

Review of “Cost Proxy Models and Telecommunications Policy”

ERIC KODJO RALPH*

EKonomics LLC

1 Introduction

This slim volume – less than 190 pages of main text – manages the astonishing feat of dragging the largely theoretical principal/agent literature, as applied to regulation, into the domain of robust practicability.¹ A clear implication for the authors is that cost proxy models need to become an essential tool of regulatory policy. However, this book does more than merely introduce and convey the value of cost proxy modelling to regulatory economists. It single-handedly provides a roadmap for the practical development of a decade’s worth of regulatory technology. The authors’ over-riding vision of making theory work in practice, a quality often missing in the academic literature, is a tribute to them. Despite covering a huge range of very complex material, the book is clearly written and presented. As an added bonus, the data and cost models used are provided on a CD so readers can extend the authors’ results or verify them through repetition and sensitivity tests.

2 Cost proxy models

Traditional econometric studies of regulated industries based on time-series data – too often with the sample being largely confined to a regulated monopoly – simply cannot generate the kind of data necessary to consider out-of-sample industry structures. Consequently, they cannot even answer such basic questions as whether, for example, in telecommunications, there are economies of scope between public switched services and private line services, or what are efficient forward-looking costs. In contrast, proxy cost models can address such questions, as the authors demonstrate, for example, in their examination of economies of scope between switched and private access lines discussed below.

The more central concern of the authors is that regulation effects firms’ willingness to reveal and reduce their costs, which hinges “on very detailed properties of the cost function of a representative firm... For example, ... on the detailed relationship of derivatives of the cost function with respect to the technology and effort variables” (177).² The use of cost proxy models provides a framework in which such effects can be modelled.

* EKonomics LLC, 2323 Audubon St, New Orleans, LA 70125-4117, USA. EKonomicsLLC@cox.net

¹ Farid Gasmi, D. Mark Kennett, Jean-Jacques Laffont, William W. Sharkey, 2002, *Cost Proxy Models and Telecommunications Policy*, MIT Press: Cambridge, MA. ISBN 0-262-7237-8.

² Un-referenced numbers in brackets refer to pages of the reviewed book.

While the authors argue that cost proxy models have a long tradition in economics (15-16), they were unable to cite any models that were sufficiently sophisticated to undertake analysis on a par with that provided in this book. Unlike earlier cost proxy models which merely mimic observed processes, LECOM, developed by Gabel and Kennet (1991) and described in the second chapter of the book, is an engineering model that engages in several layers of optimisation to establish a least-cost telecommunications network architecture. Input prices and outputs, including location of customers, demand for access lines and traffic, and types of traffic (switched or private line) are inputs to the model.

The model's outputs are a network architecture and its cost. Thus by holding input prices fixed and varying outputs a cost function is generated, though not one that is fully analytically representable. Such a cost function is sufficient to develop the first result of the book.

Chapter 3, based on Gabel and Kennet (1994), provides the simplest demonstration of the value of the proxy model approach over time-series based econometric approaches. LECOM is used to build fully optimised networks for all possible combinations of local and long distance switched services and local and long distance private services. The chapter convincingly demonstrates that public switched telephony in the form of the option to make and/or receive local and long distance telephone calls is likely to be most cheaply supplied through a unified production process. However, adding the provision of private line services likely increases production costs compared with independent switched and private line supply. In short, there are important economies of scope between public switched local and long distance calls sufficient to say these should be produced jointly, but that a marriage of switched and private line services is not productively efficient. Similarly, there appear to be economies of scope between local and long distance private line services (compare the 6th and 8th entries of column b in Table 3.2, 45). In contrast, conclusions of earlier time-series based analysis are shown to be wanting.

