WHY DO EX-MONOPOLISTS KEEP A HIGH MARKET SHARE AFTER DEREGULATION?

LIBOR DUŠEK[1]

(working paper, as of winter 2003)

SUMMARY

The paper studies the evolution of market shares in the retail markets for electricity, gas, and telecommunication services after deregulation. In most cases, the dominant position of the used-to-be monopoly retailer erodes slowly despite the absence of barriers to entry. I discuss several alternative explanations, of which the models of switching costs seem to fit the observed data best.

Collecting enough data turned out to be a rather long-term project. Therefore I restrict myself to a discussion of theories and "stylized facts" and postpone econometric analysis until the later stages of the research.

1. INTRODUCTION

The market shares of dominant firms erode, but they erode slowly. This fact is documented in the literature. For example, in their study of manufacturing trusts formed at the turn of the century, Caves, Fortunato, and Ghemawat [1985] report that the average trust had a market share of 69% in 1905, which declined to 45% by 1929, implying an annual erosion of mere 1%. Yamawaki's [1985] paper on U.S. Steel reports that U.S. Steel accounted for 65.7% of production of steel ingots and castings in 1901 (the year it was founded), and its market share declined gradually to 41.2% in 1930, implying an annual erosion of about 0.8%.

These studies generally conform to the theory of dynamic limit pricing (Gaskins [1971]), which explains the decline in the market share of the dominant firm as a consequence of rational pricing strategy: while charging the short-run monopoly price would invite quick entry and charging the entry deterring price would eliminate most of short-run profits, the dominant firm selects a price path somewhere in between so that entry does occur, and the industry converges to its long-run equilibrium. The long-run equilibrium market share of the initially dominant firm is determined by its cost advantage over actual or potential rivals (if it has any). The dynamic limit pricing theory requires some entry barriers, otherwise entry would be instantaneous and the industry would immediately reach the long-run equilibrium. Industries covered in the cited literature (manufacturing, steel) certainly satisfy the assumption of significant barriers to entry.

The recent wave of deregulation of retail markets for electricity, gas, and telecommunication industries around the world provides new and somewhat troubling evidence into this area. Since I am most familiar with electricity deregulation, I will demonstrate the ideas and institutional background only on electricity, although the concepts can be immediately applied to gas and telecommunications. Under regulation, a selected firm (henceforth referred to as ex-monopolist) has an exclusive franchise to distribute and sell electricity to end-users in a given geographical area. This firm may generate its own electricity or buy it from other generators - this distinction is not essential for the analysis of the retail market. When the retail market is deregulated, the distribution and retailing of electricity are "unbundled", at least conceptually. The ex-monopolist still remains the only distributor, but it does not have a monopoly on selling electricity to end-users. They can switch to alternative retailers, who pay

[1] I am grateful to Sam Peltzman, Jeffrey Campbell and Dennis Carlton for useful suggestions, and to Clare Spottiswoode, Ian Wood and Stephen Littlechild for many inspiring ideas.

a (regulated) access charge to the distributor for using its network. The alternative supplier can be a generator, who sells its own power directly to end-users, a power marketer, who does not own any physical assets but buys power at the wholesale market and sells it at retail, or an ex-monopolist from a different geographical area who acts as a power marketer in other areas.

What happens to the market share of ex-monopolist after it is suddenly threatened by competitive entry? The results so far (discussed in section 2) show that the market share of the ex-monopolist deteriorates slowly (although perhaps faster than in manufacturing). Yet, in the business of retailing electricity, the barriers to entry are virtually nil.[2] This might suggest that the barriers to entry are not sufficient to explain why the dominant position does not disappear overnight; the goal of this paper is to discuss alternative explanations and suggest possible ways of testing them. Section 3 discusses the hypothesis of superior cost advantage, section 4 the maximum efficient scale hypothesis, section 5 the raises the question whether access charges are unfairly high, and section 6 explains the data by the presence of switching costs. I also derive a simple model of switching costs with heterogeneous consumers.

2. "STYLIZED FACTS"

I have assembled some information on the market shares of the ex-monopolist retailers in electricity, gas and telephone markets in the United States and other countries. The data are not perfectly comparable, since the market share is sometimes measured by number of consumers, sometimes by physical output and sometimes by sales. Also the classification of consumers into "residential", "small commercial and industrial (C&I)", and "large C&I" varies from place to place. The most detailed are the data on electricity in California and Massachusetts where I have the cumulated number of consumers who switched up to the end of each month and their consumption (Charts 1-6). For Rhode Island, I have only the number of consumers by the end of each quarter (Chart 7). Chart 8 depicts the erosion of the British Gas market share year by year. Chart 9 shows the longest available time series, AT&T.

Unfortunately, the information on other markets that I found is very piecemeal and involves only a few observations on residential electricity consumers in Norway, residential and C&I electricity and telecom consumers in the UK, and all consumers of Deutsche Telecom. It is presented in Table 1. Two broad patterns emerge:

1. The ex-monopolist keeps a high market share and remains the largest supplier.
2. The ex-monopolist keeps much higher market share among residential than industrial consumers. In fact, switching of residential electricity consumers is so far minimal in all markets. Within each customer category, the market share measured by number of consumers is greater than the market share measured by quantity, implying that in each category the ex-monopolist keeps the smaller consumers.

The only exceptions are the UK electricity and gas markets for C&I consumers, where British Gas and the 12 regional electricity companies (RECs) gradually lost about 80% of the market in about 6 or 8 years. On the other hand market share of AT&T took 13 years to fall below 50%. So even when the ex-monopolist eventually becomes "small", it takes quite a while.

Why is this so? Why should the ex-monopolist's market share be any higher than the market shares of new entrants, if there are no significant barriers to entry? And why such a difference between residential and industrial customers, which certainly is not predicted by the limit

[2] The retail business needs to be distinguished from generation of electricity. While often being deregulated at the same time, generation obviously does require substantial investment in physical capital and so the entry patterns in generation could be related to entry patterns in manufacturing. The same applies to selling gas at retail versus drilling gas, or selling long-distance phone services versus installing cables and satellites. I want to focus only on the retail business in order to study the slow erosion of market share despite low barriers to entry.

pricing theory or by the Viner's theory of firm size? I think these patterns of ex-monopolist's market share are general enough across countries and industries so that they are ripe for explanation.

3. SUPERIOR COST ADVANTAGE

If the ex-monopolist is more efficient than the new entrants, than it obviously keeps a higher market share. Although the data are consistent with the superior cost advantage hypothesis, it is a rather cheap hypothesis since we could explain any industry configuration by appropriately conjecturing cost differences. Since we see a high market share in all markets, we would have to require that the ex-monopolists are more efficient than the new entrants in all markets. This requirement is very unlikely to hold. In fact, we often think that the ex-monopolists are less efficient than the new entrants due to x-inefficiencies inherited from regulation (overemployment, strong unions, bureaucratic management, poor customer service, etc.). Hence their market shares should be lower than the market shares of new entrants. Perhaps some ex-monopolists are indeed more efficient than new entrants but it is hard to believe that all of them are. Therefore we may reject the superior cost advantage hypothesis, and we need an explanation that predicts a high market share even if the ex-monopolist does not enjoy any cost advantage and possibly is at a cost disadvantage.

4. MAXIMUM EFFICIENT SCALE

This is a direct application of Viner's theory. We need to assume is that the average cost curve is decreasing up to Q_{MIN}, flat over the interval $[Q_{MIN}, Q_{MAX}]$, and strictly increasing beyond that. Q_{MIN} is the minimum efficient size and Q_{MAX} the maximum efficient size. The minimum average cost is AC_{MIN}, and the demand at $p=AC_{MIN}$ is Q. Under regulation, the monopolist serves the entire market, which is more than the maximum efficient size. After deregulation, the competition forces the ex-monopolist to shrink to at least Q_{MAX}, and the market price is AC_{MIN}. The configuration in which the ex-monopolist sells Q_{MAX} is a possible equilibrium, but other configurations in which it produces anything between $[Q_{MIN}, Q_{MAX}]$ are also possible equilibria. There has to be some adjustment cost that stops the convergence to equilibrium at Q_{MAX}.[3] Once the price falls to AC_{MIN}, any firm that changes its quantity only bears the adjustment cost without any benefit. This leaves a "residual market" for new entrants, in which the lower bound for the size of each new entrant is the minimum efficient size Q_{MIN}, and the upper bound for the number of new entrants is $(Q - Q_{MAX})/ Q_{MIN}$.[4]

There are several problems with the maximum efficient scale explanation. First, there is no particular reason why the average cost curves should look the way drawn above. We know very little about the costs of retailing of electricity. The previous cost studies (for example, Nerlove [1963], Christensen and Greene [1976]) focused on the existence of scale economies in distribution and generation; no attention was paid to retailing. In fact, some of the activities currently involved in retailing (advertising, tailoring contracts to particular customer's needs, hourly metering of electricity, risk management) were not being done when the industry was regulated. There are undoubtedly some scale economies in retailing at low number of customers; on the other hand, the fact that many suppliers survive in these markets suggests that this is not a natural monopoly, but we do not have evidence on what the maximum efficient size is.

