: Ordnance Survey Financial Summary for Year 2006/2007 (in £000s). '% Public Sector' gives the percentage of revenue coming from the Public Sector.

and at a similar level to the UKHO (though as discussed below the UKHO has a rather more complex relationship to its data than either the Met Office or Ordnance Survey).

5.4.2 Data

For the purposes of analysis there were some important limitations to the data obtained from OS.

- In most cases, there was no revenue breakdown beyond the category levels given in Table 5.11.
- No quantity information (and without a product breakdown this would be of limited value).
- While a relatively detailed breakdown of costs by activity (e.g. data collection, IT, etc.) was provided most costs could not be allocated to a given product or product category. To a great extent this was to be expected – many products are derived from a few central databases and there are substantial economies of scope in data collection.

- No direct information on marginal costs was available. That said, for most digital products a zero marginal cost could be assumed a priori without much difficulty – as Ordnance Survey themselves suggested.[25]

Table 5.12, provides information on the products associated with each OS's categories. As should be apparent, in most cases there is a fairly large number of products within a given category with a fairly substantial heterogeneity. Thus, it would have been useful to have a revenue breakdown beyond the category level. Unfortunately however, either such a breakdown was not available or there was not sufficient time for it to be provided.

This has several consequences. Most obviously, without a product breakdown, quantity information has little value – even if it were provided. In any case, the nature of Ordnance Survey's sales strategy would have made it difficult to get one uniform 'quantity' figure for each product. Specifically Ordnance Survey tends to sell a given product via one of three distinct channels:

- 'Direct': Framework Direct Licenses (licenses for direct internal business use to end user businesses) and income from sales of retail products such as paper maps.

- Collective Purchase Agreements (CPAs): this is where a particular group of organizations club together to purchase access to various datasets in exchange for a fixed fee (though perhaps with annual payments). This category includes items such as the Pan-Govt Agreement, the Mapping Service Agreement etc.

- Partners: Revenue from those who reuse data in their own products and services via Specific Use Contracts (SUCs).

[25]According to Ordnance Survey 'The cost for distributing data to a single "marginal" user might be negligible.' Of course as OS also pointed out the fixed costs of distribution, might be substantial, but this would not affect the marginal costs.

	Type	RW/VA	R/UR	Description (Comments)
Large Scale Topo	DD	RW	UR	Land-line / OSMasterMap Topography Layer (replaces Land-line)
1:10k Products	3P/DD	VA	R	1:10,000 Scale Raster / OS Street View / OS Landplan Data / Map Return Scheme / OS Locator
Mid & Small Scale Digital	DD	RW/VA	UR/R	1:25,000 Scale Colour Raster (Unrefined) / 1:50,000 Scale Colour Raster (Unrefined) / Gazetteer Data / 1:250,000 Scale Colour Raster / 1:250,000 Data / Historical Data / Meridian 2 / Miniscale Products / Strategi
Mid & Small Scale Consumer Mapping	3P/ND	RW/VA	R	OS Explorer Map / GB Routeplanner / Guides & Wall Maps / Historical Maps / OS Landranger Map / OS Select / Road Maps / Tourist Maps
Transport Network Products	DD	RW	UR	Oscar / OS MasterMap Iintegrated Transport Network Layer (replaces Oscar)
Address	3P/DD	RW	UR/UR	ADDRESS-POINT / OS MasterMap Address Layer Products
Imagery & Height	3P/DD	VA	R	Aerial Photographs / Landform Products / OS MasterMap Imagery Layer / Pictometry
Other Products & Services	3P/S/DD	RW/VA	UR/R	Boundary Products / Code-Point / OSMM Boundary Layer / Copyright Licences / Joint Venture Settlements / Consultancy / Reprographic Services / Surveying Services / Developer Programme Licence / Contract Carto / Positional Services / OS Sitemap / Siteplan / Superplan

Table 5.12: Ordnance Survey Category Information. DD refers to digital data, ND to non-digital, 3P to products containing 3rd Party Data and S to services. R refers to refined products, UR to unrefined as classified by the OS (where R and UR are both used it indicates that there are products of both types in that category. NB: OS wished it to be particularly emphasized that their discussions with OFT in regard of the exact classification for their products were still ongoing and hence that any classification should be taken as preliminary and subject to future revision.

Total Revenue	Direct	CPAs	Partners	Total
CONFIDENTIAL				

Table 5.13: Revenue by Data Channel for the Ordnance Survey (in £000s)

Table 5.13 details the breakdown in revenues from each channel. For all categories[26] it would, in theory, be possible to have quantity information, however it could not be obtained, at least for this report – again, either because systems were not designed to record this information or because Ordnance Survey were not able to provide the information in the time available. Table 5.14 does list price information provided via Framework Direct Licences on some of representative products in the digital category while Table 5.15 gives price information on one of the twelve SUCs available prior to December 2007[27] (that related to navigational products). The main thing to take from this is that, at least, for some products prices are reasonably substantial, and certainly on a par with some of the examples considered in Section 4.3.[28]

Also noteworthy is that price options, especially for SUCs, are relatively complex. Not only are twelve different SUCs, each tailored for a particular market, but within each of these there are a wide variety of products. One would imagine that this must have some increase in costs both for potential reusers and for the OS itself (keeping the relative prices across SUCs consistent, recording demand etc.[29]

[26]One might think that CPAs allowed unrestricted usage and hence that for those contracts quantity information would not be relevant. However OS have advised us that this is not the case and quantity restrictions can apply to CPAs just as for other categories.

[27]As of December 2007 there are fourteen SUCs.

[28]For example Land-Form PROFILE contour data (a height product) costs around £43k for the whole of the UK and Boundary-Line costs around £7.1k for the same area. For comparison New Zealand's boundary product mentioned in section 4.3 cost NZ$25k – approximately £9.6k at the current exchange rate. (It should be emphasized that this comparison only applies to price and is not intended to suggest any further similarity between the products).

[29]Because each SUC was limited to a 'specific use' this could cause problems when a new customer arrives whose intended activity did not fit well within the existing categories. This issue was mentioned several times in discussions with individuals and organizations outside of the OS. Clearly dealing with these sorts of issues would also add to the general transaction costs on all sides.

Product	Item	Price
OS MasterMap Topography Layer	Low Density Tiles (0-140 TOIDS per Km2)	1.2
OS MasterMap Topography Layer	Medium Density Tiles (141-5,300 TOIDS per Km2)	18.35
OS MasterMap Topography Layer	High Density Tiles (5,301+ TOIDS per Km2)	120.25
Land-Form PROFILE Digital Terrain Model	Great Britain National Cover	43820.0
Land-Form PROFILE Digital Terrain Model	Tile (5 Km x 5 Km)	4.2
Land-Form PROFILE Contour Data	Great Britain National Cover	43820.0
Land-Form PROFILE Contour Data	Tile (5 Km x 5 Km)	4.2
Meridian 2 Full Dataset	Great Britain National Cover	17548.0
Meridian 2 Communications theme	Great Britain National Cover	13161.0
Meridian 2 Topographic theme	Great Britain National Cover	4387.0
Strategi	Great Britain National Cover	5852.75

Table 5.14: Selected Ordnance Survey Framework Direct Prices. Prices are for a Corporate User, Single year licence

	Perpetual Licenses	**Annual Subscription**	**On-demand Services**	**On-demand Services**
CONFIDENTIAL				

Table 5.15: Ordnance Survey License Fees for SUC 'Navigational Products'

Returning to Table 5.12, categories been classified as Raw or Value Added. The Ordnance Survey considered raw data to be synonymous with the data it was required to create as part of its public task.[30] However one should note that the definition of value-added used in this report required 'raw' information to be *both* produced as part of a public-task and to have had minimal further processing. Ordnance Survey's approach, by contrast, is based on their understanding of the Treasury definition of 'raw' data as "information that is central to Government's core responsibilities *processed to the extent that Government use requires*" [emphasis added]. Ordnance Survey therefore considered that the datasets in its Public Task document are processed to that extent only and are therefore are 'raw data'. Whether this is the correct interpretation is not a matter to be decided here – it does appear to differ from the interpretation from some other Trading Funds but this serves more to illustrate the lack of clarify in relation to the definition than the merit of otherwise of the Ordnance Survey's approach.

Categories have also been classified as refined or unrefined based on informal discussions on what could be considered upstream data.[31] In some cases the category includes products of both types. For example this is the case for Mid & Small Scale Digital and for Other Products & Services. One implication is that, given the absence of per product information, it would not be possible to distinguish between refined and unrefined products when performing welfare calculations.

[30]Specifically the OS consider the following products as part of their public task, and so could be considered raw: OS MasterMap Topography Layer; Land-Line; OS Sitemap; Strategi; OS MasterMap Integrated Transport Network Layer; OS MasterMap Address Layer 2; OS MasterMap Address Layer; ADDRESS-POINT; Land-Form PROFILE Plus; Land-Form PROFILE; Land-Form PANORAMA; 1:25,000 Scale Colour Raster; 1:50,000 Scale Colour Raster; OS Explorer Map (1:25,000); OS Landranger Map (1:50,000); 1:250,000 Scale Colour Raster; Boundary-Line. All other products would then be value-added.

[31]One should also note that were the OS to provide further direct access to its basic databases (beyond what it already provides), it is likely that more of OS's products would be classified as refined (as the data would now be available directly at a more basic level). It should also be noted that the OS is currently in the process of subsuming the data in DDMS, CMS and Boundary Line Production System into a new database, called MAIA. Once this has occurred, and depending on how access is provided, the current classification of products as 'unrefined/refined' may need to be revisited.

Similar issues where a particular product category includes different types of products e.g. services and digital data. As discussed in Section 5.1.1 individual service or non-digital products must be excluded from the analysis. Hence where there are multiple product types in a product category, without per product information, the whole category must be excluded. This is most obviously the case for 'Other Products & Services', where the inclusion of consultancy, repographic services, surveying services, contract carto and positional services excludes the other data products from the analysis. In addition 'Mid & Small Scale Consumer Mapping' is excluded because it includes a large number of non-digital products, Imagery and Height is excluded because it includes non-digital (aerial photographs are offered in print form) and 3rd party material. The presence of 3rd party data (as determined by the existence of royalty payments) also leads to the exclusion of 'Address', 1:10k Products, Mid & Small Scale Mapping. Making these exclusions this leaves the following datasets to be analysed:

- Large scale topo (unrefined)
- Transport Network Products (unrefined)
- Mid & Small Scale Digital (refined and unrefined)

As shown, 'Mid & Small Scale Digital', includes both refined and unrefined products. As discussed above, the combination of refined and unrefined products in the same category makes the analysis rather tricky. Furthermore, without per product revenue information, when doing the analysis it will be necessary to apply the same parameters to the entire category. Clearly, this is problematic when there is significant product heterogeneity – exactly the case for this product category (see Table 5.12). Thus, for these reasons, it seems prudent to exclude also 'Mid & Small Scale Digital' category. This leaves just 'Large Scale topo' and 'Transport Network Products'. In addition to being pure 'unrefined' these both have the advantage of

essentially consisting of a single product. Furthermore, together they account for nearly £70m (61%) of Ordnance Survey's revenue (and almost all of their unrefined revenues).

Lastly, it is important to consider costs. As mentioned above, mosts costs could not be allocated to a product, or even product category. This was to be expected given that most products derive ultimately from a few basic databases (in fact, in future, there may only be one). What was available was a fairly thorough breakdown of costs by activity. This breakdown is shown in Table 5.16, along with information on the costs that could be allocated to an individual product category.

5.4.3 Analysis

The basic form of the analysis has already been laid out in detail in Section 5.1.4. What remains is to set values for g, the proportion of revenue from government, and for the elasticity of demand: ϵ, and the multiplier λ. Lacking per-product information here it will be necessary to apply parameters to a given product category – though for the two categories being considered the category is approximately identical with a single product.

Obtaining a value for g is fairly straightforward since one can simply use the 'public sector' proportions in Table 5.11.

Turning to the elasticity of demand, comparison with the data in Section 4.3 suggest using the medium or high range. Given the complexity and level of current price, the wide number of potential users both in business and among the general public, as well as the experience elsewhere (for example in Australia), it seems the high range seems the most appropriate to use here.

Turning to the multiplier, as ever, the situation is made more difficult by the greater uncertainty. Just as for weather data, geographic information offers very significant scope for reuse and recombination whether in new datasets or services.

Category	Data Collection & Management	IT Infrastructure	IT Services	Product Distribution	Sales & Marketing	Corporate Overhead	Total
CONFIDENTIAL							

Table 5.16: Ordnance Survey Cost Breakdown (in £000s)

Product	**Rev**	g	ϵ	λ	ΔB	ΔG	ΔW	**AC/MC**	**RC**
Large Scale Topo	66097	0.58	2.0	3.0	159898	-11554	148343	MC	True
Transport Network Products	3263	0.53	2.0	3.0	8014	-721	7293	MC	True
Total	69360				167912	-12276	155636		

Table 5.17: Ordnance Survey Analysis (in £000s). For meaning of headings see caption to analogous Companies House Table in section 5.5.