3 Incentive regulation: monopoly and duopoly

The analysis of scope is interesting, and helps convince the reader of the value of LECOM and cost proxy models more generally, but as already suggested, the ability to better analyse the implications of incentive regulation is the real game in the eyes of the authors. For this kind of analysis, however, much more is needed than the cost function outlined. In particular, analytically tractable cost, demand and disutility of effort functions are called for, as well as a means of characterising regulatory ignorance and a measure of the cost of government funds. A dataset was created by holding prices fixed and varying outputs (which included numbers of access lines, access line locations, traffic per line – which typically also varied by location – and sometimes allowing for both switched and private lines). Analytical cost functions were estimated using a fully flexible translog cost function or a quadratic cost function where variables took a value of zero.

Effort by the firm is represented by a single variable, the cost of labor – when a firm is 'lazy' it takes more 'expensive' labor to achieve the same outcomes, thus its labor costs are higher. The cost of capital is used by the authors to provide means of representing uncertainty on the part of the regulator about the firm's cost structure (for a given level of effort). The regulator is assumed not to know the firm's cost of capital, but to understand its likely distribution. As a result, the basic dataset included variations over the price of

labor and the price of capital. The author's results do not appear sensitive to the choice of probability functions chosen for the price of capital.

Demand for traffic is assumed to take the form of an exponential demand function and calibrated through elasticity and average annual spend assumptions (the base case assumed an elasticity of -0.2 and an average annual revenue of \$246) (96-7). The default cost of government funds for developed countries was set at 0.3, and considerably higher for developing countries, but wide variations of this value were tested for virtually all simulations. A disutility of effort function was said to take an increasing convex quadratic form and it was calibrated by nominating a minimum price of labour corresponding to zero effort, and assuming full effort reduced labor costs by 40% (97-98). This function, and its zero point, being virtually arbitrary, is the weakest part of the modelling exercise. However, the 40% estimate has some basis in industry changes post-privatisation and market deregulation, but only a single newspaper source was cited for it when more supportive evidence could have been found. Ideally, greater sensitivity testing of the effort function should have been done.

The use of the price of labour and capital respectively as proxies for effort and 'efficiency' and the calibration of demand are also not entirely compelling, and clearly call for future work. Despite this, sensitivity tests suggest, at least within the context of the chosen functional forms, that the authors' results are robust. Moreover, it would be wrong to quibble about these matters in such an ambitious and pioneering study.

With demand, cost, disutility of firm effort, and characterisation of regulatory knowledge specified, Chapter 5 begins by considering the question of economies of scale in the supply of publicly switched telephone traffic. What distinguishes this from the measurement economies of scope discussed above is that it is not conducted in the terms of the industry cost function, but rather in the context of economic efficiency taking account of the impacts of regulation. Four situations are compared: an unregulated monopolist, unregulated Cournot duopolists, a monopolist regulated optimally *à la* Laffont and Tirole (cost is observable *ex post*, but not effort – 53-57) and 'yardstick' competition where the regulator can discover how cost efficient each duopolist is (71-72). In setting up this comparison, Chapter 4 of the book provides a brilliant, concise, cohesive and relatively accessible summary of all the major forms of regulation in thirteen pages. The book is worth purchasing for this summary alone.

In a first set of comparisons where usage is varied, it is found that for all levels of examined traffic levels,

- (1) despite very large economies of scale in production, unregulated duopoly dominates unregulated monopoly;
- (2) regulation of either monopoly or duopoly improves outcomes; and
- (3) the regulated monopoly brings the greatest benefits by far.

These results imply that the efficiency gains from economies of scale achieved by the monopoly dominate the cost of the information rents associated with it (82-84). These results were supported by sensitivity analyses (84-5). However, sensitivity to the assumptions on the firms' effort decisions were not tested – no trivial omission since, as discussed, the modelling of firm effort is perhaps the most speculative aspect of the book.

This particular modelling exercise is, however, a little peculiar. Perhaps in order to focus solely on the question of economies of scale in the supply of traffic minutes, the

competitive scenario involves two identical and overlapping networks that each reach every customer. Such networks are, of course, highly unlikely to ever be observed in practice. The presence of heterogeneous firms without a completely overlapping customer base would strengthen the relative position of the regulated monopolist, since yardstick regulation would not be as effective on heterogeneous firms, and some interconnection costs (both physical and regulatory) would be incurred under the duopoly.