Second, competition between ex-monopolists does not fit well into the maximum efficient size hypothesis. For example, the two largest electric utilities in California, Southern California Edison (SCE) and Pacific Gas and Electric (PGE), compete not only with the new

[3] One special case of adjustment costs are switching costs, discussed in a greater detail in section 6.
[4] Beyond that, the number of new entrants and their size distribution is indeterminate.

entrants, but also with each other. The maximum efficient size hypothesis does not say anything about matching particular customers and firms, it tacitly assumes that customers are randomly assigned to firms. For example, if adjacent monopolists A and B served 100 customers each (so the entire market has 200 customers) but the maximum efficient size is 90, 20 customers should be served by new entrants, 45 customers of firm A should be assigned to firm B and 45 should be randomly assigned to stay with firm A, and similarly for firm B. We do not observe this. We observe that out of 100 original customers of firm A, say 88 stay with firm A, 10 switch to new entrants and 2 switch to firm B. Firm A in turn takes 2 customers away from firm B. Hence we need to explain why most customers stick with the ex-monopolist.

Third, the maximum efficient size hypothesis is at odds with an observed variety of sizes among ex-monopolists. The percentage reductions in market shares are similar, yet the initial sizes of ex-monopolists are quite different. For example, the utilities in Rhode Island and Massachusetts are much smaller than SCE and PGE. British Gas used to be the monopoly gas supplier for the entire England and Wales, while electricity supply was fragmented into 12 RECs. If the maximum efficient size hypothesis is true, the larger ex-monopolists should be losing more customers than the smaller ones. In fact, if the small retailers were smaller than the maximum efficient size, they should not be losing any customers at all. However, the switching patterns seem to be independent of firm size; large variation in size of electricity retailers persists.

The *minimum* efficient size of firm can be important for explaining market shares of small ex-monopolists. In many areas of Germany and Norway, local distribution and sale of electricity has been carried out by municipal utilities. Some of them serve big cities (Bremen, Dusseldorf) but many of them are very small.[5] It is conceivable than many of them are below the minimum efficient size. If this is the case, they should be losing their customers faster than the bigger ex-monopolists and eventually go out of business or merge with bigger suppliers or with one another. This is potentially testable; unfortunately I did not find any data on these small utilities. In the long run one should be eventually able to pin down the minimum efficient size: it should be the smallest size that survives in the market.

5. RAISING RIVAL'S COSTS: UNFAIR ACCESS CHARGES

The ex-monopolist retailers also distribute electricity and collect a regulated access charge from competing retailers for access to the distribution network. Obviously, if the access charge is set too high (i.e., above the socially optimal level), the new entrants are at a competitive disadvantage; if it is too low, the ex-monopolist is at competitive disadvantage. Regulators try to prevent the ex-monopolist from disadvantaging its competitors by requiring some separation (unbundling) of retailing and distribution. The degree of unbundling ranges from accounting separation (keeping separate accounts as if these were two firms) to corporate unbundling (having a separate distribution and retail company within a holding company). But the money is still under one roof and the incentives for ex-monopolists to take advantage of the distribution bottleneck are not eliminated.[6]

[5] In the Southern and Western United States, municipal utilities and rural electric co-operatives are also common. In California, however, they are not required to provide open access to competitors due to jurisdictional reasons.

[6] I am aware of only one two examples (albeit big ones) where the ex-monopolist retailer is not integrated into distribution. First is British Gas, which split itself into a pipeline company (BG Group) and a trading company (Centrica) three years after the retail market was open for C&I consumers. BG did so voluntarily without any intervention or even informal pressure from the regulator. The second is AT&T, which could not provide the local service and in turn the local operators were until very recently prevented from providing long-distance service.

High market shares of the ex-monopolists are consistent with the hypothesis that the access charges are too high. Why should they be too high? One simple explanation lies in the basic positive theory of regulation (Stigler [1971]), according to which the regulation serves the distributor's interest and hence the regulated price is higher than the socially optimal price. More subtle explanation uses the pressure group theory (Becker [1983]), in which the access charge is determined through political pressures by the party that gains from high charges (the distributor) and the parties that lose from high charges (consumers, competitors). The pressure group approach yields an interesting prediction: The access charges should be higher for distributors that are integrated into retailing than for those who are separated. The reason is that the integrated distributor/retailer gains from higher access charge in two ways: First, it collects more money for electricity transmitted across his lines, and second, high charges give it an advantage over other retailers, and so it makes higher profits in the retail business. For the separated distributor, only the first effect is present. Therefore the integrated distributor is willing to invest more into political pressure, and the resulting access charge is higher. In theory this can be tested, in practice the sample of non-integrated ex-monopolists contains so far only two observations. .

How much the access charges matter for the ex-monopolists' market share can also be potentially tested. The speed at which they lose the market share should be negatively correlated with the access charge, after controlling for cost factors that affect the access charge.

The last benchmark that enables us to study the effect of access charges could be other deregulated industries that are not networks in nature and have low barriers to entry. In these industries, access charges are simply not an issue. Hence we can compare the speed at which the market share of ex-monopolist erodes between the network industries and non-network industries. If it is lower in the former, access charges are probably too high and they do matter.[7]

Unfair access charges can certainly help the integrated ex-monopolist and hurt the new entrants. Whether they are indeed too high is an important policy question that is worth empirical study. But one aspect of the data cannot be explained by unfair access charges: why the ex-monopolist's market share remains always higher among residential consumers than among industrial consumers.[8] Unfair access charges may probably matter for the overall market share, but not for its composition across consumer classes.

6. SWITCHING COSTS

I think that the concept of switching costs is the most appropriate for explaining the observed pattern of market shares. The theory of switching costs was formulated in a static and two-period version by Klemperer (1987) and in dynamic version by Beggs and Klemperer (1992). Assume throughout this section that all producers have the same unit cost c and that they cannot price discriminate. The consumers are assumed to have previously purchased the good

[7] The only good example I am aware of is the auto insurance in the Czech Republic, where the monopoly provider, Ceska Pojistovna, lost its statutory monopoly in January 2000. New entrants were insurance firms already established in other insurance products. Old policies were not grandfathered, so all consumers had to sign a new contract with some insurance provider. Ceska Pojistovna kept 49% market share (measured by number of cars insured).

[8] Unless we conjecture that the access charges for supplying industrial consumers are relatively lower than the access charges for supplying residential consumers - relatively to the socially optimal level for each group of consumers. This seems an odd assumption, given what we know about the regulators' tendency to cross-subsidize the high-cost consumers (Peltzman [1976]). In electricity distribution, the households are the high cost consumers, so we should rather expect the opposite: relatively higher access charges for supplying industrial consumers.

from a certain producer in period one and have non-zero costs s of switching to another producer in any subsequent period. These costs may involve time costs, direct pecuniary outlays, brand loyalty or uncertainty about the quality of the other producer's good. Products that are homogenous in real, functional way are made heterogeneous after they are first purchased: the consumer who bought quantity q from firm A in period 1 is willing to buy a product from firm B in period 2 only if $p_B *q + s < p_A *q$.

When switching costs are present, the equilibrium prices and profits are higher in all periods except the initial period. Many switching cost models study how firms try to build up a high market share in the initial period in order to reap higher profits in subsequent periods. This issue is not relevant for the deregulated industries. The 100% market share of the ex-monopolist in the initial period (the day before deregulation) can be regarded as exogenous since it was not gained by any superior performance (lower prices, patent for a new product) but was given from the government.[9] Thus we have a good natural experiment of the market behavior in subsequent periods.