Furthermore geographic information is a key input into activities, such as local and regional planning, flood prediction by insurers and government, transport use and planning, that clearly have associated externalities. As such, it seems appropriate to assign these product categories multipliers to the medium range.

These assignments for ϵ and λ, along with the resulting values for the outcome variables, are shown in Table 5.17. With the values used, for the two product categories under analysis, a marginal cost regime would be preferable, and this result is robust to the usual checks (see Section 5.1.4 for more on robustness).

As for the Met Office there are questions about whether lowering prices for these product categories might then affect revenues for products in other categories. However such reductions in revenue would be matched by an increase in surplus for external users and consumers. Furthermore, a reduction in revenues would be matched, at least to some extent by a reduction in costs.[32] Moreover the size of the likely welfare gains shown in Table 5.17 are such that even if the impact on other product categories were very large (and thus the impact on government expenditure more severe) the change to marginal cost would still be welfare improving.

In addition there are a couple of further issues which merit attention. First the presence of CPAs might result in some upwards bias in the estimates in Table 5.17. Specifically, CPAs allow an organization access to a defined range of products for a single fixed tariff. Thus, including the CPAs in the usual demand curve approach

[32]For example fewer paper maps to print imply reduced printing and overhead costs.

Product	**Rev**	g	ϵ	λ	ΔB	ΔG	ΔW	**AC/MC**	**RC**
Large Scale Topo	20545	0.0	2.0	3.0	58519	-14613	43905	MC	True
Transport Network Products	1644	0.0	2.0	3.0	4682	-1169	3513	MC	True
Total	22189				63202	-15783	47418		

Table 5.18: Ordnance Survey Analysis without CPAs (in £000s).

might be slightly misleading.[33] Thus, as a further 'robustness' check Table 5.18 repeats the analysis stripping out revenues from the CPAs and setting $g = 0$ for the remaining products (on the assumption that all government usage comes via the CPAs). As can be seen the general conclusions remain the same though obviously the net welfare gains are lower.

Finally, the lack of internal transfer pricing of data mean there is no great certainty that revenues equal costs for a given product category.[34] In particular costs could be higher than revenues for the two product categories under consideration. Specifically, one could argue that, because most costs are not product specific, the costs incurred for these two product categories are almost identical to the costs for all products. **... CONFIDENTIAL ...** Using these higher cost figures would obviously reduce the net welfare benefits (and increase costs to government). Of course, this approach envisions that most production, management and distribution costs would now have been paid. Hence, either the surplus on the remaining product categories would have gone up or one could reduce prices there as well (generating increases in general welfare and taxation). The overall effect would then be ambiguous and to investigate it properly would require a more detailed cost and revenue breakdown than is available and so it is not pursued further here.[35]

[33]However, given that CPAs still have some quantity restrictions and that there is likely to be a range of customers for CPAs all with different willingness-to-pays, the standard demand curve approach should still apply reasonably well.

[34]**... CONFIDENTIAL ...**

[35]For more on the data available and the restrictions it places on the analysis see the previous section.

5.4.4 Summary

Given the data available the analysis has focused down on two main product categories: 'Large Scale Topographic' and 'Transport Network Products', both of which consisted solely of unrefined products. Together the products considered account for around £70m of the Ordnance Survey's £114m of revenue in 2006/2007 (and would certainly account for the vast majority of revenue from unrefined products). The following discussion is therefore premised on the assumption that the charging policy for other products remains unchanged.[36]

For these two product categories, the analysis suggested that a change from an average cost to a marginal cost regime would increase welfare. Specifically, gross benefits would be around £168m a year while net costs to government would be around £12m. Overall this implies an overall net benefit to society of 156m.[37]

The actual increase in subsidy required would obviously be higher than the £12m net impact on government as this figure includes the benefits of increased tax revenues. Taking the simplest approach the increase in subsidy would be equal to the loss of revenue from non-government sources, which for these two categories combined would be around £30m. Adding this to existing payments from government would make a grand total of £85m. By comparison it should be noted that the Met Office currently receives around £125m from government each year. The comparison with the Met Office is instructive in other ways here as well. There, the CAA (Civil Aviation Authority) contributes to the subsidy for the PWS (Public Weather Service) in addition to the government. Looking at the Ordnance Survey, if some sort of analogous PGS (Public Geodata Service) subsidy were being considered it

[36]That said, one could probably extend the conclusions to all unrefined products without the results changing very substantially. This is because the great bulk of revenue from unrefined products is already accounted for by these two categories. Thus, even if the data had been available to extend the analysis to all unrefined products, the results would have remain largely unchanged.

[37]Since the amounts involved here are substantially above those for most other trading funds it is worth emphasizing that while, necessarily sensitive to the particular parameter values chosen even using the most conservative plausible parametrisations a change to marginal cost pricing delivers a net welfare gain.

might also be possible to have contributions spread across a variety of organizations (and departments within government).[38]

5.5 The UK Hydrographic Office (UKHO)

5.5.1 Introduction

The UKHO is a trading fund producing charts of the world's oceans for navigation. Unlike the Met Office the UKHO does not itself produce the survey data, but obtains the data from other organisations such as national hydrographic offices or UK ports. This is important as it means that the UKHO does not necessarily have sufficient rights to do what it (or government) wishes with the data it is supplied with.

The UKHO main efforts are thus directed towards assessing and compiling the data. The outputs of this process are paper charts of the UK and the rest of the world. These are also digitalised into Electronic Navigational Charts. Table 5.19 summarises the costs and revenues of each of its key activities. As can be seen the UKHO has apportioned revenue (and costs) to UK data (which it is in a position to distribute as it wishes)[39] and international data (which is true 3rd party data in the sense that it has limited ability to redistribute/resell as it wishes). Furthermore in Table 5.19 it can be seen that international data accounts for the lion's share of UKHO revenue (and costs). That is out of total revenues from charts, publications and Notices to Mariners of £69m, £46m (67%) is derived from non-UK data. The

[38] Concretely, a 'PGS subsidy' could start out based on the existing CPAs which cover government (e.g. the PGA and the MSA) but also include other large current or potential users (e.g. the NHS, large utility companies etc) who have an interest in maintaining the quality and availability of UK geographic information.

[39] The exact position of the UKHO with regard to UK data is not entirely clear. Much of this data is collected by third parties such as port authorities who then 'give' the data to the UKHO – the incentive being that without suitably charted waters ships will not be able to use those ports – but do not necessarily formally transfer the rights in the data. Nevertheless, given the interest that most of these third-parties have in the UKHO making such data available (as widely as possible), as well as the position of the UKHO as the premier chart provider for UK waters, it seems likely that the UKHO would be obtain the necessary permissions to distribute UK data as it wished (or, at least at marginal costs – for example).

Sales/Cost Area	Sales	Costs	Comments
CONFIDENTIAL			

Table 5.19: UKHO Financial Summary for Year 2006/2007 (in £000s).

subsequent third party royalties are likely to account for a significant proportion of total costs. Out of a total of £78M in costs, £38M (49%) falls in the 'other commercial costs category', which includes 3rd party royalty payments. In addition out of a total revenue in 2006/7 of £83M, £12M (15%) came from government though most of that revenue (approx £9.1M) did not come from purchase of data but provision of 'defence-related' services.

5.5.2 Data

As discussed in Section 5.1.1 this study only focuses on data which the UKHO is in a position to distribute as it wishes – essentially all data related to UK waters. Electronic and Paper Charts are available on a chart by chart basis,[40] so where the product includes the possibility of purchasing charts from the rest of the world as well as the UK, the UKHO have apportioned the revenues and costs accordingly. However the division does not always fall cleanly. UK charts are also included in wider area disks and bulk discounting is in use. This means that some of the revenue will originate from customers who bought the wider area disk, but did not require UK charts. However given that the UK takes up a relatively large area of the wider area disk it is likely that it would cost less to purchase individual non-UK charts, so these customers are likely to be small. It is therefore still worthwhile to analyse digital charts, but the results should be treated with caution.

Some of the publications (both digital and paper), such as the Tide Tables and

[40]ENCs can be purchases on a unit by unit basis. See the ENC brochure for more details available at http://www.ukho.gov.uk/content/amdAttachments/brochures/enc_brochure.pdf. ARCS coverage is on a chart by chart basis. See the ARCS brochure for further details available at http://www.ukho.gov.uk/content/amdAttachments/brochures/arcs06.pdf.

Light Lists, also include other regions in addition to UK waters. As the UK remains a significant proportion of the area they can still be analysed. The UKHO may only be able to lower the price of the UK component. This might lower $\frac{\Delta p}{p}$ from the baseline assumption that $\frac{\Delta p}{p} = 1$. Taking this into account would favour average cost pricing. For the time being this effect will not be considered, but if the analysis finds that these products should be charged at marginal cost, the results should be treated with caution.[41]

Again no specific marginal cost data was provided so the analysis is restricted to digital products (where it is fair to make the approximation that marginal costs of supply are close to zero), this leaves 'Digital Charts', 'Digital Publications' and 'Licensing and Miscellaneous' categories. Table 5.20 provides a summary of the products selected for analysis and other paper products of interest. The data available for these products is summarised in Appendix B.1.

5.5.3 Analysis

The basic form of the analysis has already been laid out in detail in Section 5.1.4. What remains is to set values for g, the proportion of government revenue, and for the elasticity of demand: ϵ, and the multiplier λ. For g, the 'conservative' assumption has been made that all of government expenditure on UKHO was on 'services', that is the expenditure on the 'pure' data products was zero (so $g = 0$).[42]

Turning to the elasticity a low range is used for most products. This is because according to Chapter V of the International Convention for the Safety of Life at Sea 1974 (SOLAS V), UK registered vessels over a certain size are required to carry specific Nautical Charts and Publications (in certain cases only UKHO charts

[41] In fact as will be seen below this issue does not arise, because average cost turns out to be preferable for these particular products.

[42] In any case government sales totals £11.7m, £9.1m of which is defense specific services in 2006. This suggests that government sales of 'pure' data products was only £2.6m, which is relatively small compared to total revenues of £82.5m, and so assuming $g = 0$ is likely to have a negligible effect on any results.

Product Category	Product	RW/ VA	RF/ URF	UK/ AP	Product Type	Notes
ARCS Navigator	World Series	VA	RF	AP	DD	Commercial Admirality (electronic) Raster Chart. Compatible with ECDIS.
ARCS Skipper	World Series	VA	RF	AP	DD	Leisure Admirality (electronic) Raster Chart. Compatible with ECDIS.
ECDIS	5AF (SAtlan, Ascen, Falk)	VA	RF	UK	DD	Electronic Chart Display and Information Systems. Combination of ARCSs and ENCs.
ECDIS	5EC (EngChan (Chan Isle))	VA	RF	UK	DD	Electronic Chart Display and Information Systems. Combination of ARCSs and ENCs.
ECDIS	5SN (Scot, NE Coast, NthSea)	VA	RF	UK	DD	Electronic Chart Display and Information Systems. Combination of ARCSs and ENCs.
ECDIS	5SW (Scot,W Coast)	VA	RF	UK	DD	Electronic Chart Display and Information Systems. Combination of ARCSs and ENCs.
ENC	English Channel	VA	RF	UK	DD	Electronic Navigational (vector) Charts. Compatible with ECDIS.
ENC	World Series	VA	RF	AP	DD	Electronic Navigational (vector) Charts. Compatible with ECDIS.
Digital Publications	ADRS6	VA	RF	UK	DD	Admirality Digital Radio Signals Volume 6 (substitute fo ALRS).
Digital Publications	Digital List of Lights	VA	RF	AP	DD	Digital version of paper lights list.
Digital Publications	Digital TotalTides	VA	RF	AP	DD	Digital version of paper tide tables.
Paper Publications	ALRS	VA	B	UK	ND	Admirality List of Radio Signals. Information on maritime communications.
Paper Publications	Lights List	VA	B	AP	ND	Listings of all lighthouses, lightships, etc.
Paper Publications	Mariners Handbook	VA	RF	UK	ND	Essential maritime information on charts, operational information and regulation, tides, currents, etc.
Paper Publications	Sailing Directions	VA	RF	UK	ND	Essential information on all aspects of navigation.
Paper Publications	Tide Tables	VA	RF	AP	ND	Daily predictions of the times and heights of high and low waters for over 230 standard and 6,000 secondary ports
Licensing	Copyright	-	-	UK	L	No VA/RW and RF/URF assignment due to range of products licensed.