The second exercise of Chapter 5 addresses the question of more realistic entry, but shifts its focus to customers (access lines). Average traffic is held constant and the efficiency impact of different access line shares evaluated. Two forms of entry are assumed, both of which are probably best understood as being imposed by licensing conditions on the entrant. The first is uniform entry over the entire geographical market, so the entrant's customer base is randomly selected from anywhere in the market; and the second is targeted entry over a representative fraction of the entire market, so if the entrant has a 50% market share, it is the unique supplier of a representative half of the market (as in Argentina). In these exercises, the LECOM model is stretched a little to estimate the physical cost of interconnection, though no estimates are made of the regulatory cost of interconnection.

Again, unregulated duopoly outperforms unregulated monopoly despite interconnection costs, loss of economies of scale and, for the uniform entry case, loss of economies of density. Regulation improves economic efficiency in all cases. More interestingly,

- targeted entry with yardstick competition outperforms the Laffont and Tirole regulated monopolist when the cost of government funds is low and is equal in performance to the regulated monopoly for a high cost of funds; and
- yardstick regulated uniform entry is dominated by both regulated targeted entry and regulated monopoly reflecting the importance of economies in production.

4 Incentive regulation: Monopoly

The book also seeks to address some practical implementation issues of incentive regulation (Chapter 6). It demonstrates, *for the simple characterisation of efficiency and effort chosen by the authors*, that Laffont and Tirole regulation of a monopolist can be substantially simplified. In particular,

- the regulator can in practice treat its two instruments independently – (a) transfers to the firm to provide incentives to the firm to optimally reduce costs, that is, exercise effort; and (b) price regulation to maximise allocative efficiency (which the authors misleadingly refer to as social welfare), by application of the Ramsey rule subject to the cost of government funds (101-104, especially 103); and
- the optimal regulatory transfer function can be implemented by means of a menu of two-part tariffs (101-107). Indeed, even a simple menu of three choices works very well (107-109, especially 109).

Not surprisingly, these results do not hold for more general characterisations of effort, in particular if effort is not independent of output (109-111), as one would expect, for example, when network expansion is called for. Consequently, while the chapter seeks to address issues of implementation, it ultimately does not have any direct practical application.

Chapter 7 compares a wide range of monopoly regulatory mechanisms. The results in terms of economic efficiency are not surprising. The Laffont and Tirole mechanism outperforms all others, but interestingly is followed by Baron and Myerson optimal regulation which assumes the regulator not only cannot observe effort, but also cannot see cost (58-59). Both these mechanisms are sophisticated and require the possibility of transfers to, as well as from, the regulated firm. The third best mechanism is a price cap that also allows for transfers from the firm to the government (60-62), followed by the standard price cap (59-60). Cost-plus regulation with transfers in both directions (62-63) lies below this, with standard cost-plus regulation (62) being the least efficient of the mechanisms considered (122, Table 7.11).

Perhaps the most interesting results of this chapter are related to distribution, in particular that

- the ordinary price cap grants firms the greatest level of profit, providing a positive explanation for the support incumbents have given price cap regulation (122, Table 7.11);
- cost-plus regulation, with or without transfers, is by far the most favourable to consumers, even though it is the least efficient (126 ff; see especially 127, Table 7.2); and
- if the Laffont and Tirole mechanism is generalised to allow for different weights being applied to producer surplus, then, as less weight is given to producer welfare, leading to regulatory solutions that favour consumers, increasing amounts of producer surplus must be foregone to gain an additional increment of consumer surplus (127-129). (An argument also appears to be made for “constitutionally imposing” a price cap with transfers, being that if firms and consumers fought over what weight should be assigned to producer welfare, with firms having a 25% influence and consumers 75%, then the expected outcome would be Pareto inferior to a price cap with transfers. It is, however, not clear as to what political forces would come into play to produce this outcome (129).)