The switching costs put the new entrants at a disadvantage since they must charge a lower price than the ex-monopolist if they are to make any sales. Specifically, they must charge $p_B < p_A - s/q$ in order to sell to consumers witch switching costs s and consumption q.

If the consumers are identical and the cost of supplying consumers is c (assumed to be the same for the ex-monopolist and the new entrants), the competitive equilibrium is very simple: the ex-monopolist charges $p_A = c + s/q - \varepsilon$, and no consumers switch.

I derived a simple model in which all consumers have the same switching costs but differ in the quantity consumed. (see Appendix for derivation). Then for any pair of prices p_A, p_B, $p_A \geq p_B$ there is a cutoff level of $q^* = s/(p_A - p_B)$ such that consumers with $q<q^*$ stay with the ex-monopolist and consumers with $q>q^*$ switch to a new entrant. The quantity is distributed in the population according to a density function $f(q)$. In equilibrium, the ex-monopolist charges price somewhat above $p_B + s/q_{max}$ and therefore gives up some of the largest consumers to new entrants. On the other hand, he makes a higher profit on the consumers that do not switch. The total quantity sold by the ex-monopolist, Q_A, is greater than the quantity sold by new entrants, Q_B. The following holds for the other measures of the ex-monopolist's market share:

$$p_A Q_A/(p_A Q_A + p_B Q_B) > Q_A/(Q_A + Q_B).$$

$$N_A/(N_A+N_B) > Q_A/(Q_A + Q_B)$$

In words, the market share measured by revenue is greater than if measured by quantity since $p_A > p_B$ and $Q_A > Q_B$. Also the market share measured by the number of consumers is greater than if measured by quantity since the B's consumers are larger than A's consumers, and $Q_B<Q_A$. This simple model explains very well why the ex-monopolist's market share erodes more quickly among the industrial consumers than among the residential consumers, and why the market share measured by output is always smaller than market share measured by

[9] An obvious question arises as to why some firms were given the monopoly franchise and not others. Indeed, the monopoly franchises were always given to firms that were selling electricity, gas and phone services in the franchise territory in the times when these industries were in the infant stages, and hence were probably in some sense more efficient than competitors. However, it was a long time ago. In the USA, the monopoly franchises in electricity industry were imposed by most states during 1910's, in most European countries, the electricity monopolies were given their current form by nationalization after WWII, and so on. While the 100% market shares may have not been exogenous at the time when the monopoly franchise was granted, I think it is safe to treat it as exogenous when the monopoly franchise was removed some 50 or 80 years later.

number of consumers. It does not explain why some small consumers switch at all and why some large consumers stay with the ex-monopolist.

Heterogeneity in switching cost can reconcile this contradiction. If we assume distribution of switching costs $g(s)$ that is independent of q and $g(s) > 0$ for all $s \in (0, s_{max})$ (so there are some consumers with infinitesimally low switching costs), then for each q there will always be some consumers who will switch to new entrants, and there will be more such consumers as q rises.[10]

The importance of switching costs can be documented on some regulatory policies. Many institutional details affect the switching costs, especially for residential consumers. An extreme example is Norway, where all electricity consumers have had the opportunity to change retailers since 1991. Initially, the distributors were permitted to charge a fee for changing retailers (capped by the regulator at NOK 246, or about $36). Moreover, all switchers were required to install hourly meters, while the staying consumers could keep the traditional meters. The distributors could charge an annual fee up to NOK 5000 (later reduced to NOK 4000, or about $585) to cover the costs of hourly metering. Not surprisingly, virtually no residential customers switched, the savings were not worth the costs.[11] The regulator was frustrated that competition did not step down to the residential market, and in 1997 he eliminated both the fees for changing retailers and the metering fees. Only consumers with annual consumption over 400 MWh are now required to have hourly meters. Residential consumers finally started switching: 1.75% switched in 1997 and another 2.75% in 1998, so the ex-monopolists lost 4.5% of residential consumers during 2 years. I do not have fresher data; presumably the process has not stopped here, but erosion by just 4.5% in two years is not very fast either.[12]

Switching costs appear to be high for residential consumers. In a survey of residential consumers in the UK, 86% of respondents said that they would require 10% or more reduction in their electricity bill in order to switch. Full 27% of respondents would be induced to switch only if the savings exceeded 25%.[13]

California probably made switching unappealing by imposing a stranded costs tax on all consumers (so called competition transition charge) and by mandating a 10% reduction in residential electricity bill on the day the market was opened. This did not leave enough margin for new entrants to undercut the ex-monopolists and induce consumers to switch. Our neoclassical models are silent about how the equilibrium prices are arrived at: whether the new entrants come in and undercut the ex-monopolist who has little choice than to respond by matching the price (and adding a little bit to exploit the switching costs), or whether the ex-monopolist sets the new equilibrium price at the outset (as they probably did in California with the "assistance" from the government). Yet it intuitively does matter for the market outcome - one would expect more vigorous and dynamic competition in the former case, and a faster erosion of the market share. It appears that the mandated rate reduction was a good deal for the California's ex-monopolists: the government did for them what the market would force them to do anyway, and saved them almost all residential customers.

[10] In reality, switching costs are probably positively correlated with consumption. The model with independently distributed switching costs therefore predicts more switching among high q consumers and too little switching for low q consumers than the more realistic model with correlated switching costs.

[11] A rare exception was Jan Moen, the chief electricity regulator, who is known to have been changing suppliers very frequently for the sheer fun of exercising the freedom to choose.

[12] NVE Report [1998].

[13] MORI report [1999].

One final note on switching costs concerns the long-run dynamics. In the Beggs and Klemperer (1992) model, new consumers are entering the market each period and some old consumers are leaving. In the deregulated industries, all consumers are old when the market is deregulated. They first make a decision whether to stay or switch; after they do, the industry is in a "short-run" equilibrium and the ex-monopolist loses some consumers once and for all. The new consumers do not bear any switching costs when choosing their retailer for the first time and so the market shares of all firms (including the ex-monopolist) should be evenly distributed across the new consumers. The speed at which the market share of the ex-monopolist declines is determined by the rate of turnover of consumers. High rate of turnover (a lot of people moving or starting new households and businesses) implies faster decline. The long-run equilibrium is reached the day when all consumers who were old at the date of deregulation will have left the market. Therefore we should predict that the market share of the ex-monopolist should keep declining even after the "short run" equilibrium is reached. We do not know whether it takes months or years to reach the "short-run" equilibrium, and most of the deregulations presented here are too recent to demonstrate this dynamic effect. A good exception is AT&T, which was deregulated in 1984 and whose market share (as measured by revenue) has been constantly declining by about 3.2 % a year. Disentangling empirically the switching of old consumers and first purchases of new consumers may be extremely hard since the market share data do not distinguish these two groups of consumers.

CONCLUSIONS

Despite no barriers to entry, dominant firms do not disappear overnight. Deregulation makes the retail markets contestable rather than perfectly competitive, and switching costs appear to be the factor that blocks the ex-monopolist's market share from eroding quickly.

The unanswered question is what determines the differences in the speed of erosion between markets. The electricity and gas markets in the UK saw much faster erosion than, say, the telecommunications market in the UK or the electricity market in California. The design of market and regulatory institutions at the time of deregulation appears to matter and therefore a better understanding of this phenomenon is of interest not only to the prospective ex-monopolists and their competitors in other markets that are on the verge of deregulation (gas in the EU, electricity in many US states, telecommunications in Central Europe), but also to policy makers. I think it is a potential direction for future research.

REFERENCES

Becker, Gary S.: A theory of Competition Among Pressure Groups for Political Influence, Quarterly Journal of Economics, Vol. 98, No. 3. (August 1983), pp. 371-400.

Beggs, Alan, and Paul Klemperer: Multi-Period Competition with Switching Costs, Econometrica Vol. 60, No. 3 (May 1992), pp. 651-666.

Caves, Richard E., Michael Fortunato, Pankaj Ghemawat: The Decline of Dominant Firms, 1905-1929, Quarterly Journal of Economics, Vol. 99, No. 3. (Aug., 1984), pp. 523-546.

Christensen, Laurits R., and William H. Greene: Economies of Scale in the U.S. Power Industry, Journal of Political Economy, Vol. 84, No.4 (August 1976), pp.655-676.

Gaskins, Darius W.: Dynamic Limit Pricing: Optimal Pricing under the Threat of Entry, Journal of Economic Theory, Vol. 3 (1971), pp. 306-322.