Table 5.20: Available Data UKHO (2006). RW=Raw, VA=Value Added, RF=Refined, URF=Unrefined, B=Both Refined and Unrefined UK=UK Only Products AP=Apportion to the UK by the UKHO, DD=Digital Data, ND=Non-Digital Data, L=Licensing

can be employed). This applies to all products analysed except for 'Copyright Licensing'.[43] It would therefore seem reasonable to assign legally required products demand elasticities from the low range, that is all products other than copyright licensing. In this last case it seems reasonable to use an elasticity from the medium range as there is no reason to expect it to be particularly high or low.[44]

For the multiplier the low range is employed for all the products. This seems appropriate for charts and publications as these are designed for and sold directly to large shipping vessels, so there seems to be limited scope for organisations to re-use the data and add-value. Turning to licensing, a large fraction of companies to which the data is licensed serve the merchant shipping industry or produce leisure products. While there clearly is some scope for reuse here the opportunities seem more limited than in most cases and so, in order to be 'conservative' a low multiplier has been used.

The assignments for ϵ and λ, along with the resulting values for the outcome variables, are shown in Table 5.21. With the values used, a marginal cost regime would only be preferable for Copyright Licensing, and all results are robust to the usual checks (see Section 5.1.4 for more on robustness).

5.5.4 Summary

The analysis focused on Digital UK Charts, Digital Publications and Licensing, and for all other products it was assumed that the pricing regime would remain unchanged.

Focusing on these products, the results suggested that under the likely range of

[43] ADRS6 is designed to meet SOLAS requirements and UKHO is in the process of attaining regulatory approval, so it is considered as if this is the case.

[44] The UK Hydrographic Office informed us that they had reduced prices substantially in this area over the last few years and had simultaneously seen a significant increase in demand. Unfortunately no precise quantitative estimates of these effects were available though it should be noted that, even if they were, the large technological changes also taking place over the same period would render interpretation difficult (i.e. it would not be clear to what extent changes in demand were driven by price reductions versus technological advance).

Product	**Product Category**	**Rev**	g	ϵ	λ	ΔB	ΔG	ΔW	**AC/MC**	**RC**
World Series	ARCS Navigator	XXX	0.0	0.25	1.5	XXX	XXX	-25673	AC	True
World Series	ARCS Skipper	XXX	0.0	0.25	1.5	XXX	XXX	-528	AC	True
5AF	ECDIS	XXX	0.0	0.25	1.5	XXX	XXX	-978	AC	True
5EC	ECDIS	XXX	0.0	0.25	1.5	XXX	XXX	-3271	AC	True
5SN	ECDIS	XXX	0.0	0.25	1.5	XXX	XXX	-2172	AC	True
5SW	ECDIS	XXX	0.0	0.25	1.5	XXX	XXX	-1784	AC	True
English Channel	ENC	XXX	0.0	0.25	1.5	XXX	XXX	> 0	MC	True
World Series	ENC	XXX	0.0	0.25	1.5	XXX	XXX	-2896	AC	True
ADRS6	Digital Publications	XXX	0.0	0.25	1.5	XXX	XXX	-1802	AC	True
Digital List of Lights	Digital Publications	XXX	0.0	0.25	1.5	XXX	XXX	-1883	AC	True
Digital TotalTides	Digital Publications	XXX	0.0	0.25	1.5	XXX	XXX	-2953	AC	True
Copyright	Licensing	XXX	0.0	1.0	1.5	XXX	XXX	337739	MC	True
Total		1956101				1082013	-744274	337739		

Table 5.21: Welfare Analysis for UKHO (2006/2007). For meaning of headings see caption to analogous Companies House Table in section 5.5. Note that the zero revenue for ENC English Channel leads to spurious classification into the MC category and this should probably be ignored (of course since revenue is zero here it makes no difference to the calculations). Note also that the totals other than for revenue are only for those products for which a change in pricing appears preferable (i.e. Copyright Licensing only). XXX indicates that the contents of a cell have been removed for confidentiality reasons.

parameter values a change from an average cost to a marginal cost regime would be welfare improving only for 'Copyright Licensing'. Specifically, the calculations showed that adopting a marginal cost pricing policy for Copyright Licensing (and maintaining average cost pricing for all other products) would result in gross benefits of around £1.082m with government incurring net costs of around £744k (on a direct subsidy of 854k). Overall this means that the net welfare gain to society would be around £338k.

5.6 HM Land Registry

5.6.1 Introduction

HM Land Registry holds the official register of titles to land for England and Wales. The register also records dealings (for example, sales and mortgages) with registered land, and the prices for which land (and the property thereon) was sold. Table 5.22 provides a summary of HM Land Registry's activities. Revenues of the Land Registry are by far the largest of the Trading Funds studied. Preliminary (inspection) services accounts for the majority of revenues from Land Registry's data provision activities. However Land Registry's data provision activities account for a small proportion of its total revenue. As with Companies House the majority (86%) of its total revenue is from collecting the data itself, through obligatory registrations (substantive applications).

5.6.2 Data

At a disaggregated level the Land Registry provided revenue and cost information for preliminary (inspection) services, an annual summary of which is found in Table 5.23. The Land Registry considers all these products value-added and refined. One could however argue that the data is refined given that this is the most upstream

Activity	Revenue	Cost	Notes
Day List	Nil	-	
Preliminary (inspection) services	68322	100428	
Substantive applications (transactions)	422475	-	
Property Data Services	1298	-	
Consultancy Services	553	-	
Local Office Costs		278011	Includes the cost of manual production and postage for Preliminary Services.
Information Systems Costs		65,912	Includes the electronic processing, storage and issue costs for Preliminary Services.
Totals	492,648	-	

Table 5.22: Land Registry Financial Summary for Year 2006/2007 (£000).

data available to other users and is not contestable. Each document is available either electronically or in paper format (in which case it is posted) through Land Registry Direct. As discussed in Section 5.1.1 it is only possible to analyse the digital data. However as the revenue data could not be split between digital and paper items, it would be difficult (if not impossible) to analyze the vast majority of digital items in the preliminary (inspection) services category (including Official Copies).

Furthermore many of these products had a significant service component as they were offered in conjunction with a search facility, and the breakdown between search and data access is complex (for example, no fee is charged for a Search of the Index Map, unless more than 10 title numbers are revealed in the result). In addition revenue data for Register Views, Title Plan Views and Document Views was not collected individually for each product category, but only for the combined total of all three categories. These remaining three categories are therefore excluded from the analysis.

It should also be noted the quantity data was available only in terms of trans-

actions and not sales of each item. This may explain why the average price per transaction is higher than might be expected given the prices set in the Fees Order[45] and suggests that more than one item was sold per transaction. Recording data in this way makes it difficult to establish actual demand. Finally, there was also no breakdown of costs at the product level, but only for the whole organisation.

Together these data limitations make it difficult to conduct a review of pricing policy for most of the Land Registry products.[46] Fortunately however, the Land Registry were able to for some other data services, specifically 'Property Price Data'[47] and 'Polygons',[48] which fall into the 'Property Data Services' category. Their respective revenues were £893k, the majority of which was from a bulk form of the product, and £405k. Analysis therefore focuses on these two items.

[45]The 2006 order is available at `http://www.opsi.gov.uk/si/si2006/20061332.htm`

[46]Such limitations do not seem inevitable – for example Companies House, whose activities are in many ways similar to the Land Registry, were able to provided much more detailed accounting information.

[47]Land Registry wished it to be noted that some of their Property Price Information (in non-bulk, summary form) was produced with the aid of consultants who adjusted the data to make it more robust. The particular metrics used are the trade-secret of the consultants and not of the Land Registry and thus there might be some issues if the terms under which the data was to be provided were changed. However this would only affect the summaries (e.g. price indices per region or per postcode) and would not affect the bulk data product which permits direct access to the full price information database on all residential property sales in England and Wales.

[48]Polygons are the electronic shapes (GIS vector data) of registered ownership. There will be some third-party (Ordnance Survey) IPR in the 'Polygons' data, as Land Registry explained: "Here we share IPR with Ordnance Survey. All of our title plans are based (as the legislation requires) on Ordnance Survey detail. Therefore, when we look to sell ownership polygons, the data include Ordnance Survey IPR. For example, on a single, fenced house plot, the ownership polygon will share a 100% co-incidence rate with the Ordnance Survey map detail. On a large public authority ownership terrier, the co-incidence rate will be substantially lower, because there may be no recorded features (there may be open spaces, verges and the like). Because it would be unproductive to work out the precise co-incidence rate in each case (the manual counting and comparison of thousands of polygons with the OS map would be impractical) we have worked out an 'average' co-incidence rate with OS and we pay them a 'royalty' for each sale based on that." However since the external data (where present) comes from another trading fund, it has not been considered as truly '3rd Party', that is, as grounds for excluding the Poylgons from the analysis due to the difficulties of obtaining a license from an external source in order to supply at e.g. marginal cost.

Product	Year	Avg. Price	Quant.	Rev.	Tot. Cost	Product Type	R/ VA	RF/ URF
Official Copies	2004	5.63	5091	28685	42593	DD	VA	R
Official Copies	2005	5.19	6168	31983	42350	DD	VA	R
Official Copies	2006	6.63	7067	46869	42323	DD	VA	R
Searches of the Index Map (charged for)	2004	0.13	3353	424	24917	S	VA	R
Searches of the Index Map (charged for)	2005	0.14	3253	460	26969	S	VA	R
Searches of the Index Map (charged for)	2006	0.16	3302	516	30197	S	VA	R
Register Views	2004	1.54	5902	9724	11723	DD	VA	R
Register Views	2005	1.71	6812	13481	13624	DD	VA	R
Register Views	2006	2.25	7805	20938	23414	DD	VA	R
Title Plan Views	2004	1.54	380	9724	754	DD	VA	R
Title Plan Views	2005	1.71	988	13481	1977	DD	VA	R
Title Plan Views	2006	2.25	1214	20938	3643	DD	VA	R
Document Views	2004	1.54	36	9724	72	DD	VA	R
Document Views	2005	1.71	105	13481	210	DD	VA	R
Document Views	2006	2.25	284	20938	851	DD	VA	R

Table 5.23: Preliminary (Inspection) Services Data. Quantity (Quant), Revenue (Rev) and Total Costs (Tot. Costs) in £000s. N.B Revenues and average price given for Register Views, Title Plan Views and Document Views are totals (or averages) for all three products, as revenue data was not collected for each product individually. Hence they are the same value in the same year.

5.6.3 Analysis

As noted above formal analysis has been limited to 'Property Data Services'. It would seem reasonable that products in this category have elasticities and multipliers within the medium ranges, as there is no reason to expect them to be particularly high or low. As the data is available in digital form one can approximate $\Delta p/p$ as close to 1. Since there is no information on g, the proportion of government revenue, the conservative assumption that it is zero is made. The assignment for ϵ and λ, with the resulting values for the outcome variables, are shown in Table 5.24. With the values used, for all the products under analysis, a marginal cost regime would be preferable and the results are robust to the usual checking (see Section 5.1.4 for more on robustness).

It is also important to note that there is no bulk product. There does not appear to be any obvious reason why the Land Registry Direct could not make available its data in bulk form (e.g. an ftp download of a snapshot of their database or an RSS feed of daily changes). Currently the only way to get access to the whole database is to download and pay for each data item individually. However, given that there are 21 million items each priced at £3, this would have a total price of £63m and with no update service the value of the data is likely to depreciate rapidly. In addition providing the data one item at a time would likely be very inefficient both for the Land Registry and for customers. It is therefore likely that the provision of a bulk product would significantly improve welfare and comparison with Companies House suggests that pricing this product at marginal cost is likely to be the preferable policy.

5.6.4 Summary

The Land Registry offers no bulk data downloads for all of its data. As there are no other feasible ways to access all (or substantial portions) of the Land Registry's data

Product	**Rev**	g	ϵ	λ	ΔB	ΔG	ΔW	**AC/MC**	**RC**
Property Price Data	893	0.0	1.0	3.0	1602	-730	871	MC	True
Polygon GIS Data	405	0.0	1.0	3.0	726	-331	395	MC	True
Total	1298				2328	-1061	1266		

Table 5.24: Land Registry Welfare Analysis (in £000s). Based on rough annual figures. For meaning of headings see caption to analogous Companies House Table in section 5.5.

service bulk data access and reuse is effectively prohibited for the core parts of the Register. This could restrict competition with the Land Registry's search services and hold back the innovative use of the data. A comparison with Companies House would indicate that introducing bulk data provision and charging at marginal cost is likely to raise social welfare significantly. However no precise calculation could be made due to limits of the available data. Concerns may be raised concerns over privacy, however these are questions about whether data should be made available at all, rather than the price at which data it is provided. Furthermore, it is likely that many of these concerns could be addressed by adopting some form of anonymisation (as usesd by the DVLA) and specific license conditions to prohibit forms of reuse that are unwanted (for example unsolicited direct mailing).