These results are again reasonably robust to parameter value changes, at least in the case of the cost of government funds and minutes-of-use demand elasticity. As either of these rise, mechanisms with transfers become relatively less efficient. This is relatively easy to understand. For a given mechanism involving transfers, social welfare *in telecommunications* is U-shaped in the cost of government funds. If the cost of government funds is very low, then they can be efficiently used to improve telecommunications welfare, but, as the cost of government funds rises, the benefits generated by their use initially declines. However, at some point, it becomes valuable to tax telecommunications use as the welfare loss generated by this is less than the cost of raising an additional dollar elsewhere. Such taxes become valuable earlier, the lower elasticity of telecommunications demand. It turns out that the minimum point for social welfare for all the mechanisms considered when the demand elasticity cost of government funds is low (-0.2) is reached

for a cost of government funds of 0.3. For higher demand elasticities (-0.7) the minimums vary more and this allows for a slight change of ordering when the cost of government funds rises. In particular, under the high elasticity case, the Baron and Myerson approach is outperformed by a price cap with transfers when the cost of government funds reaches 1, but not so in the low elasticity case. Intriguingly, in the high elasticity case, cost-plus regulation starts out better than price caps with and without transfers, but eventually ends up worse than both, while in the low elasticity case it starts out worse than both, and ends up better.

One problem with the Laffont and Tirole mechanisms is that, in general, it requires transfers from the governments' budget. As a result, the authors also indicate how the mechanism could be approximately implemented without transfers, through the setting of menus of two-part tariffs to consumers. While relatively few elements would likely be required for such a menu (see Burns and Walsh, 1985 and Wilson, 1998, 91 ff), strictly the optimisation of the Laffont and Tirole mechanism and hence the authors' earlier reported efficiency results, do not apply in this case. Under the standard Laffont and Tirole approach, government funds are used in the transfers, the marginal cost of which is the social cost of raising funds. If no transfers to or from the government are allowed, and instead the regulator imposes such 'transfers' through taxes or 'grants' to consumers via a regulated menu of two-part tariffs, then the marginal cost of funds is related to the welfare losses generated, because consumers who would be willing to pay the marginal price are not also willing to pay the fixed fee of the two-part tariff. These would most likely be low income customers (52, and Chapter 7, footnote 12, 235).

An important caveat to these comparisons of monopoly regulation, which the authors only briefly raise, is that the analysis is quite static (132). Regulators announce the rules, subject to various constraints, such as the kinds of transfers allowed and whether cost can be observed *ex post*, and then firms optimise. End of game. This raises questions about whether the authors' reported results are likely to follow in a dynamic environment. For example, under the Laffont and Tirole mechanism, costs, but not effort, are observed after supply and a transfer is made. At this point, the regulator can also back out the firm's actual cost structure, though this is of no value to it in a one-shot game. However, if the game is repeated, firms in general would want to disguise their cost-type, otherwise they would face tougher regulation in subsequent periods. The question then is how one *ex ante* commits to transfers based on costs without creating the obvious incentives to game.

Perhaps this is less problematic than it first seems, but it would have been valuable to hear the authors' views on the question. My reasoning is as follows. The effectiveness of a price cap is known to depend on the extent that the regulator can credibly commit to an exogenously determined cap (that is, at what ceiling and/or floor rates-of-return is the regulator likely to abandon the cap); and the length of the regulatory commitment. Thus the price cap's capacity to induce effort in the authors' static models can be extended to a dynamic model where a credible commitment to an exogenous cap can be made over a period of years. While it is true that the dynamics of mechanisms such as those of Laffont and Tirole, and Baron and Myerson are less well-known, similar though more complex commitments to those necessary for effective price caps are presumably possible. In particular, the promised transfers and pricing rules in these more complex cases would be based on *ex ante* knowledge and would be fixed over the period of the regulatory contract, including the specification of exogenous indices for adjustment of cost-bases, transfer amounts and pricing rules.