Klemperer, Paul: Markets with Consumer Switching Costs, The Quarterly Journal of Economics, Vol. 102, No. 2 (May 1987), pp. 375-394.

Nerlove, Marc: Returns to Scale in Electricity Supply, in C. Christ et al.: Measurement in Economics, Stanford University Press 1963.

Peltzman, Sam: Toward a More General Theory of Regulation, Journal of Law and Economics (January 1976), pp. 211-240.

Stigler, George: The Theory of Economic Regulation, Bell Journal of Economics (April 1971), pp. 3-22.

Yamawaki, Hideki: Dominant Firm Pricing and Fringe Expansion: The Case of the U.S. Iron and Steel Industry, 1907-1930, The Review of Economics and Statistics, Vol. 67, No. 3. (August 1985), pp. 429-437.

DATA SOURCES

Centrica, plc

DASR Reports, California Public Utilities Commission

Electricity and Gas Competition Review, research study conducted by MORI for OFGEM (September 1999). (MORI report)

Federal Communications Commission, Common Carrier Bureau (web site)

Massachusetts Department of Public Utilities (web site)

MF Dnes (Czech daily paper), January 3, 2000.

Opening the Power Market to End Users in Norway 1991 - 1999, Norwegian Water Resources and Energy Administration (December 1998). (NVE report).

Prof. Stephen Littlechild

Review of the Development of Competition in the Domestic Gas Market, OFGEM, December 1999.

Rhode Island Public Utilities Commission (web site).

Welfens, Paul J., and Cornelius Graack: Telecommunication in Western Europe: Liberalization, Technological Dynamics and Regulatory Developments, in Welfens, Paul J., and George Yarrow: Telecommunications and Energy in Systemic Transformation, Springer, Berlin 1997.

Appendix
Simple Switching Cost Model Where Consumers Differ by Quantity Consumed

Assume that in period 0 all consumers have purchased the good from firm A (the ex-monopolist). In period 1, firm B enters and offers price p_b. Both firms have the same unit costs c. If a consumer switches from A to B, he bears switching cost s, which is identical for all consumers. The consumers differ in the quantity consumed q, which is independent of price. Quantity is distributed across consumers according to a continuous density function $f(q)$ over an interval $[0, q_{\max}]$.
A consumer will switch from A to B if
$p_b q + s \leq p_a q$

This defines a cutoff value of q^*, such that consumers with $q < q^*$ stay with firm A and cosnumers with $q > q^*$ switch to firm B:
$q^* = \frac{s}{p_a - p_b}$

Obviously p_b has to be less than p_a if the firm B is to make any sales. The qunatities sold by both firms as a function of prices are given by
$Q_a = \int_0^{q^*} q\, f(q) dq, \;\; Q_b = \int_{q^*}^{q_{\max}} q\, f(q) dq$

The total quantity sold does not depend on prices:
$Q = Q_a + Q_b = \int_0^{q_{\max}} f(q) dq$

Profits of both firms are given by

$\pi_a = (p_a - c) Q_a, \;\; \pi_b = (p_b - c) Q_b$

A pure strategy equilibrium in prices is guaranteed to exist if profit functions π_a, π_b are continuous (they are, since $f(q)$ is continuous) and if they satisfy the strategic complementarity condition:

$\frac{\partial^2 \pi_a}{\partial p_a \partial p_b} \geq 0, \;\; \frac{\partial^2 \pi_b}{\partial p_a \partial p_b} \geq 0$

Check this condition for π_a :

$$\frac{\partial^2 \pi_a}{\partial p_a \partial p_b} = \frac{\partial Q_a}{\partial p_b} + (p_a - c)\frac{\partial^2 Q_a}{\partial p_a \partial p_b} =$$
$$= \frac{\partial Q_a}{\partial p_b} + (p_a - c)\frac{q^{*2}}{(p_a - p_b)^2} f(q^*) - (p_a - c)\frac{q^{*3}}{(p_a - p_b)} f'(q^*)$$

The first two terms are positive, the third term is non-negative if $f'(q^*) \leq 0$, i.e., if $q^*$is at the decreasing tail of the density function. This I think is an innocuous assumption to make since the real life distribution of quantity consumed is heavily left skewed: there are many small consumers and very few large consumers. Hence the q^*should be on the decreasing tail.
An interior solution in which some consumers do switch requires that $f(q_{\max})$ is not "too big", i.e., there is not a big mass on the upper tail that would make it attractive for firm A to keep the biggest consumers by charging a price that firm B cannot beat ($p_a \leq c + s/q_{\max}$). If these conditions hold, we may characterize the equilibrium from the first order conditions:

$$Q_a + (p_a - c)\frac{\partial Q_a}{\partial p_a} = Q_a - (p_a - c)\frac{q^* f(q^*)}{p_a - p_b} = 0$$
$$Q_b + (p_b - c)\frac{\partial Q_b}{\partial p_b} = Q_b - (p_b - c)\frac{q^* f(q^*)}{p_a - p_b} = 0$$

which after some rearranging gives

$$\frac{Q_a}{Q_b} = \frac{p_a - c}{p_b - c}$$

Hence the ratio of sales is proportional to the equilibrium price-cost markups. Since $p_a > p_b$, this implies that $Q_a > Q_b$. The ex-monopolist charges a higher price and sells more output at the same time. This implies that its market share measured by revenues is greater than if measured by output. Since the ex-monopolist keeps the smaller consumers and $Q_a > Q_b$, the market share measured by number of consumers is also greater than if measured by output.

Chart1

California, Electricity
Residential Consumers Taken Away by Competitors since April 98

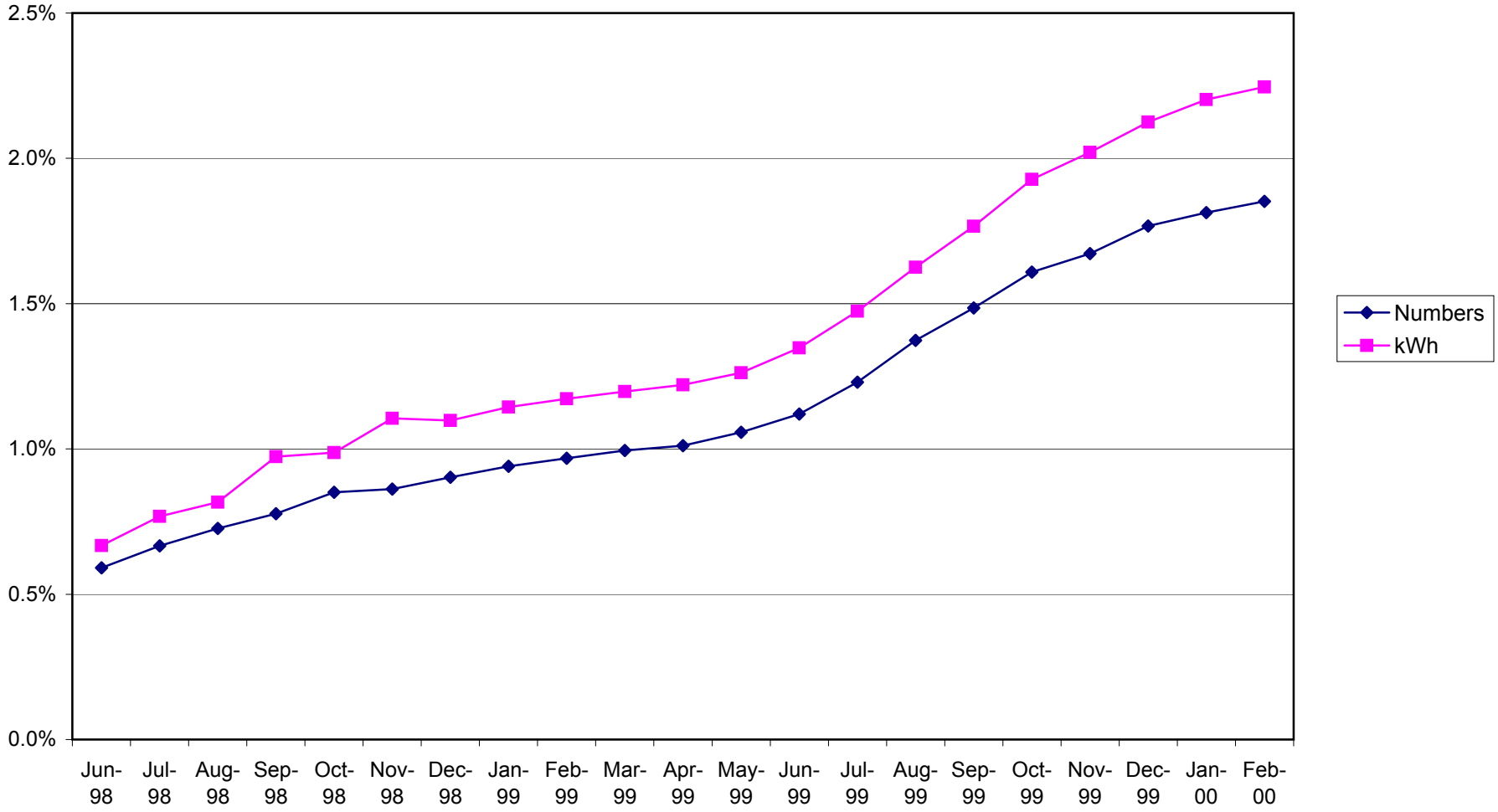


Chart2

California, Electricity
Medium Commercial and Industrial Consumers Taken Away by Competitors since April 98