Analysis for Property Price data and Polygon GIS data was possible. This indicated that under the likely range of parameter values a change from an average cost to a marginal cost regime for these individual products would be welfare improving. Specifically the calculations showed that lowering the price to marginal cost would result in gross benefits of around £2.3m with government incurring net costs of around £1.1m (£1.2m gross) if it were to finance the policy. Overall this means that the net welfare gain to society would be around £1.2m.

These figures assume that the fall in revenue of £1.2m at the Land Registry would be made up by government. This does not take into account the possibility of

	2006-07	Comments
Total Income from Operations	613	
– o/w Data Provision (estimate)	7	Revenues cover direct costs of providing specific data and contribute to the DVLA's cost of updating the records.
Total Operating Expenditure	488	

Table 5.25: DVLA Summary 2006-7 (millions).

making up the deficit via charges for registrations. In this case the financial burden would not fall on government. If the elasticity of demand for registrations is low, which is likely to be the case when registration is compulsory, this would reduce the costs (and increase the net benefits) of moving to marginal cost pricing.

5.7 The Driver and Vehicle Licensing Agency (DVLA)

5.7.1 Introduction

The DVLA is responsible for vehicle registration and the collection and enforcement of Vehicle Excise Duty. The vast bulk of its income comes from its registration activities and Excise activities. However it does make some of the data it collects subsequently available (without any personalised information unless there is consent from the license holder). Estimated revenues from data provision are low compared to the rest of DVLA's operations, as can be seen in Table 5.25, accounting for only about a ninetieth of DVLA's total income last year.

5.7.2 Data

Data for each data product was provided including prices, and customers/transactions and is presented in Table 5.26 along with this study's estimates of the revenues.[49]

[49]No details were provided on whether the products were raw/value-added or refined/unrefined.

No specific cost information was provided. As the data was provided at a very late stage of the study only an initial analysis is presented. As discussed in Section 5.1.1, only Anonymised Data, Bulk Data, and Mileage Data is considered as it is data in digital and bulk form.

The mosaic data is excluded as it appears to be customised data for one client. Vehicle Fee Paying Enquiries, the Driver Entitlement Checking Service and Drivers' Fee Paying Enquiries are not analysed as this contains personal data (which in the latter two cases requires consent by the license holder), so it is likely that there has restrictions over their dissemination by the DVLA. This leaves anonymised data, bulk data and mileage data.

5.7.3 Analysis

As bulk digital products are being considered one can make the approximation that $\Delta p/p = 1$. Given the high price and hence low demand for the bulk and anonymised data, as well as the variety of possible uses, it would seem reasonable to use elasticities in the high range. That is one would expect a substantial increase in demand on moving to marginal cost pricing. As the mileage data is at a lower price and it is difficult to perceive what additional use there would be for such a narrow data set, its elasticity is assigned to the medium range.

'Bulk data' is provided to HPI (Hire Purchase Information) companies. As HPI checks indicate if a vehicle is stolen, providing these checks for free may provide the additional benefit of increasing the return of such vehicles. Preventing the sale of stolen vehicles may in turn discourage their theft. Taking these further benefits into account it would seem reasonable to use a range of λ from the high category. For the remaining products the medium range of λ is used as there is no reason to expect it to be particularly high or low for these categories.

No information was provided on g, the proportion of government revenue, but it

Information Released	Price	Quant.	Revenues	Use	Notes
Anonymised Data	100000	5	500000	Statistics, commercial siting and lifestyle analysis.	Price per customer
Bulk Data	85000	6	510000	Hire Purchase Information Checks	Price per customer
Mosaic	50000	1	50000	Mosaic codes provided by the one client to conduct demographic analysis.	Price per customer
Mileage Data	3000	4	12000		Price per customer plus 1p per record (no transaction information provided)
Driver Entitlement Checking Service (DECS)	3	42500	127500	Validation of employee driver entitlement.	Price per Transaction. Consent must be given by the license holder.
Vehicle Fee Paying Enquiries	3.75	1379946	5174798	Investigation of vehicle offences.	Price per Transaction (average of £2.5 and £5 fees charged for different data requirements). Fees cover cost of providing the service.
Drivers' Fee Paying Enquiries (written)	5	173690	868450	Entitlelment to Drive Information	Price per Transaction. Consent must be given by the license holder. There is also a premium telephone service, however no price information was available.

Table 5.26: DVLA Product Information 2006-07. 'Quant' is the number of customers/transactions.

Product	**Rev**	g	ϵ	λ	ΔB	ΔG	ΔW	**AC/MC**	**RC**
Anonymised Data	500	0.0	2.0	3.0	1424	-355	1068	MC	True
Bulk Data	510	0.0	2.0	7.0	2886	-217	2668	MC	True
Mileage Data	12	0.0	1.0	3.0	21	-9	11	MC	True
Total	1022				4332	-582	3749		

Table 5.27: DVLA Analysis 2006-07 (in £000s). For meaning of headings see caption to analogous Companies House Table in section 5.5.

seems reasonable (and 'conservative') to assume that it is zero for the products being considered. The assignment for ϵ and λ, with the resulting values for the outcome variables, are shown in Table 5.27. With the values used, for all the products under analysis, a marginal cost regime would be preferable and the results are robust to the usual checks (see Section 5.1.4 for more on robustness).

5.7.4 Summary

The analysis focused only on Anonymised data, Bulk data and Mileage data, and for all other products it was assumed that the pricing regime would remain unchanged.

For these examined categories, the results suggested that under the likely range of parameter values a change from an average cost to a marginal cost regime would be welfare improving. Specifically, the calculations showed that adopting a marginal cost pricing policy for these particular products would result in gross benefits of around £4.3m with government incurring net costs of around £582k if it were to finance the policy (gross costs would be around £1m). Overall this means that the net welfare gain to society would be around £3.7m.

In the case of the DVLA central Government need not necessarily provide any funds itself. Specifically it might be preferable to finance lower priced data from registrations (or from Excise Duty) as demand for these services is likely to be (relatively) more inelastic than for data. With the loss of revenue in moving to marginal cost at just 0.16% of total revenues (1m out of 613m) and the DVLA's

rate of return on capital to March 2007 of 16.6% this seems a very viable option.[50]

[50]See DVLA's 2006-07 annual accounts available at http://www.dvla.gov.uk/media/pdf/publications/annual_accounts0807.pdf

Chapter 6

Commitment, Incentives and Regulation

6.1 Introduction

The terms of reference stated that particular attention be paid to the question of how changes in charging policies would impact on "data collection, maintenance and production", and whether this would result in "any changes to data quality". In this section these questions will be considered in the wider context of how best to address the regulatory, commitment and incentives issues that arise in relation to trading funds. These issues are dealt with here because, though important, they cannot be easily addressed either theoretically or empirically in the 'formal' framework laid out above.

With the development of the 'knowledge' economy, driven in large part by improvements in digital technology, the supply of data by trading funds can be seen as an analogous activity in 'information' sector to the supply of physical infrastructure in the form of power and electricity, transport (roads, trains etc), and telecommunications. This comparison is illuminating in a variety of ways.

First, existing utilities often have similar cost structures where large fixed costs are combined with low marginal costs. Related to this, many of them, at least in some areas of their activities, have 'natural' monopolies just as trading funds may do in some areas of their business. Utilities are usually providing 'essential' infrastructure which, if not directly essential to government, are essential to the general economy – this could be seen as similar to the 'public task' of trading funds. For a combination of these reasons many of these utilities are regulated and have been now for some time and one might think that these regulatory experiences would have something to offer when considering the situation of trading funds (few, if any, of which have any independent regulation at the present time).

The analysis is complicated however, when compared with many other 'regulated' industries, by the fact that government takes multiple roles in relation to trading funds. In particular, Government acts as both a shareholder (via the Shareholder Executive), regulator/parent (each trading fund sits in a particular government department), and customer. Furthermore Government's customer role is far more prominent in relation to trading funds than in relation to any other 'utility' – Government is often by far the largest customer for trading fund data and in some cases account for over 50% of sales.[1] This close relationship is reflected in the status of trading funds which have no separate legal identity from their parent departments. This means, for example, that while trading funds can draft detailed 'Memorandums of Understanding' or 'Customer Supplier Agreements' with government it is not clear whether these are legally enforceable contracts – it is not possible after all for government to sue itself.[2]

[1]For example the Met Office income from government whether via sales or subsidy is over 80% of revenue. Even for Ordnance Survey where the proportion of revenue coming from Government has been falling the proportion is close to 50% and for particular product ranges may be well over that. At the same time, for some other trading funds, especially those which registration-based, the proportion of income from Government is very low (approximately zero in the case of the Land Registry, for example).

[2]This point would also be of some importance when considering how best to address the commitment issues discussed below.

Additionally the government-to-trading fund relationship remains rather opaque and it is often unclear exactly what a given party can and can not do – though the current situation might be seen as a substantial improvement in transparency compared to the past.

For example, one of the advantages frequently mentioned of average-cost or profit-maximizing charging policies is the greater freedom and certainty it provides for trading funds (this is discussed further below). However, given the Government's role as the major purchaser of data from these organizations it is entirely possible for the Government to use its role as a monopsonist to reduce suddenly its payments in lean years (just as the Government might choose to reduce a subsidy). Conversely, it is not clear what would necessarily prevent a trading fund using its position as a sole supplier of some data products to raise charges to Government very sharply. Obviously, in practice, neither of these outcomes are particularly likely, precisely because of the close connection between trading funds and government. This connection is clearly very important but is, as yet, largely unformalized.[3]

Finally, a crucial point to bear in mind is that many of the trading funds enjoy a near-monopoly on at least some of their data, a monopoly furthermore made possible or strengthened by government activity. For example, in the case of 'registration-based' trading funds such as Companies House, the Land Registry or the DVLA it is a statutory requirement to deposit data with them. In the case of the Met Office, in addition to the natural monopoly afforded by the high fixed costs of data collection, the government provides substantial funding for the PWS.[4] Furthermore,

[3]Several of the trading funds contacted for this study pointed that there was no explicit definition, written or otherwise, of what exactly their 'public task' was. This was of some importance since the Treasury definition of 'raw' information makes explicit reference to 'public task' (as does Regulation 5 sub-section (1) of the *Re-use of Public Sector Information Regulations (SI 2005 No 1515)*). Some of the trading funds had sought clarification for the purposes of responding to this study and others indicated that they had substantially revised (and narrowed) their view on what their 'public task' was as a result of the APPSI appeal report (Board, 2007).

[4]The Met Office have sought to address some of the problems these my cause from a competition perspective by maintaining a clear division between their 'wholesale' and 'retail' arms with the same access terms applied to all, including their own retail division, when purchasing data from

in most cases the data marketplace in which trading funds operate have a clear upstream/downstream structure with the trading fund 'monopoly' most prominent in the upstream market. As discussed at length in the OFT report (Annex C) this presents a whole raft of competition issues, particularly in relation to tying, exclusionary dealing, predatory pricing and the like. As a result it would seem clear that some form of price/access regulation would be necessary if abuses of market power were to be avoided, and adequate competition and innovation be encouraged (at least downstream).[5] It would also make it extremely difficult to permit trading funds pursue a profit-maximization (monopoly-pricing) strategy in the absence of some form of regulatory oversight.

6.2 Commitment

In the analysis presented above it has been explicitly assumed that government would provide any necessary subsidy to maintain trading fund incomes at their present levels (should a charging policy be chosen that resulted in trading funds income dropping below costs). This *implicitly* assumes an ability for government to *commit* to payments both now and in the future. Such an ability cannot be taken for granted. Governments, both in the UK and elsewhere, have frequently demonstrated the difficulty of making such commitments and the impact of political considerations on infrastructure investment.[6] Sudden fluctuations, or simply reductions, in the

the 'wholesale' arm.

[5] In some cases such price regulation is already explicitly required by law. For example, in the case of Companies House, it has been suggested that the Amended First Directive (Council Directive 68/151/EEC amended by Council Directive 2003/58/EC implemented into UK law by s1086(2) of the Companies Act 2006) requires that the cost of obtaining a copy of a whole or part of any document may not exceed the administrative cost thereof. Meanwhile the Capital Taxes Directive (Council Directive 69/335/EEC) and subsequent case law mean that registration fees cannot exceed the costs of the registration.

[6] For example, in 1991, the UK government promised an extra 750 million pounds to the Tube to do renovation work only to have to reverse this commitment a year later due to sudden pressure on the national finances. (LRB, Vol. 27 No. 9, 5 May 2005). See also the discussion of the Land Registry's experience in the early 1990s below.

level of subsidy would be likely to have substantial negative effects on the ability of trading funds to maintain both the range and quality of their products. Clearly, the issue of commitment is an important one to consider.