5 Incentive regulation: cross-subsidies and entry

Another issue that regulators face is the problem of entry into more lucrative urban areas which undermines the capacity of an incumbent to cross-subsidise rural services. Chapter 8 compares targeted urban-only entry with uniform entry across low and high cost areas (as in Argentina). Targeted urban-only entry is intended to approximate market outcomes in countries such as the US and Australia. Under the targeted entry scenario, urban prices are set by competition to cost-covering levels and these prices are imposed on the incumbent in rural markets. As a result, the incumbent faces a deficit which either is covered directly from the budget or by a telecommunications-sector tax. Under uniform (Argentine) entry two geographical monopolists are yardstick regulated to set cost-covering uniform prices imposed over both low and high cost areas with the implied cross-subsidies.

The authors' read their results as suggesting that the urban-only approach, especially with external funding, may be superior to Argentine entry where the cost of government funds is reasonably low. In my view, no strong conclusion can be drawn about the relative efficiency of these two approaches on the evidence given, though it does lean toward the authors' position.

The chapter first assumes complete information and finds that the costs of cross-subsidy are likely to be low. Perhaps this is as one would expect given a low elasticity of demand for telecommunications services, but this was not necessarily or even commonly the view taken of the costs of cross-subsidies when the possibility of introducing competition was debated in many developed countries. As a result, Argentine entry that supports the possibility of cross-subsidy is superior to urban-only entry which requires directly paying for the incumbent's deficit.

Governments, however, are not typically willing to directly fund cross-subsidies, raising the question about the efficiency of telecommunications-based taxes to fund the cross-subsidy. If such taxes are used effectively, then the efficiency effects of urban-only entry would be the same as Argentine entry, absent a technical advantage to one form of entry (and ignoring the need to cover costs and instead setting first best prices, it is first shown that urban-only entry is slightly more productively efficient than Argentine entry (144-145)). However, if taxes raised from the industry are misused the Argentine entry approach again is preferred, since the cross-subsidising firm, given perfect information, is more efficient in distributing its internally raised taxes).

Of course, the more interesting situation arises when regulators do not have perfect information. This raises the cost of Argentine entry because the regulator is not as well-informed (urban entry does not require a regulator). As a result, assuming the urban entrant is highly efficient, then urban entry with uniform prices is preferred whenever the cost of government funds is less than 4.3 or so, which would likely apply to most developed economies. In this case, if the cross-subsidy in the urban-only entry case is funded by raising taxes within the telecommunications sector, then it is better than Argentine entry (due to the efficiency losses engendered from regulation), unless tax wastage exceeds approximately 20%.

If urban-only entry is not highly efficient, its mild superiority would be undermined, though the authors do not provide a sensitivity test (which presumably could be done by a very keen reader using the supplied software). The authors do provide sensitivity tests based on demand elasticities, and their results are not much affected by changes in rural demand elasticity, but are sensitive to urban-demand elasticities. An urban demand

elasticity of 0.1 (compared with the base case of 0.2) makes it likely that Argentine entry dominates urban-entry with budget financed subsidies, while if the elasticity of urban demand is 0.6, the reverse is true.

Another issue that entry raises and which is covered in Chapter 9 is whether an incumbent, operating as a monopoly in a regulated segment of the overall market, might behave in ways that would harm entrants in more competitive market segments. However, while appearing to suggest that such actions might have anti-competitive intent, the chapter is more concerned with the possibility that optimisation may incidentally result in preventing entry or harming an entrant. Put another way, no equilibrium outcomes appear to be modelled, but rather simply the possibility raised that the incumbent's natural response to the regulatory constraints it faces might lead it to excluding entry without specifying the full implications of this. As a result, the chapter seems to be less about 'strategic cross-subsidies' designed to harm an entrant, and more about the unintended consequences regulation can have when applied to an incumbent monopolist that also competes in less regulated or unregulated markets. The chief practical lesson is not original, but is extended to the more complex environment where firms can conceal effort and regulators do not know firms' costs. Specifically, that high powered incentive schemes, such as price caps, should be applied to the regulated sector, rather than lower powered schemes, for instance, cost-plus regulation, as this reduces the basic driving force for cost shifting, namely the desire to be reimbursed in the regulated sector for unregulated costs.