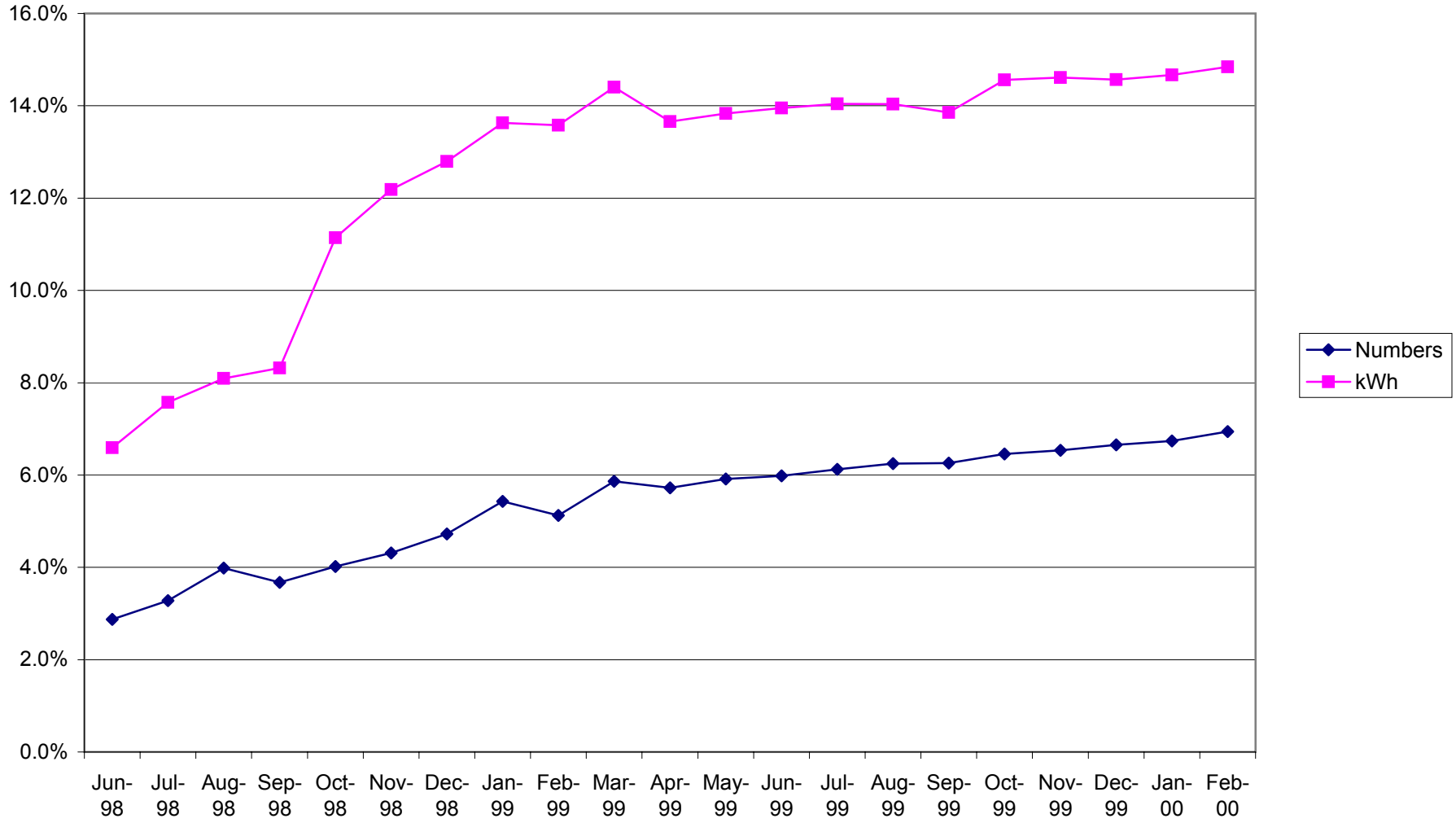


Chart3

California, Electricity
Large Commercial and Industrial Consumers Taken Away by Competitors since April 98

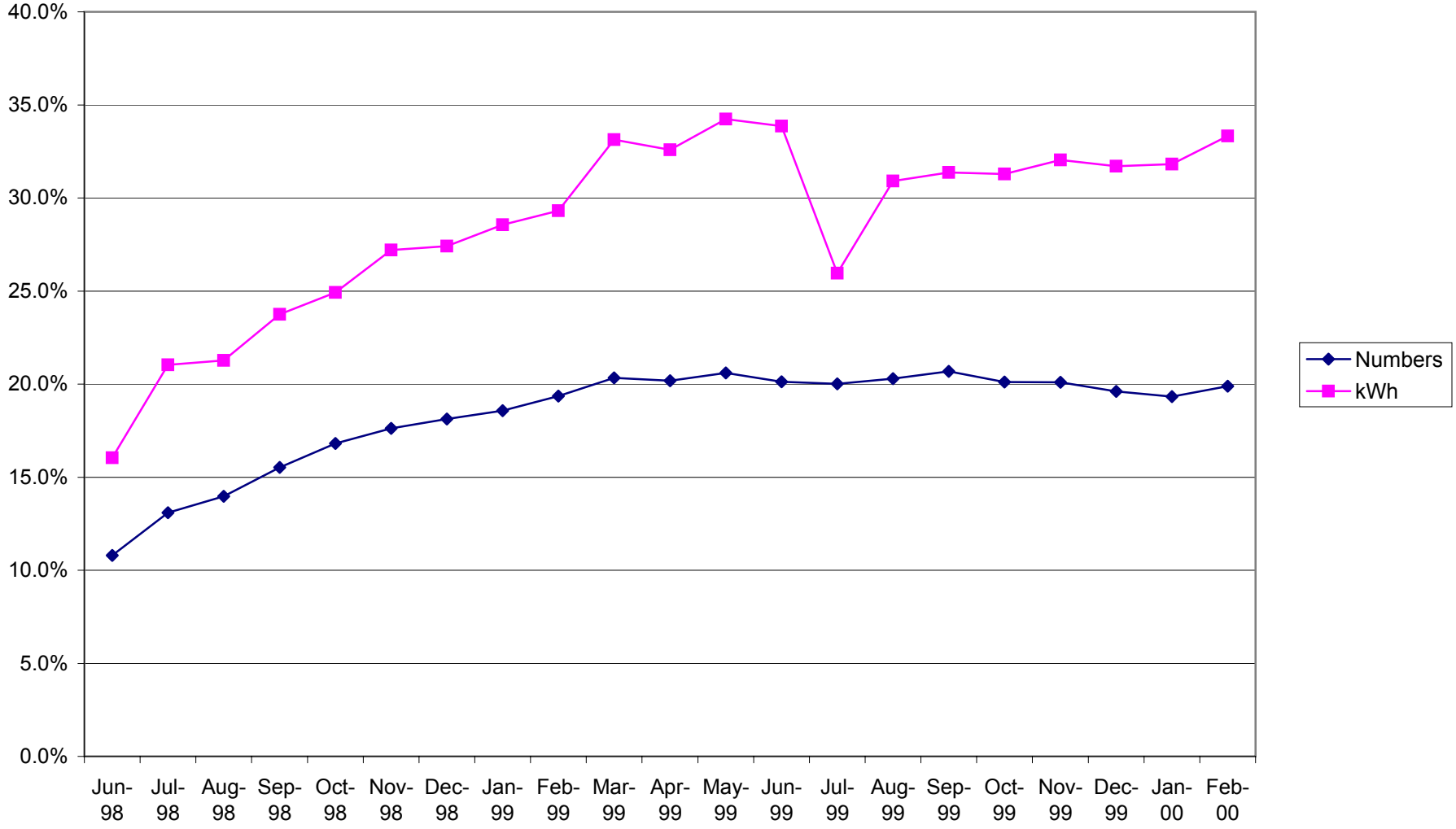


Chart4

Massachusetts, Electricity
Residential Consumers Taken Away by Competitors Since March 98