The issue of commitment is not solely confined to the case where subsidies are being provided. Consider, for example, the hypothetical situation where a trading fund is following a policy of profit-maximization but still retains its current institutional setup where it sits within a given government department. Suppose then that the trading fund decides that one obvious way to increase profits is to increase charges to central and local government, perhaps to the extent that some sections are no longer able to purchase the data. In this case there might be substantial pressure brought to bear by government on the trading fund to price more 'reasonably', or the government might consider amending the trading funds charging policy. In either case the government would have reversed its 'commitment' to allow the trading fund to pursue a policy of profit-maximization. Thus it should also be clear that while the 'commitment' issue may be most prominent in the case where government is providing funds it arises in relation to all of the possible pricing policies. In fact, as discussed further below, the commitment issue relates more to the institutional and regulatory structure in which trading funds operate than to the chosen charging policy.[7]

6.3 Incentives

In addition to the basic commitment issues it is also the case that different charging policies, and the associated different relationships with central government, might result in different incentives faced by trading funds. In particular charging policy could affect incentives for responsiveness, innovation (development of new products), cost reduction and general performance.

[7]See also the discussion of the government's multiple roles above.

For example, a trading fund which has been mandated to price data products at marginal cost may have reduced incentives to develop new products as it will not be able to reap any particular benefits from doing so.[8] Conversely, if marginal cost pricing was combined with some kind of per unit output subsidy this could result in incentives for *over-investment* in quality and capacity improvements because, by over-investing, the trading fund stimulates demand and obtains a larger subsidy.

In terms of responsiveness an organization operating a more 'commercial' pricing policy (e.g. profit-maximizing) might lead a trading fund to be more customer oriented – more responsive to complaints and more concerned about general service quality.

Similarly, wherever a trading fund is regulated (i.e. in all cases except profit-maximization) it may lack adequate incentives to reduce costs – because any reduction in costs may be partially appropriated by the regulator (either in the form of a lower subsidy or lower prices).

6.4 Information and Regulation

All of the charging policies considered with the exception of profit-maximization require some form of regulation (by government or otherwise) to ensure compliance. Even in the case of profit-maximization the government's role as sole shareholder would necessitate some form of oversight.[9]

One might assume that marginal cost (and zero-cost) pricing would require more information (and more effort on the regulator's part) than average cost pricing. In particular as it is unlikely that the level of investment is constant over time there will be important questions as to how subsidies (and price regulation) were allowed

[8]The same could be true in theory from average cost pricing though this depends somewhat on the degree to which the organization engages in cost recovery at the organizational rather than the per product level.

[9]The Shareholder Executive could be seen as currently acting in this role.

to change over time to reflect these needs.

However, as already alluded to above, under cost-recovery managers may have an incentive to 'over-invest' since higher costs can be covered by increasing revenues ('gold-plating'). Additionally, with the ability to set prices in at least some areas trading funds could also behave inefficiently, for example, investing in poor projects, while still complying with cost-recovery at the organizational level since losses could be made up by raising prices or cross-subsidies from other parts of the business. The information needed by a regulator to avoid these outcomes is similar to that required when monitoring a marginal-cost or zero-cost regime – in particular the regulator will need to monitor investments in order to ensure that they are at the efficient level.

Leaving aside these investment questions it is certainly true that different pricing regimes provide different information about the demand curve (and therefore implicitly about surplus). This has already been discussed extensively in the theory chapter above (see figure 3.1 in particular) but it is worth reiterating here. Specifically, if the given pricing policy is being pursued at the per-product level,[10] then profit-maximization and average-cost both have the advantage that they guarantee that a given product is only produced if the surplus from doing so is positive. By contrast under marginal cost (or zero cost) pricing it is possible for a product to be produced (and subsidised) whose net surplus is negative. However it should be noted that this particular point can be taken both ways. A profit-maximization or average-cost regime ensures that a product is produced if and only if the producer surplus is positive (i.e. revenues are larger than costs). Thus there may be products whose total (consumer plus producer) surplus is positive but whose producer surplus is negative – products which (depending on the subsidy structure) might well be produced under a marginal cost or zero price regime.

[10] A moot question given the cost structures of trading funds and the actual evidence in the form of the data provided.

6.5 Summary

The main aspects of the previous discussion have been drawn together in the summary table (6.1) below. One important point to bear in mind when reading this, and when considering these issues in general, is the likelihood that any given charging rate might be applied selectively. For example, different charging policies could be applied to unrefined and refined data – for example marginal cost for unrefined and average cost, or profit maximization, for refined. Thus rather than situating a trading fund in a single column it is important to keep in mind that it could be 'spread' across several, with different parts of a trading fund's operations under different charging policies. The table attempts to reflect this, at least to some extent, by explicitly noting where a particular point relates only to data with particular properties.

6.6 Discussion

There are two lessons to draw from the preceding sections. First, that there is no direct linkage of charging policy to governance issues – in fact governance questions are best seen as orthogonal to pricing ones. In particular, all policies (other than perhaps profit-maximization)[11] require some form of regulation to function well. The second, and related lesson, is that charging policy is not the central issue when considering problems such as commitment and incentives which are themselves the primary determinants of overall performance in terms of data quality, investment and efficiency. Rather, charging policy is best seen as secondary, and dependent upon, the primary matter of the regulatory/governance structures under which data provision (and collection) by trading funds occurs.

[11] Unregulated profit-maximization is likely to have so many problems from a competition perspective as to be infeasible. Thus, even there significant regulation is likely to be required – in which case it will start to resemble one of the other options depending on the form of price and access regulation used.

6.6.1 Commitment

To illustrate consider, a concrete example, provided by the Land Registry. In discussion, they mentioned their situation in the late 1980s and early 1990s just prior to becoming a trading fund. At that point they also operated a cost-recovery regime in which charges were set to cover costs. However, they did not control their revenues but rather returned them to central government. The Land Registry management would then go 'cap in hand to Treasury' to negotiate their budget for the next financial year. In the late 1980s this resulted, they said, in some degree of underfunding, which made it impossible to deal with the level of applications they were receiving. As a result a large backlog of applications built up with all the attendant problems.

In 1993 they became a trading fund, in part because of the problems that had been encountered. Since then this sort of problem has not recurred and, in their opinion, the greater autonomy provided by being a trading fund means that investment can be planned better and they are less subject to the vagaries of 'vote-funding'.[12] Note that throughout the basic charging policy was unchanged with cost-recovery both before and after trading fund status was obtained. Hence, here, it would seem clear that if the improvements in service quality were due to anything, they were due to changes in the regulatory environment, in particular the greater certainty and autonomy provided by the trading fund structure.

To take this point further, under the present regulatory structure, given the close relationship of government and trading funds, there are potential commitment issues under all pricing regimes (see discussion above). Moving to a different regu-

[12]Though interestingly all of their fees are still set by Government through fees orders (more precisely the fees are set by the Lord Chancellor and then approved by HM Treasury). Thus Government still largely controls their year to year revenues (and hence, one would imagine, their investment levels and incentives). This suggests that, in this case, the major benefit of Trading Fund status was not to reduce the level of (central) Government control but to reduce the risk that Government would, especially in 'difficult times', take too great a share of Land Registry revenues for other purposes leaving the Land Registry with insufficient funds to carry on its operations. In this sense Trading Fund status could be seen as a form of 'ring-fencing' in relation to the Land Registry budget.

latory structure could improve this. For example, if trading funds were more legally independent it would permit the creation of arm's length legally-binding contracts regarding both subsidies and purchases. Combined with independent and transparent regulation this sort of structure would go a long way to eliminating concerns about the ability of government to deliver on subsidy and purchase promises and eliminate fears about the effects of such risks on the quality and availability of trading fund data.

In particular, it should certainly be emphasized that a change in charging policy, for example to use marginal cost pricing for some part of a trading fund's products, does *not* require removing their trading fund status or a reversion to 'vote-funding'. In fact, as just suggested, such a change would optimally be combined with improvements in the independence and transparency of the governance structures to provide trading funds (and government) with *more* certainty, clarity and independence than they currently have. An obvious example in this respect is provided by the case of the Train Operating Companies (TOCs), where the government has been able to agree subsidies as well as payments for long-term investment. While trading funds obviously differ from the TOCs in several respects, notably by not being privately owned companies, it would not be very difficult to design mechanisms for trading funds which that could provide a similar degree of certainty.

6.6.2 Incentives and Performance

Coming to the question of incentives and performance, the differences between charging regimes are, if anything, even less significant than when considering commitment. As already discussed, without adequate regulatory/governance structures in place, all charging regimes can result in poor incentives, inefficiency and overall poor performance.[13] Conversely with a good regulatory/governance structure in place any

[13]If the trading fund is still government owned profit-maximization here is no different since the monitoring role usually played by shareholders and the market is now the responsibility of

of the charging policies could be implemented without jeopardising the incentives, efficiency, and performance of a trading fund.

Consider the current situation, which roughly approximates to capital-based regulation – a trading fund is expected to cover costs and make some specified return on capital. As is well known, this approach has obvious problems from an incentives and efficiency perspective. First, and most obviously, the organization no longer has incentives to minimize costs but rather seeks to match costs to revenue. Furthermore, given the market power trading funds have, at least in some markets, overspending can always be addressed by raising prices and increasing revenue. Second, and relatedly, the organization now seeks to equate average costs and average revenue rather than marginal costs and marginal revenue. As a result there will be 'gold-plating' and over-investment in quality.[14] Third, and more subtly, this pricing policy provides incentives to over-invest in order to extend (inefficiently) the capital base since this then allows an increase in revenues.

These are all fairly serious issues. Thus, the Government, in its role as owner and regulator of a trading fund, needs to exert a substantial degree of effort to try and reduce or eliminate these risks. In particular, to correct these potential biases in a trading fund's behaviour it would likely need both to put in place some form of incentive scheme, and associated monitoring mechanisms. This has been the approach in other areas, for example Network Rail (which replaced the privately-

government.

[14]This is distinct from the previous point in that, for any given project, the costs may be at their, optimal, minimal level for the quality chosen, but that quality will be at inefficiently high level. To put this in terms of a simple example, suppose a purchase of a computer system is being considered. Suppose furthermore there are two manufacturers M and N and both offer a high and low quality system: MH, ML, NH, NL, and that the corresponding M and N system are equally good but the N one costs more. In addition suppose the high quality system is four times as expensive but that revenue is only twice as much. Suppose also that the resulting revenue from buying the high-quality system is just sufficient to cover its costs. Then inefficiency in the first sense, would be to choose the N system over the M system (perhaps because one failed to do enough research about what was on offer) – so more money would would have been spent than was needed. Inefficiency in the second sense would be choose the high quality system – profits are zero in this case but would have been equal to a quarter of this revenue if the low quality system were chose (costs fall to a quarter but revenue by only a half).

owned RailTrack), while run as a not-for-profit company limited by guarantee has put in place a fairly complex incentives package for managers and is also monitored by the Office of Rail Regulation.[15]

A similar approach could also be taken if a marginal cost pricing regime were adopted. Here too there are issues though it will be assumed that a subsidy can be provided in a transparent and commitable way (see the extensive discussion above). Specifically, just as with average-cost pricing, the government (or the regulator if independent) would need to think carefully about providing incentives for (efficient) reduction in costs (while keeping investment at the optimal level). To put this in more concrete terms, for those products priced at marginal cost the government (or regulator if distinct from government) would need to be setting a subsidy level. This subsidy would likely be tied to (previous) output and expenditure in some manner. One option would be to set the subsidy to equal fixed costs in the last period. This would result in poor incentives to lower costs. Similarly setting a straight per output subsidy might lead to over-investment.

However with a little effort and combining these two approaches one could develop something a lot better. To provide just one example, one could estimate a particular periods fixed costs using previous periods fixed costs (multiplied perhaps by a deflator), and then use this, together with some estimate of the value of usage, to set the per unit of output subsidy. The reason for incorporating output is that this would ensure a trading fund has incentives to get their data used (whether by making it easy to use, publicizing it etc etc). Additionally, incorporating output measures this approach makes it easier to allow for the introduction of a new data products – which is an important factor to consider when managing marginal cost pricing.[16]

[15]To some extent such mechanisms are already being used in relation to trading funds. For example, we were informed by HM Treasury that in some cases a trading fund's sponsoring department within Government ties expenditure to measures of efficiency gains.

[16]There are other ways to address this. For example one could follow a system used by the

This is just one example, and clearly that the regulator would need to consider in more detail. However it should be sufficient to demonstrate that the problems are not insurmountable, and are, in many ways, little different from the issues confronting government when it uses a cost-recovery approach.[17] What is clear in both cases is that there are incentives questions to be addressed, and if they are not, there would likely be serious detrimental impacts on efficiency and general performance. However as long as reasonable thought and effort are put into dealing with these issues, in particular by designing a robust governance/regulatory regime, these negative consequences can be avoided.