The chapter begins by demonstrating that if cost-plus regulation is applied to one segment of the market then the incumbent has strong incentives to seek cost allocations that push as much shared costs as possible to the regulated sector (154-156). As is well-known, this allows the collection of monopoly rents in the regulated sector and gives a cost advantage to the incumbent in the competitive sector. However, it remains unclear whether the net effect of profit maximising actions on the part of the incumbent would be to exclude competition. For example, if it faces no regulation in the 'competitive' market *only* because there is entry, it would be better to maintain a price that allowed some entrants some degree of market share.

More interesting is the possibility that optimisation of the firm's unobserved effort might also harm entry (156 ff). In the first example, an accounting rule is fixed and the incumbent and an entrant are equally efficient (one interpretation of which is that the entrant relies on the incumbent's infrastructure for certain inputs at an effectively regulated price) (158-160). It turns out entry could be blocked if, through optimising effort applied in the monopolised and competitive market segments, the incumbent ends up setting an effort level in the competitive segment that exceeds that which is worthwhile to the entrant. In effect, the regulated firm's choices of effort levels across its production processes shifts observed costs toward the regulated sector.

Despite this, there is no anti-competitive intent here. The firm's goal is not exclusion of the entrant, but is merely an attempt to maximise profit. Nor is it clear what the actual market equilibrium would be, since exit would likely change the regulatory parameters. However, there is no doubt the authors' example raises a problem: under certain parameterisations, if the incumbent were *not* to act strategically then entry would be excluded. Moreover, it goes without saying that the regulator would not want to be in a position where strategic action on the part of the regulated firm is necessary to maintain the presence of an entrant.

Because the results just described rely on the authors' effort functions, a second example is considered which fixes firm effort. Similar results are observed – the incumbent, by optimally responding to the regulatory restraints it faces, may prevent entry. Grounding this possibility in a situation where both switched and private lines are supplied, the likelihood of entry prevention is increased as the degree to which the incumbent is reimbursed for its costs in the regulated sector rises (since this increases the gain from shifting costs), and to the extent the entrant's operations are small compared with the incumbent's. As before, the implications of this are less that such exclusion would be undertaken for anti-competitive intent or even occur, but that, under certain parameterisations, regulation could have awkward unintended consequences.

6 Minor quibbles and conclusion

Despite overall being well-written and organised, with excellent internal cross-referencing, I had some minor organisational quibbles with the book. While endnotes used to cross-reference material elsewhere in the text were very helpful, other endnotes could have profitably been left in the text, most especially given the unfortunate decision to use endnotes rather than footnotes. It was also the case that some repetition was possibly unnecessary, but in a book this complex the authors erred correctly in leaning towards too much recapping rather than too little. However, I think the material on calibrating the model at pages 96-98 should have replaced the earlier calibration description at pages 80-81.

In conclusion, this book was a pleasure to read despite its difficult subject matter. It demonstrates that recent regulatory theory can be practicable, and provides a platform on which a modern regulatory practice can be built. To my mind that is an important, much needed and even astonishing achievement.

7 References

Burns, M.E. and C. Walsh (1985) "Public goods with price exclusion: Market segmentation and allocative efficiency," in R. Pethig (Ed) *Public Goods and Public Welfare*, Peter Lang: New York, 1-39.

Gabel, D. and M. Kennet (1991) "Estimating the cost structure of the local telephone exchange network," Report 91-16, National Regulatory Research Institute, Columbus, OH.

Gabel, D. and M. Kennet (1994) "Economies of scope in the local exchange market," *Journal of Regulatory Economics*, 6: 381-398.

Wilson, R.B. (1993) *Nonlinear Pricing*, Oxford University Press: Oxford, UK.