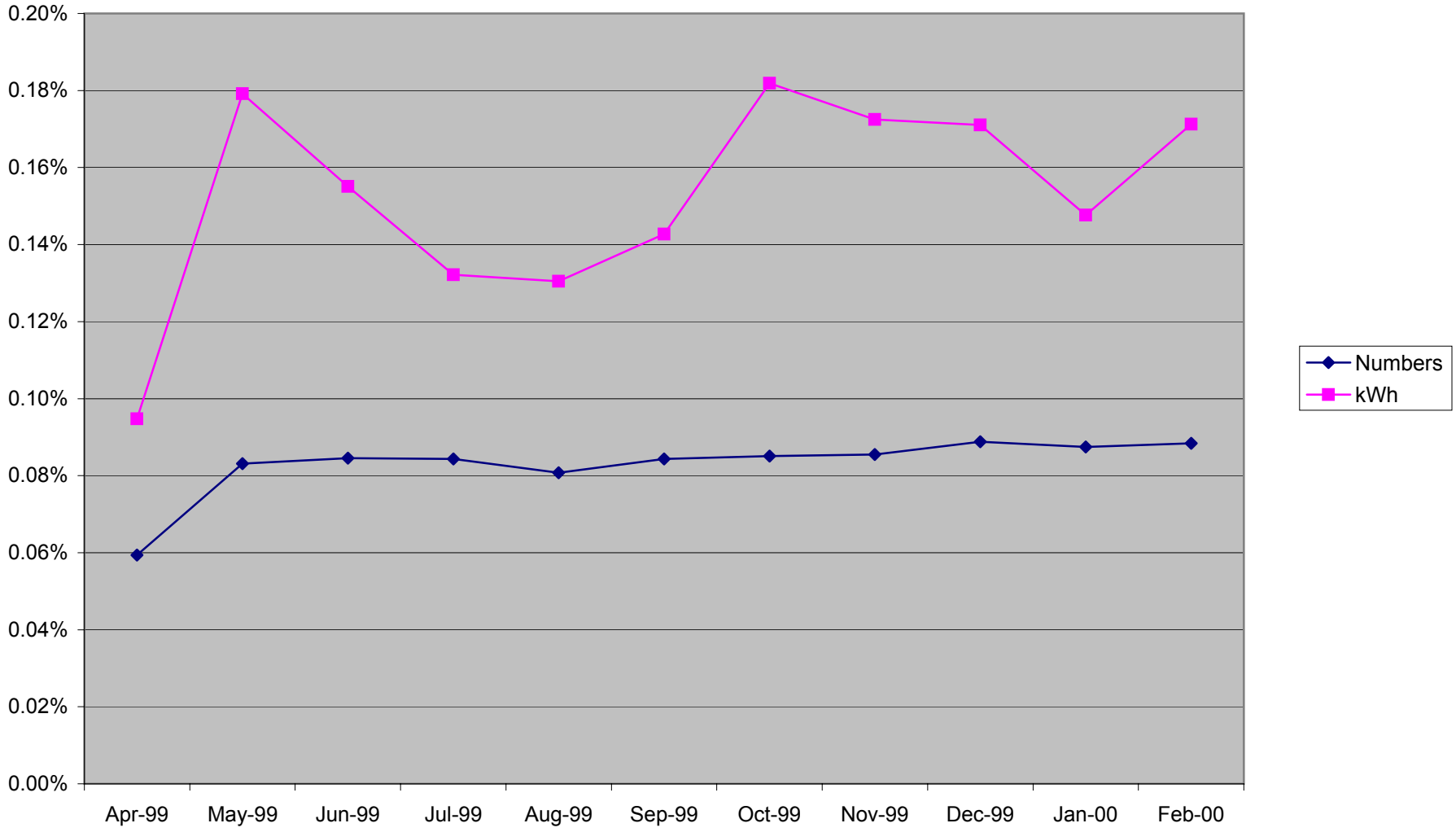


Chart5

Massachusetts, Electricity
Small Commercial and Industrial Consumers Taken Away by Comeptitots Since March 98

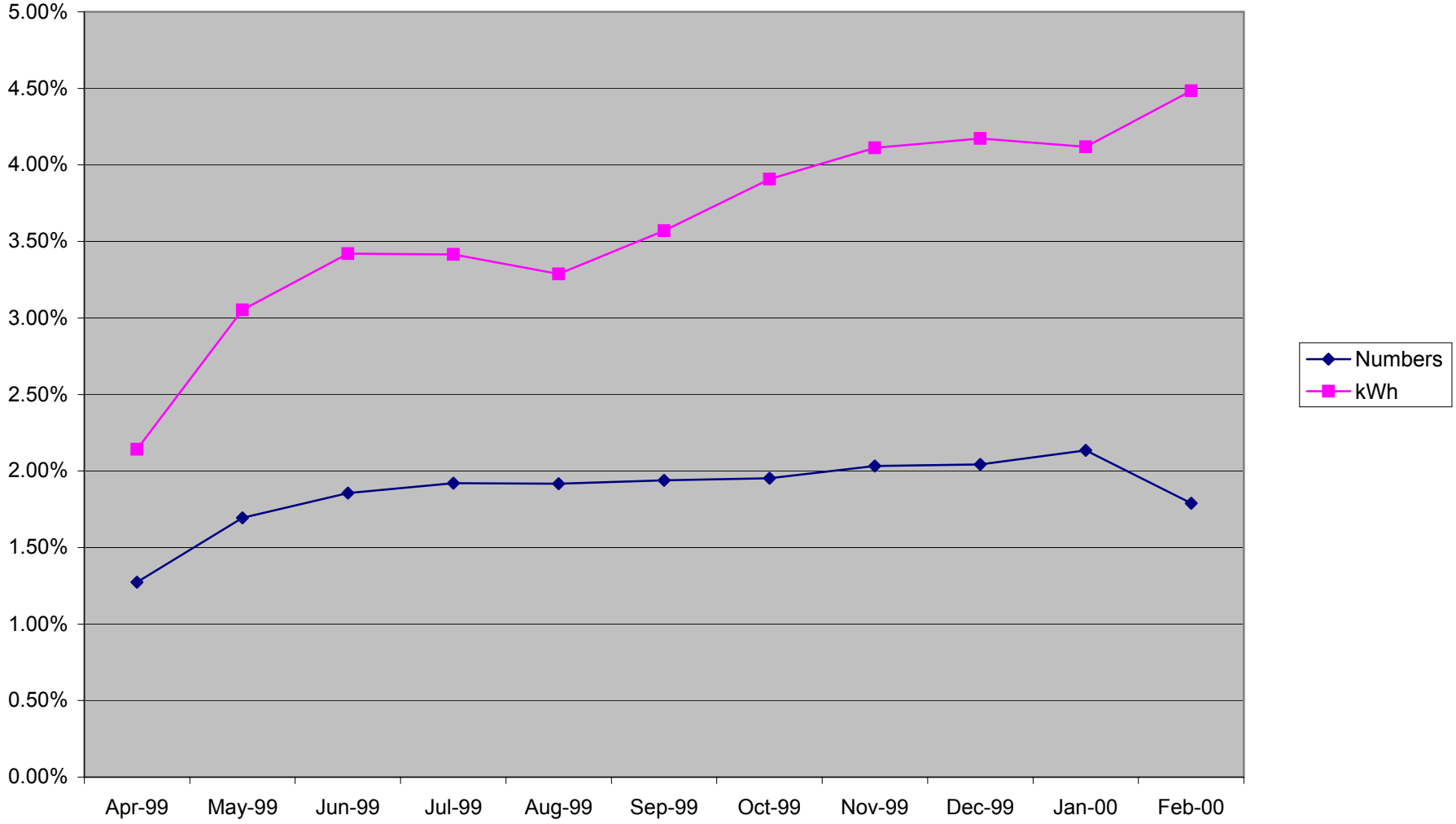


Chart6

Massachusetts, Electricity
Large Commercial and Industrial Consumers Taken Away by Competitors Since March 98