6.7 Conclusion

Much of the concern about the impact of a change in charging policy (particularly to marginal cost or zero cost) is based on a misidentification of charging policy with regulatory structure. Moving a trading fund back 'inside' Government, and thereby making it dependent on year-to-year 'vote-funding', might well have substantial negative impacts – but it would do so whatever charging policy was being followed. Conversely, any of the charging policies under consideration could be followed successfully, and without these kinds of negative effects, if a independent, transparent and coherent governance structure were in place. In this regard charging policy can

TOCs who present a 'shopping-list' to Government of possible capital improvement projects which Government then chooses form. Alternatively one could provide some way for users to feed back requirements to trading fund regarding new datasets to collect. This is also a major advantage to having a trading fund retain a 'Retail' arm in additional to any marginal cost 'Wholesale' arm as 'Retail' can pass on feedback regarding their requirements to 'Wholesale' (in fact, the Met Office stated that something like this already occurs with their 'Retail' division passing back feedback to 'Wholesale' as to what new kinds of data would be useful in the provision of their own products and services).

[17] Though the very fact that, because of the need for a subsidy, these kind of calculations are more out 'in the open' is a significant advantage of marginal cost pricing. Such an increase in transparency benefits all concerned, and, furthermore, requires that a regulator have access to the relevant cost and output data from a trading fund on a regular basis, and thus, could be seen as a way of credibly committing the government to a more transparent and active governance/regulatory regime in future.

largely be seen as *orthogonal* to the question of trading fund performance – whether evaluated in terms of quality, responsiveness or efficiency. Moreover, the importance of having an adequate governance structure – whatever charging policy is chosen – cannot be overemphasized.

Much of that structure is already in place though, as already discussed, there are likely several important ways in which it could be extended in pursuit of delivering on the key goals of transparency, certainty, independence and incentivisation. If such an adequate governance/regulatory structure in place – and there seems every reason to be confident that it would be given Government's substantial previous experience in these matters – then there is every reason to be confident that any of the major pricing policies considered in this report can be implemented without adverse effects on the efficiency and performance of the Trading Funds affected.

Issue	Profit Maximization	Average Cost/Cost Recovery	Marginal Cost/Zero Cost
Commitment	Good.[a]	Largely dependent on regulatory / governance structure.	Largely dependent on regulatory / governance structure.[b]
Incentives	Optimal for trading fund though likely non-optimal for other market participants (see next item).	Risk of over-investment and inefficiency (costs too high). Monitoring required of investment, quality and costs.	Risk of either over or under performance depending on subsidy function. Monitoring required of investment, quality and costs.
Distortion of Competition	Unrefined/upstream: major issue given dominant position of trading funds. Refined/downstream: minor as long as cross-subsidy is limited.	Significant issue if trading fund provides internal access to upstream material on different terms to external firms (esp. if cost allocation between upstream and downstream is opaque).[c]	Minor.[d]
Information	Not relevant as no regulation.	Single point on demand curve where revenue covers costs.[e] At aggregate level know trading fund covers total costs.	Single point on demand curve where price equals marginal cost.

[a]Though could depend on relationship of government and trading fund – particularly risk that profits are ex-post 'appropriated'

[b]Could be a greater issue than under 'average-cost' because here the government may be providing subsidies.

[c]Oversight would still be required here to prevent the use of discriminatory tariffs. For example a trading fund could set as tariff a very large one-off fee for all its data. This might exclude, for example, external users who only need a small part of that data. Similarly without transparent cost allocation under average cost pricing a trading fund might have an incentive to overcharge for upstream access to exclude downstream entrants – a problem familiar from the telecommunications literature, see e.g. Farrell (2003).

[d]Though the provision of subsidy may retard entrants who wish to compete directly with the trading fund in the provision of data. However, as long as the marginal cost data provision were largely confined to those datasets of which the trading fund was sole provider this would not become an issue.

[e]Though where a trading fund performs cost-recovery only at the aggregate level the exact relation of revenue to costs for a given product may be unclear.

Table 6.1: Charging Policies and Regulatory/Governance Issues.

Chapter 7

Conclusion

7.1 General

This study has analyzed the impact of adopting different models for the provision of public sector information by trading funds. Its basic task has been to examine the cost and benefits for society, and the effects on government revenue, of four different charging policies: profit-maximization, average cost (cost-recovery), marginal cost and zero cost; both on their own and when interacted with various data distinctions such as raw versus value-added, and unrefined versus refined.

The study focused on the six largest trading funds by data provision: the Met Office, Ordnance Survey, the UK Hydrographic Office, the Land Registry, Companies House and the Driver Vehicle Licensing Agency. Starting from the general theoretical framework set out in Chapter 3, Chapter 4 reviewed the general empirics in preparation for Chapter 5 which analysed each trading fund individually.

There is a general proposition[1] that public sector goods and services should be offered at efficient prices, unless there are compelling reasons to depart from efficiency. In the absence of beneficial spill-overs, the efficient price is marginal cost (with supply adapted such that the short and long-run marginal costs are equal).

[1] Discussed further in Appendix A.1.

One reason for departing from efficient pricing is that the marginal cost is below the average cost, and that the benefits of a hard budget constraint outweigh the distortionary costs of raising the revenue to make up the short-fall, not from general taxation but from raising the price of the products supplied.

For many public utilities the difference between average and marginal cost is quite small,[2] and the distortionary costs are a small price to pay for the benefits of commercial pressure. In other cases two-part tariffs allow any short-fall from marginal pricing at marginal cost to be recovered from a connection charge, for which the demand is much less elastic. But the costs of raising the revenue to cover the short-fall between average and marginal cost (i.e. the fixed costs) increase as the square of the margin required, so if marginal costs are considerably below average costs (as for information goods), the deadweight losses are likely to be high.

In some cases, notably the Land Registry, Companies House and the DVLA, something close to two-part tariffs are possible, with the charges for registration being a charge to recover the fixed costs, and the provision of the resulting data being at marginal cost. This is in accordance with good Ramsey pricing principles that if distortionary mark-ups are necessary to cover or contribute to fixed costs, they should be higher for inelastically demanded goods and lower for elastically demanded services (in simple cases, the mark-up divided by the price should be inversely proportional to the elasticity of demand).

This report has shown that the case for pricing no higher than marginal cost (which, for most digital data will be zero) on basic data products is very strong, for a number of complementary reasons. First, the distortionary costs of average rather than marginal cost pricing are likely to be high, for several reasons. The mark-up to cover fixed costs is high, as marginal costs are such a low fraction of average costs. The demand for digital data as with other information services is likely to be high

[2]Thus the long-run marginal cost of expanding the high pressure gas network is quite close to the average cost allowed by the regulator.

and growing. Finally, there are likely to be large beneficial spill-overs in inducing users to innovate new services based on the data, as is evidently the case for other ICT services. Second, the case for hard budget constraints to ensure efficient provision and induce innovative product development is weak for public enterprises not subject to regulation and providing monopoly services without fear of competition. It would be far better to address issues of incentives, regulation and commitment explicitly rather than indirectly through budget constraints. Finally, for several services, the Government is already providing effectively a large contribution to fixed costs, without allowing the public to enjoy the benefits of efficient pricing.

The report has, however, followed the injunction to quantify the costs and benefits of moving from cost-recovery pricing to marginal cost and/or zero pricing without stressing these more general considerations. Where it is found that cost-recovery pricing may not be very damaging compared to the preferable solution of efficient pricing, that has been stated clearly, and the report is biased against the (strong) presupposition that marginal cost pricing ought always to be preferable, allowing the benefit of the doubt to those who would argue for continuing the present regime unless the arguments against are almost irresistible.

7.2 Individual Trading Funds

The general results of the analysis have already been presented in the Executive Summary above and the points set out there will not be repeated here. Instead summaries for each trading fund are provided, detailing for which products a change in pricing regime would be welfare improving and the overall associated benefits and costs. These summaries are excerpted from the full analysis provided in Chapter 5 and for further details the reader should see the relevant trading fund section of that chapter.

7.2.1 Companies House

The analysis focused on Companies House's 'bulk' ('wholesale') data, specifically their 'CD ROM' and 'Bulk Data and Image' product categories, and for all other products it was assumed that the pricing regime would remain unchanged.

Focusing on these two categories, the results suggested that under the likely range of parameter values a change from an average cost to a marginal cost regime would be welfare improving. Specifically, the calculations showed that adopting a marginal cost pricing policy for these particular products would result in gross benefits of around £2.6m with government incurring net costs of around £681k. This figure is calculated on the basis that Government would make up the loss in revenue of 946k suffered by Companies House itself.[3] However given the registration-based nature of Companies House it is feasible that this shortfall could be covered by the registration side of its operations.[4] Being conservative and leaving aside this possibility for the time being, and using the figures for benefits and costs just enumerated, the overall net welfare gain to society from the change envisaged would be approximately £1.9m.

7.2.2 The Met Office

Given the data available the analysis has focused on the Met Office's 'wholesale' products.[5] The focus was further narrowed down by the exclusion of service products and those which consist primarily (or entirely) of third party data (see Section 5.1.1

[3]Note that for these product categories Companies House's costs here were actually substantially lower than revenues (653k versus 946k). Thus if Government were simply to make up costs rather than revenues the figure would be lower.

[4]Since these would be costs incurred in maintaining and created the registration database it is likely that there would not be any conflict with the existing legal restrictions on Companies House charging policy (see footnote on page 110) – though this is something that might need to be investigated further.

[5]The Met Office is unusual among trading funds in explicitly dividing its upstream (wholesale) arm from its downstream (retail) arm with retail purchasing from wholesale on the same terms as any other user.

for the reasoning behind this). While only accounting for a very small proportion of overall costs and revenue (1% or below) these wholesale data products make available almost all of the basic data the Met Office collects itself.

For this set of products, the results of the analysis suggested that under the likely range of parameter values a change from average cost to marginal cost regime would be welfare improving. In particular, gross benefits would be around £1.2m with costs to government of only around £260k. The actual additional payment required to the Met Office from government would be equal to the loss in revenue which in this case is £390k. This would be on top of the existing £68m in direct subsidy and £57m in sales coming from government. Thus the change in policy would require a 0.3% increase in current expenditure by government in this area.

Finally, putting together the benefits and the costs implies an overall net benefit to society of £1.03m and a return on investment of approximately 500%.[6]

7.2.3 Ordnance Survey

Given the data available the analysis has focused down on two main product categories: 'Large Scale Topographic' and 'Transport Network Products', both of which consisted solely of unrefined products. Together the products considered account for around £70m of the Ordnance Survey's £114m of revenue in 2006/2007 (and would certainly account for the vast majority of revenue from unrefined products). The following discussion is therefore premised on the assumption that the charging policy for other products remains unchanged.[7]

For these two product categories, the analysis suggested that a change from an

[6]Adding in the loss of ECOMET royalties – and assuming no corresponding benefit to society since these gains accrue to non-UK citizens elsewhere in the EU – would increase government costs by 56k. Subtracting out the NWP product to leave only unrefined items would reduce government costs by 193k to only 66k but would also reduce gross benefits to 866k and net gains to 800k.

[7]That said, one could probably extend the conclusions to all unrefined products without the results changing very substantially. This is because the great bulk of revenue from unrefined products is already accounted for by these two categories. Thus, even if the data had been available to extend the analysis to all unrefined products, the results would have remain largely unchanged.

average cost to a marginal cost regime would increase welfare. Specifically, gross benefits would be around £168m a year while net costs to government would be around £12m. Overall this implies an overall net benefit to society of 156m.[8]

The actual increase in subsidy required would obviously be higher than the £12m net impact on government as this figure includes the benefits of increased tax revenues. Taking the simplest approach the increase in subsidy would be equal to the loss of revenue from non-government sources, which for these two categories combined would be around £30m. Adding this to existing payments from government would make a grand total of £85m. By comparison it should be noted that the Met Office currently receives around £125m from government each year. The comparison with the Met Office is instructive in other ways here as well. There, the CAA (Civil Aviation Authority) contributes to the subsidy for the PWS (Public Weather Service) in addition to the government. Looking at the Ordnance Survey, if some sort of analogous PGS (Public Geodata Service) subsidy were being considered it might also be possible to have contributions spread across a variety of organizations (and departments within government).[9]

7.2.4 The UK Hydrographic Office

The analysis focused on Digital UK Charts, Digital Publications and Licensing, and for all other products it was assumed that the pricing regime would remain unchanged.

Focusing on these products, the results suggested that under the likely range of parameter values a change from an average cost to a marginal cost regime would

[8]Since the amounts involved here are substantially above those for most other trading funds it is worth emphasizing that while, necessarily sensitive to the particular parameter values chosen even using the most conservative plausible parametrisations a change to marginal cost pricing delivers a net welfare gain.

