

Outline of Presentation for ACCC Regulatory Conference
“Revisiting the Rationale for Regulation”
Session on Economics and Regulation
Gold Coast, Queensland, July 24, 2008
Russell Pittman*

Developments in economic analysis have contributed massively to improvements in regulation – as well as to deregulation and restructuring – in the past 50 years. I would argue, however, that some of the economic lessons we have learned are wrong, or at least not quite right, or at least right and useful and valuable in some circumstances but not others. I will discuss four such lessons here.

1. Profit margins tell us something about monopoly power.

Profit margins – defined here as $(total\ revenues - variable\ costs)/total\ revenues$, or some variation thereof – may provide useful information to regulators and competition law enforcers under certain circumstances. Changes, or lack of changes, in profit margins over time in response to policy changes or natural experiments may suggest answers to questions about market definition. Profit margins are necessary inputs into “critical loss” analyses, which have become widely used both in market definition and competitive effects analysis.

But profit margins are not what firms – or the invisible hand, for that matter – are assumed to maximize. (That is profit **rates**, i.e. $[revenues - costs]/assets$.) And I would argue that, despite statements to the contrary in textbooks and journal articles, cross-industry comparisons of profit margins tell us little about differences in market power.

It is true that the well-known Lerner equation shows a direct relationship between $(P - MC)$ and the elasticity of demand facing a firm. But I suggest that there are 2 reasons why we cannot translate this into the frequent assumption of a direct relationship between observed profit margins and market power.

First, as anyone knows who has worked on a predation investigation, it may be surprisingly difficult to translate measured variable cost into theoretical marginal cost. Like the ISIC codes, accounting categories and measures are not set up for the convenience of industrial economists or regulators. Franklin Fisher and others have pointed to several specific measurement issues, from the failure to capitalize advertising expenditures to the treatment of quasi-rents as profits. One possible additional complication here – I do not claim to fully understand it yet myself – is a possible systematic divergence between measured variable cost and true marginal cost depending on the capital intensity of the firm. Or could the problem be partly that the Lerner equation implicitly concerns *long-run* marginal cost, so that short-run variable cost is of necessity often a poor proxy?

* Antitrust Division, U.S. Department of Justice, and New Economic School, Moscow. The views expressed are not those of the U.S. Department of Justice.

Second, I don't know of any empirical demonstration of this point, and I've been unable to find anything in the management literature on it, but my observation is that a firm's profit margin tends to be systematically related to its capital intensity. Indeed it is hard to see how that could not be the case, since higher capital intensity requires higher returns to capital for firm viability and survival, and those returns come directly from the profit margin. Thus when there is divergence between capital intensity and the profit margin required for its maintenance, one assumes that firm entry or exit restores the balance. (Note that in this case there are no measurement problems with profit margin as a source of returns to capital – though of course the actual measurement of the capital stock, required depreciation, and so on are clearly matters of art.)

The problem, as I see it, is that this observation coupled with the Lerner equation would seem to imply a systematic relationship between the capital intensity of an industry and the market or monopoly power of the firms in the industry. Do we see such a relationship? I suppose we would all agree that significant economies of scale may cause an industry to become more "oligopolistic" and less "competitive", all else equal, but a capital intensive industry still may support a number of competing firms. I know of no study that shows that firms in capital intensive industries have more market power in firms with less. (In fact this brings to mind the old studies of Bain, Pratten, and Scherer, et al., which seemed to demonstrate that concentration in a number of industries was significantly greater than that required by scale economies.)

The lesson I take from all this is that for reasons of both measurement and theory, the use of profit margins in regulatory and antitrust analysis are better honored in the breach than in the observance.

2. Regulators and antitrust enforcers should ignore distributional considerations.

The thinking on this point is standard neoclassical economics by now. If the prohibition on interpersonal utility comparisons does not clinch the argument, then the traditional "division of labor" argument of Musgrave, Harberger, and others – that distribution is best addressed by public finance tools such as the graduated income tax and transfer payments – may carry the day.

I have disputed this point in detail concerning merger enforcement elsewhere. Here I want to make just a few broader points. First, not all economists, and certainly not all social scientists, are quite so reticent about making interpersonal utility comparisons, at least regarding income. There is a standard empirical subset of the public finance literature now that seems to support the hypothesis that the marginal utility of income declines with income – in the language of this literature, that the "Frisch parameter" varies inversely with expenditure. In addition, the assumption of a declining marginal utility of income is standard in some areas of the public finance literature, for example that using voting models to examine political party behavior.