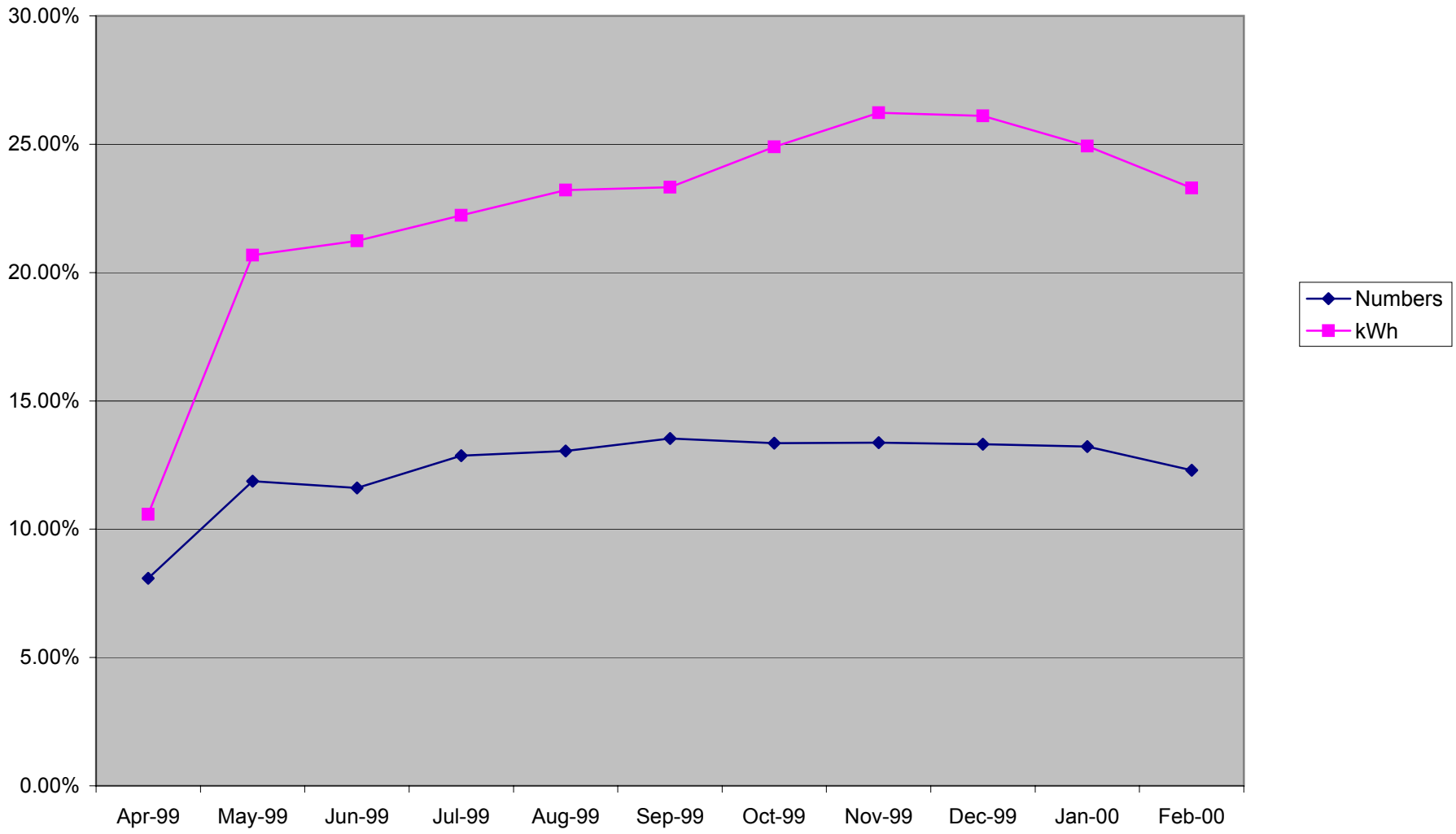


Chart7

Rhode Island, Electricity
Number of Customers Taken Away by Competitors Since January 98

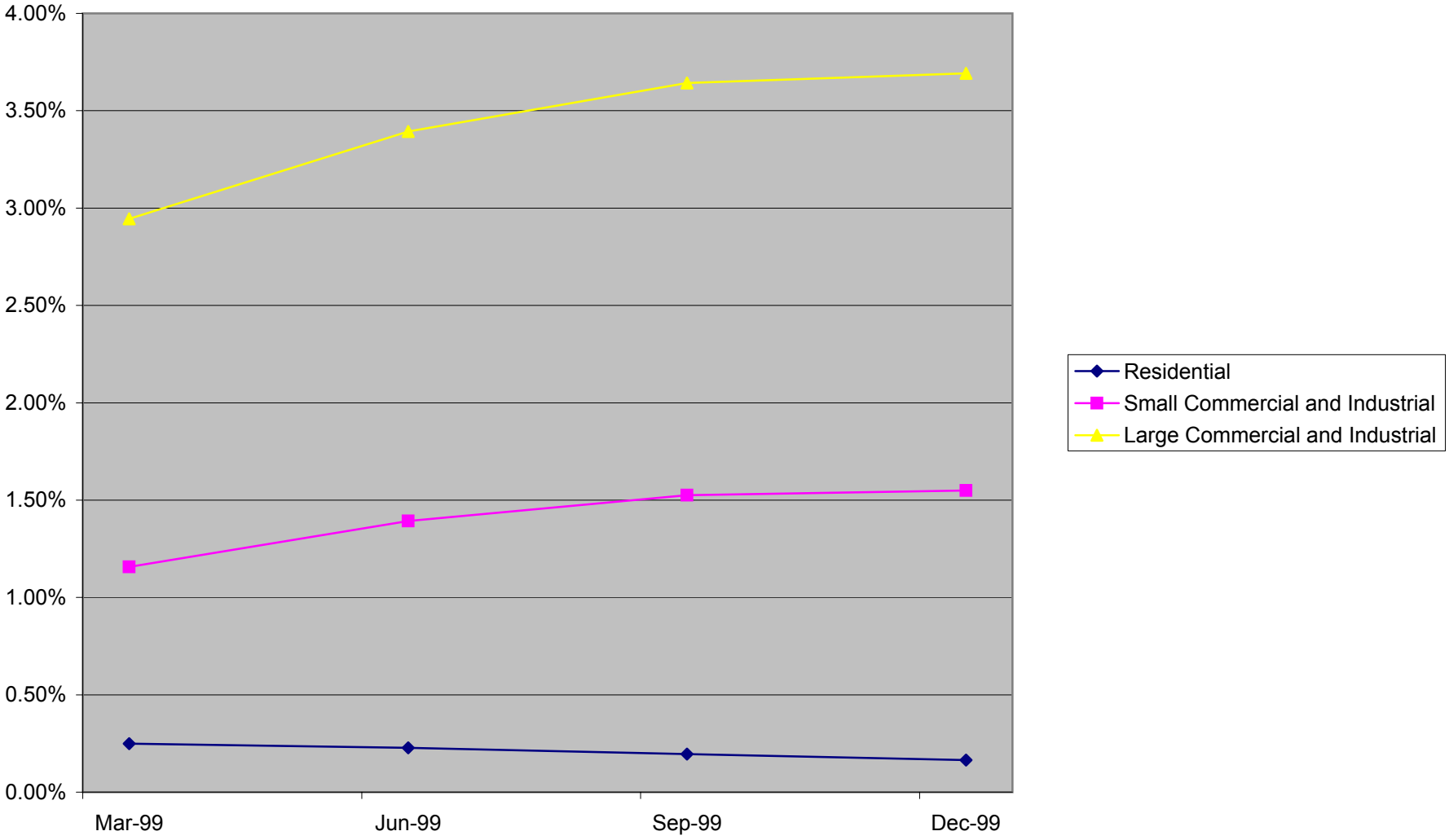
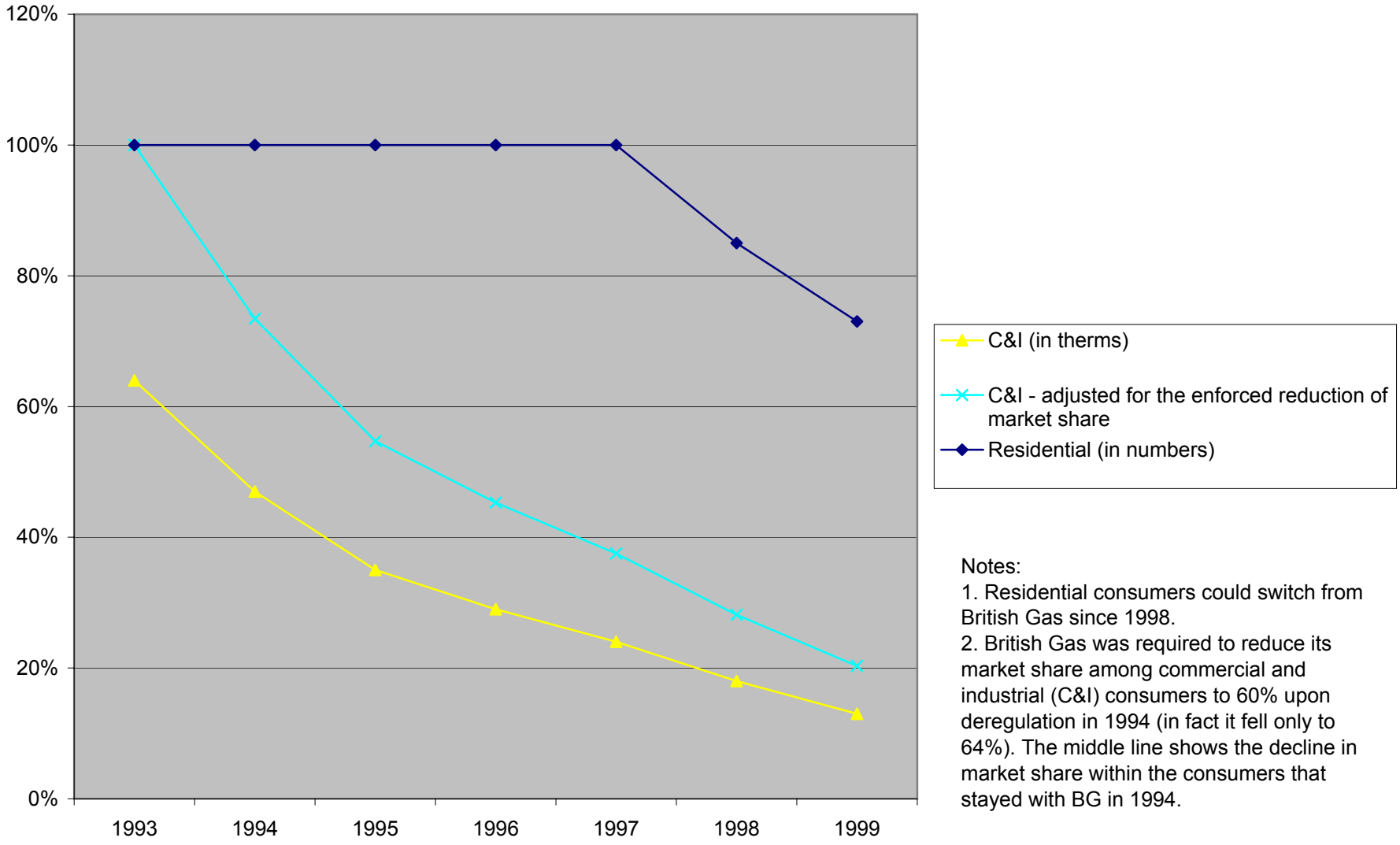