[9]Concretely, a 'PGS subsidy' could start out based on the existing CPAs which cover government (e.g. the PGA and the MSA) but also include other large current or potential users (e.g. the NHS, large utility companies etc) who have an interest in maintaining the quality and availability of UK geographic information.

be welfare improving only for 'Copyright Licensing'. Specifically, the calculations showed that adopting a marginal cost pricing policy for Copyright Licensing (and maintaining average cost pricing for all other products) would result in gross benefits of around £1.08m with government incurring net costs of around £744k (on a direct subsidy of 854k). Overall this means that the net welfare gain to society would be around £338k.

7.2.5 The Land Registry

The Land Registry offers no bulk data downloads for all of its data. As there are no other feasible ways to access all (or substantial portions) of the Land Registry's data service bulk data access and reuse is effectively prohibited for the core parts of the Register. This could restrict competition with the Land Registry's search services and hold back the innovative use of the data. A comparison with Companies House would indicate that introducing bulk data provision and charging at marginal cost is likely to raise social welfare significantly. However no precise calculation could be made due to limits of the available data. Concerns may be raised concerns over privacy, however these are questions about whether data should be made available at all, rather than the price at which data it is provided. Furthermore, it is likely that many of these concerns could be addressed by adopting some form of anonymisation (as usesd by the DVLA) and specific license conditions to prohibit forms of reuse that are unwanted (for example unsolicited direct mailing).

Analysis for Property Price data and Polygon GIS data was possible. This indicated that under the likely range of parameter values a change from an average cost to a marginal cost regime for these individual products would be welfare improving. Specifically the calculations showed that lowering the price to marginal cost would result in gross benefits of around £2.3m with government incurring net costs of around £1.1m (£1.2m gross) if it were to finance the policy. Overall this means

that the net welfare gain to society would be around £1.2m.

These figures assume that the fall in revenue of £1.2m at the Land Registry would be made up by government. This does not take into account the possibility of making up the deficit via charges for registrations. In this case the financial burden would not fall on government. If the elasticity of demand for registrations is low, which is likely to be the case when registration is compulsory, this would reduce the costs (and increase the net benefits) of moving to marginal cost pricing.

7.2.6 The Driver and Vehicle Licensing Agency (DVLA)

The analysis focused only on Anonymised data, Bulk data and Mileage data, and for all other products it was assumed that the pricing regime would remain unchanged.

For these examined categories, the results suggested that under the likely range of parameter values a change from an average cost to a marginal cost regime would be welfare improving. Specifically, the calculations showed that adopting a marginal cost pricing policy for these particular products would result in gross benefits of around £4.3m with government incurring net costs of around £582k if it were to finance the policy (gross costs would be around £1m). Overall this means that the net welfare gain to society would be around £3.7m.

In the case of the DVLA central Government need not necessarily provide any funds itself. Specifically it might be preferable to finance lower priced data from registrations (or from Excise Duty) as demand for these services is likely to be (relatively) more inelastic than for data. With the loss of revenue in moving to marginal cost at just 0.16% of total revenues (1m out of 613m) and the DVLA's rate of return on capital to March 2007 of 16.6% this seems a very viable option.[10]

[10]See DVLA's 2006-07 annual accounts available at `http://www.dvla.gov.uk/media/pdf/publications/annual\_accounts0807.pdf`

Appendices

Appendix A

General

A.1 A General Argument for Selling Public Sector Products at Marginal Cost

It is a standard economic argument, formalised by Mirless and Diamond, that if the private sector is competitive and undistorted, then the public sector should sell any goods and services that it produces at the efficient price, unless there are compelling reasons for insisting on breaking even (such as providing adequate incentives and commitment, discussed in chapter 6). The efficient price is the marginal cost, unless there are positive externalities, in which case the efficient price is below marginal cost.

The argument is that the distortions and losses caused by taxes should be confined to the consumption side of the economy, leaving production undistorted, and hence delivering the level of demand (at distorted consumer prices) at least resource cost. If production is inefficient, and if consumer taxes (including labour income taxes) are potent, and can be adjusted, then it should be possible to increase output and consumption without sacrificing any tax revenue. The main proviso is that it should be possible to isolate consumer income (if necessary by adjusting consumer

taxes) from any changes to production (e.g. when removing distortions). Put simply, if changes in production efficiency do not give leverage over raising revenue or redistributing income that could not otherwise be achieved by changing consumer taxes, then it is best to aim at overall production efficiency. That in turn means that if private production is competitive (more narrowly, the relevant public production has no direct leverage over the any departures from efficient pricing in the private sector) then public production should also price efficiently, i.e. at marginal cost (or below with beneficial spill-overs).

Put another way, taxing public production (by the difference between price and marginal cost) is inefficient if the production is an input into production, and unlikely to be part of an optimal commodity tax system when sold as a final good. Indeed, Diamond and Mirrlees (1971) also characterise the optimal commodity tax system, which should be a Value Added tax except for corrective excises (to deal with such externalities as pollution) and as such would exempt the production sector -which is why pricing above marginal cost for inputs into production is inefficient. The optimal VAT on final consumption goods will be uniform unless taxes on specific goods directly affect the labour-leisure choice - and it is difficult to identify such goods. Certainly it is hard to believe that taxing any PSI products would increase consumers willingness to undertake taxed labour activities, or that reducing their price would lead to an increase in leisure at the expense of paid employment.

If this argument is accepted then the case for selling PSI at marginal cost is direct, and it is hardly necessary to estimate the extra social value that would be generated by moving from average to marginal cost pricing. Indeed, it is generally quite difficult to do that, as changes in prices in a full equilibrium model of the kind needed for tax *system* design will typically lead to changes in demands and supplies elsewhere, and these will have impacts on total tax revenue, and/or the budget. An attempt is made in this report, but it is far from a comprehensive general equilibrium

analysis with a full description of the existing tax and benefit system, and as such must be treated with some caution. The attraction of a general argument such as this is that it rests on a theory that properly treats the rest of the tax system, and the ability of the Government to change other taxes and/or expenditures at the same time as changing its policy towards PSI pricing.

A.2 Calculating θ

The social value of a reduction in the price of a trading Fund PSI depends on the social valuation of the extra consumption that users enjoy. If we use the Treasury's (2003) *Green Book* implicit social welfare approach, implied by the annexes on distributional impacts (Annex 5) and discounting (Annex 6), a key parameter is the elasticity of the marginal utility of consumption, μ in the *Green Book* (more properly called the social marginal utility of consumption, and equal to the coefficient of inequality aversion, ν). The UK Government attaches importance to the distributional consequences of its actions, many of which are justified by the beneficial impact they have on distributional outcomes. Indeed, any Government that was unconcerned with equity would choose taxes that were least distortionary, and these would be on inelastically demanded goods and services, i.e. on necessities; and in extreme cases by lump sum taxes. Once it is recognised that equity is of concern, then the social values of equal transfers to those at different standards of living are no longer equal - and one cannot treat £1 of consumption as equally valuable no matter who receives it.

"Consumption" is therefore not a well-defined numeraire without further specification (whose consumption?). It is in any case more logical to choose uncommitted public funds (i.e. funds that can be used either for public investment, redistribution, or funding activities such as the provision of PSI products) as the numeraire, particularly as such funds will need to be used to subsidise the production of PSI

products if they are to price below average cost, as recommended here. It is then necessary to find the social value of transfers to different consumers. The welfare weight to attach to a consumer with total expenditure c_h can be derived from the Social Welfare Function, which has the general form

$$W = \sum U(c_h) = \frac{kc_h^{1-\nu}}{1-\nu},$$

(where we have replaced the *Green Book* μ with ν to avoid confusions with mean values of distributions that we shall need later). In the *Green Book* $\mu = \nu = 1$ corresponding to $U(c_h) = \ln(c_h)$. The social marginal utility of an extra £1 for consumer h is then $\beta_h \equiv dU/dc_h = kc_h^{-\mu}$, or in the *Green Book* case, $\beta_h = k/c_h$. If at the margin the best way for the Government to redistribute tax revenue for distributional purposes is through equal lump-sum grants,[1] then the social value of giving £$1/n$ to each of n consumers is $\frac{1}{n}\sum \beta_h = \frac{1}{n}\sum kc_h^{-\nu}$, and its value must be £1, its value as uncommitted funds in the hands of the Government. This effectively defines the value of k and hence β_h:

$$\beta_h = \frac{c_h^{-\nu}}{\frac{1}{n}\sum c_h^{-\nu}}.$$

The reference level of consumption, $\widehat{c}$ is defined as that level of consumption for which $\beta = 1$:

$$\widehat{c} = \left(\frac{1}{n}\sum c_h^{-\nu}\right)^{-1/\nu}.$$

The reference level of consumption would be such that the Government would be happy to make £1 transfer out of uncommitted funds to people at and below that level (if to do so had no other effects on incentives, etc.). It can be readily

[1] This seems a reasonable implication of observed Government expenditure, for pensions, health, and most education is provided equally to all. Most other ways of making transfers (e.g. by subsidising particular goods or services such as rail travel) gives relative more subsidy to those with higher incomes (as income elasticities are positive for the overwhelming share of expenditure.

calculated from household expenditure data as in the *Green Book* p94. It simplifies the value of β_h to $\beta_h = (\widehat{c}/c_h)^\nu$. In the *Green Book* case where $\mu \equiv \nu = 1$, $\beta_h = \widehat{c}/c_h$, inversely proportional to consumption.

We can now derive the social value of any consumer surplus generated by lowering the price of a PSI product. Let household h have total consumption c_h and demand for a PSI product $q_h = AQc_h^\eta$ where Q is total demand and $A = 1/\sum c_h^\eta$, constant. If prices fall, then each consumer benefits in proportion to their demand q_h, i.e. by δc_h^η. and each £1 of total consumer surplus will therefore be worth $\theta = \delta \sum \beta_h c_h^\eta$, where $\delta \sum c_h^\eta = £1$, or

$$\theta = \frac{\widehat{c}^\nu \sum c_h^{\eta-\nu}}{\sum c_h^\eta}.$$

If, as is a good approximation empirically, c_h is distributed log-normally with parameters (μ, σ), and if we relace summations by expectations, then since for the log-normal distribution

$$Ec^a = \exp(\alpha\mu + \frac{1}{2}\alpha^2\sigma^2),$$

we can first evaluate

$$\begin{aligned} \widehat{c}^\nu &= \left(Ec_h^{-\nu}\right)^{-1}, \\ &= (\exp(-\nu\mu + \frac{1}{2}\nu^2\sigma^2))^{-1} \\ &= \exp(\nu\mu - \frac{1}{2}\nu^2\sigma^2). \end{aligned}$$

The value of θ is then

$$\begin{aligned} \theta &= \frac{\widehat{c}^\nu Ec_h^{\eta-\nu}}{Ec_h^\eta} = \frac{\widehat{c}^\nu \exp((\eta-\nu)\mu + \frac{1}{2}(\eta-\nu)^2\sigma^2)}{\exp(\eta\mu + \frac{1}{2}\eta^2\sigma^2)} \\ &= \exp(-\nu\eta\sigma^2). \end{aligned}$$

In the special case of $\nu = 1$

$$\theta = \exp(-\eta\sigma^2).$$

For the UK an estimate of $\sigma = 0.47$, and on average goods have unit income elasticity so $\eta = 1$, in which case $\theta = 0.8$, $1/\theta = 1.25$. If PSI goods are typical goods then this is a reasonable estimate, but if they have higher income elasticities, say $\eta = 1.5$, $\theta = 0.7$, $1/\theta = 1.4$.