Second, it is not as if the public finance tools of income distribution are perfect, frictionless, and innocent of causing their own distortions. Most economists would agree that a graduated income tax has at least some undesirable incentive properties; the debate is typically over the rate range at which these bad incentives become significant and/or so powerful as to outweigh the benefits of redistribution perceived to flow from this policy. Similarly, a standard part of any debate or policy design regarding the provision of “social safety nets” is how to minimize the incentives that any such services provided to the vulnerable may provide for people to remain in the targeted vulnerable status – in order to receive the services – rather than working as hard as possible to lift themselves out of it.

Third, although there are certainly regulatory and antitrust cases where the existence and/or effects of transfers are ambiguous, there are just as certainly others where one can make at least a good educated guess. As I have argued elsewhere, transfers to owners of firm stocks and bonds from final consumers can usually be safely assumed to be regressive. Furthermore, using standard economic thinking, transfers to sellers of intermediate goods from buyers of those goods can usually be assumed to be passed on down the chain to the next buyers, and ultimately to final consumers – and so to be regressive as well.

Finally, agencies and judges already factor distributional issues into their decision making in some specific cases – in evaluating “litigation risk” if nothing else. The concept of a “sympathetic plaintiff” or “sympathetic victim” is used in antitrust discussions as well as in other litigation contexts. The judge in the Antitrust Division’s *Syufy* case famously characterized the Justice Department as “suspect[ing] that giant film distributors like Columbia, Paramount and Twentieth Century-Fox had fallen prey to Raymond Syufy, the canny operator of a chain of Las Vegas, Nevada, movie theatres”, and “intent on rescuing this platoon of Goliaths from a single David.” Challenges to proposed hospital mergers must consider whether a judge can be convinced that if this local non-profit care provider raises prices, the victims will be the local citizens rather than the giant insurers that pay the bills directly. I wasn’t there, but one can imagine some concerns within the FTC about how much pain a judge would feel at the thought of the customers of Whole Foods being forced to pay higher prices for their arugula, free-range eggs, and artisanal cheeses.

3. Price caps are good. Rate of return regulation is bad.

Professor Kleindorfer has already discussed clearly some of the perils involved in taking this sometimes persuasive concept too far. He has also touched on a broader concept that is relevant here: that price caps, and their ilk, are at one end of a spectrum of the “power” of incentives that has as its other end rate of return regulation, and its ilk, and that there may be advantages and disadvantages to operating at any particular point on the spectrum.

I want to focus on two separate issues here, both of them perhaps most important in the developing country context.

First, the powerful incentives promised by price cap regimes may rely on government commitments that are in fact not credible. As even a developed country like the UK discovered, no matter how efficiently a water company operates, and no matter if prices to customers are falling in real terms, the government may simply not find it politically acceptable to allow a company that provides necessary public services to the general population to earn high profits. In the South American context, we have even seen renationalization in this context. Moreover, as Levy and Spiller emphasize, it may also not be credible – in any country, let alone a poor country – for the government to threaten to allow a provider of necessary public services to go bankrupt if it cannot meet its price cap. If the government refuses to revisit the concession and allow higher rates than originally agreed to, the provider may win that particular game of “chicken”. (Luis Guasch’s results on Latin American concessions are particularly valuable here.)

More broadly, there are other goals of regulation besides the efficiency toward which price caps are designed to lead. (I believe Darryl will discuss this broad point in a moment.) Particularly in developing countries, regulatory reforms are often, and appropriately, more concerned with attracting investment, increasing coverage, and insuring against future expropriation, than they are with pure efficiency. At this point in our experience, it is not at all clear that an imperfect price caps regime is superior to an imperfect rate of return regime in aiming at those targets; in fact the reverse seems more likely to be the case.

And finally...

4. Vertical separation is the way to create competition in network industries.

I argued against this position when I had the honor and pleasure of participating in the ACCC regulatory conference in 2005, so I will be brief here.

The idea that absolute separation of the “competitive” from the “network” portions of the railway, electricity, natural gas, etc. sectors is the best and perhaps only way to achieve a real competitive outcome is a familiar one. It relies on the (quite defensible) assumption that it is difficult for regulators to prevent discrimination by vertically integrated network operators in favor of their subsidiaries in the competitive “upstream” sector, and the (hard won) preference of economists and policy makers for competition over regulation whenever the former seems a feasible outcome.