Chart8

British Gas Market Share since deregulation



Notes:
1. Residential consumers could switch from British Gas since 1998.
2. British Gas was required to reduce its market share among commercial and industrial (C&I) consumers to 60% upon deregulation in 1994 (in fact it fell only to 64%). The middle line shows the decline in market share within the consumers that stayed with BG in 1994.

Chart9

AT&T Market Share Since 1984

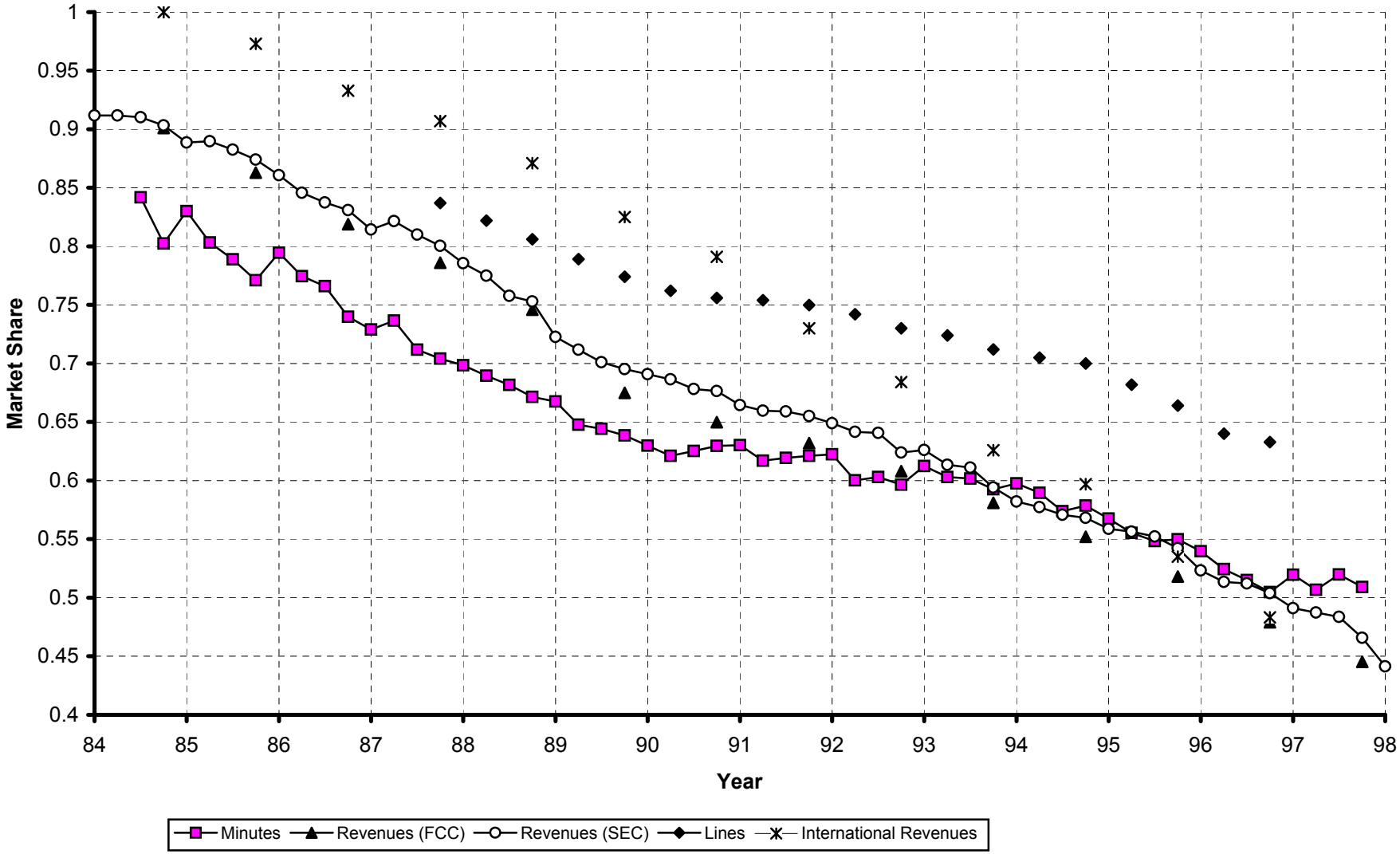


Table 1

Other selected industries
Market shares of ex-monopolists

Industry	Country	Customer category	Years since deregulation 1	2	3	5	7	8	9	Measured by
Electricity	UK	>1MW	60%					20%	20%	kWh
	UK	100kW - 1MW	75%			40%				kWh
	UK	<100kW	98%	92%						number
	Norway	residential	99%+			99%+	98%	96%		number
Telecoms	UK	residential	99%+	99%	98%			88%		revenue
	UK	C&I	94%	90%	87%					revenue
	Germany	all	65%							revenue