A.2.1 The marginal cost of public funds, MCPF

The marginal cost of public funds is a term often encountered in social cost benefit analysis, and purports to measure the full cost including the deadweight loss to consumers of raising £1 to public funds using distortionary taxation. We have already noted that "consumers" need further characterisation before one can measure the social value of this full cost, and once we have done this, we are effectively comparing the social value of £1 in the hands of this consumer with £1 as uncommitted public funds. The *Green Book* (p94) is rather vague on the reference level of consumption to use in computing distributional weights, but other analysts often use average consumption as the reference point. In terms of our public funds numeraire, the social weight of average consumption is $\beta = (\widehat{c}/Ec_h)^\nu$ and using the log-normal assumptions above, this is

$$\begin{aligned} \beta &= \frac{\exp(\nu\mu - \frac{1}{2}\nu^2\sigma^2)}{(\exp(\mu + \frac{1}{2}\sigma^2))^\nu} \\ &= \exp(-\frac{1}{2}\nu(\nu+1)\sigma^2) \end{aligned}$$

which for $\nu = 1$ is $\beta = \exp(-\sigma^2) = 0.8$. The MCPF in terms of average consumption is then $1/\beta = 1.25$, which is close to other estimates derived by a variety of more or less reasonable techniques. Note that if inequality aversion is not equal to unity, then the relationship between these various estimates becomes more complicated,

although consistency with *Green Book* avoids these complications. Note also that $\beta = \theta$ for the case of unit income elastic goods and $\nu = 1$. The MCPF is sometimes written as $1 + \alpha$, where α is often described as the distortion cost of raising the revenue, whereas in fact it is a combination of equity and distortionary costs in a distributionally sensitive optimal tax system. In terms of average consumption levels,

$$\alpha = \exp(\frac{1}{2}\nu(\nu+1)\sigma^2) - 1 \approx \frac{1}{2}\nu(\nu+1)\sigma^2.$$

A.3 Review of the estimates of the MCPF

The marginal cost of public funds measures the total cost of raising an addition pound of tax revenue. There are generally two different approaches in which to estimate the marginal cost of public funds: under differential analysis (D) government revenues are returned lump sum to tax payers or the marginal tax is compared against a lump sum tax; under balanced budget analysis (BB) the revenue is not returned but instead contributes to government expenditure that has no further effect on private utility or the labour supply. The latter approach tends to estimate lower marginal costs of public funds; the labour supply increases in response to lower private incomes, which generates additional revenue, and so the distortionary costs are less for a given rise in taxes.

Raising government revenue to subsidise trading funds falls between the two approaches, since on one hand, the revenue is not returned lump sum to tax payers, but on the other hand, the lower cost of PSIs will raise the public's utilities. Estimates will also rise with greater degrees of progressivity of any tax rise, higher parameter values such as elasticity of labour supply and due to further specifications to the modelling. Taking these factors into consideration, the literature has been surveyed and the results have been summarised in Table A.1.

On examination of the literature it would appear that a reasonable range for the

Author(s)	**Approach**	**MCPF Estimates**	**Detail**
Stuart (1984)	D BB	1.21-1.57 1.07-1.43	US study
Ballard, Shoven, and Whalley (1985)	BB	1.12-1.23 (wage tax) 1.16-1.31 (income tax) 1.18-1.46 (capital tax)	US study
Browning (1987)	D BB	1.30-1.47 1.14-1.20	US study
Fullerton and Henderson (1989)	BB	1.06-1.17 (wage tax) 1.14-1.25 (income tax) 1.26-1.31 (corporate tax) 1.25 (capital tax)	US study
Thirsk and Moore (1991)	D BB	1.22-1.37 (prop. wage tax) 1.42-1.81 (prog. wage tax) 1.18-1.27 (prop. wage tax) 1.30-1.45 (prog. wage tax)	Canadian study
Ballard and Fullerton (1992)	BB	0.936-1.147 (prop. wage tax) 1.025-1.315 (prog. wage tax)	US study
Ruggieri (1999)	D	1.13	Models MCPF in a small open economy, which is appropriate to the UK as opposed to close economy models suited to the US. [2]
Feldstein (1999)	D	1.32-1.52	US study incorporating tax avoidance.
Parry (2002)	D/BB	1.3-1.5	US study incorporating tax avoidance
Parry (2003)	D	1.16-1.51	UK study
Allgood and Snow (2006)	D BB	1.15-1.21 (data set A) 1.06-1.07 (data set B) 1.07-1.13 (data set A) 1.04-1.05 (data set B)	US study incl. human and physical capital. A uses higher ranges of avg. and marginal tax rates and elastities of labour supply than B

Table A.1: Estimates of the Marginal Cost of Public Funds

marginal cost of public funds is 1.15 to 1.35. That is to raise 1 of tax revenue, costs an additional 15 to 35 pence. Where necessary a point estimate of 1.25 will be used. This is very similar to an estimate used by HM Treasury in Appendix C of Treasury (2000) who, based on Ruggieri (1999), use a figure of 1.2 to 1.3.[3] This range does not include some of the particular high or low estimates such as those in Feldstein (1999) and Ballard and Fullerton (1992).

Feldstein (1999) models the deadweight loss of changes in taxable income (as opposed to changes in labour supply) in the US. This includes distortions to the amount spent on housing and medical services as these are tax deductable or exempt in the US. These estimates may not be so applicable to the UK where these tax subsidies have been phased out Parry (2003).

Ballard and Fullerton (1992) find in some cases a MCPF of less than one as a consequence of using negative uncompensated elasticities of labour supply. However there is no clear consensus on whether this elasticity is positive or negative (Blundell, 1992). They use a negative value to illustrate the following misunderstanding:

'Under one view, the marginal cost of public funds must be greater than one. However, under an alternative view, the MCF can actually be less than one. We will illustrate this possibility using numerical examples for labor taxes. In this case, the MCF reduces the costs of the project.' Ballard and Fullerton (1992) p118. It would therefore appear that their intention was not to provide a representative estimate of the MCPF, and hence why it is not within our chosen range. Ballard and Fullerton (1992) also point out that 'The MC(P)F ultimately depends not just on the tax, but also on the nature of the government expenditure under consideration.' This is a particularly salient point in the case of government revenue subsidising trading funds in order to offer below average cost pricing. As an example, the lower cost of trading fund data may lead to greater innovation. On the one hand this could

[3]The inclusion of more recent empirical studies, such as Parry (2003), gives reason to employ a larger range than in Treasury (2000).

result in higher corporate incomes, which would contribute to subsequent higher government revenues and hence a lower MCPF. On the other hand the lower costs of trading fund data may be passed onto lower final goods prices. This case would leave the public with more income to spend on other goods and services, and could weaken incentives to supply labour. This time the lower government revenue would raise the MCPF.

Appendix B

Supplementary Data

B.1 UKHO Costs Analysis

SECTION OMITTED FOR CONFIDENTIALITY REASONS

Bibliography

Sam Allgood and Arthur Snow. Marginal Welfare Costs of Taxation with Human and Physical Capital. *Economic Inquiry*, 44(3):451–464, July 2006.

Charles L Ballard and Don Fullerton. Distortionary Taxes and the Provision of Public Goods. *The Journal of Economic Perspectives*, 6:117–131, 1992. ISSN 08953309.

Charles L Ballard, John B Shoven, and John Whalley. General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States. *American Economic Review*, 75(1):128–38, March 1985.

Ravi Bedrijvenplatform. Economische effecten van laagdrempelige beschikbaarstelling van overheidsinformatie, 2000. Private sector members of the Dutch Geographic Data Committee.

James Bessen and Eric Maskin. Sequential Innovation, Patents, and Innovation. NajEcon Working Paper Reviews 321307000000000021, www.najecon.org, May 2006.

APPSI Review Board. Report in relation to requests by Intelligent Addressing Limited and Ordnance Survey to review certain recommendations made in the Report of the Office of Public Sector Information of 13 July 2006 relating to a complaint by Intelligent Addressing Limited (SO 42/8/4), 2007.

Edgar K Browning. On the Marginal Welfare Cost of Taxation. *American Economic Review*, 77(1):11–23, March 1987.

F A Cowell and K Gardiner. Welfare Weights. Economics Research Paper 20, STICERD, London School of Economics, August 1999.

James B. Davies and Al Slivinski. The Public Role in Provision of Scientific Information: An Economic Approach. University of Western Ontario, RBC Financial Group Economic Policy Research Institute Working Papers 20051, University of Western Ontario, RBC Financial Group Economic Policy Research Institute, 2005.

Joseph Farrell. Integration and Independent Innovation on a Network. *AER Papers and Proceedings*, pages 420–424, 2003.

Martin Feldstein. Tax Avoidance And The Deadweight Loss Of The Income Tax. *The Review of Economics and Statistics*, 81(4):674–680, November 1999.

Office for National Statistics. Family Spending: A Report on the 2003-04 Expenditure and Food Survey, 2005.

Don Fullerton and Yolanda Kodrzycki Henderson. The Marginal Excess Burden of Different Capital Tax Instruments. *The Review of Economics and Statistics*, 71 (3):435–42, August 1989.

Austan Goolsbee. The Value of Broadband and the Deadweight Loss of Taxing New Technology. *Contributions to Economic Analysis & Policy*, 5(1):1505–1505, 2006.

Austan Goolsbee and Peter J. Klenow. Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet. *American Economic Review*, 96(2):108–113, May 2006.

Peter Hackl and Anders H. Westlund. Demand for international telecommunication time-varying price elasticity. *Journal of Econometrics*, 70(1):243–260, January 1996.

K. N. Hampton. Neighborhoods in the network society: the e-Neighbors Study. In *Information, Communication and Society*. 2007.

Jerry A Hausman. Exact Consumer's Surplus and Deadweight Loss. *The American Economic Review*, 71:662–676, September 1981. ISSN 00028282.

Jerry A. Hausman, Ariel Pakes, and Gregory L. Rosston. Valuing the Effect of Regulation on New Services in Telecommunications. *Brookings Papers on Economic Activity. Microeconomics*, 1997:1–54, 1997.

F. J. Hellinger. Focus on Research: HIV Disease. Technical report, Agency for Healthcare Research and Quality, United States Department of Health and Human Services, 2002.

HM Treasury. The Green Book, 2003.

Adam Jaffe. Real Effects of Academic Research. *American Economic Review*, 79: 957–970, 1989.

Donald Kridel, Paul Rappoport, and Lester Taylor. The Demand for High-Speed Access to the Internet: The Case of Cable Modems. In D.G. Loomis and Kluwer L.D. Taylor, editors, *Forecasting the Internet: Understanding the Explosive Growth of Data Communications*. Academic Publishers, Dordrecht, 2002.

Jeffrey K. Lazo and Lauraine G. Chestnut. Economic Value of Current and Improved Weather Forecasts in the US Household Sector. In *Report prepared for the National Oceanic and Atmospheric Administration by Stratus Consulting Inc. Boulder, Colorado*. 2002.

Dr Grace Lomax. Presentation to Patient Compliance, Adherence and Persistence Conference, 2005.

Roger Longhorn and Michael Blakemore. Re-visiting the Valuing and Pricing of Digital Geographic Information. *Journal of Digital Information*, 4(2), 2004.

E. Mansfield. Academic Research Underlying Industrial Innovations: Sources, Characteristics and Financing. *Review of Economics and Statistics*, 77:55–65, 1995.

Ed Mayo and Tom Steinberg. The Power of Information, 2007.

James Mirless and Peter Diamond. Optimal Taxation and Public Production: I. *American Economic Review*, 61:8–27.

OECD, Working Party on the Information Economy (Directorate for Science Technology and Industry). Digital Broadband Content: Public Sector Information and Content. Technical report, 2006.

Office of Fair Trading. The Commercial Use of Public Information, 2006.

OFT. Commercial Use of Public Information; Annexe G: Economic value and detriment analysis, December 2006.

Ian Parry. Tax Deductions and the Marginal Welfare Cost of Taxation. *International Tax and Public Finance*, 9(5):531–552, September 2002.

Ian W H Parry. On the Costs of Excise Taxes and Income Taxes in the UK. *International Tax and Public Finance*, 10(3):281–304, May 2003.

PIRA. Commercial Exploitation of Europe's Public Sector Information. Final Report for the European Commission, Directorate General for the Information Society. Technical report, 2000.

Rufus Pollock. Cumulative Innovation, Sampling and the Hold-up Problem. DRUID Working Papers 06-29, DRUID, Copenhagen Business School, Department of Industrial Economics and Strategy/Aalborg University, Department of Business Studies, 2006.

S. Rodgers and Q. Chen. Internet community group participation: psychosocial benefits for women with breast cancer. *Journal of Computer-Mediated Communication*, 10(4), 2005.

Giuseppe Ruggieri. The marginal cost of public funds in closed and small open economies. *Fiscal Studies*, 20(1):41–60, March 1999.

Spatial Information Industry Action Agenda. Positioning for Growth. Technical report, 2001.

Charles E Stuart. Welfare Costs per Dollar of Additional Tax Revenue in the United States. *American Economic Review*, 74(3):352–62, June 1984.

W. Thirsk and J. Moore. The social cost of Canadian labour taxes. *Canadian Tax Journal*, 39:55466, 1991.

HM Treasury. Cross Cutting Review of the Knowledge Economy, 2000.

HM Treasury. Charges for Information: When and How, 2001.

Peter Weiss. Borders in Cyberspace: Conflicting Government Information Policies and their Economic Impact. In G. Aicholzer and Burkert H., editors, *Public Sector Information in the Digital Age: Between Markets, Public Management and Citizens' Rights*. 2004.

Robert D Willig. Consumer's Surplus Without Apology. *The American Economic Review*, 66:589–597, September 1976. ISSN 00028282.

S. Ziebland. The importance of being expert: the quest for cancer information on the internet. *Social Science and Medicine*, 9(59):1783–1793, 2004.