Nevertheless there have been enough failures of restructuring using the vertical separation methodology by now to suggest caution. UK railways and California electricity are the two most well known cautionary examples, but in addition these two sectors in particular are characterized by at least significant minorities of commentators questioning more broadly whether the benefits obtained by restructuring have been – so far, at least – commensurate with the considerable costs.

Two problems that may have been inadequately considered in the heady early days of the restructuring debates are a) the potential for competition and competitive outcomes in the upstream sector, and b) the presence of economies of vertical integration that are lost with separation. Let us consider these two in turn.

In railways, it is not clear yet that competition will break out if entry is freed, or that competition will be sustainable in the long run if it does break out. For passenger service, the main problem is that basic services, at least, typically must be subsidized if operations are to take place at all. (Otherwise autos take the shorter distance traffic and airplanes the longer distance.) Governments often agree to these subsidies in order to reduce congestion and air pollution. It may be that this subsidization is more transparent if operations and infrastructure are separated – this was indeed the rationale for the first modern vertical separation of the rail sector, in Sweden – and that “competition for the market” is somehow superior for the train service alone as compared to for the vertically integrated operation – for example, because the required investment is smaller. But the outcome remains highly regulated, and the advantages created by vertical separation are in any case much more subtle than as presented in the traditional argument.

For rail freight service, on-track competition among train operators is at least feasible, and takes place currently in, for example, the UK, Poland, and Romania. One problem here is that there are apparently economies to density in freight railway operations, so that even with free entry an incumbent may have strong first-mover advantages, and there may be room for only a very small number of competitors. Just as important – especially in the eyes of the companies – is the fact that rail is a sector with high fixed and sunk costs, and it may be difficult or impossible to recover those costs without discriminating among customers with different elasticities of demand (i.e. different intermodal options). If, as would be expected and has been the experience in the countries named, market entrants go after the high margin traffic, other traffic may be left covering not much more than its variable costs (in order to keep it off the roads), and the capital in this capital intensive sector may deteriorate. (A related question, raised by the BTRE in its superb report on rail infrastructure pricing, is whether price discrimination may be easier at the shipper level than at the train level.)

In the electricity sector, serious problems for competition have been raised by the presence of different generation technologies with different degrees of cost and flexibility, exacerbated by the inelasticities of demand and supply that characterize this sector. This was allegedly one of the causes of the problems in California, and the Antitrust Division recently confronted the issue in the proposed PSEG/Exelon merger. As described for this merger by Wolak and McRae (and for California by Borenstein and Bushnell), one serious concern is that any one firm’s ownership of both baseload (low MC) and peakload (high MC) generation plants may provide it with both the incentive and ability to withhold output at times of peak demand, thus forcing up wholesale price and giving quasi-rents to its baseload plants. In addition, the fact that, in any given geographic market, the same firms may play the competition game on an hourly basis, 24 hours a day and 7 days a week, provides them with ample opportunities for learning how to interact cooperatively rather than competitively.

Finally, it seems somewhat remarkable in retrospect – I certainly claim no more real-time insight than others – that during the same period when “transactions cost economics” was developing so nicely and economists were coming to better understand the advantages of vertical contracting, vertical integration, and similar arrangements, those of us focusing on regulation, deregulation, and restructuring were working under the apparent assumption that sectors as complex as railways and electricity could be cut down the middle – “unbundled” – with little or no loss in efficiency. (I think Paul Joskow was the principal exception here.) As I have argued elsewhere, just a look at the management and engineering literature makes clear how much the costs and efficiency of railways depends on getting things exactly right at the interface of wheel and rail – “where steel meets steel” – which is of course exactly the point where vertical separation takes place as well. Similarly in the electricity sector, the physics of the intersection between generation and transmission are quite unforgiving of error, and this has led to the creation of new regulatory institutions such as systems operators who are to replace the coordination previously provided by the vertically integrated provider and who are somehow assumed to be welfare-maximizing technocrats.

I will not dwell here on an important related question: whether vertical separation in these capital intensive sectors may reduce or at least muddy the incentives for maintaining and improving the network infrastructure. I commend to you the analyses of Newbery, von Hirschhausen, and others – including, again, the BTRE – on this point.

I think many or most would agree that the jury is still out on the ability of vertical separation to deliver competition in these and other infrastructure sectors, and especially to deliver sufficient competitive benefits to outweigh the costs and complications of restructuring. (Even such a proponent as William Hogan characterizes the “glass” as at best “half full”.) I note that the World Bank has finally backed away from its general insistence that vertical separation is the best restructuring strategy for developing country infrastructure, to a more nuanced focus on particular conditions in particular settings. This is the right move, and the rest of us should follow.